ERROR ANALYSIS AND DATA REDUCTION FOR INTERFEROMETRIC SURFACE MEASUREMENTS

by

Ping Zhou

A Dissertation Submitted to the Faculty of the

COLLEGE OF OPTICAL SCIENCES

In Partial Fulfillment of the Requirements
For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2009

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Dissertation Committee, we certify that we have read the dissertation prepared by Ping Zhou

entitled Error Analysis and Data Reduction for Interferometric Surface Measurements

and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy

_____ Date: 11/19/2009
James H. Burge

_____ Date: 11/19/2009
James C. Wyant

_____ Date: 11/19/2009
Hubert M. Martin

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

_____ Date: 11/19/2009
Dissertation Director: James H. Burge

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at the University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

SIGNED: Ping Zhou

ACKNOWLEDGEMENTS

I always feel fortunate that I enrolled in the program at the College of Optical Sciences and spent five years working with my advisor Dr. James H. Burge. Thanks to Dr. Burge for teaching me a lot in optics, guiding me in my research and encouraging me to pursue a successful career. I feel very grateful that I have Dr. Burge as part of my education. His enthusiasm in optics and challenges to modern technologies greatly inspire me.

I also would like to thank my committee – Professor Hubert M. Martin and Professor James C. Wyant – for spending hours carefully reviewing my dissertation and providing me with valuable suggestions and helpful advices.

There are many people I want to thank. This dissertation would not have been possible without their help. Thanks to Chunyu Zhao, Bob Parks and Peng Su for discussing many problems I run into in my research and offering me thoughtful insights and help. Thanks to Stacie Manuel, Rene Zehnder and Tom Zobrist for regularly reading my dissertation, giving me helpful comments, as well as encouraging me along the way. Thanks to Matt Novak and many other people at the optical shop, for their generous help on my experiments. Thanks to all my colleagues and friends at the College of Optical Sciences.

Most importantly, I would like to thank my family. My parents are always supporting me to pursue my career. They offer tremendous help whenever I need them. My husband, Yuxiang, is always there to support me, help me, and love me. My son, Jaxon, brings me lots of happiness and joy. His smile means the whole world to me.

To my husband Yuxiang

and

my son Jaxon

TABLE OF CONTENTS

TABLE OF CONTENTS - CONTINUED

TABLE OF CONTENTS - CONTINUED

LIST OF FIGURES

LIST OF FIGURES - CONTINUED

LIST OF FIGURES - CONTINUED

LIST OF FIGURES - CONTINUED

LIST OF FIGURES - CONTINUED

LIST OF FIGURES - CONTINUED

LIST OF FIGURES - CONTINUED

LIST OF TABLES

LIST OF TABLES - CONTINUED

ABSTRACT

High-precision optical systems are generally tested using interferometry, since it often is the only way to achieve the desired measurement precision and accuracy. Interferometers can generally measure a surface to an accuracy of one hundredth of a wave. In order to achieve an accuracy to the next order of magnitude, one thousandth of a wave, each error source in the measurement must be characterized and calibrated.

Errors in interferometric measurements are classified into random errors and systematic errors. An approach to estimate random errors in the measurement is provided, based on the variation in the data. Systematic errors, such as retrace error, imaging distortion, and error due to diffraction effects, are also studied in this dissertation. Methods to estimate the first order geometric error and errors due to diffraction effects are presented.

Interferometer phase modulation transfer function (MTF) is another intrinsic error. The phase MTF of an infrared interferometer is measured with a phase Siemens star, and a Wiener filter is designed to recover the middle spatial frequency information.

Map registration is required when there are two maps tested in different systems and one of these two maps needs to be subtracted from the other. Incorrect mapping causes wavefront errors. A smoothing filter method is presented which can reduce the sensitivity to registration error and improve the overall measurement accuracy.

Interferometric optical testing with computer-generated holograms (CGH) is widely used for measuring aspheric surfaces. The accuracy of the drawn pattern on a hologram decides the accuracy of the measurement. Uncertainties in the CGH

manufacturing process introduce errors in holograms and then the generated wavefront. An optimal design of the CGH is provided which can reduce the sensitivity to fabrication errors and give good diffraction efficiency for both chrome-on-glass and phase etched CGHs.

CHAPTER 1

INTRODUCTION

1.1 Background

High-precision optical systems are generally tested using interferometry, since it often is the only way to achieve the desired measurement precision. Interferometry is the technique that superimposes two wavefronts – the reference and test wavefronts – coherently, which results in an intensity pattern, called an interferogram. The interferogram contains the information of the amplitudes of the two interfering wavefronts, and the phase difference between them. The intensity pattern on the detector is

$$\begin{aligned} I(x,y) &= A_t^2 + A_r^2 + 2A_tA_r\cos\left[\phi(x,y)\right] \\ &= I_t + I_r + 2\sqrt{I_tI_r}\cos\left[\phi(x,y)\right] \end{aligned}, \qquad (\mathbf{1.1})$$

where A_t and A_r are the amplitudes of the test and reference wavefronts and $\phi(x,y)$ is the phase difference between them. Of these three, the phase difference between the two interfering beams $\phi(x,y)$ is of the most interest because it has the surface information of the test part. I_t and I_r are the intensities of the test and reference wavefronts.

In mid 1970s, the first interferograms were obtained with some tilt fringes and recorded on films. A human operator used a graphics tablet to analyze interferogram fringe data. The repeatability of the fringe analysis suffers from the human variability and

randomness. In late 1970s, standard video technology was used to image either photographically recorded interference patterns, or live fringes at the output of an interferometer [**Augustyn 1979; Womack 1979**]. This technology can analyze the fringes electronically by differentiation or level comparison [**Koliopoulos 1981**].

Either graphics tablets or video systems have to record interferograms and then digitize them by finding the fringe centers. This static fringe analysis method limited the measurement precision. To take full advantages of the accuracy available in an interferometric test, the direct phase has to be measured. The development of the solid-state detector and microprocessors made this possible [**Creath 1988**]. A detector array is used to sense fringes, and a phase-shifting device is placed in the reference beam to provide a phase change. A number of measurements with known phase changes ϕ_i are taken to solve the phase difference between the reference and test wavefronts, so the intensity pattern on the detector can be written as

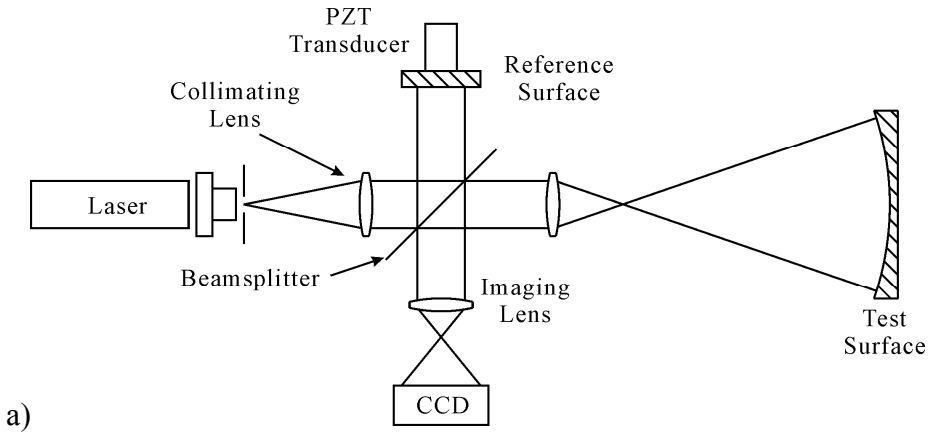
$$I_i(x,y) = I_t + I_r + 2\sqrt{I_t I_r}\cos\left[\phi(x,y) + \phi_i\right]. \qquad (\mathbf{1.2})$$

Because there are three unknowns in the interferometric data, at least three measurements must be made. Typically the phase is changed by 90^o between consecutive intensity measurements. If four measurements are taken with the phase change of $0^o, 90^o, 180^o$ and 270^o, then the phase difference between the two interfering beams $\phi(x,y)$ can be derived as

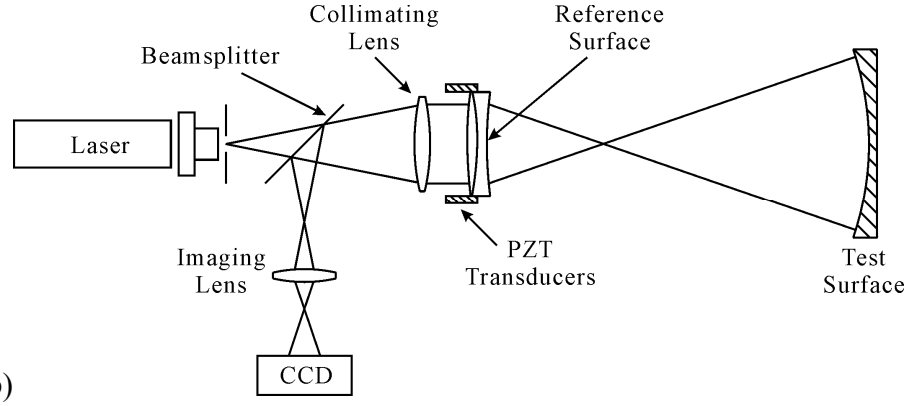
$$\phi(x,y) = \arctan\left[\frac{I_4(x,y) - I_2(x,y)}{I_1(x,y) - I_3(x,y)}\right]. \qquad (\mathbf{1.3})$$

The data from the phase-shifting interferometer (PSI) are more precise, and the tests can be repeatable to a hundredth of a wavelength. Phase-shifting interferometry was introduced for optical testing nearly 40 years ago [**Malacara 1991**]. It has been a very powerful tool that is used in numerous industrial, research and development applications.

The most common method used to introduce the phase shift in a PSI system is to translate one of the mirrors or optical surfaces in the interferometer with a piezoelectric transducer (PZT). The phase shift configurations for the two common interferometer types – the Twyman-Green and the laser Fizeau – are shown in Figure **1.1**. The Twyman-Green interferometer achieves the phase shift by translating one of the mirrors in the reference arm of the interferometer, while the transmission sphere or flat is translated for the phase shift in the Fizeau interferometer. The PSI calculates the wavefront phase at every measurement location from the time-varying intensity measured at that point.



a)



b)

Figure **1.1**: Two common configurations of the phase-shifting interferometry: a) Twyman-Green; b) Laser Fizeau.

The method of using a PZT to introduce the phase shift, so-called temporal phase shifting, is very sensitive to vibration. It is because the various phase shifted frames of interferometric data are taken at different times and vibration causes the phase shifts between the data frames to be different from what is desired. This fundamental difficulty with conventional PSIs brought about research into a newer type of phase-shifting interferometry: Dynamic Phase-Shifting Interferometry (DPSI). In DPSI, all the phase-

shifted data is collected in a single exposure in order to minimize time-varying environmental effects. The DPSI is also known as spatial phase-shifting interferometry.

One natural method to achieve DPSI is to use four CCD cameras to obtain four phase shifted interferograms simultaneously using waveplates. The use of four separate CCD cameras has some limitations because the calibration and the alignment requirements of the four cameras are very critical [**Brock 2005**]. Another approach is to use a holographic element to split the interferogram into four separate beams and then introduce the phase shift to each beam by waveplates. This approach has four interferograms falling on a single CCD camera so it suffers in measurement resolution [**Millerd 2005**]. Spatial carrier technique is another DPSI method. In this method, a series of tilt fringes are purposely introduced to the data with known spatial frequency and orientation. Fourier methods are then used to extract the phase from a single shot interferogram [**Takeda 1987**]. Recently the pixelated phase-mask dynamic interferometer was demonstrated which uses a pixelated phase-mask to impart a discrete phase-shift on each pixel [**Millerd 2004**]. The phase-mask is a micropolarizer array and can work with any type polarization interferometer.

Dynamic phase-shifting interferometry makes it possible to take accurate and low noise measurements. Interferoemters generally measure wavefront deviation from spherical or planar wavefronts. When testing a highly aspheric surface or even free-form surface with an interferometer, a null corrector must be used to generate a wavefront that matches the test wavefront [**Burge 1993**]. Beside null correctors, computer-generated holograms (CGHs) are also often used to generate a desired wavefront. CGHs are

diffractive optical elements synthesized with the aid of computers, and they are capable of producing optical wavefronts with virtually any desired shape. Lohmann and Paris described the first computer-generated holograms [**Lohmann 1967**] and soon a CGH was used to test an aspheric optic [**Wyant 1972; Wyant 1974b**]. These early applications were limited by the accuracy with which computer-generated holograms could be produced using available plotters. Advances in micro-lithographic technologies have made it possible to manufacture holograms with high precision and accuracy. As manufacturing techniques improved, the use of CGHs in optical testing expanded widely; they are now used routinely for testing aspheric and other complex surfaces. Figure **1.2** shows a commonly used interferometer configuration which has a CGH in the test arm.

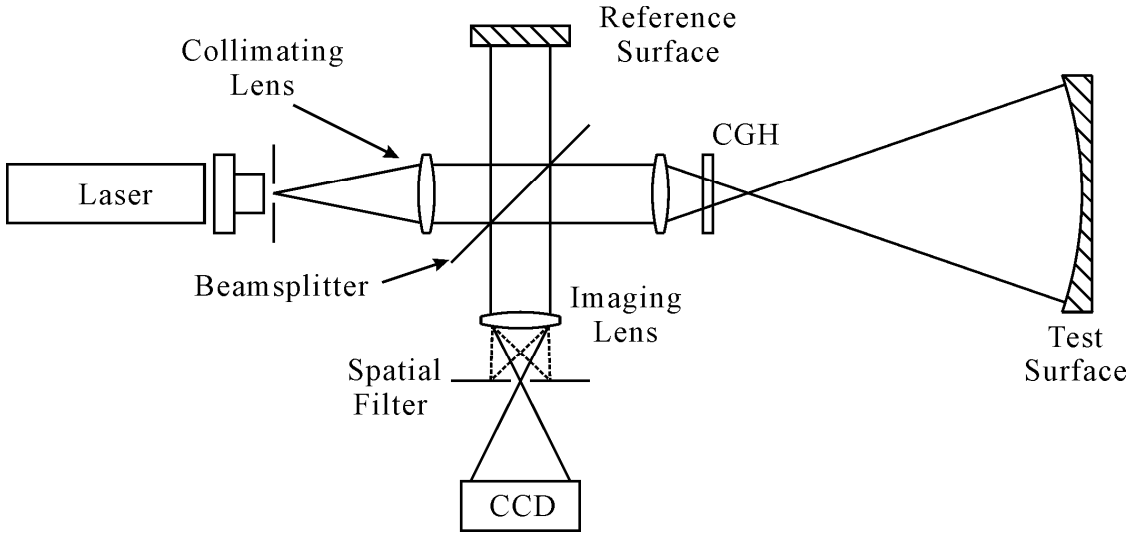


Figure **1.2**: Twyman-Green interferometer with a CGH in the test arm.

1.2 Motivation for current work

The motivation of this dissertation is the need for highly accurate and low noise interferometric measurements. The measurement accuracy of a current commercial phase-shifting interferometer is about λ/100, given that the interferometer is well-calibrated. In order to push the accuracy to the next order of magnitude, λ/1000, each error source must be studied thoroughly and calibrated out. There are many sources of errors in phase-shifting interferometry that can degrade the measurement accuracy, and they are summarized below:

- **Incorrect phase shift:** this can be due to many sources such as incorrect phase shifter calibration, vibration or air turbulence; the measurement error caused by the incorrect phase shift is sinusoidal with a frequency equal to twice the frequency of the interference fringes [**Wyant 1998**].
- **Vibrations:** the PSI is very sensitive to vibration because it uses time-dependent phase shifts; it is hard to correct because the optimum algorithm depends upon the frequency of the vibration present as well as the phase of the vibration relative to the phase shifting [**Brophy 1990; de Groot 1996**]; the DPSI is much less sensitive to vibration.
- **Detector non-linearity:** this refers the nonlinear response of the detector to irradiance; generally the detectors used in the PSI have adequate linear response to irradiance.

- **Quantization errors:** this arises when digitizing the analog detector signal output for the computer; generally 8 bits (256 levels) are used in the digitization, but sometimes more bits can be used to reduce quantization errors [**Brophy 1990**].
- **Source frequency instability and intensity fluctuations:** instable source frequency introduces the phase shift error between adjacent data frames and then causes the measurement errors [**Takeda 1987; Ishii 1987**]; the amplitudes of the test and reference wavefronts are unknowns for the PSI, so intensity fluctuations of the laser source will directly affect the amplitudes of the two interfering beams [**Bruning 1978; Wyant 1975**].
- **Air turbulence:** pockets of warm or cold air moving around within the testing environment can introduce errors in the measurement; errors due to air turbulence can be minimized by averaging many measurements.
- **Stray reflection:** this is a common problem in laser interferometers; the best way of reducing the error due to stray light is to use a short coherence light source [**Wyant 1998; Ai 1988**]. It is common to let the laser pass through a spinning ground glass to destroy the spatial coherence of the laser beam. Unlike most other error sources, error due to stray light is hard to correct and can not be reduced by averaging many measurements.

The error sources listed above have been studied thoroughly, and current commercial phase-shifting interferometers have minimal amounts of these errors. These

errors are not in the scope of this dissertation. The errors addressed in this dissertation are:

- **Random noise:** this can be from some error sources described above, like laser source instability, mechanical vibrations, air turbulence and detector noise, etc.; random noise can be reduced by averaging many data sets.
- **Retrace error:** this appears when the test is not a null test because, strictly speaking, the test and reference wavefronts no longer share a common path when there are fringes in the test [**Gardner 2005; Liu 2009**].
- **Imaging distortion:** this is a non-linear mapping from the test optic to the interferogram and occurs especially when using fast transmission spheres or null lenses [**Murphy 2000**].
- **Diffraction effects:** because wavefront aberrations of the test and reference beams change as they propagate, interferometers suffer errors due to wave propagation; the diffraction effects include phase smoothing and edge diffraction.
- **Inaccurate reference surface:** errors from the reference surface can be obtained by an absolute test; it can often be subtracted from the subsequent tests to improve accuracy.
- **Interferometer modulation transfer function (MTF):** this describes the effect of spatial frequencies on the phase measurement; interferometers have relatively good phase MTF at low spatial frequency and low phase MTF at middle/high spatial frequency; information in the middle/high spatial frequency region will be

modulated by the interferometer phase MTF and become smaller than they actually are.

1.3 Dissertation overview

This dissertation consists of seven chapters. Chapter 1 gives an introduction to phase-shifting interferometry and computer-generated holograms.

Chapter 2 discusses intrinsic errors in interferometric surface measurements from three aspects. First is measurement noise, which can be reduced by averaging many data sets. A method to quantify the noise in a single measurement and the averaged data is provided, which helps determine the number of measurements required to achieve the desired accuracy. Then a review of geometric errors in the interferometry is given, including the retrace error, imaging distortion and shear effect. These three errors can be studied using a geometrical approach. In addition to geometric errors, interferometers also suffer errors due to diffraction. Phase smoothing and edge diffraction are the two main diffraction effects and can explained using Talbot imaging theory.

Interferometer measurements rely on accurate and well-known reference wavefronts. Errors from the reference optic should be characterized and subtracted for accurate measurements. Chapter 3 gives a review of different interferometer calibration methods. We use the random ball test as an example to describe remaining errors due to the limitations of the calibration.

Chapter 4 presents the concept of interferometer phase modulation transfer function and provides a method to calculate it. With knowledge of the interferometer

phase modulation transfer function, we propose a Wiener filter method which can be applied to the interferometric data to recover the middle spatial frequency content in the measurement.

In interferometric surface testing, it is very common to use a reference optic to test another optic. To remove the errors from the reference optic requires correctly registering the two data maps. Chapter 5 presents a smoothing-filter method to reduce the sensitivity to registration errors and improve the overall measurement accuracy.

Computer-generated holograms are often used to test aspheric surfaces with interferometers. Chapter 6 gives a review of computer-generated holograms and introduces a parametric model to calculate the wavefront sensitivity to fabrication errors. Then the CGH fabrication errors are outlined and an example of estimating the wavefront errors due to CGH fabrication is given. In the end, an optimal design of CGH to give good diffraction efficiency and limited sensitivity to manufacturing errors is provided.

This dissertation concludes in Chapter 7 with reviews and discussion of several important findings. Suggestions for future research in this area are also provided.

CHAPTER 2

ERRORS IN INTERFEROMETRIC SURFACE MEASUREMENTS

This chapter discusses errors in interferometric surface measurements. First, an approach to estimate the measurement noise is described. Second, geometric errors, including the retrace error, imaging distortion, and shearing effect, are reviewed. Finally, errors due to diffraction effects are addressed, and discussed extensively in this chapter.

2.1 Noise estimation

Interferometric measurements suffer from noise, or random non-repeatability from one measurement to the next. The reason for this can be laser source instability, mechanical vibrations, air turbulence, detector noise and other sources. Averaging a number of measurements reduces the random noise in the test and improves the measurement accuracy. The variation in the data can be used to estimate the noise in individual maps and the magnitude of the residual error in a map created by averaging multiple measurements.

The noise is quantified using the repeatability of the measurements [**Griesmann 2005**]. Since the true surface figure of the test optic is unknown, an average of many measurements (usually much more than 10) is used instead as an estimate of the surface figure. The difference between a single measurement and the average map provides an estimate of the measurement noise.

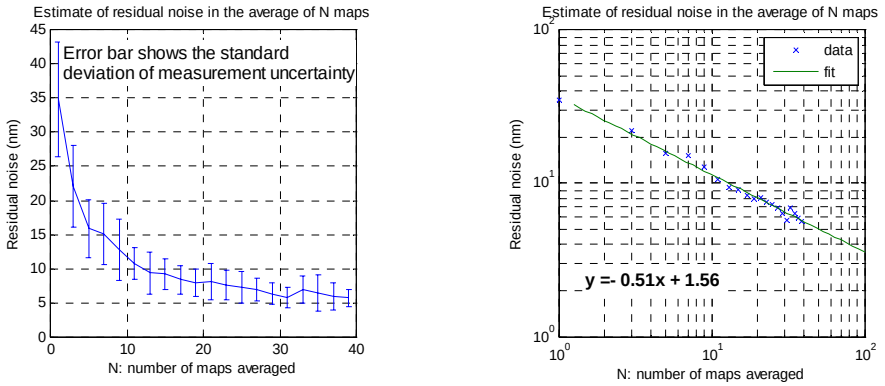
An example of estimating the measurement noise for a reference sphere with a radius of curvature of 16 m is provided here. The radius of curvature of the mirror is so long that the air turbulence causes a large amount of noise. Figure **2.1 (a)** shows a map of a single measurement, with an RMS surface error of 47.7 nm. The astigmatism shown in the picture is mainly due to the turbulence. The average of 70 measurements, $\overline{W}$, shown in Figure **2.1 (b)**, is more rotationally symmetric and the RMS error is 28.9 nm. The average of 70 measurements is certainly less noisy, and the noise in the average map can be estimated using the method described below.

a) b)

Figure **2.1**: Map of a) a single measurement and b) an average of 70 measurements.

From the 70 measurements, N ($N < 70$) maps are chosen at random and averaged. This average map, denoted as $\overline{W}_N$, is subtracted from $\overline{W}$, and the RMS of the average is calculated as RMS($\overline{W_N} - \overline{W}$) . This process is repeated several times, and the mean and the standard deviation of RMS($\overline{W_N} - \overline{W}$) are calculated. This is then repeated for various values of N (from 1 to 39 in this case), and the results are plotted in Figure **2.2 (a)**. The curve follows the mean, and the error bars represent the standard deviation. The mean of RMS($\overline{W_N} - \overline{W}$) becomes significantly smaller as the number of averaged measurements

increases. This plot gives the magnitude of residual noise in the average of N maps. For example, the noise is about 10 nm rms when 15 measurements are averaged. Figure **2.2 (b)** is the same curve plotted in a log-log scale. A straight line with a slope of -0.51 is fitted to the data, which means that the noise in the measurements is random since it is dropping off by $1/\sqrt{N}$. By extrapolating, the noise in the average of 100 maps is reduced to about 3.5 nm.



a) b)

Figure **2.2**: Measurement noise as a function of the number of measurements averaged. a) normal scale; b) log scale showing that the noise is dropping off by $1/\sqrt{N}$.

A plot like Figure **2.2** is very useful, because it can be used to estimate the residual error in an average map and the number of measurements needed to achieve the desired accuracy. As conditions improve, the slope of the log-log plot remains 0.5, but the entire curve shifts down. The y-intercept of this curve represents the quality of the testing environment and the accuracy of a single measurement.

2.2 Geometric effects

Geometric errors are intrinsic to the interferometer. They include retrace error, imaging distortion and shearing effects. All the geometric errors can be studied by geometrical ray-tracing. If the interferometer is completely characterized, the geometric errors can be compensated. Unfortunately, it is difficult to characterize an interferometer without knowing the distances and optics properties inside. The first order geometric errors can be estimated without knowing the details of the interferometer and will be discussed in Chapter 3. This section will provide an overview of retrace error, imaging distortion and shearing effect.

2.2.1 Retrace error

Retrace error is caused by wavefront mismatch between the reference and test beams because, strictly speaking, they no longer share a common path. When the test and reference beams take different paths back through the interferometer, they accumulate different aberrations. The resulting retrace error scales with the test surface curvature, and becomes significant when testing surfaces have small radii of curvature, as shown in Figure **2.3**. In this figure, the test surface is convex, but the retrace error behaves the same for concave surfaces as well.

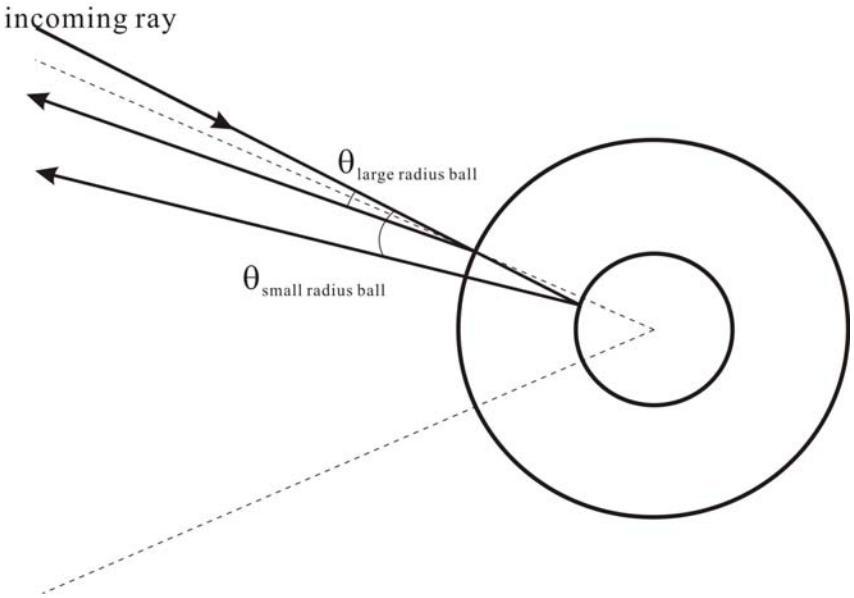


Figure **2.3**: Retrace error depends on the curvature of the test optic. A test optic with smaller radius of curvature has larger retrace error.

Sykora proposed an analytic model to characterize the retrace errors based on interferometer cavity geometries [**Sykora 2008**]. The model assumes the interferometer image is focused precisely at the surface under test and the reference ray is reflected at normal incidence from the reference surface. Sykora derived a formula to approximate the retrace error in the wavefront using small-angle approximations:

$$W_{\text{retrace}} \approx \alpha^2 R_1 \left(\frac{\rho + 1}{\rho} \right). \qquad \textbf{(2.1)}$$

In this equation, α is the interferometer wavefront slope error, and $\rho = -\frac{R_2}{R_1}$, where R_1 is the radius of curvature of the reference surface and R_2 is the radius of the surface under test. When the radius of the test surface is much smaller than that of the reference surface

(small ρ), retrace errors will be large. The retrace error is zero when the interferometer has a perfect null test.

2.2.2 Imaging distortion

Imaging distortion, the non-linear mapping from the test optic to the interferogram, occurs especially when using fast transmission spheres or null lenses. Null lenses are optimized to generate an aspheric wavefront from a spherical one, and not necessarily to provide uniform mapping from the test object to the image plane. The imaging distortion from a null lens can be quite severe [**Martin 2006**], and will cause two problems: one is that the surface defects appear shifted, and the other is that lower order alignment errors appear as higher order wavefront errors, such as third order coma, third order spherical and fifth order spherical. A discussion of interferometric measurement errors introduced by imaging distortion is given by Selberg [**Selberg 1990**].

The measurement error due to imaging distortion depends on the surface deviation from null [**Burge 1993**]. In a null test, the imaging distortion only causes the surface defects to be shifted, and can be corrected by remapping. In a non-null test, which has large phase difference between the reference and test arms, additional wavefront errors will be introduced. In this case, the interferometer no longer works in a common-path configuration. The wavefront error can be predicted and compensated if the interferometer is completely characterized [**Murphy 2000**]. The imaging distortion is

described by mapping functions, which can be determined by putting fiducial marks at known positions on the surface under test. Mapping functions are discussed in Chapter 5.

2.2.3 Shearing effect

When using a Fizeau interferometer to measure a surface, it is always important to have the transmission sphere or flat aligned perpendicular to the optical axis of the interferometer. The outgoing beam from the interferometer is collimated before the transmission sphere. When the transmission sphere is not correctly aligned, the outgoing collimated beam will hit the transmission sphere at an angle, causing some off-axis aberration like coma. It also causes shearing between the test and reference wavefronts, which is an effect that is frequently overlooked.

Many interferometers have an alignment mode that switches to the interferometer focus plane and allows adjustment of the transmission sphere by overlapping reference spots onto the center of a crosshair. For most applications, this will allow adequate alignment of the transmission sphere. However, when the interferometer wavefront has high slope errors or the test optic has a short radius of curvature, the measurement errors will be very sensitive to misalignment of the transmission sphere because of shearing effect as demonstrated in Figure **2.4**.

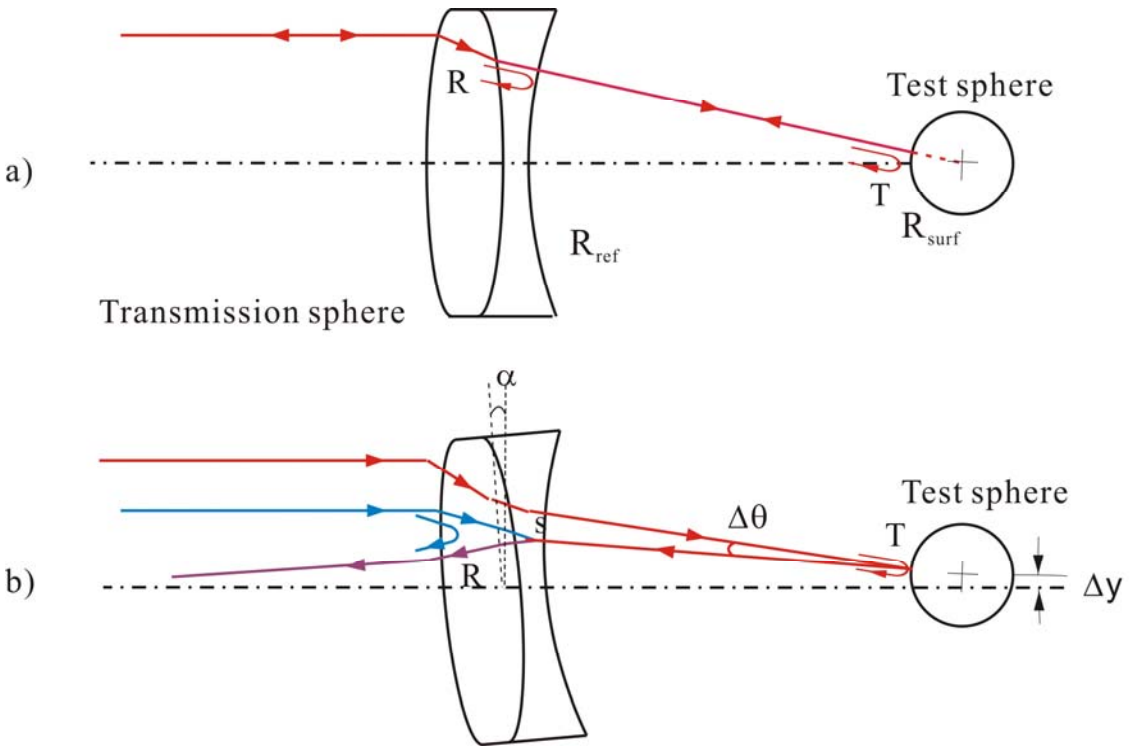


Figure **2.4**: a) The transmission sphere is correctly aligned to the optical axis of the interferometer; b) The transmission sphere is tilted, which causes a shear in the test. Letters R and T represent reference beam and test beam, respectively. Letter s is the shear displacement.

