Toward Nano-accuracy in Scanning Beam Interference Lithography

by

Juan Montoya

B.S., University of Texas at Arlington (1999)
S.M., Massachusetts Institute of Technology (2002)

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2006

Author .....
Department of Electrical Engineering and Computer Science
May 25, 2006

Certified by .....
Mark Schattenburg
Senior Research Scientist
Thesis Supervisor

Accepted by .....
Arthur C. Smith
Chairman, Department Committee on Graduate Students

Toward Nano-accuracy in Scanning Beam Interference Lithography

by

Juan Montoya

Submitted to the Department of Electrical Engineering and Computer Science
on May 25, 2006, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

Scanning beam interference lithography is a technique developed in our laboratory which uses interfering beams and a scanning stage to rapidly pattern gratings over large areas (300x300 mm^2) with high precision. The repeatability of the system $\approx \pm 3$ nm is an important precursor for obtaining nanometer accuracy. The R&D award winning tool developed in our laboratory, referred to as the nanoruler, uses scanning beam interference lithography to pattern large gratings with periods on the order of 574 nm at velocities approaching 100 mm/s.

In this thesis, I will present techniques which I developed to improve the accuracy of the nanoruler. These techniques include mirror mapping, which allows one to characterize the reference mirrors used for stage scanning. In addition, I will present characterization techniques which include translation and rotation tests to measure the distortion present in our system. In order to correct for the measured distortion, I have implemented an on the fly phase-lookup technique in which the phase of the interfering beams are modulated to correct for the system distortion.

Several potential applications of this technology require not only high phase fidelity, but uniform linewidth as well. Toward this end, I have presented a detailed analysis of the relationship between the exposure dose contrast, beam geometry, phase modulation, and stage scanning parameters. In addition, I have implemented novel scanning techniques which have allowed for patterning more general periodic structures. For example, a technique referred to as Doppler writing will allow one to scan the stage perpendicular to the interference fringes. This technique may be utilized to create several overlapping strips of grating, each with a different period, allowing one to obtain a chirp in a direction parallel to the interference fringes.

Furthermore, I developed a patterning technique referred to as beam-blanking. While conceptually simple, the challenges for implementing this writing strategy includes synchronization of high speed electronics with the stage motion to phase-lock the interfering beams to the stage at high stage velocities. By combining all of the latter techniques: namely the ability to phase-lock, turn off the writing beams, implement generalized scanning with phase-look up on the fly, several more generalized

geometries of interest for applications including photonic Bragg devices, metrology, and X-Ray telescopes may be patterned at high speed, over large distances, with precision and accuracy.

Thesis Supervisor: Mark Schattenburg
Title: Senior Research Scientist

Acknowledgments

I would like to thank my advisor, Mark Schattenburg, for his leadership, vision, and mentorship in this project. Mark has been generous with his time and encouragement in this project. In addition, he has provided much insight into the fascinating field of optics and metrology. He is always receptive to ideas, patient, and allows his students enough freedom to explore the realm of science independently in order to grow and progress as future scientists.

I would also like to thank my Mom and Dad, and my family for their support throughout my educational experience here. My sister, Ana and husband David Hunter, and my older brother Sergio along with my Aunt and Uncle: Camila and Andy.

I would also like to thank my girlfriend Jana Ziemba (and marathon coach). She has been very supportive throughout writing of my thesis. We have coined a unit of measure, the so called love-nano-milli-meter (equivalent to one love-picometer). Her help and love throughout this thesis will not be forgotten.

Finally, I would like to thank the many Professors who have had a big influence on my educational experience here at MIT. Notably, my master thesis advisor Qing Hu, with whom I began my graduate career. The skills passed on from the Hu lab have been very influential. Professor Fonstad, for whom I served as a teaching assistant and who also served on my thesis committee. It was truly a pleasure to assist in his microelectronics course, he provides incredible insight into a complicated world of electrons, holes, photons, and semiconductor devices. Professor Orlando, for providing a welcoming environment at MIT. My first courses at MIT were in Prof. Orlando's Quantum Mechanics (6.720) and Solid State Physics (6.732). Indeed, his enthusiasm and insightful lectures made it fascinating to learn about the "spooky action at a distance". In addition, he provided a collegial environment by hosting book readings in which well respected technical authors would discuss their books with students.

Moreover, I would like to acknowledge all who have worked on or contributed to

this project: Carl Chen, Paul Konkola, Chih Hao (thanks for the SEM pics), Mirielle Akilian, Ralf Heilmann, Bob Fleming, and Yong Zhao. From the Nano-Structures laboratory: Professor Hank Smith for passing on a vast array of knowledge to his students and for organizing discussion forums where ideas may be freely exchanged. Professor Karl Berggren (associate director of the Nano-Structures Lab) for serving on my thesis committee and for his feedback, comments, and expertise. Finally, Jim Daley, Tim Savas, and the many others of NSL for their help and assistance.

I have not only learned from Professors and my research group, but from incredibly gifted colleagues. There are too many to mention, but notably Hans Callebaute, Sushil Kumar, Janice Lee, Chris Rycroft, Vivian Lei, Joe Rumpler, Kaity Ryan, Tyrone Hill, Saeed Saremi, and Leeland Ekstrom.

Last but not least, I would like to thank my sponsors at NASA, and Plymouth Grating Labs. I sincerely hope for PGL's continued success, and that a fruitfull collaboration between academic research and commercialization will lead to advanced technologies. I would also like to thank Doug Smith for his generosity and support on this project, Sean Smith for his enthusiams and technical assistance, and David Chargin (with Fraunhoffer USA) for his brilliant mechanical engineering insight.

Contents

List of Figures

Chapter 1

Introduction

1.1 The goal and motivation for Scanning Beam Interference Lithography

The goal of Scanning Beam Interference Lithography (SBIL) is to produce large area (on the order of 300 mm^2) periodic patterns with high phase fidelity, uniform linewidth, at high speed. Applications for this technology include the fabrication of gratings for metrology, space instrumentation such as X-Ray telescopes, and high power laser pulse compression to name a few. In the area of metrology there is an ongoing research effort to provide metrology for small features on large length scales. Indeed, the International Technology Roadmap for Semiconductors in 2004 requires positioning ranges to cover an area of 450x450mm by 2010 to 2014 and to manufactur features on the order of 45 nm[29]. The high precision over large areas is necessary for measuring small feature sizes in integrated circuits, wafer inspection, and in new photonic and biotech devices. The gratings we are producing may address the semiconductor roadmap's metrology needs by functioning as a position encoder (otherwise known as an optical encoder).

Moreover, Schattenburg and Smith have argued that there is a "critical role of metrology in nanotechnology"[39]. Citing the importance of measurement in several industrial revolutions, an argument is presented that nanometrology is necessary for

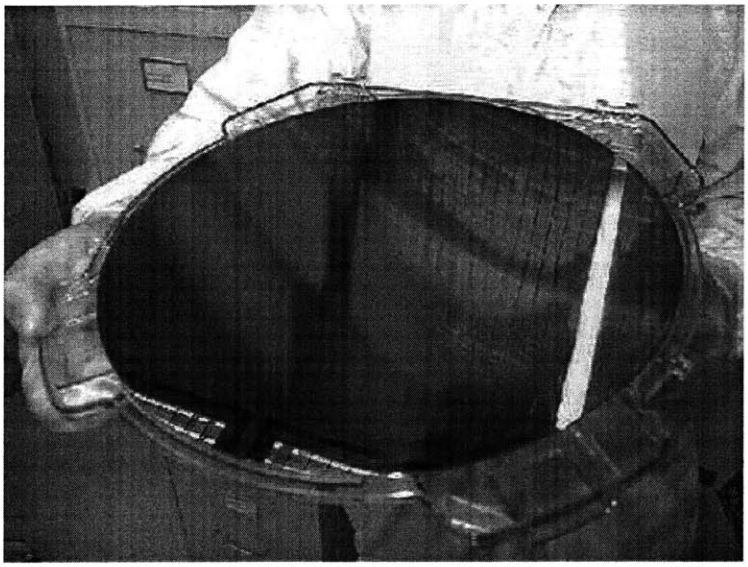

Figure 1-1: *300 mm grating fabricated in Scanning Beam Interference Lithography on a silicon substrate*

the nanotech era. Indeed, according to the National Institute of Standards and Technology (NIST), the semiconductor industry (a $200 billion dollar industry in 2003) has been driving the metrology needs for other industrial sectors [10]. A paper by Postek at NIST illustrates that the value of a nanometer control in the IC industry exceeds ten dollars per nanometer (per microprocessor) resulting in an overal 1 billion dollar savings for critical control [35]. Indeed, it is estimated that the NIST linewidth standard resulted in a savings of over 30 million dollars. Moreover, Postek observes that the rule of thumb in the semiconductor industry is "if you can't measure it, you can't make it". This unfortunate truth leaves several advanced technologies shelved in research institutions, and slows the progress of a new market.

There are several examples of emerging technology that may benefit from nanoaccurate positioning. For several years carbon nanotubes have been receiving increasing attention. A recent application of carbon nanotube nanodevices involves their uses as vertical interconnects for semiconductor technology. As current densities increase due in part to the scaling of semiconductor devices a need for replacing traditional silicon technology metal wires has made carbon nanotubes a promising candidate for

interconnects. Another example of carbon nanotube devices involves using arrays of carbon nanotube Field Effect Transistor (FET) logic circuits. However, as pointed out in the literature, a major hurdle for producing carbon nanotube FETs involves the limitation associated with positioning carbon nanotubes at specific locations on a wafer because the dimensions of the nanotubes are small and the transistor sizes are large [5, 15].

One may make a case study of carbon nanotubes because indeed they illustrate many of the nano-fabrication principles occuring in todays nano-era. For example, a proposed method for fabricating carbon nanotubes uses a photo-lithography produced template. Similarly, a technique known as templated self-assembly utilizes a template which contains long range spatial order to assist the growth process of self assembled materials. Self assembled materials exhibit short range order, but when combined with lithographic techniques long range spatial order may be achieved [24, 8, 9].

Another example in which coherent diffraction gratings are useful is in the field of photonic devices. Bragg gratings are commonly used as add-drop filters in optical communications, mirrors in laser cavitities, and as dispersion compensators [28]. This is because they actively select certain frequencies while rejecting others based on constructive or destructive feeback from reflection or transmission from a grating. When used in this manner, this process is sometimes referred to as distributive feedback (for example distributed feedback lasers). Smith et al have developed a technique known as spatial phase-locked e-beam lithography for creating long range spatially coherent Bragg devices. They have shown for add/drop channel filters that phase errors larger than 5 nm lead to undesirable consequences [28, 17].

In the field of space telescopes, large area gratings are useful since they allow for a larger collection area of shallow incidence X-Rays. Indeed, our sponsors at NASA have been funding our work in order to progress the technology to be used in future space missions. Recently, a new sponsor, Plymouth Grating Laboratories, have been applying our technology for high-power laser pulse compression. High-power pulses contain large energies over small areas, sufficiently high to damage small optical components. If the beams are expanded such that they contain less energy

per unit area (i.e. the intensity is decreased) the use of large optics such as our diffractions gratings may be employeed as an optical device, in this instance for pulse compression.

While I have listed a number of applications for this technology, it is by no means exhaustive. Currently, there is plenty of activity in nanotechnology as evidenced by the emergence of new technical journals and hot research areas in industry and academia related to the field. In order for mass production and commercialization to occur, metrological requirements must be satisfied. It is the goal of this thesis to progress towards providing these metrology needs.

1.2 History of Scanning Beam Interference Lithography

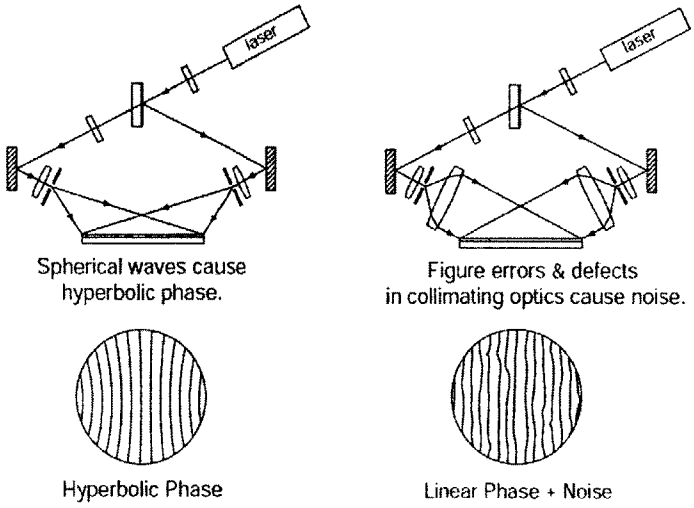

Figure 1-2: *Traditional Interference Lithography.*

The motivation for creating large area gratings is clear. The natural question arises as to why SBIL? The answer to this question is found by examining the predecessors to

scanning beam interference lithography. Traditional interference lithography involves a single exposure, and is therefore arguably fast. For covering large areas, a beam is expanded into a spherical wave. Now one may interfere this spherical wave over a large area substrate. However, the drawback of this approach is that it produces hyperbolic phase distortions [6] as shown in figure 1-2 (a). Alternatively, one may use a large lens to collimate the beams as shown in figure 1-2 (b). These lenses are notoriously difficult and costly to make and commonly introduce errors due to manufacturing imperfections, known as figure error, into the interfering wavefront which ultimately produces noise in the grating.

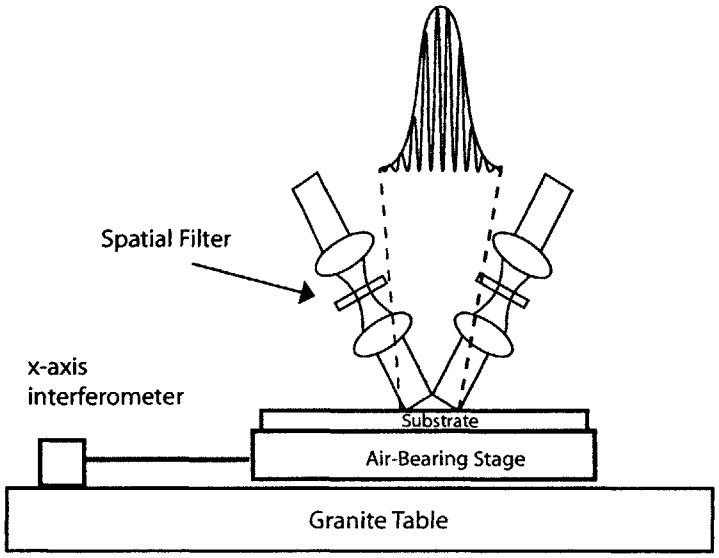

Figure 1-3: *SBIL concept showing image formation and beam path through a focusing lens, spatial filter, and collimating lens. An air-bearing stage is scanned on a granite block to expose a substrate. The stage position is monitored using X and Y axis interferometers.*

The pioneers of SBIL, Mark Schattenburg, and previous PhD students Carl Chen ,and Paul Konkola ,eveloped scanning beam interference lithography to address these issues. Scanning beam interference lithography uses a small diameter beam ($\approx$ 1mm) which traverses small commercially available optics as shown in figure 1-3. An air-bearing stage (Model: Microglide T300L) designed by Anorad Corporation (Haup-

page, NY) is used to scan the stage using frictionless magnetic forces. The stage is mounted on a precision lapped ultraflat granite base. The specification of the flatness of the stage travel is ± 1.5 microns over the full range of travel [4]. In addition, an active vibration isolation system designed by Integrated Dynamics Engineering is provided to reduce vibrations of the stage during scanning. Also shown in figure 1-3 are two incident beams which traverse symmetrical optical paths. The gaussian beams traverse a focusing lens which focuses each beam to a Fourier plane. A spatial filter is placed in the Fourier plane to remove any high spatial frequencies. A subsequent collimating lens is used to re-collimate the beams. On the substrate both beams overlap resulting in a Gaussian intensity envelope with carrier fringes. The carrier fringes inside the intensity envelope are not drawn to scale. Typically, the period is set for a 574 nm period by controlling the angle of the incident beams. The period of the resulting intensity is given by $P_f = \lambda/2\sin\theta$ where θ is the half angle between the beams. For a 2 mm diameter beam and a 574 nm period, there are over 3000 fringes within the beam diameter.

Large substrates (on the order of $\approx$ 300x300 mm^3) are exposed by scanning a stage in either a parallel scan as shown in figure 1-4 (a), or in a Doppler scan fashion as shown in figure 1-4 (b). The stage trajectory is shown in red in figure 1-4. Parallel scan is named as such because the stage is scanned parallel to the interference fringes (along the Y axis). The stage is stepped over at the end of the scan, in a direction perpendicular to the interference fringes (i.e., along the X axis) off of the substrate in order not to smear the fringes on the substrate. The scan is then repeated in the opposite Y direction in order to overlap adjacent scans. In Doppler scanning, the stage is scanned in a direction perpendicular to the interference fringes (along the X-axis). The stage is then stepped over (along Y) at a location off of the substrate and the scan is repeated in the reverse (X) direction. A uniform average exposure dose is developed by overlapping adjacent scans in both cases. In Doppler scanning, the fringes must be synchronized with the stage in order not to smear the image transferred onto the photo-resist covered substrate.

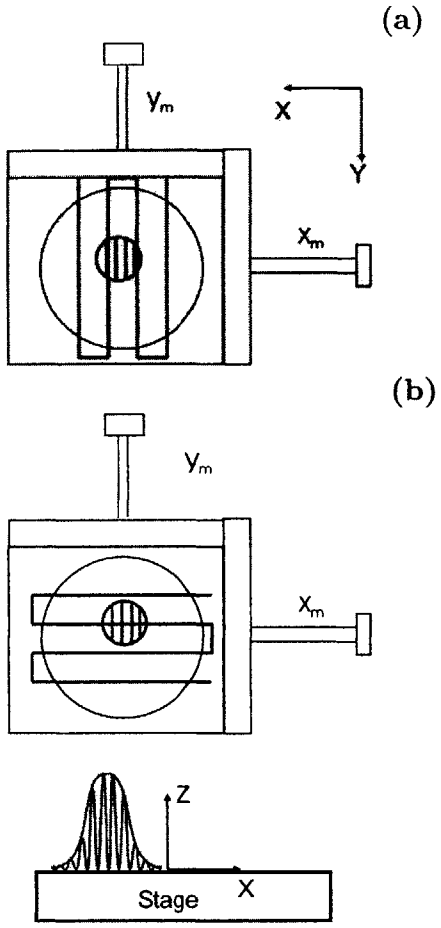

Figure 1-4: *Stage scanning methods (a) Parallel scanning in which the stage is scanned parallel to the interference fringes on the substrate. The stage scan direction is shown in red. (b) Doppler scanning in which the stage is scanned perpendicular to the interference fringes on the substrate. In both cases, the stage steps over in a perpendicular direction to the scan off of the substrate. The scan is then repeated in the reverse direction in order to overlap adjacent scans.*

1.3 Heterodyne Phase Modulation

The heart of SBIL resides in its control architecture. One anticipates that there is significant error at the nanometer level in the stage motion. The key to avoiding printing these errors involves using high-precision interferometers to monitor the stage error. Once this error is known, the interference fringes may be modulated to correct for the stage error. In this way, the interference fringes are said to be stationary in

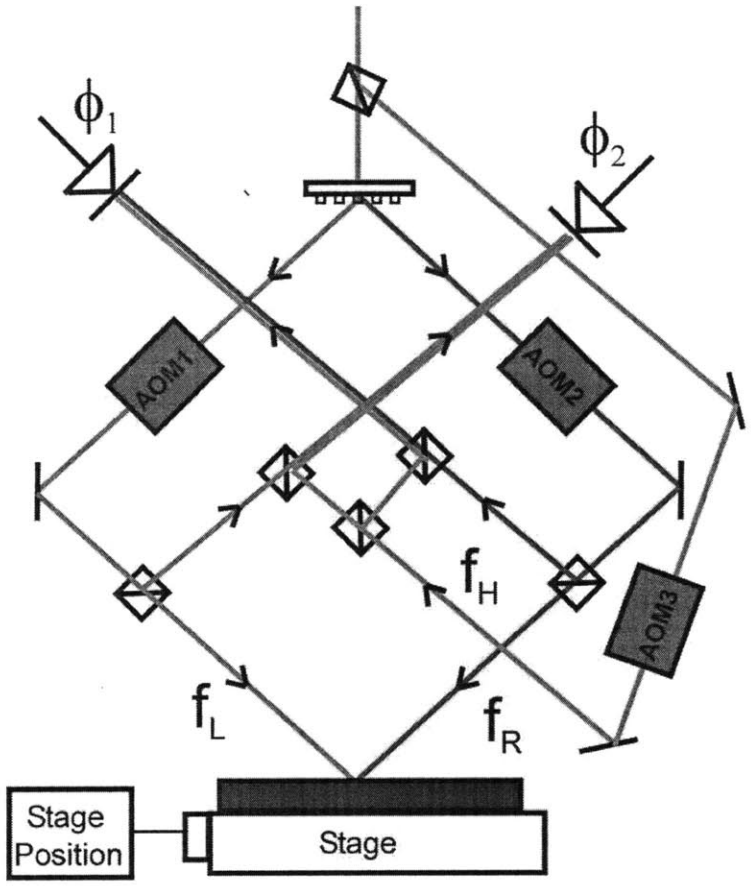

Figure 1-5: *Scanning beam interference lithography writing mode.*

the stage reference frame. In order to implement the interference fringe motion, one requires a method of modulating the interference fringes. The SBIL architecture uses acousto-optic modulators to phase modulate the left and right arm beams. The left and right arm refers to the left and right beam paths which make the interference pattern. Acousto-optic modulators are devices which act as transducers to effectively convert a radio frequency (RF) wave into a refractive index modulation in a crystal. By changing the frequency of the RF wave the period of the index modulation may be changed. The period is given by $\Lambda = v/f$, where v is the velocity of the RF wave in the crystal and f is the frequency of the wave. The index modulation may then be thought of as a grating which is traveling at the same velocity as the RF wave. This moving grating gives rise to a doppler frequency shift of the diffracted wave [21].

The end result is that the first order diffracted wave from an acousto-optic modulator obtains a frequency shift, or a phase modulation, which may be controlled by changing the RF input. In this manner, the interference fringes may be synchronized with the stage motion so that they are stationary in the stage refererence frame. Figure 1-5 illustrates this principle. Shown are three acousto-optic modulators (AOMs) which modulate the left arm with frequency f_L, the right arm with frequency f_R, and the heterodyne beam with frequency f_H by changing the RF input frequency into AOM1, AOM2, and AOM3 respectively. The nominal modulation frequencies are $f_L = 100$ MHz, $f_R = 100$ MHz, and $f_H = 120$ MHz. These modulations frequencies are of course small with respect to the optical frequencies (several hundred THz) in the beams given by $f = c/\lambda$, where $\lambda = 351$ nm is the wavelength of light and $c = 3\text{x}10^8$ m/s is the speed of light in air. However, when the left arm and the heterodyne arm interfere at phase detector ϕ_1 the resulting measurement signal contains the frequency difference $f_H - f_L$. In this way, one may recover the phase difference between the right arm θ_R and the heterodyne arm θ_H by observing the output of phase meter ϕ_1 in figure 1-2 given by $(\phi_1 = \theta_R - \theta_H)$. Similarly, phase meter ϕ_2 in figure 1-2 measures the phase between the left arm θ_L and the heterodyne arm as $\phi_2 = \theta_L - \theta_H$. Subtracting the two measurements $\phi_1 - \phi_2$ then allows us to obtain the phase difference between the left and right arms given by $\phi_1 - \phi_2 = \theta_R - \theta_L$.

In parallel writing, the stage travels along a direction parallel to the interference fringes in the Y direction. In order to overlap adjacent scans the the stage is programmed to step over an integer number of grating periods in between scans. In this case, a stage error along the X fringe direction, x_{error}, leads to a phase error in the grating. The phase of the interference fringes are changed to compensate for this stage error. This is done by modulating the phase difference between the arms such that

$$\theta_R - \theta_L = \frac{2\pi}{P_f} x_{error}, \tag{1.1}$$

where $P_f = 574$ nm is the fringe period. The concept of correcting for the stage error is clearly shown in the error equation 1.1. That is, in the case that the stage error

is zero, and the stage properly steps over an integer number of periods in between scans, the phase modulation goes to zero. In this case, the fringes of the new scan overlap with the fringes of the previous scan and optimal contrast is obtained.

During Doppler writing, a similar argument applies. The main difference, however, is that during Doppler writing the stage position is moving along the X direction during a scan. The fringes are then simply modulated to follow the measured stage X position. Explicitly, the fringe locking condition is expressed as

$$\theta_R - \theta_L = \frac{2\pi}{P_f} x_m, \tag{1.2}$$

where x_m is the measured stage position along the X-axis obtained from the output of our X-axis interferometer. Ideally, implementing Doppler writing should involve modifying the parallel writing error equation given by equation 1.1 and programming a new scan direction. Another consideration needed in achieving high contrast images during Doppler scanning involves the effect of a subtle time delay in the electronics. Since the stage typically achieves a constant velocity v during a scan in the range of 10-100 mm/s, the fringes must move at a frequency of 17-174 kHz. A time delay in the electronics leads to a phase delay in the fringes. This will be described in greater detail in the chapter on novel writing techniques (chapter 2). In the next section, I will describe a powerful feature in the nanoruler. The nanoruler also uses a heterodyne technique to read the phase of a written grating.

1.3.1 Reading Mode

The nanoruler's heterodyne reading mode scheme shown in figure 1-6. The left and right arms, modulated by frequencies $f_L = 90$ MHz and $f_R = 110$ MHz ,respectively, are combined in phase meter ϕ_3. The frequencies of the left and right arms are chosen to give a 20 MHz frequency difference in the interference signal collected on phase meter ϕ_3 [25]. These beams are split off before they interact with the grating surface, and therefore do not contain any phase information about the grating. Rather, phase detector ϕ_3 measures the phase difference between the right and the

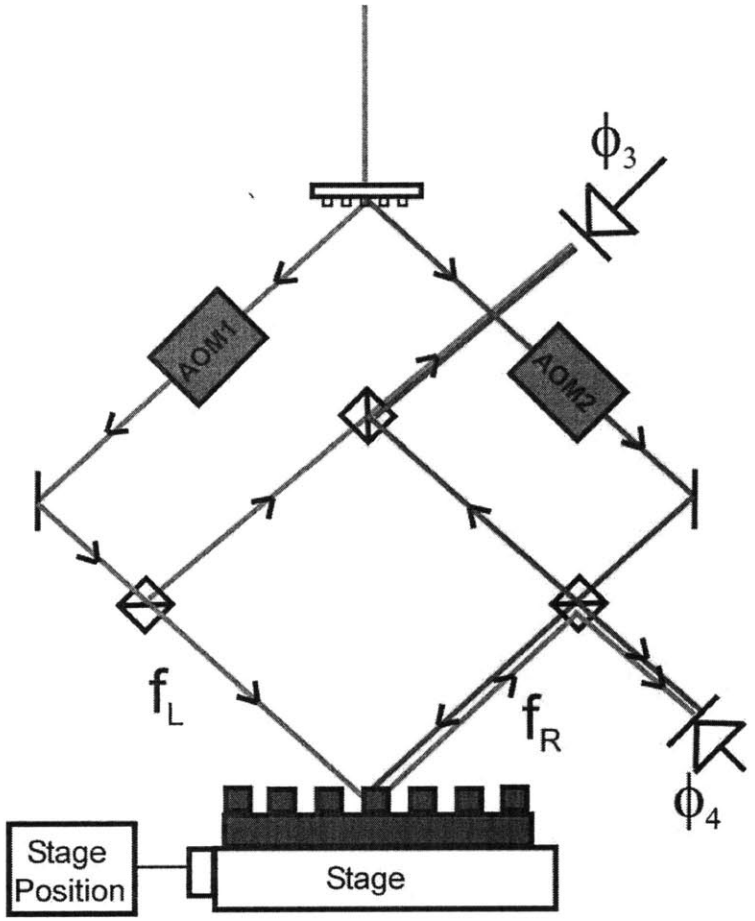

Figure 1-6: *Scanning beam interference lithography reading mode.*

left arms ($\phi_3 = \theta_R - \theta_L$). On the other hand, the reflected left arm from the grating and the back-diffracted right arm are collected into phase meter ϕ_4 . This phase meter contains the phase difference between the reflected left arm and the back diffracted right arm. The back diffracted right arm contains the phase information of the grating we seek along with the phase of the reflected left arm θ_L^r and incident right arm θ_R. The phase detected at phase detector ϕ_4 therefore contains $\phi_4 = \theta_R + \theta_G - \theta_L^r$ where θ_G is the phase of the grating. The phase of the grating is then recovered by taking a difference between the output of the two detectors as in $\phi_4 - \phi_3 = \theta_G$.

This assumes that the phase of the back diffracted right arm θ_R^d is equal to the phase of the incident right arm θ_R plus the phase of the grating as expressed by the

relation $\theta_R^d = \theta_R + \theta_G$. This also assumes that the phase of the incident left arm θ_L is equal to the phase of the reflected left arm θ_L^r. This reading mode ability will be discussed in greater detail in chapter 5. The original reading mode configuration proposed by Heilmann, and implemented by Konkola will be used in combination with an algorithm I developed to determine the grating phase in the presence of a system distortion. The original reading mode configuration, as described in this section may be used to recover the phase of a grating in an ideal system. However, a real system will contain non-ideal characteristics such as Abbe error, reference mirror distortion, and substrate expansion (due to thermal effects and vacuum chuck forces). How to recover the grating phase in a distorted system will be the highlight of my contribution. In addition, the reading mode system as described requires a specific grating orientation. Namely, the grating must be oriented with its lines parallel to the Y-axis. A dual pass reading mode technique which I propose will also be presented in chapter 5 which will allow one to read the grating in a rotated position. In addition, the surface map of a grating substrate may be recovered using this approach.

The grating phase may be read across the grating by scanning the stage. In the following section, I will describe how the stage interferometer works.

1.3.2 Stage Measurement Interferometer

The stage position must be measured with high resolution in order to achieve repeatability at the nano-meter level. This is accomplished by using Zygo's linear/angle interferometers (ZMI 2000 Series). A cartoon illustrating how the position measuring interferometer works is shown in figure 1-7. Zygo's 2000 series of interferometers use a heterodyne scheme where two cross polarized beams (S and P polarizations) with modulated frequencies (f_1, f_2) exit a Helium-Neon laser. The two beams enter a polarized beam splitter together and seperate upon exiting into two seperate paths. A particular polarization associated with frequency f_1, the P polarization undergoes reflection at the beam splitter interface and reflects from a reference mirror surface. Upon reflection from the reference mirror, it passes a retarder of $\lambda/4$ for a second time which effectively rotates the polarization of the beam by 90 degrees resulting in a

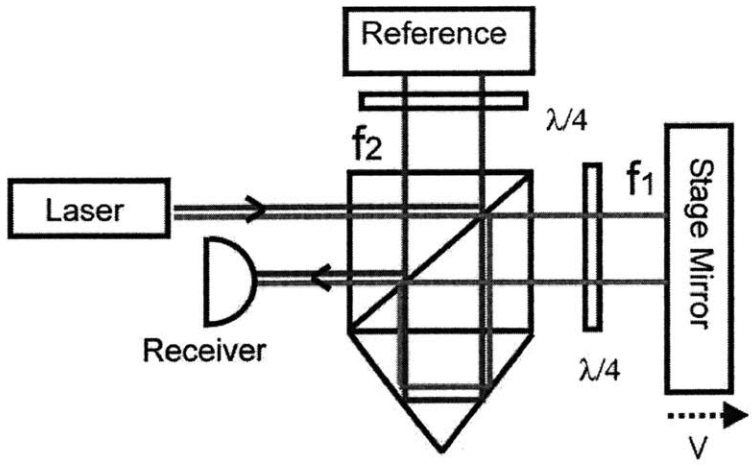

Figure 1-7: *Position Measuring Interferometer (Zygo Corporation [1]). The output of a helium neon laser includes two cross-polarized beams modulated at frequencies f_1 and f_2. Red corresponds to the beam path path of the reference arm modulated with frequency f_1. Shown in blue is the measurement arm, modulated with frequency f_2.*

beam of modulation frequency f_1 with a S polarization. This new polarization allows for the beam with frequency f_1 to transmit through the polarized beam splitter as it proceeds to the optical element which translates and retroreflects the beam. It then undergoes transmission through the polarized beam splitter for a second time. Upon reflection from the reference mirror, for the second time, it changes its polarization once again as it passes twice through the optical retarder. This returns the beam once again to its original P polarization allowing it to reflect from the beam-splitter and into a receiver.

The other beam with a frequency modulation of f_2 and a polarization of S transmits through the polarized beam splitter and reflects from the stage reference mirror. Similarly, it changes polarization after passing the $\lambda/4$ retarder and becomes P polarized allowing for it to reflect at the beam splitter interface. The beam then gets translated and retroreflected. It once again reflects from the beam splitter interface and reflects back from the stage mirror for the second time. This time it changes its polarization to S and transmits through the beam splitter where it is collected at the receiver along with the other beam of frequency f_1.

While this level of detail is provided in the Zygo manual [1], it is included in this introduction as a reference for how the stage interferometer works. The point to be taken away from this cartoon is that the beam modulated with frequency f_2 which reflects from the stage mirror does so twice. This is known as a four pass scheme, since each pass is a round trip and therefore doubles the path length. By viewing the stage or reference mirror, we will see two spots on each mirror associated with the position measurement.

We will return to the stage measuring interferometers in chapter 3 where we will discuss coordinate system errors. At that point, an additional two beams (not shown in figure 1-2) will also be shown in the stage and reference mirror, however, we will not provide a beam trace for these additional beams. These additional two beam paths are used for an angle measurement. We will utilize both the position and angle measurements in chapter 3 to measure the distortion in the stage reference mirror. It is noted that the optical layout used for the angle measurement is similar to the one shown above and is well documented in the Zygo manuals. Rather than repeating a schematic of the layout, the reader is referred to the appropriate reference. The final note in closing of this section is that the resolution of all of the phase detectors used in this chapter correspond to a least significant bit resolution of $\frac{2\pi}{512}$. For a wavelength of $\approx$ 633 nm, this corresponds to a resolution of $633/(512x4) = 0.3091$ nm, where the factor of 4 is a result of the 4 pass configuration. However, the writing and reading mode interferometers use a single pass arrangement with a UV wavelength. In that instance, the least signficant bit resolution on the phase detector electronics is 351/512 nm = 0.6855 nm. With all other considerations aside, this places a limit on the repeatability due to the resolution of our electronics.

1.3.3 Refractive Index Correction

Turbulence and temperature fluctuations change the refractive index of air. The nanoruler minimizes the error due to refractive index fluctuations by enclosing the system inside an environmental enclosure (designed by Control Solutions) which controls the temperature to ± 10 mK. The environmental enclosure is also placed within

a clean room environment. In addition, a refractometer is used to monitor the refractive index fluctuations. Paul Konkola et al. implemented the refractive index control which essentially entails using an interferometer which monitors a stationary path as a method to measure refractive index fluctuations $\Delta n(t)$ [26, 27]. The refractometer provides a measurement θ_n for $t > 0$ expressed as

$$\theta_n(t) = \frac{2\pi L}{\lambda}[n + \Delta n(t)], \tag{1.3}$$

where L is a constant which is a measure of the distance between the two interferometer arms (the reference and measurement arm). At time $t = 0$, when the measurement begins $\theta_n(0) = \theta_{n0} = \frac{2\pi nL}{\lambda}$, where n is the initial refractive index. The change in refractive index $\Delta n(t)$ relative to the refractive index at the start of the measurement n is then found to be

$$\Delta n(t) = [\theta_n - \theta_{n0}]\frac{\lambda}{2\pi L}. \tag{1.4}$$

In general, the stage error due to refractive index fluctions will be of the form

$$x_{error} = \Delta n(t)[x_m - x_0], \tag{1.5}$$

where x_m is the measured stage position and x_0 is a constant offset. Physically, the offset x_0 correspond to the position of the stage where the so called "deadpath" is zero. The deadpath is a term given to the uncompensated path length in an interferometer which is subject to index fluctuations. For example, when the reference and measurement arms pass through the same path length in a uniform medium the deadpath is zero. In order to calculate equation 1.5, one only needs to solve for two unknowns L in equation 1.4 and x_0 in equation 1.5. Paul Konkola devised an experiment which allows one to use reading mode to determine the unobservable error in SBIL. This is done by reading the phase of the grating while the stage is held stationary. Since the stage position is known using stage interferometers, and the grating phase is assumed to be linear over small distances, the noise in the grating phase measurement may be determined and is referred to as the unobservable

error. By applying a least-squares fit to the unobservable error, one may solve for the constants L and x_0. Paul Konkola found that refractive index correction will compensate for approximately 10 nm of error incurred over an hour period [26]. In the next section, I will describe the software and hardware in SBIL which takes all of the inputs previously described to fringe lock the interference fringes to the stage.

1.4 SBIL Hardware and Software Implementation

The core of scanning beam interference lithography resides in the software. It would be difficult to describe the software in any great level of detail because of its complexity. However, a basic understanding of how the software works will already provide us with much insight. Here I will present a basic overview of the software architecture, and I will try to highlight some of my specific contributions to the software development.

The basic platform for the software to be described was written by Paul Konkola and Carl Chen. I have built off of their platform and have made several contributions, some of which I will make a note of at the end of this section. Carl's software mostly used a Labview environment for non-time critical tasks such as automated beam alignment, CCD camera capture, and data analysis of intensity measurements to deduce the fringe period. His software resides on a computer, which I refer to as the Labview computer.

Paul Konkola's software was written in a code-composer C-environment. This software is used to communicate with an Ixthos Champ C6 VME board which contains four Texas Instruments TMS3206701 DSPs. Out of these three DSPs, three are currently being used. This board resides on a VME bus and is the master to several other VME input/output boards to be described. Once the C-code is compiled using Texas Instrument's compiler (Code Composer Studio), the code is downloaded to the DSPs using a JTAG interface. The code is then loaded and ran off of each DSP. The DSPs run at an internal 167 MHz speed, and each contains 16 megabytes of memory. The C compiler environment refers to these three DSPs as DSP A, B, and D. We

prefer to use more descriptive names. Since DSP B is used to pass user commands,we refer to it as the user-interface DSP.

DSP A, known as the real time DSP, contains all of the real time code which computes all of the necessary values to close our feedback loop. These computations include reading the nine inputs from the other VME devices (9 variables). Namely, these include the x and y inteferometer measured position, the refractometer measurement, the x and y stage angle interferometers measurements, the writing mode phase detector outputs ϕ_1, ϕ_2, and the reading mode outputs ϕ_3, ϕ_4. This loop speed is currently 10 KHz, and is determined by an integer multiple of the stage interferometer's laser clock frequency. Accordingly, every 0.1 ms all of the various interferometer inputs are triggered and available. The 10 KHz laser clock also generates a hardware interrupt (interupt 5) on DSP A which causes it to service an interrupt routine (internally known as the interrupt 5 routine). This then computes a phase correction to be sent to a frequency synthesizer which generates the RF phase modulation necessary to drive the acousto-optic modulators. In addition to interrupt 5, which is where all of the real time code is stored, an additional interrupt (interrupt 4) is used to monitor errors. If the stage hits a limit switch, or if the phase meters report any errors, a digital change of state board (VME-VMIC-1181) generates a VME hardware interrupt 4. DSP A then enters a interrupt 4 service routine and further investigates the cause of the error. Recently, I have added a Labview trigger which also shows up as a change of state on the VME-1181 and generates an interrupt 4 hardware interrupt. This is used for handshaking and transferring information to the Labview computer. Previously, communication did not exist between the two seperate systems. This has allowed further automation of routine tasks, and has also allowed one to implement a phase shifting interferometry algorithm discussed in chapter 4 on contrast.

DSP D, the data exchange DSP, stores all of the data which is collected during reading mode. This data is then transferred to an external computer through the JTAG interface so that it may be processed. I wrote an additional layer of software in visual basic to make this data transfer process user friendly. Visual basic is truly a user friendly environment which due to its simplicity provides for a fast turn around

time when programming graphical user interfaces. Visual basic manipulates the code composer studio environment by using "COM Object Programming" features in Microsoft Windows. For more information on how this is done, the interested reader is referred to Texas Instruments Technical Support Knowledge Base (available online) for code composer studio.

While the handshaking signals between the Labview system and the DSP environment is controlled by an interrupt, the actual data transfer is done using a VME digital input board. The VME-VMIC 2510 board provides the digital input/output to the Labview environment. In addition, this board also serves the function of providing a digital output to the frequency synthesizer which drives the AOMs. This digital word may be used to change the phase or amplitude of the RF power to the AOMs.

The other boards on the VME bus consist of four ZMI-2002 series electronic phase boards (manufactured by Zygo Corp [1]). These detect the output of the stage x and y position/angle interferometers, the writing mode signals ϕ_1 and ϕ_2, and the reading mode signals ϕ_3 and ϕ_4. In addition, a Zygo ZMI-2001 series board is used to monitor the refractometer input.

My contribution to the software development in SBIL comes in the form of implementing Doppler writing, as described previously. In addition, I wrote a Visual Basic interface to add user-friendliness and to import/export data to the DSPs. This is particularly useful for importing phase map corrections to be described in chapter 2. Another major contribution comes in the form of absolute phase correction and beam blanking. Previously, due to the design of the ZMI-2002 series of electronics turning on and off the beams would force one to lose the "absolute phase" of the beams. Essentially, this means that the phase differences in writing mode are known up to an arbitrary constant. This is not a problem for writing entire wafers. However, if one wishes to pattern wafers while blanking the beams it is necessary to recover this phase information in order to obtain high contrast. I developed a method to re-establish the phase of the beams and to blank the beams during writing by using lookup tables during writing and by taking advantage of a phase diagnostic register

in the ZMI-2002 boards.

Finally, I created a Labview/DSP communication channel using a VME interrupt as a handshaking signal and a VME digital I/O board for data transfer. This is critical for developing the phase shifting interferometry algorithms which require synchronization of the real time DSP to change the phase of the interference fringes with a Labview driven capture of CCD frames. In addition, this is particularly useful for automating certain tasks and for merging the software written by Chen and Konkola on two different platforms.

1.5 Roadmap to this thesis

In chapter 2 we will discuss general patterning techniques used in SBIL. I have implemented a novel scanning technique in SBIL known as Doppler scanning. This allows one to pattern a grating using an interference pattern while scanning a stage perpendicular to the interference fringes. This method requires synchronization of the interference fringes with the stage motion. A particular advantage of this technique involves the speed and ease with which one may write certain patterns. For example, suppose a pattern has an area W x L, where W is the width of the pattern in a dimension perpendicular to the interference fringes and L is the length of the pattern parallel to the interference fringes. Patterning by Doppler writing will requires less scans and be in general faster if $W > L$. Another advantage of Doppler writing involves the ease with which we may tune into a desired period. It will be shown in chapter four that one may choose a desired period in parallel scanning by stepping the stage over by an integer number of grating periods. In Doppler scanning, on the other hand, one only needs to change the phase modulation of the beams while scanning. This will have important implications for the next generation of scanning beam interference lithography referred to as the variable period system. Finally, if one considers other patterns which consist of a chirp along a direction parallel to the interference fringes Doppler writing will be a superior choice. We will discuss this type of pattern in greater detail in the section on Doppler writing.

In addition to the aforementioned Doppler writing scheme, I have developed a method to blank the beams in order to implement amplitude maps while patterning. While conceptually simple, the challenges involves the use of synchronizing high speed digital electronics to synchronize beam blanking, re-establish phase, and to change the amplitude of the beams while the stage is in motion. This allows for writing different patterns on different portions of a substrate. In the section on contrast, we will show that changing the amplitude of the beams will also allow for changing the linewidth of the grating. This will allow for writing different patterns on the same substrate with different diffraction efficiencies. In addition, I have developed a method to implement two dimensional phase maps which may be used for phase correction, or phase modulation. This requires transfering two dimensional phase information to memory locations which are addressed during scanning. The phase information is then sent to the frequency synthesizer which drives the acousto-optic modulators that ultimately change the phase of the interfering beams which pattern the substrate.

The nominal grating period in this thesis is 574 nm. We would like to fabricate gratings to have a phase accuracy limited only by the repeatability of the system ($\approx \pm 3$) nm. Toward this end, this thesis will address achieving this goal by identifying the major sources of error which are present in scanning beam interference lithography and developing techniques to measure these errors. These error sources include coordinate system error which may be further divided into Abbe error and reference mirror distortion which are addressed in chapter 3. A novel technique I developed will be presented to characterize the mirror distortion in our stage reference mirrors.

In addition, the sources of linewidth variation are discussed in chapter 4. In the last section of this thesis, we will devise characterization techniques which may be employed to characterize the grating distortion in SBIL. These tests are similar in scope to methods used in the optics industry for testing wavefronts and optical components. However, the novelty involves the mathematical analysis done in the context of discrete signal processing and the application for characterizing grating

distortions in SBIL using the nanoruler's unique reading mode ability.

Chapter 2

Novel Patterning Techniques

In this chapter, I will present novel techniques which I developed to pattern gratings in Scanning Beam Interference Lithography. These techniques will include a Doppler writing scheme as described in the introduction (Chapter 1). In addition, a beam blanking and absolute phase technique will be described along with implementation of 2-dimensional phasemap correction. We will begin this chapter by describing Doppler writing in great detail[31].

2.1 Doppler Writing

The Scanning Beam Interference Lithography (SBIL) system, see figure (2-1 a), includes an X-Y air-bearing stage onto which a resist-covered substrate is placed and held down using vacuum forces. A stationary interferometer attached to the optical bench and supported by the granite air-bearing block provides a grating "image" on the substrate. In order to expose large areas on the substrate, the stage is scanned using two different approaches as shown in figure 1. In both methods the interference fringes are held stationary with respect to the substrate using fringe locking electronics and acousto-optic modulators [25]. The first method (figure 2-1b) consists of a parallel scan, in which the stage scans along a direction parallel to the interference fringes. At the end of the scan, at a location outside of the substrate, the stage is stepped over by a distance of approximately half a beam diameter or less. This allows

the new scan to overlap with the previous scan and produces a uniform dose profile [6].

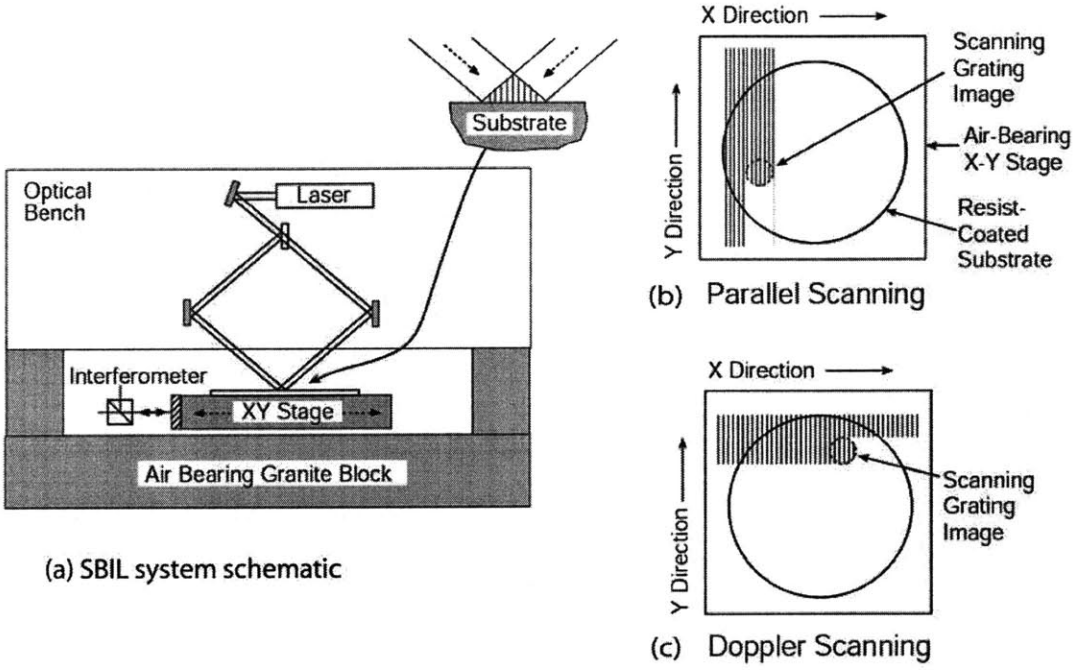

Figure 2-1: (a) SBIL schematic showing X-Y stage and stage interferometer SBIL uses two scan methods to expose a large resist-covered substrate. In both methods, the interference fringes are held stationary with respect to the stage using phase-locking electronics. (b) Parallel scanning requires the stage to scan in a direction parallel to the interference fringes. The stage steps over less than half of the beam diameter off of the substrate at the end of a scan. A new strip is then written in the opposite direction which overlaps with the previous strip. (c) In Doppler scanning, the stage scans perpendicular to the interference fringes. The stage then steps over parallel to the interference fringes at the end of a scan.

In this chapter we present novel results on the implementation of Doppler scanning. In Doppler scanning, the stage is scanned along a direction perpendicular to the interference fringes (see figure 2-1c)). High contrast gratings require the fringe locking algorithm to synchronize the interference fringes to the stage motion perpendicular to the fringes. For example, consider an X-Y coordinate system which is defined by

the reference frame of the granite block shown in figure 2-1a. Assume the interference fringes are aligned parallel to the Y axis. If the stage moves perpendicular to the fringes (X direction), and the fringes are held stationary with respect to the granite, smearing of the fringes occurs on the exposed photoresist. On the other hand, if the interference fringes are phase-locked with the X stage motion, the interference fringes move with the substrate along X and a high-contrast grating is produced in the photoresist. We will derive this result in chapter 4 where we will discuss contrast in greater detail.