Figure **2.4 (a)** shows the configuration when the transmission sphere is correctly aligned to the optical axis of the interferometer. The light strikes the reference surface and the test sphere at normal incidence, and retro-reflects back, following a common path. In Figure **2.4 (b)**, the transmission sphere is tilted, so the light is not at normal incidence on the reference surface and test sphere. To compensate, the test sphere is shifted to minimize the number of the fringes. As a result, the "red" light becomes a test ray, the "blue" light becomes a reference ray, and they interfere to provide the "purple" light in the figure. This first order shearing effect is because the reference ray interferes with the test ray from a different field point. When the transmission sphere is tilted by an

angle α (in radians), the reference beam will be deviated by 2α (ignoring the light deviation due to the other surfaces of the transmission sphere). To align the test wavefront to the reference wavefront, the test sphere has to be translated by

$$\Delta y = \alpha \cdot R_{\text{ref}}, \tag{2.2}$$

where R_{ref} is the radius of curvature of the reference surface (the last surface of the transmission sphere). It is intuitive to think that the test sphere has to be translated so that it is concentric with the reference surface because that can minimize the optical path difference in the cavity. The angle deviation between the incident beam and return beam at the test sphere is

$$\Delta\theta = \frac{\Delta y}{R_{\text{surf}}/2}, \tag{2.3}$$

where R_{surf} is the radius of curvature of the test sphere. Therefore, the shear at the reference surface, *s*, is

$$s \approx \Delta\theta \cdot \left(R_{\text{ref}} - R_{\text{surf}}\right) = \frac{2\alpha R_{\text{ref}} \cdot \left(R_{\text{ref}} - R_{\text{surf}}\right)}{R_{\text{surf}}}. \tag{2.4}$$

The shear effect introduces the wavefront error ΔW_{shear}, which is approximately the product of the wavefront slope of the illumination wavefront $\bar{\nabla} W(x, y)$ and the shear displacement

$$\Delta W_{\text{shear}}(x, y) = s \cdot \bar{\nabla} W(x, y). \tag{2.5}$$

When the illumination wavefront from the interferometer has large slope errors, the errors caused by shear will become important.

Equation **2.4** shows that the amount of shear depends on the tilt angle of the transmission sphere and the radius of curvature of the test surface. Larger tilt angles and smaller radius of curvature of the test surface will introduce more shear. To accurately align the transmission sphere, it is necessary to finish the alignment in the "interferometric mode" of the interferometer rather than just the "alignment mode". This can be done by placing an uncoated concave mirror at the focus of the converging wavefront (the cat's eye position), and then removing the tilt from the resulting interference pattern by tilting the transmission sphere. For a 6" F/3.2 transmission sphere, if there are 10 tilt fringes (λ = 632.8 nm) across the aperture at the cat's eye position, it means that the transmission sphere is tilted by 41.5 μrad. If the test sphere has a radius of curvature of 12.5 mm, then the shear on the reference surface is 1.5mm. Assuming that the detector has 600 pixels along the shear direction, there would be six-pixel shear between the test beam and reference beam.

2.3 Diffraction effects

In addition to geometric errors, interferometers also suffer errors due to diffraction. The wavefront aberrations of the test and reference beams change as they propagate. Therefore, there will be errors due to diffraction effects for interferometric measurements. Fresnel diffraction and Talbot imaging theories, which are reviewed in Section **2.3.1**, are used here to analyze the diffraction effects in interferometry. Sections **2.3.2** and **2.3.3** discuss how diffraction effects influence interferometric measurements by

smoothing actual phase irregularities and introducing phase errors due to edge diffraction. Section **2.4** illustrates how the Zernike phase terms, which are often used in optical testing, change as they propagate.

2.3.1 Analysis of Fresnel diffraction using the Talbot effect

Diffraction theory is the study of the effects of light propagation that cannot be predicted by geometrical models. There are two diffraction theories: Fresnel diffraction applies when the observation plane is relatively close to an aperture, and it also works in the far field region; Fraunhofer diffraction only works in the far field. The angular spectrum and the free-space transfer function are used to calculate complex amplitude distribution after an aperture [**Barrett 2003**]. The Talbot effect is a direct application of Fresnel diffraction, which can be used to analyze the behavior of periodic wavefronts as they propagate [**Goodman 2004**]. The use of the Talbot effect can greatly simplify the complex diffraction problem. Effective propagation distance is a useful parameter, which allows a simple calculation of the Talbot effect for converging or diverging light by transforming them to an equivalent collimated space. The Fresnel number is another parameter often used in diffraction analysis, which can simplify the geometric propagation to a single parameter. All these concepts are important for analysis of the diffraction effects. The following sections provide definitions and analysis of all these concepts.

2.3.1.1 Angular spectrum and wavefront propagation

If a wavefront with complex amplitude is known at some plane, then the wavefront at some arbitrary observation plane can be predicted by Fresnel diffraction. The concepts of angular spectrum and the free-space transfer function are often used to calculate the complex amplitude at any observation plane [**Goodman 2004**].

For a complex amplitude $u(x,y)$ at the plane $z=0$, the Fourier Transform of $u(x,y)$ is

$$P_{z=0}(\xi,\eta)=\iint u(x,y)\exp\left[-i2\pi(\xi x+\eta y)\right]dx\,dy, \qquad (\mathbf{2.6})$$

where ξ and η are the spatial frequencies along the x and y directions, respectively. Consequently, the complex amplitude $u(x,y)$ can be written as the inverse Fourier Transform of $P_{z=0}(\xi,\eta)$:

$$u(x,y)=\iint P_{z=0}(\xi,\eta)\exp\left[i2\pi(\xi x+\eta y)\right]d\xi\,d\eta. \qquad (\mathbf{2.7})$$

The complex exponential function $\exp[i2\pi(\xi x+\eta y)]$ may be regarded as representing a plane wave propagating with direction cosines

$$\alpha=\lambda\xi \quad \beta=\lambda\eta \quad \gamma=\sqrt{1-(\lambda\xi)^2-(\lambda\eta)^2}, \qquad (\mathbf{2.8})$$

so the kernel of the Fourier Transform may be regarded as a unit-amplitude plane wave propagating with direction cosines (α,β,γ). As a result, $u(x,y)$ may be regarded as a linear superposition of plane waves, each of which is traveling in a direction (α,β,γ) and weighted by $P_{z=0}(\xi,\eta)$. For this reason, the function

$$P_{z=0}\left(\frac{\alpha}{\lambda},\frac{\beta}{\lambda}\right)=\iint u(x,y)\exp\left[-i2\pi\left(\frac{\alpha}{\lambda}x+\frac{\beta}{\lambda}y\right)\right]dx\,dy \qquad (\mathbf{2.9})$$

is called the *angular spectrum* of the complex amplitude $u(x,y)$.

The angular spectrum $P_z(\xi,\eta)$ at the plane z can be calculated by multiplying $P_{z=0}(\xi,\eta)$ by the free-space transfer function $H_z(\xi,\eta)$, which is associated with Fresnel diffraction and described as

$$H_z(\xi,\eta)=\exp\left(i2\pi z\sqrt{\frac{1}{\lambda^2}-\xi^2-\eta^2}\right). \qquad (\mathbf{2.10})$$

The procedure to find the complex amplitude $u_z(x,y)$ at the plane z is:

1) Find the Fourier Transform of $u_{z=0}(x,y)$ with respect to the frequency variables ξ and η, which is the angular spectrum $P_{z=0}(\xi,\eta)$ at the plane $z=0$.

$$P_{z=0}(\xi,\eta)=\Im\left(u_{z=0}(x,y)\right) \qquad (\mathbf{2.11})$$

2) Multiply $P_{z=0}(\xi,\eta)$ by the transfer function of free space, $H_z(\xi,\eta)$ to get the angular spectrum $P_z(\xi,\eta)$ at the plane z.

$$P_z(\xi,\eta)=P_{z=0}(\xi,\eta)\cdot H_z(\xi,\eta) \qquad (\mathbf{2.12})$$

3) Find the inverse Fourier Transform of $P_z(\xi,\eta)$ with respect to the spatial variables x and y. This result is the complex amplitude on the observation plane at distance z.

$$u_z(x,y)=\Im^{-1}\left(P_z(\xi,\eta)\right) \qquad (\mathbf{2.13})$$

2.3.1.2 Talbot effect

The Talbot effect is a diffraction phenomenon that occurs for any field with a periodic complex amplitude. If a wavefront with sinusoidal complex amplitude of period p is illuminated by collimated light, then the same complex amplitude is formed by free space diffraction at integer multiples of the Talbot distance $z_T = \frac{2p^2}{\lambda}$.

The Talbot effect can be applied to wavefronts with sinusoidal phase ripples, as long as they are small in magnitude compared to the wavelength. This can be proven using the method of angular spectrum analysis [**Goodman 2004**]. Consider a 1-D sinusoidal wavefront phase with a period of p and magnitude of $m << 1$

$$u_{z=0}(x) = e^{im\cos(\frac{2\pi x}{p})} = 1 + im\cos\left(\frac{2\pi x}{p}\right) - \frac{m^2}{2}\cos^2\left(\frac{2\pi x}{p}\right) + \ldots, \qquad \textbf{(2.14)}$$

where $m = 2\pi W$ in radians and W is the magnitude of the wavefront in waves. When $m << 1$ (i.e., $W << 1/2\pi$ waves), the third term in the Taylor expansion is small and can be ignored. If only keeping the first two terms in Equation **2.14** to describe the phase ripple, the Fourier Transform of this complex amplitude is

$$\begin{aligned} P_{z=0}(\xi) &= \int_{\infty}\left[1 + im\cos\left(\frac{2\pi x}{p}\right)\right]\cdot e^{-i2\pi\xi x}dx \\ &= \delta(\xi) + \frac{1}{2}im\left[\delta(\xi - \xi_0) + \delta(\xi + \xi_0)\right], \end{aligned} \qquad \textbf{(2.15)}$$

where $\xi_0 = 1/p$. The diffraction pattern after propagating a distance z can be calculated by multiplying its angular spectrum $P_{z=0}(\xi)$ by the free-space transfer function $H_z(\xi)$

and then taking an inverse Fourier Transform. If the period of the wavefront is much larger than λ, the free-space transfer function can be written as

$$H_z(\xi) = e^{i\frac{2\pi z}{\lambda}\sqrt{1-(\lambda\xi)^2}} \approx e^{i\frac{2\pi z}{\lambda}} e^{-i\pi z\lambda\xi^2}. \tag{2.16}$$

Applying the free-space transfer function to $P_{z=0}(\xi)$ gives

$$\begin{aligned} P_z(\xi) &= e^{i\frac{2\pi}{\lambda}z}\left\{\delta(\xi) + \frac{1}{2}ime^{-i\pi z\lambda\xi_0^2}\left[\delta(\xi-\xi_0)+\delta(\xi+\xi_0)\right]\right\} \\ &= e^{i\frac{2\pi}{\lambda}z}\left\{\delta(\xi) + \frac{1}{2}im\left[\cos(\pi z\lambda\xi_0^2) - i\sin(\pi z\lambda\xi_0^2)\right]\left[\delta(\xi-\xi_0)+\delta(\xi+\xi_0)\right]\right\}. \end{aligned} \tag{2.17}$$

Dropping the constant phase term $e^{i\frac{2\pi}{\lambda}z}$, the complex amplitude after propagating a distance z can be given by taking the inverse Fourier Transform of $P_z(\xi)$:

$$\begin{aligned} u_z(x) &= 1 + im\cos\left(\frac{2\pi x}{p}\right)\left[\cos(\pi z\lambda\xi_0^2) - i\sin(\pi z\lambda\xi_0^2)\right] \\ &= 1 + m\sin(\pi z\lambda\xi_0^2)\cos\left(\frac{2\pi x}{p}\right) + im\cos(\pi z\lambda\xi_0^2)\cos\left(\frac{2\pi x}{p}\right). \end{aligned} \tag{2.18}$$

Since the Talbot distance is $z_T = \frac{2}{\lambda\xi_0^2} = \frac{2p^2}{\lambda}$, Equation **2.18** can be written as

$$u_z(x) = 1 + m\sin\left(2\pi\frac{z}{z_T}\right)\cos\left(\frac{2\pi x}{p}\right) + im\cos\left(2\pi\frac{z}{z_T}\right)\cos\left(\frac{2\pi x}{p}\right). \tag{2.19}$$

This complex amplitude can be written in the form of amplitude and phase, $u_z(x) = A_z(x)e^{i\psi_z(x)}$. The phase of the complex amplitude is the arctangent of the ratio of the imaginary part to the real part

$$\psi_z(x) = \arctan\left[\frac{m\cos\left(2\pi\frac{z}{z_T}\right)\cos\left(\frac{2\pi x}{p}\right)}{1+m\sin\left(2\pi\frac{z}{z_T}\right)\cos\left(\frac{2\pi x}{p}\right)}\right], \quad \textbf{(2.20)}$$

and the amplitude is

$$A_z(x) = \sqrt{1+2m\sin\left(2\pi\frac{z}{z_T}\right)\cos\left(\frac{2\pi x}{p}\right)+m^2\cos^2\left(\frac{2\pi x}{p}\right)}. \quad \textbf{(2.21)}$$

When the wavefront propagates to a distance of $z = \frac{nz_T}{2}$ (n is an integer), Equation **2.19** becomes

$$u_z(x) = 1 \pm im\cos\left(\frac{2\pi x}{p}\right) \approx e^{\pm im\cos\left(\frac{2\pi x}{p}\right)}. \quad \textbf{(2.22)}$$

The diffraction pattern has constant amplitude and pure phase variation. At a distance $z = \frac{(2n+1)z_T}{4}$, Equation **2.19** reduces to

$$u_z(x) = 1 \pm m\cos\left(\frac{2\pi x}{p}\right), \quad \textbf{(2.23)}$$

so the diffraction pattern has pure amplitude variation. As a sinusoidal phase pattern propagates, it will cycle through a reverse contrast amplitude pattern, a conjugate phase pattern, a pure amplitude pattern, and then back to the original phase pattern.

With the small-angle approximation and the fact that the magnitude of phase ripple $m << 1$, Equation **2.20** reduces to

$$\psi \approx \tan(\psi) \approx m\cos\left(2\pi\frac{z}{z_T}\right)\cos\left(\frac{2\pi x}{p}\right). \quad \textbf{(2.24)}$$

The magnitude of the phase ripple is approximately modulated as a cosine function [**Zhao 2007a**]. If the magnitude of the phase ripple is W in waves, then the attenuation of the phase ripple as it propagates a distance z can be described as

$$W' = W\cos\left(2\pi\frac{z}{z_T}\right) = W\cos\left(\frac{\pi z\lambda}{p^2}\right). \qquad (\mathbf{2.25})$$

The pure phase ripples will also create amplitude ripples accordingly, which have sinusoidal distribution. If the amplitude of the original field is unity and the higher order variation in the amplitude is ignored in Equation **2.21**, then the amplitude distribution as it propagates is

$$A' = 1 + m\sin\left(2\pi\frac{z}{z_T}\right)\cdot\cos\left(\frac{2\pi x}{p}\right) = 1 + 2\pi W\sin\left(\frac{\pi z\lambda}{p^2}\right)\cdot\cos\left(\frac{2\pi x}{p}\right). \qquad (\mathbf{2.26})$$

This effect was derived from the Fresnel transfer function for the case when the complex amplitude contains only a single sinusoidal variation. Note that this does not include edge effects that come from an aperture. For a phase ripple with the magnitude of ±0.05λ, the phase and amplitude distributions as it propagates the Talbot distance two times are shown in Figure **2.5**.

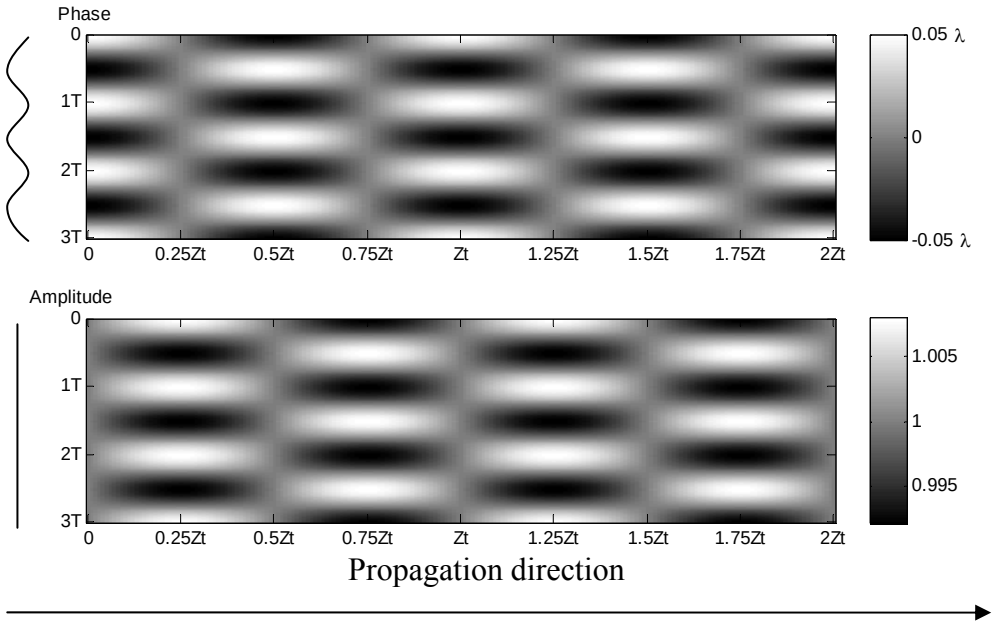


Figure **2.5**: A wavefront with a periodic structure changes its phase and amplitude as it propagates. This is called the Talbot effect. In the figure, the sinusoidal wavefront propagates from left to right.

The physical insight of the Talbot effect is the three-beam interference. The form of Equation **2.15** is that of three plane waves, as shown in Figure **2.6**. One plane wave is propagating along the *z* axis, and the other two are off axis and symmetric about the *z* axis. The resulting complex field is the addition of these three plane waves. This three-beam interference causes the amplitude and phase to vary periodically in space.

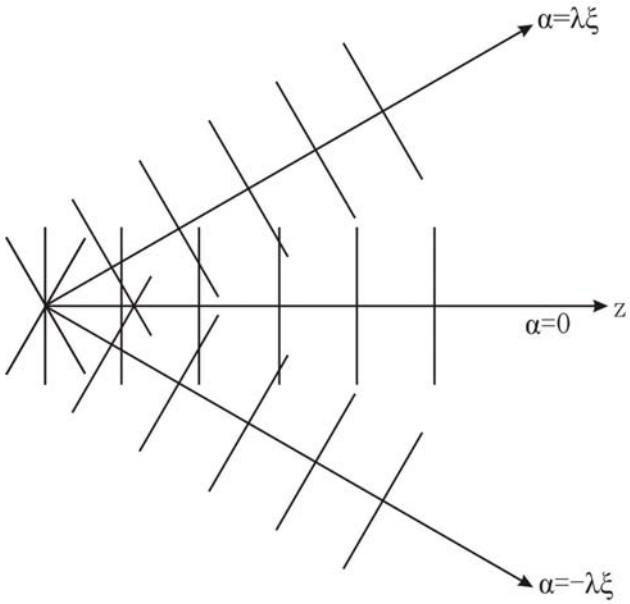


Figure **2.6**: The Talbot effect illustrated as three-beam interference. The angular spectrum of a laser beam transmitted through a weak phase grating consists of three plane waves, which combine together to form the periodic replication of the original phase distribution along the z axis.

The diffraction effects are calculated for sinusoidal ripples of arbitrary spatial frequency. In general, any phase distribution can be decomposed to a collection of sinusoidal ripples with different spatial frequencies.

The equations derived above assume that the magnitude of the phase ripple is small ($m << 1$). When m is close to 1 or even larger than 1, the third and higher order terms in Equation **2.14** can not be neglected, and those terms will generate harmonics as they propagate.

2.3.1.3 Effective propagation distance

The Talbot distance z_T is defined for a collimated beam. For a diverging or converging beam, it is convenient to convert it into an equivalent collimated beam and then use the Talbot effect to predict how wavefronts change due to propagation. The diffraction pattern for a non-collimated wavefront is the same as that for a collimated beam, except that the diffraction pattern occurs at the *effective propagation distance* L_e, and it is scaled in the transverse dimension.

As shown in Figure **2.7**, a converging wavefront starting with radius of curvature R_1, diameter $2a_1$, and ripples with period p_1, propagates to a position where it has radius of curvature R_2, diameter $2a_1\frac{R_2}{R_1}$, and ripples with period $p_1\frac{R_2}{R_1}$. To convert this propagation to equivalent propagation in a collimated space, a lens with an arbitrary focal length f is used to convert the light into collimated light. Then we use geometric imaging relationships to calculate equivalent dimensions.

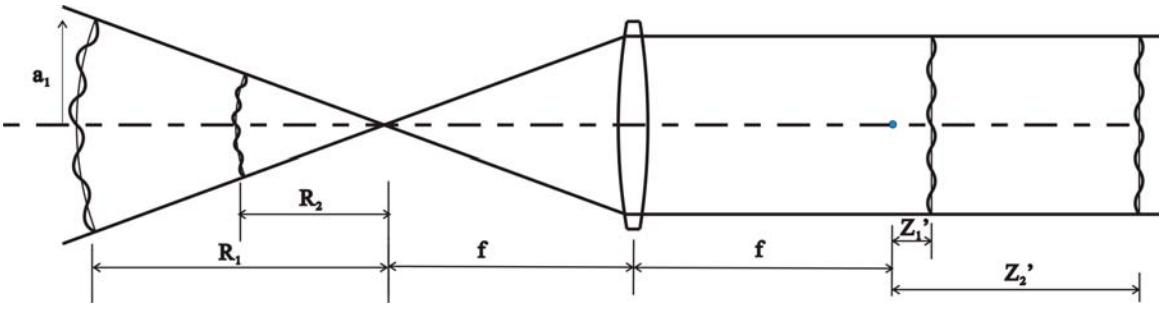


Figure **2.7**: Propagation in a converging space is converted to equivalent propagation in a collimated space when the two wavefronts have the same sign of curvature.

In the collimated space, the equivalent propagation distance is

$$\Delta Z' = Z_2' - Z_1' = f^2(\frac{1}{R_2} - \frac{1}{R_1}), \quad \textbf{(2.27)}$$

and the phase ripple has a period of

$$p' = \frac{f}{R_1} p_1 . \quad \textbf{(2.28)}$$

The attenuation of the phase ripple in the equivalent collimated space becomes

$$W' = W \cdot \cos\left(\frac{\pi\lambda \cdot \Delta Z'}{p'^2}\right) = W \cdot \cos\left(\frac{\pi\lambda R_1 \cdot (R_1 - R_2)}{R_2 \cdot p_1^2}\right). \quad \textbf{(2.29)}$$

Comparing Equation **2.29** with Equation **2.25** shows that $\frac{R_1(R_1 - R_2)}{R_2}$ corresponds to z. Therefore, the effective propagation distance L_e can be defined as

$$L_e = \frac{R_1(R_1 - R_2)}{R_2}, \quad \textbf{(2.30)}$$

and Equation **2.29** can be written as

$$W' = W \cdot \cos\left(\frac{\pi\lambda \cdot L_e}{p_1^2}\right). \quad \textbf{(2.31)}$$

The effective propagation distance can be proven to be the same for spherical wavefronts with different signs of curvature, which is provided in Appendix A.

2.3.1.4 Fresnel number

The period of the phase ripple p_1 varies as it propagates in a converging or diverging beam. To avoid the scaling issue, the ripple period can be normalized by $2a_1$, the diameter of the aperture. Thus, the normalized frequency

$f_{\text{normalized}} = 2a_1 / p_1$ [cycles/diameter], remains unchanged as the wavefront propagates in any type of illumination. By replacing the period p_1 with the normalized frequency $f_{\text{normalized}}$, Equation **2.31** becomes

$$W' = W\cos\left(\frac{\pi\lambda L_e f_{\text{normalized}}^2}{4a_1^2}\right). \qquad (\mathbf{2.32})$$

Notice that the Fresnel number N_f is

$$N_f = \frac{a_1^2}{\lambda L_e}, \qquad (\mathbf{2.33})$$

so Equation **2.32** can be further simplified to

$$W' = W\cos\left(\frac{\pi f_{normalized}^2}{4N_f}\right). \qquad (\mathbf{2.34})$$

Physically, the Fresnel number represents the geometric difference between the distance from the edge of the aperture to the observation point, and the distance from the center of the aperture to the observation point, divided by λ/2. Equation **2.34** shows that the smoothing effect of the periodic structure only depends on the Fresnel number, which is a function of effective propagation distance, aperture size and wavelength. The larger the Fresnel number, the less the smoothing effect is. For any given geometry, the Fresnel number can be calculated, allowing the change of the wavefront magnitude due to propagation to be estimated at any given frequency. All the diffraction effects can be analyzed with Equation **2.34**. A transfer function describing the magnitude of the phase smoothing can be defined as

$$TF = \frac{W'}{W} = \cos\left(\frac{\pi f_{normalized}^2}{4N_f}\right). \qquad (\mathbf{2.35})$$

Another useful form of the transfer function is to describe it in terms of F-number F_n, which is $R_1/2a_1$ in this case. Rewrite Equation **2.29** with F_n and $f_{\text{normalized}}$ and we have

$$TF = \cos\left(\frac{\pi\lambda\cdot\Delta Z'}{p'^2}\right) = \cos\left(\frac{\pi\lambda R_1^2}{p_1^2}\left(\frac{1}{R_2}-\frac{1}{R_1}\right)\right) = \cos\left(\pi\lambda F_n^2 f_{\text{normalized}}^2\left(\frac{1}{R_2}-\frac{1}{R_1}\right)\right). \quad \textbf{(2.36)}$$

2.3.2 Diffraction effect: phase smoothing

The diffraction effects for interferometric surface measurements include phase smoothing and edge diffraction. In this section, the phase smoothing due to diffraction is discussed. First, a physical explanation of errors due to phase smoothing for the test wavefront, reference wavefront and common wavefront is given. Then a computer simulation is performed to verify this explanation. In the end, examples of the Fresnel number calculation and the phase smoothing effect are provided.

2.3.2.1 Physical explanation

Phase smoothing comes from two diffraction effects, which can both be analyzed using the Talbot effect. One effect is due to wavefront propagation, where the diffraction effects are different for different propagation distances. The other is due to imaging, where the diffraction errors will be generated if the wavefront is not in focus. In the following sections, the phase smoothing effect, when measuring a surface with an

interferometer, will be discussed in three aspects. They are diffraction effects in the test wavefront, the reference wavefront and the common wavefront.

Diffraction effect in the test wavefront

Errors in the test wavefront are caused by the null optics (if they exist) and the test surface. A simple case is to consider only errors from the test surface itself. Interferometers usually focus the test surface onto the detector to mitigate diffraction effect. Improper imaging causes the phase ripples on the test surface to distort, and causes diffraction "ripples" around the edge of the mirror. This edge diffraction will be discussed in detail in Section **2.3.3**. As long as the test surface is focused onto the detector, the diffraction effect from the test wavefront is avoided. Note that imaging aberrations violate this assumption and lead to errors in the test, which can be large when using a null corrector [**Zhao 2007a**].

Diffraction effect in the reference wavefront

Errors in the reference wavefront are caused by imperfections from the optics in the reference arm. A simple case is to assume that the reference surface figure is the sole source of errors in the reference wavefront.

Interferometers measure the wavefront difference between the test and reference wavefronts. Any errors in the reference wavefront will appear as errors in the test surface unless the interferometer is calibrated. The reference wavefront also suffers from diffraction effects since the reference surface is not in focus. However, errors from the reference wavefront, including diffraction effects, can be calibrated with an absolute test.

For Twyman-Green interferometers, errors due to diffraction in the reference surface only affect the reference beam, and have no impact on the test beam. Fizeau interferometers are a special case, in which the reference surface acts as a beam splitter. Errors in the beam splitter affect both the test and reference wavefronts, so for Fizeau interferometers, errors in the reference surface will also affect the test wavefront. The diffraction effect of the reference surface error on the test wavefront cannot be calibrated out with an absolute test, and it can be treated as the common wavefront errors discussed next.

There are many different methods discussed in the literature for calibrating and backing out the reference surface errors from the surface measurement. Some of these calibration methods will be reviewed in Chapter 3. Because the diffraction effects depend on test geometry, the calibration will only be accurate if the calibration configuration is the same as the test configuration.

Diffraction effect in the common wavefront

The common wavefront refers to the wavefront from the illumination optics in an interferometer. Reference and test beams carry the same common wavefront information right before they are split from each other. Figure **2.8** shows a Fizeau interferometer testing a plane mirror. The common wavefront keeps propagating a distance L forward to the test optic, so that the total round-trip propagation distance difference between the test and reference arms is $2L$. Because the diffraction effect varies with the propagation distance, the propagation distance between the two beams will therefore introduce errors to the final wavefront map.

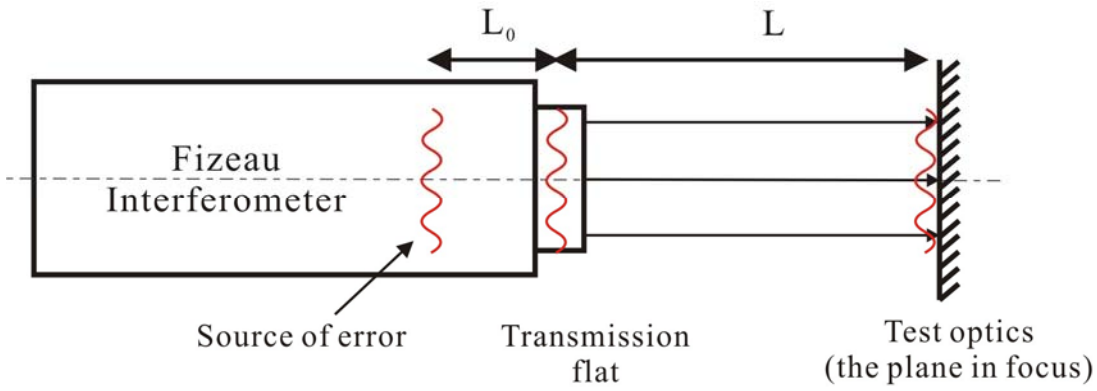


Figure **2.8**: A plane mirror, a distance *L* away from the transmission flat, is tested with a Fizeau interferometer.

Note that precise knowledge of the interferometer configuration and properties of the illumination optics are required to calculate the diffraction effects of the common wavefront in both arms. This is because the diffraction effect varies as a cosine function instead of linearly as shown in Equation **2.25**. To find out the diffraction errors from the common wavefront ΔW_{comm}, one should know where the source of error is and then calculate the wavefronts as it propagates to the test surface and the reference surface, respectively. Figure **2.9** shows how a phase ripple with a certain spatial frequency changes with the propagation distance. The two circles in the plot represent the magnitudes of phase ripple in the reference and test arms. The difference in the wavefront magnitudes ΔW_{comm} are the diffraction errors from the common wavefront.

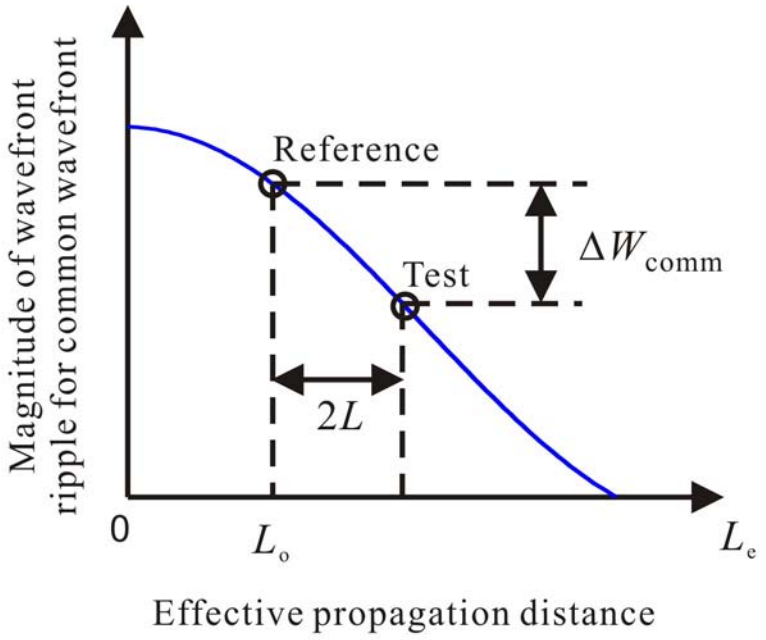


Figure **2.9**: The magnitude of wavefront ripple changes with the propagation distance, following a cosine distribution. The two circles in the plot represent the magnitudes of phase ripple in the reference and test arms. The difference in the wavefront magnitudes ΔW_{comm} are the diffraction errors from the common wavefront.

2.3.2.2 Computer simulation

The diffraction effect in interferometric surface measurements can be analyzed using the Talbot effect. A simulation, based on wave propagation, using the software ASAP (Breault Research Organization, Tucson, AZ), verifies this. Figure **2.10** shows a simple ASAP model of a Twyman-Green interferometer, with a phase plate placed in the common path. The distances between different parts are specified in the figure, and two imaging lenses are used to image the test mirror on the detector perfectly. The reference mirror is out of focus by 5 mm.

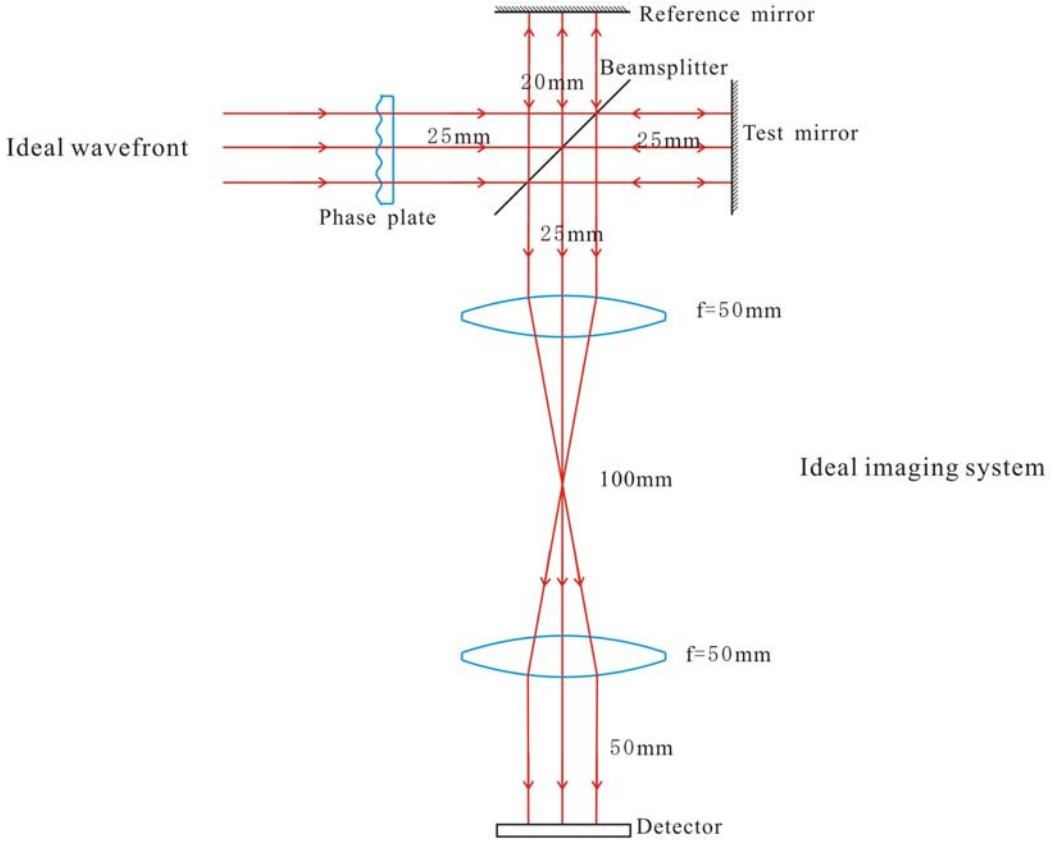


Figure **2.10**: A Twymann-Green interferometer, with a phase ripple in the common wavefront, is simulated in ASAP. An ideal imaging system is assumed.

The phase plate is used to generate a sinusoidal phase ripple with a magnitude of ± 0.05λ, and a period of 0.1 mm. The clear aperture of the beam from the interferometer is 2 mm in diameter, and the wavelength is 500 nm. The Talbot distance for this phase ripple is $z_T = \frac{2p^2}{\lambda} = \frac{2 \cdot 0.1^2}{0.5 \cdot 10^{-3}} = 40\text{mm}$. For the three separate simulations, this phase plate is placed at the test mirror, at the reference mirror and in the common path, respectively, while the ASAP software is used to perform the wave propagation analysis. The results of the three simulations are shown in Figure **2.11**, Figure **2.12** and Figure **2.13** respectively. There are three plots in each figure, and they show the reference wavefront,

the test wavefront and the measured wavefront. The measured wavefront on the detector equals the difference of the test and reference wavefronts on the detector. The simulation is compared with theoretical calculation using Equation **2.25**.

With the phase ripple on the test mirror and no errors from the reference mirror and the common wavefront, the phase ripple on the test mirror should be correctly measured since it is focused by the interferometer. In the ASAP simulation, there is a slight difference between the wavefront on the detector and at the test mirror (Figure **2.11**) due to simulation errors. In general, the ASAP simulation matches our prediction.

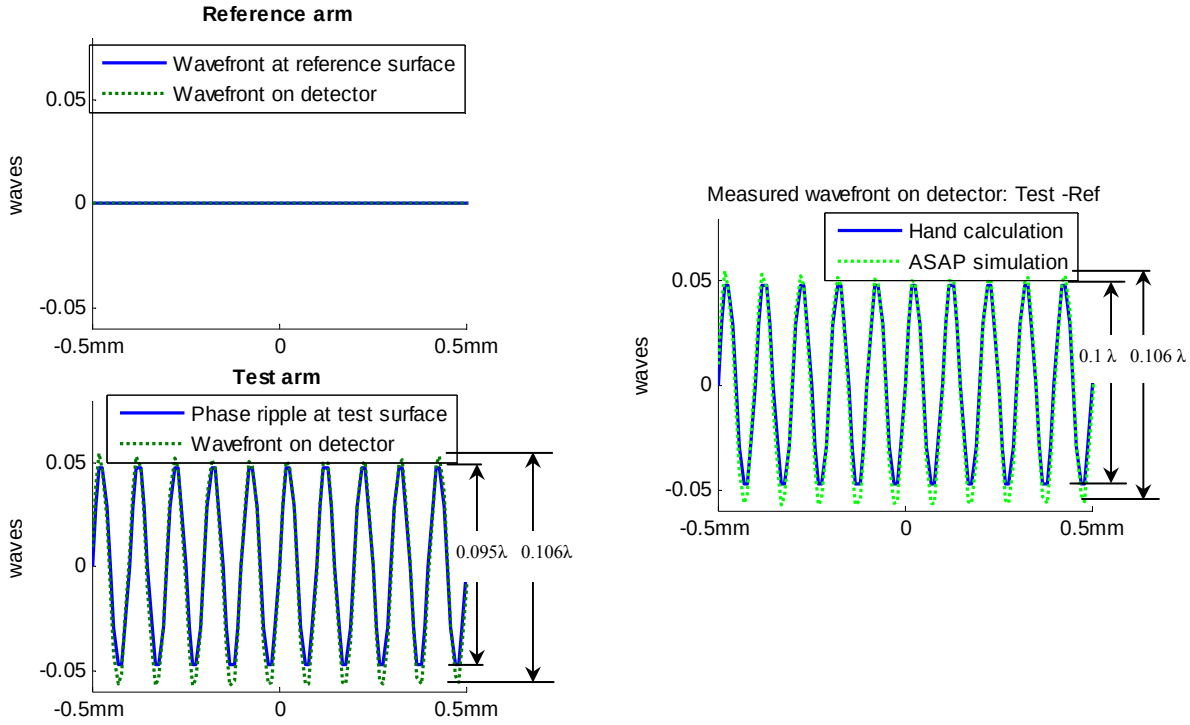


Figure **2.11**: Wave propagation analysis by ASAP when the phase ripple of 0.1λ PV is added at the test surface. The measured wavefront on the detector is calculated as the difference between the test and reference wavefronts on the detector. The phase ripple on the test surface can be correctly measured. Only the 10 cycles at the middle of the aperture are shown in the pictures.

With the phase ripple on the reference mirror and no errors from the common wavefront and the test surface, this phase ripple will be smoothed by diffraction at the detector plane since the reference surface is out of focus by 5 mm. The smoothing effect from the reference surface can be calculated using Equation **2.31**, where L_e becomes the effective propagation distance from the reference surface (or optical conjugate plane of the reference surface) to the test surface. The magnitude of the phase ripple at the detector, calculated using the Talbot effect, is $W'_r = 0.05\lambda \cdot \cos\left(2\pi \frac{5\text{mm}}{40\text{mm}}\right) = 0.035\lambda$.

The ASAP simulation (Figure **2.12**) shows that the magnitude of this phase ripple on the detector is $\pm 0.037\lambda$ (0.073λ/2), which is very close to our calculation.

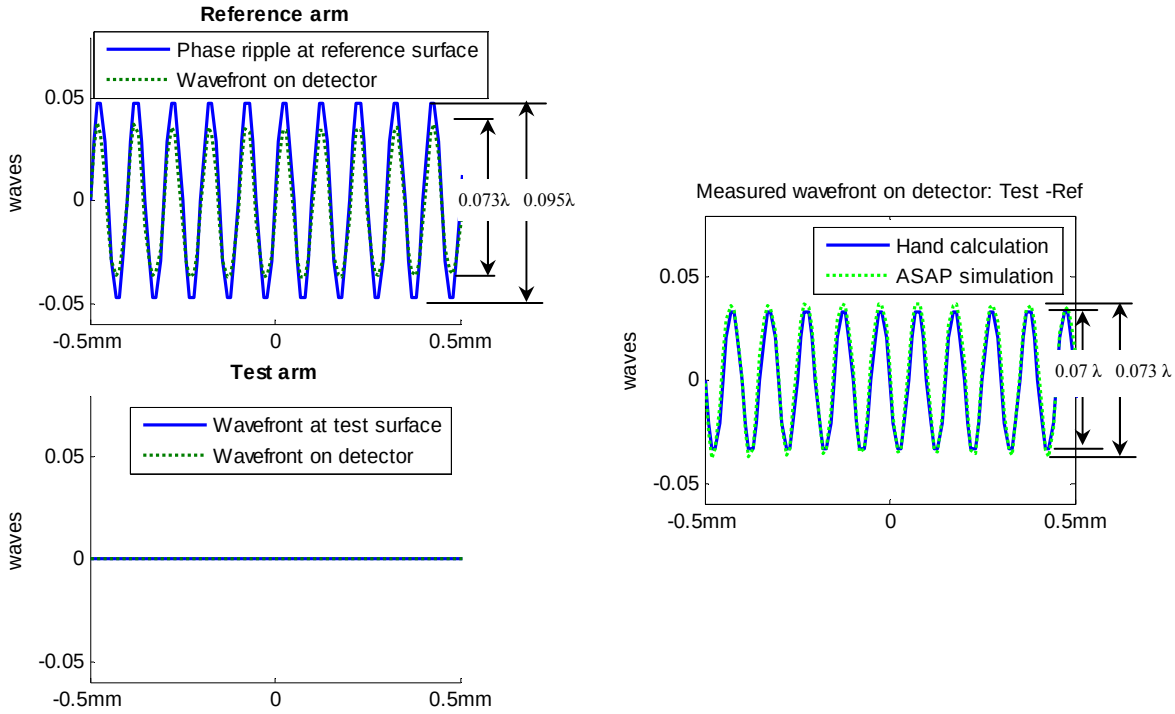


Figure **2.12**: Wave propagation analysis by ASAP when the phase ripple of 0.1λ PV is added at the reference surface. The measured wavefront on the detector is calculated as the difference between the test and reference wavefronts on the detector. The phase ripple at the reference surface is smoothed because the reference surface is out of focus by 5 mm. Only the 10 cycles at the middle of the aperture are shown in the pictures.

When the phase ripple is in the common path as labeled in Figure **2.10** and there are no errors in the reference and test surfaces, the ASAP simulation is shown in Figure **2.13**. As this phase ripple propagates to the reference surface, its magnitude becomes $W_r = 0.05\lambda \cdot \cos\left(2\pi \frac{45\text{mm}}{40\text{mm}}\right) = 0.035\lambda$ (ASAP has $0.069\lambda/2 = 0.035\lambda$). If we unfold the light path in the reference arm, then the phase ripple in the common path is 40 mm away from the plane of focus. Therefore, the magnitude of the phase ripple on

the detector should be $W'_r = 0.05\lambda \cdot \cos\left(2\pi \frac{40\text{mm}}{40\text{mm}}\right) = 0.05\lambda$ (ASAP has $0.1\lambda/2 = 0.05\lambda$).

The phase ripple on the test surface can be calculated as $W_t = 0.05\lambda \cdot \cos\left(2\pi \frac{50\text{mm}}{40\text{mm}}\right) = 0$,

and the phase ripple should have zero magnitude at the detector because the interferometer focuses on the test surface. ASAP also shows that the phase ripple has almost zero magnitude.

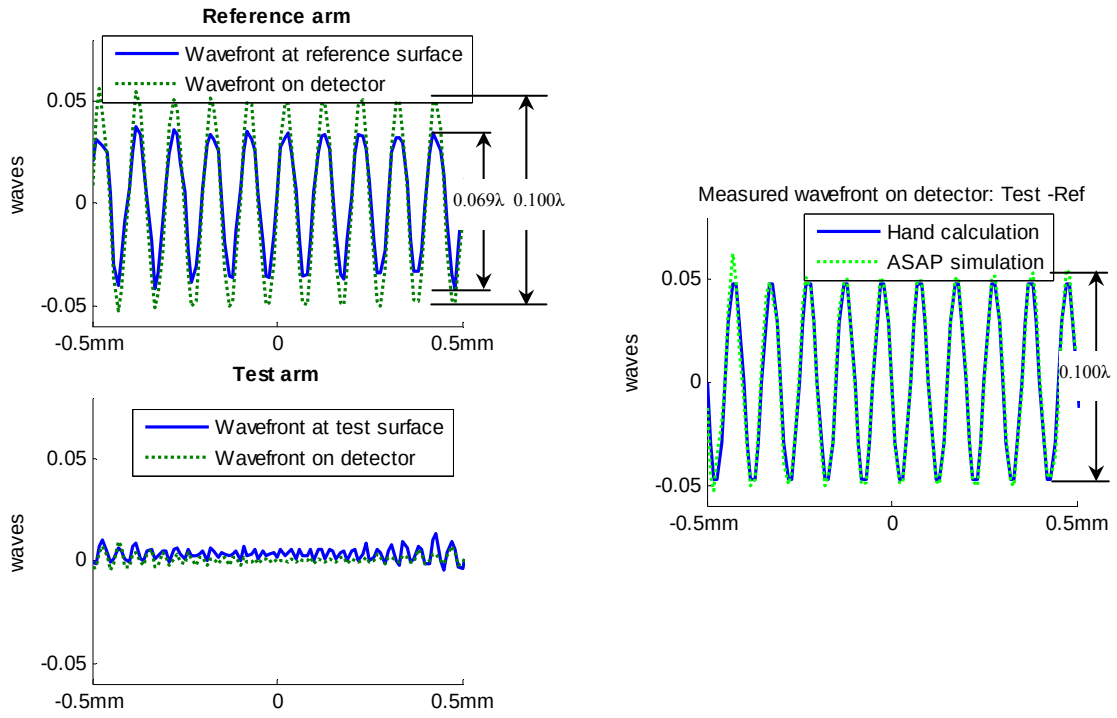


Figure **2.13**: Wave propagation analysis by ASAP when the phase ripple of 0.1λ PV is added in the common path. The measured wavefront on the detector is calculated as the difference between the test and reference wavefronts on the detector. Only the 10 cycles at the middle of the aperture are shown in the pictures.

The above three cases show that calculations using the Talbot effect agree with the ASAP simulation based on wavefront propagation, so it can be used to analyze the

phase smoothing effects due to diffraction and greatly simplify the complex wavefront propagation problem. Note that if the distances and optics properties inside the interferometer are known, the diffraction effect from each optic can be accurately simulated based on wavefront propagation or calculated using the Talbot effect.

2.3.2.3 Example

In this section, two examples of calculating the Fresnel number and the phase smoothing effects from the reference surface are given.

Test of a plane mirror with a Twyman-Green interferometer:

Figure **2.14** shows a flat mirror tested with a Twyman-Green interferometer. The aperture size of the reference beam is 7 mm in diameter and the test optic is 300 mm away from the plane optically conjugate to the reference surface.

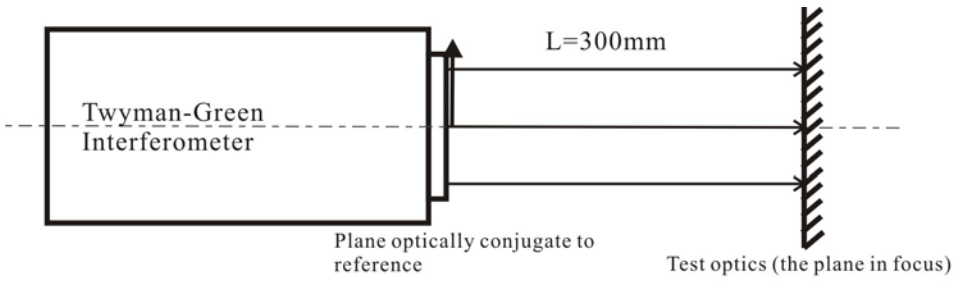


Figure **2.14**: Use of a Twyman-Green interferometer to test a flat mirror. The flat is 300 mm away from the plane optically conjugate to the reference surface.

The Fresnel number describing propagation from the plane optically conjugate to the reference surface to the test optic for a wavelength of 633 nm is

$$N_f = \frac{a^2}{\lambda L} = \frac{(7/2)^2}{0.633 \cdot 10^{-3} \cdot 300} = 64.5 .$$

Test of a concave mirror using a Twyman-Green interferometer:

When testing a concave or convex surface, the effective propagation distance can be calculated such that it is equivalent to propagation in a collimated beam. An example of testing a concave mirror with radius of curvature of 100 mm is given below, shown in Figure **2.15**. The optically conjugate plane of the reference surface is at the transmission sphere, which has an F-number of 7 and the beam diameter of the interferometer is 7 mm. Therefore R_1 is calculated to be 49 mm. The effective propagation distance from the plane optically conjugate to the reference surface to the test optic is

$$L_e = \frac{R_1(R_1 - R_2)}{R_2} = \frac{49 \cdot (49 + 100)}{-100} = -73\text{mm} ,$$

and its Fresnel number is

$$N_f = \frac{a^2}{\lambda L_e} = \frac{-3.5^2}{0.633 \cdot 10^{-3} \cdot 73} = -265 .$$

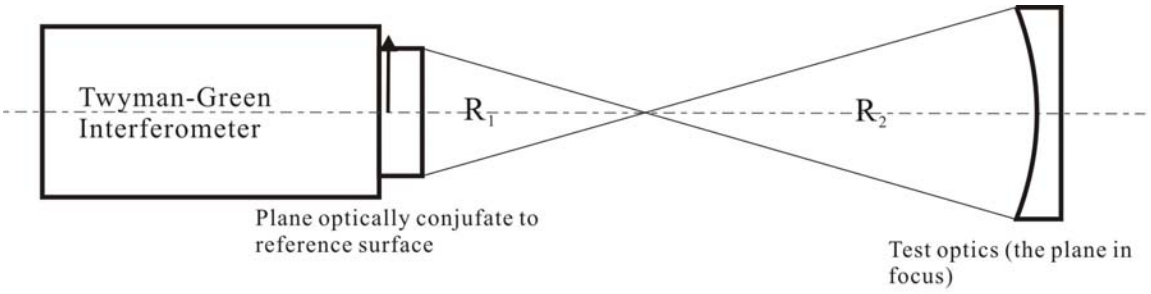


Figure **2.15**: Use of a Twyman-Green interferometer to test a sphere. The sphere has a radius of curvature of 100 mm.

The transfer functions of the reference wavefront for the two test setups above, calculated using Equation **2.35**, are shown in Figure **2.16**. They are close to unity at the low frequency region, which means that there is not much smoothing effect in this region. However, the phase smoothing effect becomes significant at the high frequency region, especially when the magnitude of the Fresnel number is small. It is useful to check the Fresnel number of a test setup, which gives a good understanding about how accurate the measurement will be up to a certain spatial frequency. Usually, the reference surface of an interferometer is of high quality and errors from the reference surface are small and of low spatial frequency. Therefore, the smoothing effect of the reference wavefront is likely to be small. Moreover, errors from the reference wavefront, including diffraction effects, can be calibrated with an absolute test and this will be discussed in detail in Chapter 3.

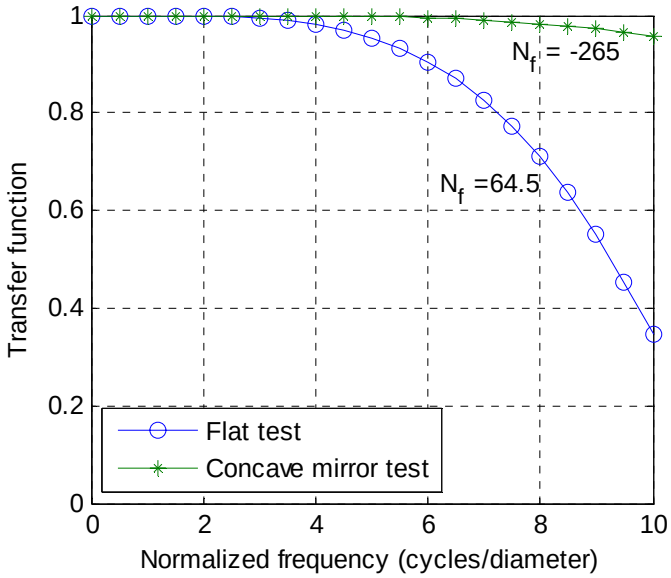


Figure **2.16**: Transfer functions of the reference wavefront vs. spatial frequency for the two test setups shown in Figure **2.14** and Figure **2.15**.

2.3.3 Diffraction effect: edge diffraction

Edge diffraction affects the measurement when the aperture of an interferometer is not in focus. The limiting aperture of the interferometer itself, which is the transmission sphere or flat, will generally not be in focus since interferometers often image the surface under test on the detector to correctly represent any errors in the test surface. The edge diffraction pattern from the aperture of the transmission optics can be calculated as a spherical wavefront propagating from this aperture to the test optic (plane of focus). The severity of edge diffraction from the transmission optics depends on the Fresnel number, which represents the propagation geometry from the transmission optics

to the test optic. The larger the Fresnel number, the denser the diffraction ring pattern and the less edge diffraction effect is observed.

To evaluate the phase error introduced by edge diffraction, a MATLAB (The MathWorks, Inc., Natick, MA) simulation with a 1024 × 1024 sampling is performed for different Fresnel numbers. This is for the case when the limiting aperture is at the reference surface. Figure **2.17** illustrates that systems with a smaller Fresnel number have larger RMS phase errors due to edge diffraction at the wavelength 632.8 nm. The RMS phase error for the Fresnel number of 62 is 10.7 nm. Therefore, for a measurement requiring high accuracy, it is important to check the Fresnel number of the test setup to make sure that the errors due to diffraction are small.

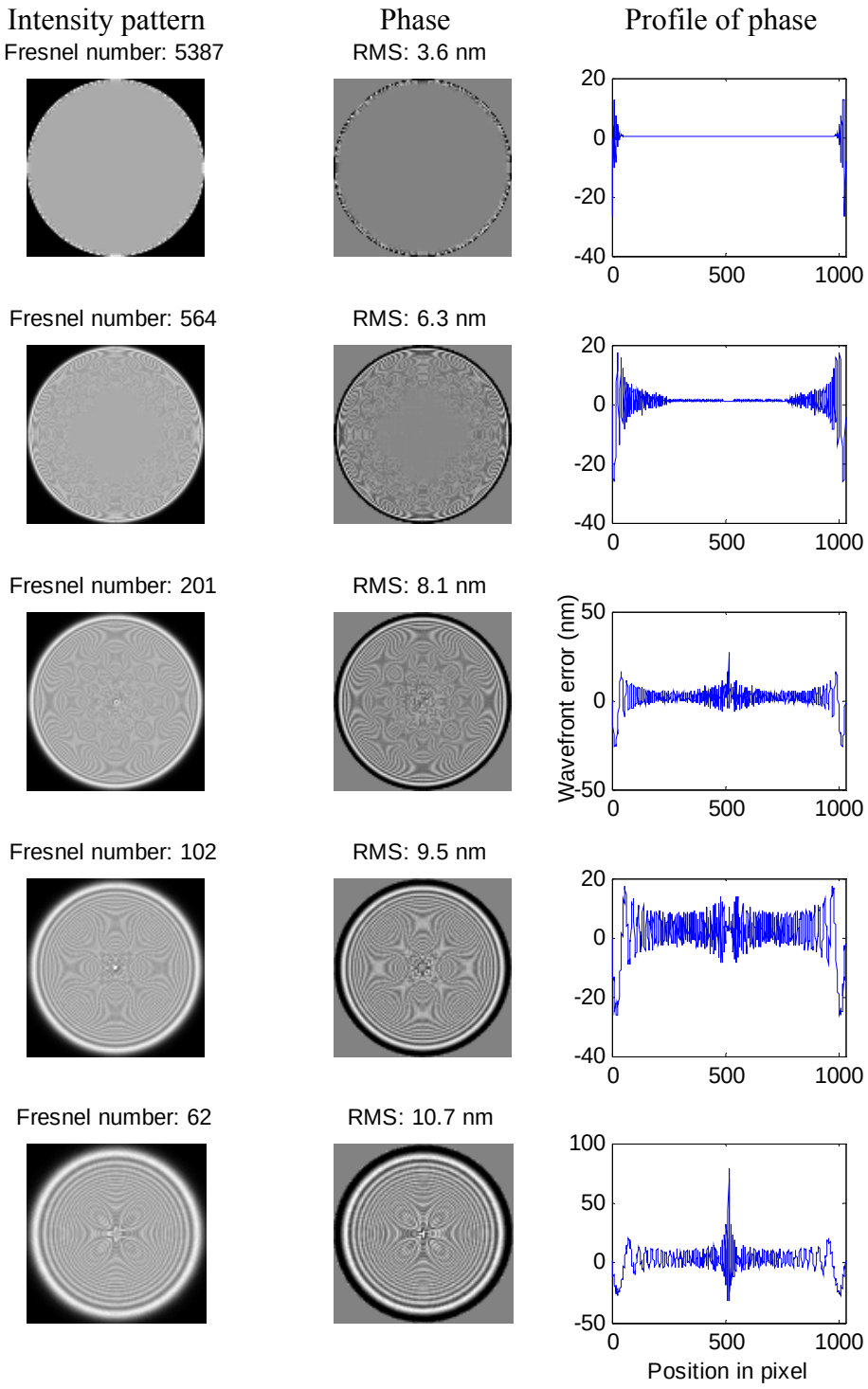


Figure **2.17**: Intensity pattern (left), corresponding phase pattern (middle) and the line cross-section along the diameter (right) due to edge diffraction at different Fresnel numbers. The wavelength is 632.8 nm.

2.4 Propagation of Zernike polynomial wavefronts

Zernike polynomials are often used in optical testing since they have the same form as the types of aberrations often observed in optical tests. It is useful to see how diffraction effects influence phase changes for each Zernike term. The definitions of Zernike polynomials and the first 37 terms are shown in Appendix B. Unlike sinusoidal wavefronts, Zernike polynomials do not have a single spatial frequency. Therefore, it is difficult to use the Talbot effect to estimate the phase smoothing of Zernike polynomials. We simulated the behavior of Zernike polynomials due to wave propagation in a collimated beam with MATLAB.

The simulation is performed following the procedure described in Section **2.3.1.1**. The input function at the plane $z = 0$ is each Zernike term. The free-space transfer function is multiplied by the Fourier Transform of the input Zernike function, then an inverse Fourier Transform is performed to get the wavefront phase after propagating a distance L. This can be described as

$$u_L(x, y) = \Im^{-1}\left(U_{z=0}(\xi, \eta) \cdot H_L(\xi, \eta)\right). \qquad (\mathbf{2.37})$$

The simulation includes both diffraction effects: phase smoothing and edge diffraction. As shown in Figure **2.18**, if the original input of the Zernike polynomial Z_i has a magnitude α_{in} , the output Zernike term will have a magnitude α_{out} after propagating a distance of L, and this is the smoothing effect. Because each Zernike polynomial has more than one spatial frequency and each frequency component has a

different smoothing effect, there will also be residual errors which could not be fit by the original Zernike term.

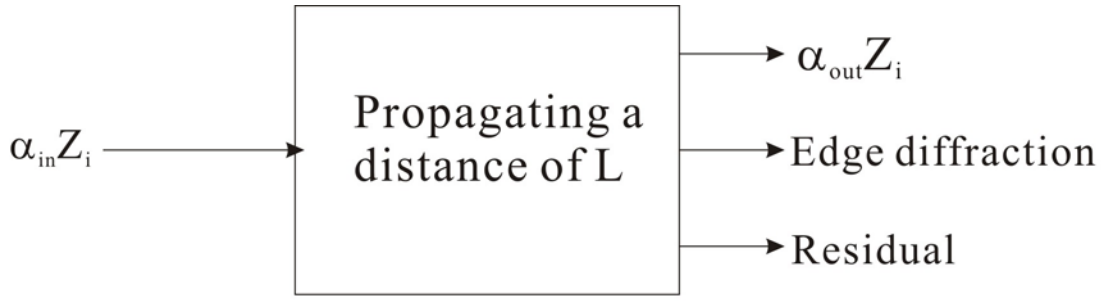


Figure **2.18**: As a Zernike term propagates a distance *L*, there will be smoothing effect, edge diffraction and some residual errors since Zernike terms are not of single spatial frequency.

Figure **2.19** shows the simulation results for standard Zernike terms from 35 to 39 with a Fresnel number of 50. The first column shows the input Zernike functions, all of which have an error of 0.1λ rms. The second column is the output wavefront after propagating a distance of *L* in a collimated space. The third column shows the output wavefront after removing the edge diffraction effect. The fourth column is a fit of the input Zernike polynomial to the output wavefront with edge diffraction removed. It shows how much the input Zernike term is smoothed due to propagation. The higher Zernike terms have a larger smoothing effect because they are of higher order spatial frequencies. The fifth column is the residual error, which is the difference between the third and the fourth columns. The residual errors are introduced by diffraction and are uncorrelated with original wavefronts.