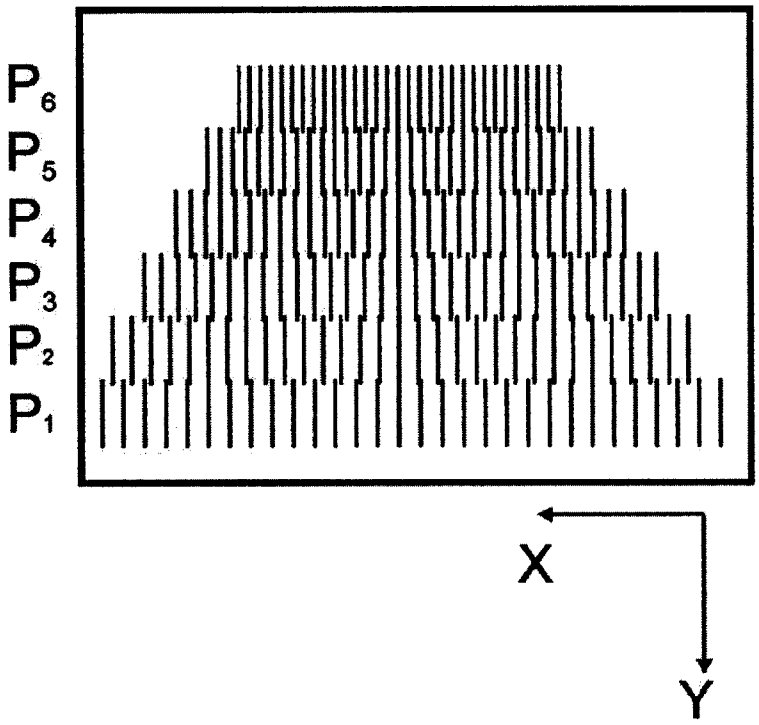

Figure 2-2: Geometry well suited for Doppler writing

It would not be correct to say that Doppler writing is in general superior to Parallel writing, because there are situations in which each offers a particular advantage. Parallel scanning offers a continous scan along the Y with discrete steps along the X axis. Doppler scanning on the other hand, offers discrete steps along the Y axis, but continous scanning along the X axis. Therefore, neither offers an advantage in terms of continous scanning because each technique scans continously in orthogonal

directions. However, Doppler writing offers a unique advantage in patterning in certain geometries. Consider the grating shown in figure 2-2. This pattern consists of constant periods along the X direction. However, as one increases along Y, the grating periods decrease. We may say, loosely speaking, that the grating is chirped along Y although we note that the lines are parallel to the Y axis. In Doppler writing, this particular geometry would be relatively easy to pattern. For example, each period P_i may correspond to an individual Doppler scan. In between scans, the fringe period may be adjusted by changing the angle of the incident beams. The Doppler scan may then be repeated. Since the pattern is chirped along a direction "parallel" to the fringes, this particular pattern would be difficult to write using Parallel mode.

While the latter pattern may at first appear to be quite abstract, it may be useful for a number of practical applications. Several photonic devices may benefit from this type of structure. For example, consider a grating directional coupler device which is described in standard texts on photonic devices [42]. Distribute feedback may be used in these waveguide devices to spatially select certain frequencies for use in optical fiber communication applications. Since the bandwidth of these devices is proportional to $1/N$ where N is the number of periods, long gratings could offer narrow bandwidth filtering. Doppler writing allows us to generate a mask for these devices which include several spatial frequencies. This will allow for coupling out several wavelengths for fiber optical communication applications.

Another application for this structure includes its use in x-ray telescope applications. Shown in figure 2-3 is a diagram illustrating a proposed variable period structure for a NASA x-ray telescope. It has been suggested that a grating with a chirp along the Y direction as shown in figure 2-2 will serve as a good approximation for the structure in figure 2-3. The geometry of figure 2-3 consists of a radially chirped grating. Accordingly, the contours of constant period reside on circles of radius ρ. The details of patterning this type of structure will be the thesis topic of another student in our group. However, we mention it here because it is a good motivation for Doppler writing.

For the remainder of this section, we will describe the details of implementing

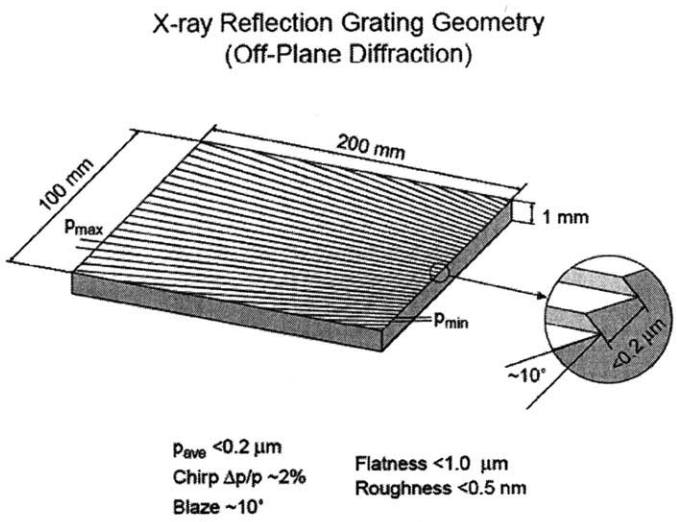

Figure 2-3: Variable period structure for x-ray telescope application

Doppler writing. We will return to Doppler writing in chapter 4 where we will discuss contrast issues. In chapter 4, we will show that the period which is printed in a grating is determined by the fringe modulation in Doppler scanning, whereas it is determined by the stage step-over distance in parallel scanning.

Implementing Doppler writing requires synchronous motion of the interference fringes with the stage of our system. In order to accomplish this motion, high speed electronics are required. Here we will discuss the time delay associated with the time it takes to update the phase of the interference fringes from the time the stage position is acquired. Compensating for this time delay is necessary to implement Doppler writing. In this context, the time delay should not be confused with the temporal equivalent of an optical path length difference in the interfering arms.

In order to implement Doppler writing our control loop uses five phase-meter inputs. Two interfering arms using a $\lambda_{uv} = 351.1$ nm wavelength produce a grating image on the substrate, as shown in figure 2-4. The left arm with frequency f_L is interfered with a heterodyne beam with frequency f_H in phase meter 1 (PM1), which outputs Φ_1. The right arm with frequency f_R is interfered with the heterodyne

reference in phase meter 2 (PM2), which outputs Φ_2. The remaining three phase meters are necessary for stage control: the stage X-axis interferometer phase meter outputs the x stage position Φ_x in units of 2π phase corresponding to a distance $\lambda_{HeNe} = 632.8$ nm, the Y-axis interferometer outputs the y stage position Φ_y in the same phase units, and a refractometer phase meter which monitors variations in the index of refraction that lead to changes in optical path length. A detailed discussion of the refractometer phase meter is available elsewhere [26] and is described in the introduction to this thesis.

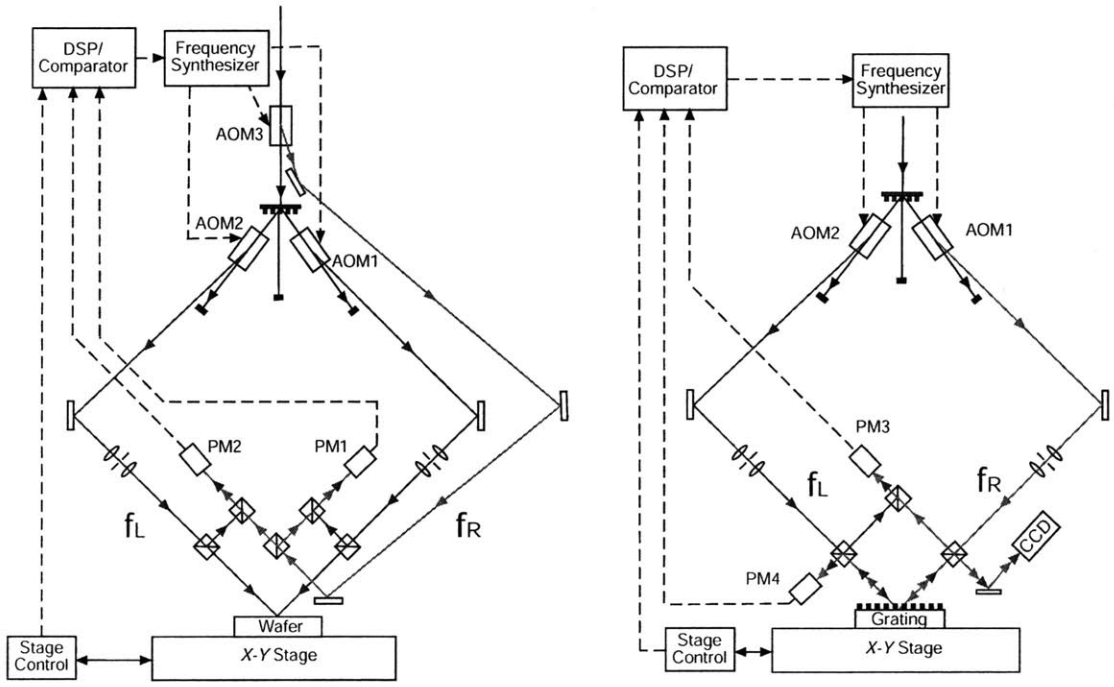

Figure 2-4: Metrology and writing modes used in scanning beam inteference lithography.

The difference between phase meter 1 and phase meter 2 in figure 2-4 contains the phase difference $\Delta\Phi = \Phi_L - \Phi_R$ between the left and the right arms [25]. In order to implement Doppler writing, this phase must be synchronized with the stage motion. This requires significant frequency differences Δf between the arms, where

$\Delta f = \frac{1}{2\pi}\frac{d}{dt}\Delta\Phi$. In order to describe the stage motion, we will define an X-Y reference frame to be that of the granite table on which the stage moves, as shown in figure 2-1.

For example, consider the case in which the image fringes formed by the left and right arms are parallel to the Y-axis and periodic along the X-axis with period $P_f = \lambda_{uv}/(2\sin\theta)$, where θ is the half-angle between the beams. The image intensity pattern will be of the form

$$I(x,t) = A\{1 + \gamma cos[\frac{2\pi}{P_f}x - \Delta\Phi(t)]\}G(x), \tag{2.1}$$

where A is the amplitude of the interference pattern, γ is the image contrast, x is the coordinate along the X-axis, and $\Delta\Phi(t)$ is the phase difference of the left and right arms at time t. Here $G(x) = exp[-8(x - x_0)^2/D^2]$ is the Gaussian envelope of the image, where x_0 is the image center and D is the e^{-2} diameter. During parallel scan mode $\frac{d}{dt}\Delta\Phi \approx 0$, leading to fringes essentially stationary with respect to the Gaussian envelope.

Now consider the reference frame of the moving stage. The stage reference frame X'-Y' is traveling at velocity $v(t)$ with respect to the stationary frame, which we assume is along the X-axis. The moving stage coordinate relative to the stationary reference is therefore given by

$$x' = x - x_s, \tag{2.2}$$

where $x_s = \int v dt$. As an aid in understanding this relative coordinate relation, let us consider an object on the stage located at a coordinate in the stage reference frame x'_0. Now let us assume that the stage is stationary (with respect to the laboratory frame) such that $x_s = 0$. An object at a position x'_0 in the stage coordinate frame may also be described as being at a position $x_0 = x'_0$ in the laboratory frame consistent with our notation in equation 2.2. Now let us allow the stage to move relative to the laboratory frame with a velocity v. An object on the stage is still located at the same position x'_0 in the stage frame. However, in the laboratory frame, this object has moved to position $x_0 = x' + x_s$ due to the motion of the stage. This confirms are

sign convention in equation 2.2.

The intensity in the stage reference frame may then be written by substituting $x = x' + x_s$ into equation 4.45 as written below:

$$I(x',t) = A\{1 + \gamma cos[\frac{2\pi}{P_f}(x' + x_s) - \Delta\Phi(t)]\}G(x' + x_s). \tag{2.3}$$

As the image envelope in the stage reference frame moves with time we want to frequency-shift the left and right arms so that the fringes are stationary in the stage coordinate x'. We accomplish this by shifting the frequency of the two arms using acousto-optic modulators such that

$$\frac{2\pi}{P_g}x_s = \Delta\Phi(t), \tag{2.4}$$

which can be written,

$$\frac{2\pi}{P_g}\int vdt = 2\pi\int \Delta f dt, \tag{2.5}$$

where P_g is the desired grating period. For example, in the constant velocity case, $x_s = v \cdot t$ and the condition of fringe locking becomes $\Delta f = v/P_g$. In order to obtain a high-contrast latent image in the resist, the difference between the fringe period P_f and the grating period P_g must be small otherwise fringe smearing will occur and contrast will degrade. This condition is met when $\frac{\Delta P}{P_g} << \frac{P_g}{D}$ where $\Delta P = P_f - P_g$ and D is the image diameter. This condition is provided here without proof. We will prove this relation in chapter four where we devote our discussion to contrast.

2.1.1 Electronic Time Delay

Early implementations of Doppler writing yielded gratings with poor contrast at certain velocities. This was determined to be due to the time delay in the control loop which results in phase offsets between overlapping scans. This is illustrated by

including a time delay term t_d in equation (2.3), assuming constant velocity, obtaining

$$I(x',t) = A\{1 + \gamma cos[\frac{2\pi}{P_f}(x' + x_s) - \frac{2\pi}{P_g}v \cdot (t - t_d)]\}G(x' + x_s). \tag{2.6}$$

The phase offset between adjacent scans is given by

$$\Phi_d = \frac{4\pi}{P_g}|v| \cdot t_d, \tag{2.7}$$

which increases with the velocity. In general, the above equation can be generalized to a two-dimensional case with the grating arbitrarily aligned with respect to the stage axes and with an arbitrary velocity vector. Note that at certain velocities where the phase offset $\Phi_d = (2n + 1)\pi$, where n is an integer, the contrast in the printed image goes to zero and no grating results. This is because the adjacent scans overlap with a π phase shift. In order to compensate for the phase delay, we must determine the time delay in our system. In the next section, I will describe an experiment which was performed to measure this time delay.

2.1.2 Measuring the Time Delay

We begin by identifying the various sources of time delay in our system. A set of five synchronized phase meters update internal phase registers at a 20 MHz clock rate (50 ns). Every 2000 cycles (100 μs) the phase meters simultaneously latch out their phase values making them available to externally readable registers, and generate a bus interrupt corresponding to a 10 kHz interrupt rate. A digital signal processor (DSP) services the interrupt at time Ti as shown in the timing diagram in figure 2-5.

The phases read by the DSP have a data age on the order of 335 ns as a result of time delays in electronics and digital filters [1, 11]. In figure 2-5, the upper left inset shows the time Ti when the phase meters are latched and the interrupt routine begins. The phases which are latched at time Ti, however, corresponds to phase of the measurement signal at time Tp (335 ns delay). The comb in the inset corresponds to 50 ns events. Every 50 ns a new phase measurement occurs and the phases are

summed and stored in internal phase registers [1].

During the interval from Ti to Tc, the DSP reads the phase registers, calculates the phase error and outputs the phase correction to the digital frequency synthesizer as shown in figure 2-5. At time Ts the synthesizer outputs an RF frequency correction to the acousto-optic modulator (AOM). The propagation delay associated with the RF cable from the synthesizer to the AOM and the velocity of sound in the AOM transducer adds additional time. Finally, at time Ta the AOM phase-shifts the beam and the beam propagates to the wafer. At time $Ti' = Ti + 100\mu s$, the next DSP interrupt occurs and the cycle repeats. The overall time $t_d = Ta - Tp$ is the total time delay of the system measured to be 99.5 μs, which is just a little shorter than the DSP loop rate of 100 μs. Rather than determining the time delay of the various components individually, a series of experiments were used to determine the overall time delay of the system.

In order to measure the time delay of our control loop, a grating is mounted on the stage and the system is placed into metrology mode (see figure 2b). In open-loop metrology mode SBIL uses two phase meters to measure the phase of gratings which have been previously written. Phase meter 4 measures the interference of the left arm back-diffracted off of the grating and the right arm zero-order reflected from the grating, obtaining

$$\Phi_4(t) = \Phi_L(t) - \Phi_R(t) - \Phi_g(t) = \Delta\Phi(t) - \Phi_g(t), \tag{2.8}$$

where Φ_L and Φ_R are the phase of the left and right arm, respectively, and Φ_g is the phase of the grating [25].

Phase meter 3 measures the interference of the left and right arms which are picked off before they reach the grating, obtaining

$$\Phi_3(t) = \Delta\Phi(t) = \Phi_L(t) - \Phi_R(t). \tag{2.9}$$

Subtracting the output of phase meter 4 from phase meter 3 results in the phase of the grating Φ_g. The stage can be scanned in a parallel fashion, or in a Doppler

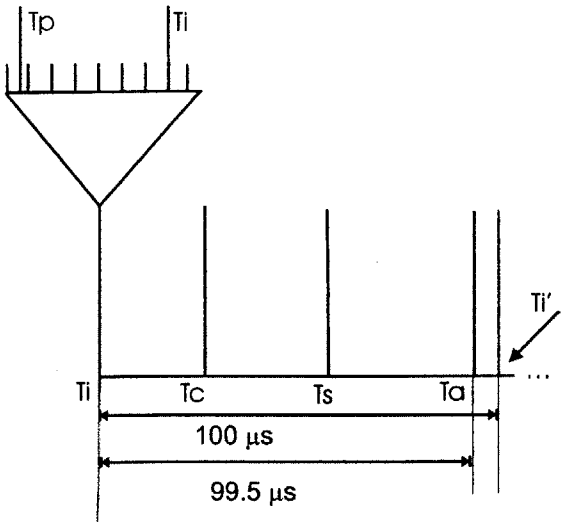

Figure 2-5: Timing diagram of the SBIL control loop. Shown in the upper left is a zoomed-in view when the phase measurement occurs. At time Ti a DSP interrupt occurs. The phases which are available at Ti have a 335 ns delay time ($Ti - Tp =$ 335 ns) due to electronic propagation delay. The tick marks shown on the zoomed-in scale correspond to a period of the 20 MHz phase meter clock (50 ns). At time Tc, the DSP computes a phase correction and it is sent to the frequency synthesizer. At time Ts, the synthesizer outputs RF power to the AOM. At time Ta, the AOM frequency shifts the beam. The overall time delay $Ta - Tp$ is 99.5 μs. At time $Ti' = Ti + 100$ μs the next DSP interrupt occurs and the cycle repeats.

scan mode. We developed a novel "closed-loop" Doppler configuration which allows measurement of the overall time delay of our feedback loop. In closed-loop Doppler metrology mode the stage is scanned perpendicular to the stage and a feedback loop synchronizes the phase of the image grating with the stage motion.

To illustrate this principle, let us confine ourselves to a one-dimensional example where the grating lines are periodic along the X-axis of the stage. Furthermore, we scan the stage along the X-axis. The phase of a constant-period grating in the stage frame is given by $\Phi_g = \frac{2\pi x'}{P_g}$. The phase of the grating in the stationary frame is given by

$$\Phi_g(t) = \frac{2\pi}{P_g} \cdot (x - x_s), \tag{2.10}$$

where x_s is the grating position at time t in the stationary reference frame. In closed-loop Doppler metrology mode, the phase of the left and right arms are controlled by

a frequency synthesizer. The phase difference between the left and the right arms are set by the DSP at time t to be

$$\Delta\Phi_{DSP}(t) = 2\pi\frac{x_s}{P_m}, \tag{2.11}$$

where P_m is a constant to be determined. Taking the derivative of equation (2.11) yields

$$\Delta f = \frac{1}{2\pi}\frac{d}{dt}\Delta\Phi_{DSP} = v/P_m. \tag{2.12}$$

In the absence of a time delay in the electronics, the phase difference between the left and the right arms would be given by $\Delta\Phi(t) = \Delta\Phi_{DSP}(t)$. This simply means that the DSP code which updates the frequency difference between the left and right arms would instantaneously modulate the left and right beams. To account for a time delay, we note that the phase difference between the left and right arms may be expressed as $\Delta\Phi(t) = \Delta\Phi_{DSP}(t-t_d)$. The output of phase meter 4 in the presence of this time delay, from equation 2.8 is given by $\Phi_4(t) = \Delta\Phi_{DSP}(t-t_d) + \Phi_g(t)$. From a Taylor series expansion, we know that we may express $\Delta\Phi_{DSP}(t) \approx \Delta\Phi_{DSP}(t-t_d) + t_d\{\frac{d}{dt}\Delta\Phi_{DSP}(t-t_d)\}$. If we further assume that the stage velocity v does not change over the time delay interval (i.e., the linear expansion is valid), we may express the derivative term in the taylor series as $\frac{d}{dt}\Delta\Phi_{DSP}(t-t_d) = 2\pi v/P_m$ where we have used equation 2.12. Using these relations, we may arrive at an expression for Φ_4 as

$$\Phi_4(t) = \Delta\Phi_{DSP}(t) - \frac{d}{dt}\Delta\Phi_{DSP}(t-t_d)\cdot t_d + \Phi_g(t) \tag{2.13}$$

$$= \frac{2\pi}{P_m}(x_s - vt_d) + \frac{2\pi}{P_g}(x - x_s). \tag{2.14}$$

The first term in equation 2.13, $\Delta\Phi_{DSP}(t)$, is defined by equation 2.11. The second term, $\frac{d}{dt}\Delta\Phi_{DSP}(t-t_d)$ is given by equation 2.12 where we use the assumption that the stage velocity does not change over a time interval t_d. The final term $\Phi_g(t)$ is substituted with equation 2.10 giving rise to equation 2.14. Note that the terms involving x_s cancel if the modulation constant $P_m = P_g$. Note also that the offset in

phase $\frac{2\pi}{P_g}vt_d$ depends on the product of the velocity and the time delay t_d.

With these considerations, we may express the output of phase meter 4 as a function of the position x of the grating in the laboratory frame. That is, $\phi_4(x) = \frac{2\pi}{P_m}(x - vt_d)$. Figure 2-6 shows the results of our closed-loop Doppler metrology mode experiment. The output of phase meter 4, Φ_4, is scaled by $\frac{P_m}{2\pi}$ to give units of distance. A hysteresis curve is obtained as the stage is scanned repeatedly back and forth, starting at rest at point A and accelerating in the positive direction to point B, then traveling with a velocity of +60 mm/s for $\approx$ 60 mm to point C corresponding to the upper portion of the hysteresis curve, then decelerating from point C to point D at rest, then accelerating in the negative direction from point D to point E, then traveling with a velocity of -60 mm/s for $\approx$ 60 mm to point F corresponding to the lower portion of the hysteresis curve, then finally decelerating to rest at starting point A. The offset from the center corresponds to the scaled phase delay, $\Phi_d \cdot \frac{P_g}{2\pi} = vt_d$, in units of length. This phase delay is proportional to the velocity of the stage and the time delay. Once the time delay is corrected for in the control loop, the hysteresis disappears and one obtains the dashed line in the center of the hysteresis loop shown in the figure. Ideally, the line should contain a zero slope. The slope in the line is a consequence of not matching the modulated constant P_m to the grating period P_g, resulting in a moiré beat between the two periods.

2.2 Conclusions

In this section, I have implemented a novel scanning technique in SBIL referred to as Doppler Writing. While the concept of Doppler scanning has been previously described by Schattenburg et al[40], I was the first to implement this writing scheme in SBIL. The challenge involved in implementing Doppler writing requires an understanding of the time delay in the electronics. An experiment was performed in which the time delay was found to be 99.5 μs. Correcting for this time delay in the control loop requires compensating for a phase offset which is a function of velocity. When the measured time delay was corrected for, optimal contrast was obtained in

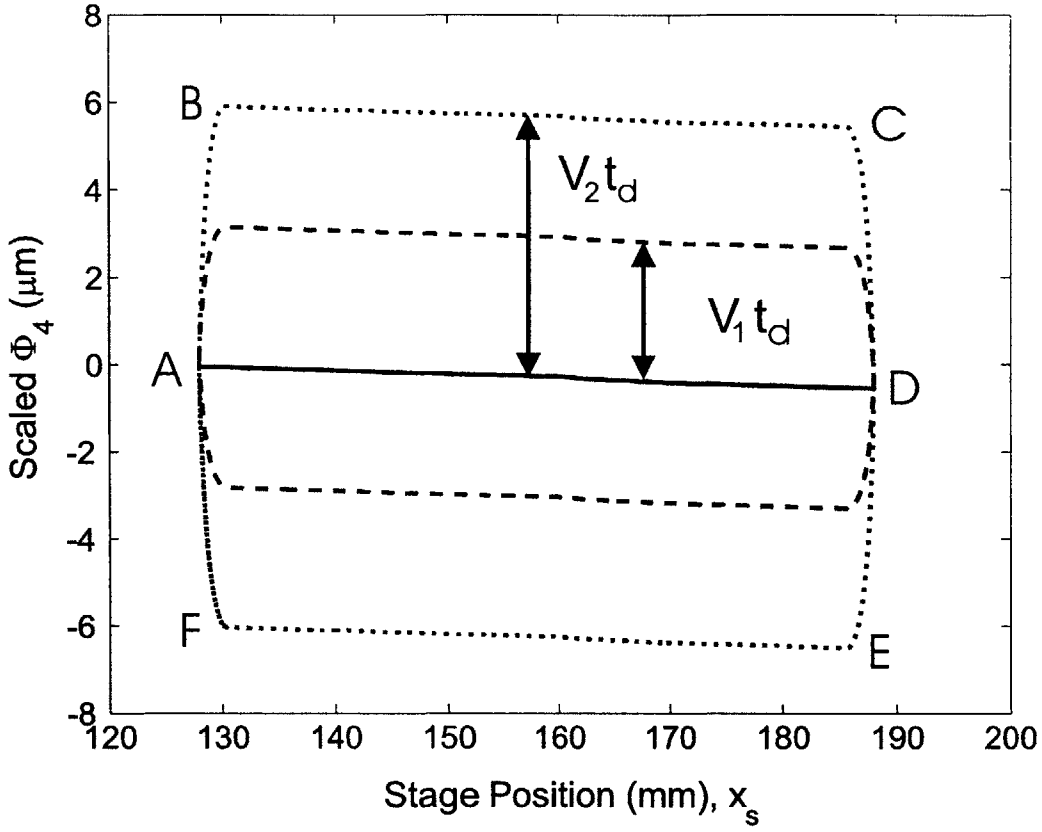

Figure 2-6: Measured hysteresis curve of the Doppler metrology mode experiment used to determine the time delay. The upper portion of the curve (dotted line from B-C) corresponds to a positive stage velocity of v_2=60 mm/s for 60 mm. The bottom portion (dotted line from E-F) corresponds to a negative velocity of 60 mm/s as the stage reverses its scan direction. Similarly, the dashed line corresponds to a velocity v_1=30 mm/s. The offset from the center is proportional to the product of the velocity and time delay. Once the time delay is corrected for, the solid line in the center results in which the hysteresis is removed. The output of phase meter 4 is scaled by $P_m/(2\pi)$ to give units of distance, where $P_m = 574$ nm.

the developed photoresist grating. By changing the time delay from the measured value, contrast degradation occured as measured with the aid of a scanning electron microscope. This gives us further confidence in our measurement of the time delay. It is also noted that prior to correcting for the time delay, certain velocities resulted in gratings of poor contrast. Properly accounting for the time delay solved this problem, making Doppler scanning highly repeatable for all velocities.

In addition, I have described some novel applications for Doppler scanning. Indeed, it shows promise for writing periodic structures with a chirp in a direction parallel to the interference fringes. These structures may be useful for photonic devices which utilize distributive feedback, x-ray telescope applications, and it is my hope that new applications will become available.

2.3 Beam Blanking and Absolute Phase

In this section, I will describe how I blank the beams and re-establish phase during patterning a wafer. Beam blanking refers to turning the beams on and off during patterning. While conceptually simple, synchronizing the stage motion, the intensity amplitude, and the phase locking electronics requires a detailed understanding of the various hardware components throughout the system. While it is not particularly instructive to discuss all of the software details, nor to provide an in depth hardware description of the various components involved, some of the more subtle aspects will be presented here. The software will be made available to future users, and armed with knowledge of the more subtle details presented in this section one should be able to follow it.

The beam blanking problem is broken down into two parts : (1) turning off/on the power to the acousto-optic modulators, and (2) re-establishing fringe locking. The power to the acousto-optic modulators is derived from an RF output of a frequency synthesizer. The frequency synthesizer is controlled by a 32 bit address line and a 32 bit data line. The frequency synthesizer contains three main registers, namely a power register and a frequency register (and a phase register which is not being used). During fringe locking, the frequency register is addressed. In order to turn off the AOM's, a digital word must be sent out on the address lines to select the power register. The zero amplitude must then be sent out on the data lines. This is done by using a digital I/O board on the VME bus (VME 2510B).

Turning the beams back on requires one to write the desired amplitude value to the synthesizer's power registers. Next, the frequency registers must be addressed and

loaded with the nominal frequency values for the left and right arms. Now finally, the phase meter boards must be reset. This is done after verifying that there is sufficient signal strength on the phase meters by checking a status register.

The challenge involves re-establishing fringe locking. A pre-requisite for implementing fringe locking requires a fundamental understanding of how the Zygo phase meter electronics measure the phase. Here we will describe the operation of the phase meters in the context of phase meter 3, which outputs Φ_3 as described in equation 2.9.

The Zygo phase meter electronics are designed to measure the phase of a heterodyne signal. This heterodyne signal has a nominal frequency of 20 MHz. For the purposes of our discussion, the measured signal will consist of ϕ_3 which contains the 20 MHz frequency difference between the left and the right arms. As is common to all methods of phase detection, the phase of this measured signal must be measured with respect to a reference signal. The reference signal to the Zygo phase meter electronics consists of a 20 MHz clock signal which triggers all of the phase meter measurements. There are several specialized circuits available for detecting the phase of a heterodyne signal. These include digital and analog phase-locked loops, zero-crossing detectors, and discrete-signal processing based specialized circuits which compute the FFT of a signal. The latter technique becoming available with the advent of fast analog to digital conversion technology. However, it is beyond the scope of this thesis (and not our design) to discuss the phase-detection circuitry.

During patterning in writing mode or reading in reading mode, the phase output (in our example ϕ_3) is read from a phase register, referred to as a position register in the Zygo literature. Unfortunately, the Zygo 2002 series of phase meter electronics start up with a phase value of zero in the position register.

Ideally, the phase difference between the measurement signal ϕ_3 and a reference signal (usually referred to as the reference clock) would be stored in the position register. While this phase difference is computed, at startup, this phase difference is lost because the Zygo phase meters initiates the position register to zero. From this point of initiation onwards the position register unwraps the phase difference

between the measurement signal and the reference clock. Here what we mean by phase unwrapping is simply that the phase difference is not restricted to be between 0 and 2π. To see why this is important, one only needs to consider the position of the stage. We are interested in the relative stage position and the total distance it has traveled, not the distance modulo the period of the helium neon laser. For the purposes of fringe locking, an unwrapped phase is required because our controller differentiates this phase to generate a frequency word to send to the synthesizer. A wrapped phase (restricted to 0-2π) contains many discontinuities creating problems for differentiation. We would therefore like to use the position register because it contains a filtered unwrapped phase. In order to use the position register and recover the absolute phase, we must calibrate for the phase lost. To be clear, we define the absolute phase of a signal to be the phase difference of the signal between $0-2\pi$ with respect to a reference clock.

We may express the phase read from the position register as

$$\phi_3(t) = \phi_{3abs}(t) + \theta_{offset}, \tag{2.15}$$

where θ_{offset} is a value added to initialize the position register to 0 for $t = 0$. If we can recover this offset, we may subsequently subtract its value from future $\phi_3(t)$ measurements to obtain the signal $\phi_{3abs}(t)$ we seek, that is, the absolute phase $\phi_{3abs}(t)$ defined as the phase difference of the signal into phase meter 3 and the phase of a reference clock at time t.

The trick to recovering the phase difference between the measurement arm and the laser clock resides in using a phase diagnostic register on the Zygo 2002 boards. As a suggestion to the reader, understanding of this section is not necessary for a broad understanding of the work that is to follow in subsequent sections. A detailed discussion is provided here for the next generation of engineers to work on SBIL. The actual code for this section is provided in Appendix C, and a discussion is provided in this section to provide a summary of how the algorithm works.

The phase diagnostic register contains the absolute phase difference between the

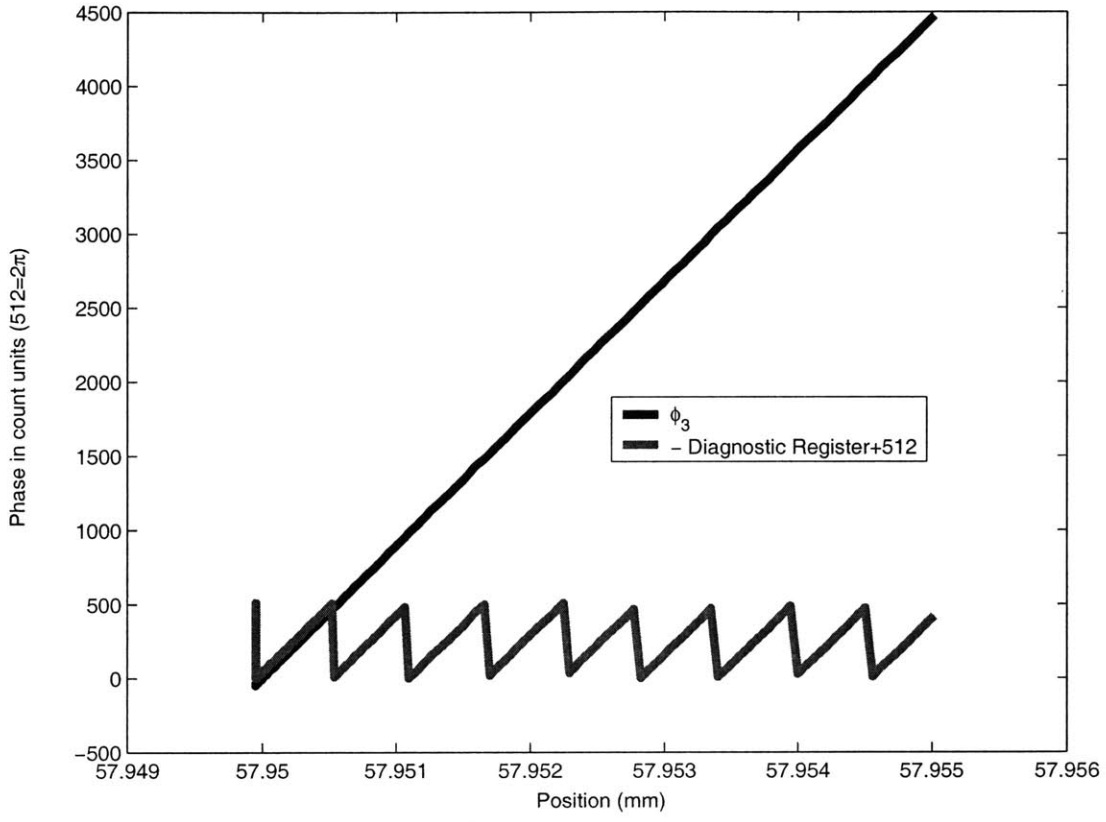

Figure 2-7: *Measured output of the phase meter 3 position register and the phase diagnostic register (scaled by -1 and shifted by 2π) are shown. The units are in counts where 512 counts corresponds to 2π. Note that the output $\phi_{3pos}(t)$ is unwrapped, while the phase of the diagnostic register is wrapped (i.e., restricted to a range of 0-2π).*

measurement arm and the laser clock. So, one simply needs to read this register for each phase meter and the phases can then be compared to a commmon reference. The last consideration one must take into account is that the phase which is read from the phase diagnostic register is the wrapped phase. In addition, the phase diagnostic register has an opposite sign convention from that of the position registers. Accordingly, the phase diagnostic register output $\phi_{3diag}(t)$ has a range from 0 to 2π. This may be considered the modulo 2π value of $-\phi_{3-unf}(t)$, which may be expressed as

$$\phi_{3diag}(t) = -mod2\pi[\phi_{3-unf}(t)], \tag{2.16}$$

where $mod2\pi[]$ is the modulo 2π operator. We have emphasized the fact that the phase diagnostic register is unfiltered. That is, the position register signal $\phi_3(t)$ is a filtered version of the measurement signal, whereas the phase diagnostic register corresponds to the negative modulo 2π version of the unfiltered measurement signal ϕ_{3-unf}.

Now to calculate the phase offset in equation 2.15 we simply perform a modulo 2π operation on equation 2.15 and add equation 2.16. This may be expressed as

$$\theta_{offset} = mod2\pi[\phi_{3pos}(t)] + \phi_{3diag}(t) \tag{2.17}$$

In order to restrict θ_{offset} to be between 0 and 2π we perform a final modulo 2π operation on equation 2.17. A measurement was performed in reading mode. The output of phase meter 3 and the phase diagnostic register are shown in figure 2-7. The phase diagnostic register has been scaled by -1 and shifted by 2π in order to restrict its value between 0 and 2π. The output of both curves are displayed in units of counts, where 512 counts = 2 π. If we wrap the phase of the position register output ϕ_{3pos}, and subtract the scaled diagnostic register the result would be a constant phase offset θ_{offset} which is the absolute phase that we seek.

A final note of interest involves averaging the measurement of θ_{offset}. It is critically important that the phase of the interference fringes are controlled with low phase noise in order to obtain optimal contrast. Since the phase diagnostic register on the ZMI-

Figure 2-8: Beam blanking and absolute phase were implemented to apply an amplitude map to the wafer

2002 series are unfiltered, I have applied averaging of the calculated absolute phase offset. The averaging that is performed corresponds to a sum of all N previously calculated absolute phase values (i.e., $\phi_{3abs}[0] + \phi_{3abs}[1] + ...\phi_{3abs}[N]$) divided by N, where the brackets are used to specify the loop iteration number. As pointed out in the Zygo literature, one must be careful when averaging wrapped phases since the average of 0 and 512 is 511 (in units of counts where 512=2π). In other words, the average of 0 and 2π is 0 rather than π. An example for where this would be a problem is if a measured value for θ_{offset} is close to 2π and contains noise. My strategy for dealing with this involves creating a convention to deal with this scenario. The problem resides in averaging two phase values which reside in quadrant I defined as $0 < \theta_{offset} < \pi/2$ or quadrant IV defined as $3\pi/2 < \theta_{offset} < 2\pi$. A convention is made when this occurs to force all subsequent values to be in one of the two quadrants by adding or subtract 2π. Here is an example.

Suppose a current value is in quadrant IV on the unit circle, and passes our loop criteria such that ($\theta_{offset} < 3\pi/2$). Further suppose that a previously stored value $\theta_{offset-old}$ computed in the last interrupt cycle resided in quadrant I as in ($\theta_{offset-old} < \pi/2$). A decision is then made to force the current value and all subsequent values to quadrant I. This is done by subtracting 2π from the current value forcing it to become negative. This negative number is suitable for averaging since the average of a positive value in quadrant I and a negative value does not pose a problem.

At the end of the loop iteration, $\theta_{offset-old}$ gets updated and becomes a negative value. In the subsequent loop iteration, $\theta_{offset-old}$ will be a negative value which strictly speaking is not in quadrant I. However, our loop criteria which considers all ($\theta_{offset-old} < \pi/2$) as belonging to Quadrant I will treat it as such. In subsequent loop iterations, if the present value resides in Quadrant I, averaging is performed without further consideration. If the present value resides in Quadrant IV, 2π is once again subtracted and the process repeats.

The result of implementing absolute phase is shown in figure 2-8. For this figure, a lookup table was created in memory where an amplitude value is retrieved while scanning. The map for the amplitude was created using a bit map file and transferred into

the DSP memory. The fringe locking was re-established during scanning at a velocity of 40 mm/s. The characters MIT were spelled and the remaining exposed substrate contained uniform high contrast, as visible to the eye. This particular pattern is chosen to portray the capability of the system, however, it is not particularly useful. Applications where beam-blanking would be useful include maskless patterning of a substrate. For example, different sections of the substrate may be patterned with varying linewidths without additional processing steps. This would lead to further study of the relationship between linewidth and diffraction efficiency. This in itself may be useful if one wishes to amplitude modulate a diffracted beam by modulating the diffraction efficiency across a substrate.

Furthermore, absolute phase recovery may be used in SBIL's reading mode configuration. The nanoruler is the only grating patterning tool with its own reading capability to characterize the phase of a grating. A particular problem that occurs in reading mode involves signal loss from varying reflectivity or diffraction efficiency across the substrate. This may occur if for example a large particle is lodged somewhere on the substrate, or if the material is not uniform. Ordinarily, a scan would have to be programmed to avoid these areas on the substrate. By using the absolute phase feature, and some of the beam blanking features (such as checking for signal strength before re-establishing fringe lock), the reading mode performance may be greatly enhanced.

2.4 Phase Map

In this section, I will describe how I implement phase map correction in scanning beam interference lithography. This is a critically important feature for implementing self calibration. The nanoruler has the ability to measure several sources of error. In the following chapters, I will discuss novel techniques which I developed to measure the system distortion in the nanoruler which is the lump sum of all the error sources which contribute to phase distortion during patterning. In addition, in the subsequent chapter I present a technique which I developed to characterize the mirror

distortion in the nanoruler. Many of these sources of error may indeed be difficult to remove from the system. For example, the mirror distortion is due for the most part to manufacturing tolerances. The mirrors where provided from a vendor with a specification of approximately $\lambda/10$ or roughly 60 nanometers. Furthermore, the stage must rotate during scanning due to torque from the magnetic drive forces or non-straightness of the guiding ceramic surfaces. Phase map correction allows one to correct the phase due to all these error sources by modulating the interference fringes.

Implementing phase map correction with the existing infrastructure really turns into a software challenge. As noted in the introduction, the nanoruler uses three Texas Instruments TMS32006701 DSPs denoted (A,B,D) on an Ixthos Champ C6 board. A visual basic program transfers data from a windows operating system onto the program memory of DSP D (the data exchange DSP) using an interface standard known as a JTAG interface. Since DSP A is the real-time DSP, which computes all of the fringe locking operations, it must be in control of a local PCI bus. A diagram of the DSP architecture, modified from the Ixthos manual, is shown in figure 2-9. An important point to take notice of is that DSP's A and B form a cluster. These two DSPs can share their memory directly through the IXstar chip. However, DSP D (the data exchange DSP) must transfer memory to DSP A or B with the through the use of an additional chanel, the PCI-PCI bridge chip. The general strategy then for performing memory transfers is to have the real time DSP, DSP A, in charge of the Ixstar PCI bus. This means that DSP A should initiate writes and reads from DSP B. Transfering memory maps from Visual Basic then involves using the JTAG to transfer the data to DSP D, followed by a data transfer from DSP D to DSP B by means of the PCI-PCI bridge. Subsequently, DSP A may then read DSP B's memory at a non-critical time. This memory routing scheme has been observed to give reliable results. Prior attempts by Konkola, and the author to directly write into DSP A have resulted in sporadic errors causing hang ups.

Once the memory has been appropriately transferred, we need to devise a method to look up the memory map for phase correction. This is done by taking a two-

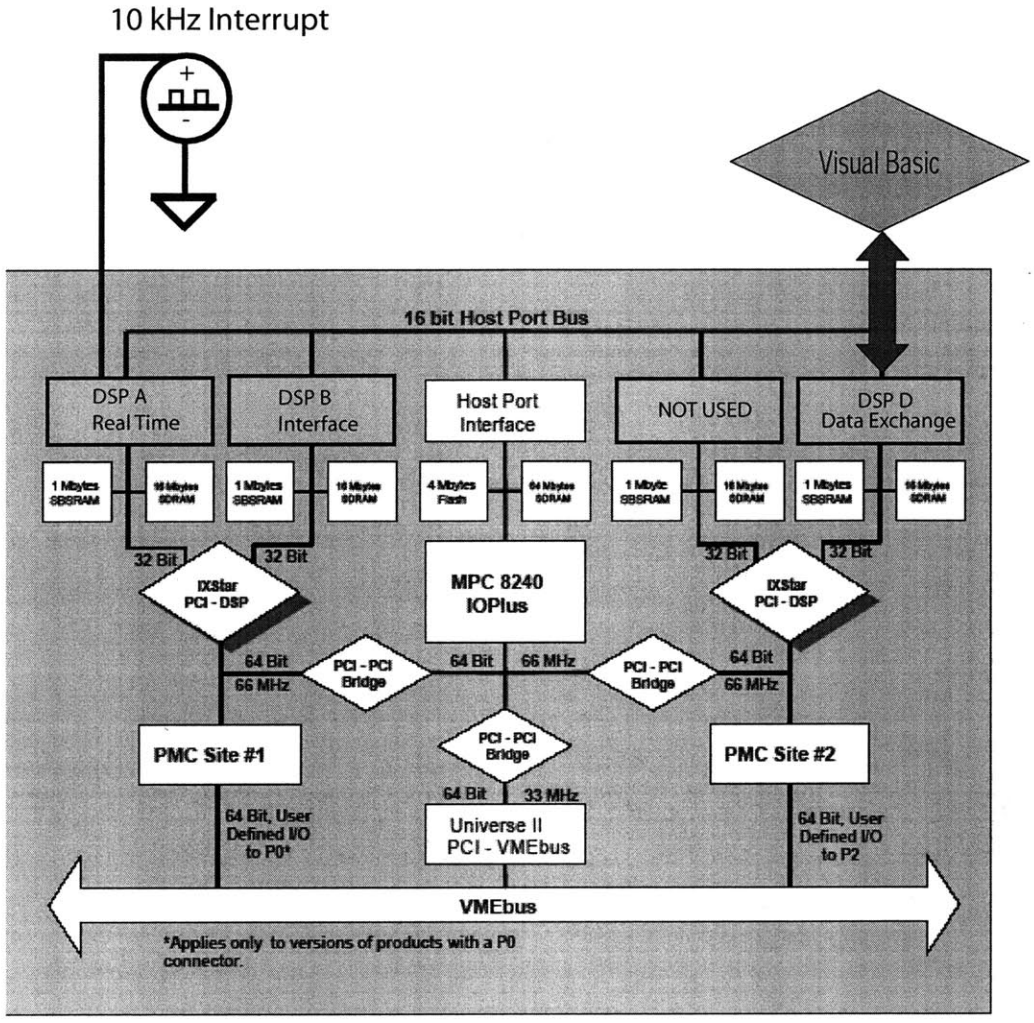

Figure 2-9: DSP architecture. DSP A is the real time DSP which uses a Zygo reference clock to initiate an interrupt service routine. An additional interrupt (not shown) is used to monitor errors and for handshaking with labview. The phase maps are transferred from Visual Basic to DSP D using a JTAG interface.

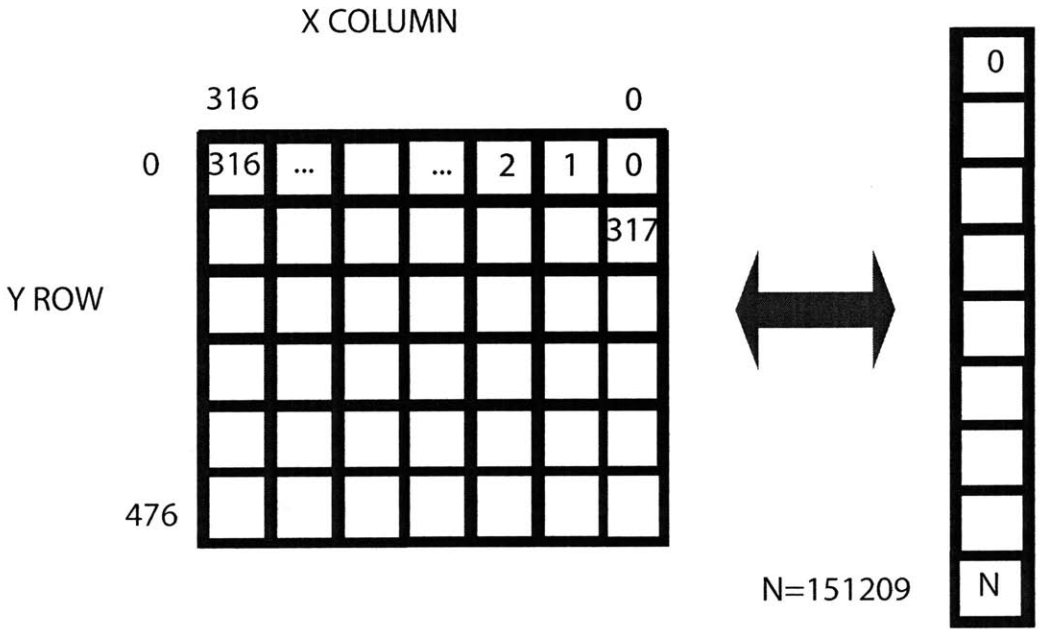

Figure 2-10: A two-dimensional phase map is converted into a 1-D array lookup table

dimensional phase correction map described by $\theta_{Map}(x, y)$ and discretizing it by placing it into 2-D pixels of area 1x1 mm^2. The phase map should be arranged into a 2-dimensional array of size 477x317 to allow for the full range of travel. Accordingly, the stage has a range of travel of 316 mm in the X direction and 476 mm in the Y direction. Since the memory locations in the DSP use linear addressing, it is more convenient to convert this two dimensional phase map into a 1 dimensional array. This is done by mapping the (i, j) element in the two dimensional array to an element k in the one dimensional array where $k = i317 + j$. This way, a 2-D array of size 317x477 gets mapped into a 1-D array of length 151,209. This mapping scheme is illustrated in figure 2-10.

Once the lookup table is generated and transferred into a linear memory address, the phase may easily be retrieved while the stage is scanning. This is accomplished by reading the (x,y) positions of the stage and rounding them to the nearest millimeter. This provides the corresponding indices (i,j) of the two dimensional phasemap position. The memory offset address is then computed by the relationship: Offset Address=i317 + j. The phase values are then retrieved and sent out to the frequency synthesizer for phase modulation.