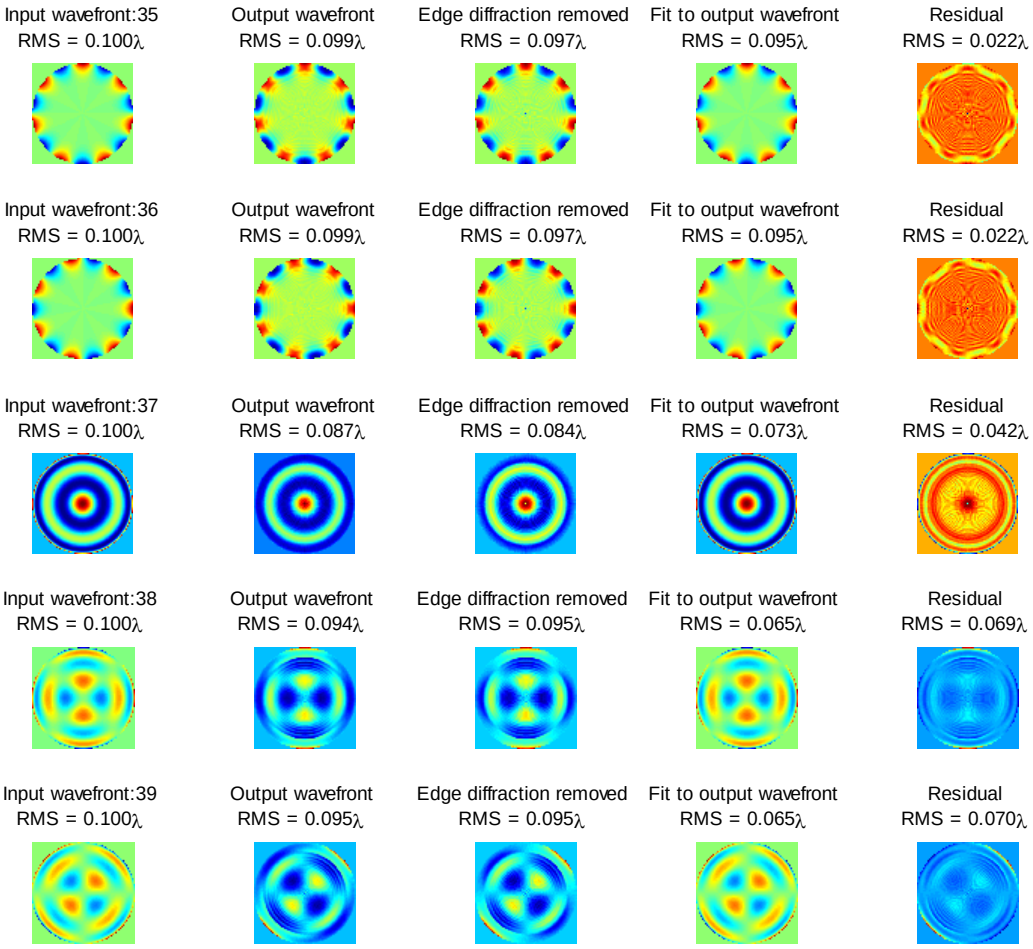


Figure **2.19**: Computer simulation of wavefront propagation for standard Zernike terms from 35 to 39 for a Fresnel number of 50. The first column is the input Zernike functions with an error of 0.1λ rms. The second column is the output wavefront after propagating a distance of L in a collimated space. The third column shows the output wavefront after removing the edge diffraction effect. The fourth column is a fit of the input Zernike polynomial to the output wavefront. The fifth column shows the difference between the third and the fourth columns.

Figure **2.20** shows plots of the RMS values of the input wavefront (Column 1 in Figure **2.19**), the remaining input wavefront (Column 4), and the residual (Column 5). All of the numbers are scaled such that the input wavefront has an RMS error of 1. The plot shows a number of low points in RMS errors of the residual maps, which correspond

to the Zernike terms having a form $r^n \begin{pmatrix} \cos n\theta \\ \sin n\theta \end{pmatrix}$ such as Z35 and Z36 shown in Figure **2.19**. It is because these Zernike terms contain less curvature variations (less high spatial frequencies) than the others.

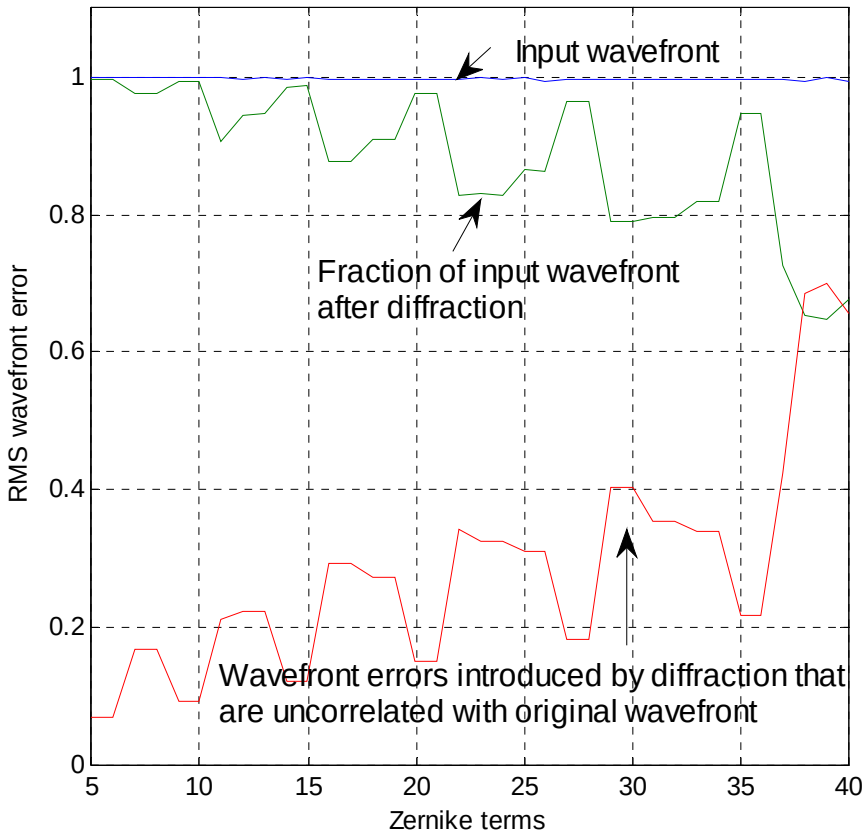


Figure **2.20**: RMS wavefront errors of the input wavefront, the fit of output wavefront to the input wavefront after removing edge diffraction, and the residual.

The ratio of remaining RMS wavefronts α_{out} to input RMS wavefronts α_{in} can be described as transfer efficiency for each Zernike polynomial. Figure **2.21** shows the transfer efficiency of each Zernike term for different Fresnel numbers. This helps

estimate how much each Zernike term is smoothed in optical testing. It is apparent that the test geometries with larger Fresnel numbers have less smoothing effect.

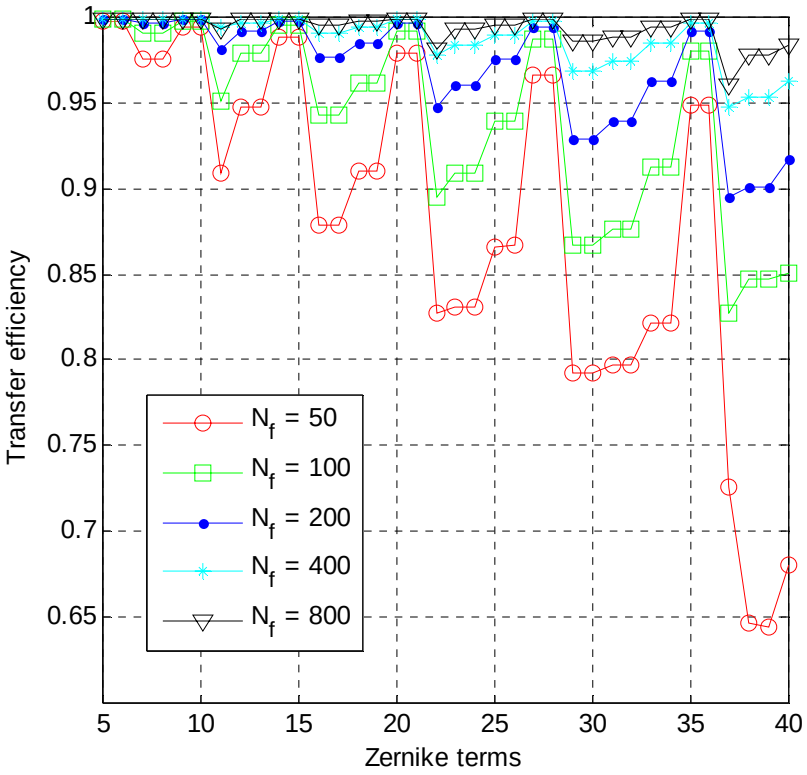


Figure **2.21**: The transfer efficiency for each Zernike term at different Fresnel numbers, defined as the ratio of remaining RMS wavefronts α_{out} to input RMS wavefronts α_{in}.

Figure **2.22** shows the generation of higher order residual error, which is defined as the ratio of the RMS residual wavefront to the input RMS wavefront. It is shown that the systems with small Fresnel numbers will introduce more high order residual errors.

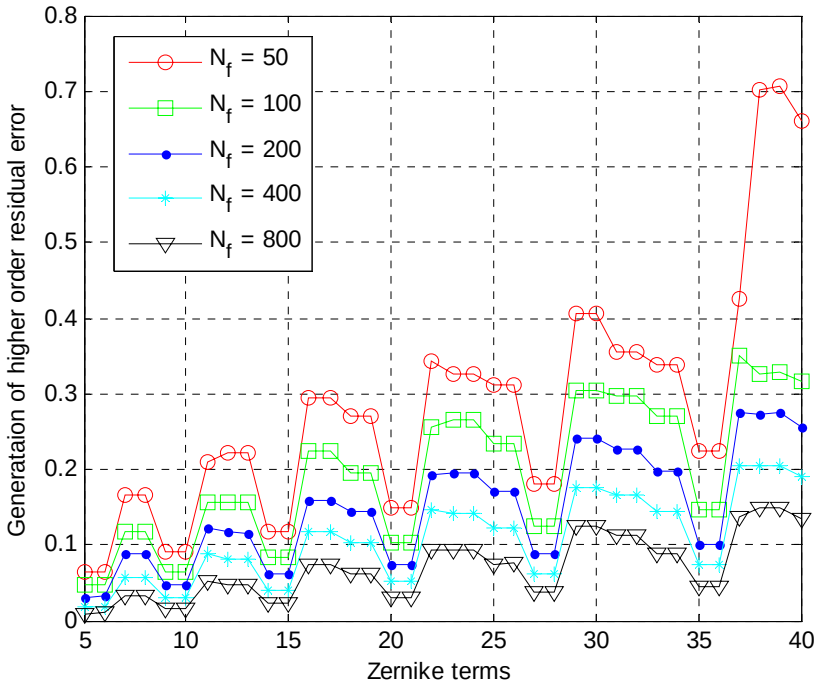


Figure **2.22**: Generation of higher order residual error, defined as the ratio of the RMS residual wavefront to the input RMS wavefront.

2.5 Conclusion

Interferometric surface measurements suffer from uncertainties from random errors and systematic errors. Random errors can be reduced by averaging a number of measurements, while systematic errors are classified into geometric errors and errors due to diffraction effects. Retrace error, imaging distortion and shearing effects are the three geometric errors, and are discussed in this chapter. Diffraction effects include phase smoothing in the mid/high spatial frequencies and edge diffraction. This chapter presents

a method of analyzing the diffraction effects using the Talbot effect, which can greatly simply the complex diffraction problem.

CHAPTER 3

INTERFEROMETER CALIBRATION ERRORS

Interferometric measurements rely on accurate and well-known reference wavefronts. For measurement of precision surfaces and wavefronts, the errors from the reference surface are often of the same order of magnitude as the errors of the surface being measured. To achieve measurement accuracy better than the reference surface, errors from the reference surface must be characterized and subtracted from the measurement result. This chapter first discusses different interferometer calibration methods, and then describes in detail the errors remaining due to the limitations of the calibration for the random ball test.

3.1 Interferometer calibration methods

The accuracy of interferometric measurements is usually limited by the errors of reference optics; however, some optical components need to be measured more precisely for demanding applications. This necessitates interferometer calibration to measure errors from reference optics, and then remove those errors from test surface measurements. These procedures are also called absolute tests. Over the years, a variety of calibration methods have been developed for measuring surfaces. Of all these methods, the three-flat test, the three-position test and the random ball test are the most common.

3.1.1 Three-flat test and three-sphere test

The three-flat test [**Schultz 1976**] is the most popular technique to measure flats in an absolute sense with a Fizeau interferometer. Three flats A, B, and C are measured against each other using the sequence outlined in Figure **3.1**. The first letter in a pair, such as BA, refers to the reference flat, while the second letter refers to the test flat. Four measurements are required to find an absolute measurement of the x and y profiles of three flats. If the shape of the entire surface is desired, multiple profiles can be obtained by repeating the process.

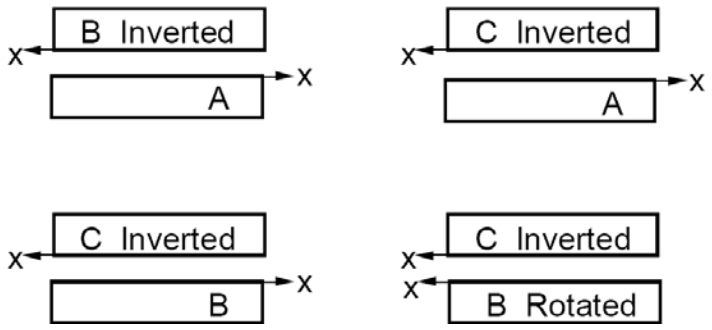


Figure **3.1**: Three-flat test. The flat on the top in the pair refers to the reference flat, while the bottom one refers to the test flat. The measurements BA, CA and CB give the y profile in absolute sense. Rotating B by 180^{o} gives the x profile.

The three-flat test philosophy can also be applied to the absolute test of a sphere. This has been carried out by Harris [**Harris 1971**] and Hopkins [**Hopkins 1971**]. Similar to the three-flat test, three unknown spherical surfaces A, C (concave surfaces), and B (convex surface) are tested in pairs as shown in Figure **3.2**. As is the case for flats, this method only delivers the absolute deviations along x and y profiles of three spheres. Unlike the flat test, the two combinations of the concave surfaces A and C with the

convex surface B result in a short cavity, while the third surface combination AC results in a large cavity length. Therefore, a wave propagation problem arises that limits the achievable accuracy.

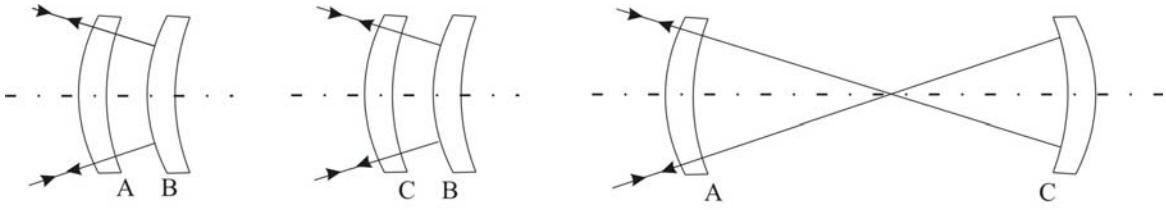


Figure **3.2**: Absolute test of spheres. The figure shows the side view of three two-by-two combinations of spherical surfaces A, B, C in a Fizeau interferometer.

3.1.2 Further investigation on three-flat test

The classical three-flat test can only solve the profile along a straight line, and there have been a number of efforts to extend it to the entire surface. One approach is to rotate one of the flats and use four measurements, but fit the Zernike polynomials to the wavefront and represent the surfaces by Zernike polynomials [**Parks 1978; Fritz 1984**]. Methods based on Zernike polynomials, although powerful, are limited to errors that are well represented by the order of Zernikes fitted.

Another approach uses the symmetry properties of odd and even functions [**Ai 1993**]. The flat is decomposed into four components: odd-even even-odd, even-even, and odd-odd. This method requires six measurements to derive the three surfaces, and unlike

the use of Zernike polynomials, it solves the surface map on a pixel-by-pixel basis and has the ability to solve the high-spatial-frequency information.

The N-position test [**Evans 1996**] is another technique, which requires rotating the test piece to N equally spaced positions. The average of the N measurements contains three errors: all the errors in the reference optic, the rotationally invariant errors in the test part and the non-rotationally symmetric errors of azimuthal order Nk (where k is any integer) in the test piece. Subtracting the average data from one of the individual measurements will return the nonrotationally symmetric terms for the interferometer and the test piece, respectively. The N-position test can also be used in an absolute test, and the disadvantage is that it is blind to the errors of azimuthal order Nk.

Recently, the three-flat test has been investigated in the presence of deformations caused by the support mechanism for the flats, which is a significant problem for large, heavy flats [**Griesmann 2007**]. The solution is derived based on rotational symmetry and mirror symmetry, which provides the flat deformation along with the wavefront flatness errors for the three flats.

3.1.3 Three-position absolute measurement

The technique for absolute measurements of spherical surfaces described by Jensen requires three separate measurements of the surface being tested [**Jensen 1973**], depicted in Figure **3.3**. The first measurement is with the test surface at the focus of the diverger lens (also known as the cat's eye position). The second measurement is with the test surface positioned so that its center of curvature is at the focus of the diverger lens

(also known as the confocal position). The third measurement is taken after rotating the test surface 180º about the optical axis. These three measurements can then be used to solve the test surface by

$$W_{surf} = \frac{1}{2}\left(W_{0^0} + \overline{W}_{180^0} - W_{\text{focus}} - \overline{W}_{\text{focus}}\right). \qquad \textbf{(3.1)}$$

A bar over a wavefront W indicates a 180º rotation of that wavefront. The errors in the reference surface can be obtained by calculating

$$W_{\text{ref}} = \frac{1}{2}\left(W_{0^0} - \overline{W}_{180^0} + W_{\text{focus}} + \overline{W}_{\text{focus}}\right). \qquad \textbf{(3.2)}$$

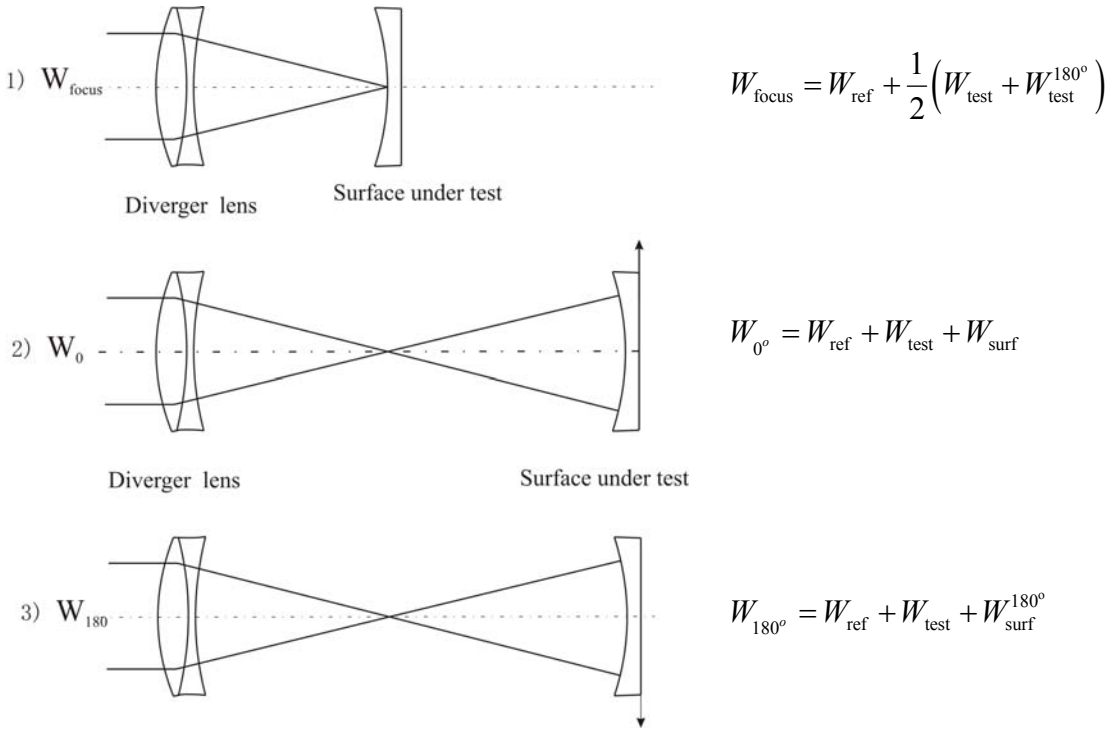


Figure **3.3**: Three-position absolute measurement of spherical surfaces. The spherical surface is tested at the cat's eye position and two confocal positions. W_{focus}, W_{0^o} and W_{180^o} are the wavefront obtained in the test. W_{ref}, W_{test} and W_{surf} refer to the reference arm wavefront, test arm wavefront and the test surface wavefront. The number on the superscript means the orientation of the wavefront.

This technique can be used to calibrate both Fizeau and Twyman-Green interferometers. The main disadvantages are that it requires critical alignment of the test surface and an extremely good rotation stage for 180º rotation, and rotating the optic may cause the shape to change. The three-position test suffers additional errors due to wave propagation since the measurements at cat's eye and confocal positions have different propagation distances.

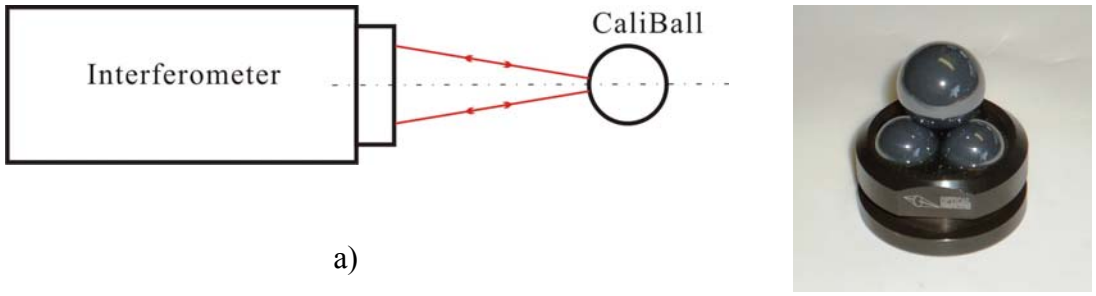
3.1.4 Absolute test with the Maximum Likelihood (ML) method

Recently, Su demonstrated an absolute test based on the maximum likelihood estimation [**Su 2006**]. Unlike the other absolute test, this method involves an optimization process. Multiple measurements need to be taken where both the test and reference surfaces are translated and rotated, which provides sufficient information to reconstruct the figure of both the reference and test surfaces. There is no need for additional optics for this method. The optimization characteristics of this method make it general and easily applied for multi-interferogram testing.

3.1.5 Random ball test

The Random Ball Test (RBT) has recently become a popular technique for calibrating spherical wavefronts. In this test, a high-quality silicon nitride ball, called a CaliBall, is measured against the reference surface at its confocal position in a number of random orientations shown in Figure **3.4 (a)**. This procedure averages out the errors in

the ball and converges toward the stationary error in the reference sphere [**Parks 1998**]. Practical experience with the RBT indicates that the number of ball orientations required for a reference calibration good to 1 nm rms is about 10 to 15 measurements. To maintain alignment while rotating the ball, the ball can be set on a kinematic support consisting of three smaller balls arranged in an equilateral triangle, as shown in Figure **3.4 (b) [Parks 2006]**. The 25 mm CaliBall are Grade 5 (meaning round to better than 125 nm) and have a surface finish of about 1.3 nm rms. The reflectivity is about 4% in the visible.



a) b)

Figure **3.4**: Random ball test. a) The ball is measured at the confocal position; b) The kinematic support for RBTs.

The RBT is quite simple and easy to perform, and does not require additional optics unlike the three-flat method.

3.1.6 Comparison of interferometer calibration methods

A comparison of the advantages and disadvantages of all the calibration methods mentioned above is listed in Table **3.1**.

Table **3.1**: Comparison of different interferometer calibration methods

Test	Advantages	Disadvantages
Three-flat or three-sphere	-Absolute tests of three flats/spheres at a time can be obtained.	-Only a line profile is measured - Multiple measurements are required if entire surface shape is desired. - Wave propagation problem arises for absolute sphere test. - Three flats/spheres are needed including the one under test.
Three-position	- No additional optics are required besides the transmission sphere.	- Critical alignment of the test surface is required. - Extremely good rotation stage is needed. - Non-common path measurement at cat's eye position may cause aberrations.
ML method	- No additional optics are required besides the transmission sphere or flat - A general absolute method for testing flat and sphere	- Do require rotation and translation of the test piece or reference piece.
Random ball	- Measurements are easy to perform. - No additional optics are required. - Kinematic support is used.	- Relatively large number of measurements is needed

3.2 Remaining errors in the RBT

Absolute tests can provide an error map of the reference surface, which can then be subtracted from measurements of the surfaces being tested. Besides measurement noise, the calibration will not be perfect if the interferometer cavities of the absolute test and the surface measurement are different, because retrace errors, imaging distortion, and

errors due to diffraction depend on the test geometry. Unlike random measurement noise, these errors are systematic.

In this section, the RBT is used as an example to study the remaining errors due to the limitation of the calibration [**Zhou 2009**]. A 25 mm diameter CaliBall is used to calibrate a 6" F/3.2 transmission sphere on a Wyko interferometer, unless otherwise stated. Measurement errors discussed in Chapter 2 will be studied for the RBT. The goal is to provide a procedure that one can follow to estimate the remaining errors in an absolute test.

3.2.1 Measurement uncertainty

The measurement uncertainty in the RBT consists of the noises in each test and the errors due to an insufficient number of measurements. Figure **3.5** shows an average of 30 random ball tests, which has a surface irregularity of 6.5 nm rms. This map will be used as a calibration map and subtracted from the surface measurement.

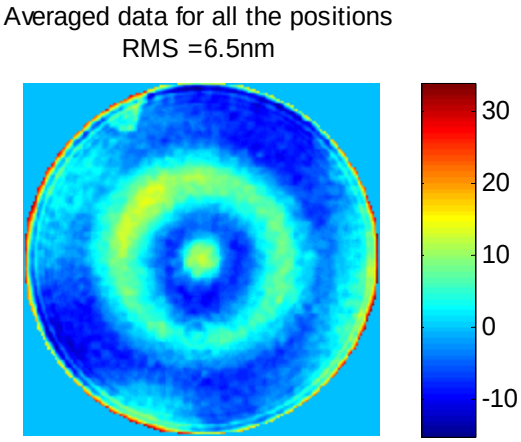
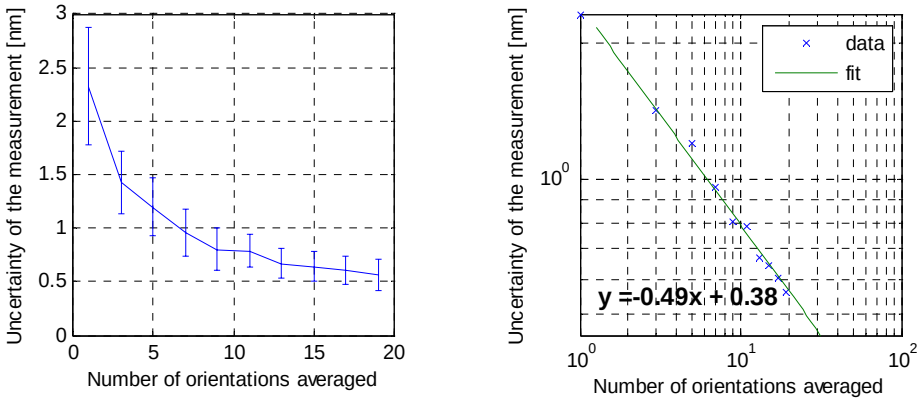


Figure **3.5**: An average of 30 RBTs. A 25 mm CaliBall is used to calibrate a Wyko interferometer with a 6" F/3.2 transmission sphere.

The variation in the data sets can be used to estimate the contribution to the calibration uncertainty from the ball surface irregularity and the measurement noises, as described in Chapter 2 Section **2.1**. Figure **3.6** gives the measurement uncertainty versus the number of random ball tests, and the log-log plot shows that the noise in the average of 30 maps is about 0.5 nm rms. The slope of the fitted straight line is -0.49, which means that the errors are random since it is dropping off as $1/\sqrt{N}$, where N is number of RBTs.



a) b)

Figure **3.6**: Measurement noise versus the number of the RBTs. a) normal scale, error bar shows the standard deviation of the measurement uncertainty; b) log scale showing that the noise is dropping off as $1/\sqrt{N}$, which means that the noise is random.

3.2.2 Geometric errors

Geometric errors are intrinsic to the interferometer. As discussed in Chapter 2, geometric errors include retrace error, imaging distortion and shearing effect. The shearing effect can be avoided if the transmission sphere is aligned correctly. Geometric errors can be compensated if the interferometer is completely characterized [**Murphy 2000**]. Unfortunately, it is difficult to completely characterize the interferometer without knowing the distances and optical properties inside the interferometer. However, the first-order geometric errors can be estimated using sensitivity functions by

$$\Delta W_{\text{geo}} = S_x(x,y)\nabla_x M + S_y(x,y)\nabla_y M \,, \qquad (\mathbf{3.3})$$

where M is the surface measurement, and $\nabla_x M$ and $\nabla_y M$ are the slopes of the measurement in the x and y directions. S_x and S_y are the sensitivity functions to the x slope and y slope respectively, which quantify the first order effect from retrace error and imaging distortion. The sensitivity functions can be described as linear combinations of Zernike polynomials

$$S_x(x,y) = \sum_i C_{xi} \cdot Z_i(x,y), \quad S_y(x,y) = \sum_i C_{yi} \cdot Z_i(x,y), \tag{3.4}$$

where C_i represents the coefficients of the corresponding Zernike polynomials $Z_i(x,y)$. The sensitivity functions have units of $\text{nm}/\mu\text{rad}$, and can be multiplied by the local slopes of a measurement to calculate the geometric errors. The procedure to measure the geometric errors is:

1. Set up the interferometer, align the part to achieve minimum fringes (or null measurement), and take a measurement denoted as M (average to reduce noise). Also at every pixel, calculate the measurement local slopes in the x and y directions denoted as $\nabla_x M$ and $\nabla_y M$, which have units of μrad.

2. Introduce various amounts of tilt along both the x axis and y axis, and record the Zernike coefficients. Any change from the null test, after alignment terms of tilt and power are removed, is attributed to geometric errors.

3. Perform a linear fit on the recorded Zernike coefficients with varying amounts of tilt to find the sensitivity coefficients C_{xi} and C_{yi}.

4. Multiply the sensitivity functions by the slope of the surface measurement M, and calculate the geometric errors.

The reason to use lower order Zernike polynomials to find sensitivity coefficients is that they act as a low-pass filter and are less sensitive to measurement noise. This method of estimating the geometric errors has limitations. It only estimates the first-order geometric error, which is linearly related to the wavefront slope. It also requires that the surface under test is of high quality, so that its figure errors have small contribution to the geometric errors when introducing tilt to find the sensitivity functions.

A ZEMAX (ZEMAX Development Corp., Bellevue, WA) simulation was run to verify this method of calculating geometric errors, and the layout is illustrated in Figure **3.7**. Two phase plates with diameters of 150 mm are separated by 500 mm, and the surface types of both plates are Zernike Standard Phase. One wave of spherical aberration is put in the first plate, and the first 37 terms on the second plate are optimized to eliminate the residual error. The stop is on the second phase plate, and the system is set to an afocal imaging space with real ray aiming. This system has imaging distortion because the rays uniformly distributed on the first plate become non-uniform on the second plate due to propagation of the spherical aberration.

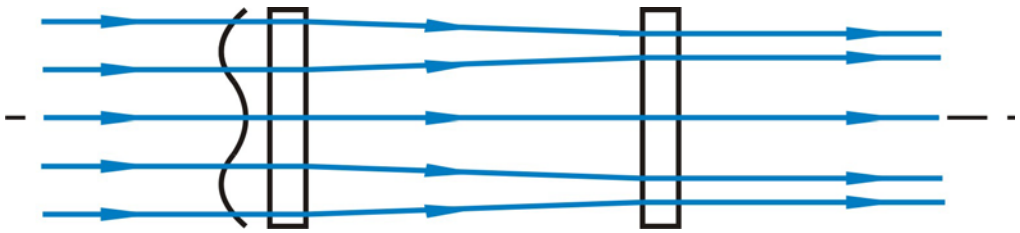


Figure **3.7**: ZEMAX simulation of geometric errors. One wave of spherical aberration is put in the first plate, and the first 37 terms on the second plate are optimized to eliminate the residual error. Then tilt and power are added in the first plate and the errors in the residual wavefront are checked.

As one to four waves of *x* tilt, *y* tilt and power are added separately to the first plate, the residual geometric errors in the form of standard Zernike polynomials (Z5 to Z11) are given in Table **3.2**. It is shown that the coma terms (Z7 and Z8) are dominant in the residual errors, and they vary linearly with the corresponding tilt added in the first phase plate. The added power causes spherical aberrations in the residual error to increase proportionally.

Table **3.2**: The residual geometric errors in the form of standard Zernike polynomials (Z5 to Z11) as one to four waves of *x* tilt, *y* tilt, and power are added separately in the first phase plate. These numbers are from the ZEMAX simulation.