In order to test the phase map, a two dimensional sinusoidal phase was written with a 30 mm period as shown in figure 2-12. The sinusoidal amplitude of the phase map was choosen to be equal to 2π. This corresponds to a period of 574 nm. The units in figure 2-12 are shown in nm, where we have scaled the output phase by a factor of $P_f/2\pi$ where P_f is the fringe period of 574 nm. It is common in the field of optics to express the phase in terms of wavelength or period, which will be done frequently throughout this thesis. The grating was then read back in reading mode, which resulted in the scaled output shown in 2-12. The results are rather remarkable. Not only can we use phase map correction to correct for phase distortions caused by our system, but we may verify that these phase corrections work by using the nanoruler's reading mode.

Before leaving this section, I would like to make a few additional comments. Phase correction may also be applied to provide spatial phase modulation to a diffracted

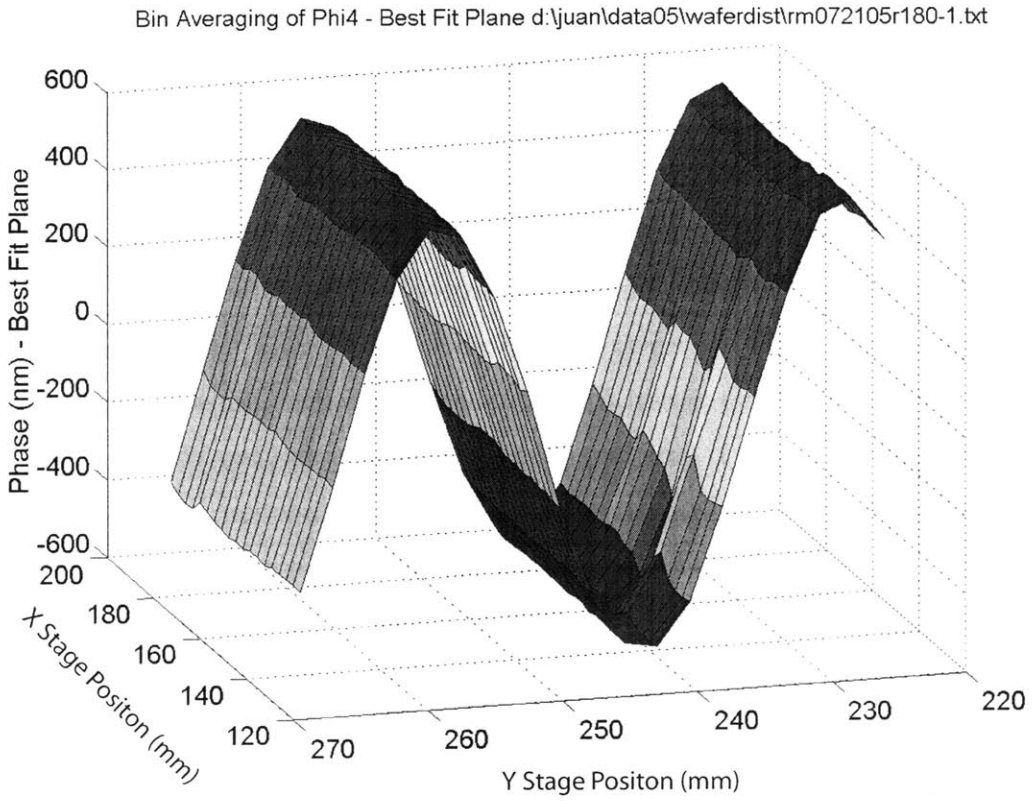

Figure 2-11: A sinusoidal phase was written into a grating using a 2-D phase map and read back using the nanoruler's reading mode. The phase is scaled by $\frac{P_f}{2\pi}$ to give units of nanometers. P_f is the nominal grating period of 574 nm. This confirms that our phase-correction works.

wavefront. This phase modulation may serve several purposes. If one wishes to apply a phase correction for the substrate of a surface for example, the period may be varied to compensate for the substrate non-flatness.

Secondly, I would like to point out another interesting technique which may be employeed to verify the phase map correction. When two exposures are done on to a photo-resist covered substrate, a moire pattern is formed. This may be understood by recognizing that areas in the resist where the second exposure is in phase with the first exposure, high contrast results. Where the two exposures are a 180 degrees out of phase, poor contrast is formed. The resulting image in the resist where high contrast is formed and a grating is visible then corresponds to regions where the sum or difference of the phase of the two exposures is a constant. Two exposures were made to generate the pattern shown in figure ??. In the first exposure, the grating period was written with a period $P1 \approx 574$ nm where the periodicity is in the X direction. The second exposure was made with a period $P2 = P1 + .01$ nm in the X-direction with a with a phase modulation in the Y direction given by $\phi(y) = 2\pi sin(2\pi y/30mm)$. The moire pattern consists of two overlapping sinusoids in the y direction, corresponding to a moire pattern with a Y phase $\pm\phi(y)$ variation and a period in the X-direction of approximately 30 mm. This double exposure moire technique may also be used to verify that the phase maps work. The advantage of this approach is that one does not have to place the system into reading mode, and requires only two exposures followed by development. In addition, the moire technique may be used as a test of the repeatability of our system over extended periods of time. For example, a first exposure may be performed days before the second exposure. Any changes in the system which causes phase distortion result in a phase shift of the moire pattern.

The latter technique is useful for a visual estimate of system distortion, and in our example of verifying a phase map. However, I have developed superior approaches which will be described in chapter 5 that provide quantitative results. In the next chapter, I will describe the major sources of coordinate system error in the nanoruler.

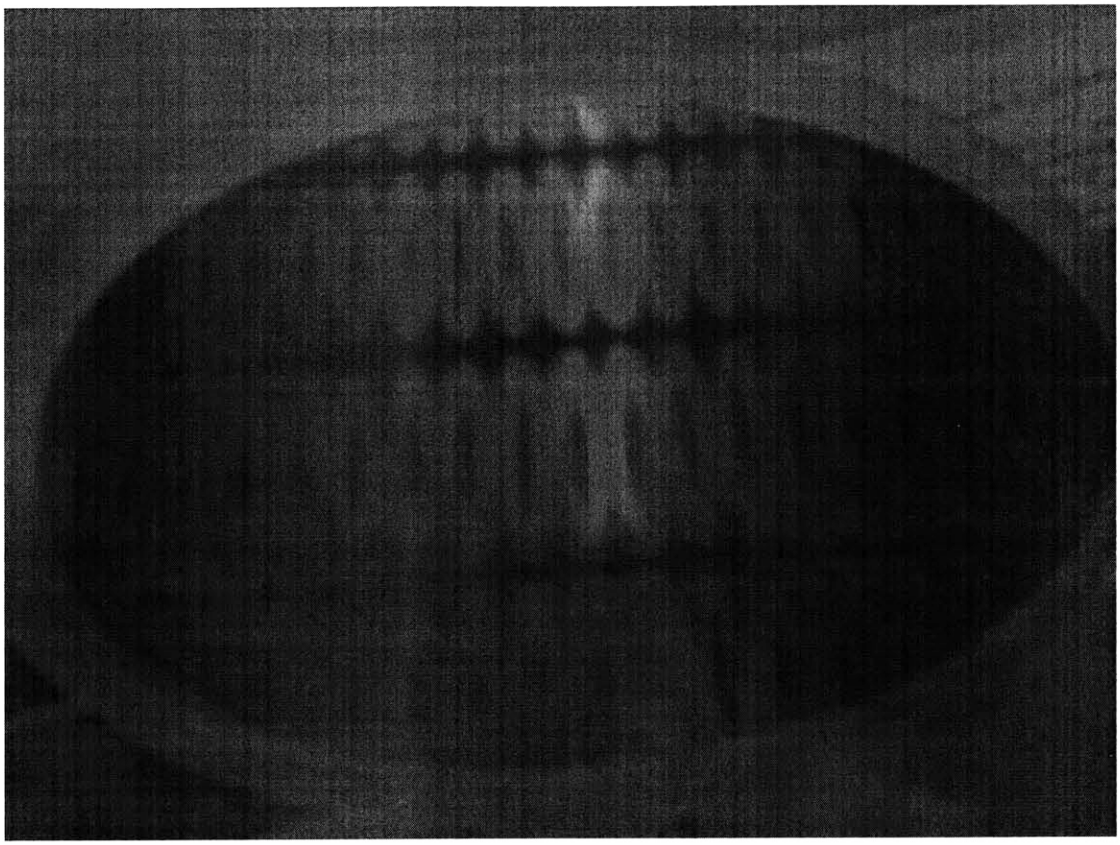

Figure 2-12: Two exposures were used to generate this moire pattern. In the first exposure, a grating period of $P_1 \approx 574$ nm along the X axis was patterned . In the second exposure, a period of $P_2 = P_1 + .01$ nm along the X was also written into the grating. During the second exposure, a phase map was also used to add a sinusoidal phase modulation along the Y direction in the form of $2\pi \sin(2\pi y/30mm)$. The resulting moire period is shown where the X axis runs from left to right (horizontally) in the figure, and the Y axis from top to bottom (vertically).

Chapter 3

Coordinate System Considerations

In this chapter we will discuss errors in measuring the true stage location. We will begin this chapter by describing a class of errors known as Abbe errors. These errors are created by improper alignment of the coordinate measuring axis, and rotations of the object to be measured. A simple way to illustrate this process is to consider the "grating image" which is formed by the ultra-violet beams. If the stage rotates about the grating image it has two potential consequences will be described seperately. The first potential problem involves loss of contrast.

If the dark and bright fringe periods are being printed onto a photo-resist covered substrate a small rotation will cause the pattern to expose areas of the resist which are not meant to be exposed. This would effect the linewidth of the pattern, but not the phase error. Certainly, the fringe period of the image grating may change. However, it will be shown in the chapter on contrast that the period of the grating which is printed into the resist (and therefore its phase) will be determined by the step-over distance in the scan (for the case of parallel scanning) or by the phase modulation in the case of doppler scanning. The effect of the actual image grating period and orientation will affect the contrast of the printed grating, but not the phase error.

The primary goal of scanning beam interference lithography is to achieve high phase accuracy. Uniform linewidth is also of high priority, however, it may be achieved by controlling the exposure dose and the image contrast.

The second problem caused by rotations involve translations. If the stage is ro-

tated about the origin of the "grating image" any translation of the stage, as a result of the rotation, would lead to a phase error. For example, the grating image may be described as a function $f(x, y)$, with a centroid at a location on the stage which we take to be the origin $(x, y) = (0, 0)$. As the stage rotates about this centroid, the origin of the stage will not translate. However, improper alignment of the interferometer beams will lead one to conclude that the stage origin has translated, when in fact it has not.

A claim which is made, but is by no means obvious, is that proper alignment of the interferometer beams would minimize this translational error. We will show that misalignment of the interferometer beams would therefore lead to translational error, and elaborate more on this important result. We will consider rotations only in the $(X - Y)$ plane, however, the techniques developed may be extended to out-of-plane rotations. The reader is encouraged to read Konkola's thesis for a detailed analysis of out-of-plane rotations [27].

Another contributor of coordinate system error involves the stage mirrors themselves. These mirrors serve as reference planes which define the X and Y axis of our stage. Any deviation from nonflatness of the stage mirrors will lead to a deviation in the linearity of our coordinate system. The particular mirrors which we use have a nonflatness of approximately $\lambda/10 \approx 60$ nm. This will lead at best to gratings with 60 nm of phase distortion. It is the goal of this chapter to develop a measurement algorithm that will allow us to measure the stage mirror distortion. Once the distortion is known, the phase may be corrected by using a phase lookup table which I developed in chapter 2.

Before embarking on the adventure to measure rotational errors, we will develop mathematical operators which will alleviate the need for making several geometric arguments based on sketches. These operators may be found in text books on precision design and elsewhere [41]. The power of these operators comes about because they immediately allow one to make generalized statements. They also reveal key properties which are useful in this analysis. For example, they allow one to derive the coupling which occurs between rotation about a point and translation.

3.1 Translation and Rotation Operators

The stage motion is extremely critical for good performance in scanning beam interference lithography. Rotation of the stage could lead to contrast degradation among other sources of error. In addition, the X axis interferometer and Y axis interferometer must be aligned properly such that the intersection of the two measurement axis meet at the location of the "grating image." If this is not the case, then a class of errors known as "Abbe" errors occur. Dr. Ernst Abbe made the following statement regarding the design of precision microscopes "if errors of parallax are to be avoided, the measuring system should be placed coaxially with the axis along which the axis is to be measured on the workpiece" [41, 12].

It is the aim of this section to develop the formalism to treat our mirror measurement process. In doing so, we will gain the additional benefit of understanding the source of "Abbe" errors, which will be a powerful tool in understanding the need for precision design in the Nanoruler.

We will restrict our discussion to translations and rotations in the $X-Y$ plane. There are a set of linear matrix operations which will make the discussion more convenient. First, let us consider an arbitrary position vector $\overrightarrow{r}$ and describe its coordinates in polar form such as

$$\overrightarrow{r} = \rho cos(\theta)\overrightarrow{x} + \rho sin(\theta)\overrightarrow{y}. \tag{3.1}$$

Now suppose we wish to rotate this vector by an angle α. This new vector may be written as

$$\begin{aligned} \overrightarrow{r'} &= \rho\cos(\theta+\alpha)\overrightarrow{x} + \rho sin(\theta+\alpha)\overrightarrow{y} \\ \overrightarrow{r'} &= \{\rho\cos(\theta)\cos(\alpha) - \rho\sin(\theta)\sin(\alpha)\}\overrightarrow{x} + \{\rho\sin(\theta)\cos(\alpha) + \rho\cos(\theta)\sin(\alpha)\}\overrightarrow{y}. \end{aligned} \tag{3.2}$$

It is convenient to express this rotation operation in its equivalent matrix form as

$$\begin{bmatrix} r'_x \\ r'_y \end{bmatrix} = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix} \begin{bmatrix} r_x \\ r_y \end{bmatrix}. \tag{3.3}$$

The rotation matrix $R_\alpha(\alpha)$ is then defined to be a two-by-two matrix which operates on $r = r_x\overrightarrow{x} + r_y\overrightarrow{y}$ to produce r' by the relation $r' = R_\alpha r$. As one could see, the transformation matrix does not allow for translations. In order to allow the additional degree of freedom for translation one must increase the dimensionality of the matrix by one [41]. This is commonly known as a homogenous transformation. In preparation for the next section, we will define two fundamental operators. The first is the homogenous rotation operator

$$R_H(\alpha) = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) & 0 \\ \sin(\alpha) & \cos(\alpha) & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{3.4}$$

which allows for rotations about an origin (0,0). The second operator is a homogenous translation operator which allows for a translation of $(\Delta x, \Delta y)$

$$T(\Delta x, \Delta y) = \begin{bmatrix} 1 & 0 & \Delta x \\ 0 & 1 & \Delta y \\ 0 & 0 & 1 \end{bmatrix}. \tag{3.5}$$

One may verify that the inverse of the translation operator $T^{-1}(\Delta x, \Delta y)$ reverses the translation to an increment $(-\Delta x, -\Delta y)$ such that $T^{-1}T = T(-\Delta x, -\Delta y)T(\Delta x, \Delta y) = I$, where I is the identity matrix. The operation of the translation operator $T(\Delta x, \Delta y)$ on a position vector $\overrightarrow{r} = x_0\overrightarrow{x} + y_0\overrightarrow{y}$ may be summarized as $T(\Delta x, \Delta y)\overrightarrow{r} = (x_0 + \Delta x)\overrightarrow{x} + (y_0 + \Delta y)\overrightarrow{y}$. In order to confirm our understanding of the use of

these operators, we will express the equivalent operation in matrix form as

$$\begin{bmatrix} x_0 + \Delta x \\ y_0 + \Delta y \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & \Delta x \\ 0 & 1 & \Delta y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \\ 1 \end{bmatrix}. \tag{3.6}$$

The usefullness of defining this formalism will become apparant in the next section where we will apply the rotation and translation operators to determine the effect of Abbe errors in our system.

3.1.1 Stage Rotations and Abbe Error

Paul Konkola performed the original design work on the Nanoruler and discussed several sources of Abbe errors present in the system [27]. In his thesis he considers the rotations of several components of the Nanoruler system. He provides a thorough analysis of the rotation of a subsystem where all of the phase meters reside (also known as the "metrology block"). Furthermore, he also considers "out-of-plane rotations" (the pitch and roll) of the various subsystems including the interferometer head.

While it is not appropriate to repeat that analysis here, we will elaborate on some of the critical findings in his thesis. Specifically, Konkola's thesis presents a small section (section 4.3.1) on the consequence of rotational stage errors [27]. We would like to derive some of the statements made in that section since it is fundamental to the understanding of the most important accuracy limitations in the Nanoruler.

Eliminating Abbe errors are crucial to the accuracy of the Nanoruler and it has therefore earned itself a section in this thesis. It is the sincere hope that at the end of this analysis, a reader without the insight gained from experience in the area of precision design will understand the source of Abbe error. At the very least, it will confirm the results Dr. Konkola achieved and serve as a good background. In any event, the analysis in this section will allow us to develop an argument for a novel mirror distortion algorithm to be presented later in this chapter.

Before taking advantage of the mathematical operators defined in the previous

section, it is important to point out some general principles. Refering to figure 3-1 we will notice three different orientations of the stage.

In figure 3-1 (a) the stage is in a zero degree orientation and the stage X axis interferometer measures the distance to the point P on the stage mirror. The grating image which is formed by the ultra-violet beams is assumed to be fixed at the origin of the coordinate system $X - Y$ denoted by point O in the figure.

Now consider a stage rotation by an angle α about the origin as shown in figure 3-1 (b). If the stage rotates about the origin (point O) as shown in figure 3-1 (b) it is observed that the origin is invariant to rotation. Since the "image grating" remains fixed to the origin the stage should not translate. However, as shown in figure 3-1(b) the point P on the stage mirror translates by an amount δ_x in x and δ_y in y. This results in a new position measurement in the X-axis interferometer. The interferometer now measures the x coordinate of point S (in the unrotated coordinate system). The following analysis will show that for small rotations (α) and the correct choice of P this error will be neglible. For an incorrect choice of P, that is one in which the measurement axis of the interferometer does not include the origin, there will be a considerably larger error which we have been referring to as an Abbe error.

The analysis will be made simpler by considering a key property of rotations. Namely, a rotation about the origin may be viewed as a rotation about any point plus a translation. The use of the operator formalism may be used to prove this. Pictorally, as shown in figure 3-1 (c), the stage is rotated around point P. In this case, P is invariant to the rotation. We may verify that the orientation of the stage in figure 3-1 (b) may be obtained by performing the rotation in figure 3-1 (c) followed by a translation (δ_x, δ_y).

Incidentally, I would also like to point out that I am neglecting the fact that the reflected beam from the stage mirror will also change its angle upon reflection from point P. This angle change with respect to a "reference beam" at worse would lead to a fringe pattern at the detector which will be averaged out over the spot size of the beam. The path length of the reflected beam will also change as a result of the angle of reflection, but results in a cosine error of the form $L(1 - \frac{\theta^2}{2})$ where L is

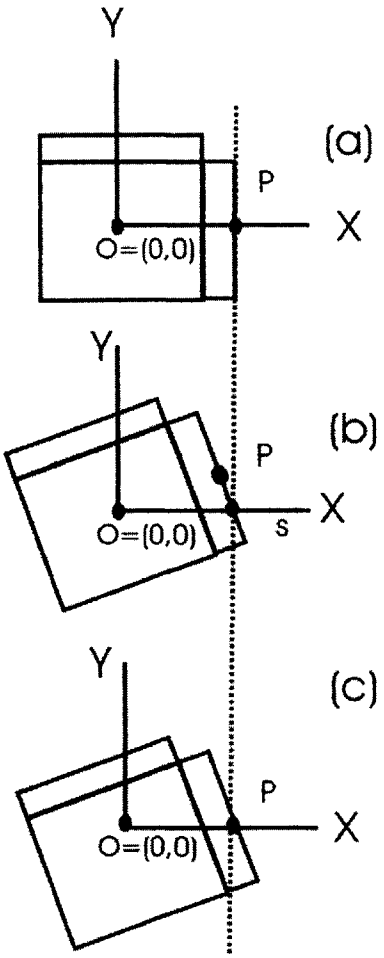

Figure 3-1: *Top view of the stage showing an X-axis mirror and a Y-axis mirror. Three different orientations for the stage are shown with an $X-Y$ coordinate system defined by the measurement axis of the X and Y axis interferometers. The origin labeled O is the location where the measurement axis of the X and Y interferometers meet. In order to eliminate Abbe error, the grating image must be located at the origin O. (a) the stage is in its nominal position and the inteferometer measures the X-axis stage mirror at point P. (b) the stage is rotated by α at the origin 0 and (c) the stage is rotated by α at the point labeled P.*

the pathlength in the case of reflection from normal incidence. Since θ, the angle with which the reflected beam changes with respect to normal incidence is small, we may neglect this cosine error. We are interested only in path length changes along the measurement axis of the interferometer. For this reason, the angle change upon reflection is not of significance for small angles.

This suggests a useful way to keep track of the location of point P due to a rotation about the origin. The process is broken down into two steps. First, rotate around P which does not result in any translation of point P and therefore does not cause a phase change on the interferometer measurement. Secondly, translate by some amount (δ_x, δ_y).

The first step of rotation around P is simple enough. We will need to work a little harder to determine the translation δ_x along the X-axis and $\delta_y$along the Y axis.

A convenient method to determine the translation (δ_x, δ_y) is to apply the rotation operator around P, and compare it to the rotation operator applied at the origin O. Using the operator formalism, the rotation about a point $P = (p_x, p_y)$ in a coordinate system $X - Y$ may be expressed as $R' = T^{-1}(-p_x, -p_y)R_H T(-p_x, -p_y)$. This result may require some thought, and could be found in a standard text on linear transformations. The rotation operator is defined to operate about an origin in some reference frame. We may attribute a reference frame to point O, and a reference frame to point P. We know how to apply the rotation operator in either reference frame. We now need to express the operator which operates in the reference frame of P to its equivalent in the reference frame of O.

Mathematically, we can find the equivalent operator (rotation about P) in the reference frame of O by transforming the reference frame of O to P (i.e., translating), applying the rotation in P, and transforming the reference frame back to O.

An alternative way to look at this result is to realize that we first translate a point P back to the origin by applying $T(-p_x, -p_y)$. We rotate about the origin by applying the second operator (R_H)(which we know preserves the origin). Finally, we apply the third operator $T(p_x, p_y)$ which translates the point P back. Throughout the process the coordinates of point P are preserved which we know must be the case

since P is invariant to any rotation α about point P. The resulting transformation for a rotation about $P = (p_x, p_y)$ in the coordinate system $(X - Y)$ with an origin O is shown below:

$$R' = T^{-1}(-p_x, -p_y)R_H T(-p_x, -p_y) = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) & -cos(\alpha)p_x + \sin(\alpha)p_y + p_x \\ \sin(\alpha) & \cos(\alpha) & -sin(\alpha)p_x - \cos(\alpha)p_y + p_y \\ 0 & 0 & 1 \end{bmatrix} \quad (3.7)$$

We can observe that the operation of rotating around the origin R_H in equation 3.4 may be expressed as a rotation about P with a translation of (δ_x, δ_y). This statement may be compactly written as $R_H = T(\delta_x, \delta_y)R'$ where δ_x and δ_y may be found by inspection from equation 3.7. Notably, we find $\delta_x = cos(\alpha)p_x - \sin(\alpha)p_y - p_x$ and $\delta_y = sin(\alpha)p_x + \cos(\alpha)p_y - p_y$.

With all the prelimaries aside we are now ready to make some powerful conclusions on the effect of rotations. For small angles α one may make the approximation $\cos(\alpha) \approx 1 - \frac{\alpha^2}{2}$ and $\sin(\alpha) \approx \alpha$. This allows us to rewrite the translation of point P as $\delta_x = -\frac{\alpha^2}{2}p_x - \alpha p_y$ and $\delta_y = \alpha p_x - \frac{\alpha^2}{2}p_y$. Typically, the angles of rotation on the Nanoruler are on the order of 10 microradians. This allows us to eliminate all the second order terms with regards to α. The interpretation of this result may be aided by an illustration. Figure 3-2(a) depicts a rotation of the stage about the point P. Next, the stage is translated by $\delta_y \approx \alpha p_x$. This step changes the distance of the X-axis interferometer measurement by a negligible amount (ϵ) as shown in in figure 3-2 (b). The geometry in figure 3-2 shows that the distance change (ϵ) in the X axis interferometer due to a translation of δ_y as a result of a rotation α about the origin is given by $\epsilon \approx \alpha^2 p_x$. The most important contribution to the error in the X axis interferometer measurement is due to the translation of δ_x.

The translation δ_x was derived to be $\delta_x \approx \alpha p_y$. This is where the *Abbe* error comes in and is the most important point in this analysis. If the axis of the X-axis interferometer is aligned with the origin (which is taken to be the location of the image grating) then $p_y = 0$. However, if there is an offset of an amount $p_y = Y_\Delta$, the

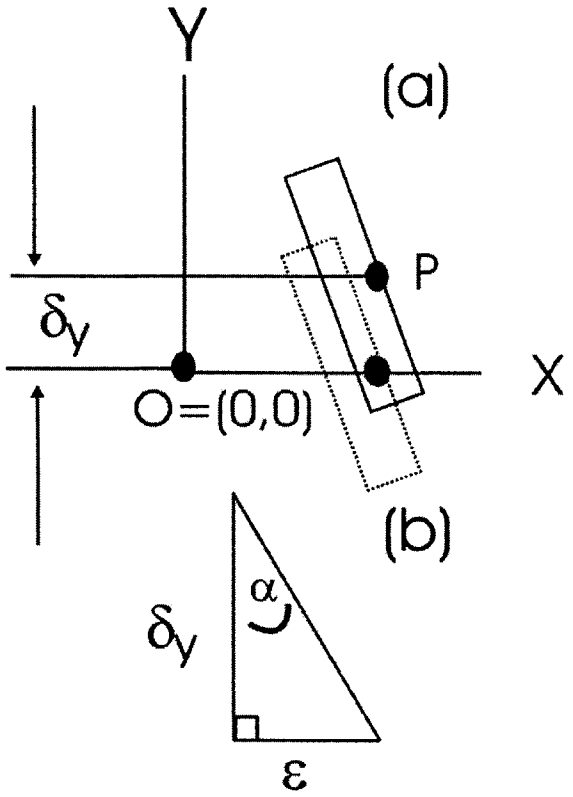

Figure 3-2: *A rotation about the origin may be considered as a rotation about P (which leaves P invariant), plus a translation of (δ_x,δ_y). (a) the stage mirror is rotated by α and translated by δ_y. (b) The error in the X-axis interferometer measurement as a result of translating δ_y is negligible for small α. This error is found to be $\epsilon \approx \alpha\delta_y$*

magnitude of the Abbe error in the X axis will be given by

$$Abbe_{error-x} = Y_\Delta \alpha. \tag{3.8}$$

An analogous argument may be made for the Y-axis interferometer. In this case ideally $p_x = 0$. However, if the Y measurement axis is translated from the origin by $p_x = X_\Delta$, then the important contribution to the Y-axis error comes about from a translation δ_y. In this case, $\delta_y = X_\Delta \alpha$. We give this error a special name of $Abbe_{error-y}$ and highlight this important result as

$$Abbe_{error-y} = X_\Delta \alpha. \tag{3.9}$$

Figure 3-3 clearly defines the offset X_Δ and Y_Δ. As we have seen, proper alignment of

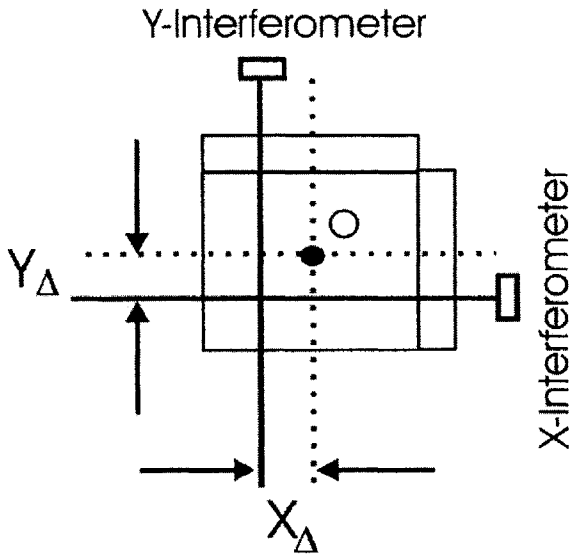

Figure 3-3: *The origin is defined at point O where the image grating resides. The X-axis interferometer is shown with its measurement axis in solid lines. The measurement axis has an offset from the origin by an amount Y_Δ. A rotation of the stage about the origin by (α) would produce an Abbe error in the X axis of $Abbe_{error-x} = \alpha Y_\Delta$. Similarly, an offset of the measurement axis of the Y-interferometer by X_Δ would lead to an Abbe error $Abbe_{error-y} = \alpha X_\Delta$*

the stage interferometers axis is critical for accuracy in the Nanoruler. Unfortunately, the Nanoruler does not have at its disposal good metrology on the intersection of the interferometer beams. An offset of 1 mm for a stage rotation of 10 microradians could lead to an Abbe error of 10 nanometers. The alignment for the interferometer beams with respect to the location of the grating image must be made before the stage is installed which in turn requires metrology on squareness of the stage mirrors and on the squareness of the interferometer beam paths.

On the other hand, the Nanoruler has very good metrology on the placement of the grating image. For the current design of the Nanoruler, however, large translations of the grating image location (taken to be our origin) are not practical. The mirrors which guide a split-off portion of the ultra-violet beams which compose the grating image are only a few millimeters in diameter. Small translations on the order of a

mm would require major realignment of the optical components to avoid clipping.

An improved design of the Nanoruler will allow for translations of the grating image, or of the interferometer beams. One may conceive of an experiment where the phase error in a grating is read for different positions of the grating image. One must be careful in interpreting the results when reading the phase of a grating. For example, if a grating is written in a distorted coordinate system and the grating phase is read back in the same coordinate system, one would conclude there is no error. An offset of the grating image would produce a relative Abbe error. Chapter 5 addresses this problem by designing a grating rotation test which will allow one to draw conclusions about the systematic error.

In the following section, we will discuss another source of coordinate measurement error. This error is due to a distortion in the stage mirrors. An ideal mirror would have all of the properties of a geometrical plane. It would be perfectly flat over the length of interest. Real mirrors have non-flatness due to manufacturing errors and distortions due to mounting stresses.

3.2 Mirror Nonflatness

Typically, high accuracy two-axis (x,y) stages use displacement interferometers to determine stage position and velocity. A stage controller uses these quantities as feedback to minimize the error between the desired and actual stage travel. This method, however, is prone to error due to the non-ideal flatness of the stage mirrors. In order to compensate for the mirror nonflatness, knowledge of the mirror profile over the length scale of interest is required. We have developed a method to measure the stage mirror nonflatness by using two linear/angular interferometers mounted on opposing axes. By performing discrete sampling of the stage mirrors along the range of travel, we show that both the stage rotation and a filtered mirror profile can be obtained. Here we present the measurement transfer function which describes the spatial frequencies of the mirror profile that can be recovered using this approach. A novel inverse transfer function was developed which eliminates the instability due

to zeros in the transfer function. Moreover, we will apply this technique to measure nonlinearities of a 300 millimeter stage mirror. In the context of this measurement, we will describe the assumptions and limitations of this approach.

Techniques for calibrating mirror nonflatness exist in the literature. One promising method uses a self-calibration approach known as the three-flat test [10]. This method uses three reference flats which are compared to each other interferometrically. By comparing three combinations of the reference flats, followed by a fourth combination in which one of the flats is rotated, a resulting set of equations are obtained. These equations may be solved to obtain the nonflatness of the reference flats [23].

Once a reference flat is calibrated, it can subsequently be employed to calibrate an interferometer, which itself is subject to non-ideal optical components. While these measurement approaches have been demonstrated successfully, the equipment required may be expensive or require extensive design. For a 300 mm stage mirror, a nonflatness measurement would require a pre-calibrated interferometer with a large aperture. Another approach would be to perform a 3-flat test on the mirror itself, which would require extensive design for rotation and probing equipment. Moreover, there is also the possibility that distortion will be added to the mirror by the installation procedure [38].

Here we will present an alternative technique which allows us to measure the mirror nonflatness after it is mounted on a stage. This eliminates the measurement uncertainty resulting from the mounting process. The measurement procedure involves using two opposing axis (X-Y) measurements. Each axis measurement uses a Zygo angular/displacement interferometer. Here, we will focus on measuring the X-stage mirror. The Y-stage mirror may be measured similarly.

3.3 Measuring the X-Stage Mirror

We will begin by considering the stage drawn in figure 3-4. This figure illustrates the X and Y stage mirrors. The X-stage mirror is measured at four different points labeled a, c, d, f, which are separated by a distance h. The stage is located in a

laboratory frame with coordinates (x,y). For convenience, we have chosen the same definition for the laboratory axis as is currently implemented in SBIL.

The X axis interferometer produces a measurement

$$X_m(x,y) = \frac{m_x(x,y) + m_x(x,y+h)}{2}. \tag{3.10}$$

for a given stage position (x,y) in the laboratory frame. For a linear mirror this has the geometric interpretation of monitoring the average of the displacements of the two locations on the mirror labeled "a","c". Accordingly, this would be equivalent to measuring the displacement of the stage mirror at a point labeled "b" the average of "a","c". For a nonlinear mirror, this geometric interpretation will not be entirely correct as we shall see.

Equation 3.10 simply states that when the stage is located at a position (x_0, y_0) in the laboratory frame, point "a" corresponds to the displacement measurement $m_x(x_0, y_0)$, and the measurement of point "c" corresponds to the displacement measurement $m_x(x_0, y_0 + h)$. If the stage is translated to a position $(x_0, y_0 - h)$, one may verify that the displacement measurement of point "a" is equivalent to the displacement measurement of point "c" when the stage was at location (x_0, y_0). This aids in understanding the nomenclature used in the equation 3.10. It also allows us to become familiar with the laboratory reference frame, and confirms that we are using the correct sign "$+h$" in equation 3.10.

The zygo linear/angular displacement interferometers use a configuration which allows for optical differencing. Essentially, an additional two interferometric arms monitor the points "d, f" shown in figure 3-4. This may be viewed as a measurement $X_m(x, y + 2h)$ which has the geometric interpretation of monitoring the point labeled "e" in figure 3-4 for a linear mirror. The angle measurement of the stage is determined by

$$\theta_x(x,y) = \frac{X_m(x,y) - X_m(x,y+2h)}{2h}. \tag{3.11}$$

This has the geometric interpretation of measuring the slope between the points labeled b and e in figure 3-4 for a linear mirror. This is equivalent to measuring

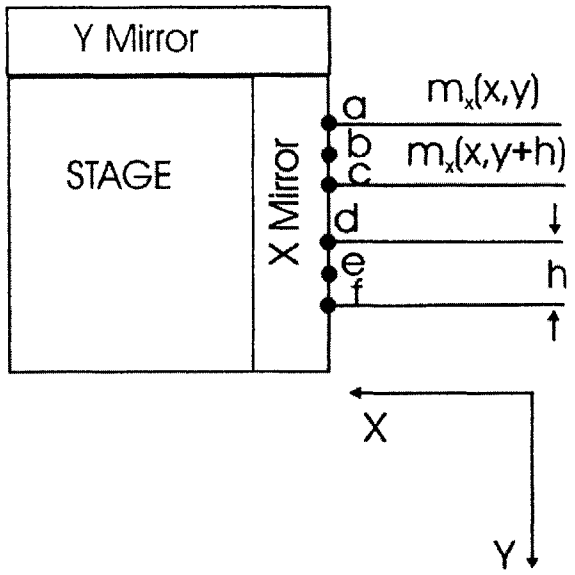

Figure 3-4: (X-Y) stage showing the orientation of the X-Y stage mirrors. For the X-stage mirror, the interferometers measure for different points on the mirror labeled a, c, d, f. These for points are separated in the Y direction by a distance h.

$tan(\theta) \approx \theta$ where θ is the angle the mirror makes with the Y-axis.

Let us proceed to construct our measurement process by now allowing for a nonflat reference mirror and a small stage rotation α. We will assume an X-axis stage mirror which will produce a displacement measurement of the form

$$m_x(x, y) = x + D_x(y), \tag{3.12}$$

where x is the stage position in the laboratory reference frame and $D_x(y)$ is a nonlinear distortion which results from the nonflatness of the stage mirror. In the absence of a distortion $D_x(y)$, $m_x(x, y)$ describes a straight line for each location x_0. For example, substituting $x = x_0$ in equation 3.12 yeilds $m_x(x_0, y) = x_0$ which describes a vertical line (i.e., flat mirror) at a distance x_0. If we apply a rotation about the origin, such that $x' = cos(\alpha)x - sin(\alpha)y$, and $y' = sin(\alpha)x + cos(\alpha)y$, we find that $m_x(x', y') = cos(\alpha)x - sin(\alpha)y + D\{\sin(\alpha)x + \cos(\alpha)y\}$.

We could interpret this result by identifying the two components of the stage

mirror measurement

$$m_x(x', y', \alpha) = m_{x-linear}(x, y, \alpha) + m_{x-nonlinear}(x, y, \alpha). \tag{3.13}$$

The linear portion of the mirror under rotation by α results in a displacement measurement $m_{x-linear}(x, y, \alpha) \approx x - \alpha y$. One may verify that in the case $m_x(x, y) = m_{x-linear}(x, y, \alpha)$, equations 3.10 and 3.11 result in the measurement $\theta_x(x, y) = \alpha$. This is equivalent to adding a wedge in the mirror and measuring the slope of the wedge.

The nonlinear distortion of the mirror $m_{x-nonlinear}(x, y, \alpha) \approx D_x(\alpha x + y)$ may be interpreted as the nonlinear distortion measurement of the X-stage mirror as a function of y. There is a translation $\delta_y = \alpha x$ of the distortion function $D_x(y)$ as a result of the rotation α occuring at the origin. However, the rotations in the Nanoruler are on the order of $\pm$ 10 microradians. This will result in a worse-case translation $\delta_y = 4$ micro-meters. Our measurement process is blind to variations in the X-mirror on this length scale since the beam diameter of the interferometer beams are on the order of 3 mm. In other words, the interferometer measures a moving average version of the mirror profile, which averages out any disturbances with frequencies higher than 1/spot size.

We can therefore neglect the effect of the translation δ_y on the nonlinear mirror measurement. This results in $m_{x-nonlinear}(x, y, \alpha) \approx D_x(y)$, where we see that the rotation of α has negligable effect on the nonlinearity contribution of the mirror to the measurement process.

With these considerations in mind, we can use equations 3.10 and 3.11 to write our measurement θ_x for a nonlinear mirror as

$$\theta_x(x, y) = \frac{D_x(y) + D_x(y + h) - D_x(y + 2h) - D_x(y + 3h)}{4h} + \alpha(y). \tag{3.14}$$

In order to obtain an equation as a function of the X-axis mirror distortion, $D_x(y)$, we must subtract out the angle of rotation $\alpha(y)$ as shown in equation 3.14. This is accomplished by performing a simultaneous measurement on the Y-axis mirror using

the Y-axis interferometer.

3.3.1 Measuring α using the Y-Axis Interferometer

If the stage translates along the line $x = x_0$, the Y-axis interferometer will provide a similar measurement. I have provided a drawing of the displacement measurent of the Y-axis interferometer in figure 3-5. The Y-axis interferometer produces a displacement measurement of the form

$$Y_m(x,y) = \frac{m_y(x,y) + m_y(x-h,y)}{2}, \tag{3.15}$$

which has the geometric interpretation of measuring the average displacements at the points labeled g, i on the Y-stage mirror. Incidentally, the X and Y axis interferometers are not identical. Notice the definitions of the displacements on the Y-axis measurement are not the equivalent of rotating the X-axis interferometer by 90 degrees. The interferometer head was customized for the use in the Nanoruler, but the drawing in figure 3-5 is a faithful representation of our current setup.

The Y-axis interferometer also provides an angular measurement which is defined as

$$\theta_y(x,y) = \frac{Y_m(x,y) - Y_m(x-2h,y)}{2h}. \tag{3.16}$$

An analogous argument to the one used for the X-axis interferometer will allow us to consider the Y-displacement measurement $m_y(x,y) = y + D_y(x)$. The nonlinear distortion in the Y-mirror is given by $D_y(x)$. This allows us to express the Y-displacement measurement as $m_y(x,y) = m_{y-linear}(x,y) + m_{y-nonlinear}(x,y)$, where $m_{y-linear}(x,y) = y$ and $m_{y-nonlinear}(x,y) = D_y(x)$. We are once again interested in observing the effect of a rotation on the Y-mirror displacement measurement $m_{y-linear}(x,y)$. For small α the linear displacement becomes $m'_{y-linear}(x,y,\alpha) \approx \alpha x + y$. This is the equivalent of a wedge in the Y-axis mirror, and one may verify that in the case $m_y(x,y) = m'_{y-linear}(x,y,\alpha)$, equation 3.16 results in the measurement $\theta_x(x,y) = -\alpha$. Special attention must be made to confirm the negative sign of

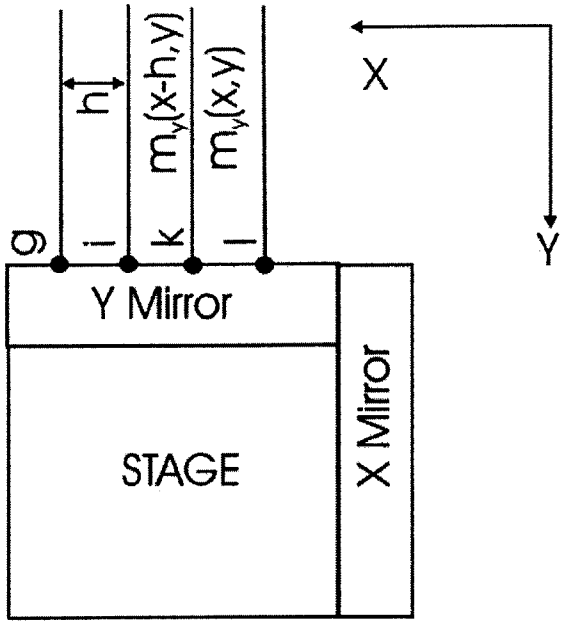

Figure 3-5: (X-Y) stage showing the orientation of the X-Y stage mirrors. For the Y-stage mirror, the interferometers measure four different points on the mirror labeled g, i, k, l. These four points are separated in the X direction by a distance h.

α. The geometry of the Y-axis interferometer is different from the X-axis interferometer. Had the X and Y axis interferometer heads been identical, a positive sign for α would result.

Under a rotation of α, the nonlinear distortion becomes $D_y'(x, y, \alpha) = D_y(x - y\alpha)$, where the prime notation denotes the distortion which occurs after the rotation, and unprimed denotes the distortion prior to rotation. We note that a translation of $\delta_x = y\alpha$ occurs due to a rotation at the origin (0,0). We may once again neglect this translation since for small rotations α the translation will be much smaller than the beam spot size of 3 mm. The maximum translation is on the order of $\delta_y \approx 2$ μm. We therefore conclude that the Y-axis mirror distortion does not change with rotation and may be written as $D_y'(x, \alpha) = D_y(x)$.

Moreover, since we are holding $x = x_0$ constant we expect to be sampling a

constant part of the Y-stage mirror nonlinear distortion. This allows us to write

$$\begin{aligned}\theta_y(x = x_0, y) &= \tfrac{D_y(x_0)+D_y(x_0-h)-D_y(x_0-2h)-D_y(x_0-3h)}{4h} - \alpha(y). \\ &= c - \alpha(y).\end{aligned} \tag{3.17}$$

In equation 3.17 we have recognized that the terms involving D_y do not vary with y and we have lumped their sum into a constant c. We are only interested in knowing α up to an arbitrary constant and we will therefore omit this constant in the analysis.

3.4 Measurement Transfer Function

The method used to extract the X-mirror profile from the angle measurements involves translating the stage while holding the x coordinate of the stage constant. The θ_x measurement on the stage X-interferometer provides information about the stage mirror nonlinearity at four equi-distant points, plus a rotation α. In order to extract the mirror profile information, we add the Y-axis interferometer measurement given by equation 3.17. This results in the quantity

$$\begin{aligned}\Delta\theta(x_0, y) &= \theta_x(x_0, y) + \theta_y(x_0, y) \\ &= \tfrac{D_x(y)+D_x(y+h)-D_x(y+2h)-D_x(y+3h)}{4h} + c,\end{aligned} \tag{3.18}$$

where we may subtract out the constant from our measured result. This will allow us to define a measurement transfer function. From linear systems theory, we can describe this measurement process by an impulse response

$$h(y) = \frac{1}{4h}\{\delta(y) + \delta(y + h) - \delta(y + 2h) - \delta(y - 3h)\}, \tag{3.19}$$

where $\delta(y)$ is the Dirac unit impulse function. We recognize that the output quantity $\Delta\theta(x_0, y)$ may be expressed as

$$\Delta\theta(x_0, y) = h(y) * D_x(y), \tag{3.20}$$

i.e., a convolution of the input consisting of the X-axis mirror's nonlinear distortion $D_x(y)$, and the impulse response $h(y)$. This is a powerful representation in that all of the key properties we need to characterize the measurement process are described by the impulse response. For example, the impulse response has an odd symmetry about the point $y = -3h/2$. We can therefore represent the impulse response as an odd function with a linear delay of $-3h/2$. The Fourier transform of the impulse response results in a measurement transfer function of the form

$$H_c(e^{j\omega}) = -2e^{j\omega 3h/2}[2jcos(\omega h/2)sin(\omega h)]\frac{1}{4h}. \tag{3.21}$$

The transfer function described by equation 3.21 contains the properties which we may have anticipated by inspecting the impulse response. For example, the delay of $-3h/2$ shows up as the complex exponential $e^{j\omega 3h/2}$. Furthermore, we note that the remaining term inside the brackets is completely imaginary as we expect for the Fourier transform of a function with odd symmetry. The zero locations of the continous transfer function may also be observed from inspection of the magnitude of the transfer function response shown in figure 3-6 obtaining $\omega_n = n\pi/h$. This may also be found from inspecting equation 3.21. The component $cos(\omega h/2)$ has zeros at $\omega_n = (n + 1/2)2\pi/h$ where n is an integer. This results in zeros for odd multiples of π/h as in $\omega = \{\pi/h, 3\pi/h, 5\pi/h...\}$ with a first zero location at $\omega = \pi/h$. This corresponds to a spatial frequency of $f = 1/2h$ or a spatial period of $\Lambda = 2h$. The $sin(\omega h)$ results in zeros for all multiples of π/h as in $\omega = \{\pi/h, 2\pi/h, 3\pi/h...\}$ The result of the product of the cosine and sine term results in double zeros at odd multiples of π/h.

A plot of the continuous transfer function is provided in figure 3-6 for a normalized frequency of ω/ω_1 where $\omega_1 = \frac{\pi}{h}$.

This transfer function has a nice physical interpretation. Accordingly, any spatial frequencies present in the mirror with a period of $2h$ which corresponds to the distance between points "b" and "e" in figure 3-4 will not be detected by the measurement process. Periods $\Lambda = 2h/n$ corresponding to the higher order harmonics of the

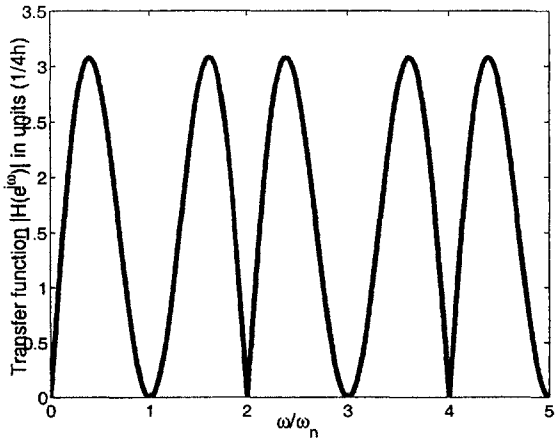

Figure 3-6: $|H_c(e^{j\omega})|$ ploted as a function of ω/ω_1 where $\omega_1 = \frac{\pi}{h}$.

fundamental spatial frequency ($f = nf_0 = n/2h$), will also not be detected by the measurement process. In addition, the DC term will be lost.

In order to recover the input $D_x(y)$, we must apply the inverse transfer function. The corresponding transfer function will have a magnitude in the frequency domain of $|H^i(e^{j\omega})|$ such that $|H^i(e^{j\omega})| = |H_c(e^{j\omega})|^{-1}$. This may be easily seen if we take the magnitude of Fourier transform of the output $\Delta\theta(y)$, which we will denote as $|\overline{\Delta\theta(e^{j\omega})}| = |\overline{D(e^{j\omega})}||H_c(e^{j\omega})|$, and then multiply it by $|H^i(e^{j\omega})|$, resulting in $|\overline{D(e^{j\omega})}| = |\overline{\Delta\theta(e^{j\omega})}||H^i(e^{j\omega})|$. Unfortunately, we do not have an implementable inverse filter in the continuous domain. An equivalent transfer function may be found in the discrete domain . The distortion for the stage mirror will then be found by sampling the output $\Delta\theta$ which results in a discrete signal. The inverse transfer function $H^i(e^{j\omega})$ will then be applied to the output in the discrete domain. In the next section, we will develop this process in greater detail.

3.4.1 Discrete Sampling of the Mirror

In the previous section, we described a measurement process in which the input is the mirror distortion $D_x(y)$ and y is the stage coordinate along the Y-axis. The interferometer beams sample the mirror along discrete points, and this process may

be viewed as a convolution of the input with an impulse response $h(y)$. To recover the input, the inverse filter with the property $|H^i(e^{j\omega})| = |H_c(e^{j\omega})|^{-1}$ must be applied. In this section, we will show how to implement this inverse filter digitally.

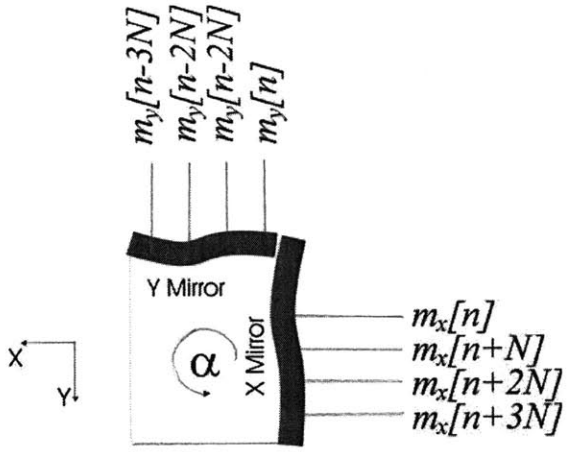

Figure 3-7: The X-Y interferometers use a four-pass scheme to measure stage angle and displacement. The four measurement points are shown for each stage mirror. As the stage travels along Y, keeping the X displacement constant, small rotation errors α may occur.

In order to progress, we will discretize the former equations and introduce a standard notation. Figure 3-7 illustrates our discrete measurement process. The discretization occurs because our controller samples the interferometers at finite intervals during the stage travel. If we represent the discretized stage coordinate $y = n\Lambda_s$ and remove the redundant notation we may rewrite equations (3.10) - (3.17) as

$$X_m[n] = \frac{m_x[n] + m_x[n+N]}{2} \tag{3.22}$$

$$\theta_x[n] = \frac{X_m[n] - X_m[n+2N]}{2h} + \alpha[n] \tag{3.23}$$

$$Y_m[n] = \frac{m_y[n] + m_y[n-N]}{2} \tag{3.24}$$

$$\theta_y[n] = C - \alpha[n]. \tag{3.25}$$

We note that we have chosen to write the discrete equation 3.23 slightly different than its continuous counterpart in equation 3.11. Equation 3.11 was derived for conceptual purposes in the case that the mirrors were oriented at a 90 degree orientation with respect to the measurement axis. Equation 3.23 allows for a linear wedge in the mirror profile before a rotation α occurs (i.e., non-orthogonal mirrors), and is therefore slightly more general. In any event, once we recover the mirror profile by applying the inverse measurement transfer function, we will remove any linearity leaving us with the nonlinear mirror distortion which we seek.

The discretization is facilitated by choosing a sampling period such that $\Lambda_s = h/N$. This insures that the measurement points are located at integer sampling distances from each other. This results in the discrete equations 3.22 - 3.25 having "integer delays" of multiples of N. As we will see, the impulse response of the measurement transfer function will lead to a simple algorithm to recover the mirror's nonlinear distortion.

Certainly, we are not restricted by any physical laws which prevent us from sampling at any rate other than ($f_s = \Lambda_s^{-1} = N/h$). Indeed, discrete filters which have the overall effect of producing a non-integer delay in the frequency response may be represented by using an infinite impulse response (IIR) filter [33]. We choose not to implement such a filter because implementing the inverse of an IIR filter adds considerable complexity and does not allow for the one-to-one correspondance between the continous time finite impulse response filter described by equation 3.19 and a discrete finite impulse response.

Another constraint on the sampling period Λ_s involves satisfying the Nyquist criteria. Namely, we must sample at sampling rate $1/\Lambda_s$ which is greater than twice the maximum frequency content in the mirror distortion to avoid aliasing. The beam spots in our interferometer which sample the mirrors have diameter size $D = 3$ mm. This spot size may be considered as a moving average, since the detector averages the intensity distribution over the spot size (i.e., it detects the power in the beam). This effectively attenuates spatial frequencies present in the mirror larger than $1/D$. As an upper limit, we conservatively acknowledge that our measurement is bandlimited

to within $1/D$. Since the mirrors are of optical quality, they are likely "smoother" and may be treated as bandlimited to within a smaller frequency range.

Investigation of the frequency spectrum of our measuremnents indicates that the mirror does not have significant power for frequencies ¿ $1/D$. For our purposes, we have chosen a sampling rate of $\Lambda_s = h/2 = 3.2$ mm for a seperation $h = 6.4$ mm. This will satisfy the Nyquist criteria if the mirrors are bandlimited such that highest spatial frequency in the mirror spectrum is less than $1/6.4$ mm^{-1}.

3.5 Discrete Measurement Transfer Function

With the desired notation in place for the various quantities, discrete Fourier analysis will allow us to investigate the spatial frequency spectrum of the mirror profile. By holding the X-displacement constant, equation (3.23) for the θ_x measurement results in a measurement of the mirror profile with an error term due to the stage rotation. The rotation may be removed by adding the Y-axis angle measurement. Adding equation (3.25) to equation (3.23) and ommiting the constant offset results in

$$\Delta\theta[n] = \frac{m_x[n] + m_x[n+N]}{4h} - \frac{m_x[n+2N] + m_x[n+3N]}{4h}, \tag{3.26}$$

where $\Delta\theta[n] = \theta_x[n] + \theta_y[n]$. Note, we are only interested in solving equation 3.26 up to an arbitrary constant. By losing information about a constant offset, we are disregarding any linear variation (i.e. a wedge) in the mirrors.

The discrete measurement transfer function may then be defined in terms of the discrete Fourier transforms $\Delta\Theta(e^{j\omega})$, and $M_x(e^{j\omega})$ of the quantities $\Delta\theta[n]$ and $m_x[n]$ respectively. To relate these quantities, we define the discrete measurement transfer function $H_d(e^{j\omega})$ as

$$H_d(e^{j\omega}) = \frac{\Delta\Theta(e^{j\omega})}{M_x(e^{j\omega})} \tag{3.27}$$

$$H_d(e^{j\omega}) = \frac{(1 + e^{+j\omega N} - e^{+j\omega 2N} - e^{+j\omega 3N})}{4h}. \tag{3.28}$$

The measurement transfer function may be interpreted as follows. When we add the stage rotation given by the Y-interferometer θ_y to the X-interferometer θ_x, we obtain a filtered mirror profile. The filter that is applied to the mirror profile is given by the measurement transfer function. A key property of the measurement transfer function is that it does not have a dc (zero frequency $\omega = 0$) component. This component would provide information about the X-displacement of the stage mirror, which is not of interest. Moreover, we would like for our measurement to be immune to the stage X-displacement since considerable stage error occurs while attempting to maintain a constant displacement $x = x_0$.

A few points may be made about the transfer function in equation 3.28. First, the exponents of the complex exponentials are postive, which indicates that the filter is non-causal. This is expected, because the X-axis interferometer measures the points "c","d", and "f" in figure 3-4 which are ahead of point "a" in the measurement. In order to implement this filter in software such as matlab which requires causal filters we must advance the mirror by $3N$. This may be shown mathematically by rewriting the filter's frequency response as

$$H_d(e^{j\omega}) = -e^{+j\omega 3N}\{(\frac{(1 + e^{-j\omega N} - e^{-j\omega 2N} - e^{-j\omega 3N})}{4h}\}, \tag{3.29}$$

where the terms inside the braces {} correspond to a causal filter. The complex exponential denotes an advance of the mirror by $3N$. The procedure to recover the mirror distortion requires an application of the inverse to the causal filter, followed by a delay of $3N$ to the result. The causal-discrete filter may then be described as a Z-transform with the following coefficients

$$H_{d-causal}(z) = -\frac{(1 + z^{-N} - z^{-2N} - z^{-3N})}{4h}. \tag{3.30}$$

We note the negative sign in the transfer function and sign of the mirror distortion to be consistent with our definition of the displacement axis. Accordingly, a positive distortion is to be interpreted as a distortion which increases the path length of the interferometer's beam. A negative distortion is one in which shorten's the interfer-

ometers beam, and would consist of a distortion in which the nonflatness protrudes toward the interferometer.

The pole and zero locations of the discrete causal filter are shown in figure 3-8. We notice that the first zero location (beyond DC) resides at a location corresponding to a normalized frequency of $\omega = \pi/2$. This corresponds to a spatial frequency of $f_s = (\pi/2)1/(\Lambda_s 2\pi) = 1/(4\Lambda_s) = 2/h$. It is not a coincidence that this corresonds to the same zero location of the continuous transfer function. In effect, by sampling the mirror at discrete intervals we have found a discrete equivalent representation of an effective continuous transfer function, $H_{c-eff}(e^{j\omega}) = H_c(e^{j\omega})$ within the frequency range $(-\frac{1}{2\Lambda_s} < f < \frac{1}{2\Lambda_s})$. The interested reader is referred to Oppenheim and Schafer's section on "Discrete Time Processing of Continous-Time Signals" for the sufficient conditions in which this approximation is valid [33]. Moreover, from the pole-zero plot we observe locations where there are double zeros. We notice that we observed the same double zero locations in our discussion of the continuous transfer function.

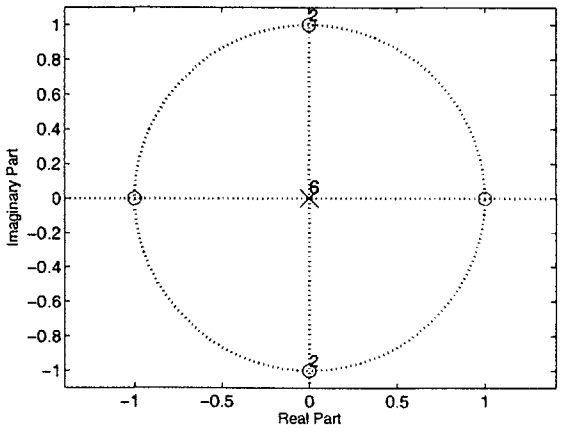

Figure 3-8: Zeros of the Z-transform of the causal discrete transfer function $H_{d-causal}$. The first zero beyond DC corresponds to a normalized frequency of $\omega = \pi/2$ which corresponds to a spatial frequency of $f_0 = 1/(2h)$. The remaining zeros are higher order harmonics nf_0. The sampling period is $\Lambda = h/2$. The numbers in the figure indicate the order of the zeros and poles

To recover the mirror profile, we must apply the inverse of the discrete transfer function $H_d(e^{j\omega})^{-1}$. Unfortunately, investigation of the zeros of $H_d(e^{j\omega})$ reveal that

they reside on the unit circle, which by definition does not have a stable inverse transfer function. In order to obtain a stable inverse, we must find an approximation to $H_d(e^{j\omega})$ with zeros inside the unit circle. Subsequently, we may take the inverse of this approximate function and apply it to $\Delta\theta$ to recover the mirror profile. I would also like to point out, that if one were to apply the unstable inverse transfer function in Matlab the result would "blow up." Figure 3-9 shows an approximate transfer function for $N = 2$.

The transfer function shown in the red dashed curve figure 3-9 is an approximation to the transfer function $H_d(e^{j\omega})$. This transfer function has a stable inverse since it's poles and zeros shown in figure 3-10 reside within the unit circle. The magnitude of the approximate transfer function corresponds well with the actual transfer function with the exception of the zero location at a normalized frequency of $\omega = 0.5$ and at DC. Any power at this frequency will be lost. Moreover, the phase shown in the bottom of figure 3-9 is approximately linear (which corresponds to a constant group delay) with the exception of a distortion occuring at the normalized frequency ($\omega = 0, .5, 1$) which is not of interest since these locations correspond to lost frequency content.

3.6 Experimental Results

So far we have outlined a process in which we can determine the nonlinear distortion of a mirror. In the experiments to be described in this section, the stage is scanned along a constant x_0. The particular position x_0 that is chosen is not significant, since the measurement process is not a function of this parameter. As the stage is scanned along the Y axis, both stage mirrors are sampled using high speed electronics (i.e., the Zygo ZMI-2000 series measurement boards). The angular output of the Y-Axis interferometer is added to the angular output of the X-Axis interferometer producing the signal $\Delta\theta[n]$. The resolution of the Zygo 2000 series electronics is $\frac{\lambda}{4}/512$, where $\lambda = 633$ nm corresponds to the wavelength of a helium neon laser, and the factor of 4 is a result of the multiple passes in the beam path. This corresponds to the resolution

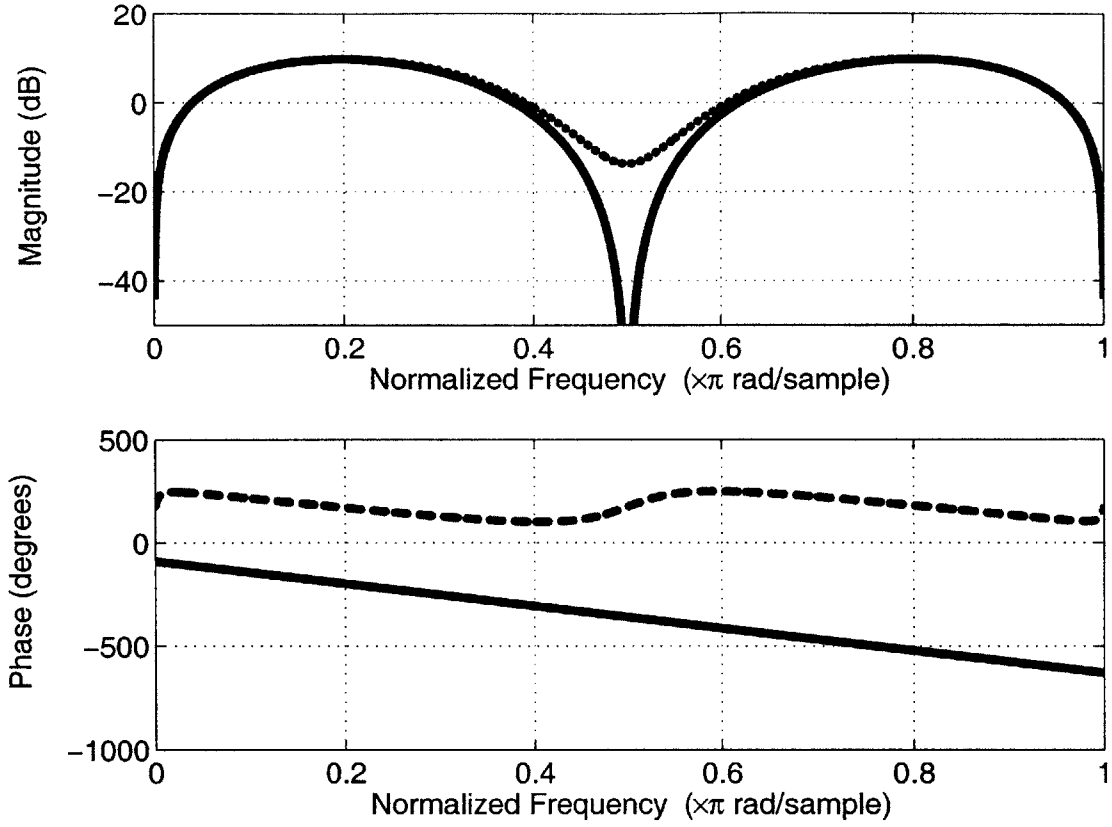

Figure 3-9: The solid line shows the frequency response of $H_d(e^{j\omega})$, while the dashed line corresponds to an approximation $H_a(e^{j\omega}) \approx H_d(e^{j\omega})$ where the double zeros corresponding to a normalized frequency of 0.5 have been moved inside the unit circle. In addition, the zeros corresponding to a normalized frequency of $\omega = 0$ and $\omega = \pi$ have been moved slightly within the unit circle. These zeros would result in an unstable inverse transfer function. Any spatial frequencies in the mirror corresponding to the zeros in $H(e^{j\omega})$ will be lost. For the case of $N = 2$, the zero location at a normalized frequency of 1 corresponds to spatial periods equal to the spacing h between the four beam spots on the mirror.

of the lowest bit in the phase measurement.

The angular measurements of the X and Y axis interferometers are shown in figure 3-11(a) and (b) respectively. We notice from the Y-axis interferometer measurement that the total rotation is on the order of $\alpha \approx \pm 10$ μ rads. This is within the specification provided by the manufacturer. Furthermore, the stage yaw (α) has been found to be on the same order of magnitude by utilizing a commercially available

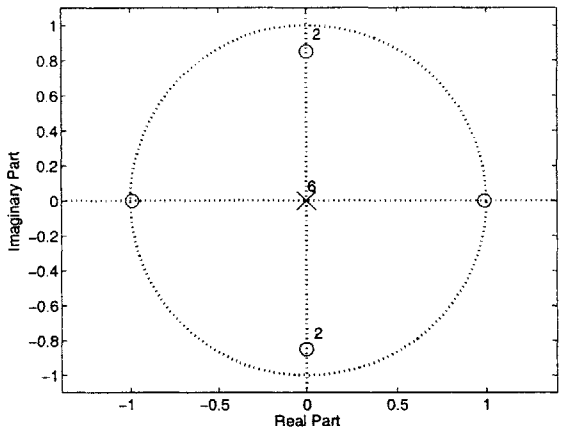

Figure 3-10: The zeros for the approximate transfer function $H_a(e^{j\omega}) \approx H_d(e^{j\omega})$ are shown. Since all of the zeros fall within the unit circle, the approximate transfer function H_a has a stable inverse. The zero locations for the approximate transfer function are $\{0.99, -0.99, z_0, -z_0, z_0, -z_0\}$ where $z_0 = 0.85e^{j\pi/2}$.

autocollimator. While there may be various sources for the yaw rotation of the stage, one potential cause is a result of deformations in the ceramic surface. The stage is an *Anorad* microglide T300L (Anorad Corporation, Hauppauge NY) model which floats on a granite surface using air-bearings. The stage is accelerated using frictionless magnetic forces. Deformations of the surfaces (granite and ceramic) on which the stage glides with the use of air-bearings, in addition to torques caused by magnetic forces may lead to small rotations of α.

For the purposes of our measurement, summing θ_x and θ_y will cancel the rotation α leaving only a function of the mirror distortion to be measured. In deriving the equations for θ_x and θ_y, we noted that each measurement would contain the rotation α but of a different sign. This is clearly seen in the data in figure 3-11, where both θ_x and θ_y measurements have a linear slope in common but of different sign.

The frequency content in our measurement $\Delta\theta$ is shown in the bottom of figure 3-12. A discrete Fourier transform is applied to the the output $\Delta\theta$. The frequency axis is normalized to the same dimensionless units as the transfer function in figure 3-9. One may observe that the majority of the power in the measurement resides within a normalized frequency of $-0.3 < \omega_n < 0.3$. For any arbitrary mirror shape, we can not

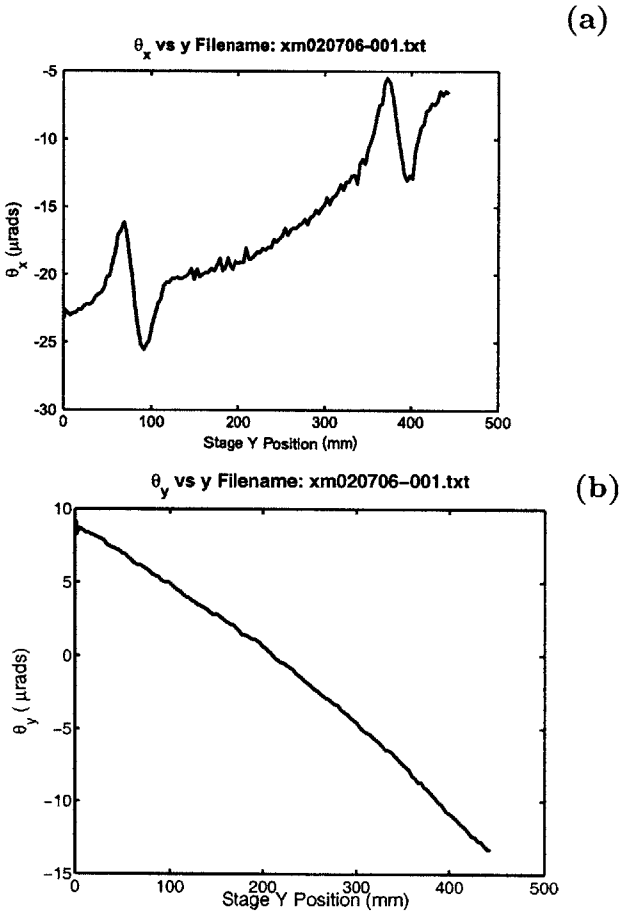

Figure 3-11: Measurements (a) $\theta_x(y)$ and (b) $\theta_y(y)$ are shown as a function of the stage coordinate y in units of microradians ($\mu rads$)

make any general statements about aliasing by simply expecting the data. However, for a mirror of optical grade we may make a few general statements. Namely, since the mirrors are polished we expect smoothing of high frequencies. Therefore, we conclude that most of the power would reside at low frequencies, with a shape in the power spectrum that decreases with increasing frequency. Inspection of the Fourier spectrum of $\Delta\theta$ in figure 3-9 suggests that the data is bandlimited to a normalized frequency of $\omega_{n-max} = 0.3$, which corresponds to a spatial frequency of $f_s/6 = \frac{1}{3h}$. This gives us confidence that we are sampling with a sampling rate that satisfies the Nyquist critera, since $f_{max} < f_s/2$.

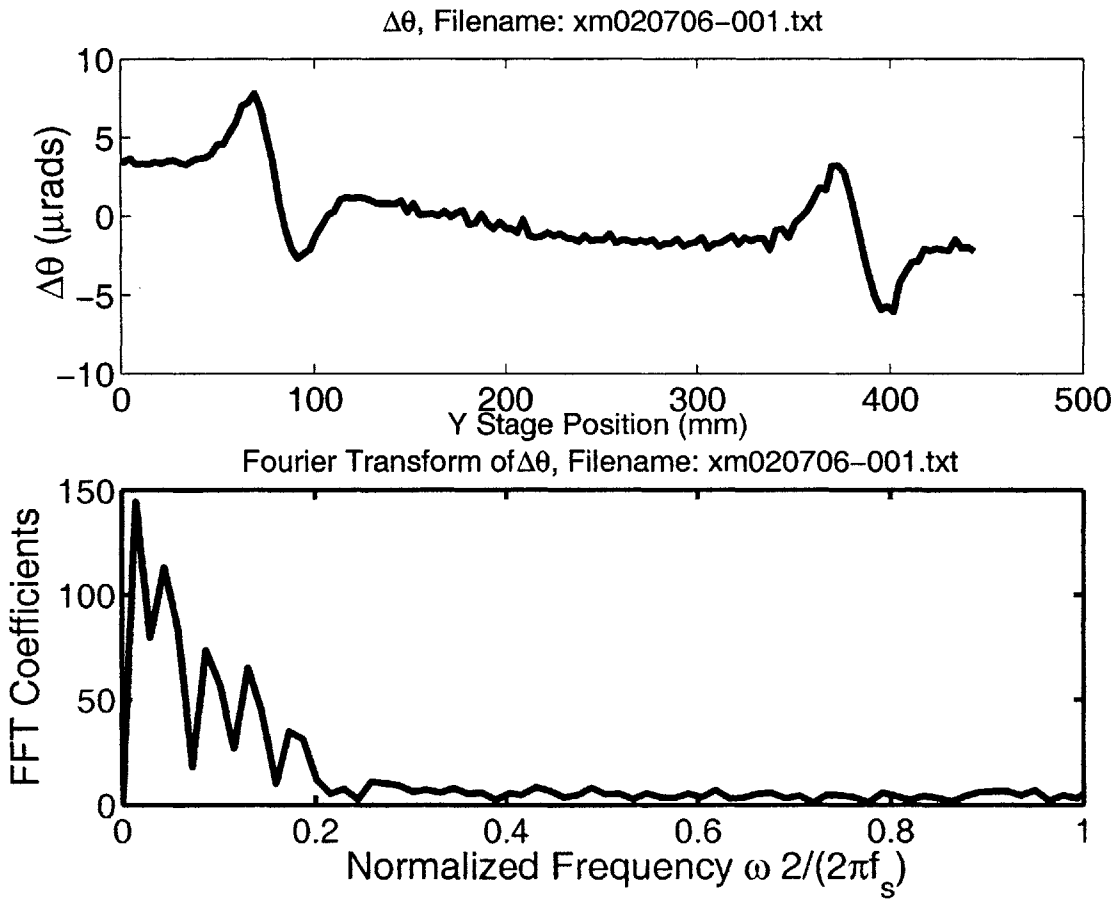

Figure 3-12: Top: Output $\Delta\theta = \theta_x + \theta_y$. Bottom: Fast Fourier transform of $\Delta\theta$. Most of the power resides below a normalized frequency of .3 which satisfies the Nyquist sampling criteria.

Applying the inverse filter, $H_a^i(e^{j\omega}) = |H_a^{-1}(e^{j\omega})| \approx |H_d^{-1}(e^{j\omega})|$ is a relatively simple process to implement in Matlab. One may construct a finite difference equation and apply the filter using a recursive loop. Alternatively, one may make use of the matlab filter function and input the filter coefficients. The inverse filter may be placed into the form:

$$H^i(e^{j\omega}) = G\frac{-4h}{a_0 + a_1 z^{-1} + a_2 z^{-2} + a_3 z^{-3} + a_4 z^{-4} + a_5 z^{-5} + a_6 z^{-6}} \tag{3.31}$$

where the filter coefficients are given by $a_0 = 1, a_1 = 0, a_2 = 0.4649, a_3 = 0, a_4 = -0.8942, a_5 = 0, a_6 = -0.5116$, and the gain constant is .7367. One may verify that this filter is the inverse to the approximate filter with the pole-zero diagram shown in figure 3-10. Applying this inverse filter to the discrete output $\Delta\theta[n]$ results in the the reconstructed mirror shown in figure 3-13. A best fit line is subtracted from the reconstructed mirror in order to obtain the mirror nonflatness. There are two pronounced features in the data at 75 and 375 mm. These features correspnd to locations where holes were made in the mirror for mounting purposes. In addition, there appears to be a small high frequency oscillation in the data. If one observes the Fourier transform of the data in the bottom of figure 3-13, a slight feature appears at a normalized frequency of $f_n = 0.5$. This is a direct consequence of the inverse filter.

The measurement transfer function which is applied to the data should remove any spatial frequencies corresponding to this normalized frequency of $\omega = 0.5$ as seen from the zero locations in the pole-zero plot in figure 3-10, and from the large attenuation shown in the transfer function in figure 3-9. However, one may observe that there is a background noise level in the data (noise floor). This uniform power spectrum is typically referred to as white noise. Temporal fluctations and other random processes may contribute to this background noise level. When the inverse filter is applied, any noise present in this portion of the power spectrum becomes amplified. In an ideal situation, the zeros of the measurement transfer function and the poles of the inverse transfer function would cancel each other out. Due to the finite precision of

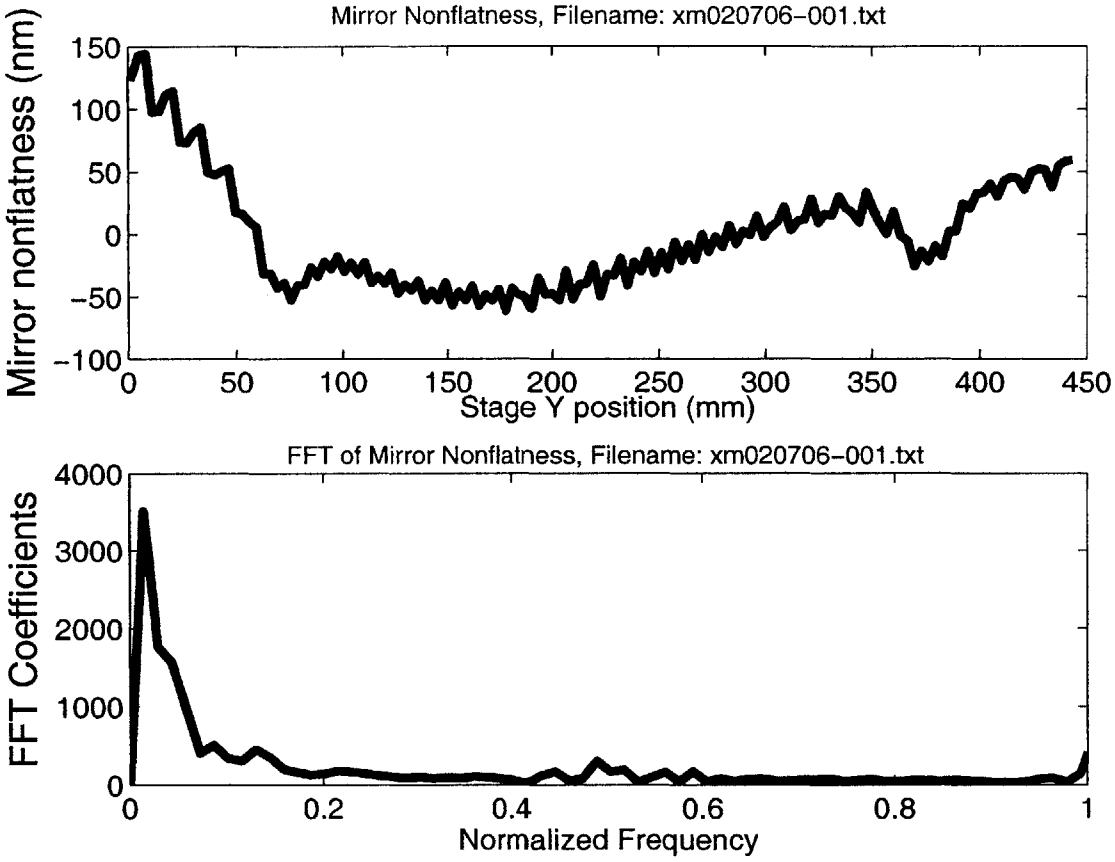

Figure 3-13: (Top) Measured mirror nonflatness after applying an inverse filter $H^i e^{j\omega}$ to the output $\Delta\theta$ and subtracting out the best fit line. (Bottom) Discrete Fourier Transform of the measured results.

our discrete signal processing and measurement noise, however, the poles which reside on the unit circle lead to an unstable inverse transfer function. The pole locations of the inverse transfer function are brought in closer to the origin of the unit circle (i.e., the radius is reduced) in order to reduce the gain of the transfer function.

If one wishes to remove the feature which causes the ripples in the data in figure 3-13 one may move the zeros of the approximate transfer function, at a normalized frequency of 0.5, further inward away from the unit circle. This leads to an inverse transfer function with less gain at these spatial frequencies. Alternatively, an inverse filter may be applied to the measured data followed by a low pass filter to reduce the

large gain at the normalized frequencies corresponding to the pole locations. Since we are dealing with linear systems the order of the cascade is not of importance. One may also apply the low-pass filter followed by an application of the inverse filter. This has an alternative interpretation of removing the noise at the unstable frequencies. The result of applying a low pass filter to the data in figure 3-13 is shown in figure 3-14.

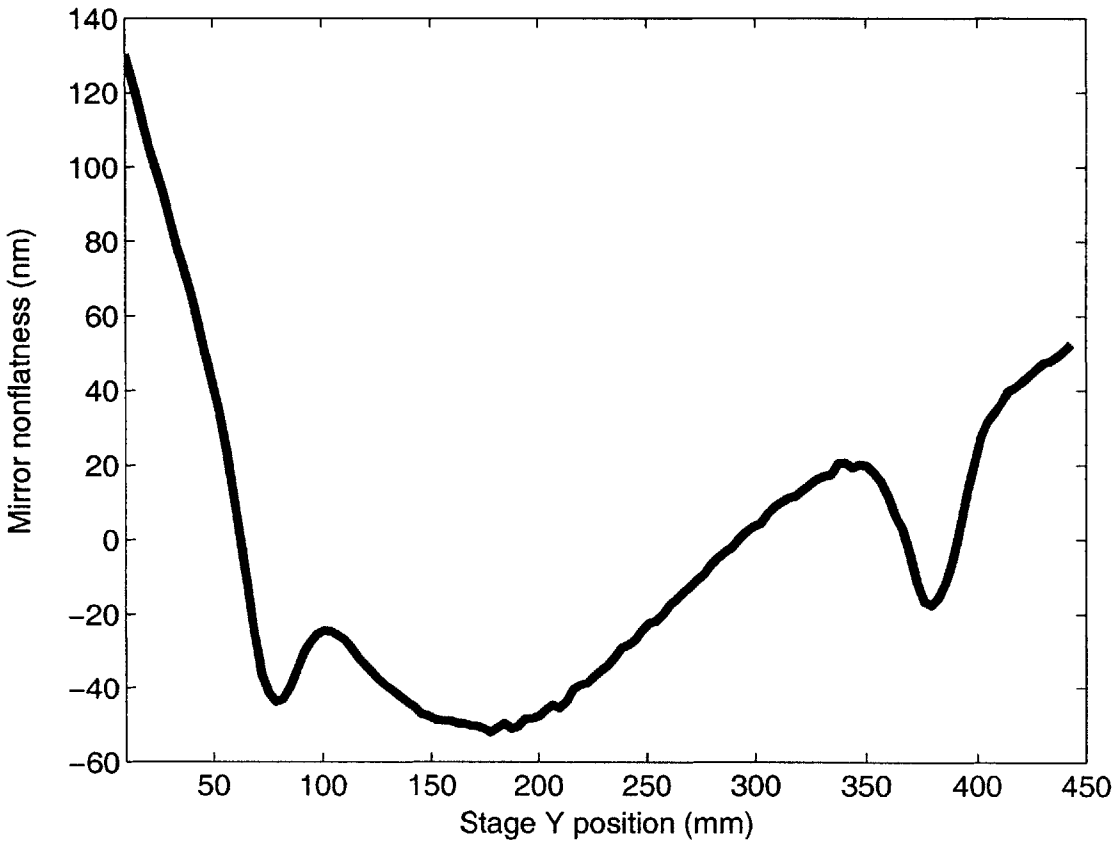

Figure 3-14: A low pass filter consisting of a 4 point moving average is applied to the reconstructed mirror profile. This filter removes the high frequency errrors introduced by the noise present at the pole locations of the inverse transfer function.

In summary, our discussion has allowed one to taylor a technique for reconstructing the mirror profile from angular measurements using angular/displacement interferometers [30]. A recipe has been provided which allows one to reconstruct the mirror profile by applying a stable inverse measurement transfer function. The exact form

of this transfer function has been given in equation 3.31. Alternative inverse transfer functions may be found which may provide a better approximation to the true inverse measurement transfer function as a result of minimizing an error criteria. As a result of the background noise in our measurements, one must be careful to reduce the noise floor by pre-filtering the output $\Delta\theta$ in order to reduce the large gain of the inverse transfer function at the pole locations.

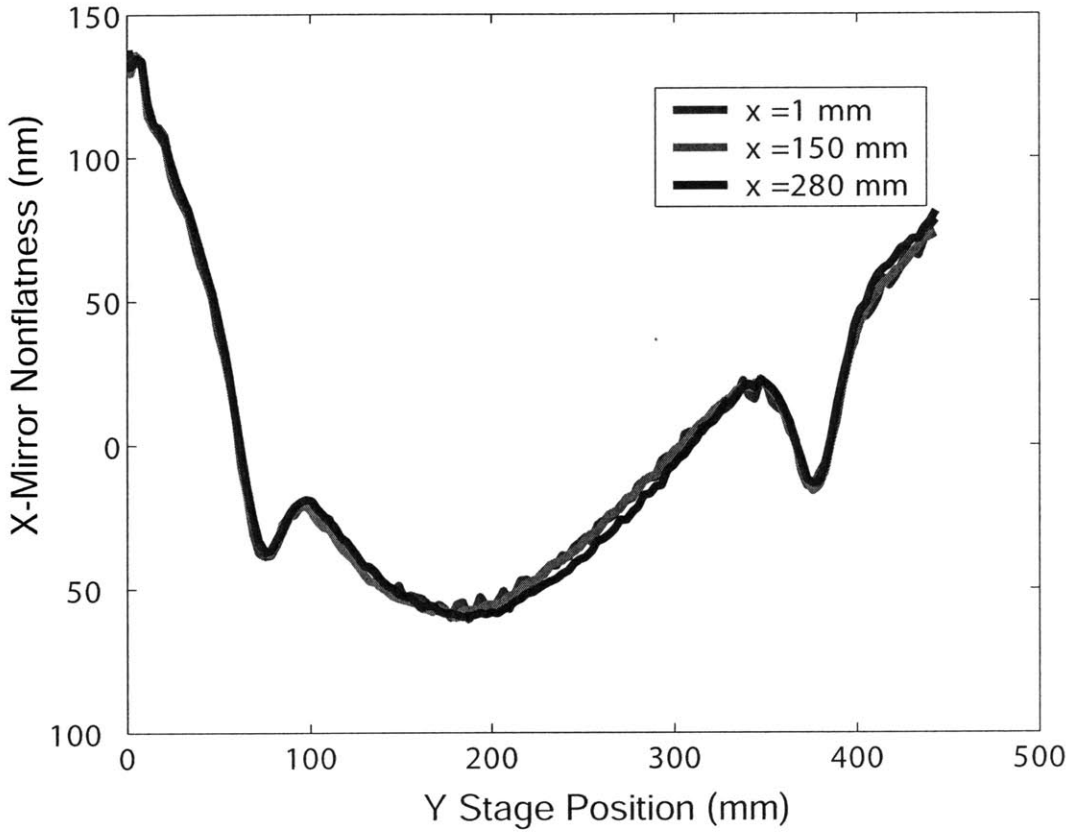

Figure 3-15: The X-axis stage mirror was scanned along the Y-axis for three different stage locations (x_0=1 mm, x_0 = 150 mm, x_0 = 280 mm). For each scan along Y, the mirror is discretely sampled using the X and Y axis interferometers. The results look similar for the various X locations as we expect.

In any event, our recipe for computing the mirror works well and results in a fairly smooth mirror profile as we would expect. From the analysis, we have shown that the exact x coordinate in which the X-axis mirror is scanned does not effect our data. In

order to test this claim, we scanned the stage X mirror along the Y-axis at various x locations. The result of our measurements for three different x coordinates are shown in figure 3-15. The results indeed look similar for the three different x locations as we would expect.

In the range of $y = 200 - 300$ mm, however, the scan corresponding to a stage location of $x = 280$ mm has an order of 10 nm difference from the other scans. Part of this error may be the result of air-refractive index variations which were not corrected for during the acquisition of this data. Furthermore, averaging was not used and would further reduce the noise in the measurement. Since the measurements overlap at various locations of the scan, we are overall satisfied and believe the X-mirror nonlinearity $D_x(y)$ and our measurement process to be invariant to the various stage locations under which the data is acquired to within a 10 nm tolerance.

Chapter 4

Contrast Considerations

The relationship between linewidth and dose in photolithography has been known for some time. Minimizing the variations in linewidth is important for a number of applications. For example, for diffractive grating applications the linewidth determines the diffraction efficiency of the diffracted wave. Similarly, distributed feedback photonic devices such as waveguides and photonic bandgap device's performance will largely depend on coherence (phase fidelity) and linewidth control (for efficiency). In this chapter, we will therefore study the various parameters which influence the linewidth in scanning beam interference lithography.

We will start off with an overview of a simple model which has been frequently used in photolithography. This model is known as the binary resist clipping model, and provides insight into the general relationships one would expect for an ideal photoresist behavior. More sophisticated modeling would include modeling of the volume of the resist structure. For example, anti-reflection coatings are used to prevent back reflection from the bottom of the substrate. This has been shown to eliminate standing waves normal to the surface which lead to "scalloping" of the sidewalls. In addition, post exposure baking has also been shown to reduce the roughness on the developed resist. For the purposes of this chapter, we will ignore these higher order effects which are modeled in the IC industry using commercial software such as PROLITH.

As we proceed, we will develop models to determine the contrast of the dose. Scanning beam interference lithography, as opposed to traditional interference litho-

graphy, uses a scanning method to create an exposure dose to cover large areas. This produces a set of stringent requirements unique to SBIL on scanning. It also provides a unique opportunity in that one could tune in to a certain period by choosing a set of scan parameters. We will find that in the case of parallel scanning, one may choose the period by appropriately choosing the step over distance. While this remarkable result was first discovered by one of my colleagues Dr. Carl Chen, a novel model will be presented in this section which will provide some keen insight. The former model requires a simulation in the spatial domain which hides many interesting details. In this chapter, we will develop a model in the Fourier domain, which will allow us to draw many conclusions without having to run any simulations. For example, I will show the relationship between the interfereing beams' spot size, the shape of the intensity, fringe period, and step over distance on the contrast of the dose.

In a previous chapter we described a new scanning technique known as Doppler writing. This method of writing requires that the fringes in the interference pattern synchronize with the stage motion, as the stage scans perpendicular to the interference fringes. In order to move the fringes in the interference pattern, the phase of the interference beams are modulated using acousto-optic modulators. For the first time, I will show how we can control the period using Doppler writing by choosing the appropriate phase modulation. Moreover, the relationship between the fringe period, phase modulation, and contrast will be presented. This has important implications for the next generation of SBIL, the so called variable period SBIL, in which Doppler writing will play an important role.

At the end of the chapter we will discuss some non-idealities in the wavefronts which may be incorporated into our model. For example, an assumption throughout the chapter is that the wavefronts of the interfering beams are planar. That is, the phase of the wavefronts are at most linear. For Gaussian beams, this condition is only met when the beams interfere at their waists. Previous studies by Chen have shown the effect of the placement of collimating lenses on the phase of the Gaussian wavefronts. In this section, we will discuss some predictions made by our model on the relationship between contrast and on wavefronts with nonlinear phase. In addition,

we will present some experimental results on measuring the wavefronts and we will compare algorithms used to extract the wavefront phase from an interference pattern.

4.1 Introduction: The relationship between contrast, dose, and linewidth

In this section, we will discuss the relationship between the average dose, the resist properites, and the resulting linewidth in scanning beam interference lithography. The dose will be denoted as $D(x)$ and will be defined to be the integral of the total intensity over time $D(x) = \int I(x)dt$. The image grating will be described as the intensity $I(x)$ which results from the interference pattern of two beams.

The dose contrast is a function of several parameters in scanning beam interference lithography. For example, the image-grating's shape, which is formed by the interference of two interfering beams, will have a large impact on the contrast of the exposure. In traditional interference lithography, two large beams interfere and the image grating is directly transfer into a photoresist-covered substrate. In scanning beam interference lithography, an image grating which exists within the overlap region of two interfering beams (typically 2mm) is scanned, stepped, and repeated until the entire wafer is exposed. The step size, and the influence of these parameters on the contrast will become clearer in the following sections of this chapter. For the purposes of an introduction, we will consider a clipping model where the image grating is assumed to be sinusoidal and we will reserve our discussion of scanning and the Gaussian nature of the interfering beams until later in this chapter.

The latent image is a term given to the pattern which is transfered onto photoresist. The latent image characteristics (such as shape) and linewidth will depend on the contrast of the exposure dose. When we speak of contrast, unless stated otherwise, we will for the purposes of this chapter be describing the contrast of the exposure dose. In this particular section, in which case scanning is not considered, the dose will be proportional to the intensity $I(x)$ in the image grating. By definition, the dose

is the integral of the intensity over time, and since our image grating is stationary the the proportionality constant between dose and intensity is the exposure time. So in this sense, the dose contrast will be proportional to the contrast of the intensity. However, as we proceed further in the chapter we will assume the contrast of the image grating is perfect, and study the unique features of SBIL that effect the contrast of the dose. If we wish to account for the non-ideal contrast of the image grating, it will be easy to incorporate into the analysis by scaling the resulting contrast with the image contrast.

Let us now clearly define what we mean by dose contrast. We will use a common definition for the contrast of a dose $D(x)$ with variation along the x-axis. Namely, the contrast γ may be defined in the following manner

$$\gamma = \frac{D_{max} - D_{min}}{D_{max} + D_{min}} \tag{4.1}$$

where D_{max} is the maximum dose and D_{min} is the minimum dose. For a sinusoidal dose with an offset of D_{avg} and an amplitude D_A descibed by

$$D(x) = D_{avg} + D_A sin(\frac{2\pi}{P_f}x), \tag{4.2}$$

the contrast γ defined in equation 4.1 is found to be $\gamma = D_A/D_{avg}$. Typically, a sinusoidal dose will be written to reflect this definition in a form

$$D(x) = D_{avg}[1 + \gamma \sin(\frac{2\pi}{P_f}x)]. \tag{4.3}$$

With the definition in place for contrast, we would like to now evaluate how the contrast affects the linewidth of our latent image. Strictly speaking, the latent image will require development in order for the pattern transfer to be complete. Similar indeed to the role of film in photography. Interestingly enough, the first commercial photo-resists were developed by Kodak. Nonetheless, we may attribute a linewidth to describe the spatial width of the region in which the light induced reaction occurs. In order to do so, we will use a simple model for the resist (known as the binary resist

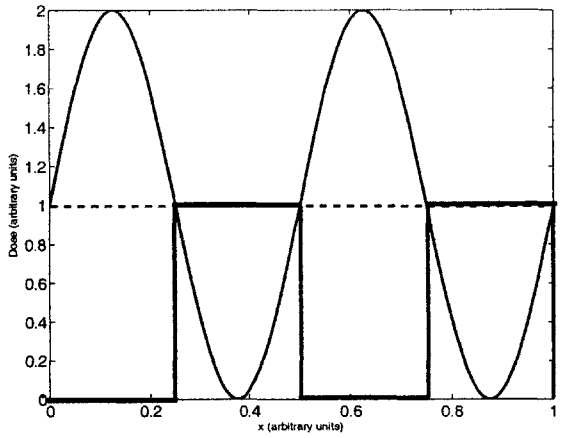

Figure 4-1: *Binary Resist model. A threshold dose of 1 (dimensionless units) is shown. After post-exposure development, any resist exposed with a dose greater than the threshold dose is removed. The result is the square profile shown, with a linewidth equal to half the period for a sinusoidal exposure with a contrast γ=1.*

or clipping model). Namely, any region in the photo-resist which is exposed above a threshold dose is removed by subsequent application of the photo-resist developer (for positive photo-resist). Figure 4-1 is an illustration of the clipping model.

Accordingly, a sinusioidal dose with a constant offset corresponding to $D_{avg} = 1$ in normalized units is shown in figure 4-1. This sinusoidal dose is also chosen to have a contrast $\gamma = 1$. The photoresist threshold D_Γ (which is usually determined empirically) is shown to equal the average dose $D_{avg} = 1$. Regions in the resist which are exposed with a dose greater than the threshold dose $D > D_\Gamma$ are removed by applying a developer. The regions in the resist which are exposed below the threshold remain. The rectangular pulse train which remains may be described by its width within a period which we refer to as the linewidth.

There are two basic parameters for a given resist dose threshold D_Γ that will affect the linewidth in the latent image. These parameters are the average dose D_{avg} and the dose contrast γ. Figure 4-2 illustrates the effect of changing the average dose D_{avg} on the latent image linewidth for a constant γ, in this case $\gamma = 1$. For the purposes of illustration, the threshold dose is shown to to be 0.5 in normalized units. The average dose for the three curves are greater than the threshold dose. The change in

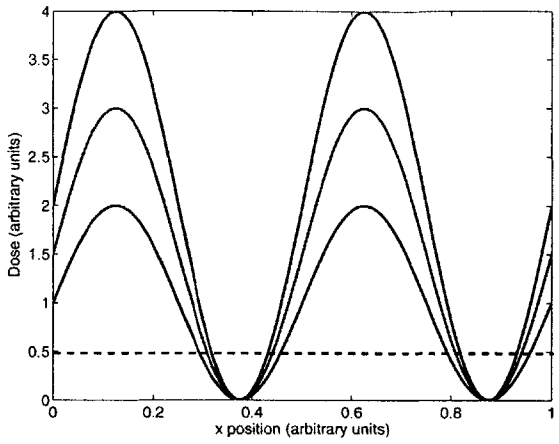

Figure 4-2: *Effect of Dose. Three different exposures with an average dose of 1, 1.5, and 2 for a fixed contrast of* $\gamma = 1$. *The threshold dose is shown in the dotted line. A change in linewidth will result in the developed latent image as the intersection of the dose and the resist's threshold dose changes.*

linewidth that results may be visualized by noting the change in intersection of the dose curves with the threshold dose.

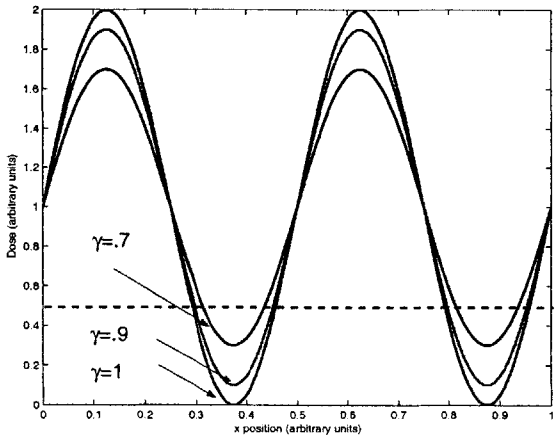

Figure 4-3: *Effect of Contrast. Three different dose profiles which have the same average dose but different contrast* ($\gamma = 0.7, \gamma = 0.9, \gamma = 1$) *are shown. The intersection with the dose threshold (dotted line) will separate regions which are removed upon development.*

In figure 4-3 I show the effect of varying the contrast for three exposures with the

same average dose D_{avg}. Shown are three dose profiles corresponding to $\gamma = 1, 0.9, 0.7$ each with an average dose of $D_{avg} = 1$. Similar to the case in which we changed the dose with a fixed contrast, changing the contrast with a fixed dose will also change the linewidth. This may be visualized by once again noting the change in the intersection of the dose with the threshold dose. For illustration purposes, a threshold dose of $D_\Gamma = 0.5$ is shown in figure 4-3 in a dotted line. For emphasis, we recall that any dose levels greater than the threshold dose will be removed during development.