Unit: waves	Z2 (Tilt x)				Z3 (Tilt y)				Z4 (Power)			
	1	2	3	4	1	2	3	4	1	2	3	4
Z5	0	0	0	0	0	0	0	0	0	0	0	0
Z6	-1.40E-07	-5.40E-07	-1.22E-06	-2.17E-06	-1.40E-07	-5.40E-07	-1.22E-06	-2.17E-06	0	0	0	0
Z7	0	0	0	0	6.94E-04	1.39E-03	2.08E-03	2.78E-03	0	0	0	0
Z8	6.94E-04	1.39E-03	2.08E-03	2.78E-03	0	0	0	0	0	0	0	0
Z9	0	0	0	0	0	0	0	0	0	0	0	0
Z10	0	0	0	0	0	0	0	0	0	0	0	0
Z11	0	0	0	1.00E-08	0	0	0	1.00E-08	1.50E-03	3.01E-03	4.50E-03	6.00E-03

One wave of standard Zernike tilt is equivalent to a slope of 2 wave/radius, so the sensitivity functions to the *x* and *y* slopes can be found using Table **3.2**, as

$$S_x(x,y)=\frac{0.000694}{2}Z8(x,y)\,,\ S_y=\frac{0.000694}{2}Z7(x,y)\,. \qquad (\mathbf{3.5})$$

Figure **3.8** (a) and (b) show the sensitivity functions simulated in ZEMAX by adding 0.5 λ of tilt in *x* and *y* directions, respectively. Figure **3.8** (c) gives the residual errors when one wave of power is added in the first plate. The ZEMAX simulation shows the residual errors with power removed (Figure **3.8** (d)) is dominated by spherical aberration with an error of 0.0015 λ rms.

The residual geometric errors from one wave of added power can also be calculated by Equation **3.3** using the sensitivity functions,

$$\begin{aligned}\Delta W_{\text{geo}} &= S_x(x,y)\nabla_x M + S_y(x,y)\nabla_y M \\ &= 0.000347\cdot Z8\cdot\nabla_x Z4 + 0.000347\cdot Z7\cdot\nabla_y Z4 \\ &= 0.000347\cdot\left[24\sqrt{6}(x^4+y^4)+48\sqrt{6}x^2y^2-16\sqrt{6}(x^2+y^2)\right]. \\ &= 0.000347\cdot\left[4\sqrt{\frac{6}{5}}Z11+48\sqrt{6}x^2y^2+8\sqrt{6}(x^2+y^2)-4\sqrt{6}\right] \\ &= 0.0015\cdot Z11+....\end{aligned} \qquad (\mathbf{3.6})$$

So the calculated spherical aberration (0.0015λ) introduced by one wave of added power matches the ZEMAX simulation listed in Table **3.2** and Figure **3.8 (d).** This means that the sensitivity functions can be used to calculate the geometric errors.

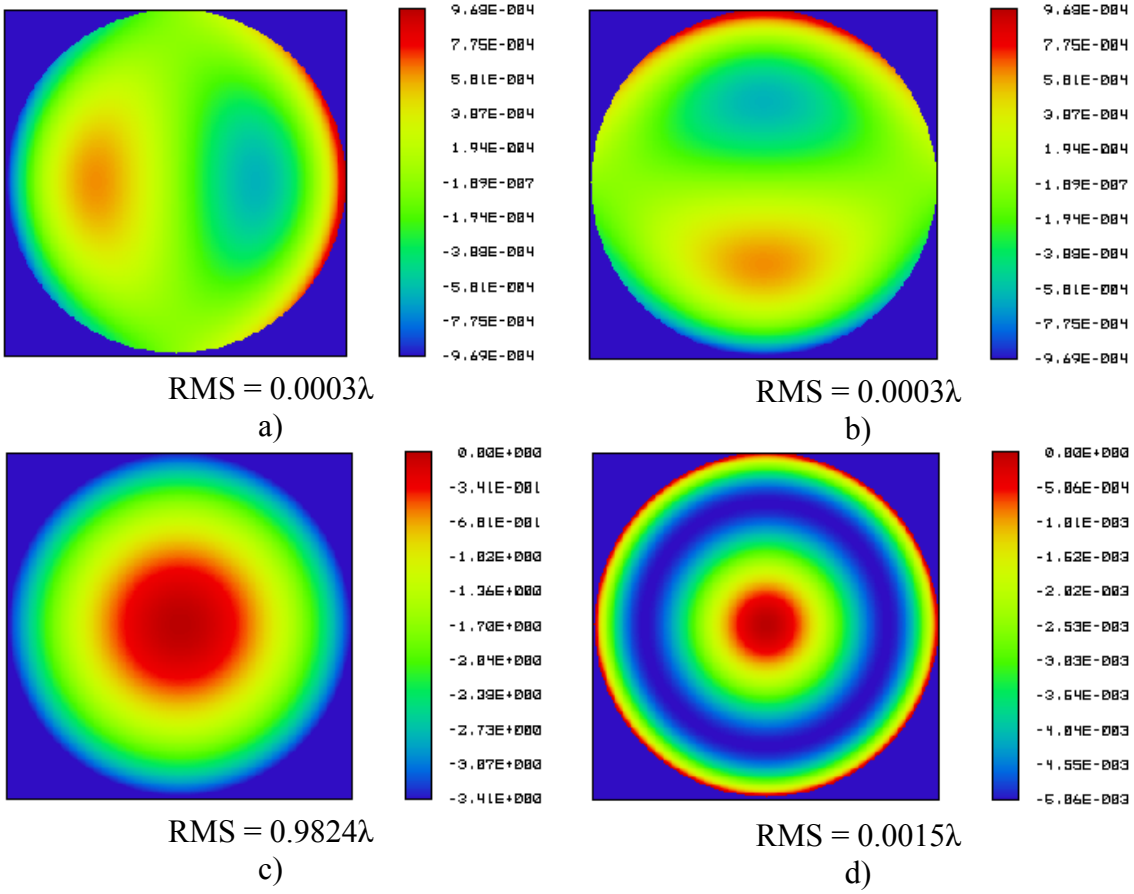


Figure **3.8**: a) The sensitivity function to *x* slope; b) The sensitivity function to *y* slope; c) Residual errors when one wave of power is added in the first plate; d) Residual errors with power removed when one wave of power is added in the first plate. All the figures are from ZEMAX simulation.

Figure **3.9** shows an example of estimating the first order geometric errors in the RBT. Following the procedure described above, the measurement errors are plotted in terms of lower-order Zernike polynomials as a function of tilt.

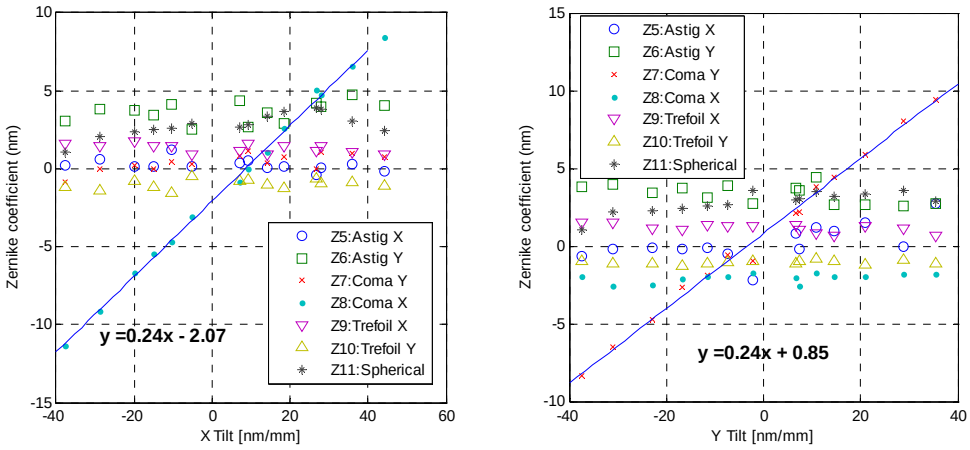


Figure **3.9**: Several lower-order Zernike coefficients plotted as a function of *x* and *y* tilts for the RBT.

Among the residual geometric errors, the coma *x* varies linearly with tilt *x* and the coma *y* varies linearly with tilt *y*. The other terms are almost unchanged. The sensitivity coefficients for the *x* and *y* coma are both 0.24 nm/μrad. Keeping only the single coma term that changes with tilt linearly, the sensitivity functions become

$$S_x(x,y) = 0.24 \cdot Z8(x,y), \quad S_y(x,y) = 0.24 \cdot Z7(x,y), \tag{3.7}$$

where $Z7(x,y) = \sqrt{8}\left(3y^3 + 3x^2y - 2y\right)$ and $Z8(x,y) = \sqrt{8}\left(3x^3 + 3xy^2 - 2x\right)$. The plots of sensitivity functions are given in Figure **3.10**. If some power is added to the RBT by displacing the CaliBall, then spherical aberration will be introduced and the amount of spherical aberration can be predicted using Equation **3.7** for a given power.

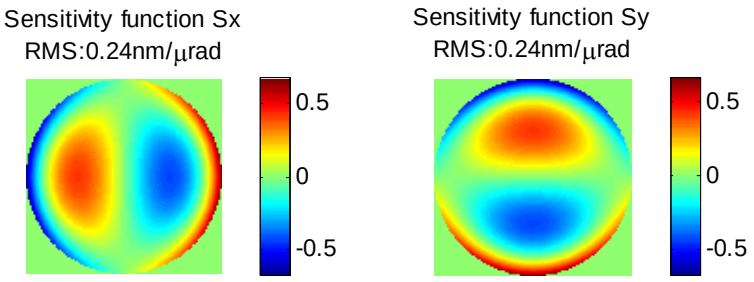


Figure **3.10**: The plots of the sensitivity functions S_x and S_y for the RBT.

The measurement results due to different alignment errors in the RBT are shown in Figure **3.11**. Since the slope data is very sensitive to edge diffraction, a small portion of data around the edge is clipped before this processing. Case (a) shows the data when the CaliBall was adjusted to have minimum fringes. Some tilt and power were introduced intentionally in case (b) and (c). The first column is the original data with tilt and power in the map, while the second column shows the data with tilt and power removed. From the second column, we can see that the measurements suffer from the geometric errors and they change with alignment errors. The next two columns show the maps of Slope X and Y, which are calculated based on the original data (the first column). The estimated geometric errors are given in the fifth column, which are subtracted from the original data. Removing tilt and power gives the compensated map shown in the last column. The compensated maps show that the geometric errors can be removed from the measurements using the sensitivity functions. The geometric error in the null test (case a) is mainly due to the local slope on the surface, and is about 0.7 nm rms.

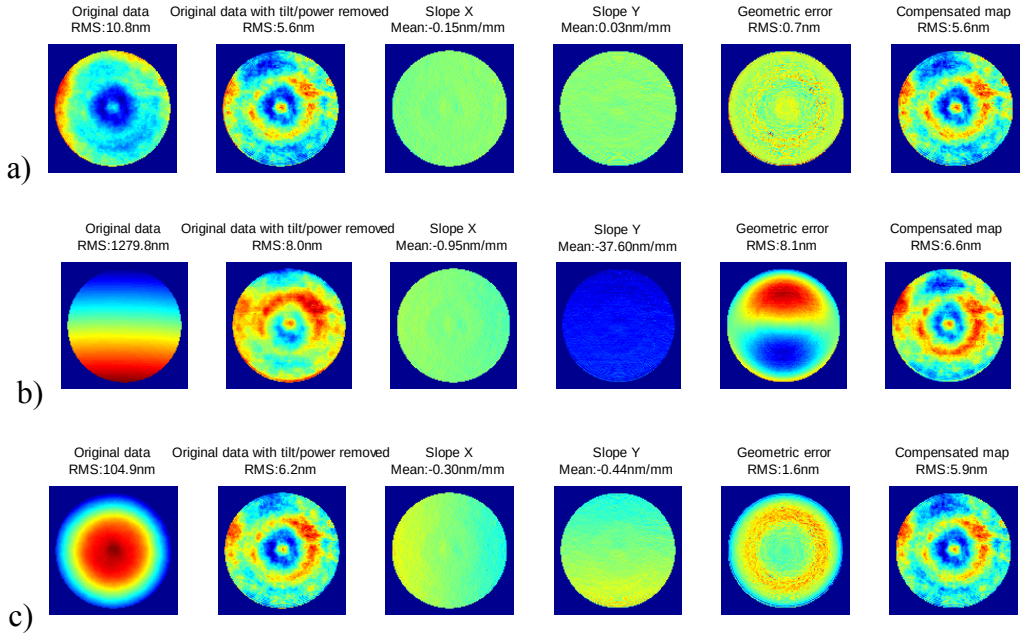


Figure **3.11**: Geometric errors when there are alignment errors. a) Test with null fringe; b) Test with -2.25λ of y tilt; c) Test with -0.2λ of power. The coefficients for tilt and power are calculated using Zernike Standard Polynomials.

3.2.3 Diffraction effects

In addition to geometric errors, the RBT also suffers errors due to wave propagation. It is best to perform the RBT after testing the surface of interest without changing the zoom or focus. This is because when the interferometer is used to test an optic, the imaging focus and zoom will be adjusted accordingly so that the interferometer images the test optic on the detector. The RBT can be performed afterwards without changing anything inside the interferometer. There are two tests: one is the RBT and the other is the test having the information of the surface of interest. To distinguish between

them, the latter one is called the surface measurement. Errors due to diffraction can be discussed from two aspects: phase smoothing and edge diffraction.

3.2.3.1 Phase smoothing for RBTs

Diffraction effects cause a decrease in high spatial frequency components, which we call phase smoothing. Phase smoothing can be studied using Talbot imaging theory [**Goodman 2004**], and it has been discussed in Chapter 2. Like Chapter 2, the remaining diffraction errors in the RBTs are discussed in three aspects: diffraction effects from the test wavefront, the reference wavefront, and the common wavefront.

Diffraction effect from the test wavefront:

In the surface measurement, the diffraction effects from the test surface are avoided because the surface is focused onto the detector. In the RBT, the CaliBall is not in focus unless it has the same radius of curvature as the test surface. However, the CaliBall is rotated to many orientations and the average of all the measurements can be regarded as a perfect spherical surface. Therefore, one can assume there is no diffraction error from the CaliBall surface either.

Diffraction effect from the reference wavefront:

In both the RBT and the surface measurement, the reference wavefronts are the same since optics inside the interferometer are unchanged. The diffraction effects on the reference wavefronts in both tests can be cancelled, even though the reference surface may not be in focus. This is only true if the focus and zoom are unchanged between calibration and the use of the system.

Diffraction effect from the common wavefront:

The common wavefront refers to the wavefront from the illumination optics in an interferometer. After passing through the beam splitter inside an interferometer, the common wavefront splits into two beams, one in the reference arm and the other in the test arm. In both the RBT and the surface measurement, the common wavefront behaves the same in the reference beam, while it usually propagates different distances in the test beams as shown in Figure **3.12** . When the test surface has the same radius of curvature as the CaliBall, the common wavefront propagates the same distance in the two tests and will not introduce any errors due to diffraction effects.

Different propagation distances cause different phase smoothing of the common wavefront, and introduce errors. The transfer function, as defined in Chapter 2, predicts the filtering of high frequency phase ripples as the wavefront propagates from the transmission sphere to the test optic. If the wavefront at the transmission sphere is known, its phase attenuation as it propagates to the test optic (plane of focus) can be evaluated. In the RBT, it is a little complicated because there is a CaliBall in between. We can first find the conjugate plane of the wavefront at the transmission sphere through the CaliBall's imaging, and then propagate that wavefront to the test optic. The complex amplitude is reproduced at the wavefront's conjugate plane. The equations for calculating transfer functions for the surface measurement and the RBT are provided below.

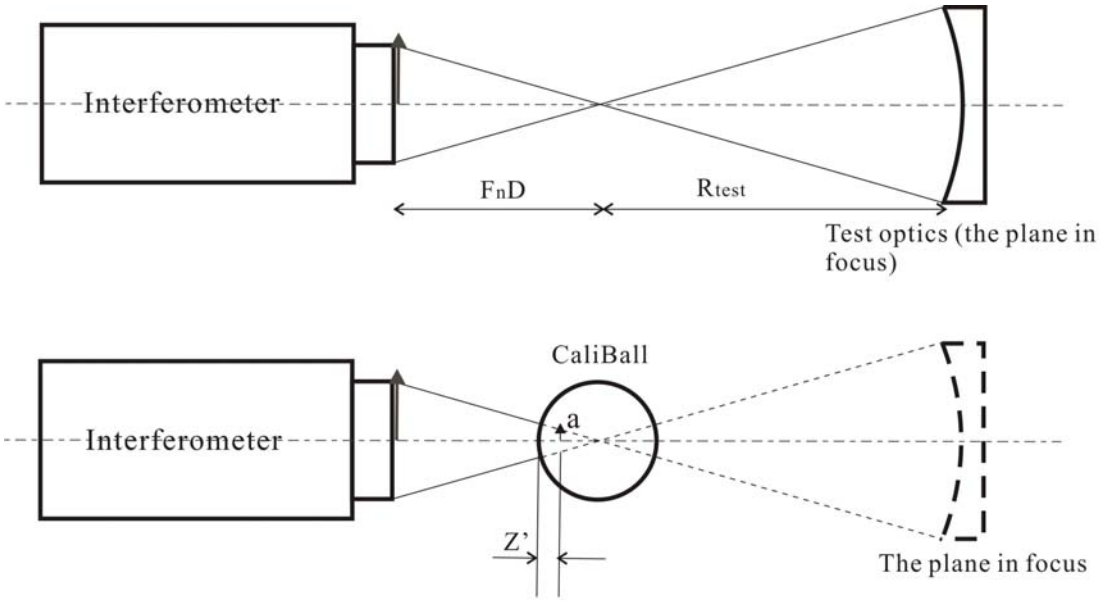


Figure **3.12**: Propagation of the common wavefront in the surface measurement (top) and in the RBT (bottom). The arrows represent the common wavefront and its conjugate through the CaliBall imaging.

Assume that the source of error in the common wavefront is at the transmission sphere. In the surface measurement, as this wavefront propagates to the test optic, its phase is smoothed, and it can be estimated using the effective propagation distance and Fresnel number discussed in Chapter 2. The effective propagation distance from the transmission sphere to the test optic L_{test} and the corresponding Fresnel number $N_{\text{f\_test}}$ are

$$\begin{aligned} L_{\text{test}} &= \frac{F_n D \cdot (F_n D - R_{\text{test}})}{R_{\text{test}}} \\ N_{\text{f\_test}} &= \frac{D^2}{4\lambda L_{\text{test}}} = \frac{D \cdot R_{\text{test}}}{4\lambda F_n (F_n D - R_{\text{test}})} \end{aligned}, \qquad (\mathbf{3.8})$$

where F_n and D are the F-number and the diameter of the transmission sphere respectively, and R_{test} is the radius of curvature of the test optic. Note that R_{test} is negative for concave surfaces. Therefore, the transfer function becomes

$$TF_{\text{test}} = \cos\left(\frac{\pi f_{\text{normalized}}^2}{4N_{\text{f\_test}}}\right) = \cos\left(\frac{\pi\lambda F_n f_{\text{normalized}}^2 (F_n D - R_{\text{test}})}{DR_{\text{test}}}\right), \quad \textbf{(3.9)}$$

where $f_{\text{normalized}}$ is the normalized frequency with units of cycles per diameter. If $|R_{\text{test}}| >> F_n D$ ($F_n D$ equals to the radius of the reference surface), the transfer function can be simplified as

$$TF_{\text{test}} = \cos\left(\frac{\pi\lambda F_n f_{\text{normalized}}^2}{D}\right). \quad \textbf{(3.10)}$$

In the RBT, the image of the common wavefront through the CaliBall is at distance z' from the vertex of the CaliBall:

$$\begin{aligned} \frac{1}{z'} &= \frac{1}{-F_n D + R_{\text{CaliBall}}} + \frac{-2}{R_{\text{CaliBall}}} \\ z' &= \frac{(R_{\text{CaliBall}} - F_n D) R_{\text{CaliBall}}}{2F_n D - R_{\text{CaliBall}}} \end{aligned}. \quad \textbf{(3.11)}$$

The effective propagation distance L_{RBT} and the image size inside the ball $2a$ (in diameter) are

$$\begin{aligned} L_{\text{RBT}} &= \frac{(R_{\text{CaliBall}} + z' - R_{\text{test}}) \cdot (R_{\text{CaliBall}} + z')}{R_{\text{test}}} \\ 2a &= \frac{R_{\text{CaliBall}} + z'}{F_n} \end{aligned}, \quad \textbf{(3.12)}$$

where R_{CaliBall} is the radius of curvature of the CaliBall. Then the Fresnel number in the RBT is

$$N_{\text{f\_RBT}} = \frac{a^2}{\lambda L_{\text{RBT}}} = \frac{R_{\text{CaliBall}} R_{\text{test}} D}{4F_n \lambda \left[F_n D(R_{\text{CaliBall}} - 2R_{\text{test}}) + R_{\text{CaliBall}} R_{\text{test}} \right]}. \quad \textbf{(3.13)}$$

So the transfer function of the common wavefront in the RBT becomes

$$TF_{\text{RBT}} = \cos\left(\frac{\pi f_{\text{normalized}}^2}{4N_{\text{f\_RBT}}} \right) = \cos\left(\frac{\pi \lambda F_n f_{\text{normalized}}^2 \left[F_n D(R_{\text{CaliBall}} - 2R_{\text{test}}) + R_{\text{CaliBall}} R_{\text{test}} \right]}{D R_{\text{CaliBall}} R_{\text{test}}} \right). \quad \textbf{(3.14)}$$

When $\left| R_{\text{test}} \right| > F_n D R_{\text{CaliBall}}$, Equation **3.14** can be simplified as

$$TF_{\text{RBT}} = \cos\left(\frac{\pi \lambda F_n f_{\text{normalized}}^2 \left[2F_n D - R_{\text{CaliBall}} \right]}{D R_{\text{CaliBall}}} \right). \quad \textbf{(3.15)}$$

If $2F_n D >> R_{\text{CaliBall}}$, then Equation **3.15** can be further simplified to

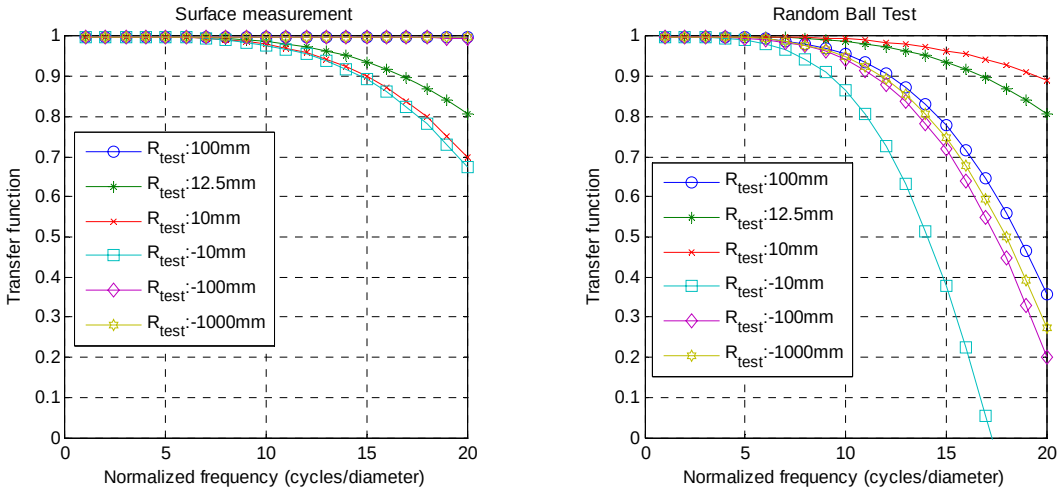
$$TF_{\text{RBT}} = \cos\left(\frac{2\pi \lambda F_n^2 f_{\text{normalized}}^2}{R_{\text{CaliBall}}} \right). \quad \textbf{(3.16)}$$

A sensitivity function, defined as the difference between the transfer functions of the surface measurement and the RBT, gives the measurement error due to the diffraction effect as a function of spatial frequency of the common wavefront, when using the RBT to calibrate the interferometer:

$$\Delta TF = \cos\left(\frac{\pi f_{\text{normalized}}^2}{4N_{\text{f\_test}}} \right) - \cos\left(\frac{\pi f_{\text{normalized}}^2}{4N_{\text{f\_RBT}}} \right). \quad \textbf{(3.17)}$$

It is a function of F-number, diameter of transmission sphere, the CaliBall size and the radius of curvature of the test optic.

Figure **3.13** shows the transfer functions of the common wavefront as it propagates from the transmission sphere to the test optic (plane of focus) in the surface measurement and the RBT. For the RBTs, the focus is still set at nominal test. The six lines in each plot represent different radii of the test surface from 100 mm to -1000 mm, respectively. The negative sign means the test surface is concave and a positive number means a convex surface. The smoothing of the common wavefront in the surface measurement is small when the radius of curvature of the test surface is large, for both concave and convex surfaces. The RBT generally suffers from phase smoothing more than from the surface measurement.



a) b)

Figure **3.13**: Transfer functions vs. spatial frequency in a) the surface measurement, b) the RBT test. A 25 mm CaliBall is used to calibrate an F/3.2 transmission sphere with a beam diameter of 150 mm. The test optic has a radius of curvature of 100 mm, 12.5 mm (same as CaliBall radius), 10 mm, -10 mm, -100 mm and -1000 mm respectively.

Figure **3.14** gives the sensitivity function which is the difference between Figure **3.13 (a)** and Figure **3.13 (b)**, and it shows that the measurement errors due to diffraction are small when the frequency of phase ripples is less than 10 cycles/diameter. For a concave test surface, the change in sensitivity function decreases dramatically as the radius of the test surface becomes larger. For a convex surface, the errors due to phase smoothing cancel out when the test optic has the same radius of curvature as the CaliBall. At spatial frequencies higher than those shown in Figure **3.13**, the transfer functions oscillate between 1 and -1 due to cosine dependence. Therefore, the sensitivity functions oscillate between 2 and -2 at high spatial frequencies.

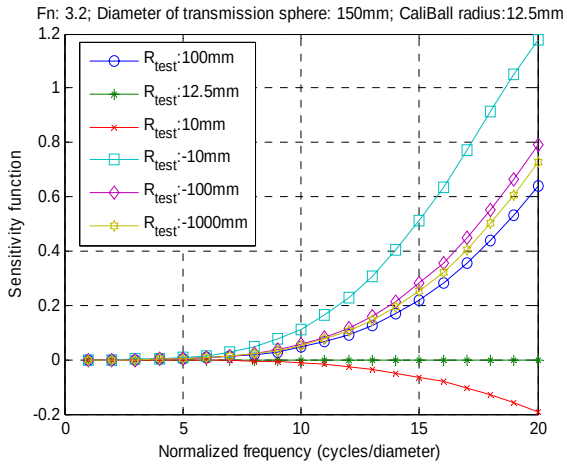


Figure **3.14**: The sensitivity function, defined as the difference between Figure **3.13 (a)** and Figure **3.13 (b)**, gives the measurement error due to diffraction as a function of the spatial frequency of the common wavefront, when using the RBT to calibrate the interferometer.

At spatial frequencies of 10, 15 and 20 cycles/diameter, the sensitivity function versus the radius of curvature of the test surface is plotted in Figure **3.15**. It shows that as the radius of curvature increases for both concave and convex surfaces, the sensitivity

function is approaching an asymptote. For small radii of curvature, the sensitivity increases dramatically except for the radius of curvature that matches the CaliBall radius, where the errors due to diffraction become zero. When the radius of curvature of the test surface is large, the horizontal asymptote of the sensitivity function can be calculated using Equation **3.10** and Equation **3.15**

$$\lim_{R_{\text{test}}\to\infty}\Delta TF=\cos\left(\frac{\pi\lambda F_n f_{\text{normalized}}^2}{D}\right)-\cos\left(\frac{\pi\lambda F_n f_{\text{normalized}}^2\left[2F_nD-R_{\text{CaliBall}}\right]}{DR_{\text{CaliBall}}}\right). \quad (\mathbf{3.18})$$

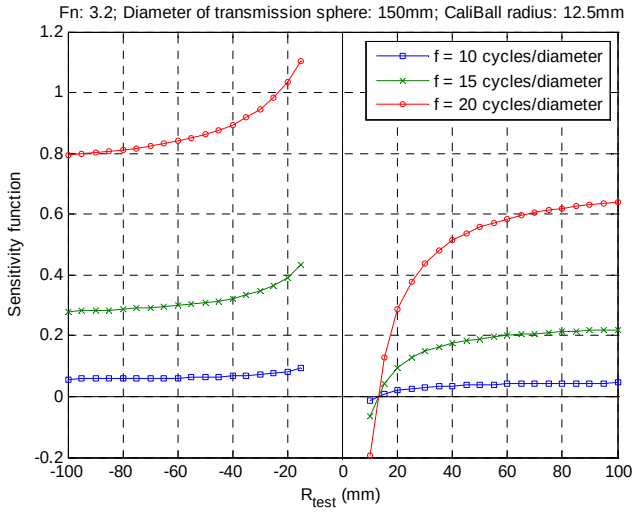


Figure **3.15**: The sensitivity functions versus the radius of curvature of the test surface at spatial frequencies of 10, 15 and 20 cycles/diameter.

To investigate the effect of F-number and CaliBall size on the phase attenuation, we assume that the test optic is concave with a radius of curvature of 1000 mm and the diameter of the transmission sphere is 150 mm. Figure **3.16** gives the sensitivity function

at different F-numbers of the transmission sphere, and it shows that faster transmission spheres are less susceptible to error from the common wavefront.

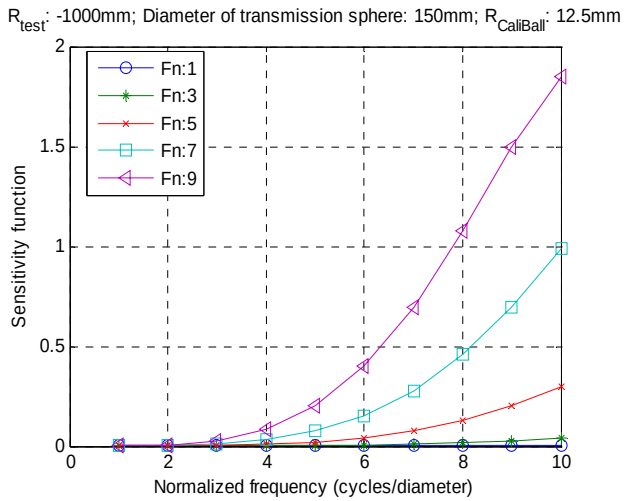


Figure **3.16**: The sensitivity functions vs. spatial frequencies at different F-numbers. The fast transmission spheres are less susceptible to error from the common wavefront.

The sensitivity function versus spatial frequency with different CaliBall sizes is plotted in Figure **3.17**. The F-number is 3.2 for all three curves. It shows that larger CaliBalls generally introduce less error, so it is better to choose a larger diameter CaliBall for calibration if the CaliBall's radius of curvature cannot match that of the test surface.

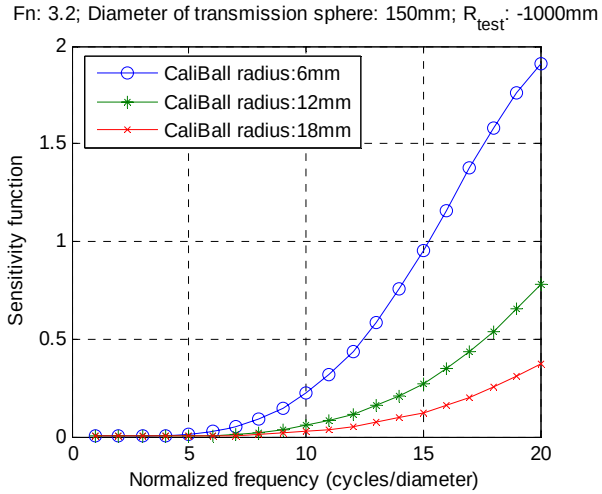


Figure **3.17**: The sensitivity functions vs. spatial frequency with different CaliBall sizes. Larger CaliBalls generally introduce less error due to diffraction.

The diffraction errors from the common wavefront depend on the radii of curvature of the test surface and the CaliBall, and the F-number. In general, it is true that

- A concave test optic with a larger radius of curvature will have less error introduced from the common wavefront; a convex test optic with the same radius of curvature as the CaliBall will have no errors due to diffraction.
- A faster system will have less error from the common wavefront.
- A larger diameter CaliBall has smaller wavefront errors when the radius of curvature of the test surface is much larger than that of the CaliBall.

Phase smoothing only becomes noticeable when the phase ripple in the common wavefront has high spatial frequency content. Usually, the wavefront coming out of an interferometer is dominated by low spatial frequencies. Even if there are some high

frequency components, they are often small in magnitude. Therefore, errors due to phase smoothing are usually small.