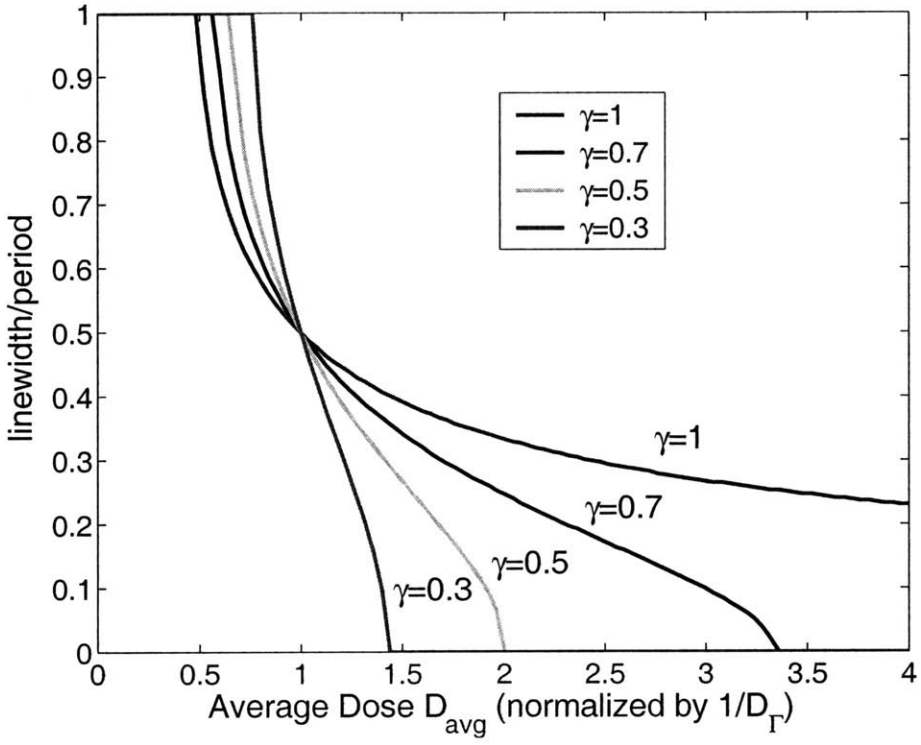

Figure 4-4: *Duty cycle (ratio of linewidth/period) as a function of the average dose, shown for different contrasts* γ

It is also worth noting from figure 4-3 that for each contrast the important parameter is the average dose relative to the threshold dose. For each contrast, we may then generate a family of curves by either increasing the threshold dose or by changing the average dose. One may also gain insight by examining three limiting cases of interest. In the case that the threshold dose D_Γ is equal to the average dose D_{avg},

a 50% duty cycle is achieved for all contrast variations. This is also shown in figure 4-4 which plots the duty cycle vs dose for different contrasts. In the case that the threshold dose is below the minimum dose, all of the latent image is removed during development. This is the case in which $D_{min} = D_{avg}(1-\gamma) > D_\Gamma$. We may also then solve for the minimum average dose in which complete overexposure occurs (i.e., the linewidth goes to zero) resulting in $D_{avg} = D_\Gamma/(1-\gamma)$. The latent image may then be said to be completely overexposed. In the ideal case, the linewidth will not go to zero when $\gamma = 1$ as shown in figure 4-3. Finally, in the case in which the maximum dose is below the threshold dose $D_{max} = D_{avg}(1+\gamma) < D_\Gamma$, all of the resist remains after exposure and the image is said to be overexposed. The maximum average dose in which this occurs may be expressed as $D_{avg} = D_\Gamma/(1+\gamma)$.

For emphasis, we may label the over-exposure limit $D_{over-expose}$ and the underexposure limit as $D_{under-expose}$ as

$$D_{over-expose} = \frac{D_\Gamma}{1-\gamma} \tag{4.4}$$

$$D_{under-expose} = \frac{D_\Gamma}{1+\gamma}. \tag{4.5}$$

The benefit of exposing with a high contrast is that the sensitivity to the average dose in the exposure to achieve a given linewidth is improved. For example, from equation 4.4 we find that as we increase γ the over-exposure dose in which the linewidth goes to zero is decreased. Thus far, the relationships we have described are common to resists and photolithography and serve as good introductory material. In the following sections, we will describe how contrast is achieved in SBIL. Obtaining good contrast is highly dependent on satisfying certain conditions on the stage scanning as we will see.

4.2 Parallel Writing Contrast

In this section, we will illustrate the effect of the stage scanning on the dose contrast. For simplicity, we will consider a one-dimensional intensity $I(x)$. For scanning beam

interference lithography we acknowledge that the intensity is in fact within a two-dimensional Gaussian envelope which is a function of both x and y. However, in the case of parallel scanning the y dependence is integrated out of the problem by scanning along a continuous y direction as shown in figure 4-5. At the end of the scan, the stage is translated in the x direction by a discrete step amount S and the scan is repeated now in the opposite y direction of the previous scan. This scanning mode is known as parallel scanning because the stage is scanned along a direction parallel to the interference fringes. Therefore, the problem is truly becomes one-dimensional. The first scan creates a dose profile along x given by $D(x) = \int (I(x, y - y_s)\frac{dy}{v}$, which is then stepped over along x to expose the entire wafer [7]. In order to keep the notation more managable, we will assume we have a one-dimensional intensity. Lets begin by assuming that we have two electric fields in the left and right arms given by $\overrightarrow{E_L}$ and $\overrightarrow{E_R}$, as written below:

$$E_L = A_L(x)\exp(jk_{Lx}x + jk_{Lz}z)\hat{y} \tag{4.6}$$

$$E_R = A_R(x)\exp(jk_{Rx}x + jk_{Rz}z)\hat{y}. \tag{4.7}$$

For the time being, we do not have to make any assumptions about the amplitudes $A_L(x)$, and $A_R(x)$, although we expect them to be Gaussian. Furthermore, we assume the x-components of the $\overrightarrow{k}$ vectors of the left and right arms are of the form $k_{Lx}\hat{x} = -k_{Rx}\hat{x} = \frac{2\pi\sin(\theta)}{\lambda}\hat{x}$, where $\lambda = 351$ nm and θ is the half angle between the beams, resulting in a period of $P_f = \frac{\lambda}{2\sin(\theta)}$. The z-components of the $\overrightarrow{k}$ vectors are equal $k_{Lz}\hat{z} = k_{Rz}\hat{z}$ and therefore do not play a role in the intensity. The half angle is currently chosen for a period of $P_f = 574$ nm. The average intensity $I(x)$ is calculated by taking the complex conjugate product of the sum of the fields $(E_L+E_R)(E_L+E_R)^*$, where E^* denotes the complex conjugate of E as shown below

$$I(x) = \|E_L\|^2 + \|E_R\|^2 + E_L E_R^* + E_R E_L^* \tag{4.8}$$

$$= \|E_L\|^2 + \|E_R\|^2 + 2Re\{E_L E_R^*\} \tag{4.9}$$

$$= \|A_L(x)\|^2 + \|A_R(x)\|^2 + 2A_L(x)A_R(x)cos(\frac{2\pi}{P_f}x) \tag{4.10}$$

Now we define the local contrast of the image grating component with spatial period P_f to be

$$\gamma_{image-grating}(x) = \frac{2A_L(x)A_R(x)}{\|A_L(x)\|^2 + \|A_R(x)\|^2}. \tag{4.11}$$

By local contrast we mean the contrast within a few periods P_f of location x in the intensity $I(x)$. This requires $A_L(x)$ and $A_R(x)$ to be slowly varying with respect to $\cos(\frac{2\pi}{P_f}x)$. This will be satisfied, for example, if the highest spatial frequency f_{max} in $A_L(x)$ and $A_R(x)$ is much less than $1/P_f$.

The more interesting question is what is the contrast of the dose when we step over, or sum the intensity. This could be represented mathematically as a summation, or equivalently by a convolution with an impulse train as shown in equation 4.12. Since the step over is along the direction x, we are assuming that the dose along the x direction of a single scan is proportional to the intensity. The proportionality constant τ is equal to the time T the intensity dwells at location x before the stage is stepped over for our one-dimensional treatment. For the full two-dimensional treatment of an intensity $I_{2d}(x, y)$, the one-dimensional dimensional intensity may be modified by $I(x) = \int I_{2d}(x, y)dy$ and the proportionality constant may be replaced by $\tau = 1/v$ where v is the velocity of the scan in the y direction. This functional replacement would provide the correct results. We will continue however with our one-dimensional treatment with this knowledge in mind.

The dose $D(x)$ may be expressed as a convolution with an impulse train $d(x)$ as

$$D(x) = \tau I(x) * d(x) \tag{4.12}$$

$$d(x) = \sum_{n=-\infty}^{\infty} \delta(x - nS) \tag{4.13}$$

where S is the step-over distance, and the impulse train expresses the process of translation and summation. In order to calculate the contrast of the dose, $D(x)$, we simply need to look at the ratio of the radian spatial frequency component $\omega = 2\pi f$ at $2\pi/P_f$ to the DC component. Since the transform of a convolution is the product

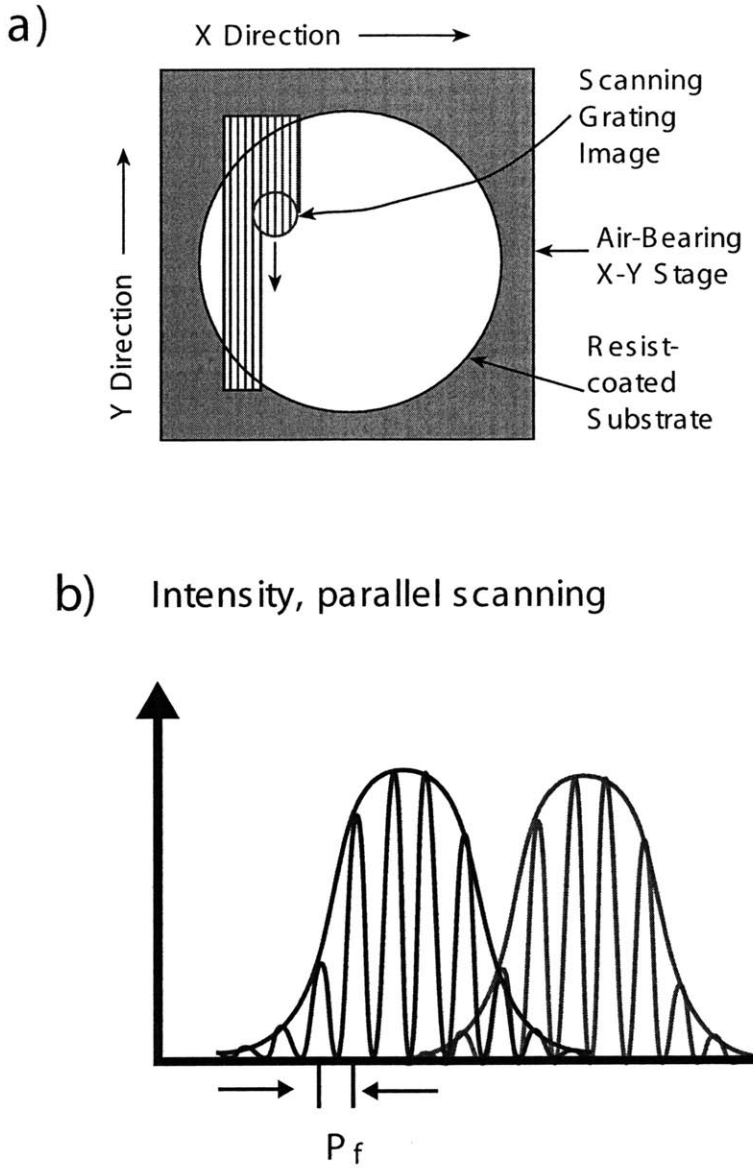

Figure 4-5: *(a) Parallel scanning concept. The image grating is scanned in the y direction. At the end of the scan, the stage steps over a distance S in the x direction. The scan direction y is then reversed and the new scan overlaps with the previous scan. (b) The image grating consisting of a Gaussian envelope modulated with carrier fringes. The stage is stepped over a discrete amount. For optimal contrast, the step over distance should be an integer number of grating periods so that the fringes of the new scan overlap with the previous scan.*

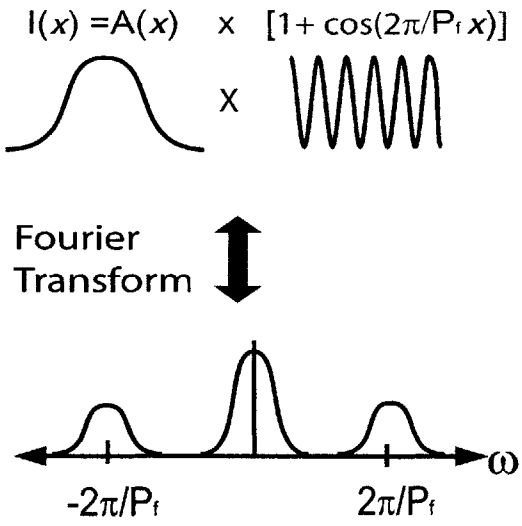

Figure 4-6: *Cartoon illustrating the Fourier transform of the intensity.*

of the transforms, and since the transform of a comb of spatial delta functions is a comb of frequency delta functions,

$$\bar{D}(\omega) = \bar{I}(\omega)\bar{d}(\omega) \tag{4.14}$$

$$\bar{d}(\omega) = \frac{2\pi}{S}\sum_{n=-\infty}^{\infty}\delta(\omega - \frac{2\pi n}{S}), \tag{4.15}$$

where the overscore denotes the Fourier transform. The Fourier transform of the intensity may be found with the aid of figure 4-6. For the case $A_L(x) = A_R(x)$, the intensity of the form given in equation 4.10 may be visualized as the product of $A(x) = 2|A_L(x)A_R(x)|$ with a function of the form $[1 + cos(\frac{2\pi}{P_f}x)]$. In frequency space, the cosine term within the brackets translates and scales the Fourier transform of A(x), $F\{A(x)\}$, to spatial frequencies $\frac{2\pi}{P_f}$, $-\frac{2\pi}{P_f}$. The scale factor is 1/2. The constant offset 1 within the bracketed term transforms into an impulse and therefore the convolution of the Fourier transform of $F\{A(x)\}$ with an impulse results in the identity system. The overal result shown in figure 4-6 is that the Fourier transform $F\{A(x)\}$ is just a replicated at spatial frequencies 0, $2\pi/P_f$, $-2\pi/P_f$, and scaled by a factor of $1/2$ at spatial frequencies $2\pi/P_f$, $-2\pi/P_f$.

Thus, the frequency spectrum of $D(x)$ is discrete since the product of the Fourier transform of $I(x)$ and a discrete impulse train $\bar{d}(\omega)$ is discrete. The impulse train therefore selects the spatial frequency components $\omega = \frac{2\pi n}{S}$ of the Fourier transform $F\{I(x)\}$ as shown below

$$\bar{D}(\omega) = \bar{D}(\frac{2\pi n}{S}) = \tau \frac{2\pi}{S} \bar{I}(\frac{2\pi n}{S}), \tag{4.16}$$

where n is an integer and S is the step-over distance, and the overscore denotes the Fourier transform. It is worth noting that it is a general property of Fourier transforms that the Fourier transform of a periodic signal is discrete. If the step-over distance is chosen correctly, such that it is an integer number of periods $S = NP_f$, the spatial frequency spectrum of the dose will have a component at $\omega = 2\pi/P_f$ (i.e., for $n = N$ in equation 4.16). Otherwise, we note that we could slightly change the fundamental period of the dose by modifying the step-over distance. Let us assume that we choose the step over distance to be an integer number of grating periods $S = NP_f$. The contrast at $\omega = 2\pi/P_f$ is then found to be

$$\gamma = \frac{2\bar{D}(\frac{2\pi}{P_f})}{\bar{D}(0)}. \tag{4.17}$$

The Fourier transform $2\bar{D}(\frac{2\pi}{P_f})$ could easily be verified to be

$$2\bar{D}(\frac{2\pi}{P_f}) = \frac{2\pi}{S}\tau < 2A_L(x)A_R(x) > \tag{4.18}$$

where $<>$ denotes the DC component given by $< f(x) >= \int_{-\infty}^{\infty} f(x)dx$. In light of figure 4-6 and our previous discussion, the Fourier transform component of $D(x)$ at spatial frequency $\frac{2\pi}{P_f}$ is just the translated and scaled (by a factor of π/S) replica of the DC component of the Fourier transform of $A(x) = 2A_L(x)A_R(x)$.

Interestingly then, one could have guessed the solution by taking the spatial average of the image grating contrast given in equation (4.11). In conclusion, the contrast

for given envelopes $A_L(x)$, and $A_R(x)$ as a result of discrete scanning is given by

$$\gamma = \frac{< 2A_L(x)A_R(x) >}{< \|A_L(x)\|^2 + \|A_R(x)\|^2 >} \tag{4.19}$$

Prior to this analysis, it would have been difficult to calculate the resulting constrast for different beam profiles $A_L(x)$ and $A_R(x)$. Before moving onto examples in which this Fourier analysis will be helpful (such as in beam overlap, and period error) I should make a few comments about the step over distance. The contrast given in equation 4.19 will be meaningful if there is only one spatial frequency component $\omega = \frac{2\pi}{P_f}$ in the dose. To insure that this is the case, the step over distance should satisfy the Nyquist limit such that $2\pi/S$ is larger than the maximum spatial frequency of $A(x)$. This is always satisfied in our scanning if we approximate $A(x)$ to be bandlimited to $2\pi/L$ where L is the beam diameter. Therefore, S must always be chosen such that $S < \frac{1}{L}$. This will insure that a uniform dose is achieved, and that only one spatial frequency component is present. Finally, we will also make a note about the limits in the integrals. We have chosen to use an impulse train $d(x)$ that extends over all space where the summation limits are over infinity. In doing so, we are obtaining what we refer to as the steady state response and are therefore neglecting the transients at the beginning of the scan which are not of interest. Physically, this corresponds to the dose one would expect at the center of the substrate where the dose corresponds to a summation of the nearest neighbor scans. Therefore, a dose in the center will be representative of a steady state response since the dose at the center is equivalent to a dose calculated using the infinite limits in the sum.

4.2.1 Example: Beam Overlap Error

In this section, we will carry forward our Fourier analysis to determine the contrast due to two displaced Gaussian envelope functions. This example will be useful in considering the contrast degradation and alignment tolerances on our interfering beams.

Consider the case where $A_L(x)$ and $A_R(x)$ are two displaced Gaussian functions:

$$A_L(x) = E_L exp(\frac{-x^2}{R^2}) \tag{4.20}$$

$$A_R(x) = A_L(x - x_0) \tag{4.21}$$

where R is the $1/e^2$ radius (typically 1 mm) of the product $A_L(x)A_R(x) = E_L^2 exp(\frac{-2x^2}{R^2})$ when the displacement $x_0 = 0$. The contrast as a function of the displacement is given below. One could verify that evaluating the Fourier transform in equation 4.17 results in a Gaussian contrast function of radius $R = 2L$. This could be found by using the Fourier transform relationship of a Gaussian function given by,

$$e^{\frac{-x^2}{R^2}} \Leftrightarrow \sqrt{\pi} R e^{\frac{-\omega^2 R^2}{4}}, \tag{4.22}$$

and by using the convolution property. For example the numerator of Equation (4.19), $2 < A_L(x)A_R(x) >$, is calculated by using the convolution property along with the fact that $A_L(x)$ is even symetric. In other words,

$$< A_L(x)A_R(x) >= \int_{-\infty}^{\infty} A_L(x)A_R(x)dx \tag{4.23}$$

$$= \int_{-\infty}^{\infty} A_L(x)A_L(x - x_0)dx \tag{4.24}$$

$$= \int_{-\infty}^{\infty} A_L(x)A_L(x_0 - x)dx \tag{4.25}$$

$$= A_L(x_0) * A_L(x_0) \tag{4.26}$$

where the last operation denotes convolution. In going from equation 4.24 to 4.25 we have used the even-symmetry of the Gaussian function $A_L(x)$. Furthermore, by recognizing that a convolution in space corresponds to inverse Fourier transform of the product of Fourier transforms in frequency, we have

$$A_L(x_0) * A_L(x_0) = F^{-1}\{\bar{A}_L(x_0)\bar{A}_R(x_0)\} \tag{4.27}$$

$$= F^{-1}\{\pi R^2 e^{\frac{-\omega^2 R^2}{4}} e^{\frac{-\omega^2 R^2}{4}}\} \tag{4.28}$$

$$= \pi R^2 F^{-1}\{e^{\frac{-\omega^2 R^2}{2}}\} \tag{4.29}$$

$$= \sqrt{\frac{\pi}{2}} R e^{\frac{-x_0^2}{2R^2}}. \tag{4.30}$$

We could proceed to calculate the denominator of Equation (4.19), but we recognize that it will result in a constant. Furthermore, for $x_0 = 0$ we note that the contrast γ equals one. This allows us to determine that the denominator of Equation (4.19) will equal $2\sqrt{\frac{\pi}{2}}R$, the value of the numerator in Equation 4.19 when $x_0 = 0$. This analysis allows us to conclude that for a beam overlap error of x_0 the contrast is given by

$$\gamma(x_0) = exp(-x_0^2/2R^2). \tag{4.31}$$

The value of x_0 which results in a contrast of $1/e^2$ is related to the beam spot size radius by $x_0 = 2R$. Physically, this result makes sense because when the adjacent scans are stepped over by a diameter, the beams no longer overlap as shown in figure 4-7 (a). This in turn could be used as a design criteria to calculate the depth of focus. For example, given that the beams overlap at a particular location on a substrate as shown in figure 4-7 (b), one is interested in knowing the tolerable substrate height variation h to obtain good contrast. Using the relationship $tan(\theta) = \frac{x_0}{2h}$, where h is the height of the substrate and θ is the half angle of the interfering beams, we find the displacement x_0 is equivalent to $x_0 = 2h\tan^{-1}(\theta)$. For a period of 574 nm, the half angle is $\theta = 0.3057$. We therefore find $h = 0.5x_0/0.30$. At a displacement $x_0 = 229.6$ μm, the contrast has degraded by 10 percent ($\gamma = .9$). This corresponds to a change in substrate height of $h = 382.7$ microns for a 1 mm beam radius R.

4.2.2 Example: Period Error Tolerance

In this section I will find the relationship between the step over error and contrast tolerance for two overlapped Gaussian beams. Specifically, I will assume that the envelopes for the left and right arms are equal as in $A_L(x) = A_R(x)$ so that the intensity may be described in the form $I(x) = A(x)[1 + cos(\frac{2\pi}{P_f}x)]$, where $A(x) = e^{-2x^2/R^2}$. If the step over is chosen to be an integer number of grating periods, $S = NP_f$, the contrast may easily be determined by making use of equation 4.17.

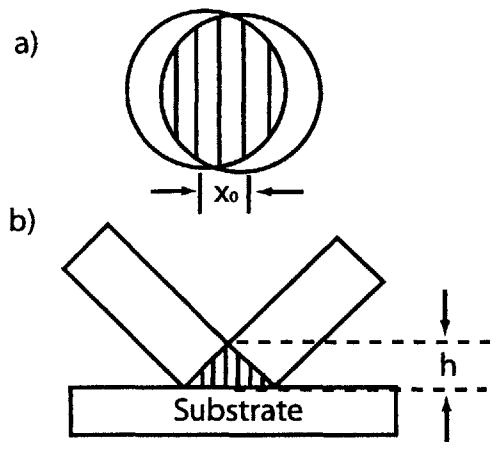

Figure 4-7: *(a) Top view of the substrate where the beams are translated by x_0. A grating exists only where the beams overlap. (b) View showing the depth of focus h. If the substrate thickness varies by h, the beams will no longer overlap on the substrate.*

Specifically, the frequency component $\bar{D}(\frac{2\pi n}{S})$ for $n = N$ is equivalent to the DC component of $A(x)$ scaled by a factor of $\tau\pi/S$. This leads to an ideal contrast $\gamma = 1$, since the DC component $\bar{D}(0)$ is equivalent to the DC component of $A(x)$ scaled by a factor of $\tau 2\pi/S$. Now we are interested in the case in which the step over distance is not equal to an integer number of grating periods. For example, we may change the period in the grating by simply stepping over by a distance $S = N(P_f + \Delta P)$. However, the contrast will suffer.

The tolerance on ΔP may be determined by making reference to figure 4-8. In figure 4-8, we illustrate in the dashed curve that if the fringes of the period were equal to $P_f + \Delta P$, our step over distance would sample the DC frequency component of $A(x)/2$. However, since there is an error of ΔP, our step over distance samples the Fourier component $F\{A(x)\}/2$ at frequency $\omega = \frac{2\pi\Delta P}{P_f^2}$. These considerations allow

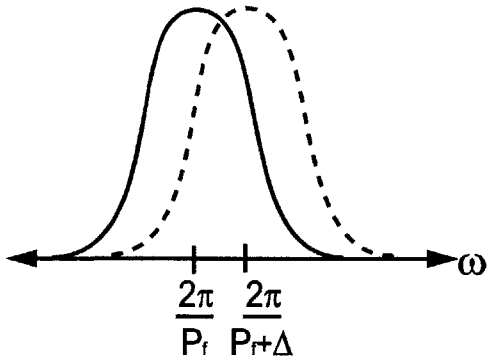

Figure 4-8: *A step over distance is chosen such that* $S = N(P_f + \Delta P)$, *where the period error is given by* ΔP. *For small* ΔP, *this corresponds to sampling the Fourier transform of* $A(x)/2$ *at* $\omega = \frac{\Delta P}{P_f^2}$

us to express the contrast as a function of period error. We will use the property that the Fourier transform of a Gaussian beam with $1/e^2$ intensity radius is

$$e^{\frac{-2x^2}{R^2}} \Leftrightarrow \frac{\sqrt{2\pi}R}{2} e^{\frac{-\omega^2 R^2}{8}}. \tag{4.32}$$

Interestingly, the $1/e^2$ frequency response corresponds to $\omega_{1/e^2} = 4/R$ or $f_{1/e^2} = 4/(2\pi R)$. This may also be expressed in terms of the beam diameter L as in $f_{1/e^2} = 4/(\pi L) \approx 1.27/L$. The back of the envelope calculation that tells us the frequency content of a beam of diameter L is bandlimited to $f = 1/L$ is not a bad approximation in light of this analysis. In any event, the contrast as a function of a period error ΔP

may be expressed by evaluating the ratio

$$\gamma = \frac{2F\{A(x)\}|_{\omega=2\pi\Delta P/P_f^2}}{\hat{D}(0)}, \tag{4.33}$$

resulting in

$$\gamma(\Delta P) = e^{-\frac{(2\pi\Delta P/P_f^2)^2 R^2}{8}}. \tag{4.34}$$

Now let us apply some numbers to this example. I will explore the limit for the value of ΔP in which $\gamma(\Delta P) = 1/e^2$. I have already showed that this occurs for a spatial frequency $f = \frac{\Delta P}{P_f^2} = \frac{1.27}{L}$. This leads to a design rule $\frac{\Delta P}{P_f} < \frac{1.27 P_f}{L}$. For $P_f = 574$ nm, a beam spot diameter $L = 2$ mm, $\frac{\Delta P}{P_f} = 1.27\frac{P_f}{L} = 365$ parts per million or ppm. This tells us that the ratio $\frac{\Delta P}{P_f} < 365$ ppm as the limit for $1/e^2$ contrast degradation, or a period error of $\Delta P = 0.2097$ nm.

As we alluded to in the introduction, Chen has performed an analysis of contrast error (which he refers to as dose amplitude error) using a spatial domain approach [7]. His technique does not provide a closed form relationship, and requires a simulation of the multiple step over distances to produce a dose error plot. Here we offer an alternative solution which provides a closed form equation. The advantage in our solution is that we explicitly can determine the dependence of the radius R, fringe period P_f, and period error ΔP on the contrast γ. Moreover, my model is consistent with the contrast predictions of my colleague's simulation.

4.2.3 Experimental Results for Period Error

In this section, we will discuss the results of an experiment which is performed to examine the relationship between period error and dose contrast. In order to obtain an expression for the average dose, we will use a two-dimensional intensity of the form $I_{2d}(x, y)$. As we scan along the first strip, the average dose will be given by $D(x) = \int I_{2d}(x, y)\frac{dy}{v}$. The subsequent sampling along x may be considered as a convolution with an impulse train $d(x)$ which results in the total dose given by $D_{total} = D(x) * d(x)$.

The average dose is then found to be

$$D_{avg}(x) = \frac{1}{S}\int_{-S/2}^{S/2} I_{2d}(x,y)\frac{dydx}{v} \tag{4.35}$$

This may be expressed in units of $\mathrm{mJ/cm^2}$ in terms of the total power P in the beam as

$$D_{avg} = \frac{100P}{Sv}, \tag{4.36}$$

where P is the total power in the interfering region given in mW, v is the stage velocity in mm/s, and S is the step-over distance between scans in mm.

I simulated a step over error as a result of a period error by adding a phase error between subsequent scans. Notably, stepping over by the correct amount corresponds to stepping over by an integer number of periods such that the fringes of the new scan overlap with the fringes of the previous scan. If a period error exists, there will be a phase error between subsequent scans. This phase error may be examined by evaluating the phase at the end of the step over distance S as in

$$\frac{2\pi S}{P_f + \Delta P} \approx \frac{2\pi S}{P_f} + \frac{2\pi S}{P_f}\frac{\Delta P}{P_f}, \tag{4.37}$$

where the step over distance was chosen to be $S = NP_f$, but the actual period in the image grating is given by $P_f + \Delta P$. This may be expressed as an equivalent phase error between scans as

$$\theta_\epsilon = \frac{2\pi S}{P}\frac{\Delta P}{P_f}. \tag{4.38}$$

Typically one prefers to scale the phase θ_ϵ by a factor $w = \frac{2\pi}{P_f}$ to produce a phase in nanometers. An experiment was performed corresponding to a phase error of $w\theta_\epsilon = 0, 30, 60$ nm. For a period of $P_f = 574$ nm, this corresponds to an equivalent phase error of $\frac{\Delta P}{P_f} = 0, 70, 140$ parts per million respectively for a step over distance $S \approx 0.4287$ mm (rounded to the nearest integer multiple of P_f). For a beam spot of radius $R = 1mm$, this corresponds to a contrast degredation of 0,0.07, and 0.25 respectively (i.e., $\gamma = 0.93, 0.75$). The results of these experiments are shown in Figure ?? taken from [31]

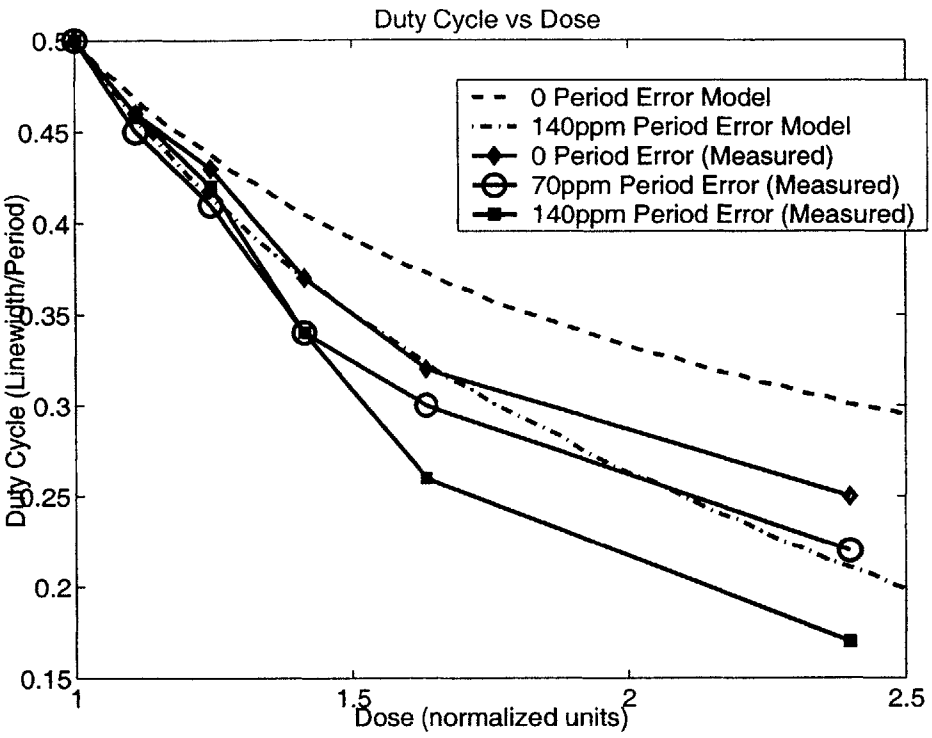

Figure 4-9: *Measured resist duty cycle (linewidth/period) vs dose for different period errors for 574.303 nm period gratings. The dashed and dotted lines show the result of a simulation where a period error of 0 and 140 parts per million are used. The solid lines indicate the results of measurements for a 0, 70, and 140 ppm period error. All plots intersect at the clearing dose, where the duty cycle is 50 %.*

The slope of the linewidth vs dose curve gives an indication of the dose contrast. For high contrast, it is possible to overexpose beyond the clearing dose with less sensitivity to dose. For poor contrast, a more rapid decrease in linewidth is observed for overexposing beyond the clearing dose. For these experiments, wafers coated with 78 nm-thick Brewer i-con-4 ARC, and 570 nm-thick Sumitomo PFI-88 resist were used.

The experimental results and simulations show the trend one would expect. There is some discrepency, however, and the measured linewidths do not overlap perfectly with the simulations. One explanation for this involves using an oversimplified resist model. For this simulation, a binary resist model is used. Improved simulations

would involve modeling higher-order effects in the resist, such as absorption, back reflections, and resist chemistry. In addition, other sources of contrast degradation may be present which are not being accounted for. Moreover, we have estimated many parameters such as the beam spot size, the shape of the Gaussian beams, and have assumed the image grating contrast to be 1. In additon, we have made the assumption that we know the period P_f to well within the uncertainty required to conduct these experiments. The repeatability of our period measurements are on the order of 2 parts per million.

4.3 Doppler Writing Contrast

4.4 Introduction

In this section we will consider the equivalent of a period error for Doppler writing. Recall that Doppler writing is a patterning technique in which the stage is scanned perpendicular to the interference fringes. Similar to the section in parallel writing, we will begin by assuming that we have two electric fields polarized along the Y axis in the left and right arms given by E_L and E_R

$$E_L = A_L(x)\exp(jk_{Lx}x + jk_{Lz}z)\hat{y} \tag{4.39}$$

$$E_R = A_R(x)\exp(jk_{Rx}x + jk_{Rz}z + j\Delta\Phi)\hat{y}. \tag{4.40}$$

We have allowed for a phase offset of $\Delta\Phi$ in equation 4.40. This phase will be controlled by our frequency synthesizer which drives an acousto-optic modulator. For the time being, we do not have to make any assumptions about the amplitudes $A_L(x)$, and $A_R(x)$, although we expect them to be Gaussian. Furthermore, we assume the x-components of the $\overrightarrow{k}$ vectors of the left and right arms are of the form $k_{Lx}\hat{x} = -k_{Rx}\hat{x} = \frac{2\pi\sin(\theta)}{\lambda}\hat{x}$, where $\lambda = 351$ nm and θ is the half angle between the beams, resulting in a period of $P_f = \frac{\lambda}{2\sin(\theta)}$. The z-components of the $\overrightarrow{k}$ vectors are equal $k_{Lz}\hat{z} = k_{Rz}\hat{z}$ and therefore do not play a role in the intensity. The half angle is

currently chosen for a period of $P_f = 574$ nm. The average intensity is calculated by taking the complex conjugate product of the sum of the fields $(E_L + E_R)(E_L + E_R)^*$, where E^* denotes the complex conjugate of E.

In order to find the intensity, we proceed to calculate

$$I(x) = \|E_L\|^2 + \|E_R\|^2 + E_L E_R^* + E_R E_L^* \tag{4.41}$$

$$= \|E_L\|^2 + \|E_R\|^2 + 2Re\{E_L E_R^*\} \tag{4.42}$$

$$= \|A_L(x)\|^2 + \|A_R(x)\|^2 + 2A_L(x)A_R(x)cos(\frac{2\pi}{P_f}x - \Delta\Phi). \tag{4.43}$$

We will define Doppler writing as patterning in the direction perpendicular to the interference fringes (i.e., x), at a velocity v (see the chapter on Doppler writing for a detailed explanation). The stage reference frame X' is moving with respect to the laboratory reference frame X at velocity v yielding the relation $X' = X + \int v dt$. We will define $x_s = \int v dt$ to be the position of the stage reference frame with respect to the laboratory frame. Notably, an object located at position x in the laboratory reference frame may be expressed as being at position

$$x' = x - x_s \tag{4.44}$$

in the stage frame since the origin of the coordinate frame X is located at distance $-x_s$ with respect to the origin of X'. The author acknowledges that constantly changing between reference frames may appear to be unecessarily confusing. A reference diagram is provided in figure 4-10 for clarity. The only result we will need is the relation given in equation 4.44 which allows us to relate an object's coordinate in the stage frame with coordinate x' or in the laboratory frame with coordinate x. This will be useful in describing the intensity which is stationary in the laboratory reference frame. As shown in the diagram in figure 4-10 as the stage moves along positive X in the laboratory frame, the intensity appears to move to the left along negative X' in the stage frame.

The intensity in the stage reference frame may then be written by simply substi-

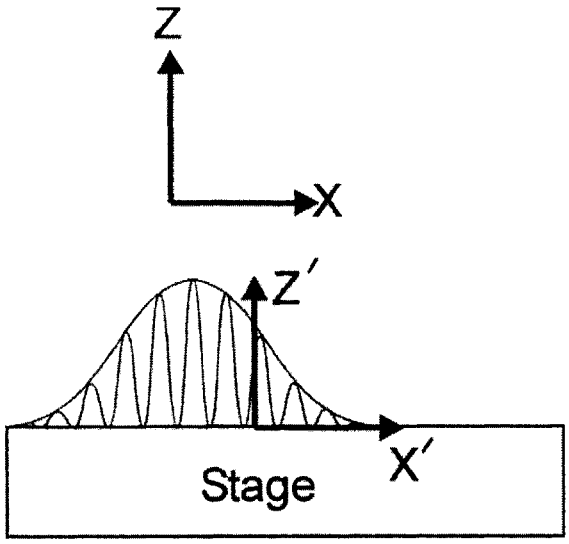

Figure 4-10: *The laboratory frame $X-Y-Z$ and the moving stage frame $X'-Y'-Z'$ are shown. For an observer in the laboratory frame, the intensity appears to be stationary. For an observer in the stage frame, the intensity appears to be moving to the left for a positive velocity v of the stage as measured in the stationary frame.*

tuting $x = x' + x_s$ for x in equation 4.43 resulting in

$$I'(x',t) = \|A_L(x'+x_s)\|^2 + \|A_R(x'+x_s)\|^2 + 2cos[\frac{2\pi}{P_f}(x'+x_s) - \Delta\Phi]G(x'+x_s). \quad (4.45)$$

where

$$G(x'+x_s) = A_L(x'+x_s)A_R(x'+x_s). \quad (4.46)$$

We note that the sign is correct in equation 4.45 since the intensity appears to shift to the left in the stage frame (along negative x') for positive values of x_s.

In order to calculate the dose, we must integrate over time and change variables to x_s [27]. Explicitly, we may write

$$D(x') = \int I'(t)dt \quad (4.47)$$

$$D(x') = \int I(x'+x_s)\frac{dx_s}{v}, \quad (4.48)$$

where the variable change of dt to dx_s is accomplished by recognizing $\frac{dx_s}{dt} = v$. As

a side note, we may gain insight by comparing the continuous Doppler scan case to that of a discrete step over in parallel scan. In the case of a discrete scan, we found that the dose resulted in a summation as given by equations 4.12 and 4.13. The summation was represented as a convolution of the intensity with an impulse train $d(x) = \sum_{n=-\infty}^{\infty} \delta(x-nS)$, where S is the step over distance. In Doppler mode, the scan is no longer discrete but continous which is the limit in which the step over distance goes to zero ($S \approx dx_s$). In the continous step over case, the summation becomes an integration, and the convolution with an impulse train becomes a convolution with a unit step function . The unit step function $u(x')$ is a function defined as $u(x') = 1$ for $\{x' > 0\}$, and $u(x') = 0$ for $x' < 0$. In order to be consistent with our choice in coordinate systems, the integral in equation 4.48 may be expressed compactly by a convolution $D(x') = I(x') * u(-x')1/v$, where the negative sign associated with the coordinate x' is necessary since the intensity envelope appears to move towards negative x' for positive x_s.

I will show that this integral may be written explicitly as a convolution in the following steps:

$$D(x') = I(x') * u(-x')1/v \tag{4.49}$$

$$= \int_{x_s=-\infty}^{0} I(x'-x_s)u(-x_s)\frac{dx_s}{v} \tag{4.50}$$

$$= \int_{x_s=0}^{\infty} I(x'+x_s)u(x_s)\frac{dx_s}{v}, \tag{4.51}$$

$$= \int_{x_s=0}^{\infty} I(x'+x_s)\frac{dx_s}{v}. \tag{4.52}$$

In going from equation 4.50 to 4.51 we changed the sign on the integration variable x_s and reversed the limits of integration. It is also noted that in equation 4.52 the unit step function was replaced by unity (by definition). This exercise served to illustrate that Doppler writing is a limiting case of a parallel scanning with a discrete step over distance that approaches zero, as we would expect. It also confirms that we have been consistent in our choice of representation.

We will now express the integral in equation 4.48 explicitly below

$$D(x') = \int \{|A_L(x' - x_s)\|^2 + \|A_R(x' - x_s)\|^2 + 2cos[\frac{2\pi}{P_f}(x' - x_s) - \Delta\Phi]\}G(x' - x_s)\frac{dx_s}{v}. \quad (4.53)$$

The first two terms in equation 4.53 may be expressed as a convolution $\{|A_L(x')|^2 + |A_R(x')|^2\} * u(-x')1/v$. We note that this convolution approaches a constant in steady state. This will contribute to a constant offset of our dose. What we will do in the following analysis is to observe the effect of $\Delta\Phi$ on the third term in equation 4.53. When $\Delta\Phi$ is equal to $\frac{2\pi}{P_f}x_s$, the x_s terms cancel out, and the *cos* function could be taken out of the integral. What remains in the integral is $G(x')$ convolved with $u(-x')/v$. If $A_L(x') = A_R(x')$, perfect contrast results. This can be seen by writing the result of equation 4.53 as $D(x') = F[1+\cos(\frac{2\pi}{P_f}x')]$ where $F = |A_L(x')|^2 * u(-x')/v$.

Now let us consider a more general case where there is an error in our phase modulation $\Delta\Phi$. We would like to determine the dose when $\Delta\Phi = \frac{2\pi}{P_f+\Delta P}x_s$. This corresponds to a linear phase modulation where we have used a period (i.e., $P_f + \Delta P$) different then the fringe period P_f. We may approach this by examining the integral

$$\int 2\cos[\frac{2\pi}{P_f}(x' + x_s) - \frac{2\pi}{\Delta P + P_f}x_s]G(x' + x_s)\frac{dx_s}{v} \quad (4.54)$$

$$= 2\int Re\{e^{\frac{j2\pi}{P_f}(x'+x_s)}e^{-j2\pi(\frac{x_s}{\Delta P+P_f})}\}G(x' + x_s)\frac{dx_s}{v} \quad (4.55)$$

$$= 2\int Re\{e^{\frac{j2\pi}{P_f}(x'+x_s)}e^{-j2\pi(\frac{x_s}{P_f})(1-\Delta P/P_f)}\}G(x' + x_s)\frac{dx_s}{v}, \quad (4.56)$$

where in equation 4.56 we have taken a Taylor series expansion of the denominator term in the phase of the complex exponential in equation 4.55. We also notice that we may cancel the terms involving x_s/P_f leading to further simplification as written below:

$$\int 2Re\{e^{j\frac{2\pi}{P_f}x'}e^{\frac{j2\pi\Delta P/P_f}{P_f}x_s}\}G(x' + x_s)\frac{dx_s}{v}. \quad (4.57)$$

For all practical purposes, we will impose the condition that $G(x_s)$ is real so that we may take the real part out of the entire integral. For mathematical convenience, we will impose the additional condition that $G(x_s)$ is even so that its Fourier transform is real. In addition, by the sifting theorem we recognize the Fourier transform $F\{G(x_s + x')\}$ to be equivalent to $F\{G(x_s)\}e^{+j2\pi fx'}$.

The reward in this analysis is that the integral in equation 4.57 may be expressed as the real part of the Fourier transform of $G(x_s + x')$ with respect to x_s evaluated at a spatial frequency $f = -\frac{\Delta P/P_f}{P_f}$. With these conditions the integral in equation 4.57 may be evaluated as

$$2Re\{e^{j\frac{2\pi}{P_f}x'}e^{(\frac{j2\pi fx'}{P_f})}|_{f=-\frac{\Delta P/P_f}{P_f}}F\{G(x_s)\}|_{f_x=-\frac{\Delta P/P_f}{P_f}}\}1/v \tag{4.58}$$

$$= 2cos(\frac{2\pi(1-\Delta P/P_f)}{P_f}x')F\{G(x_s)\}|_{f_x=-\frac{\Delta P/P_f}{P_f}}\}1/v \tag{4.59}$$

$$= 2cos(\frac{2\pi}{P_f+\Delta P}x')F\{G(x_s)\}|_{f_x=\frac{\Delta P/P_f}{P_f}}\}1/v, \tag{4.60}$$

where in going from equation 4.59 to 4.60 we have used the property that the Fourier transform of $G(x_s)$ is complex conjugate symmetric since $G(x_s)$ is real. In addition, in equation 4.60 we have used the property that the Fourier transform of $G(x_s)$ is real since $G(x_s)$ is even. The exciting result of equation 4.60 tells us that the contrast change is proportional to the spectrum of $G(x_s)$ evaluated at a spatial frequency $f = \frac{\Delta P/P_f}{P_f}$. Furthermore, it also illustrates that the period of the grating will be equal to the period $P_f + \Delta P$ determined by the phase modulation $\Delta\Phi$. This remarkable result proves that the period of the grating is not limited by the period of the interference fringes P_f but rather by the period of the modulation. On the other hand, the contrast will depend on the Fourier transform of the envelope evaluated at frequency $f = \frac{\Delta P/P_f}{P_f}$.

In terms of a linear system, the results make sense. If the input to a linear system is a sinusoid of spatial frequency $\frac{2\pi}{P+\Delta P}$, the systems transfer function must have some gain at this spatial frequency. Roughly speaking, the beam spot profile $G(x)$ determines the attenuation (or contrast degradation) that will occur. In the limit

that the spot size goes to zero any spatial frequency may be written this would be analogous to patterning with a fine tip pen versus a marker.

We could use the Fourier transform relationship for a Gaussian function to determine a closed analytical expression for the contrast due to a Gaussian beam. For a Gaussian envelope function $G(x')$ with a $1/e^2$ radius of R, the Fourier transform relationship is given by:

$$e^{\frac{-2x^2}{R^2}} \Leftrightarrow \sqrt{2\pi}R/2e^{\frac{-\omega^2R^2}{8}} \tag{4.61}$$

Using this property, we could then proceed to find the contrast degradation due to changing the period by ΔP. This could be written as

$$\gamma(\Delta P) = e^{-\frac{(2\pi\Delta P/P_f^2)^2R^2}{8}} \tag{4.62}$$

The $1/e^2$ contrast degradation occurs when $\frac{\Delta P}{P_f} = \frac{4P_f}{2\pi R}$, for a .7 mm beam radius R, 574 nm fringe period, the maximum period change before a $1/e^2$ lost of contrast is .3 nm. With these set of parameters, the contrast may be rewritten as

$$\gamma(\Delta P) = e^{-\frac{2\pi\Delta P^2}{\alpha^2}} \tag{4.63}$$

where $\alpha = .3$ nm and is defined such that when $\Delta P = \alpha$ the contrast $\gamma = 1/e^2$. A plot of the contrast for various values of ΔP assuming $\alpha = 4P_f^2/(2\pi R) = .3$ nm is plotted in figure 4-11.

Interestingly, the relationship we have found for the contrast provides for an insightful interpretation. Let us consider a discrete case in which we sum sinusoidal interference patterns of spatial frequency $1/P_f$ with different phases (given by a phase modulation $\Delta\Phi$). For pure sinusoidal signals, we would not expect the resulting pattern to produce an output of spatial frequency $\frac{1}{P_f+\Delta P}$ since the operation of summation is linear. However, since we have a Gaussian envelope with a broad bandwidth

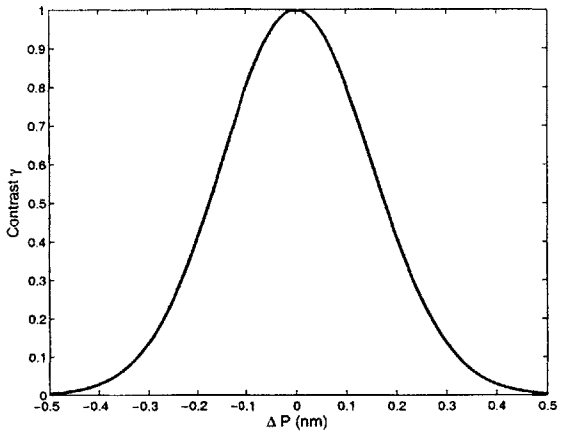

Figure 4-11: Contrast for a Gaussian function

modulated with sinusoidal carrier fringes of period $1/P_f$, the spatial frequency $\frac{1}{P_f+\Delta P}$ exists in the power spectrum. By properly applying a phase modulation $\Delta\Phi$ and by overlapping scans we can supress the other spatial frequencies in the patterned grating. The contrast suffers, because the peak spatial frequency in the Gaussian beam resides at $1/P_f$.

If the spot size were to shrink to a delta function, we could write any period we desire simply by modulating the beams with $\Delta\Phi$ because of the high spatial frequency content present. Unfortunately, this is not a viable option because the spot size is diffraction-limited and smaller spot sizes require more scans. It is also interesting to note that one may alter the shape of the envelope to allow for a greater selection of acceptable period changes. For example, one may modulate the Gaussian envelope by multiplying it with a sinusoid which will allow one to obtain a much greater range of ΔP. This has the physical interpretation of effectively take the beam spot with Gaussian envelope $G(x, y)$ and dividing it into two smaller beams. At the moment, these concepts are not being pursued but it does shed light into this remarkable conclusion.

One thing that this analysis shows is that we can continually chirp the grating by using Doppler writing. For high contrast, the period change should be much less than 1.3 nm for a beam radius of 0.7 mm and a fringe period of 574 nm. Currently in the

design phase, is a new generation of scanning beam interference lithography which will change the fringe period on the fly. This analysis will not incorporate such dynamic behavior. However, a good design rule may be found by considering the phase of a beam with a fringe period of $\frac{2\pi}{P_f+\Delta P}$. The change in phase $\delta\theta$ at the edge of a beam with diameter L may be found by a Taylor series expansion resulting in $\delta\theta \approx \frac{2\pi\Delta P}{P_f^2}L$. Setting this change in phase to be much less than π results in a design rule $\frac{\Delta P}{P_f} \ll \frac{P_f}{R}$ which is similar to the result we achieved through Fourier considerations.

4.4.1 Experimental Results

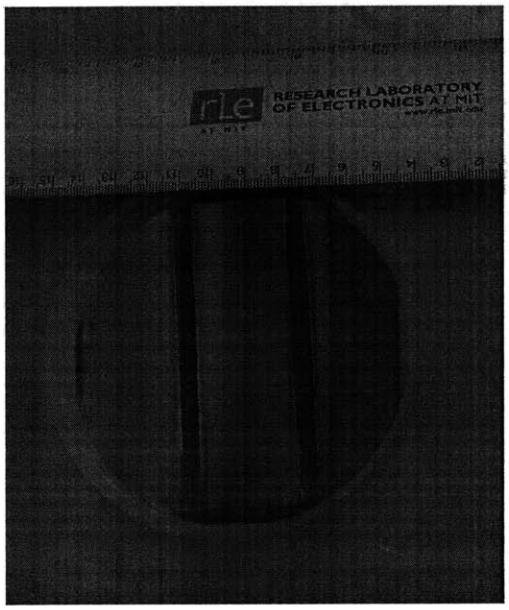

Figure 4-12: Moire experiment consisting of two exposures with periods $P_{f1} = 574.668$ nm and $P_{f2} = 574.678$ nm.

I performed an experiment which allows one to test the period modulation concepts discussed in this section for Doppler writing. Accordingly, two linear gratings

were written into a photo-resist substrate with two different fringe periods P_{f1} and P_{f2}. We expect the exposure dose on the photoresist to be of the form

$$D = D_{avg} + D_1 sin(\frac{2\pi}{P_{f1}}x) + D_2 sin(\frac{2\pi}{P_{f2}}x) \tag{4.64}$$

The average dose in units of (mJ/cm^2) for a single Doppler scan (the term D_{avg}) may be estimated by using the relationship

$$D_{avg} \approx 100\frac{P}{L}\frac{1}{v}, \tag{4.65}$$

where L is the beam diameter, P is the average power in the interfering beams, and v is the velocity of the beams. If we overlap these scans by stepping over in a direction parallel to the interference fringes, the average dose will be scaled by a factor L/S where S is the step over distance in the direction parallel to the interference fringes. This results in the same formula for the average dose for Doppler scanning as in parallel scanning. In these experiments, the average dose D_{avg} over the entire substrate resulting from two exposures was calculated to be 63 mJ/cm^2. The total power in the beams used was 10.8 mW.

The resulting dual exposed grating pattern is shown in figure 4-12. When the two sinusoids are in phase, the individual exposures constructively interferere resulting in a high exposure dose. When the two sinusoids are completely out of phase, poor contrast results as the two exposures destructively interfere with each other. The two sinusoids are in phase when $(\frac{2\pi}{P_{f1}} - \frac{2\pi}{P_{f2}})x = n2\pi$, where n is an integer. Solving for x, we find that the phase of each exposure constructively interferes at $x = n(\frac{P_{f1}P_{f2}}{P_{f1}-P_{f2}})$. The period $\Lambda = \frac{P_{f1}P_{f2}}{P_{f1}-P_{f2}}$ is known as the moireṕeriod, and for our choice of P_{f1},P_{f2} it is equal to 3.283 cm. As shown in figure 4-12 the bright fringes are separated by a distance consistent with our moireṕeriod. This confirms our understanding that the period of the grating is determined by the phase modulation $\Delta\phi$, and not necessarily by the fringe period of the interfering beams.

4.5 Wavefront Errors

So far in my analysis I have considered the intensity due to Gaussian wavefronts which are modeled as being real. This allowed us to consider the beam as a plane wave with a Gaussian wavefront as in $I(x) = A(x)[1 + cos(\frac{2\pi}{P_x}x)]$ where $A(x)$ was assumed to be real. In this section, we will make some observations of a case in which the Gaussian wave $A(x)$ may contain a non-planar wavefront. In general, if two Gaussian beams interfere at their "waist" location the phase of the wavefronts may be modeled as being ideal plane waves. At any other location, however, the Gaussian beams will have nonlinear phase. The intensity will then be of the form $I(x) = |A(x)| + Re\{A(x)e^{(\frac{2\pi}{P_x}x)}\}$, where $A(x)$ may be complex as in $A(x) = G(x)e^{j\theta(x)}$, where G(x) is a real Gaussian envelope given by $G(x) = Ae^{x^2/R^2}$ and $Re\{\}$ corresponds to taking the real part.

Now let us make some general observations. In the case that $\theta(x)$ is linear with respect to x, as in $\theta(x) = 2\pi\Delta Px$, we note that this results in the equivalent of a period error which will reduce the contrast. For example, multiplication by a linear phase complex exponential in the spatial domain corresponds to a frequency shift in the frequency domain. However, the Fourier transform of a Gaussian envelope with a Gaussian phase term $e^{j2\pi x^2/R^2}$ results in a function with a Gaussian magnitude and Gaussian phase in the frequency domain. It is well known that the Fourier transform of a Gaussian is another Gaussian. We have seen in the section on parallel contrast that stepping over by a discrete amount corresponds to multiplication in the frequency domain by an impulse train. Since this impulse train only selects one spatial frequency spectrum in the Fourier transform of the intensity $I(x)$ we reason that a nonlinear Gaussian phase error in the frequency domain will be of no consequence.

In summary, linear phase errors lead to contrast reduction and may be viewed equivalently as a period error. Second order nonlinear phase errors (i.e., of the form $e^{j\pi x^2/R^2}$ result in a Gaussian phase in the frequency domain. What will follow is an analysis which will provide us some insight into abberations, and general nonlinear phase errors.

4.5.1 Abberations and nonlinear Wavefront Errors

Let us again consider in intensity of the form $I(x) = G(x) + G(x)\cos(\frac{2\pi}{P_x}x + \theta_\epsilon(x))$ where $\theta_\epsilon(x)$ is a phase error due to the nonlinear phase in the wavefront. This may also be expressed as $I(x) = G(x) + Re\{G(x)e^{j\theta_\epsilon(x)}e^{\frac{j2\pi}{P_f}x}\}$, where $G(x) = Ae^{x^2/R^2}$ is a Gaussian envelope as before but now will will express the phase error as

$$\theta_\epsilon(x) = \beta\sin(\omega_m x), \tag{4.66}$$

where the coefficient β allows us to investigate the effect of a phase error of different magnitude. Since $e^{j\beta\sin(\omega_m x)}$ is periodic we may expand it into a Fourier series [44]. This results in

$$e^{j\beta\sin(\omega_m x)} = \sum_{n=-\infty}^{\infty} J_n(\beta)e^{jn\omega_m x}, \tag{4.67}$$

where $J_n(\beta)$ are the Fourier coefficients which are known as Bessel functions of the first kind. The solution for the coefficients is not in closed form, but they have been tabulated. For small values of β on the order of $\beta < 0.2$ the zero-order coefficient dominates and has a value $J_0 = 0.99$, which will predict minimal contrast loss. For larger values of β, i.e., for $\beta = 1$, the higher order Fourier series coefficients are significant, and the zero-order produces significant contrast loss (at least 25 % considering only the zero-order coefficient).

While this analysis has been carried out for a specific form of the phase error $\theta_\epsilon(x) = \beta\sin(\omega_m x)$ it provides some insights into the constraints on the amplitude and frequency of the phase errors in our wavefronts. For example, the above analysis suggests that wavefront distortions should be controlled such that $\theta_\epsilon < 0.15(2\pi)$ or 15% of a wave.

4.5.2 Phase Shifting Interferometry and Wavefront Detection

In the last section, we established that a Gaussian wave with a Gaussian phase produces no error in contrast nor in the phase of the printed grating. This is quite a

remarkable statement, which was originally suggested by means of a simulation by Chen [7] using a model derived by Konkola [27] which involved carrying out summations in the space domain through simulations. The statement is rather remarkable, and my frequency domain analysis confirms their results. Moreover, their model does offers some explanation as to this amazing result in the space domain. Namely, in a Gaussian amplitude function, the peak of the amplitude provides the most weight to the resulting printed phase. The phase error that is transfered onto the latent image may then be considered an amplitude weighted phase error. Since the peak of the amplitude function contains zero phase error, this phase dominates in the average. The frequency domain picture suggests that there are several spatial frequencies present. Stepping over by an integer number of grating periods selects only one spatial frequency while canceling out the others. Moreover, in practice the wavefronts are never ideally Gaussian nor plane-like. Indeed, abberations or wavefront distortions are typically present and may lead to contrast degradation as suggested by the previous section using a sinusoidal phase expansion.

In order to investigate the wavefront distortion, a technique known as phase shifting interferometry is used. There are several algorithms available for phase shifting interferometry, which offer different interpretations of the same result. Some of the most popular involve utilizing complex polynomials [2]. My particular favorite explanation for phase shifting interferometry utilizes a Fourier windowing point of view as originally proposed by De Groot[22].

The idea behind phase shifting interferometry is to solve for the phase $\theta_\epsilon(x,y)$ due to an interference pattern of the form

$$I(x,y,t) = G(x,y) + G(x,y)\cos(\frac{2\pi}{P_f}x + \theta_\epsilon(x,y) + \omega t). \tag{4.68}$$

The interference pattern is collected on a CCD camera, so one may operate on a particular pixel at at time. For example, the phase at a pixel (x_0, y_0) is given by

$$I_0(t) = I(x_0,y_0,t) = G_0 + G_0\cos(\frac{2\pi}{P_f}x_0 + \theta_0 + \omega t). \tag{4.69}$$

Now, it is particularly simple to solve for the nonlinear phase θ_0 by using the discrete Fourier transform. If we rewrite the intensity in its discrete form where $T_s = 2\pi/N\omega$ is the sampling rate, then the intensity may be rewritten as

$$I_0[n] = G_0 + G_0 \cos(\frac{2\pi}{P_f}x_0 + \theta_0 + \omega T_s n) \tag{4.70}$$

$$= G_0 + G_0 \cos(\frac{2\pi}{P_f}x_0 + \theta_0 + \frac{2\pi}{N}n). \tag{4.71}$$

In this form, the phase θ_0 may be recovered by computing the phase of the first order coefficient of the discrete Fourier transform of $I_0[n]$. The first order coefficient is found by calculating

$$\bar{I}_0(1) = \sum_{n=0}^{N-1} I[n]e^{j\frac{2\pi n}{N}} \tag{4.72}$$

and the phase may be written explicitly as

$$\theta_0 + \frac{2\pi}{P_f}x_0 = tan^{-1}\frac{\sum_{n=0}^{N-1} I[n]sin(\frac{2\pi n}{N})}{\sum_{n=0}^{N-1} I[n]cos(\frac{2\pi n}{N})}, \tag{4.73}$$

noting not interested in the linear phase term, if it is present. This algorithm is known as the Fourier transform algorithm and has a nice interpretation. There are several algorithms available in the literature. Indeed, at a Precision Engineering Conference Dr. Peter De Groot, a scientist at Zygo Corporation and the first to offer a windowed Fourier transform interpretation, stated that everyone should write their own algorithm at one point in their career. Accordingly, I have applied a 100 step Fourier transform algorithm. The experiment consists of interfering a zero-order reflected beam and a backdiffracted beam such that they interfere on a CCD camera. The two beams propagate parallel to one another, so the carrier fringe period $P_f = 0$. In addition, the various $I[n]$ are summed on the CCD camera by applying a phase shift $\frac{2\pi}{N}$ to one of the interfering arms. This phase shift is applied by using our acousto-optic modulator. The results for this experiment are shown in figure 4-13

In view of De Groot's analysis, several aglorithms may be generated by applying a window $w[n]$ to $I[n]$ before taking the discrete Fourier transform. In this instance,

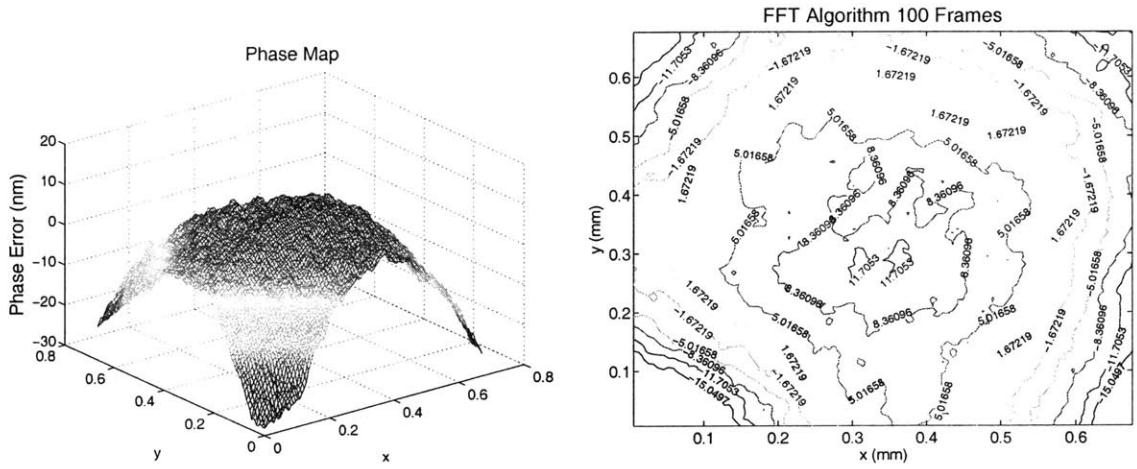

Figure 4-13: *100 Frame Fourier Transform Algorithm illustrating (a) 3-D mesh plot and (b) contour plot.*

the first-order Fourier component is given by

$$\bar{I}_0(1) = \sum_{n=0}^{N-1} I[n]w[n]e^{j\frac{2\pi n}{N}}. \tag{4.74}$$

Several choices of windows are given and are published in both signal processing and in the phase shifting interferometry literature. De Groot also points out that perhaps the most commonly known phase shifting interferometry algorithm, known as the Hariharan algorithm, may be viewed equivalently as windowing before taking the Discrete Fourier transform. The Hariharan algorithm is convenient because it uses only five frames.

The use of the Hariharan algorithm for the nanoruler was the subject of Chen's thesis work [7]. Accordingly, the phase may be obtained by setting the phase step $n\frac{2\pi}{N}$ with N=4, and performing the calculation

$$\theta = tan^{-1}\frac{2(I[3] - I[1])}{I[0] - 2I[2] + I[4]}. \tag{4.75}$$

The result of applying the Hariharran five step algorithm is shown in figure 4-14.

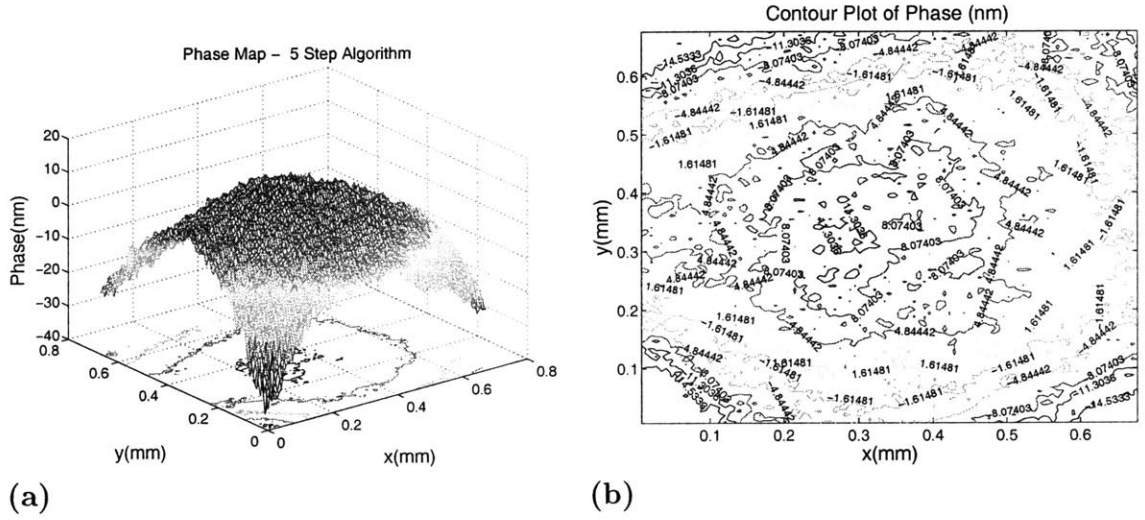

Figure 4-14: *Five frame Hariharan Algorithm illustrating phase error in nm (A)3-D Mesh Plot (B) Contour Plot*

Chapter 5

Grating and System Error Characterization

In a previous chapter, we identified two sources of error in the nanoruler. Namely, these errors consisted of Abbe error and stage mirror distortion. By modulating the phase of the interference fringes during patterning, one may in principal correct for these errors. In addition to the known errors, there may be other sources of error due to thermal expansion, compression, air turbulence, and the limited bandwidth of our controller. In this chapter we will develop a series of techniques to determine the total error in order to "self calibrate" our patterning tool.

The techniques developed in this chapter are novel in the sense of their application to characterizing the phase error in gratings. We will borrow from similar techniques found in the literature for the characterization of wavefronts. In inteferometry, relative phase differences are measured by comparing two wavefronts commonly referred to as the object wavefront (W_o) and reference wave W_r. The reference wavefront traverses non-ideal optical components and is therefore subject to distortion. In order to calibrate the reference wavefront, several algorithms which involve translations and rotations have been proposed. Indeed, interferometers known as rotational shear interferometers allow for rotation of the object wavefront with respect to the reference wavefront. These rotational tests allow one to determine the non-rotationally invariant components of the reference and object wavefronts.

Another calibration technique in interferometry involves translations. Lateral shear interferometers have been developed to compare the phase of a wavefront with a translated replica of the wavefront. For small displacements Δ, the phase difference between a translated wavefront with respect to its nominal position $W_o(x+\Delta, y) - W_o(x, y)$ may be viewed as an approximation of calculating the partial derivative of the wavefront in the shear direction. The inverse process integration, will then allow one to determine the original wavefront. This technique is similar to the one we used for calculating the mirror distortion. In this chapter we will apply the concept of lateral shearing as a characterization technique of our grating and system distortion.

In the nanoruler, reading mode allows us to compare the phase of the diffracted beam from our grating with a reference beam. The phase difference between the reflected and diffracted beams result in the phase of the grating plus a system distortion as will be shown. In this sense, we are not particularly interested in the phase of the reference wavefront. After all, the reference and diffracted wavefronts have approximately two millimeter diameters which are much smaller then the length scale of our patterned gratings. On the other hand, when we read the phase of our grating we are comparing its phase to a coordinate reference frame. In this case, the coordinate reference frame may be thought of conceptually as a reference wavefront. We may therefore perform translations and rotations of our grating with respect to this coordinate reference frame in order to determine the system distortion which is fixed to the coordinate reference frame (for example, Abbe errors and distorted reference mirrors).

We will begin this chapter by describing the nanoruler's reading mode in detail. We will then proceed to describe a classic test known as the three flat test. This is an example of an"absolute test" in which three flats may be compared to each other interferometrically to determine the absolute surface of the individual flats. This introduction will serve as a background for the methods we will propose for determining the nanoruler's systematic distortion. We will then proceed to develop two tests which may be used to determine the nanoruler's distortion. These tests will involve lateral translations and rotations.

At the end of the chapter, we will propose a novel reading mode configuration in the nanoruler which allows us to read the grating in a rotated position. A particular problem with rotating a grating is that the diffracted beams change angle with the grating orientation. Our dual pass reading mode method allows for ease of alignment which will be immediately implementable in the current nanoruler system. Measurements of a grating in its 0, 45, and 90 degree orientations will be presented for the first time.

5.1 Reading the Phase of the Grating: Traditional Reading Mode

The simplest model of reading mode contains all of the necessary elements for the experiments to be described. We will treat each beam as a plane wave and therefore disregard its gaussian nature and its finite size. Furthermore, any beam may be decomposed into a series of plane waves which will work well within the framework of our model.

In its basic configuration, two beams are used in reading mode. The two incident beams can be described as having an electric field with a given amplitude (A_L^i, A_R^i) and phase (θ_L^i, θ_R^i). The electric field of the incident left arm E_L^i and the incident right arm E_R^i will be described by

$$E_L^i = A_L^i exp(j\theta_L^i) = A_L^i exp(jk_L^i z_L + j\omega_L t) \tag{5.1}$$
$$E_R^i = A_R^i exp(j\theta_R^i) = A_R^i exp(jk_R^i z_R + j\Delta\Phi_{R-L} + j\omega_R t)$$

where z_L is a coordinate along the axis parallel to the propagation direction of the left beam. Similarly, z_R is a coordinate along the propagation direction of the right beam given by its propagation vector (k-vector). At the intersection of the two planes of constant phase described by $k_L^i z_L = k_R^i z_R$, the phase difference between the two arms are given by $\omega_R t - \omega_L t + \Delta\Phi_{R-L}$. Similarly, the reflected electric fields from the

grating interface may be written as

$$E_L^r = A_L^r exp(j\theta_L^r) = A_L^r exp(jk_L^r z_L + j\omega_L t) \tag{5.2}$$
$$E_R^r = A_R^r exp(j\theta_R^r) = A_R^r exp(jk_R^r z_R + j\omega_R t + j\Delta\Phi_{R-L})$$

We will concentrate on the right arm and refer to figure 5-1 although a similar argument applies to the left arm. Upon negative first-order diffraction a portion of the power along the direction of propagation of the right reflected arm (k_R^r) changes. If the diffraction condition is satisfied the k-vector of the diffracted right arm (k_R^d) becomes the sum of the the grating vector and the zero-order reflected k vector as in $k_R^d = k_R^r + k_g$. The grating vector $k_g = \frac{2\pi}{P_g}$ in this case is the negative fundamental spatial frequency of the grating where P_g is the period of the grating. What makes reading mode work is the Doppler shift that occurs when the grating is moved.

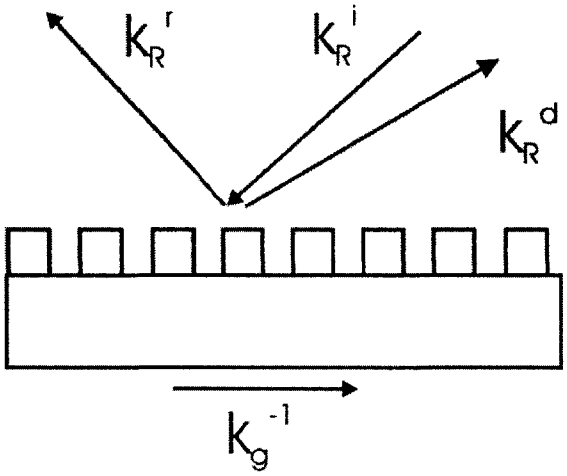

Figure 5-1: *Reading Mode showing the propagation direction of the incident right arm, zero-order reflected left arm, back-diffracted right arm, and a grating. The direction of the corresponding k-vectors are shown.*

We will digress a little to discuss how the phase of a diffracted arm changes due to a Doppler shift. Since the diffracted beam changes momentum from $\hbar k_R^d = \hbar(k_R^r - k_g)$, energy must also be conserved resulting in $\hbar\omega_d = \hbar(\omega_R + \omega_g)$ [43]. The directions of the k vectors are given in figure 5-1. Furthermore, the phase of the grating is given

by $\theta_{grating} = \frac{2\pi}{P_g}x - \omega_g t$. The temporal frequency of the grating phase is given by $\omega_g = k_g v$ where the velocity of the grating is v. The diffracted electric field therefore becomes

$$E_R^d = A_R^d exp(j\theta_R^d) = exp(jk_R^d z_{Rd} + j(\omega_R + \omega_g)t + j\Delta\Phi_{R-L}) \tag{5.3}$$
$$E_R^d = A_R^d exp(j\theta_R^d) = exp(jk_R^d z_{Rd} + j\omega_R t + jk_g vt + j\Delta\Phi_{R-L})$$
$$E_R^d = A_R^d exp(j\theta_R^d) = exp(jk_R^d z_{Rd} + j\omega_R t + jk_g x_o + j\Delta\Phi_{R-L}),$$

where the last relation in equation 5.3 comes from recognizing the displacement of the grating to be $x_0 = vt$. In the more general case $x_0 = \int vt$.

An alternative derivation would involve a Fourier optics approach. One may model the grating as a periodic amplitude mask. The field gets multiplied by this amplitude mask in real space and becomes convolved with the Fourier transform of the amplitude mask in the spatial frequency domain. The end result is that the field picks up the fundamental frequency component of the amplitude mask along with a phase shift given by the phase of the mask.

The important point to be taken away is that the diffracted beam includes a phase which is a function of the local period of the grating at position x_o.

The phase of the grating may be found by extracting $k_g x_0$ from the diffracted beam E_R^d. We will now show how the phase of the grating has been traditionally calculated [27].

In essence, the diffracted right arm and reflected left arm are recombined and their phase difference is recorded and we label this difference Φ_4. From equation 5.3 we note that we may express $\theta_R^d = \theta_R^r + k_g x_0$ by inspecting θ_R^r in equation 5.3. This requires making a number of justifiable assumptions. Namely, if we unfold the beam paths such they all travel along a parallel straight line we note that $k_R^d z_{Rd} = k_R^r z_r$. This assumes that the magnitudes of k_R^d and k_R^r are equal. In order for us to be abled to unfold the beams and provide a useful comparison, a further assumption must be made that the reflected and back diffracted beams travel in the same medium.

This corresponds to the phase difference between equation 5.3 and 5.2 as shown

below

$$\Phi_4 = \theta_R^r + k_g x_o - \theta_L^r \tag{5.4}$$

In our experiment, we also perform an independent measurement of the phase difference between the left arm and the right arm. This difference which we label Φ_3 is expressed

$$\Phi_3 = \theta_R^r - \theta_L^r, \tag{5.5}$$

which assumes that $\theta_R^r - \theta_L^r = \theta_R^i - \theta_L^i$. The difference $\Phi_4 - \Phi_3$ then provides us with the phase of the grating

$$\Phi_g(x_0) = \Phi_4 - \Phi_3 = \frac{2\pi}{P_g} x_0 \tag{5.6}$$

we seek. A few general comments may be made regarding this analysis. While the grating phase has been uniquely determined a number of assumptions have been made. Our independent measurement of the phase difference between the left and right arm at the location where Φ_3 is taken has been assumed to be a faithful representation of the difference in Φ_4 up to perhaps a constant. While efforts have been made to keep similar optical paths in place, vibrations and refractive index variations between optical paths may cause this assumption to be invalid. Furthermore, an assumption is made that the stage is scanned in the plane of the grating. Rocking of the stage with respect to the incident left and right beams will lead to additional non-idealities. It is then the subject of the next section to determine if there are any systematic errors present in reading mode.

5.2 Introduction to Self Calibration

In this section we will describe a test known as the three flat test. It is perhaps the most well known example of a self calibration test, and has been used in the context

of optical testing to calibrate reference flats. By combining and comparing three reference surfaces an absolute calibration of each reference surface may be obtained. We will begin by briefly reviewing the three flat test and then we will proceed to describe how to perform similar tests to determine the phase of our gratings.

Suppose three flats with surface profiles described by A(x,y), B(x,y), and C(x,y) are compared with each other interferometrically. Each of the flats may be compared with each other resulting in the following three measurements M_1, M_2, M_3.

$$\begin{aligned} A(x,y) + B(-x,y) &= M_1(x,y) \\ A(x,y) + C(-x,y) &= M_2(x,y) \\ B(x,y) + C(-x,y) &= M_3(x,y), \end{aligned} \tag{5.7}$$

where $M_1(x,y)$, $M_2(x,y)$, and $M_3(x,y)$ are the results of three interferometric measurements. The inversion of the X-axis occurs because the two flats face each other in the interferometer. It is necessary for example to rotate one of the reference flats about the Y axis. In order to solve for $A(x,y)$, $B(x,y)$, and $C(x,y)$ one requires more information. One may verify that a solution only exists in equation 5.7 for a line parallel to the Y-axis where $x = 0$. This result is due to the fact that the line described by $x = 0$ on the Y axis is invariant to rotations about the Y-axis and is therefore common to all three measurements M_1, M_2, M_3.

The missing information one seeks to solve in equation 5.7 relates to how the flat behaves when its axis is inverted. For example, the missing information includes how $A(x,y)$ relates to $A(-x,y)$ and so on for the other two reference flats. If one had additional information about the axis symmetry, one may in principle solve equation 5.7.

By means of a mathematical introduction we recall that any function $f(x)$ may be decomposed into its odd and even components. In a one dimensional case we can

identify these components as

$$f_{even}(x) = \frac{f(x) + f(-x)}{2} \tag{5.8}$$

$$f_{odd}(x) = \frac{f(x) - f(-x)}{2} \tag{5.9}$$

$$f(x) = f_{even}(x) + f_{odd}(x) \tag{5.10}$$

It is convenient to decompose a function of two dimensions into four components (even-even, even-odd,odd-odd, and odd-even) as suggested by [3]. For example, to calculate the even-odd component one calculates the even component with respect to the X-axis, $f_{even,x}(x,y) = \frac{f(x,y)+f(-x,y)}{2}$. The even-odd component, $f_{even,odd}(x,y)$, is then found by taking the odd part of $f_{even,x}(x,y)$ with respect to the Y-axis. The end result is $f_{even,odd}(x,y) = \frac{f_{even,x}(x,y)-f_{even,x}(x,-y)}{2}$. The complete four components which make up the decomposition of $f(x,y)$ are listed below

$$f_{even,even}(x,y) = \frac{f(x,y) + f(-x,y) + f(x,-y) + f(-x,-y)}{4} \tag{5.11}$$

$$f_{odd,odd}(x,y) = \frac{f(x,y) - f(-x,y) - f(x,-y) + f(-x,-y)}{4} \tag{5.12}$$

$$f_{even,odd}(x,y) = \frac{f(x,y) + f(-x,y) - f(x,-y) - f(-x,-y)}{4} \tag{5.13}$$

$$f_{odd,even}(x,y) = \frac{f(x,y) - f(-x,y) + f(x,-y) - f(-x,-y)}{4} \tag{5.14}$$

$$f(x,y) = f_{even,even}(x,y) + f_{odd,odd}(x,y) + f_{even,odd}(x,y) + f_{odd,even}(x,y). \tag{5.15}$$

Wyant et al. has used such a decomposition to arrive at an elegant derivation of a three flat test. In the following sections, we will develop algorithms based on rotations and rotational symmetry. In Wyant's analysis a polar Fourier transform is used to evaluate some of the rotational symmetry properties and in determining the various components. In this thesis, we will also use a Fourier decomposition, but will deviate largely from his approach and will only use one reference (namely our diffraction grating).

5.3 Reading Mode System Distortion

In reading mode, the nanoruler has the ability to "read" the phase of a patterned grating. This "reading mode" has a remarkable repeatability on the order of 8 nanometers peak to valley (three sigma) [27]. However, the accuracy of the nanoruler's reading mode has yet to be determined. It is the aim of this section to provide some insight into the accuracy limitation associated with reading mode and the need for subsequent calibration.

In general, the phase read from a grating on a real measuring system may be expressed as the sum of a system distortion decomposed into its linear $S(x,y)_L$ and nonlinear component $S(x,y)_{NL}(x,y)$ as in $S(x,y) = S(x,y)_L + S(x,y)_{NL}$, a nonlinear grating distortion $D_{gNL}(x,y)$, and an idealized grating phase consisting of a plane $D_{gL}(x,y)$. To be concise, we define the nonlinear component of $S(x,y)$ and $D(x,y)$ to consist of the terms which do not have a linear dependence on x or y. In one dimension, for example, a polynomial expansion may be used and the zero and first order polynomial terms of x will comprise the linear expansion, and the higher order terms will make up the nonlinear expansion. Strictly speaking, we will use linearity as it's well defined meaning in systems in which superposition applies. The reader is referred to Oppenheim's text on linear systems for a detailed discussion on linear systems.

The system distortion describes a distortion which is inherent in the system performing the reading. In the case of an ideal system S(x,y) would be zero. Examples of situations which would make S(x,y) nonzero in the nanoruler would include rotation of the stage during scanning, non-flat stage reference mirrors, Abbe error, and other sources of error. The interested reader is referred to P. Konkola's thesis which analyzes and treats the various sources of error[27].

We will refer to the phase of the grating which is read in reading mode as

$$\theta_{GratingRM}(x,y) = S(x,y)_L + S(x,y)_{NL} + D_{gL}(x,y) + D_{gNL}(x,y). \tag{5.16}$$

The grating distortion function $D_{gNL}(x,y)$ describes the nonlinear phase distor-

tion in the patterned substrate. For example, deformation of the substrate after exposure due to handling and phase errors introduced due to thermal expansion and compression would constitute examples of this kind of error. In addition, during the writing process any inaccuracies in the system will also be transferred onto the substrate. The phase $\theta_{GratingRM}$ is written to be as general as possible. For example, suppose we read a grating which was written on an ideal system. This ideal system would have zero distortion and therefore $D_{gNL}(x,y) = 0$. Further, assume that the grating was handled properly and was mounted on the nanoruler such that no additional distortion is introduced to the grating and $D_{gNL}(x,y)$ remains zero. During the reading process, the phase component $S(x,y)$ would be nonzero due to the inaccuracy of our system.

As another example, assume there is a coordinate system error as described in a previous chapter. For simplicity, lets allow the X-axis coordinate x to be distorted by $x_d = x + \epsilon_x(y)$ where $\epsilon_x(y)$ is on the order of a few hundred nanometers. The desired phase of our grating has a constant phase along the Y-axis, i.e., $D_g(x,y) = \frac{2\pi x}{P_x}$. However, since there is a coordinate system error the nanoruler patterns the grating such that in the distorted coordinates (x_d, y), the grating phase

$$D_{g-DistortedFrame}(x_d, y) = \frac{2\pi x_d}{P_x}, \tag{5.17}$$

is linear. However, in the undistorted coordinate frame (x,y) the same phase is

$$D_{g-undistorted}(x,y) = \frac{2\pi x}{P_x} + \frac{2\pi\epsilon(y)}{P_x}. \tag{5.18}$$

Now suppose we read this grating back in the same distorted coordinate frame in which the grating was written. The result is

$$\theta_{GratingRM-DistortedFrame}(x_d, y) = \frac{2\pi x_d}{P_x}. \tag{5.19}$$

We may alternatively express our measurement in the distorted frame as being equivalent to a measurement in the undistorted frame plus an error. Mathematically, this

may be expressed as $\theta_{GratingRM-DistortedFrame}(x_d, y) = D_{g-undistorted}(x_d, y) + S(x_d, y)$, where we see $S(x_d, y) = -\frac{2\pi\epsilon(y)}{P_x}$ by subtracting equation 5.18 from equation 5.19 where x is substituted for x_d.

In the above example, we derived a system error $S(x, y)$ for a coordinate system error $x_d = x + \epsilon_x(y)$. The latter example was made for a case in which the grating phase is assumed to be linear in the distorted coordinate frame (i.e., $D_g(x_d, y) = \frac{2\pi}{P_x}x_d$). In general, the grating distortion may be nonlinear in the distorted coordinate frame (due to for example thermal expansion errors). For a general measurement of $D_g(x_d, y)$, we are interested in expressing the phase read in the distorted coordinate system as being equalivalent to the phase read in an undistorted system plus a phase error. This phase error may be found mathematically,by substitution $x = x_d$ for a general grating distortion $D_g(x, y)$ in the undistorted coordinate system mesurement of equation 5.18. The error in the substitution is given by $x - x_d = -\epsilon_x(y)$. We may find the consequence of this error by examining $D_g(x_d - \epsilon_x, y)$. For a slow varying nonlinear component, $D_{gNL}(x, y) = D_{gNL}(x_d - \epsilon, y) \approx D_{gNL}(x_d, y)$, so we may use x and x_d interchangeably. However, for the linear component $D_{gL}(x_d - \epsilon_x, y) = \frac{2\pi}{P_x}(x_d - \epsilon_x(y))$, where we see the error in the substitution is $S(x, y) = -\frac{2\pi\epsilon_x(y)}{P_x}$. This discussion allows us to apply our abstract discussion of system errors to an important class of errors in the nanoruler, namely the coordinate system error described in a previous chapter.

For the remainder of the analysis, we will assume our measured coordinates to be ideal. We will proceed to make the substitution ($x = x_d, y = y_d$). We lump the errors of making this substitution into the system error $S(x, y)$. In addition, we will disregard the plane components $D_{gL}(x, y) = \frac{2\pi}{P_x}x + \frac{2\pi}{P_y}y$. The linear components of $D_g(x, y)$, consists of the best fit plane to the grating phase where P_x is the nominal grating period along the X-axis, and P_y is the nominal grating period along the Y-axis (i.e., in a least squares sense). In addition, we will disregard any linear phase error in our system $S(x, y)_L$. These linear phase errors may be caused by improper alignment of the stage interferometer beams which lead to cosine errors, and a wavelength drift of the interferometer laser.

The priority in this analysis is to determine the nonlinear part of the system distortion $S(x,y)$ which will allow us to calibrate reading mode to read gratings written on any system. Subtracting out the best fit plane from equation 5.16 results in

$$\theta_{GratingRMNL}(x,y) = S(x,y)_{NL} + D_{gNL}(x,y), \tag{5.20}$$

where the subscript $RMNL$ serves to remind us that we are interested in the nonlinear grating phase read in reading mode with a nonlinear system distortion $S(x,y)_{NL}$.

In the following section we will perform a simulation of a proposed algorithm to recover the system distortion function. We will make use of an example where we will consider a hypothetical system distortion function $wS(x,y) = 50sin(\pi x/a)$ nm which is scaled by $w = P_g/2\pi$ in order to express the phase in nanometers where $a = 180$ mm as shown in figure 5-2. Note, it is common to express the phase in units of a fraction of a wavelength, or grating period, or distance in order to emphasize its dependence on spatial distances.

We will also make reference to a grating distortion given by $wD_g(x,y) = \{75[(x/b)^2 + (y/b)^3] + 50cos(2\pi y/c)\}$ nm where $b = 100$ mm and $c = 90$ mm as shown in figure 5-3. This functional form is chosen for the purposes of illustration, and is somewhat arbitrary.

5.4 System Error Characterization by Translation

In the previous section we described the phase of the grating in reading mode to consist of two distinct components: the nonlinear system distortion $S_{NL}(x,y)$ and the grating phase $D_{gNL}(x,y)$. In order to seperate the two components and to determine each one individually, we must perform an operation on one of the components. A particularly simple operation involves translating the grating. We may perform the following two reading mode measurements

$$M_1 = S(x,y)_{NL} + D_{gNL}(x,y) \tag{5.21}$$

$$M_2 = S(x,y)_{NL} + D_{gNL}(x + \Delta_x, y). \tag{5.22}$$

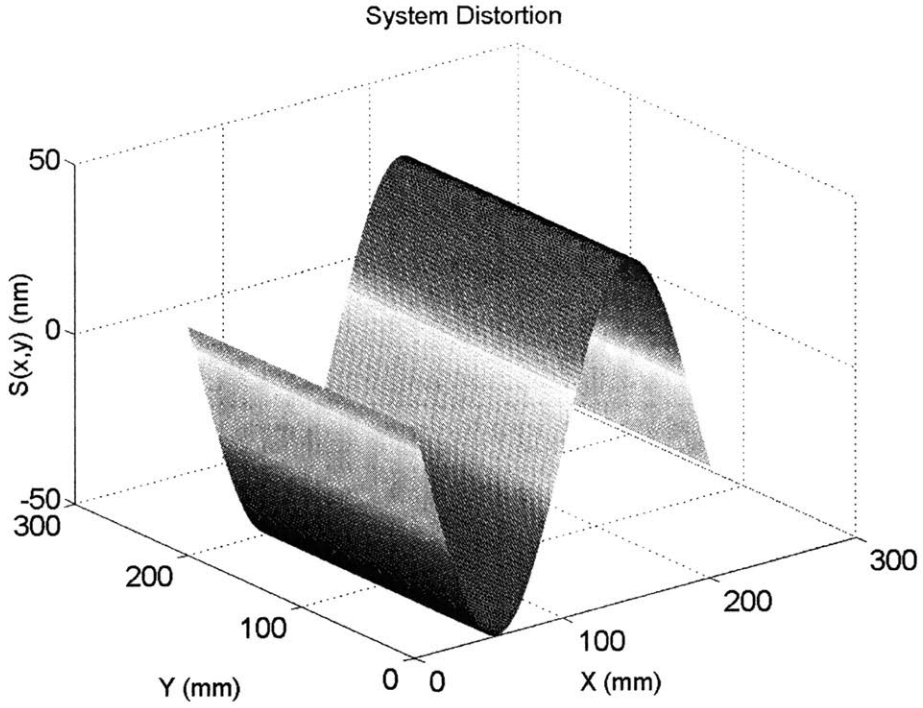

Figure 5-2: *Simulated System Distortion Function* $D_g(x, y)$

Measurement M_1 consists of a phase measurement with the grating in its nominal position. For illustration purposes, consider the output of the measurement given in figure 5-4. This measurement consists of the sum of the grating distortion $D_g(x, y)$ and the system distortion $S(x, y)$ given in figures 5-2 and 5-3. Our objective is to recover the grating distortion $D_g(x, y)$ from this measurement. We also perform a measurement M_2 consisting of translating the grating by an amount Δ_x with respect to its position in measurement M_1.

Subtracting the two quantities $M_1 - M_2$ results in

$$\bigtriangledown_x = D_{gNL}(x, y) - D_{gNL}(x + \Delta_x, y). \tag{5.23}$$

This equation is similar to the equations which described our mirror measurement process. In the special case that the grating varies along a single dimension x we could apply the same analysis which we applied to our mirror measurements. In that

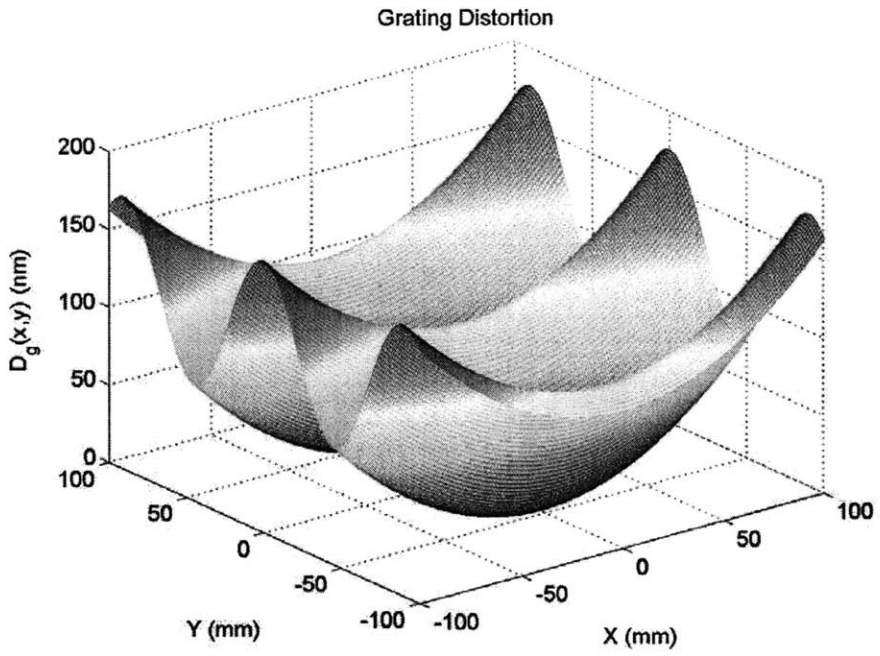

Figure 5-3: *Simulated Grating Distortion Function* $D_g(x, y)$

particular case, the output $\bigtriangledown_x = M_1 - M_2$ may be described as a convolution the grating's nonlinear phase $D_{gNL}(x, y)$ with an impulse response $h_{\Delta x}(x) = 1 - \delta(x - \Delta_x)$. The transfer function $H_{\Delta x}$ which corresponds to the Fourier transform of the impulse response contains zeros on the unit circle. From linear systems theory, we know that the inverse of this transfer function $H^i_{\Delta x} = 1/H_{\Delta x}$ is unstable. In the one dimensional case, we have shown a simple way to find an approximate transfer function $\hat{H}_{\Delta x} \approx H$ which has a stable inverse. Accordingly, we have shown in order to recover D_{gNL} from an equation of the form given in Eq. 5.23, we must simply apply the inverse transfer function $1/\hat{H}_{\Delta x}$ to the measured output $\bigtriangledown_x$. Unfortunately, in two dimensions the problem becomes more complicated as the issues of stability are not trivial and defining a closed form solution for the approximate transfer function $\hat{H}_{\Delta x}$ requires examination of poles on a 2-D surface.

One approach to apply $H^i_{\Delta x}$ in the frequency domain involves developing a thresholding algorithm which removes the singularities at the pole locations. This may be

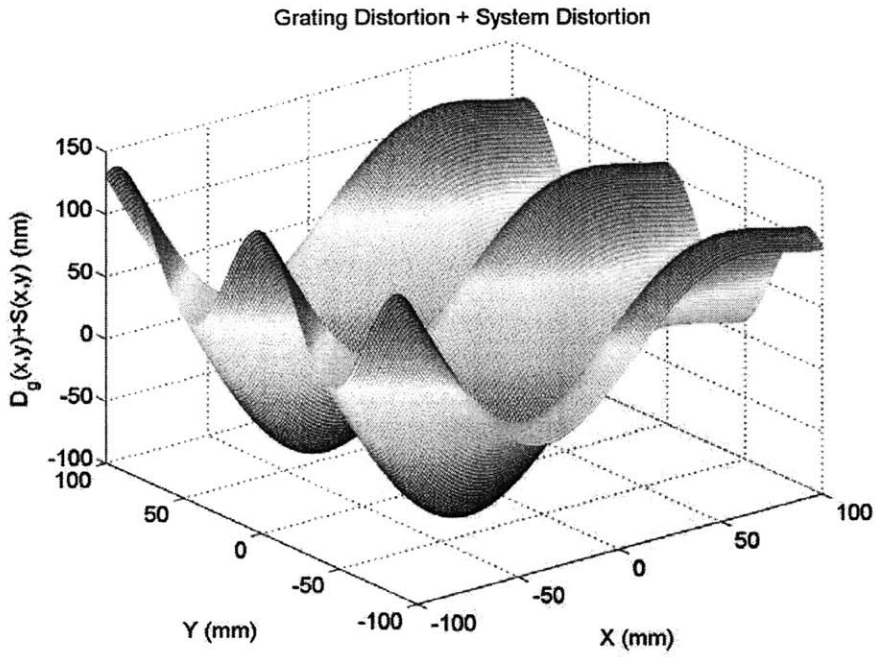

Figure 5-4: *Simulation:* $D_g(x, y) + S(x, y)$

done by setting the maximum of the inverse transfer function to a finite value. In this section, however, we will follow a different approach.

If the variables in the grating distortion are separable, for example $D_{gNL} = x^2+y^2$, then one may in effect treat the problem as a superposition of a one dimensional (1-D) problem in x, and a 1-D problem in y. However, if the variables are not separable and involve cross terms such as in $D_{gNL} = x^2 + 2xy + y^2$ we can no longer decompose the two dimensional problem into a superposition of two 1-D problems.

Interestingly, the two dimensional impulse response $h_{\Delta x}(x, y) = 1 - \delta(x + \Delta x)$ is used frequently in image processing and is known as an edge detector. It approximates the partial derivative and is therefore useful in determining the edges of an image. Instead of having the coefficients of the impulse response written as a vector, they correspond to a matrix in the two dimensional discrete case. For example, the matrix

corresponding to $h_{\Delta x}(x, y)$ may be written as

$$h_{\Delta x}(n_x, n_y) = \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}, \tag{5.24}$$

where n_x corresponds to the row index of the matrix and n_y corresponds to the column index. This formalism assumes discrete sampling of the (x,y) coordinates into a grid of pixels (n_x, n_y). While it is beyond the scope of this thesis to discuss two-dimensional systems theory in great detail, we will provide the following interpretation to allow for the jump from one dimension to two dimensions. In one dimension, the Discrete Fourier Transform (DFT) of an impulse response is defined as $H(e^{j\omega}) = \sum_{n=-\infty}^{\infty} h[n]e^{j\omega n}$, where the coefficients $h[n]$ correspond to a vector. In two-dimensions, the Discrete Fourier Transform will also be two dimensional with frequency components along two axis given by ω_x, ω_y. The two-dimensional DFT is then defined as $H_{2d}(e^{j\omega}) = \sum_{n_x=-\infty}^{\infty} \sum_{n_y=-\infty}^{\infty} h[n_x, n_y]e^{j\omega_x n_x} e^{j\omega_y n_y}$.

In addition, the translation Δx corresponds to a unit of $n_x = 1$ pixel in our discrete grid although the formalism may be modified to include larger translations. This matrix representation of the impulse response has a particularly useful interpretation in the discrete domain. It pictorally represents the weights of various pixels in the image. It shows that the output is computed by taking the difference between a pixel and its horizontally adjacent pixel. All of the information regarding any spatial frequency variations along the y axis (i.e., ω_y in the frequency domain) in the image are lost.