Assume that there are 10 nm rms errors at 8 cycles/diameter in the common wavefront, the diameter of the transmission sphere is 6 inches, and the radius of curvature of the test optic is concave and much larger than the radius of the reference sphere ($F_n D$). The measurement errors due to phase smoothing at different F-numbers and CaliBall sizes are shown in Figure **3.18**. In the case of F/4 and 12 mm radius CaliBall, the RMS error due to phase smoothing is about 0.5 nm.

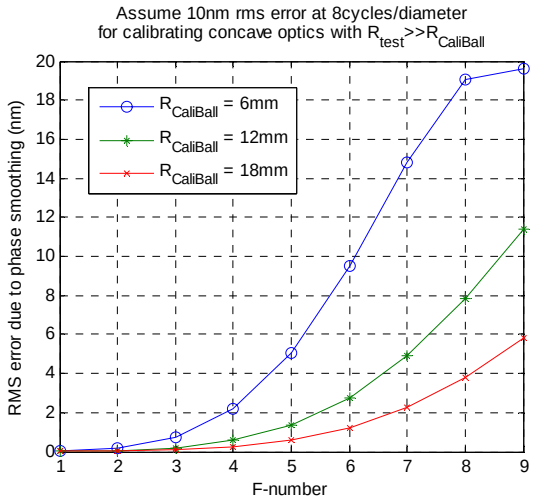


Figure **3.18**: The measurement error due to phase smoothing at various F-number and CaliBall sizes. These plots assume there are 10 nm rms errors at 8 cycles/diameter in the common wavefront, the diameter of the transmission sphere is 6 inches, and the radius of curvature of the test optic is concave and much larger than the radius of the reference sphere.

3.2.3.2 Edge diffraction for RBTs

Edge diffraction is seen when the aperture is not in focus. There are many apertures inside interferometers that cannot be in focus since interferometers often image the surface under test on the detector to correctly represent the errors in the test surface. However, edge diffraction from the limiting aperture usually has the most dominant effect in a measurement. In RBTs, one can often see the diffraction rings around the edge that are caused by the diffraction of the aperture of the transmission sphere.

The effect of edge diffraction was simulated using MATLAB. The aperture of the transmission sphere is imaged through the CaliBall at a position inside the CaliBall, then the diffraction pattern is calculated as a spherical wavefront propagating from that position to the test optic (plane of focus). We simulated this edge diffraction for the 6" F/3.2 and F/7 transmission spheres, for the case that the interferometer focuses at the test optic which is concave and has a radius of curvature of 1000 mm. The results are as shown in Figure **3.19**. Similar to phase smoothing, the slower transmission sphere has a larger RMS error due to edge diffraction. The RMS phase errors due to edge diffraction for the F/3.2 and F/7 transmission spheres are 5.5 nm and 8.1 nm, respectively.

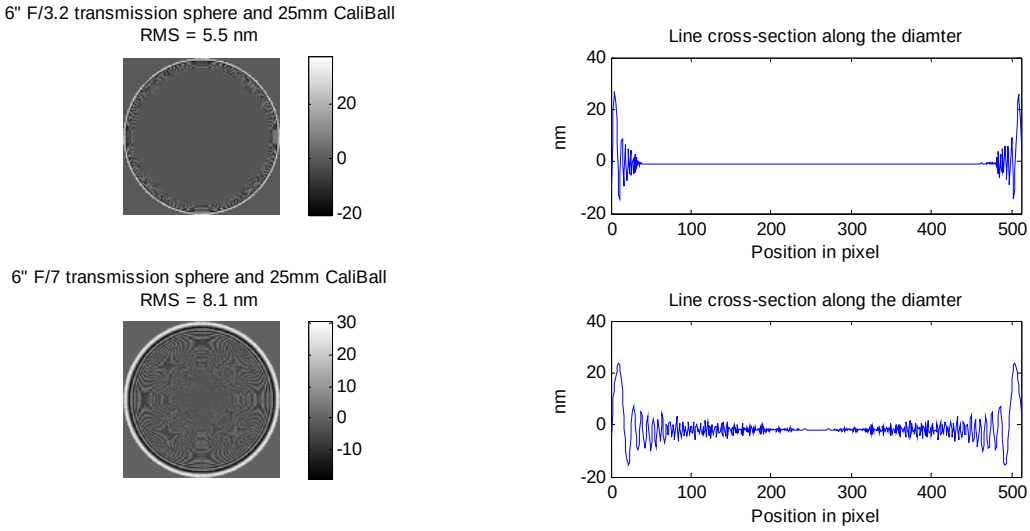


Figure **3.19**: Phase pattern due to edge diffraction (left) for the 6" F/3.2 and F/7 transmission spheres and their line profile (right). The interferometer focuses at the test optic which is concave and has a radius of curvature of 1000 mm.

If the test optic is overfilled by the beam coming out of interferometer, then only the corresponding portion of the calibration map and the edge diffraction are used. Assuming that the test optic uses 80% diameter of the interferometer, the RMS errors due to edge diffraction at various F-numbers and CaliBall sizes are given in Figure **3.20**. With an F/3.2 transmission sphere and a 25 mm CaliBall, the error from edge diffraction is about 0.2 nm rms.

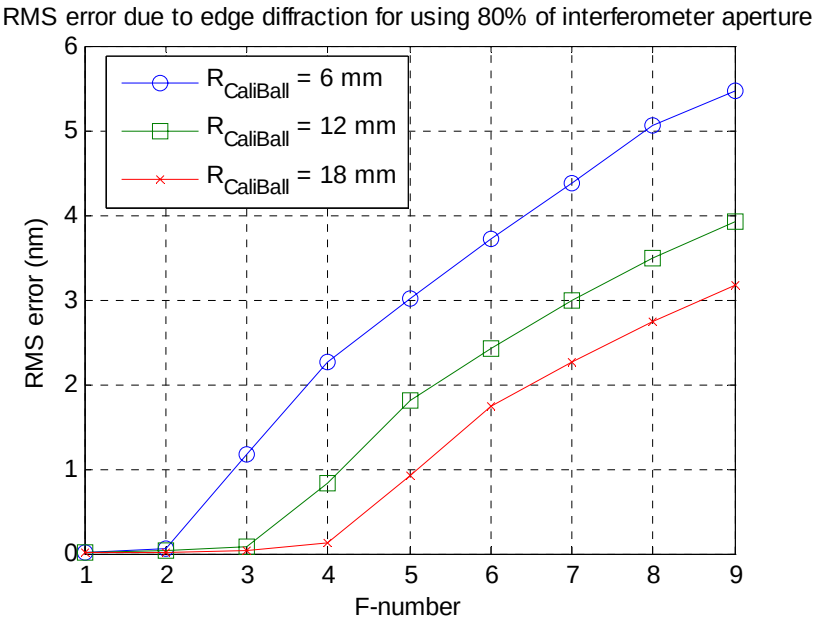


Figure **3.20**: RMS error due edge diffraction for using 80% of interferometer aperture. The diameter of transmission sphere is 6". The interferometer focuses at the test optic which is concave and has a radius of curvature of 1000 mm.

3.2.4 Example of estimating the remaining errors in the RBT

The total measurement uncertainty for the surface under test should include the measurement noise, geometric errors and diffraction errors in both the RBT and the surface measurement. If a 25 mm CaliBall is used to calibrate a 6" F/3.2 transmission sphere and the test part has a radius of curvature of 1000 mm, a reasonable breakdown of all these errors is listed in Table **3.3**. The interferometer is configured to image the test surface. Errors in the RBT are given based on the experimental data. Errors in the surface measurement depend on the part itself and the measurement environment. For the surface measurement, assuming 3 nm rms noise per measurement, then an average of

25 maps can reduce the noise to 0.6 nm rms. If the sensitivity functions of the RBT and the surface measurement are the same and the test surface has similar slope errors, the geometric error in the surface measurement will be similar to that of the RBT. The same amount of geometric error is assigned to the surface measurement. Note that the sensitivity function of the surface measurement usually is different from that of the RBT. If the common wavefront has 10 nm rms error at 8 cycle/diameter, then the error due to phase smoothing is about 0.2 nm rms. The error due to edge diffraction is 0.2 nm rms within 80% of the aperture. Given the error budget shown in Table **3.3**, the total measurement uncertainty is 1.3 nm rms.

Table **3.3**: The total measurement uncertainty for the surface under test if the interferometer is calibrated with a CaliBall.

	Error sources	**Error budget**
Random Ball Test	Measurement noise	0.5 nm rms
	Geometric error	0.7 nm rms
Surface Measurement	Measurement noise	0.6 nm rms
	Geometric error	0.7 nm rms
Diffraction Effect	Phase smoothing	0.2 nm rms
	Edge diffraction (80% aperture)	0.2 nm rms
RSS error		1.3 nm rms

3.3 Conclusion

This chapter first reviewed different interferometer calibration methods. Then the remaining errors in the RBT are extensively discussed, including the random errors and

systematic errors. The tools to quantify these errors are provided, and these tools can also be used to study the errors in other interferometric measurements. The RBT is a great test and it can calibrate the low spatial frequency errors to high accuracy by averaging enough measurements. Some errors, like geometric errors and diffraction effects, cannot be calibrated by the RBT when the radius of curvature of the test optic is different from the CaliBall radius. The geometric errors are sensitive to the wavefront slope, and can be estimated by taking some additional measurements to find the sensitivity functions. The diffraction effect becomes significant for high spatial frequency components. The diffraction effects are dominant for a slow transmission sphere and a small CaliBall. An example of estimating the remaining errors after interferometer calibration using the RBT is given in the end and it shows that the total error is about 1.3 nm rms.

CHAPTER 4

INTERFEROMETER PHASE MODULATION TRANSFER FUNCTION

In previous chapters, a number of intrinsic interferometer errors are discussed, including geometric errors and errors due to diffraction. Another intrinsic interferometer error is its phase modulation transfer capability, quantified by its modulation transfer function (MTF) [**de Groot 2006 and Chu 2008**]. This chapter discusses a method of calculating the interferometer MTF, and the use of a Wiener filter to recover the middle spatial frequency content in the measurement.

4.1 Introduction

The MTF provides a common method used to characterize the performance of an optical imaging system. It is the ratio of image contrast to object contrast as a function of spatial frequency. Special test targets with high contrast patterns of various spatial frequencies, such as Siemens star, are used to measure the modulation transfer function. Analogous to the MTF of imaging systems, the phase response of interferometers as a function of spatial frequency can be determined with a phase target. The phase MTF of an interferometer describes the effect of spatial frequencies on the phase measurement. Typically, an interferometer can measure the low spatial frequency figure errors accurately, while the high spatial frequency errors, like surface roughness, are measured with an interference microscope. Recently, the need for measuring the middle spatial

frequency content has increased, requiring the interferometer transfer function to be characterized for more accurate measurement. The middle spatial frequency errors will be modulated by the interferometer MTF, and the phase magnitude will become smaller than it actually is. Surface errors with a spatial frequency higher than the cutoff frequency of the interferometer MTF cannot be measured at all. The interferometer MTF can be affected by many parameters like the imaging optics inside the interferometer, the response of a detector, the light source, and so on.

4.2 MTF Measurement of an infrared interferometer

Unlike the test targets used for measuring the modulation transfer function of an optical imaging system, a target for measuring the interferometer MTF needs to have a known phase step with various spatial frequencies. There are several methods to determine the interferometer MTF, like three beam interference or measurement of a known phase step [**Novak 1997**]. We made a phase Siemens star with a uniform phase step of 0.55 μm as shown in Figure **4.1**. The phase pattern has a period of 2 mm at the edge. This phase plate is used in reflection to measure the MTF of a Wyko IR (CO_2 10.6 μm) Twyman-Green interferometer.

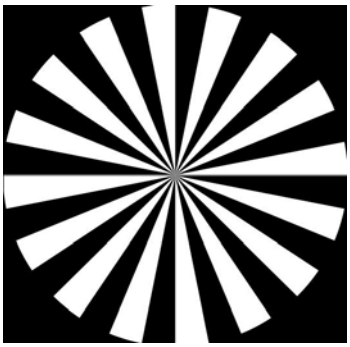
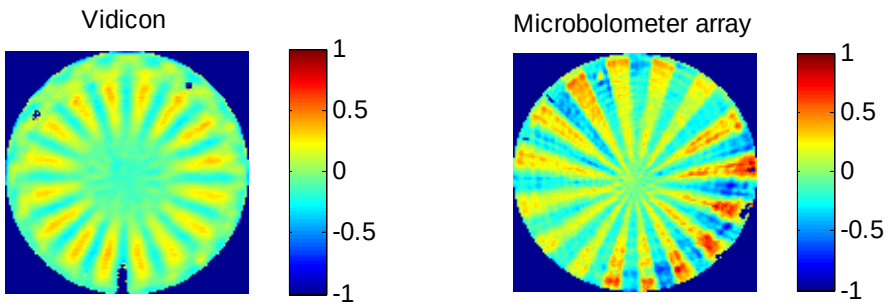


Figure **4.1**: Phase Siemens star with a uniform phase step of 0.55 μm. The period at the edge is about 2 mm.

The interferometer MTF can be determined by comparing the Fourier Transform of an actual measurement with that of the input phase at various spatial frequencies. Since the input phase is a square wave around a certain radius of the phase Siemens star, a fundamental frequency ξ_0 is used to calculate the MTF at that specific frequency, which is [**Toronto 1995**]

$$\mathrm{MTF}(\xi_0) = \left.\frac{\Im(\text{image})}{\Im(\text{object})}\right|_{\text{at fundamental frequency } \xi_0} . \qquad (\mathbf{4.1})$$

This IR interferometer initially had a Vidicon camera, which greatly limited the interferometer phase MTF; it was later replaced by a microbolometer array. In both cases, the interferometer outputs data of 640 × 480 pixels. The measurement results are illustrated in Figure **4.2**. For the Vidicon, the phase map at the edge is blurry. This may be due to the edge diffraction from some aperture.



a) b)

Figure **4.2**: Interferometric measurements of a phase Siemens star with a) Vidicon and b) micro-bolometer array.

Figure **4.3** shows the calculated MTF with the Vidicon camera and the microbolometer respectively. It is shown that the phase MTF improved dramatically with the microbolometer array.

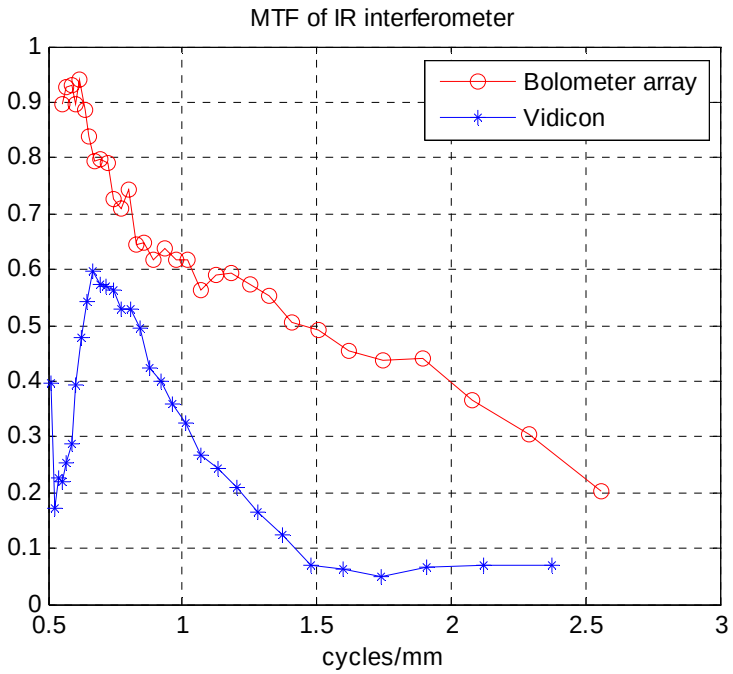


Figure **4.3**: Phase MTF measurements for a Wyko IR interferometer with a microbolometer array and Vidicon, respectively. The horizontal axis represents spatial frequency on the Siemens star.

4.3 Introduction to the Wiener filter

The interferometer MTF tells how the phase information is underestimated at a certain frequency. Once the MTF is known, some middle spatial frequency phase information can be recovered by applying a Wiener filter to the measurement. The Wiener filter is based on a statistical approach, and is a function of the MTF and the power spectra of the signal and noise. It can be defined as [**Frieden 2001**]

$$H(\xi,\eta)=\frac{\mathrm{MTF}(\xi,\eta)S_o(\xi,\eta)}{|\mathrm{MTF}(\xi,\eta)|^2\ S_o(\xi,\eta)+S_n(\xi,\eta)}, \qquad (\mathbf{4.2})$$

where ξ,η are spatial frequencies, and S_o and S_n are the power spectral densities (PSD) of the actual surface figure and measurement noise respectively.

Equation **4.2** shows that in the high noise region ($S_n >> S_o$), the Wiener filter approximates zero ($H(\xi,\eta)\approx 0$). Applying the Wiener filter to the measurement can suppress the data in this region. Conversely, the Wiener filter becomes an inverse filter ($H(\xi,\eta)\approx 1/MTF(\xi,\eta)$) in the low noise region ($S_n << S_o$), and can be applied to boost some information in this region. Interferometers usually have a high MTF in the low spatial frequency region and can measure the low frequencies accurately. In the high frequency region, they usually have low MTF and the measurement noise is often large. Therefore, the Wiener filter will not be able to boost much information in this region. It is in the middle spatial frequency region, where the interferometer MTF is not so high and the measurement noise is relatively low, that the Wiener filter can be used to "amplify" those spatial frequencies that were originally attenuated by the interferometer.

In figure measuring interferometry, the actual surface figure of the test optic is unknown. However, an average of many measurements $\overline{W}$, after calibrating the interferometer, can be a good estimate of the surface. The power spectrum of the surface then can be calculated as

$$S_o = |\Im(\overline{W})|^2 . \quad \textbf{(4.3)}$$

If each measurement is denoted as W_i, then the difference between each individual measurement and the average $(W_i - \overline{W})$ is an estimate of the noise. The power spectrum of noise can be approximated by an ensemble average of the power spectra of the noise in the individual measurements [**Barrett 2003**]. When the noise is uncorrelated from one measurement to another, the RMS noise in the averaged map is smaller by a factor of $1/\sqrt{N}$, where N is the number of measurements averaged. Therefore, the power spectrum of the noise in the average map $\overline{W}$ can be computed as

$$S_n = \frac{\left\langle |\Im(W_i - \overline{W})|^2 \right\rangle}{N} . \quad \textbf{(4.4)}$$

The Wiener filter is applied in Fourier domain. If the measured surface map $g(x, y)$ has Fourier Transform $G(\xi, \eta)$, then the best estimate of the actual surface $f(x, y)$ has Fourier Transform $F(\xi, \eta)$ given by

$$F(\xi, \eta) = G(\xi, \eta) \cdot H(\xi, \eta) . \quad \textbf{(4.5)}$$

An inverse Fourier Transform of $F(\xi, \eta)$ gives the reconstructed surface map $f(x, y)$.

The difficulty of using the Wiener filter is to estimate the power spectrum of "noise" completely. In Equation (**4.4**), the noise is from measurement non-repeatability, which changes from one measurement to another. There are also errors that are repeatable in the measurement like diffraction effects, dust and so on, which are not included in Equation (**4.4**). The Wiener filter will boost the information in the middle spatial frequencies, but may also amplify some errors due to incorrect estimate of the noise.

4.4 Application of the Wiener filter

This section gives an example of applying the Wiener filter to the measurements taken with the IR interferometer which is characterized by the phase Siemens star. The interferometer with the microbolometer array was used to test the Discovery Channel Telescope (DCT) primary mirror during the grinding stage, which has a radius of curvature of 16 m and is 4.3 m in diameter. An average of 68 measurements is given in Figure **4.4**.

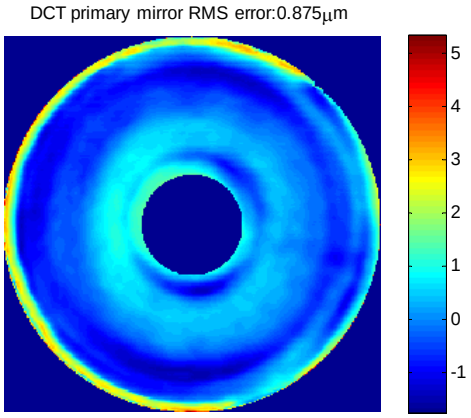
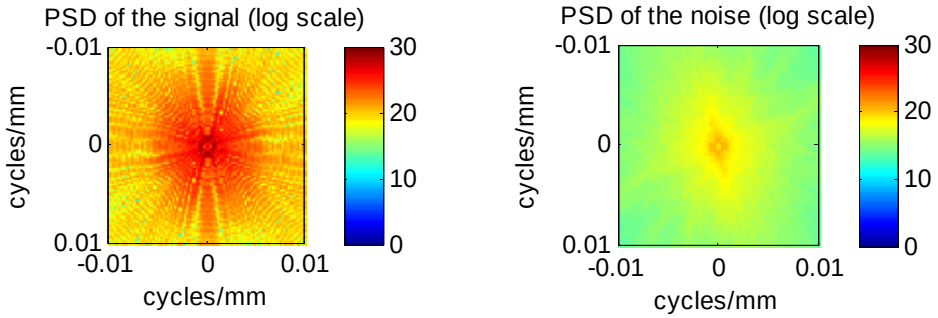


Figure **4.4**: An average of 68 maps for DCT primary mirror tested with an IR interferometer. The units of the scale are μm.

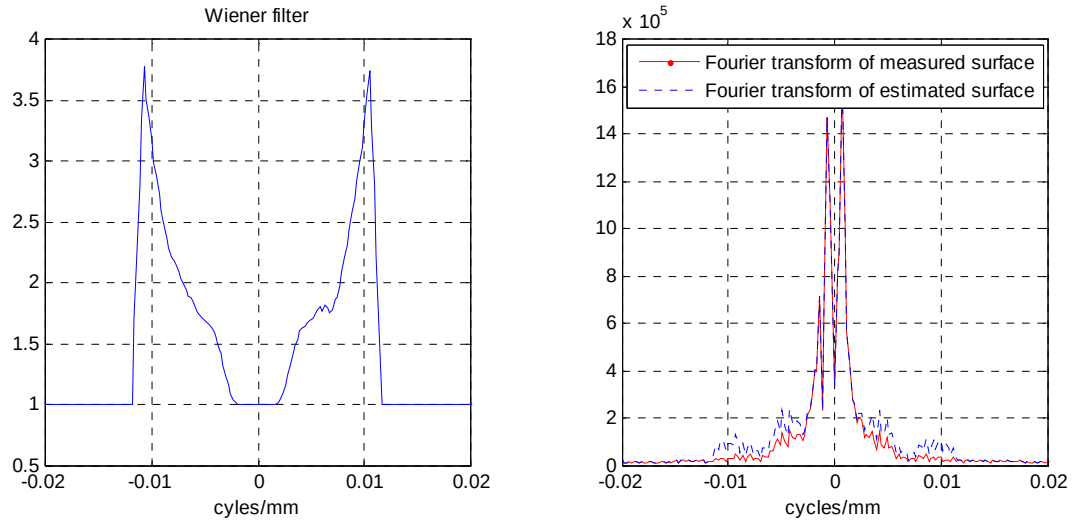
This average map was used to determine the power spectral density of the surface as shown in Figure **4.5 (a)**, and the difference maps between this average map and each individual map are used to calculate the power spectrum of the noise shown in Figure **4.5 (b)**.



a) b)

Figure **4.5**: a) PSD of the average DCT primary mirror $S_o(\xi,\eta)$; b) PSD of the noise for the DCT primary mirror measurements $S_n(\xi,\eta)$. The PSDs have units of nm^2mm^4.

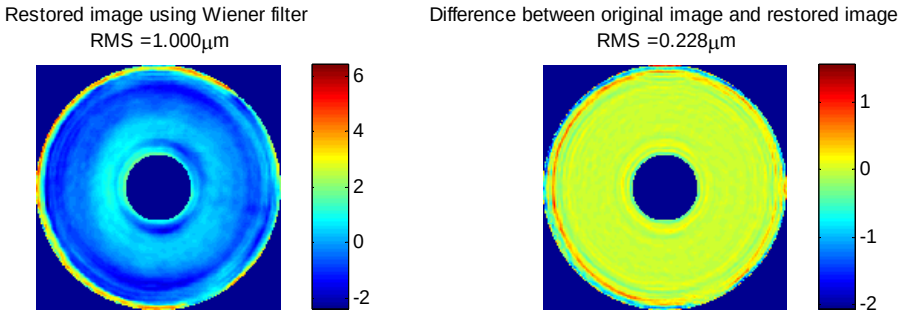
A cross-section of the 2D Wiener filter used to reconstruct the surface map is shown in Figure **4.6 (a)**. When the spatial frequency is below 0.002 cycles/mm and above 0.01 cycles/mm, the Wiener filter is unity which means the surface figure error in these frequency regions remains. The actual calculated Wiener filter is below unity when the spatial frequency is less than 0.002 cycles/mm, because there is low order measurement noise due to air turbulence. It is set to unity here in order to contain the actual low spatial frequency figure on the mirror. In the region that the spatial frequency is above 0.01 cycles/mm, the Wiener filter is unity since the interferometer MTF is only measured up to this frequency. Phase ripples with spatial frequencies between 0.002 and 0.01 cycles/mm (corresponding to 500 mm to 100 mm on the mirror) will be amplified by the Wiener filter. The Fourier Transforms of the measured average map and reconstructed map are shown in Figure **4.6 (b)**. It is clearly shown that the middle spatial frequency content is amplified by the Wiener filter.



a) b)

Figure **4.6**: a) A cross-section of the calculated 2D Wiener filter; b) The Fourier transforms of the measured average map and reconstructed map. The horizontal axis represents spatial frequency on the mirror.

The restored map after applying this Wiener filter to the average map is shown in Figure **4.7 (a)**, and the difference between the restored map and the original map is shown in Figure **4.7 (b)**. We had planned to compare the restored surface map using the Wiener filter with the visible interferometer test, which has a much better phase MTF and can provide more accurate measurement. Unfortunately, the full-aperture surface could not be resolved due to the high slopes on the mirror when we first switched from the IR interferometer to a visible one. Therefore, the Wiener filter method could not be demonstrated with experiments. This should be included in the future work.



a) b)

Figure **4.7**: a) The restored map after applying the Wiener filter; b) The difference between the restored map and the original map.

4.5 Conclusion

The example of the DCT primary mirror tested with the infrared interferometer demonstrates two things: first, the Wiener filter can be used to recover the middle spatial frequency information which is originally attenuated by the interferometer; and second, it is important to characterize the interferometer MTF if high measurement accuracy is required. This method shows great potential, but it has not been demonstrated adequately here. The fundamental difficulty of quantifying the noise provides a limitation. Developing a comprehensive technique for assessing noise and using it to develop optimal filtering for interferometer data remains an active research topic.

CHAPTER 5

MAP REGISTRATION AND SUBTRACTION

It is common in interferometric surface measurement to have two maps tested in different systems and these two maps need to be subtracted one from the other. Correctly registering the two maps is important because incorrect registration between two maps will induce errors when subtracting one map from the other. This chapter presents a method to reduce the errors due to misregistration. Section **5.2** discusses the map registration methods in optical testing. Section **5.3** presents the polynomial functions that are often used to describe the mapping, and Section **5.4** discusses the smoothing filter method, which can reduce the sensitivity to registration error and improve the overall measurement accuracy.

5.1 Introduction

In optical testing, it is common to use reference optics to test other optics. For instance, a reference sphere is used to test a flat in the Ritchey-Common Test [**Han 2003**]. To accurately measure a surface, errors in the reference optic must be characterized and removed from measurement data of the test optic. The removal of the errors from the reference optic requires two maps. The map including the test surface information is called the test map, while the other map is called the reference map. These two maps, which may be offset by translation, rotation, scale, or other complex distortion,

must be precisely mapped to each other for subtraction. Any error in mapping will couple with errors of the reference map, causing an apparent error in the test surface.

5.2 Map registration

If a reference map is to be subtracted from a test map but the mapping is incorrect, the result will contain some errors. For this reason, the mapping must be determined as accurately as possible. The process of defining the mapping is known as map registration. References placed at the surface, called fiducials, or high-frequency features on a surface are often used to facilitate accurate map registration in optical testing.

5.2.1 Fiducials

The use of fiducials is quite common in optical testing. Fiducials can be used for map registration or distortion correction. For map registration, one map is transformed so that its fiducials are aligned with those of the other map. For distortion correction, fiducial marks are placed at known positions on the surface, and then the positions of the image of these marks are measured to determine the mapping distortion. The number of fiducials depends on the initial distortion and the required mapping accuracy. Fiducials can be created by a mask attached to the surface or actual markers placed on the surface. An example of fiducial mask is given in Figure **5.1**.

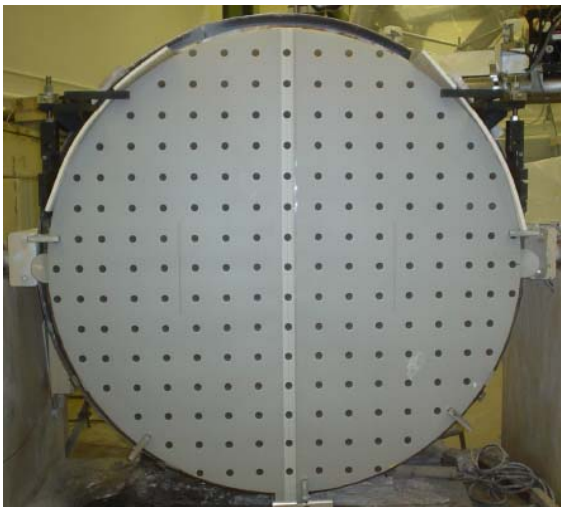


Figure **5.1**: An example of a fiducial mask with regularly spaced holes.

5.2.2 High frequency features

The other approach is to use high-frequency features on a surface for registration. These features take the form of small-scale surface imperfections. They are very stable, and can be brought out through the use of a high-pass filter. The reason for using high-frequency information is that these small-scale effects will remain constant in the presence of bending or flexing of the mirror substrate, while the low-frequency effects may vary from one test to another due to alignment or mechanical distortions. Usually, only a subset of the map is used for registration because it can reduce the time for optimizing the mapping [**Siegel 2005**].

5.3 Mapping functions

Fiducials and high frequency features are used to determine the mapping through some optimization algorithm, like a least-squares fit. The mapping can be described by polynomial functions of x and y as follows:

$$\begin{aligned} X &= \sum u_{ij} x^i y^j \\ Y &= \sum v_{ij} x^i y^j \end{aligned}, \qquad (\mathbf{5.1})$$

where X and Y are the normalized coordinates in the test map, x and y are the normalized coordinates in the reference map, and u_{ij} and v_{ij} are the coefficients of the polynomials.

These polynomial functions work fine when the mapping distortion is small and only lower-order polynomials are required to find the mapping functions. However, some optical systems cause interferometric maps to become highly distorted. When the mapping distortion is complex, it is better to use polynomials that are orthonormal over a unit circle. Recently, Zhao and Burge developed a set of vector polynomials $\vec{S}_j$, which are orthonormal over a unit circle, to represent vector quantities such as mapping distortion or wavefront gradient [**Zhao 2007b**]. Those functions are derived from the gradients of Zernike polynomials $\vec{\nabla} Z_j$. The Zernike polynomials are listed in Appendix B, according to Noll's notation and numbering [**Noll 1976**]. The $\vec{S}$ polynomials are ordered by an index j (which is a function of n and m). The integers n and m describe the power of radial coordinate and angular frequency, respectively.

For all j with $n = m$,

$$\vec{S}_j = \frac{1}{\sqrt{2n(n+1)}} \vec{\nabla} Z_j . \qquad \textbf{(5.2)}$$

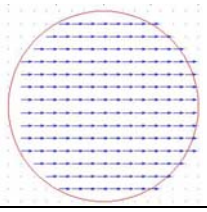
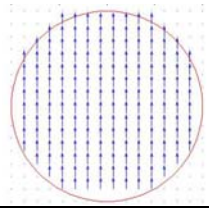
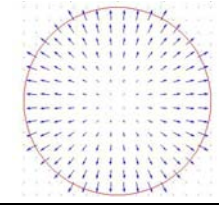
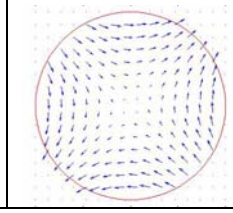
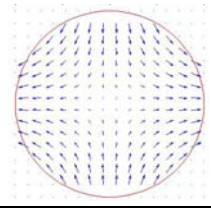
For all j with $n \neq m$,

$$\vec{S}_j = \frac{1}{\sqrt{4n(n+1)}} \left(\vec{\nabla} Z_j - \sqrt{\frac{n+1}{n-1}} \vec{\nabla} Z_{j'(n'=n-2,m'=m)} \right), \qquad \textbf{(5.3)}$$

where $j - j'$ is even when $m \neq 0$.