An impulse response corresponding to a translation along y may also be determined. For example, suppose we translate the grating along the y direction and measure

$$M_3 = S(x, y)_{NL} + D_{gNL}(x, y + \Delta y) \tag{5.25}$$

$$\bigtriangledown_y = D_{gNL}(x, y) - D_{gNL}(x, y - \Delta_y), \tag{5.26}$$

where $\bigtriangledown_y = M_1 - M_3$ corresponds to the phase difference between measurements

M_1 and M_3. The output may also be described by an impulse response $h_{\Delta y}(x,y) = 1 - \delta(x, y + \Delta_y)$ in matrix form as

$$h_{\Delta y}(n_x, n_y) = \begin{bmatrix} 1 & 0 \\ -1 & 0 \end{bmatrix}. \tag{5.27}$$

Any spatial frequency information along the x axis is lost (i.e., ω_x) with a one dimensional translation of this form. This may be viewed as an operator approximation of a partial derivative along y. The total derivative may then be approximated by a superposition of the two measurements $\nabla = \nabla_x + \nabla_y$ resulting in a superposition of the two impulse responses $h_{\Delta x \Delta y} = h_{\Delta x} + h_{\Delta y}$ producing the overall impulse response

$$h_{\Delta x, \Delta y}(x,y) = \begin{bmatrix} 2 & -1 \\ -1 & 0 \end{bmatrix}. \tag{5.28}$$

From linear systems theory, the output ∇ may be viewed as a 2-D convolution of $h_{\Delta x \Delta y}(x,y) * D_{gNL}(x,y)$. Ideally, the inverse transfer function may be found by computing the 2-D Fourier transform of the impulse response as in $H_{\Delta x \Delta y} = F\{h_{\Delta x \Delta y}(x,y)\}$, and then finding the inverse transfer function defined as $H^i_{\Delta x \Delta y} = 1/H_{\Delta x \Delta y}$. As we alluded to earlier in the discussion, the 2-D inverse transfer function is unstable, and moreover it is an infinite impulse response filter. While algorithms exist to approximate an unstable IIR filter the process is considerably complicated.

We will follow a different approach to design a reconstructive filter which will allow us to reconstruct $D_{gNL}(x,y)$. The algorithm to be used is similar to one which appears in a paper by Freischald and Koliopoulos in which a least squares modal method is used to compute a reconstruction algorithm. We will show that our reconstructive algorithm agrees with the one cited in the literature under certain assumptions. We will follow a significantly different approach using the formalism of discrete image processing. The least square criteria which is used in the literature minimizes the error between the measurement ∇ and the inverse fourier transform of a function whose coefficients are to be determined. We will use a least-squares method that

minimizes a different error which provides a more conventional metric encountered in signal processing.

The least square error reconstructive filter $\hat{H}^i$ will minimize the square error (ϵ^2) between the input $D_g(x,y)$ and the reconstructed input $\hat{D}_g(x,y) = \hat{H}^i H D_g(x,y)$. This error may be written explicitly as $\epsilon = D_g(x,y) - \hat{D}_g(x,y)$. In the presence of noise, we would like to construct a minimum mean square error reconstructive filter. This reconstructive filter has a particular advantage in that it minimizes the mean square error in the reconstructive process. For example, consider as an input a grating distortion with some additive noise in the measurement process resulting in $D_g(x,y)+n(x,y)$. The approximated inverse transfer function applied to the output, $\hat{H}^i H D_g(x,y) + \hat{H}^i n(x,y) \approx D_g(x,y)$ provides the best approximation to $D_g(x,y)$ in a minimum mean square error (MMSE) sense [20].

These class of filters are well known within stochastic signal processing and are referred to as Wiener filters (after an MIT Professor Norbert Wiener). The filter coefficients of $\hat{H}^i$ which satisfy this criteria may be expressed in the following form:

$$\hat{H}^i = \frac{H^*(\omega_x,\omega_y)}{|H(\omega_x,\omega_y)|^2 + \frac{R_{nn}(\omega_x,\omega_y)}{R_{xx}(\omega_x,\omega_y)}} \tag{5.29}$$

where (ω_1,ω_2) are the spatial frequencies along the x and y axis. The transfer function H^* denotes the complex conjugate of H, where $H = H_{\Delta x,\Delta y} = F\{h_{\Delta x \Delta y}\}$ is the overall transfer function of our measurement process corresponding to the Fourier transform of the impulse response $h_{\Delta x \Delta y}$. The remaining terms in the denominator of equation 5.29, $R_{nn}(\omega_x,\omega_y)$ and $R_{xx}(\omega_x,\omega_y)$, correspond to the power spectrum of the input in the absence of noise, and the noise respectively. The most recent authors to apply filtering concepts to sheared wavefronts (Freishclad et al.) [19], and later Poyneer et al [36, 14, 12, 37, 32] used discrete Fourier transforms and the concept of filtering in order to reconstruct a wavefront as a result of a lateral shear. However, the authors have not compared the results of their finding to a well known filter, known as a Wiener filter. The Wiener filter is used frequently in image processing applications [21, 28, 20].

If one sets the power spectrum of the noise $R_{nn}(\omega_x, \omega_y) = 0$, one obtains results similar to that of Freischlad et al. The assumption must also be made that the cross terms resulting from the magnitude operation $|H_{\Delta x}\Delta y| = |H_{\Delta x} + H_{\Delta y}|$ are negligible. Specifically, we assume that the cross terms involving $H_{\Delta x}H^*_{\Delta y}$ and its complex conjugate are not significant.

With these assumptions, the grating distortion is recovered by applying our Wiener filter in the frequency domain to obtain the relation

$$DF\{D_g(n_x, n_y)\} = \frac{(1 - e^{j2\pi/N})DF\{\bigtriangledown_x\} + (1 - e^{j2\pi/N})DF\{\bigtriangledown_y\}}{4(sin(\pi l)^2 + sin(\pi k)^2)} \tag{5.30}$$

where $DF\{\}$ denotes the 2D discrete Fourier transform operation to be defined below, and (n_x, n_y) are the discretely sampled coordinates (corresponding to pixels for example) of the continuous coordinates (x, y). The pixels are defined in a grid of size NxN. The coefficients k, l are the discrete Fourier transform coefficients given by the relation

$$\bar{D}_g(l,k) = DF\{D_g(x,y)\} = \sum_{n_x}\sum_{n_y} D_g(n_x, n_y)e^{-j(2\pi/N)kn_x}e^{-j(2\pi/N)ln_y}. \tag{5.31}$$

The distortion function $D_g(n_x, n_y)$ may then be recovered by applying the inverse 2-D Discrete Fourier transform to equation 5.30. The inverse Fourier transform relation is defined as

$$D_g(n_x, n_y) = \frac{1}{N^2}\sum_{k}\sum_{l} \bar{D}_g(l,k)e^{j(2\pi/N)kn_x}e^{j(2\pi/N)ln_y}. \tag{5.32}$$

This algorithm is tested by applying an inverse filter in the frequency domain to the discrete Fourier transform of the outputs $(\bigtriangledown_x)$ and $(\bigtriangledown_y)$ as expressed in equation 5.30. The grating distortion $D_g(n_x, n_y)$ is then recovered by inverse Fourier transforming the $\bar{D}_g(l,k)$ by using the relation for the inverse discrete Fourier transform given in equation 5.32. The reconstructed grating distortion from our example of $D_g(x,y) + S(x,y)$ is shown in figure 5-5. In our simulation, noise was not added.

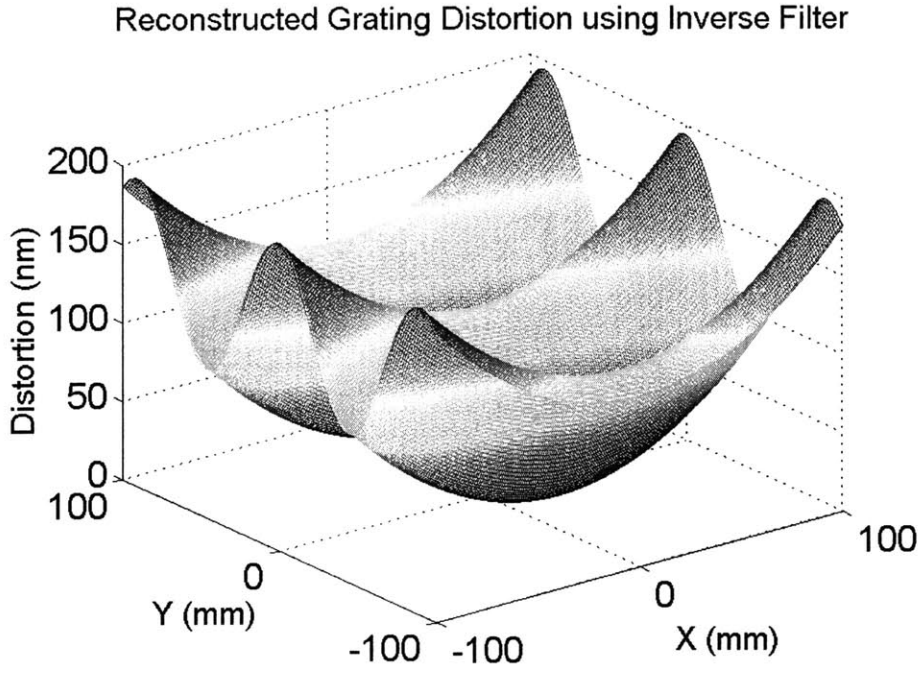

Figure 5-5: *Reconstructed grating distortion by applying inverse filter* $\hat{H}^i$

5.4.1 Discussion and Conclusions

In this section, we have presented a method for characterizing the grating and system distortion based on translating the grating. The assumption is made that the system distortion $S(x, y)$ is invariant to a translation of the grating by $(\Delta x, \Delta y)$. The measurement $\bigtriangledown_x$ described in equation 5.23 may be characterized by the transfer function $H_{\Delta x}(\omega_x, \omega_y)$ shown in figure 5-6 where we notice that any special frequencies along the Y axis $(\omega_x = 0, \omega_y)$ are lost. Similarly, the measurement corresponding to a translation along y denoted as $\bigtriangledown_y$ as given in equation 5.26 may be described by a transfer function $H_{\Delta y}$. The magnitude of this transfer function is the same as the magnitude of $H_{\Delta x}$ in figure 5-6 with the ω_x and ω_y interchanged (i.e., rotate the figure 90 degrees). If we assume the power spectrum of the noise is 0, we may minimize the mse of the output to an estimated filter $\hat{H}_{\Delta x}$ by finding the Wiener filter coefficients. The inverse of the estimated filter is denoted as $H^i_{\Delta x}$ and may be applied

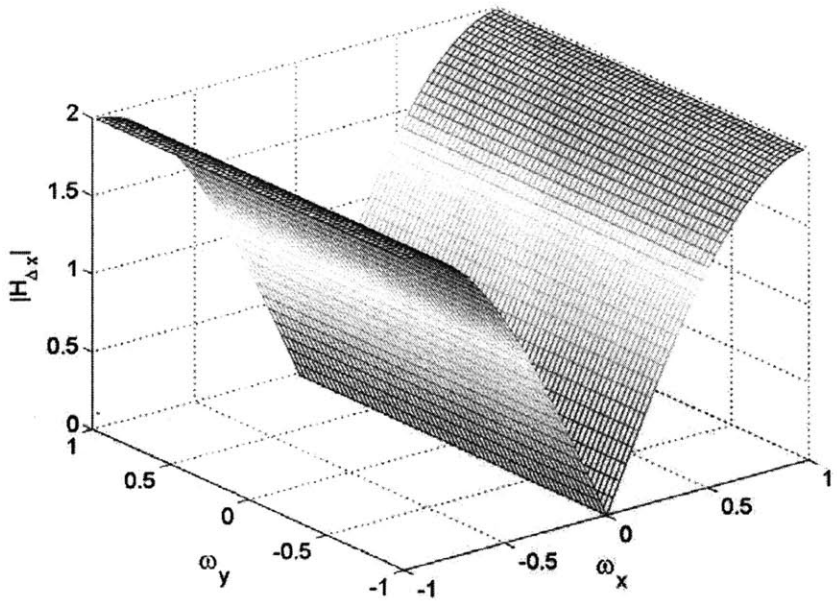

Figure 5-6: *Transfer function* $H_{\Delta x}$

to the output $\bigtriangledown_x$ in the frequency domain by

$$D_g(l,k) \approx \hat{H}^i_{\Delta x} \bigtriangledown_x = \frac{H^*_{\Delta x}}{|H_{\Delta_x}|^2} DF\{\bigtriangledown_x\} \tag{5.33}$$

Similarly, we may find the filter coefficients which minimize the mse between the output of a filter $H_{\Delta y}$ and the output of the estimated filter $\hat{H}_{\Delta y}$ in an mse sense. If we apply the inverse of the estimated filter $\hat{H}_{\Delta y}$ to the output $\bigtriangledown_y$ we obtain

$$D_g(l,k) \approx \hat{H}^i_{\Delta y} \bigtriangledown_y = \frac{H^*_{\Delta y}}{|H_{\Delta_y}|^2} \bigtriangledown_y \tag{5.34}$$

By multiplying the denominators of equations 5.33 and 5.34 to both sides of the equations and combining we obtain

$$D_g(l,k)[|H_{\Delta_x}|^2 + |H_{\Delta_x}|^2] \approx H^*_{\Delta x} \bigtriangledown_x + H^*_{\Delta y} \bigtriangledown_y . \tag{5.35}$$

From equation 5.35 we may solve for $\bar{D}_g(l,k) = DF\{D_g(n_x,n_y\}$ as

$$D_g(l,k) \approx \frac{H^*_{\Delta x} DF\{\bigtriangledown_x\} + H^*_{\Delta y} DF\{\bigtriangledown_y\}}{[|H_{\Delta_x}|^2 + |H_{\Delta_x}|^2]}, \tag{5.36}$$

which is equivalent to equation 5.30 and the results obtained in the literature ?? although obtained in a different manner. This particular derivation is insightful because it shows that the result is obtained from essentially applying a mse metric to two seperate (1-D) problems and combining the result. That this is indeed the case may be inferred by recognizing that $\frac{H^*_{\Delta x}}{|H^2_{\Delta x}|} = 1/H_{\Delta x}$. A better metric for evaluating a $2D$ distortion is obtained by a applying a 2D wiener filter to the output. In the case that the cross terms involving $H^*_{\Delta x} H_{\Delta y}$ are negligible, this is equivalent to combined 1D treatment.

While the filter of the form of equation 5.36 is suitable for all practical purposes, an improved filter may be obtained using the 2D inverse Wiener filter to a transfer function $H_{\Delta x \Delta y}$. This may be used by recognizing that in the presence of white noise, the signal to noise ratio $\frac{R_{nn}(\omega_x,\omega_y)}{R_{xx}(\omega_x,\omega_y}$ in equation 5.29 may be approximated by a constant C. The improved filter may then be applied in the frequency domain to the output $\bigtriangledown$ in the form

$$\hat{D}_g(l,k) \approx \frac{H^*_{\Delta x \Delta y} DF\{\bigtriangledown\}}{|H_{\Delta x \Delta y}|^2 + C} \tag{5.37}$$

where for our choice of the transfer function $H_{\Delta x \Delta y}$ this may be rewritten as

$$\hat{D}_g(l,k) \approx \frac{[1 - e^{j(2\pi/N)l} + 1 - e^{j(2\pi/N)k}] DF\{\bigtriangledown\}}{|1 - e^{j(2\pi/N)l} + 1 - e^{j(2\pi/N)k}|^2 + C} \tag{5.38}$$

In conclusion, we have developed an alorithm to reconstruct a grating from a difference map in the nanoruler. This algorithm is similar to Freischad's algorithm in the literature for lateral shearing of wavefronts but has some key differences. The procedure used to develop the reconstructive filter followed from applying concepts known in 2D image processing, namely Wiener filtering. This allowed us to develop a more general equation 5.37 which may be applied to the measurement process in the presence of noise by tuning a single parameter C. A simulation of the reconstructed

grating distortion obtained by applying equation 5.37 to our simulated distortion functions $S(x,y) + D_g(x,y)$ resulted in a reconstructed output $\hat{D}_g(x,y) \approx D_g(x,y)$ similar to figure 5-6 without any notable differences. The source for the matlab code is provided in the appendix.

5.5 System Error Characterization by Rotation

Analogous to the translational test, we may also develop a rotational test which will allow us to analyze the grating distortion. It should be pointed out that rotational tests are well suited for characterizing rotational symmetries. When testing optical wavefronts, rotational tests are largely employed due to their convenience of describing rotations of rotationally symmetric wavefronts by rotationally symmetric functions, such as Zernike polynomials.

One particular test which has been employeed in the nanoruler involves rotation of a grating by 180 degrees. We may apply Wyant's analysis to a rotation [3]. We find that a rotation of the grating distortion $D_g(x,y)$ by an angle θ may be expressed as $\{D_g(x,y)\}^\theta = D_g(x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta)$. It follows that a rotation of a 180 degrees inverts both the X axis and Y axis, leading to $\{D_g(x,y)\}^{180} = D_g(-x,-y)$. We may also decompose a grating distortion into its $\{even,even\}$, $\{odd,odd\}$, $\{even,odd\}$, $\{odd,even\}$ components as given by equations 5.11 - 5.15. We will use shorthand notation and define these components as $D_g = D_{gee} + D_{goo} + D_{geo} + D_{goo}$ where the (x,y) dependence is implied. Furthermore, we note that the effect of a rotation by 180 degrees results in $\{D_g\}^{180} = D_{gee} + D_{goo} - D_{goe} - D_{geo}$. Let us once again consider a reading mode measurement of the form $\theta_{gratingRMNL} = D_g(x,y) + S(x,y)$. Suppose we take the following measurements: M_1 with the grating in its nominal position, and M_2 where the grating has been rotated by 180 degrees. These measurements may be expressed as

$$M_1 = D_{gee} + D_{goo} + D_{geo} + D_{goe} + S \tag{5.39}$$

$$M_2 = D_{gee} + D_{goo} - D_{geo} - D_{goe} + S. \tag{5.40}$$

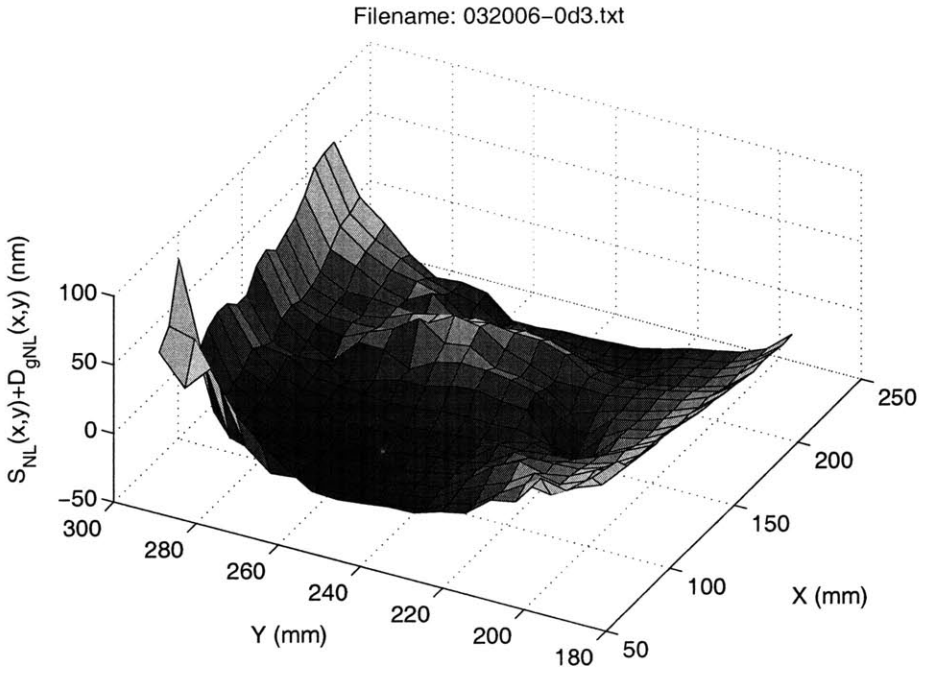

Figure 5-7: $S(x, y) + D_g(x, y)$

We note that in measurement M_2 the system distortion does not change since only the grating has been rotated. The data corresponding to measurement M_1 is shown in figure 5-7. The substrate for this measurement consists of a 220x165 mm BK7 substrate. This BK7 substrates is rather thick (on the order of 25 mm) and optically flat making it suitable for metrology purposes. An anti-reflection coating and photo-resist were spun onto the substrate. Subsequently, the substrate (substrate ID: 103039A-005-3AL) was exposed and patterned on the nanoruler resulting in a period of $P_g = 574.3$ nm. After developing the exposed resist, the substrate was coated with aluminum in order for it to be read in the nanoruler's reading mode.

The measured peaks and valleys of the data are on the order of ± 60 nm. Since we do not know the system distortion $S(x, y)$, we can not obtain any information regarding the distortion of the grating from the data shown in figure 5-7.

A second measurement corresponding to equation 5.40 in which the substrate is rotated by a 180 degrees is shown in figure 5-8. The peaks and valleys of this

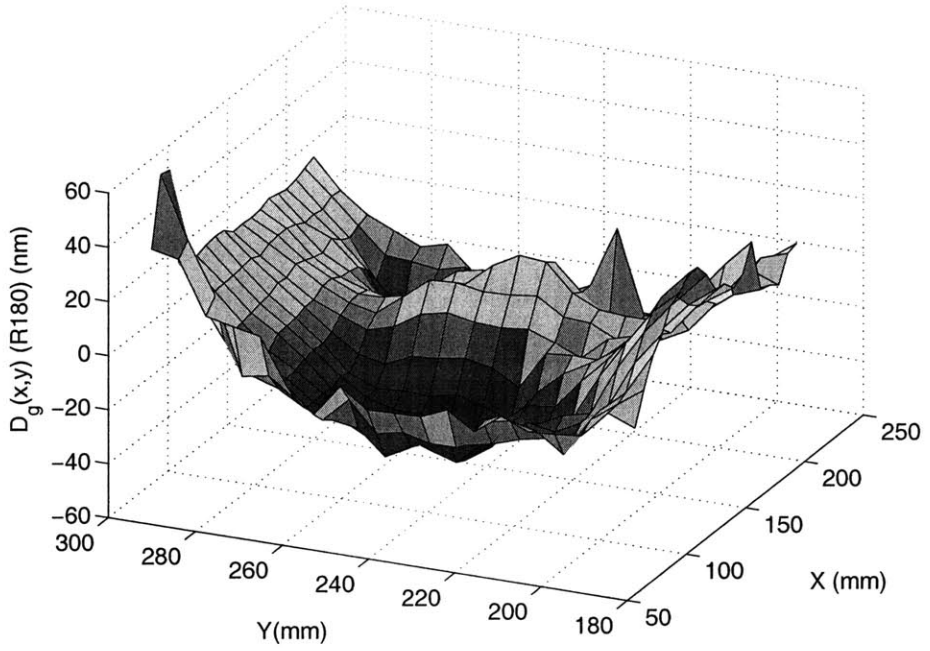

Figure 5-8: $S(x,y)+\{D_g(x,y)\}^{180}$. *The grating is rotated by 180 degrees and measured in reading mode.*

measurement are on the same order of magnitude as in the 0 degree orientation. From equation 5.40 we expect the even-odd component and odd-even component of the grating distortion to invert in this measurement. If we take the difference between measurements M_1 and M_2 we can remove all of the components of the grating distortion (and the system distortion) which are invariant to a grating rotation of a 180 degrees. This may be expressed as

$$[M_1 - M_2]/2 = D_{geo} + D_{goe}, \tag{5.41}$$

where we have subtracted equation 5.40 from eq. 5.39 and divided the result by two. The grating's even-odd and odd-even distortion as a result of applying 5.41 are shown in figure 5-9. This test is useful in the sense that it allows us to begin to make some statements regarding the distortion of the grating. The peaks and valleys of the sum of the grating's even-odd and odd-even distortion are also on the order of $\approx \pm 60$ nm.

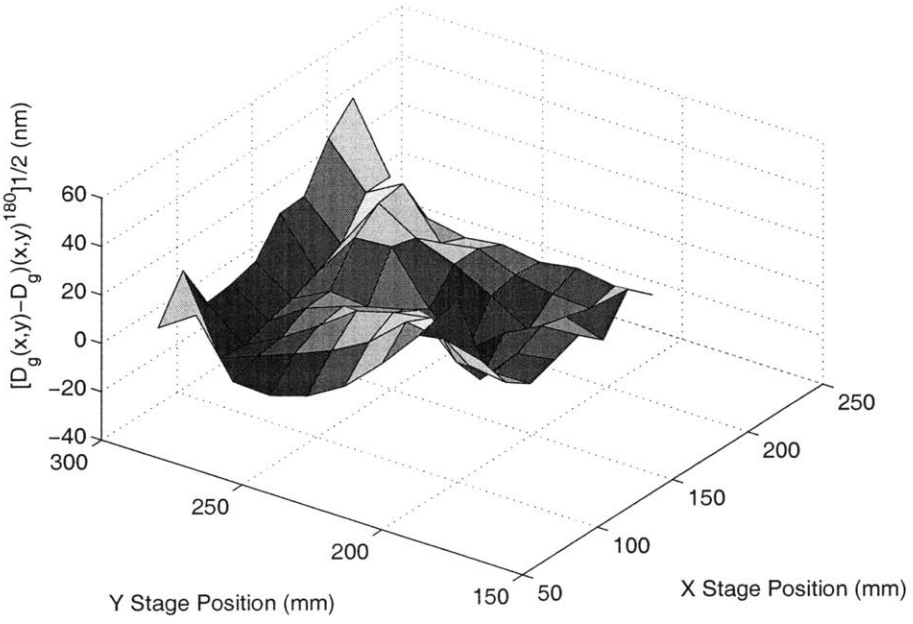

Figure 5-9: $[D_g(x,y) - \{D_g(x,y)\}^{180}]/2$

This allows us to conclude that the grating contains these errors which are of the order of magnitude we would expect for from considering the sources of coordinate error in our system (i.e., mirror nonflatness).

Furthermore, it has been suggested that vacuum forces of the stage which clamp the substrate during patterning have resulted in further grating distortion. A comparison of a grating written with the vacuum off with one in which the vacuum was on during patterning showed less distortion when measured using an external source consisting of a large aperture interferometer. However, reading mode experiments which compare the effect of the the vacuum clamping do not corroborate with these results.

Further experiments with a highly controlled parameter space must be performed in order to deduce the source of these errors. One must be careful to insure the substrate has reached thermal equilibrium before writing and reading the substrate. In addition, handling and mounting procedures when comparing data to external

(assumed to be calibrated) interferometers must also be considered. It is assuring that our reading mode experiments are repeatable to within $\approx \pm 10$ nm on these particular substrates. This is important since it allows one to correct for the system distortion by controlling the phase of the writing beams during patterning. As we have seen, rotating the grating by 180 degrees will not allow us to determine the grating distortion, nor the system distortion entirely. It is the objective of the remainder of this section to discuss rotation tests, and the information content which may be recoverd in a more general manner.

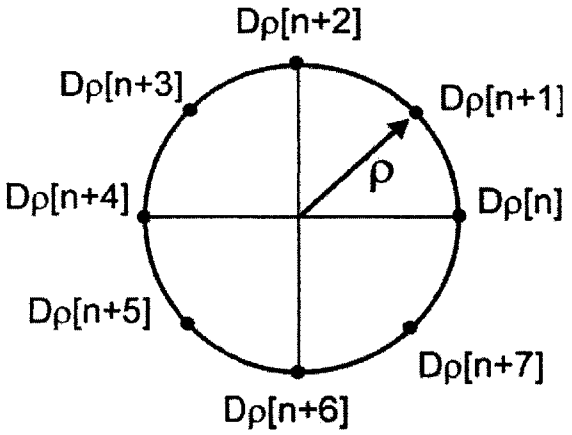

Figure 5-10: *A pixel $D_\rho[n]$ oriented at radius ρ from the center of rotation is shown. Also shown are 8 equidistant sampling points $D_\rho[n]$, for n=0,1,...,7*

Our goal in this section is to develop a test based on rotations which will allow us to determine an estimate of $D_g(x,y)$ denoted as $\hat{D}_g(x,y)$. In anticipation of rotation, we will use cylindrical coordinates to describe the grating distortion. The idea is to express coordinates (x,y,z) into a cylindrical coordinate system described by radius ρ and angle θ with height z as in (ρ,θ,z).

The convenience of using this coordinate system is that we recognize that we can express any function $f(x,y)$ in cylindrical coordinates as $f(\rho,\theta)$. This conversion results in a periodic function of θ with a fundamental period of $\theta_p = 2\pi$. This makes $f(\rho,\theta)$ particularly suitable for a fourier series expansion as $f(\rho,\theta) = \sum_m f_m(\rho) exp(j2\pi m\theta/\theta_p)$ where $f_m(\rho)$ are the fourier expansion coefficients of $f(\rho,\theta)$

with respect to θ along a radius ρ. If we restrict are discussion to a fixed radius, $\rho = \rho_0$, we observe that a function of $f(x,y)$ may be expressed as a 1-D function of θ as in $f(\theta) = \sum_m f_m exp(jm\theta)$.

Recall that during our measurement we obtain a grid of pixels consisting of dimensions NxN. For the sake of discussion, let us consider a single pixel (x_0, y_0) in our measurement. This particular pixel is located at radius $\rho = x_0^2 + y_0^2$ from the center of rotation, and $\theta = 0$. We will denote our measurement on this pixel as $\theta_{gratingRMNL} = D_\rho[n] + S_\rho[n]$. In addition, if we sample the grating distortion at M=7 discrete angles along this radius ρ given by $\theta = n\frac{2\pi}{M}$ for $n = 1, 2, ...7$ we will will obtain $D_\rho[n+1], D_\rho[n+2], ..., D_\rho[n+7]$ as shown in figure 5-10.

Our choice in notation will now become clear. Suppose we take eight measurements where the grating has been rotated by $\theta = n\frac{2\pi}{M}$, where $n = 0, 1, 2, ...7$. Let's concentrate on pixel $D_\rho[n]$. If we sum all eight measurements and divide by L=8, we will calculate an ouput on this pixel as

$$\theta_\rho[n] = \frac{1}{L}\sum_{N=0}^{N=7} D_\rho[n+N] + S[n]. \tag{5.42}$$

This measurement may also be expressed as a convolution of the input $D_g[n]$ with an impulse response $h[n]$ as in

$$\theta_\rho[n] = D_\rho[n] * h[n] + S[n] \tag{5.43}$$

$$h[n] = \frac{1}{L}\sum_{N=0}^{N=7} \delta[n+N]. \tag{5.44}$$

One may recognize that the transfer function $h[n]$ corresponds to a moving average filter. The frequency response of such a filter may be found by taking the fourier transform of 5.44 resulting in

$$H(e^{j\omega}) = \frac{1}{L}\frac{sin(\omega L/2)}{sin(\omega/2)} e^{-j\omega L/2}, \tag{5.45}$$

where ω is a normalized frequency used in the discrete fourier transform. The fre-

quency is normalized such that $\omega = 2\pi = \frac{1}{\theta_s}$, where θ_s is our sampling angle defined to be $\frac{2\pi}{L}$ in this example (i.e., 45 degrees). The transfer function given by equation 5.45 is a low pass filter which attenuates high spatial frequencies. The first zero corresponds to an angular period of $\theta = 2\pi$. Indeed, only low angular frequencies are passed. Since our function is periodic, only the zero order (DC) component of $D_g(n)$ is not attenuated in equation 5.43. This result was first present by Evans et al [16, 34] albeit by using a different method of analysis which does not emphasize the filtering approach.

I have simulated an output θ_ρ for an input of a system distortion $S(x, y)$ and grating distortion $D_g(x, y)$ in figures 5-2 and 5-3. The output in equation ?? is computed using matlab where the grating distortion is rotated by $\theta = n\theta_s$ where $n = 0, 1, ...7$ and θ_s is the sampling angle given by $\theta_s = \frac{2\pi}{8}$. Equation 5.42 applies to each pixel $D_p[n]$, in general an image is taken on a square array of pixels of size NxN. The resulting θ_ρ is shown in figure 5-11. We note that the output contains the system distortion $S(x, y)$ and the DC (zero order) angular component of the grating distortion $D_g(x, y)$. Note that the zero order component corresponds to a rotationally symmetric distortion, which shows up as a nonlinear distortion in a cartesian coordinate system containing circular symmetry.

The linear systems formalism is insightful since it illustrates the information which may be recovered after apply a filter $h[n]$ as a result of rotation. In the summation case we used a moving average filter in which case only the angular-average or DC component of the grating distortion remains. In a cartesian coordinate system, the DC component resulting from a fourier series expansion with respect to x is usually not of interest and is neglected (since we are interested only in nonlinearities). However, we note that the the DC component in a fourier series expansion with respect to θ results in a circularly symmetric distortion which by definition is nonlinear in x and y.

It is also clear that a different transfer function with an impulse response $h[n]$ may be obtained by taking a difference between measurements. For example, suppose we rotate the grating by $\theta_s = \frac{\pi}{4}$ and obtain a measurement $\theta_\rho = D_\rho[n + 1] + S[n]$. If

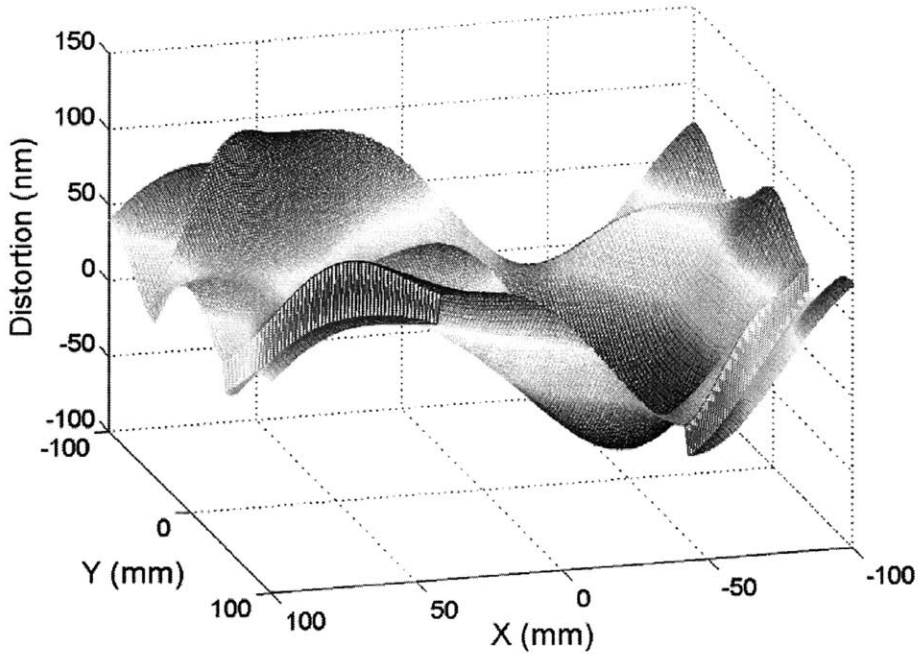

Figure 5-11: *Output θ_ρ corresponding to the average sum of eight measurements. In each measurement, the grating is rotated by an angle $\theta = n\frac{2\pi}{8}$, where $n = 0, 1, 2, ...7$ resulting in an output given by $\theta_\rho[n] = \frac{1}{8}\sum_{N=0}^{N=7} D_\rho[n+N] + S[n]$ for each pixel. The summation corresponds to computing the rotationally invariant (DC component) of $D_g(x, y)$.*

we subtract this from the non-rotated measurement $\theta_\rho = D_\rho[n] + S[n]$ and label the result $w[n]$ we obtain

$$w[n] = D_\rho[n+1] - D_\rho[n] \tag{5.46}$$

$$w[n] = D_\rho[n] * h[n] \tag{5.47}$$

where we note the impulse response describing this measurement process may be given by $h[n] = \delta[n+1] - \delta[n]$. We realize that this impulse response has a zero at the zero frequency location (for example the average of $h[n]$, $E\{h[n]\} = 0$). Therefore applying a difference equation of this form to our measurements will not allow us to recover the DC-angular frequency component.

5.5.1 Conclusion

Regardless of the algorithm we choose, it will not be possible to isolate the grating's zero order angular frequency component based solely on rotation. This makes intuitive sense since the zero-order angular frequency component is invariant to rotations by definition. Freishclad has derived a three flat test using a similar filter transfer function approach for reference flats [18]. Accordingly, if we could filter the zero order component which we will denote as $\bar{D}[0]$ of the grating distortion $D_g(x,y)$ by applying the filter $h[h]$ in equation 5.42 we would obtain $\bar{D}[0]+S[n]$. If we perform the same measurement on a second grating $E_g(x,y)$ we may obtain $\bar{E}[0]+S[n]$. The last equation one would require to solve for $S[n]$ would require a direct relation between the two grating distortions $\bar{D}[0]+\bar{E}[0]$ which we may not obtain in the nanoruler system. In other words, a three flat test requires the flexibility to compare any combination of flats to each other. This additional flexibility allows one to obtain information regarding the rotationally invariant component. In the nanoruler, we may also use three gratings but we may not compare two gratings directly to each other.

While we may carry out a similar analysis as we have done in the cartesian coordinate system for a transfer function of the form given in equation 5.47, we acknowledge that at best we will lose information about the rotationally invariant term. This term contains information about the grating's nonlinearity in a cartesian coordinate system, and it therefore makes this test ill-suited for recovering the grating's nonlinear distortion. Moreover, an inverse reconstructive filter is difficult to implement in matlab software. For a frequency domain approach, it will require taking a fourier transform along a polar grid which requires writing customized routines in matlab. Nonetheless, in some applications it may be desired to make the grating phase nonlinear and circularly symmetric. The approaches in this section would then apply. For the purposes of this thesis, the translational test is the preferred method for characterizing a grating's nonlinearity.

In the next section, I will present a technique which is developed to read the grating in a rotated position. This is useful for characterizing the system coordinate error

distortion. For example, if one wishes to characterize the y coordinate distortion, one may rotate the grating by 90 degrees and apply the translational test of the previous section. If one wishes to apply the rotational test of this section, a dual pass reading mode approach will allow for reading the grating in a rotated position. In addition, it will be shown that the substrate's surface profile may also be recovered using this dual-pass alignment scheme.

5.6 Dual Pass Reading Mode

In traditional reading mode, the angle of the incident right arm is set so that the ray of the back-diffracted negative first-order beam is parallel with the ray of the incident beam. Strictly speaking the directions are anti-parallel, or a 180 degrees apart. This configuration is commonly referred to as the "Littrow" configuration, and is shown in figure 5-12. The traditional reading mode only allows for two grating orientations

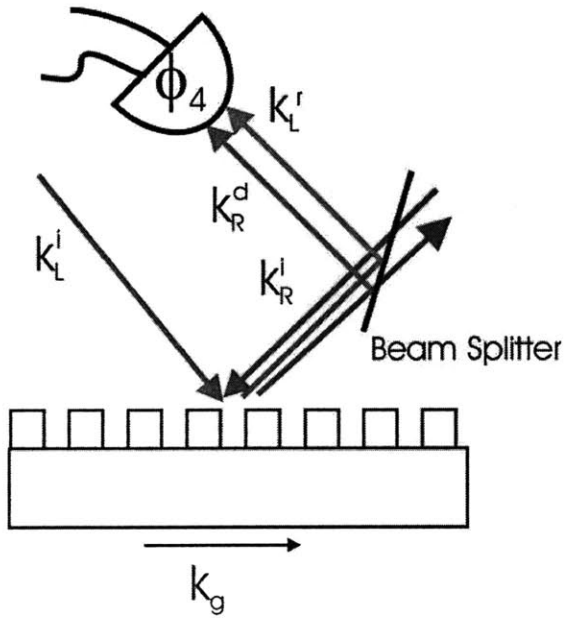

Figure 5-12: *Traditional reading mode. The incident right arm k_R^i is back-diffracted into k_R^d. The angle of incidence of the left and right arms are such that the back-diffracted right arm is parallel to the left reflected arm k_L^r. The reflected left arm and back-diffracted right arm are combined on phase detector ϕ_4.*

(0 degrees and 180 degrees). Any other grating orientation would violate the Littrow condition and cause the back-diffracted beam to change its angle. In this section, we will describe a novel reading mode technique which allows one to read the grating in any orientation. Furthermore, it will allow one to measure the surface profile of a grating which has not been possible with our existing methods. This new reading mode will be refered to as dual pass reading mode.

The idea for dual pass reading mode involves canceling out the effect of diffraction on the propagation direction of the diffracted beam. The idea of using a double pass scheme to cancel out the change in angle of the beam due to diffraction is not new. It has been found in the literature for acousto-optic modulator applications [13]. The basic principle is that upon retro-reflection a beam reverses all of its propagation vector (k) components.

Consider the right diffracted arm ($k_R^d = k_g^{-1} + k_R^r$) vector in figure 5-13 which includes the sum of the zero order (reflected) k_R^r vector and the negative first-order diffraction grating vector k_g^{-1}. We now reflect this beam onto itself such that its direction is reversed and we define its new direction as $k_{retroreflect} = -k_g^{-1} - k_R^r$. As it diffracts for a second time from the same diffraction grating the grating vector cancels out. In addition, the component of $-k_R^r$ normal to the surface changes sign. For book keeping purposes we may write $-k_R^r = -k_{R\|}^r - k_{R\perp}^r$ where $k_{R\|}^r$ and $-k_{R\perp}^r$ are the parallel and perpendicular components of k_R^r to the grating surface. This double diffraction then results in a new direction $k_R^{2d} = k_{retroreflect} + k_g^{-1} = -k_{R\|}^r + k_{R\perp}^r$. The end result is that the double diffracted vector k_R^{2d} is the same as the reflected-right incident arm with its parallel component reversed. This is exactly the direction we require, as it is equivalent to the direction of the reflected left arm (k_L^r). Both the left reflected arm and the double diffracted right arm are then detected at phase meter ϕ_4.

A few additional considerations will need to be taken into account. Namely, since the right arm undergoes diffraction twice the grating phase it incorporates is doubled. In addition, comparing the path length of the double diffracted arm k_R^{2d} with that of the reflected left arm k_L^R in figure 5-13 we notice that k_R^{2d} contains an additional optical

path. This additional path length corresponds to twice the path traversed from the surface of the substrate to the retro-reflecting mirror (i.e., it traverses this distance twice). The surface height must be properly accounted for because an increase in the surface height in the normal direction Z by Δz decreases the phase of the the diffracted arm by $\Delta\phi = 2k_z^d \Delta z$, where k_z^d is the propagation component along the Z-axis.

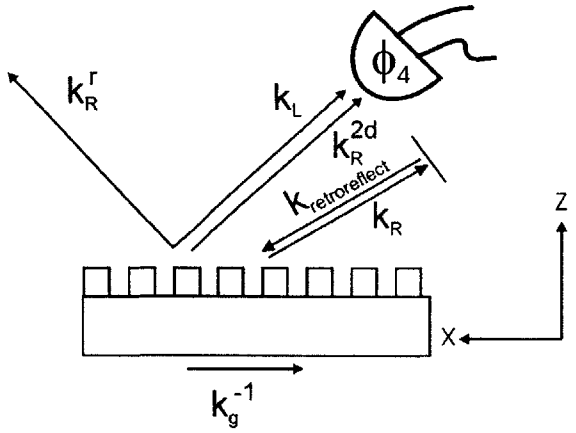

Figure 5-13: *Dual pass reading mode showing the reflected right arm k_R^r, the reflected left arm k_L^r, negative first-order grating vector k_g^{-1}. The back-diffracted right arm k_R^d, the retroreflected beam $k_{retroreflect}$ and the double diffracted beam k_R^{2d} are also shown.*

In general, an ideal grating in the X-Y plane described by coordinates (x, y) has a phase $\theta_{grating} = k_{xg}x + k_{yg}y$. Using our convention, the zero degree orientation is defined such that $k_{xg} = \frac{2\pi}{P_x}$ and $k_{yg} = 0$. During diffraction, some of the power from the reflected beam is transferred to the various diffracted orders. These orders must obey the diffraction condition

$$k^2 = (k_x^r + k_{xg})^2 + (k_y^r + k_{yg})^2 + (k_z^d)^2, \tag{5.48}$$

where $k = \frac{2\pi}{\lambda}$ and $\lambda = 351nm$ is our wavelength, and k_z^d is the z component of the diffracted beam. In addition, k_x^r is the X component of k vector of the reflected beam. Similarly k_y^r is the Y component of the reflected beam. The above equation simply states that the reflected beam gets scattered by the grating vector which transfers

momentum $\hbar k_{xg}$ and $\hbar k_{yg}$ to become the diffracted beam.

In the Littrow configuration for our grating at a 0 degree orientation $k_{xg} = -2k_x^r$ and $k_{yg} = k_y^r = 0$. To no surprise, in this condition the Z component of the back-diffracted beam k_z^d must satisfy the same energy conservation rule as the reflected beam and hence the magnitude of the Z-components are the same. For example, it is clear that the reflected beam must satisfy $k^2 = (k_x^r)^2 + (k_z^r)^2$ and from equation 5.48 the back-diffracted beam must similarly satisfy $k^2 = (k_x^r)^2 + (k_z^d)^2$ which allows us to conclude $k_z^d = k_z^r$.

The above analysis shows that we can calculate the diffraction grating Z component, k_z^d, for a given grating orientation. In the zero degree orientation, the grating axis is aligned so that $k_{xg} = \frac{2\pi}{P_x}$ and $k_{yg} = 0$. If we rotate the grating by θ, we will find a new grating vector k_g' with components $k_{xg}' = k_{xg}cos(\theta) + k_{yg}sin(\theta)$, and $k_{yg} = -k_{xg}sin(\theta) + k_{yg}cos(\theta)$. Once the diffraction grating vector components are determined, k_z^d may be found using equation 5.48. The component (k_d^z) may then be used to determine the height of the surface as will be shown in the next section.

5.6.1 Determining the surface of a grating

The technique used to determine the surface height requires an understanding of the difference between the positive and negative diffracted orders of the grating. Referring to figure 5-14, we have labeled the positive order grating vector as k_g^{+1}, and the negative grating vector order k_g^{-1}. These can be thought of as positive and negative spatial frequencies of the grating. The need for a vector is required because the dimensionality may in general be greater than one.

If the phase of the grating increases in the direction of the positive grating vector, we will obtain a positive phase shift in the corresponding diffracted beam of positive diffracted order (k_d^{+1}). The other diffracted beam (k_d^{-1}) which gets scattered by the negative spatial frequency grating vector (k_g^{-1}) obtains a negative phase shift.

Using dual pass reading mode we can choose to retroreflect either order. We will continue to use the traditional reading mode measurement (ϕ_3) described in equation 5.5 which provides the phase difference between the left and right arms. In addition,

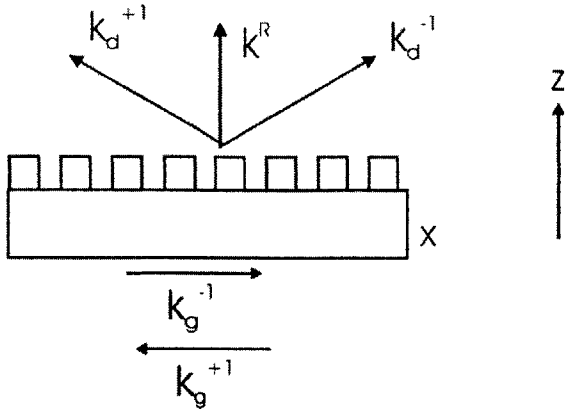

Figure 5-14: A grating with two spatial frequencies k_g^{+1} and k_g^{-1}. These spatial frequencies transfer positive and negative momentum to the reflected beam resulting in postive k_d^{+1} and negative k_d^{-1} diffracted orders. If the phase of the grating increases along the direction of the k_g^{+1}, the beam with k_d^{+1} will result in a positive phase shift, while k_d^{-1} will obtain a negative phase shift.

we will use the same principle as in traditional reading mode to determine the phase of the grating. That is, we will use phase meter ϕ_4 as shown in figure 5-13 for our dual-pass reading mode measurements. The main difference is that the signal which is collected in phase meter ϕ_4 contains information regarding the grating phase and surface profile along with the phase difference between the left and right arms. Subtracting out the phase difference between the left and right arms from the phase ϕ_4 will leave the phase information describing the surface of the grating substrate and the grating phase.

The fact that the surface information is collected in the double diffracted beam could be understood by observing the additional optical path length which is incorporated into dual pass reading mode. This is one of the major differences between the traditional and dual pass reading modes. If the surface height is raised, this additional optical path is shortened.

Now with a qualitative explanation of dual pass reading mode we will formalize the measurement process. If we retroreflect the positive diffracted beam k_R^{d-1} order and compare the phase of the double diffracted beam k_R^{2d} with the phase of the reflected

left arm we obtain at phase detector ϕ_4

$$\phi_4 = 2\theta_{grating} + 2\theta_{surface} + \phi_R - \phi_L. \tag{5.49}$$

We will define our first measurement S_{m1} to be

$$S_{m1} = \phi_4 - \phi_3 = 2\theta_{grating} + 2\theta_{surface}. \tag{5.50}$$

In a symmetric configuration in which both the positive and negative order diffracted beams k_d^{+1} and k_d^{-1} have the same Z component k_z^d the contribution of the surface of the grating to the measured phase ϕ_4 will be the same for both diffracted orders. However, the grating phase will differ by a negative sign depending on which diffracted order is used. In such a configuration, we will perform a second measurement by retroreflecting the positive diffracted order k_d^{+1}. This would result in a similar measured signal S_{m2}

$$S_{m2} = \phi_4 - \phi_3 = -2\theta_{grating} + 2\theta_{surface}, \tag{5.51}$$

where the phase of the grating has opposite sign because we have used the double diffracted beam from the positive order diffraction. Subtracting both measurements $S_{m1} - S_{m2} = 4\theta_{grating}$ will allow one to determine the phase of the grating. Summing both measurements $S_{m1} + S_{m2} = 4\theta_{surface}$ will allow one to solve for the height of the surface.

As a final note of caution, in order to cancel the effect of the surface on measurements S_{m1} and S_{m2} one requires both diffracted beams to contain the same k_z^d component. This may be accomplished by using a 90 degree orientation to be described in the next section in which both diffracted orders have symmetry about the Z axis. Once the surface is calculated, any diffracted order may be used in any orientation to determine the grating phase. In the sections to follow, we will illustrate the grating orientation for 90 and 45 degree dual pass reading mode configurations in order to perform the proof of concept experiments to determine the system error

and surface profile.