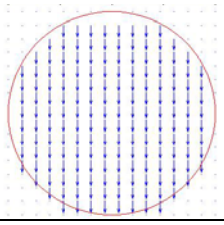
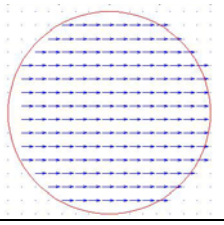
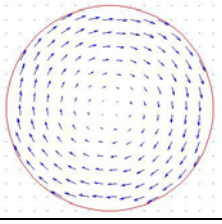
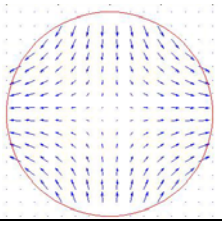
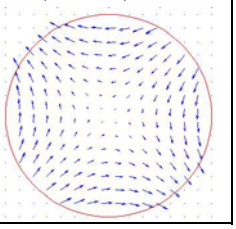
The plots of the first five $\vec{S}$ polynomials are shown in Table **5.1**. The first three polynomials vectors represent x translation and y translation, and scaling.

Table **5.1**: Plots of the first five $\vec{S}$ polynomials. The $\vec{S}$ polynomials are orthonormal over a unit circle and can be used to represent mapping distortion or wavefront gradient.

$\vec{S}_2$	$\vec{S}_3$	$\vec{S}_4$	$\vec{S}_5$	$\vec{S}_6$
				

The $\vec{S}$ polynomials are not a complete set to describe all the mapping distortion. For example, there is no $\vec{S}$ polynomial that can describe rotation, which is also one of the lowest modes for mapping. Zhao and Burge derived another set of $\vec{T}$ polynomials, which have non-zero curl. $\vec{T}$ polynomials are also orthonormal over a unit circle and there is some overlap between the two sets. Combining the $\vec{S}$ and $\vec{T}$ polynomials, we obtain a complete set [**Zhao 2008**]. The plots of the first five $\vec{T}$ polynomials are shown in Table **5.2**. $\vec{T}_4$ represents rotation, and the other terms overlap with $\vec{S}$ polynomials.

Table **5.2**: Plots of the first five $\bar{T}$ polynomials. $\bar{T}_4$ represents rotation, and the other terms overlap with $\bar{T}$ polynomials.

$\bar{T}_2(=-\bar{S}_3)$	$\bar{T}_3(=\bar{S}_2)$	$\bar{T}_4$	$\bar{T}_5(=\bar{S}_6)$	$\bar{T}_6(=-\bar{S}_5)$
				

The advantages of using these $\bar{S}$ and $\bar{T}$ polynomials are that they are orthonormal and each term has physical meaning [**Zhao 2009a**]. Unlike the polynomial functions in Equation **5.1**, the coefficient of each polynomial term remains the same as the number of terms used to fit the mapping increases.

5.4 Map subtraction – smoothing filter method

Imperfect mapping of the wavefront being subtracted will cause an error in the resulting data. To first order, the mapping error $\bar{\varepsilon}(x,y)$ and the gradient of the wavefront to be subtracted $\bar{\nabla}W(x,y)$ are coupled together and cause an error ΔW_{map} in the test surface, which can be described as

$$\Delta W_{\text{map}} = \bar{\nabla}W(x,y)\cdot\bar{\varepsilon}(x,y) = \frac{\partial W(x,y)}{\partial x}\cdot\varepsilon_x(x,y) + \frac{\partial W(x,y)}{\partial y}\cdot\varepsilon_y(x,y). \qquad \textbf{(5.4)}$$

In Equation **5.4**, $\varepsilon_x(x,y)$ and $\varepsilon_y(x,y)$ are mapping errors in the x and y directions, and $\frac{\partial W(x,y)}{\partial x}$ and $\frac{\partial W(x,y)}{\partial y}$ are the corresponding slope errors of the reference surface. To minimize the surface error due to map subtraction, both the surface slope errors and the mapping errors should be minimized during fabrication and data processing.

Martin proposed a smoothing filter method to reduce the sensitivity of wavefront errors to mapping [**Martin 2008**]. This method requires smoothing the reference map to a certain resolution first, and then subtracting the smoothed version of the reference map $W_{\text{smooth}}(x,y)$ from the test map. The smoothed version of the reference map contains all significant structure of the reference map. The difference between the smoothed map and the full resolution map is the small-scale errors, denoted as $W_{\text{residual}}(x,y)$. Any real small-scale error in the reference map will appear in the test surface measurement. The total surface errors due to subtraction of a smoothed version of the reference map include

- measurement noise in the smoothed reference map, which can be estimated using the method described in Chapter 2;
- small-scale error from the reference map – the difference between the full-resolution map and the smoothed map of the reference map;
- error caused by incorrect mapping of the smoothed map, which is equal to the product of the slope error in the smoothed map and the mapping error.

Smoothing the reference map may reduce the measurement noise and the sensitivity to the mapping errors. The wavefront errors caused by incorrect mapping

depend on the slope of the map being subtracted. A smoother reference map has a smaller slope error and then induces less wavefront error. Moreover, interferometers usually have high signal-to-noise ratio (SNR) in the low spatial frequency measurement and low SNR in the high spatial frequency measurement. The small-scale error is of high spatial frequency, so it contains more noise. The use of a smoothed reference map can mitigate the propagation of noise from the measurement of the reference surface to the measurement of test surface.

When using fiducials or high frequency features to register the reference and test map, there will be errors in determining the coordinates in both maps and fitting error due to an insufficient number of polynomial terms. Those errors will result in a random mapping error $\vec{\varepsilon}(x, y)$ which is difficult to predict; however, the average mapping errors ε_x and ε_y in the x and y directions can be estimated. One way to estimate mapping errors is to look at the residual errors in fiducial coordinates (i.e., the difference between measured coordinates and the fit). That may over- or under- estimate the mapping errors depending on the noise in fiducial coordinates, but it should be close.

If the mapping error is uncorrelated with slope error, then the root-sum-square (RSS) error can be used to describe the total error due to mapping registration and subtraction, and it can be used as a merit function to optimize the smoothing filter size. The RSS error can be described as

$$\begin{aligned} \mathrm{RSS}^2 = &\left[\mathrm{RMS}(W_{\mathrm{noise}})\right]^2 + \left[\mathrm{RMS}(W_{\mathrm{residual}})\right]^2 + \ldots \\ &\left[\mathrm{RMS}\left(\frac{\partial W_{\mathrm{smooth}}}{\partial x}\varepsilon_x\right)\right]^2 + \left[\mathrm{RMS}\left(\frac{\partial W_{\mathrm{smooth}}}{\partial y}\varepsilon_y\right)\right]^2 . \end{aligned} \tag{5.5}$$

Here is an example of applying the smoothing filter to a map of a 60 inch reference sphere with a 16 m radius of curvature. This sphere is used as a reference mirror in another test, so errors from it need to be backed out. The measurement resolution of the reference is 640 × 640, shown in Figure **5.2 (a)**. It is an average of 70 maps, and the measurement noise in the average map is 2.0 nm rms. Assuming the average mapping errors are 2 pixels in both the *x* and *y* directions and no smoothing filter is applied to the reference sphere, the surface errors due to incorrect mapping are 1.8 nm and 1.7 nm in the *x* and *y* directions, respectively, as shown in Figure **5.2 (b)** and **(c)**. In this case, the RSS error is $\sqrt{2.0^2 + 1.8^2 + 1.7^2} = 3.2$ nm .

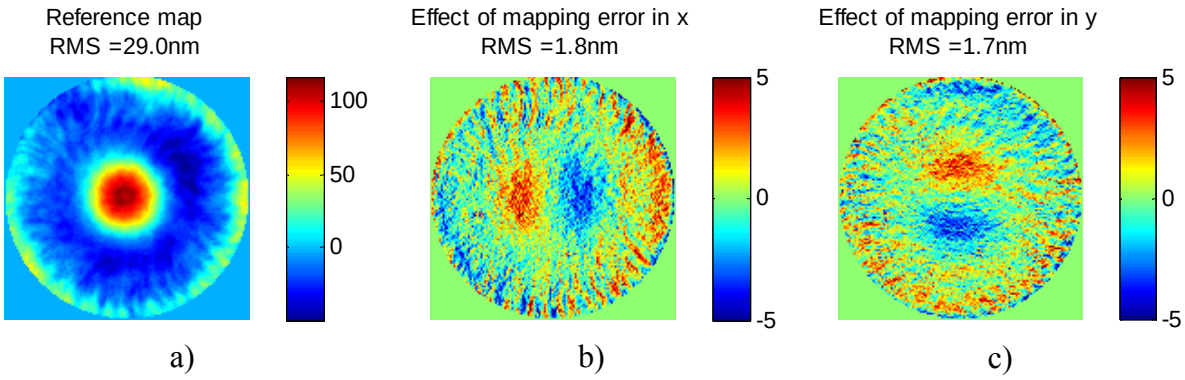


Figure **5.2**: a) Reference sphere; b) Wavefront error due to average 2-pixel mapping error in *x* direction without smoothing the reference sphere; c) Wavefront error due to average 2-pixel mapping error in *y* direction without smoothing the reference sphere.

When the reference sphere is smoothed with a 7 × 7 pixel filter shown in Figure **5.3 (a)**, the small-scale residual errors and errors due to incorrect mapping are illustrated in Figure **5.3 (b)**, **(c)**, and **(d)**. The measurement noise in the smoothed map is still 2.0 nm rms. This is because the noise is dominated by air turbulence due to the long radius of the mirror. Measurement errors caused by air turbulence are usually of low

spatial frequency, while the residual after smoothing is small-scale error. In this case, smoothing does not reduce the measurement noise in the smoothed map. However, a larger smoothing filter will generally reduce the measurement noise. It also reduces the effect of mapping error, $\vec{\varepsilon} \cdot \vec{\nabla} W_{\text{smooth}}$. The total RSS error is $\sqrt{2.0^2 + 0.8^2 + 1.5^2 + 1.4^2} = 3.0\,\text{nm}$, which is slightly smaller than that without applying the smoothing filter. If the mapping error and/or the surface slope error are large, the use of a smoothing filter method could decrease the wavefront error more dramatically.

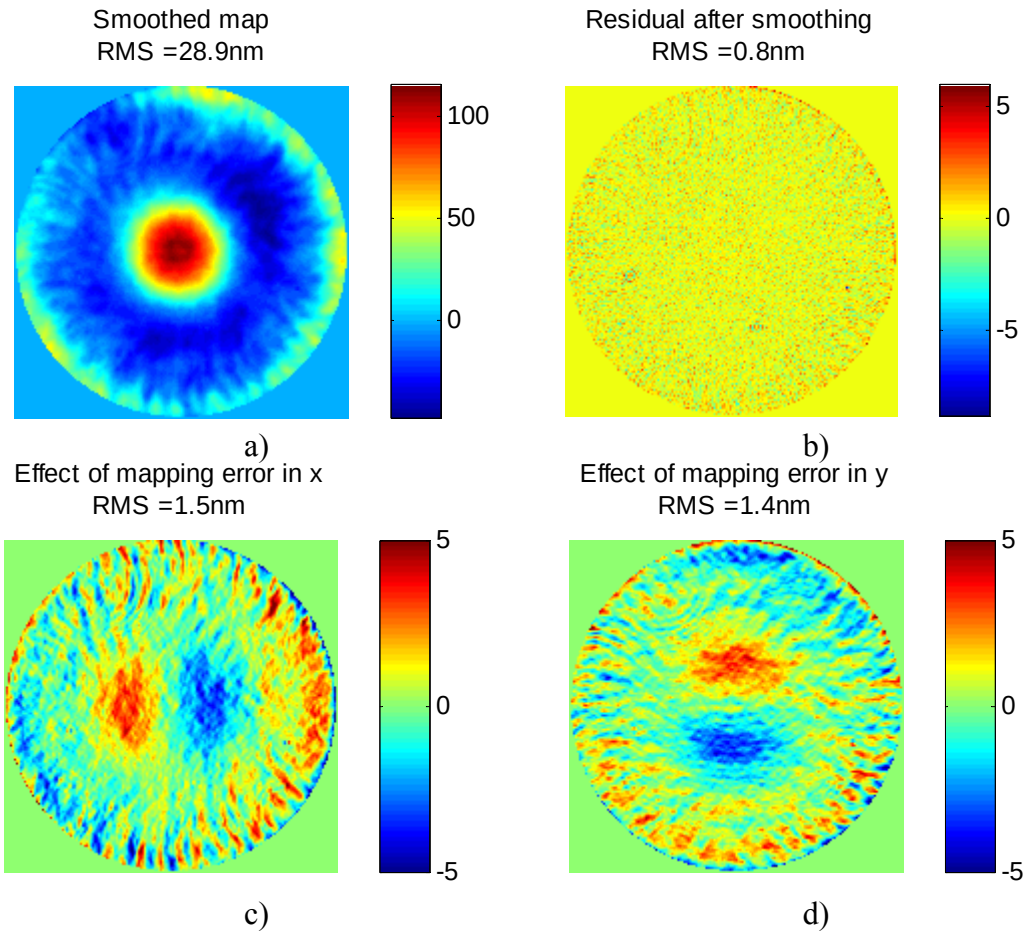


Figure **5.3**: a) Smoothed reference map with a 7 × 7 filter; b) Small-scale residual error after smoothing, showing the difference between the full resolution map and the smoothed map; c) Wavefront error due to average 2-pixel mapping error in *x* direction, d) Wavefront error due to average 2-pixel mapping error in *y* direction.

Figure **5.4** shows the RSS errors versus the filter sizes with different mapping errors for the reference sphere mentioned above. When there is no mapping error, the RSS error is only from the small-scale error due to smoothing and the measurement noise in the smoothed map. It only contains the measurement noise when no filter is applied. The RSS errors depend on the filter size. When the filter size is small, the errors due to

mapping are dominant. When the filter is large, the residual small-scale errors increase. Therefore, there is an optimal filter size that can minimize the RSS errors.

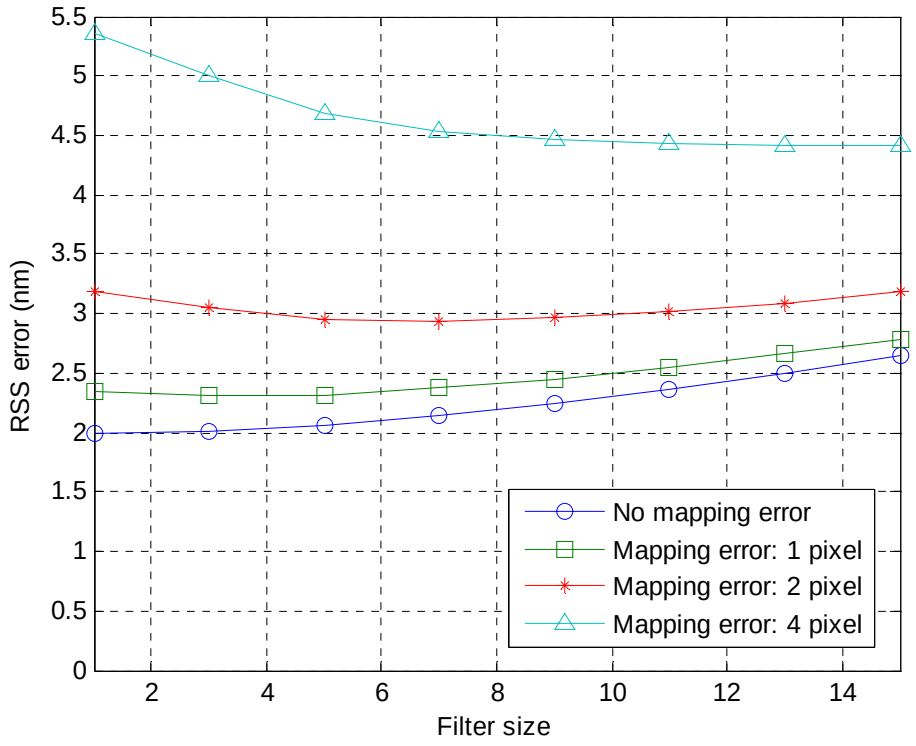


Figure **5.4**: RSS errors as a function of filter size.

To conclude this chapter, an example of estimating the measurement uncertainty of a surface is given. If the surface being tested requires subtraction of the reference surface discussed above, the uncertainty should include the RSS error from the reference map (the net error shown in Figure **5.4**), and the measurement noise for the test surface. Other sources of errors, like retrace errors and errors from the interferometer, should also be included. The total RSS error from the smoothed reference map is 3.0 nm, after applying a 7 × 7 smoothing filter. The measurement noise is 2.1 nm for the test surface

due to the measurement environment. This is calculated using the method discussed in Chapter 2. If the interferometer is calibrated to an accuracy of 1 nm, the total uncertainty of the test surface is 3.8 nm as shown in Table **5.3** after all sources of errors are summed by a RSS.

Table **5.3**: The total uncertainty in the test surface, when the test surface requires subtraction of the reference surface.

RSS error from the smoothed reference map	3.0 nm
Noise in the test map	2.1 nm
Error in the interferometer	1.0 nm
Total uncertainty of the test surface (RSS)	**3.8 nm**

CHAPTER 6

OPTIMAL DESIGN FOR COMPUTER-GENERATED HOLOGRAMS

Aspheric optical surfaces are often tested using interferometers with the help of computer-generated holograms (CGHs), because CGHs can generate reference wavefronts of virtually any desired shape. The accurately drawn pattern on a CGH provides exact wavefront control. However, uncertainties in the CGH manufacturing processes introduce errors in holograms and hence in the generated wavefront. This chapter consists of four sections, which address fabrication errors in the CGH writing process and methods to optimize the CGH design.

Section **6.1** provides an overview of computer-generated holograms. The sensitivity functions to CGH fabrication errors, such as duty-cycle, phase depth are reviewed in Section **6.2**. Section **6.3** presents methods for measuring CGH fabrication errors and an example of estimating the wavefront errors from fabrication non-uniformities. Section **6.4** describes how to design CGHs to give good diffraction efficiency and limited sensitivity to manufacturing errors for both chrome-on-glass and phase etched CGHs.

6.1 CGHs for optical testing

CGHs are diffractive optical elements synthesized with the aid of computers. They are extremely useful with a high degree of flexibility in generating complex

wavefronts with desired amplitude and phase through the use of diffraction. CGHs play an important role in a variety of applications in modern optics, include optical data storage, laser scanning, image processing and optical testing [**Gaylord 1985**].

The application of CGHs in optical interferometry allows complex non-spherical surfaces to be measured easily without using expensive reference surfaces or null lenses. Figure **6.1** shows an example of using a CGH to test an aspheric surface. The CGH is designed to create a wavefront that matches the aspheric surface under test, then the light retro-reflects from the test surface and back to the interferometer to create a null test [**Loomis 1980 and Creath 1992**]. CGHs for optical testing are usually binary, and they can be either amplitude or phase etched. Amplitude CGHs usually have chrome patterns on the glass, while phase etched CGHs have patterns etched into glass. The binary CGH has wavy line pattern, and diffraction at order m adds $m\lambda$ optical path per line.

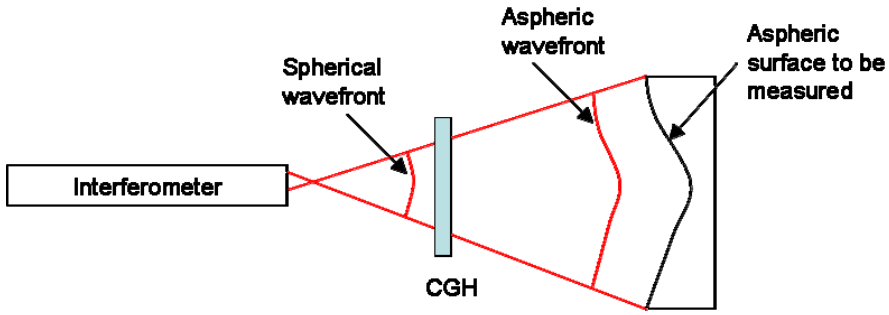


Figure **6.1**: An example of using a CGH to test an aspheric surface.

A CGH can store both the amplitude and phase information of a complex wavefront by controlling positions, widths, and groove depths of the recorded pattern. The design of a CGH is separated into two parts that allow for independent determination

of amplitude and phase of the diffracted light. The structure of the lines, such as the groove depth, or duty cycle, is set to give the appropriate diffraction efficiency. The overall pattern of lines is chosen to give the correct shape of the wavefront. For precise measurement, the influence of error must be understood and characterized. Figure **6.2** shows a drawing of a binary CGH with 1 wave Zernike spherical aberration and 36 waves of tilt.

Figure **6.2**: A binary CGH with 1 wave Zernike spherical and 36 waves of tilt.

Much research has been conducted to improve design schemes and fabrication techniques for CGHs [**Arnold 1989; Arnold 1992; Wyant 1974b; Ono 1984**]. The majority of the work has been directed toward optimizing diffraction efficiencies of the holograms. However, the wavefront phase is of most concern in optical testing. Chang's dissertation on diffraction wavefront analysis of computer-generated holograms used a binary diffraction model to derive the wavefront phase sensitivity to duty-cycle and etching depth [**Chang 1999a**]. Her parametric CGH model has been used throughout this chapter to analyze the CGH fabrication uncertainties, and optimize the design of CGHs to minimize the sensitivity to those manufacturing errors.

6.2 Sensitivity functions for binary CGHs

This section reviews Chang's binary diffraction model [**Chang 2006**]. The diffraction efficiency and wavefront phase in the far field is derived from Fraunhofer diffraction theory. Sensitivity functions for both diffraction efficiency and wavefront phase are summarized, which will be used later to estimate the wavefront errors and optimize a CGH design.

6.2.1 Parametric diffraction model

The simplest form of a hologram is a linear diffraction grating, where the spatial frequency of the grating pattern is constant over the entire hologram. A general CGH may be viewed as a collection of linear gratings with variable spatial frequencies. By controlling the spatial frequencies of these linear gratings across the CGH, the incident light can be deflected into any desired form. The performance of a CGH may, therefore, be directly related to the diffraction characteristics of a linear grating [**Swanson 1991**]. Linear gratings are often used to study CGH properties in order to avoid mathematical difficulties in modeling complicated holograms. The binary, linear grating model Chang used is illustrated in Figure **6.3**.

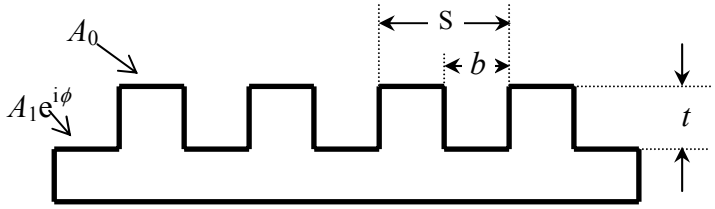


Figure **6.3**: A binary, linear grating profile.

The grating is defined by the period S and the etching depth t. Duty-cycle is defined as $D = b / S$, where b is the width of the etched area. A_0 and A_1 are the amplitudes of the output wavefronts from the unetched area and etched area of the grating, respectively. The phase function ϕ represents the phase difference between rays from the peaks and valleys of the grating structure. For a chrome-on-glass CGH used in transmission, A_0 is zero, and A_1 is unity. For a phase etched CGH used in transmission, A_0 and A_1 are both unity, and $\phi = 2\pi t(n\text{-}1)/\lambda$, where n is the refractive index of the substrate.

The wavelength of the incident light is assumed to be much smaller than the grating period S, so scalar diffraction approximations can be applied [**Gaskill 1978**]. For a planar wavefront at normal incidence, the output wavefront immediately past the grating, either reflected or transmitted, can be expressed as a simple product of the incident wavefront function and the grating surface profile function. The complex amplitude of the field, immediately after the grating, can be written as:

$$u(x) = A_o + (A_1 e^{i\phi} - A_0)\text{rect}\left(\frac{x}{b}\right) * \frac{1}{S}\text{comb}\left(\frac{x}{S}\right). \qquad \textbf{(6.1)}$$

Based on Fraunhofer diffraction theory, the far-field diffraction wavefront may be related to the original wavefront by a simple Fourier transform. Hence, the far-field wavefront function of a normally incident planar wave upon the grating described in Equation **6.1** is:

$$\begin{aligned} U(\xi) &= \Im\{u(x)\} = A_0\delta(\xi) + \left(A_1 e^{i\phi} - A_0\right)\cdot b\cdot \mathrm{sinc}(b\xi)\cdot \mathrm{comb}(S\xi) \\ &= \left\{ A_0\delta(\xi) + \left[A_1\cos(\phi) - A_0\right]\cdot D\cdot \mathrm{sinc}(DS\xi)\cdot \sum_{m=-\infty}^{\infty} \delta\left(\xi - \frac{m}{S}\right)\right\} \\ &\quad + i\left\{ A_1\sin(\phi)\cdot D\cdot \mathrm{sinc}(DS\xi)\cdot \sum_{m=-\infty}^{\infty} \delta\left(\xi - \frac{m}{S}\right)\right\}. \end{aligned} \qquad \textbf{(6.2)}$$

Looking at one order at a time gives

$$U(\xi) = \begin{cases} \left\{ A_0 + \left[A_1\cos(\phi) - A_0\right]\cdot D \right\} + i\left\{ A_1\sin(\phi)\cdot D \right\}; & m = 0 \\ \left\{ \left[A_1\cos(\phi) - A_0\right]\cdot D\cdot \mathrm{sinc}(mD) \right\} + i\left\{ A_1\sin(\phi)\cdot D\cdot \mathrm{sinc}(mD)\right\} & m \neq 0 \end{cases}. \qquad \textbf{(6.3)}$$

The spatial frequency ξ corresponds to an angle $\theta = \lambda\xi$. The diffraction orders propagate in the directions $\theta = \frac{m\cdot\lambda}{S}$ where m is the diffraction order. Equation **6.3** shows that the diffraction wavefront function $U(\xi)$ has non-zero values only when ξ is an integer multiple of $1/S$. This behavior describes the existence of multiple diffractive orders.

6.2.2 Diffraction efficiency

The diffraction efficiency of a specific diffraction order describes the efficiency of the hologram in deflecting light into a particular direction. Diffraction efficiency η is defined as the ratio of the intensity of the diffracted wavefront to that of the incident wavefront:

$$\eta = \frac{\left|U\left(\xi\right)\right|^2}{\left|U_0\left(\xi\right)\right|^2}. \qquad (\mathbf{6.4})$$

Assume the intensity of the incident wavefront $\left|U_0(\xi)\right|^2$ is unity, then η for the zero diffraction order ($m = 0$) is:

$$\eta\big|_{m=0} = A_0^2(1-D)^2 + A_1^2D^2 + 2A_0A_1D(1-D)\cos(\phi). \qquad (\mathbf{6.5})$$

For non-zero diffraction orders ($m \neq 0$):

$$\eta\big|_{m\neq 0} = \left[A_0^2 + A_1^2 - 2A_0A_1\cos(\phi)\right]D^2\mathrm{sinc}^2(mD). \qquad (\mathbf{6.6})$$

6.2.3 Wavefront phase

The CGH is designed with variable fringe spacings to produce a desired wavefront phase distribution. Additional phase variations result from unwanted variations in duty-cycle, phase depth, and amplitude. These can be determined from

Equation **6.3**. The phase function Ψ is defined as the ratio of the imaginary part to the real part of the complex wavefront $U(\xi)$:

$$\tan(\Psi) = \frac{\mathrm{Im}\{U(\xi)\}}{\mathrm{Re}\{U(\xi)\}}. \qquad (\mathbf{6.7})$$

Substituting the Equation **6.3** into Equation **6.7** leads to:

$m = 0$

$$\tan(\Psi)_{m=0} = \frac{DA_1 \sin(\phi)}{A_0(1-D) + A_1 D\cos(\phi)}, \qquad (\mathbf{6.8})$$

$m = \pm 1, \pm 2, \ldots$

$$\tan(\Psi)_{m\neq 0} = \frac{A_1 \sin(\phi)\cdot \mathrm{sinc}(mD)}{[-A_0 + A_1\cos(\phi)]\cdot \mathrm{sinc}(mD)}. \qquad (\mathbf{6.9})$$

The phase Ψ in radians can be obtained by taking the arctangent of these equations. The wavefront phase W in waves can be obtained by

$$W = \frac{\Psi}{2\pi}. \qquad (\mathbf{6.10})$$

Notice that the sinc(mD) functions are left in both the numerator and denominator of Equation **6.9** . This shorthand notation is used to preserve the sign information for the phase unwrapping process (equivalent to the atan2 function). The phase unwrapping process uses the sign information of the real and imaginary parts of the complex wavefront $U(\xi)$, which allows calculation of phase from 0 to 2π. The new phase values

are then evaluated point by point in the map. Any integer multiple of 2π can be added or subtracted from the phase value to make the phase function continuous.

6.2.4 Sensitivity functions

We have shown that the diffraction efficiency and the wavefront phase can be expressed as a function of the duty-cycle, etching depth and amplitude of the output wavefront. By taking the first derivative of diffraction efficiency or wavefront phase with respect to those parameters, we can get the sensitivity functions. The wavefront phase sensitivity functions $\partial\Psi/\partial D$, $\partial\Psi/\partial\phi$ and $\partial\Psi/\partial A_1$ are introduced to specify the wavefront error caused by small deviations in duty-cycle ΔD, phase depth $\Delta\phi$, and amplitude ΔA_1. The variation in the amplitude is due to scattering loss from the surface roughness, which varies across the substrate due to the limitations in the etching process [**Zhou 2007a**]. Table **6.1** summarizes the diffraction efficiencies, wavefront phases, and sensitivity functions for the zero and the non-zero diffraction orders. The table lists two different amplitude sensitivity functions. $\partial\Psi/\partial A_1$ can be used when only the scattered light loss in the etched area is considered, while $\dfrac{\partial\Psi}{\partial(A_0/A_1)}$ considers the scattered light loss in both etched and unetched areas.

Table **6.1**: Summary of equations for the parametric model analysis.