5.6.2 90 Degree Orientation

We will refer to figure 5-15 (a) in describing the 90 degree orientation of our grating in dual pass reading mode. In the zero degree orientation, our nominal grating vector lies along the X axis. Figure 5-15 shows the new grating orientation when it is rotated 90 degrees. The grating lines and spaces are shown along with the positive and negative spatial frequencies k_g^{-1} and k_g^{+1}. Accordingly, the incident arm k^i has a component which is projected onto the $X-Y$ plane in figure 5-15. Some of the power in the reflected arm k^r gets diffracted into two arms k_d^{+1} and k_d^{-1}. In figure 5-15(b)two mirrors are shown which retroreflect the diffracted orders back onto themselves resulting in a double diffracted beam which is anti-parallel to the incident beam. Figure 5-15 only captures the directions of the diffracted ordersx in the $X-Y$ plane. We must now calculate the diffracted components along the Z-axis. In the

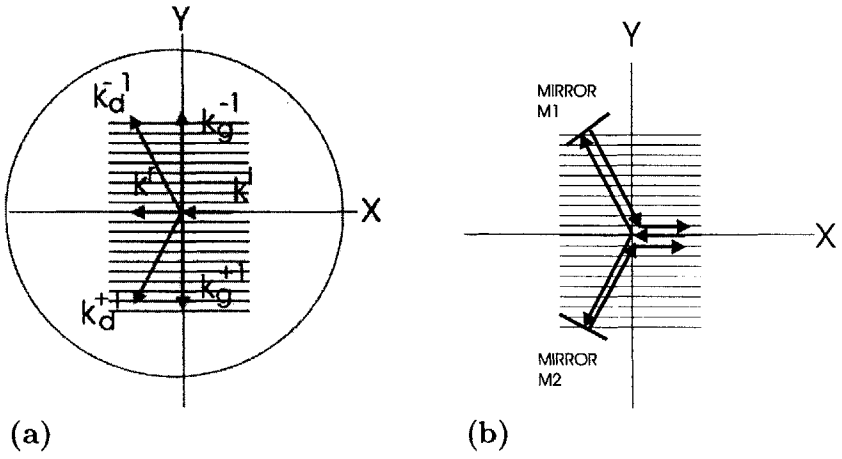

Figure 5-15: *Dual Pass Reading Mode in the 90 degree orientation: (a) An incident beam k_i interacts with the grating vector k_g^{-1} and k_g^{+1} and is diffracted into two orders k_d^{-1} and k_d^{+1}. Also shown is the reflected beam k^r (b) A mirror M1 may be used to retroreflect the negative order k_d^{-1}. Alternatively, mirror M2 may be used to retroreflect the positive order k_d^{+1}.*

previous section we described the conditions for the zero degree orienation. We stated

the grating components mathematically as $k_{xg}^{-1} = -2k_x^R$ and $k_{yg} = k_y^R = 0$. When we rotate the grating by 90 degrees, we find $k_{xg}^{-1} = 0, k_{yg}^{-1} = -2k_x^R$. Using equation 5.48 one may see that the Z component of the diffracted beam will also change. We will rewrite equation 5.48 with these new values as

$$k^2 = (k_x^r)^2 + (k_{yg})^2 + (k_z^d)^2 \tag{5.52}$$
$$k^2 = (k_x^r)^2 + (2k_x^r)^2 + (k_z^d)^2.$$

This allows us to solve for the Z component of the diffracted beam. This Z component is given by

$$k_z^d = k\sqrt{1 - 5(\frac{k_x^r}{k})^2} \tag{5.53}$$
$$k_z^d = k\sin(\gamma). \tag{5.54}$$

Substituting our value $k_x^r = 1/2\frac{2\pi}{574nm}$ we find $\sin(\gamma) = \sqrt{1 - 5/4(\frac{351}{574})^2}$ resulting in $\sin(\gamma) = 0.7299$. This number will be required to determine the surface profile of the grating in the 90 degree orientation, as described in the previous section. The surface profile, $h(x, y, z)$, may be calculated by taking the sum of the two measurements described by equations 5.50 and 5.51. This is written as

$$h(x, y, z) = 1/(2k_z^d)\frac{(S_{m1} - S_{m2})}{2} \tag{5.55}$$
$$h(x, y, z) = 1/(2k_z^d)\theta_{surface}$$

5.6.3 Experimental Results for 90 Degree Orientation

A proof of concept experiment was performed using a 165x220 mm BK7 substrate (Substrate ID # 103039A-005-3AL) which was oriented in a 90 degree orientation. The experiment performed in figure 5-16 consists of a right incident arm which undergoes diffraction once by the grating vector k_g^{-1}. The right arm is retroreflected and undergoes diffraction one more time by the same grating vector k_g^{-1}. This results in the double diffracted vector picking up twice the phase of the grating as previously

described. The double diffracted right arm with direction k_{2d} is recombined with the left reflected arm k_L^r at the phase detector ϕ_4 as shown in figure 5-13.

Shown in figure 5-16 is the measurement $(\phi_4 - \phi_3)/2$. The phase meter measurement ϕ_3 detects the phase of the same signal that is used in the traditional reading mode experiments. This phase ϕ_3 corresponds to the phase difference between the left and right arms.

The phase meter outputs (ϕ_4 and ϕ_3) are scaled appropriately so that the units are in radians.

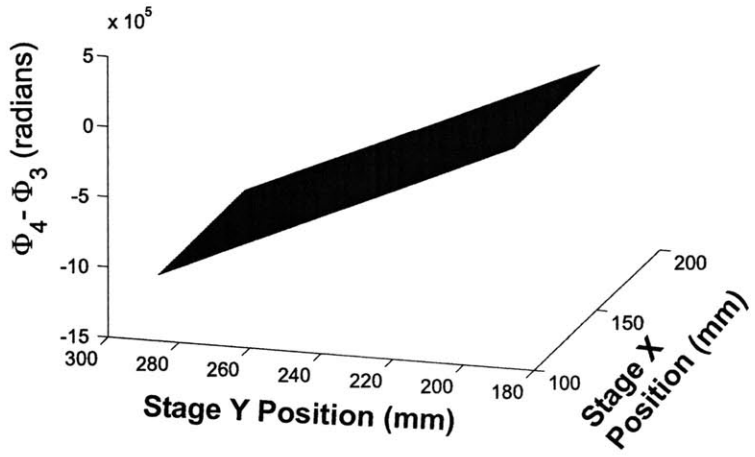

Figure 5-16: *This measurement corresponds to* $S_{m1} = 1/2(\phi_4 - \phi_3)$. *Note the plane orientation is along the* Y *axis as expected for a 90 degree orientation. The period may be extracted by using a planar fit to the data. The measured period is 574.16 nm along the* Y *axis.*

From figure 5-16 one may determine the period of the grating. This is done by finding the best fit plane to the data, using a least squares fit. The best fit plane is of the form $\theta_{fit}(x, y) = ax + by$, where the coefficients a and b are the slopes along x and y respectively. The linear phase of the grating is of the form $\theta_{grating}(x, y) = k_{xg}x + k_{yg}y$. Therefore, the period of the grating is easily determined by recognizing that $k_{xg} = \frac{2\pi}{P_x} = a$ and $k_{yg} = \frac{2\pi}{P_y} = b$.

Using the technique described to determine the period of the grating, we found

that $P_y = -574.16$ nm. The period in the x direction was found to be much larger resulting in $P_x = -287.05$ μm indicating that the alignment was fairly good. The fact that we are looking at the negative diffracted order is available from the sign of the data. In analyzing the above data, one must take care to realize that the grating axis is inverted from the stage axis. As the stage moves from left to right, one reads the grating from right to left. Furthermore, this particular grating was written at 574 nm on the nanoruler. This gives us confidence that we are indeed measuring the correct quantity, and that our scale factors are properly accounted for.

The true quantity of interest is the nonlinear phase error. A linear phase error corresponds to an error in the period of the grating which is easier to correct. In order to determine the nonlinear phase error, we proceed to subtract the best fit plane from the data in figure 5-16. The resulting nonlinear phase is shown in 5-17. A smoothing filter is applied to the data to reduce the measurement noise.

A potential cause of the measurement noise may be attributed to vibrations. In this proof of concept experiments for the sake of time and cost considerations mounting was achieved using standard laboratory optical posts and mounts. The rather bulky assembly could be customized and optimized for reducing vibrations.

With the most stable mounting however there are some additional points to consider. One aims to achieve the same noise level in dual pass reading mode as in the traditional reading mode. However, it must be realized that dual pass reading mode traverses the same optical paths as in the traditional reading and adds an extra path length. Indeed, this additional optical path length allows us to measure the surface of the substrate. On the other hand, it comes at the cost of making the measurement vulnerable to independent vibrations of the stage in the normal Z direction. In other words, the measurement process can not distinguish between a vibration of the stage and a surface feature.

The question arises as to whether or not vibrations of the stage in the Z direction are of concern for writing gratings in the system. Paul Konkola provides a detailed analysis of the effect of translations in the Z direction. To the first-order, translations of the "metrology block" which consists of the various phase meters are of inconse-

quence. However, if there is a rotation θ_Y of the stage along the $X - Z$ plane a translation ΔZ produces an error ϵ of magnitude $\epsilon = \Delta Z \theta_Y$. Since the pitch of the stage θ_Y is typically on the order of 10 microradians, writing mode is quite insensitive to vibrations of the stage in the Z direction. In addition, the specification for our stage is within 2 microns peak to valley over the total range of travel.

Since vibrations along the Z axis or not of particular concern for the nanoruler an effort has not been made measure these vibrations. However, one may rely on averaging to remove the vibrational noise from the measurement. Improved mounting of the mirror assembly would also serve to reduce the noise. In additon, a stronger intensity in the double diffracted beam would result in an improved signal strength at the detectors, which may also serve to reduce any measurement noise. Nonetheless, the results have the correct order of anticipated magnitude. Figure 5-17 contains

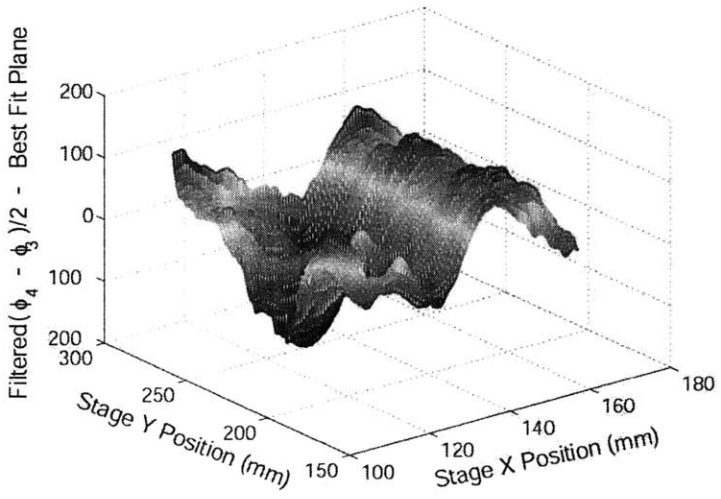

Figure 5-17: *A best fit plane is removed from the measurement* ($S_{m1} = \Phi_4 - \Phi_3)/2$ *resulting in the above nonlinear phase measurement* $S_{m1} = 2\theta_{grating} + 2\theta_{surface}$. *In addition, the data is scaled by a factor of* $\frac{P_y}{2\pi}$, *where* $P_y = 574.16$ *nm is the measured period of the grating to obtain units of nm*

phase information of both the surface and phase of the grating.

In order to remove the effect of the surface of the grating from the phase in figure 5-17 we will need to perform one more measurement. This measurement is performed

by retroreflecting the positive diffracted order k_d^{+1} in figure 5-15 (b) by using the retroreflecting mirror M2. The double diffracted beam from the positive order then gets recombined with the reflected left beam into phase meter ϕ_4.

We observe that this measurement results in $S_{m2} = -\theta_{grating} + \theta_{surface}$. If we then subtract out measurement S_{m1} from S_{m2} we cancel the effect of the surface and find the phase of the grating to be $\theta_{grating} = (S_{m1} - S_{m2})/2$. Figure 5-18 illustrates the measured grating phase in units of nanometers. Similarly, we can calculate the

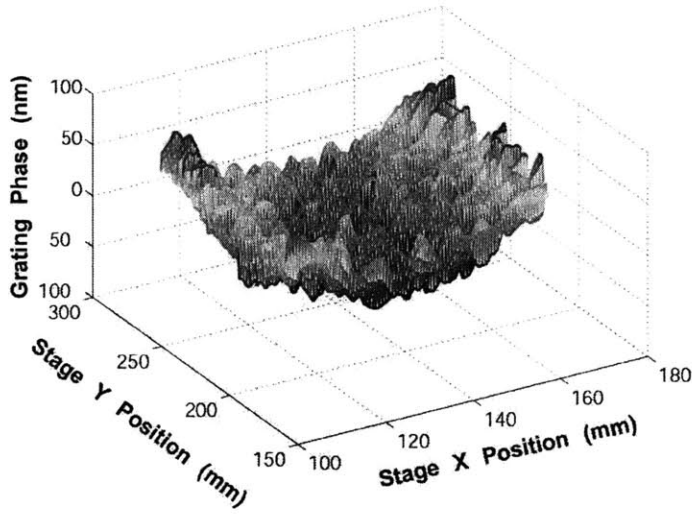

Figure 5-18: *Phase of grating in units of nanometers obtained after computing* $\theta_{grating} = \frac{1}{k_{xg}}(S_{m1} - S_{m2})/2$

surface of the grating by summing the two measurements S_{m1} and S_{m2}. The scale factor which allows one to convert the phase of the resulting measurement to a surface height is $\frac{1}{2k_z^d}$ as previously derived.

The measurement of the surface of the grating is shown in figure 5-19. These results are remarkable in the sense that this is the first time the nanoruler tool has been abled to measure the surface profile. Undoubtedly, improvements to the measurement process may be made to reduce the noise level. If all else fails, one may take several measurements and perform further averaging to reduce the noise level. One does note some surface features on the order of 200 nanometers peak to valley. These features

are pronounced macroscopically and have long spatial frequencies. For example, the data shown in figure 5-19 spans a range of approximately 70x150 mm. Furthermore, the BK7 glass substrate has been coated with aluminum which may lead to surface roughness on such a large length scale.

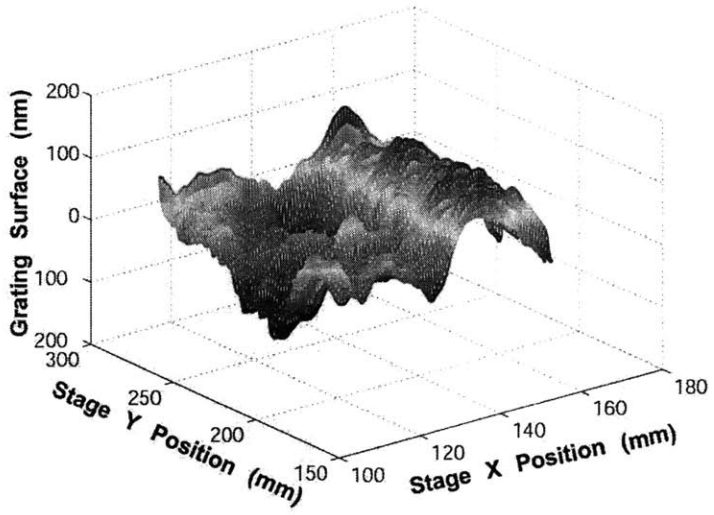

Figure 5-19: *Surface of the grating obtained after calculating* $h(x,y) = \frac{1}{2k_z^d}(S_{m1} + S_{m2})/2$

5.6.4 45 Degree Orientation

In order to complete the rotational grating test which we have developed to characterize the system distortion, it will be necessary to obtain a reading of the grating in a 45 degree orientation. Our novel dual pass reading mode technique is a convenient method which will allow us to read the grating at such an angle. It is noteworthy that any alignment scheme could be used to recombine the diffracted order at 45 degrees with the zero reflected order. However, dual pass reading mode only requires one additional component for each orientation (i.e., a retroreflecting mirror). This could be compared with having to use two additional mirrors and a beam splitter assembly, a

total of three components, which must be aligned for each grating orientation. These three components would complicate the alignment process by introducing more degrees of freedom and lead to difficult packaging requirements in the nanoruler which does not offer much real estate.

The 45 degree orientation of the grating and the diffracted orders which result are shown in figure 5-20. From the figure, we take notice that the positive diffracted order k_d^{+1} and the negative diffractive order k_d^{-1} are no longer symmetric about the Z axis. In this orientation, we may not cancel the effect of the surface as we did in the 90 degree orientation. However, since the surface profile has already been measured in the 90 degree measurement we may subtract its effect from the measurement directly.

In order to subtract the effect of the surface from the measurement one needs to calculate the appropriate scale factor which converts the surface height to units of radians. This scale factor may be verified to be twice the Z component of the diffracted order k_z^d as decribed in section 5.6.1. Following a similar procedure outlined for the 90 degree section, we recognize that in the 45 degree orientation $k_{xg} = \frac{\sqrt{2}}{2}(2k_x^r)$ and $k_{yg} = -k_{xg}$ as shown in figure 5-20. Common to all orientations, the reflected beam does not change and has an x component k_x^r, and no y component $k_y^r = 0$.

From equation 5.48 we find

$$k^2 = (k_x^r + \sqrt{2}k_x^r)^2 + (-\sqrt{2}k_x^r)^2 \tag{5.56}$$

we find $k_z^d = k\sqrt{1 - (\frac{k_x^R}{k})^2(7 + 2\sqrt{2})}$.

5.6.5 45 Degree Orientation Experimental Results

Proceeding similarly to the 90 degree orientation, we perform a measurement S_{m2} where we retroreflect the positive order k_d^{+1} of figure 5-20 and recombine the double diffracted arm with the left reflected arm into the phase meter Φ_4 as shown in figure 5-13. When we subtract out the phase of the left and right arms measured in phase meter ϕ_3, we obtain twice the phase of the grating and twice the contribution from the surface as shown in figure 5-21. Figure 5-21 contains several pieces of information.

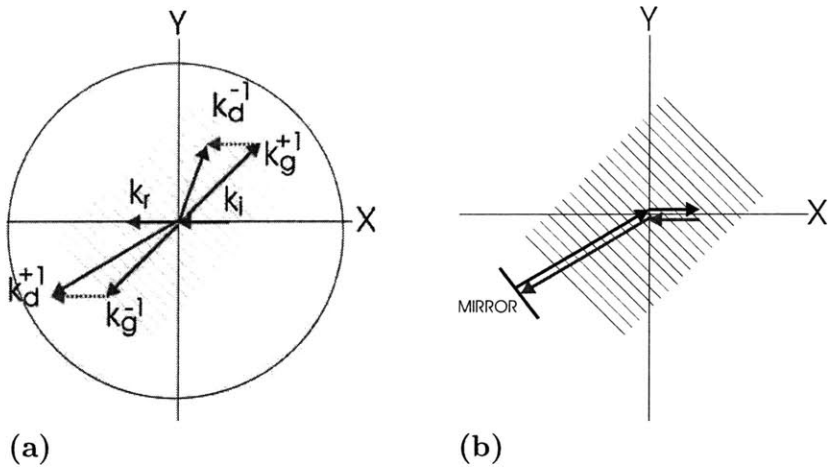

Figure 5-20: *Dual pass reading mode in the 45 degree orientation: (a) An incident beam k_i interacts with the grating vector k_g^{-1} and k_g^{+1} and is diffracted into two orders k_d^{-1} and k_d^{+1}. (b) A mirror M1 may be used to retroreflect the negative order k_d^{-1}. Alternatively, mirror M2 may be used to retroreflect the positive order k_d^{+1}.*

The best fit plane to the data provides us directly with the k components of the grating. We find the best fit plane coefficients divided by 2 produce $k_{xg} = 7.5277\text{x}10^3$ mm^{-1} and $k_{yg} = 7.9509\text{x}10^3$ mm^{-1}. This immediately tells us the grating angle α which may be found by taking the inverse tangent of the ratio of the components. We find $\alpha = atan(k_{xg}/k_{yg}) = 46.56^o$, fairly close to the 45 degree orientation we attempted to align to. The alignment for these proof of concept experiments were performed by using rulers to mark distances. Future upgrades to the system will allow for precision rotation of the grating which is not available at the time of this thesis. The grating periods in the X and Y directions deduced from the measurements resulted in $P_x = 834$ nm and $P_y = 790.25$ nm respectively.

In order to obtain the nonlinear part of the measurement, we must subtract out the best fit plane from the data. This is shown in figure 5-22

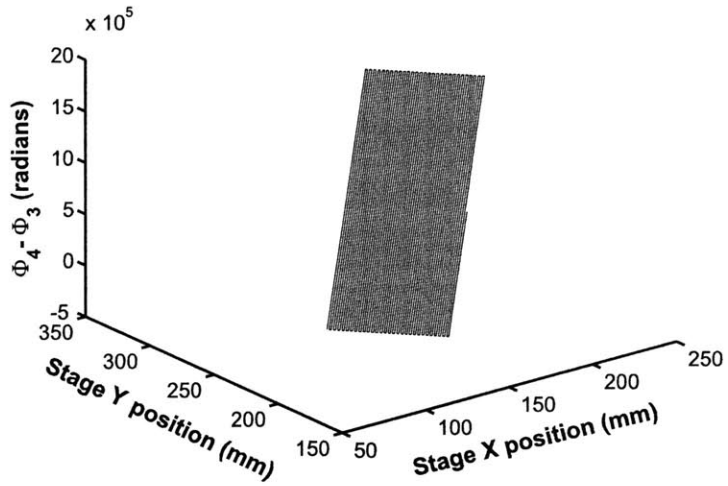

Figure 5-21: *Output from dual pass reading mode experiment in 45 degree orientation. The output consists of twice the grating phase and twice the surface contribution.*

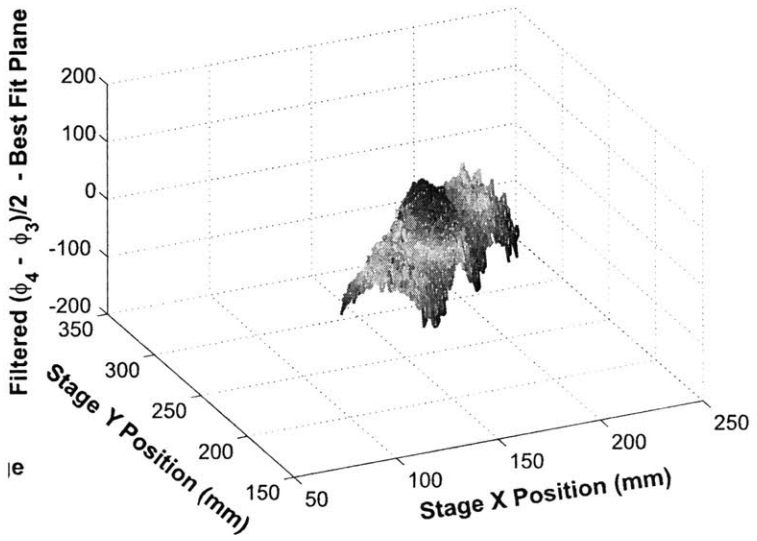

Figure 5-22: *Surface of the grating obtained after calculating* $h(x,y) = \frac{1}{2k_z^d}(S_{m1} + S_{m2})/2$.

APPENDIX A: Matlab Code For Mirror Reconstruction

```
clear all;
close all;

HENERES  = 3.236409608091024e+009; %counts/meter for stage interferometer
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
gratingperiod=574; %grating period in nm



M=4000;
N=2;
%Lets Now Find the Frequency Response
B=zeros(N+1,1);
B(3*N+1,1)=+1;
B(2*N+1)=1;
B(N+1)=-1;
B(1)=-1;



figure;
zplane(B',1);

%Now for the Approximate Transfer Function

q=roots(B)

A0=.85;
z0=A0*exp(i*pi/2);
m=[.99,-.99,-z0,z0,-z0,z0];
q2=poly(m)
```

```
q3=q2;
figure;
zplane(q3,1);

filename='xm020706-001.txt';
data = load(filename); %X Theta, Y
dataextract2;
h=6.375 %mm

figure;
[h2,w2]=freqz(q3,1,M);
[h1,w1]=freqz(B',1,M);

a=-1*abs(h2(400)/h1(400));
filtercoeff=1/a*q3;
[h2,w2]=freqz(filtercoeff,1,M);

freqz(filtercoeff,1,M);
hold on;
freqz(B',1,M);


figure;
thetaxconv = 1/HENERES/(.0254*.5)*1e6; %(it is being divided by 12.4) %for converting
y=x(:,ycol)*1/HENERES;
xd=x(:,xcol)*1/HENERES;
xtheta=x(:,thetaxcol)*thetaxconv;
ytheta=x(:,ydaccol)*thetaxconv;
Comment=''
```

```
%********************************** Theta Vs TIME
subplot(2,2,1);
plot(t,xtheta);
title(['\Theta_{x} vs Time Filename ',filename,Comment])
xlabel('time (s)');
ylabel('\Theta_{x}, micro radians')

%********************************** ThetaX Vs Y
subplot(2,2,2);
plot(y,xtheta);
title(['\Theta_{x} vs y Filename ',filename,Comment])
xlabel('y');
ylabel('\Theta_{x}, micro radians')

%********************************** Y VERSUS TIME
subplot(2,2,3);
plot(t,y); %For converting phase meter
title(['Y vs Time Filename ',filename])
xlabel('time (s)');
ylabel('Y(m)')


%********************************** ThetaY Vs Y
subplot(2,2,4);
plot(y,ytheta);
title(['\Theta_{Y} vs y Filename ',filename,Comment])
xlabel('y');
ylabel('\Theta_{Y}, micro radians')

%********************************************** Delta Theta
```

```
dtheta=(xtheta+ytheta);
figure;
subplot(2,1,1);
plot(y,dtheta);
xlim([0 1]);
title(['\Delta \theta ',filename,Comment])
xlabel('Distance(mm)')
c=median(dtheta);
dtheta=dtheta - c; %Remove Constant
hold on;
plot(y,dtheta,'r');

subplot(2,1,2);
four2=abs(fft(dtheta));
xval=linspace(0,2,length(four2));
plot(xval,four2);
xlim([0 1]);
title(['Fourier Transform of \Delta \theta ',filename,Comment])
xlabel('Distance(m)')

figure;
xreconstruct=filter(2*h,filtercoeff,dtheta*1e-3);
plot(y,xreconstruct);
%axis([0,300,0,1])
xlabel('stage position (mm)');
ylabel('Reconstructed Mirror Profile(\mum)')
title(['Reconstructed Mirror Profile with Inverse Measurement Transfer Function'])
[P,S]=polyfit(y,xreconstruct,1);
linfit=polyval(P,y);
hold on;
```

```
plot(y,linfit,'k');


mirror=(xreconstruct-linfit)*1e3;
figure;
subplot(2,1,1);
plot(y*1e3,mirror);
title(['Mirror Nonflatness: ',filename,Comment])
ylabel('Mirror nonflatness nm');
xlabel('Stage Y position (mm)')

subplot(2,1,2);
four=fft(mirror);
plot(xval,abs(four));
xlim([0 1]);
title(['FFT of Mirror Nonflatness ',filename,Comment])
Ylabel('FFT Coefficients');
xlabel('Normalized Frequency')

h=[1 1 1 1]/4;
mirrorfilt=filter(h,1,mirror);
figure;
plot(y*1e3,mirrorfilt);
xlabel('Stage Y position (mm)')
ylabel('Mirror nonflatness (nm)');
```

APPENDIX B: Matlab Code For Grating Translation Test

```
%Lets suppose we have the following system distortion:
close all;
clear all;


   xvec=-100:1:100;
   yvec=xvec;
   [X,Y]=meshgrid(xvec,yvec);
   S=50*sin(2*pi/180*X) %System Distortion
   G=((X/100).^2+(Y/100)^3)*75+50*cos(2*pi*Y/90); %Grating Distortion
   mesh(S);
   title('System Distortion');

   figure;
   mesh(X,Y,abs(G+50));  %ADD AN OFFSET SO WE ARE DEALING WITH POSITIVE QUANTITIES
   title('Grating Distortion');
   xlabel('X (mm)');
   ylabel('Y (mm)');

   figure;
   mesh(X,Y,G+S);
   title('Grating Distortion + System Distortion');
   xlabel('X (mm)');
   ylabel('Y (mm)');

   figure;

   %%%%%%%%RECOVER THE GRATING DISTORTION BY TRANSLATION %%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

```
dx=1;dy=1;
[row,col]=size(G);
hx=[-1 1;0 0]
Gfx=filter2(hx,G);
%%%%%THE LAST COLUMN NEEDS CORRECTION: See K.R. Freischlad and C. Kolipoulos%%%%
Gfx(:,col)=-1*sum(Gfx(:,1:col-1),2);
mesh(Gfx);
title('Output of 2D impulse response which corresponds to differencing');

dx=1;dy=1;
[row,col]=size(G);
hy=[-1 0; 1 0]
Gfy=filter2(hy,G);
Gfy(row,:)=-1*sum(Gfy(1:row-1,:));

N=row;
i=sqrt(-1);
Xp=X+100;
Yp=Y+100;
reconstruct=1 ./ ( 4*sin(pi/N*(Xp)).^2+4*sin(pi/N*Yp).^2) .*  ( (exp(-2*pi*i*Xp/
reconstruct(isinf(reconstruct))=0;
reconstruct(isnan(reconstruct))=0;
reconstruct(1,1)=0;
%  figure;
%   mesh(X,Y,abs(reconstruct));
w=(fft2(reconstruct));

figure;
%ADD A DC OFFSET BECAUSE ABSOLUTE FUNCTION WILL NOT WORK WITH
%NEGATIVE QUANTITIES
```

```
mesh(X,Y,1/N^2*abs(w+100*N^2));
xlabel('X (mm) ');
ylabel('Y (mm) ');
title('Reconstructed Grating Distortion using Inverse Filter');
%  figure;
% freqz2(h);


figure;
freqz2(hx);
xlabel('\omega_{x}');
ylabel('\omega_{y}');
zlabel('|H_{\Delta x}|');


figure;
freqz2(hx+hy);
xlabel('\omega_{x}');
ylabel('\omega_{y}');
zlabel('|H_{\Delta x}+H_{\Delta y}|');
figure;
%NOW LETS FIND THE TOTAL INVERSE WIENER FILTER AS DESCRIBED BY MONTOYA
k=1;
Hdxdycc=(1-exp(2*pi/N .* Xp))+(1-exp(2*pi/N .* Yp));
numerator=Hdxdycc.*fft2(Gfx+Gfy)
denominator=abs(Hdxdycc).^2+k;
Dglk=numerator./denominator;
dgreconstruct=ifft2(reconstruct);
mesh(abs(dgreconstruct + 100));
title('Obtained using Wiener 2D');
```

APPENDIX C: C Code for Beam Blanking: filename = j-sbil-bleamblnk-works.c

```
#include <ixcbsp.h>
#include "j-mfe_def.h"
#include "j-sbil_def.h"
#include "j-sbil_sdram.h"
#include "j-sbil_vme1181.h"
#include "j-vme2510b_def.h"
#include "j-sbil_zmi.h"
#include "beamblank.h"

extern int waitstate[2];
extern int turnonbusystat;
extern int turnoffbusystat;
extern int ampaddr[2][3];
extern int doseoff;
extern VME_DEVICE vme_dev_d32;
extern int datadisableaddr[2];
extern int resetzmi;
extern int fs;
extern int fr;
extern int fringescale;
extern int freqaddr[2][3];
extern double timercount;
extern int signalstatphi1;
extern int signalstatphi2;

void turnonbeams(int *ampdose1, int *ampdose2, int *velocity)
{
/*int word=0x144F0020;*/
```

```
if(turnonbusystat==1)
{

   vmex_write(&vme_dev_d32, (void *)MFE_ADDR, &datadisableaddr[1], 1, SWAP); /*DISABL

vmex_write(&vme_dev_d32, (void *)MFE_DATA, ampdose1, 1, SWAP); /*PASS AOM1 DATA*/



}
else if(turnonbusystat==2)
{
 /*ENABLE AOM1*/
vmex_write(&vme_dev_d32, (void *)MFE_ADDR, &ampaddr[1][0], 1, SWAP); /*ENABLE AOM1 */

}
else if(turnonbusystat==3)
{
/*Turn Back on AOM Power Ch2*/
/*DISABLE, DATA*/
      vmex_write(&vme_dev_d32, (void *)MFE_ADDR, &datadisableaddr[1], 1, SWAP); /*DIS

      }
      else if(turnonbusystat==4)
      {
vmex_write(&vme_dev_d32, (void *)MFE_DATA, ampdose2, 1, SWAP); /*PASS AOM2 AMP DATA *

}
/*We would like to reset ZMI_PHI4 when we have intensity*/
else if(turnonbusystat==5)
```

```
{
/*ENABLE AOM2*/
vmex_write(&vme_dev_d32, (void *)MFE_ADDR, &ampaddr[1][1], 1, SWAP); /*ENABLE AOM 2 A

}
else if(turnonbusystat==6)
{
 vmex_write(&vme_dev_d32, (void *)MFE_ADDR, &datadisableaddr[1], 1, SWAP); /*DISABLE

}
else if(turnonbusystat==7)
{
/*fs=fr + (int)(*velocity/100 * fringescale *0xffffffff/300.0);*/
fs=fr;
vmex_write(&vme_dev_d32, (void *)MFE_DATA, &fs, 1, SWAP); /* ENABLE freq1 data */

}
else if(turnonbusystat==8)
{
vmex_write(&vme_dev_d32, (void *)MFE_ADDR, &freqaddr[1][0], 1, SWAP); /* ENABLE freq
}
else if(turnonbusystat==9)
{
vmex_read(&vme_dev_d32, (void *)(ZMI_PHI1_BASE), &signalstatphi1, SWAP); /*read theta

vmex_read(&vme_dev_d32, (void *)(ZMI_PHI2_BASE), &signalstatphi2, SWAP); /*read theta

signalstatphi1=signalstatphi1>>16 & 0xFF;
signalstatphi2=signalstatphi2>>16 & 0xFF;
if ((signalstatphi1 &0x08) && ( signalstatphi2 & 0x08)) /*Tells us the measurement Si
```

```
{
/* vmex_write(&vme_dev_d32, (void *)MFE_ADDR, &datadisableaddr[1], 1, SWAP);*/ /*DISA

}
else
{
turnonbusystat=turnonbusystat-1; /*REPEAT, Phi1 and Phi2 Have No Signal*/
}
}
/*Reset and Wait 25 msecs if we have intensity*/
else if(turnonbusystat==10)
{
   /*position_reset_zmi(ZMI_PHI1_BASE);*/
   /*RESET PHASE METER BOARDS*/
vmex_write(&vme_dev_d32, (void *)ZMI_PHI1_BASE, &resetzmi, 1, SWAP); /*pass freq1 dat

vmex_write(&vme_dev_d32, (void *)ZMI_PHI2_BASE, &resetzmi, 1, SWAP); /*pass freq1 dat

}
else if(turnonbusystat==11)/*We have issued a reset and now waiting for registers to
{
vmex_read(&vme_dev_d32, (void *)(ZMI_PHI1_BASE), &signalstatphi1, SWAP); /*read theta

    vmex_read(&vme_dev_d32, (void *)(ZMI_PHI2_BASE), &signalstatphi2, SWAP); /*read t

signalstatphi1=(signalstatphi1>>16) & 0xFF;
signalstatphi2=(signalstatphi2>>16) & 0xFF;
turnonbusystat=turnonbusystat-1;
if(waitstate[0]>100)
{
```

```
if((signalstatphi1 & 0x10) && (signalstatphi2 & 0x10)) /*Tells us the position reset
{
turnonbusystat=turnonbusystat+1; /*Advance*/
  }
}
waitstate[0]++;
}
else if(turnonbusystat==12)
{
/* fs=fr + (int)(*velocity/100*fringescale*0xffffffff/300.0);*/
fs=fr;
vmex_write(&vme_dev_d32, (void *)MFE_DATA, &fs, 1, SWAP); /* ENABLE freq1 data */
turnonbusystat = -1; /*Will become 1 at the end of Loop*/
turnoffbusystat=1;
timercount=0;
}
turnonbusystat=turnonbusystat+1;
}


void turnoffbeams(void)
{

if(turnoffbusystat==1)
{
/*UNMASK INTERRUPTS*/
vmex_write(&vme_dev_d32, (void *)(ZMI_PHI1_BASE + 0x18), &doseoff, 1, SWAP); /*UNMASK
vmex_write(&vme_dev_d32, (void *)(ZMI_PHI2_BASE + 0x18), &doseoff, 1, SWAP); /*UNMASK
}
else if(turnoffbusystat==2)
```

```
{
/*DISABLE, DATA, ENABLE AOM1*/
      vmex_write(&vme_dev_d32, (void *)MFE_ADDR, &datadisableaddr[1], 1, SWAP); /*DIS

vmex_write(&vme_dev_d32, (void *)MFE_DATA, &doseoff, 1, SWAP); /*PASS AOM1  DATA (Off

}
else if(turnoffbusystat==3)
{
vmex_write(&vme_dev_d32, (void *)MFE_ADDR, &ampaddr[1][0], 1, SWAP); /*ENABLE AOM1 */

}
else if(turnoffbusystat==4)
{
/*DISABLE, DATA, ENABLE AOM2*/
 vmex_write(&vme_dev_d32, (void *)MFE_ADDR, &datadisableaddr[1], 1, SWAP); /*DISABLE

}
else if(turnoffbusystat==5)
{
 vmex_write(&vme_dev_d32, (void *)MFE_DATA, &doseoff, 1, SWAP); /*PASS AOM2  DATA */

}
else if(turnoffbusystat==6)
{
 /*NOTE DATA IS ALREADY ON DATA LINES, AND THERE IS NO NEED TO PASS DATA*/
 vmex_write(&vme_dev_d32, (void *)MFE_ADDR, &ampaddr[1][1], 1, SWAP); /*ENABLE AOM2 *
        turnonbusystat=1;  /*Phase Meter is No Longer Reset because we are turning o
 turnoffbusystat=-1;
}
```

```
turnoffbusystat=turnoffbusystat + 1; /*It takes 5 cycles to Turn off the Beams */
}

void absolutephase_fct(int *phifringe1, int *phifringe2, int *phi1offset, int *phi2of
{
vmex_read(&vme_dev_d32, (void *)(ZMI_PHI1_BASE+0x0C), absphi1, SWAP); /*read phase di
vmex_read(&vme_dev_d32, (void *)(ZMI_PHI2_BASE+0x0C), absphi2, SWAP); /*read phase di
/*phixtheta=absphi3>>12;*/
/*phiytheta=absphi4>>12;*/
/* Absolute Phase is in the Form 0xAABB C000, Where C = 9thBitA 9thBitB 0 0 */
*absphi1total=((*absphi1>> 24) & 0xFF) + 256*( (*absphi1>>15) & 0x1);
*absphi2total=((*absphi2>> 24) & 0xFF) + 256*( (*absphi2>>15) & 0x1);
*absphi1total=-1 *  *absphi1total+512;  /*Phi1 and Phi2 Diagnostic Registers count in
*absphi2total=-1 *  *absphi2total+512;


*phi1offset = ((*phifringe1 & 0x1FF) - *absphi1total) & 0x1FF;
*phi2offset = ((*phifringe2 & 0x1FF) - *absphi2total) &0x1FF;

/*TAKE CARE OF SPECIAL CASE OF QUADRANTS I & IV FOR AVERAGING */
/*IF THE PHASE CHANGES BETWEEN 0 and 512, then make number above 256 negative*/
/*SPECIAL CARE MUST BE MADE FOR FIRST ITERATION.  MAKE PHIOFFSETOLD=256*/
if(timercount>1)
{
if((*phi1offset<128) && (*phi1offsetold>384)) /*Keep the number High*/
{
*phi1offset=*phi1offset+512;
}
else if((*phi1offsetold<128) && (*phi1offset>384)) /*Keep the number Low*/
{
```

```
*phi1offset=(*phi1offset-512);
}
if((*phi2offset<128) && (*phi2offsetold>384)) /*If Difference is Greater than 256*/
{
*phi2offset=*phi2offset+512;
}
else if((*phi2offsetold<128) && (*phi2offset>384)) /*If Difference is Greater than 25
{
*phi2offset=(*phi2offset-512);
}
}
  *phi1offsetcount = *phi1offsetcount + *phi1offset;
  *phi2offsetcount = *phi2offsetcount + *phi2offset;
  *phi1offset= (int) (*phi1offsetcount/timercount);
  *phi2offset= (int) (*phi2offsetcount/timercount);
*phi1offsetold = *phi1offset;
  *phi2offsetold = *phi2offset;
    timercount=timercount+1.0;
}
```

APPENDIX D: Matlab Code For Grating Rotation Test

```
%Lets suppose we have the following system distortion:
close all;
clear all;


   xvec=-100:1:100;
   yvec=xvec;
   [X,Y]=meshgrid(xvec,yvec);
   S=50*sin(2*pi/180*X) %System Distortion
   G=((X/100).^2+(Y/100)^3)*75+50*cos(2*pi*Y/90); %Grating Distortion
   mesh(S);
   title('System Distortion');

   figure;
   mesh(X,Y,abs(G+50));
   title('Grating Distortion');
   xlabel('X (mm)');
   ylabel('Y (mm)');

   figure;
   D0=S+G;
   D1=imrotate(G,45,'neares','crop')+S;
   D2=imrotate(G,90,'neares','crop')+S;
   D3=imrotate(G,135,'neares','crop')+S;
   D4=imrotate(G,180,'neares','crop')+S;
   D5=imrotate(G,225,'neares','crop')+S;
   D6=imrotate(G,270,'neares','crop')+S;
   D7=imrotate(G,315,'neares','crop')+S;
```

```
D=D0+D1+D2+D3+D4+D5+D6+D7;
figure;
mesh(X,Y,(D)/8);
xlabel('X (mm)');
ylabel('Y (mm)');
zlabel('Distortion (nm} ');
```

Bibliography

[1] *ZYGO 2/20 Integrated Interferometer System Reference Manual OMP-0229B (Zygo Corp., Middle Field, CT.).* 1993.

[2] Design of algorithms for phase measurements by the use of phase stepping. *Applied Optics*, 35, January 1996.

[3] Chiayu Ai and James C. Wyant. Absolute testing of flats by using even and odd functions. *Applied Optics*, 32:4698–4705, 1993.

[4] New York) Anorad Corporation (Hauppauge. *Microglide T300L Motion Platform: Acceptance Test Procedure.* 1998.

[5] A. Bachtold, P. Hadley, T. Nakanishi, and C. Dekker. Logic circuits with carbon nanotube transistors. *Science*, 294, 2001.

[6] Carl G. Chen, Paul T. Konkola, Ralf K. Heilmann, G.S. Pati, and Mark L. Schattenburg. Image metrology and system controls for scanning beam interference lithography. *J. Vac. Sci. Technol. B*, 19, Nov. 2001.

[7] Carl Gang Chen. *Beam Alignment and Image Metrology for Scanning Beam Interference Lithography- Fabricating Gratings with Nanometer Phase Accuracy.* PhD Thesis, Massachusetts Institute of Technology, 2003.

[8] J. Y. Cheng, H.M. Mays, and C. A. Ross. Nanostructure engineering by templated self assembly. *Nature Materials*, 3(11), 2004.

[9] J. Y. Cheng, H.M. Mays, C. A. Ross, H.I. Smith, and G. J. Vansco. Templated self assemby of block copolymers: Effect of substrate topography. *Adv. Materials*, (15), 2003.

[10] C.J.Evans, A. D. Davies, T. Schmitz, and R. Parks. Interferometric figure metrology; enabling in-house traceability. *Proc. SPIE*, 4450, 2001.

[11] Frank C. Demarest. High-resolution, high-speed, low data age uncertainty, heterodyne displacement measuring interferometer electronics. *Meas. Sci. Technol.*, 9, 1998.

[12] C. Dong, C. Zhang, and B. Wang. Reducing the dynamic error of coordinate measuring machines. *Journal of Mechanical Design*, 125, 2003.

[13] E. A. Donley, T. P. Heavner, F. Levi, M. O. Tataw, and S. R. Jefferts. Optical electronics in modern communications. *Rev. Sci. Instrum. 76*, 76, 2005.

[14] Alfred Dubra, Carl Peterson, and Christopher Dainty. Wave-front reconstruction from shear phase maps by use of the discrete fourier transform. *Applied Optics*, 43(5), Feb 2004.

[15] Georg S. Duesberg, Andrew P. Graham, Maik Liebau, Robert Seidel, Eugen Unger, Franz Kreupl, and Wolfgang Hoenlein. Large-scale integration of carbon nanotubes into silicon based microelectronics. *Proceedings of SPIE: Nanotechnology*, 5118, 2003.

[16] Chris J. Evans and Robert N. Kestner. Test optics error removal. *Applied Optics*, 35(7), 1996.

[17] J. Ferrera, J.N. Damask, V.V. Wong, H.I. Smith, and H.A. Haus. *Optical Fiber Communication Technical Digest*, 4, 1994.

[18] Klaus R. Freischlad. Absolute interferometric testing based on reconstruction of rotational shear. *Applied Optics*, 40(10), 2001.

[19] K.R. Freischlad and C.L. Koliopoulous. Modal estimation of a wave front from difference measurements using the discrete fourier transform. *J. Opt. Soc. Am. A*, 3(11), Nov 1986.

[20] Rafael C. Gonzalez and Richard E. Woods. *Digital Image Processing.* Addison-Wesley Publishing Company, 1993.

[21] Joseph W. Goodman. *Introduction to Fourier Optics, 2nd Ed.* McGraw-Hill Series in Electrical and Computer Engineering, 1996.

[22] Peter De Groote. Derivation of algorithms for phase-shifting interferometry using the concept of a data-sampling window. *Applied Optics*, 34, August 1995.

[23] G.Schulz and J. Grzanna. Absolute flatness testing by the rotation method with optimal measuring-error compensation. *Applied Optics*, 31, 1992.

[24] Ralf K. Heilmann, Carl G. Chen, Paul T. Konkola, and Mark L Schattenburg. Dimensional metrology for nanometre-scale science and engineering:towards sub-nanometre accurate encoder. *Nanotechnology*, 15, 2004.

[25] Ralf K. Heilmann, Paul T. Konkola, Carl G. Chen, G.S. Patti, and Mark L Schattenburg. Digital heterodyne interference fringe control system. *J. Vac. Sci. Technol. B*, 19, Nov 2001.

[26] P. Konkola, C. Chen, R.K. Heilmann, C. Joo, J. Montoya, C.-H. Chang, and M.L. Schattenburg. Nanometer-level repeatable metrology using the nanoruler. *J. Vac. Sci. Technol. B*, 21, 2003.

[27] Paul Konkola. *Design and Analysis of a scanning beam interference lithography system for patterning gratings with nanometer-level distortions.* Massachusetts Institute of Technology PhD Thesis, 2003.

[28] Michael H. Lim, T. E. Murphy, J. Ferrera, J.N. Damask, and Henry I. Smith. Fabrication techniques for grating-based optical devices. *J. Vac. Sci. Technol. B.*, (17), 1999.

[29] E. Manske, T. Haussote, R. Mastylo, N. Hoffman, and G. Jagger. Nanopositioning and nanomeasuring machine for accuracy measuring procedures of small features in large areas. *Proceedings of SPIE, Optical Fabrication, Testing, and Metrology II*, 5965, 2005.

[30] J. Montoya, R.K. Heilmann, and M. L. Schattenburg. Measuring two-axis stage mirror nonflatness using linear/angular interferometers. *19th Annual Meeting of the American Society for Precision Engineering. Orlando, Fl.*, October 2004.

[31] Juan C. Montoya, Chih-Hao Chang, Ralf K. Heilmann, and Mark L. Schattenburg. Doppler writing and linewidth control for scanning beam interference lithography. *J. Vac. Sci. Technol. B.*, 23, 2005.

[32] Takashi Nomura, Seiichi Okuda, Kazuhide Kamiya, Hatsuzo Tashiro, and Kazuo Yoshikawa. Improved saunders method for the analysis of lateral shearing interferograms. *J. Opt. Soc. Am. A*, 19(10), Oct 2002.

[33] Alan V. Oppenheim and Ronald W. Schafer. *Discrete-Time Signal Processing.* Prentice Hall, 1998.

[34] Robert E. Parks, Lianzhen Shao, and Chris J. Evans. Pixel-based absolute topography test for three flats. *Applied Optics*, 37(25), 1998.

[35] Michael T. Postek. Nanometer-scale metrology. *Nanostructure Science, Metrology, and Technology*, 4608, 2001.

[36] Lisa A. Poyneer, Donald T. Gavel, and James M. Brase. Fast wave-front reconstruction in large adaptive optics systems with use of the fourier transform. *J. Opt. Soc. Am. A*, 19(10), Oct 2002.

[37] Lisa A. Poyneer and Jean-Pierre Veran. Optimal modal fourier-transform wavefront control. *J. Opt. Soc. Am. A*, 22(8), Aug 2005.

[38] T. Ruijl. *Ultra Precision Coordinate Measuring Machine Design, Calibration and Error Compensation,, PhD Thesis: Eindhoven, the Netherlands.* 2001.

[39] Mark L. Schattenburg and Henry I Smith. The critical role of metrology in nanotechnology. *Proc. SPIE, Workshop on Nanostructure Science, Metrology and Technology (invited)*, 4608, September 2001.

[40] M.L. Schattenburg, C.G. Chen, R.K. Heilmann, P.T. Konkola, and G.S. Pati. Progress towards a general grating patterning technology using phase-locked beams. *Proceedings of the SPIE: Conference on Optical Spectroscopic Techniques and Intrumentation for Atmospheric and Space Research IV*, 2001.

[41] A. Slocum. *Precision Machine Design.* Society of Manufacturing Engineers, 1992.

[42] Amnon Yariv and Pochi Yeh. *Optical Waves in Crystals.* 1984.

[43] Amon Yariv. Optical electronics in modern communications. *Oxford University Press,5th Edition, pp. 477-484*, 1997.

[44] R. Ziemer and W. Tranter. *Principles of Communications: Systems, Modulation, and Noise.* John Wiley and Sons, Inc., 1995.