	Zero order ($m = 0$)	Non-zero order ($m \pm 1, \pm 2, \ldots$)
	Diffracted wavefront	
η (Diffraction efficiency)	$A_0^2(1-D)^2 + A_1^2 D^2 + 2A_0A_1D(1-D)\cos\phi$	$\left(A_0^2 + A_1^2 - 2A_0A_1\cos\phi\right)D^2\,\mathrm{sinc}^2\left(mD\right)$
$\tan\Psi$ (ψ wavefront phase)	$\dfrac{A_1 D\sin\phi}{A_0\left(1-D\right)+A_1 D\cos\phi}$	$\dfrac{A_1\sin\phi\cdot\mathrm{sinc}\left(mD\right)}{\left(-A_0+A_1\cos\phi\right)\cdot\mathrm{sinc}\left(mD\right)}$
	Sensitivity functions	
$\partial\eta/\partial D$	$-2A_0^2\left(1-D\right)+2A_1^2 D+2A_0A_1\left(1-2D\right)\cos\phi$	$2\left(A_0^2 + A_1^2 - 2A_0A_1\cos\phi\right)D\mathrm{sinc}\left(2mD\right)$
$\partial\eta/\partial\phi$	$-2A_0A_1D\left(1-D\right)\sin\phi$	$2A_0A_1\sin\phi\,D^2\,\mathrm{sinc}^2\left(mD\right)$
$\partial\eta/\partial A_1$	$2A_1D^2+2A_0D\left(1-D\right)\cos\phi$	$\left(2A_1-2A_0\cos\phi\right)D^2\,\mathrm{sinc}^2\left(mD\right)$
$\partial\Psi/\partial D$	$\dfrac{A_0A_1\sin\phi}{A_1^2D^2+A_0^2\left(1-D\right)^2+2A_0A_1D\left(1-D\right)\cos\phi}$	$\begin{cases}\infty, & \text{for}\,\mathrm{sinc}\left(mD\right)=0\\ 0, & \text{otherwise}\end{cases}$
$\partial\Psi/\partial\phi$	$\dfrac{A_1^2D^2+A_0A_1D\left(1-D\right)\cos\phi}{A_1^2D^2+A_0^2\left(1-D\right)^2+2A_0A_1D\left(1-D\right)\cos\phi}$	$\dfrac{A_1^2-A_0A_1\cos\phi}{A_1^2+A_0^2-2A_0A_1\cos\phi}$
$\partial\Psi/\partial A_1$	$\dfrac{A_0D\left(1-D\right)\sin\phi}{A_0^2\left(1-D\right)^2+A_1^2D^2+2A_0A_1D\left(1-D\right)\cos\phi}$	$\dfrac{-A_0\sin\phi}{A_0^2+A_1^2-2A_0A_1\cos\phi}$
$\dfrac{\partial\Psi}{\partial(A_0/A_1)}$	$\dfrac{-D(1-D)\cdot\sin\phi}{(A_0/A_1)^2(1-D)^2+D^2+2D(1-D)(A_0/A_1)\cos\phi}$	$\dfrac{\sin\phi}{(A_0/A_1)^2+1-2(A_0/A_1)\cos\phi}$

The sensitivity functions can be evaluated directly to give the wavefront error due to variations in duty-cycle D, phase depth ϕ, or amplitude A_1. These functions are shown in Equation **6.11**, **6.12** and **6.13**, respectively,

$$\Delta W_D = \frac{1}{2\pi}\frac{\partial\Psi}{\partial D}\cdot\Delta D = \frac{1}{2\pi}\frac{\partial\Psi}{\partial D}\cdot\left(\frac{\Delta D}{D}\right)\cdot D\,, \qquad \textbf{(6.11)}$$

$$\Delta W_{\phi} = \frac{\partial \Psi}{\partial \phi} \cdot \Delta \phi = \frac{\partial \Psi}{\partial \phi} \cdot \left(\frac{\Delta \phi}{\phi} \right) \cdot \phi \,, \qquad \textbf{(6.12)}$$

$$\Delta W_{A_1} = \frac{1}{2\pi} \frac{\partial \Psi}{\partial A_1} \cdot \Delta A_1 = \frac{1}{2\pi} \frac{\partial \Psi}{\partial A_1} \cdot \left(\frac{\Delta A_1}{A_1} \right) \cdot A_1 \,, \qquad \textbf{(6.13)}$$

where

ΔD = duty-cycle variation across the grating,

ΔW_D = wavefront variation in waves due to duty-cycle variation,

$\Delta \phi$ = phase depth variation in radians across the grating,

ΔW_{ϕ} = wavefront variation in waves due to etch depth variation,

ΔA_1 = variation in amplitude A_1 cross the grating,

ΔW_{A_1} = wavefront variation in waves due to amplitude variation.

As long as the duty-cycle, etching depth and amplitude vary over spatial scales that are large compared to the grating spacing, Equations **6.11-6.13** can be used to determine the coupling between fabrication errors and system performance. The wavefront sensitivity functions provide a way of calculating the wavefront phase changes resulting from fabrication non-uniformities. They can be used to identify which hologram structures are the most or the least sensitive to those fabrication uncertainties.

6.3 CGH fabrication error analysis

The previous section provides the sensitivity functions to three types of CGH fabrication errors. To quantify the wavefront errors caused by fabrication errors, the fabrication non-uniformities for each error source must be determined. CGH fabrication errors can be classified into five types:

- Substrate figure errors: substrate surface variation from its ideal shape,
- Duty-cycle errors: variation in duty-cycle,
- Phase depth errors: variation in etching depth,
- Surface roughness errors: variation in surface roughness for etched regions,
- Pattern distortion errors: displacement of recorded pattern from its ideal position.

Beside these fabrication errors, another type of error associated with the CGH is the encoding induced error. Encoding is the process to approximate the smooth fringes of a CGH as polygon segments in order that it can be fabricated with e-beam machine or laser writers in a photomask house [**Kallioniemi 1997**]. This encoding error is a design error and can be controlled to be less than the interferometer noise level [**Zhao 2009b**]. This dissertation will not discuss the encoding errors and only focus on the CGH fabrication errors.

The effects of all fabrication errors are coupled together, making it difficult to decouple the individual effects of each error source on the wavefront. The parametric model discussed in Section **6.2** is used to estimate the RSS wavefront errors of the CGH due to fabrication uncertainties.

Some methods to determine the fabrication non-uniformities in duty-cycle, etching depth and surface roughness are provided in Section **6.3.1**. A wavefront error analysis for a phase CGH is given in Section **6.3.2**. The calibration of CGH substrates, which provides measurement accuracy of about 1 nm rms, is demonstrated in Section **6.3.3**.

6.3.1 Measurement of fabrication errors

To quantify the wavefront error for a CGH used in optical testing, its fabrication non-uniformities must be determined for each error source. These fabrication non-uniformities can be obtained by measuring a set of sampled points over the CGH.

Substrate measurement:

Substrate errors are typically of low spatial frequencies. Their effects on the wavefront depend on the application of the CGH. If the CGH is used in reflection, a surface defect on a CGH substrate with a peak-to-valley deviation of δs will produce a wavefront error of $2\delta s$. A transmission CGH that has the same peak-to-valley surface defect, on the other hand, will produce a wavefront phase error of $(n\text{-}1)\delta s$, where n is the index of refraction of the substrate. The substrate errors influence all diffraction orders equally [**Chang 1999a**].

One method of eliminating the effect of substrate errors in a CGH is to measure the flatness of the substrate before the grating patterns are written using an interferometer. The difficulty of this method is to correctly register the two maps (before

and after writing the patterns). Moreover, the substrate may change shape with a different mounting situation. Due to these issues, this method is not commonly used. The other method is to measure the effect of surface irregularities using the zero order diffraction from the CGH, and subtract it from the non-zero order surface measurement. A test setup is illustrated in Figure **6.4**, where the CGH is placed in a diverging beam, which is the actual working configuration for the CGH (also known as in situ measurement). The return sphere is first measured, then the wavefront with the CGH is tested. The difference of these two measurements gives the transmitted wavefront error of the CGH in zero order diffraction, which is mainly the CGH substrate error. The advantage of this in situ measurement is that there is no registration error when subtracting the zero order substrate measurement from the non-zero surface measurement, so the CGH substrate error can be correctly backed out.

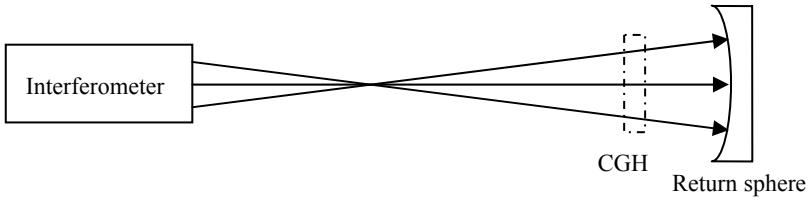


Figure **6.4**: Setup of in situ CGH substrate measurement.

Note that besides the CGH substrate error, wavefront errors introduced by non-uniformities in duty-cycle, etching depth, and amplitude also affect the zero order measurement. These errors affect the zero and non-zero orders differently. When we subtract the zero order measurement from the non-zero order surface measurement, the

CGH substrate error can be totally removed, leaving residual wavefront errors from the fabrication non-uniformities, which will be discussed later.

The setup in Figure **6.4** also introduces aberrations from CGH plate thickness, since it is illuminated in a diverging beam. The main aberration induced by the plate thickness is spherical aberration. The spherical aberration in RMS can be calculated using the following equation

$$W_{SA} = \frac{t(n^2-1)}{(F_n)^4 \cdot 128 n^3} \cdot \frac{1}{\sqrt{180}}, \qquad (\mathbf{6.14})$$

where t is the thickness of the CGH substrate, n is the index of refraction, and F_n is the F-number of the illuminated beam. A slower system has less spherical aberration, and the spherical aberration increases quickly for the faster system. For example, if a CGH with a thickness of 6.35 mm and refractive index of 1.5 is illuminated by an F/7 beam, the spherical aberration is about 0.6 nm rms. The RMS spherical aberration becomes 16.9 nm for an F/3 beam.

Usually, CGHs are designed to correct the spherical aberration in the first diffraction order or the order used for testing a surface, but there is the spherical aberration remaining in the zero order measurement for the in situ substrate measurement. To remove the spherical aberration, a synthetic map with the right amount of spherical aberration can be generated and backed out. However, there will be error due to incorrect mapping between the synthetic spherical wavefront and the zero order measurement. Figure **6.5** provides the RMS spherical aberration as a function of F-number and the RSS wavefront error caused by 0.5% mapping distortion in both x and y

directions when subtracting the spherical aberration from the measurement. It shows that the RSS wavefront error is always smaller than the RMS spherical aberration so it is worth correcting the spherical aberration.

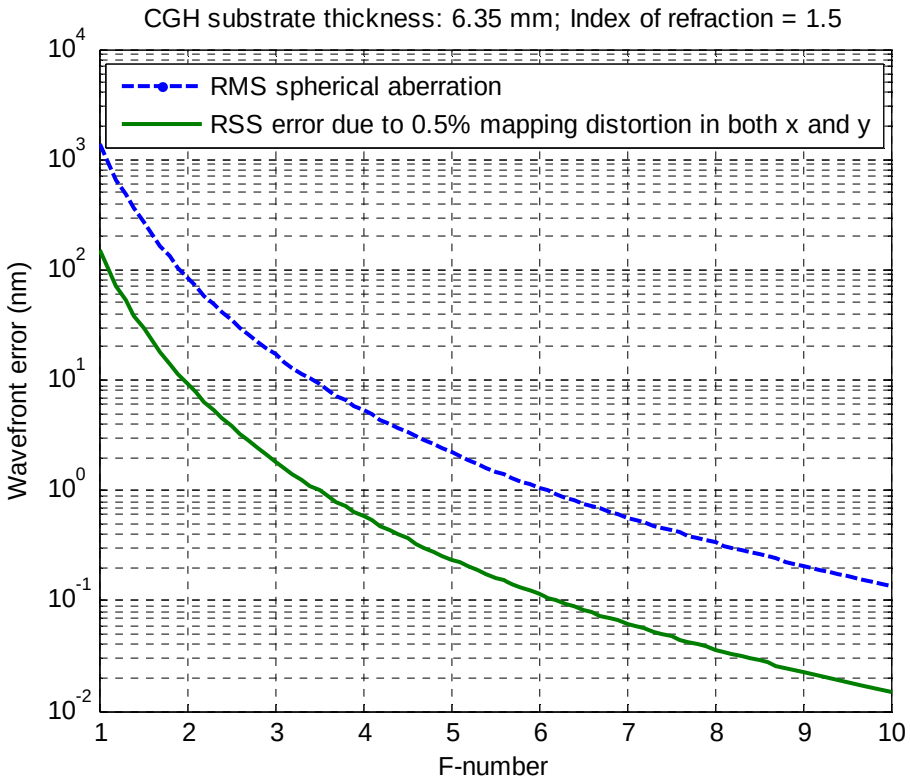


Figure **6.5**: The RMS spherical aberration as a function of F-number and the RSS wavefront error caused by 0.5% mapping distortion in both x and y directions when subtracting the spherical aberration from the measurement. Assume the CGH substrate thickness is 6.35 mm and the index of refraction of the substrate is 1.5.

For a slow system (i.e., $F_n > 6$), the RMS spherical aberration is less than 1 nm, which is very small and can often be ignored. For a fast system (i.e., $F_n < 2$), the spherical aberration is large and has to be backed out from the measurement. The RSS wavefront error due to 0.5% mapping error is larger than 10 nm even after removing the

spherical aberration in the measurement. For this case, it is better to measure the CGH substrate in a collimated beam instead of in situ, because there is no spherical aberration induced in a collimated beam.

Measuring the CGH substrate in a collimated beam has its limitation, because there will be mapping error between the zero order and non-zero order measurements when subtracting the CGH substrate error even with the use of fiducials. The induced wavefront error from this mapping distortion depends on the transmitted wavefront slope of CGH substrates. We measured many CGH substrates, and a typical one in transmission is shown in Figure **6.6**.

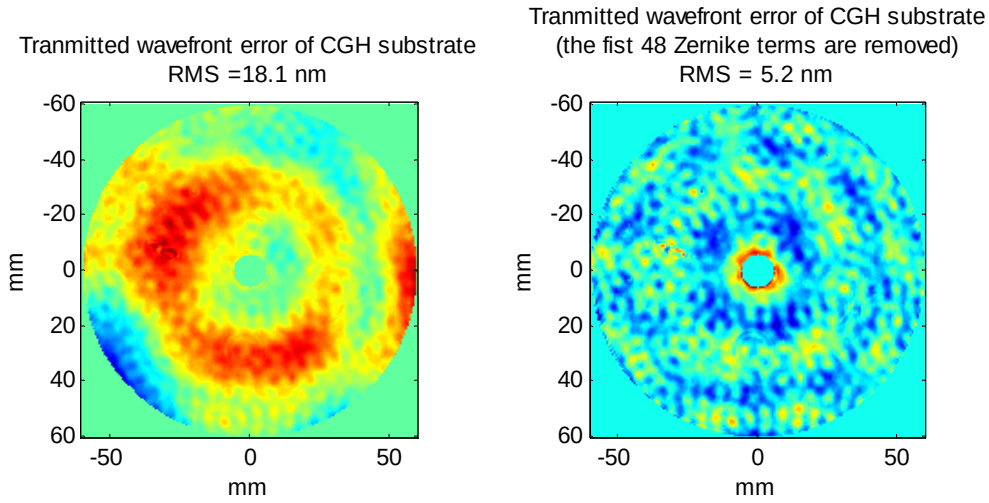


Figure **6.6**: a) a typical CGH substrate measured in transmission; b) the transmitted wavefront error of the CGH substrate with the first 48 Zernike terms removed. It shows the high-frequency turtle shell pattern on the substrate.

Assuming the same 0.5% mapping error in both x and y directions when subtracting the above CGH substrate error, the induced error are 2.8 nm and 2.2 nm in the

x and y directions, as shown in Figure **6.7**. In this case, the RSS error due to mapping is $\sqrt{2.8^2 + 2.2^2} = 3.6$ nm.

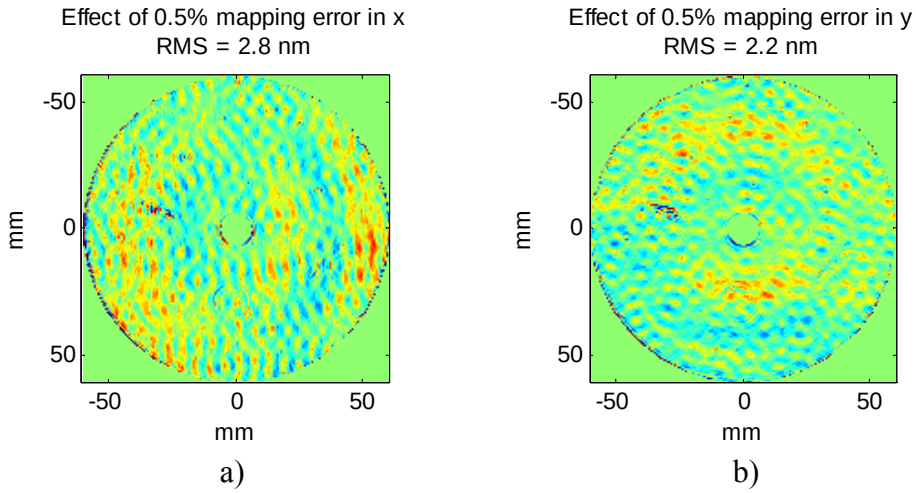


Figure **6.7**: a) Wavefront error due to 0.5% mapping error in x direction; b) Wavefront error due to 0.5% mapping error in y direction, when subtracting the CGH substrate error shown in Figure **6.6**.

Comparing the RSS error caused by the CGH substrate with that caused by the spherical aberration (Figure **6.5**), the conclusion is when the CGH is used in a system faster than F/2.5, it is better to test its substrate in a collimated beam and then back out the substrate error. When the CGH is used in a system slower than F/2.5, it can be tested in situ and then remove the spherical aberration. Note that this conclusion is based on the figure error of the CGH substrate we measured. Whether to measure the CGH substrate in situ or in a collimated beam depends on the substrate wavefront slope and the F-number of the test system. It also depends on the application or the specification of the test optic. Errors caused by removing the spherical aberration are of low spatial frequency and it may not be significant for some applications. Errors caused by

removing the CGH substrate error have some high spatial frequency content as shown in Figure **6.7**. This may be an issue when a system has a tight specification in the mid/high spatial frequency region.

If the CGH shown in Figure **6.6** is used in an F/3 system and the CGH substrate is tested in situ, then the error caused by 0.5% mapping error (when subtracting the spherical aberration from the zero order measurement) is 1.1 nm rms. If the same CGH is tested in a collimated beam, then the error caused by 0.5% map distortion (when subtracting the 0^{th} order measurement from the 1^{st} order one) is 3.6 nm rms. The total error from the CGH substrate should also include the measurement uncertainty of the substrate, the measurement noise of the return sphere, and error due to mis-registration between the return sphere measurement and the zero order CGH substrate measurement. Table **6.2** gives an example of the total error from the CGH substrate for both the in situ and the collimated tests. The total RSS errors are 3.8 nm and 1.6 nm, respectively.

Table **6.2**: The total errors from the CGH substrate when the CGH is used in an F/3 system for the in situ test and the collimated test.

	Collimated test	In situ test
Error due to 0.5% map distortion when correcting spherical aberration	N/A	1.1 nm rms
Error due to 0.5% map distortion when subtracting the 0^{th} order measurement from the 1^{st} order one	3.6 nm rms	N/A
Measurement noise in the 0^{th} order	0.5 nm rms	
Measurement noise in the return optic	0.5 nm rms	
Error due to mis-registration between return sphere measurement and the 0^{th} order measurement	1.0 nm rms	
RSS	**3.8 nm rms**	**1.6 nm rms**

Duty-cycle and etching depth measurement:

Variations in duty-cycle and etching depth can be measured with either an atomic force microscope (AFM) or an interference microscope, which provide the surface relief of the CGH. Another method is to measure the diffraction efficiencies of several diffraction orders, and then fit the model listed in Table **6.1** to determine the duty-cycle and etching depth. A set of the sampled points should be measured, and the duty-cycle and etching depth at each sampled point are determined by using a nonlinear least-squares fit. The fitted results are the average duty-cycle and etching depth over the CGH area illuminated by the small laser spot. This method has to assume amplitude variations are small enough to ignore. Otherwise the fit cannot distinguish the effects of phase depth and amplitude on the diffraction efficiency. The example in Section **6.3.2** shows this assumption is valid for a particular hologram.

To verify the feasibility of the second method, a Monte Carlo analysis was used to find the relationship between the number of diffraction orders measured and the accuracy obtainable for the duty-cycle and etching depth fit. The results shown in Figure **6.8** are for a 1000-trial Monte Carlo simulation. The nominal duty-cycle and phase depth are 0.49 and 0.33λ respectively in this simulation. These are the measured parameters of a CGH demonstrated in the next section. A_0 and A_1 are assumed to be unity, since it is a phase CGH. A uniformly distributed random error is added to each diffraction efficiency measurement. Measurement errors of up to ±1%, ±3% and ±5% for each order are assumed in the analysis. δt and δD are the standard deviations of the etching depth and

duty-cycle of the 1000 trials, which represent the measurement confidence levels of the etching depth and duty-cycle, respectively.

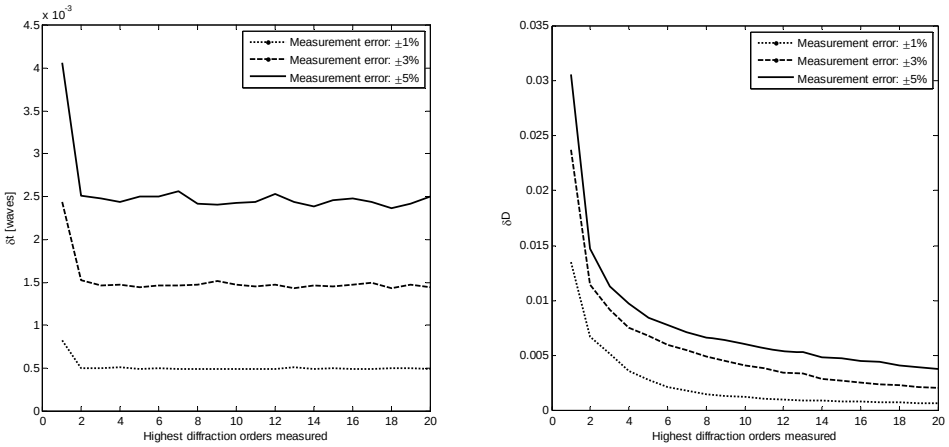


Figure **6.8**: Monte Carlo simulation result: error in estimating the etching depth vs. the number of diffraction orders measured (left) and error in estimating the duty-cycle vs. the number of diffraction orders measured (right). δt and δD are the standard deviations of the etching depth and duty-cycle.

The simulation demonstrates that δt does not change significantly by increasing the number of measured diffraction orders; it is proportional to the measurement error in diffraction efficiencies. δD decreases with an increase in the number of measured diffraction orders, and finally converges when the number of diffraction orders is sufficiently large. This can be explained by sensitivity functions of diffraction efficiency to phase depth $\partial\eta/\partial\phi$ and to duty-cycle $\partial\eta/\partial D$. Figure **6.9** shows the sensitivity functions for each diffraction order when the duty-cycle is 0.49, and the phase depth is 0.33λ. The sensitivity of diffraction efficiency to phase depth drops dramatically at higher diffraction orders, while it oscillates around zero for duty-cycle. This means the higher diffraction orders do not contribute considerably to the knowledge of the etching

depth, so the Monte Carlo simulation gives us a relatively flat curve for δt. However, higher diffraction orders can increase the measurement accuracy of the duty-cycle.

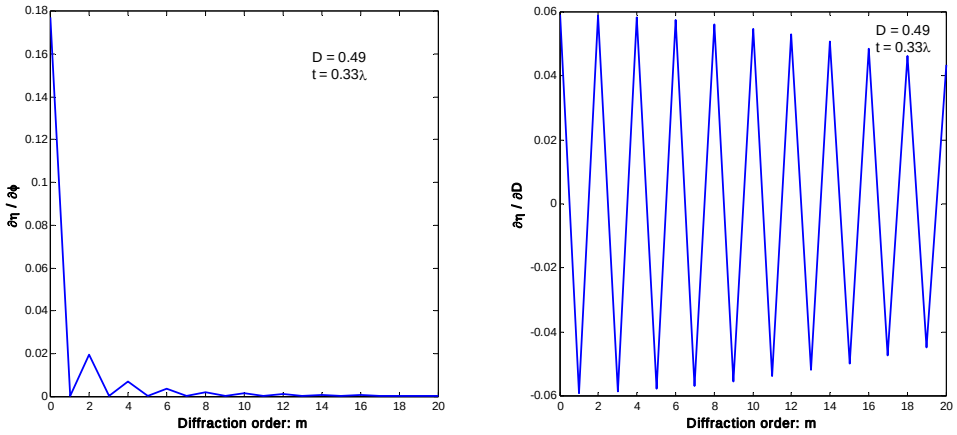


Figure **6.9**: Sensitivity of diffraction efficiencies to the phase depth (left) and the duty-cycle (right).

By careful measurement, we can control the measurement of diffraction efficiencies within ±1%. In this case, δt and δD are about 0.0005λ and 0.0015 respectively, if we measure up to the $\pm 10^{th}$ diffraction order. These accuracies are sufficient to evaluate the non-uniformities in the duty-cycle and etching depth as will be shown in Section **6.3.2**.

Amplitude variation measurement:

The amplitude of the light from the etched regions can vary due to limitations in the reactive ion etching (RIE) process. The surface roughness of the etched area may vary due to etching [**Ricks 1990**], so the light incident on the CGH scatters differently. By assuming the surface roughness is much smaller than the wavelength of incident light,

the coupling of the scattered light to the diffracted light can be ignored. However, the scattering will decrease the amplitude of the diffracted light. This leads to an amplitude ratio variation over the CGH. It is difficult to measure the intensities (or amplitudes) from the etched area and unetched area of the CGH directly. One useful method is to measure the surface roughness of both areas, and then calculate the scattered loss. The scattered loss is zero when the surface is perfectly flat, and it increases with the surface roughness. A simple formula to approximate the total integrated scatter (TIS) for the transmitted light is

$$I_{\text{scat}} \approx (2\pi\sigma)^2 = [2\pi(n-1)R_q/\lambda]^2, \qquad (\mathbf{6.15})$$

where σ^2 is wavefront variance in units of waves, n is the index of refraction of the CGH substrate, and R_q is the RMS surface roughness. The TIS for the reflected light is $I_{\text{scat}} = (4\pi R_q/\lambda)^2$. The surface roughness can be measured by an interference microscope [**Saxer 2003**]. The intensity of the light from the etched and unetched regions can be determined by

$$I = 1 - I_{\text{scat}}. \qquad (\mathbf{6.16})$$

The variation of the surface roughness causes different amount of the scattering loss, which leads to the wavefront errors.

Figure **6.10** shows how surface roughness couples into the performance for a binary, linear, phase grating. A random surface height error with a Gaussian distribution is added to the bottom surface in each period, assuming the top surface is perfectly

smooth. Since we are interested in the coupling of roughness *variation* across the substrate, we model a grating that has a sinusoid roughness distribution across the grating, so the bottom surface is rougher at the center and smoother at the edges. Assume the amplitudes A_0 and A_1 are unity if there is no scattered loss. Therefore, this rough grating is equivalent to a grating with a constant phase ϕ, duty-cycle D, amplitude A_0, and varying amplitude A_1.

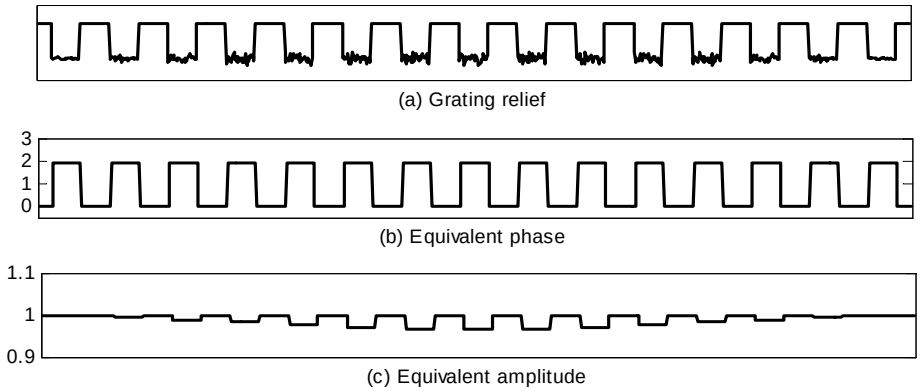


Figure **6.10**: A rough grating and its equivalent model.

The effect of the surface roughness comes in very slowly. Assume the refractive index of the substrate is 1.5, and the wavelength is 632.8 nm. The transmittance of the light at the grating interface is 96% based on the Fresnel formula. If the surface roughness of the etched area is 2 nm rms with a variation of 0.5 nm rms, the scattered loss would be 0.01% with a variation of 0.005% rms. Therefore the amplitude variation would be 0.003% rms. For the previous case of 49% duty-cycle and 0.33λ phase depth, the rms wavefront errors from the surface roughness for the zero and non-zero orders would be 0.00244 nm and 0.00086 nm, respectively.

Pattern distortion:

Another wavefront error, known as CGH pattern position error, or pattern distortion, results from the displacement of the etched pattern in a CGH from its ideal position. The wavefront error produced by the CGH pattern distortions is proportional to the diffraction order number and inversely proportional to the local fringe spacing [**Fercher 1976**]. It is expressed as

$$\Delta W(x,y) = -m\lambda \frac{\varepsilon}{S}, \qquad (\mathbf{6.17})$$

where ε is the grating position error in the direction perpendicular to the pattern. CGH pattern distortion errors do not affect the zero order diffracted beam. The pattern distortion can be measured by interferometric measurement as shown in Figure **6.11** [**Wyant 1974a and Ono 1984**]. Two plane wavefronts are incident on the hologram at the appropriate angle such that the $+N^{th}$ order produced by beam 1 and $-N^{th}$ order produced by beam 2 can pass through the spatial filter. The hologram has straight line patterns. The number of interference fringes by the $\pm N^{th}$ orders can be changed by adjusting the angles of the two wavefront. The moiré fringes formed by the interference fringes of these two wavefronts and straight lines drawn on a hologram have the information of the pattern distortion.

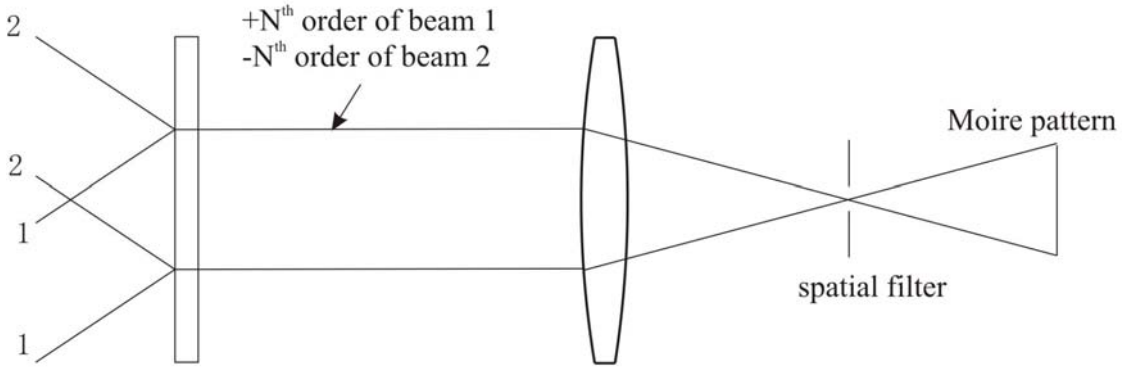


Figure **6.11**: Optical setup of measuring the pattern distortion.

Another method to measure the pattern distortion is to imbed "Nikon" marks on a CGH. These are small "+" marks 45 μm in height with bar widths between 3 and 9 μm. Their actual locations are measured with a laser interferometric X-Y measuring machine [**Arnold 1995**]. This method is often used by CGH vendors. We have made 20 CGHs and the statistics of the writing errors are listed in Figure **6.12**.

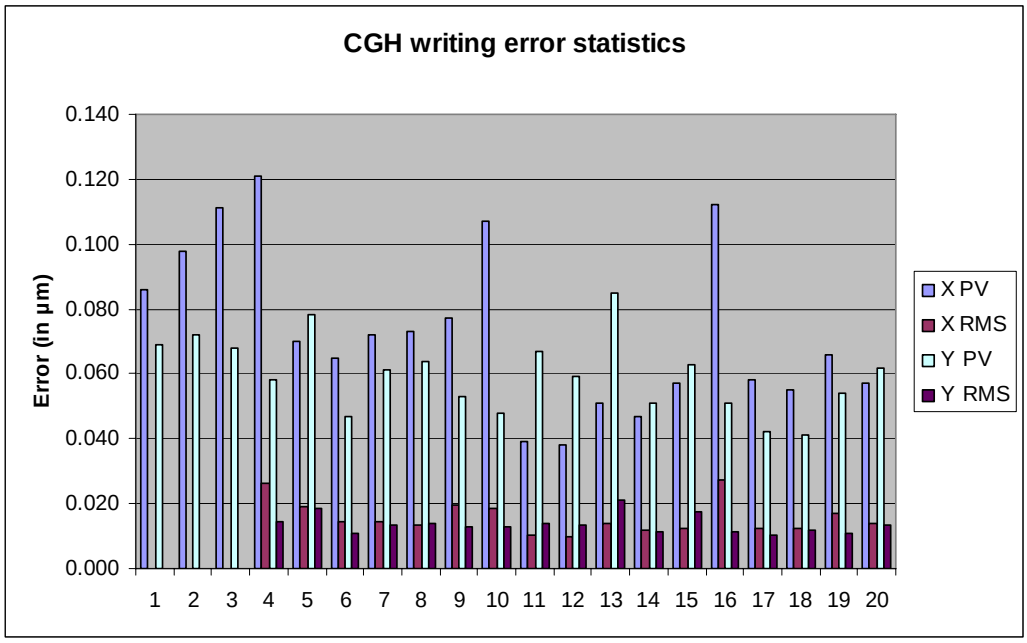


Figure **6.12**: CGH writing error statistics. The worst case for the writing errors is 0.12 µm PV and 0.027 µm rms. The first three holograms have no RMS errors provided by the vendor.

It is shown that the pattern distortion ε of 0.1 µm PV is achievable based on current CGH fabrication techniques. For a CGH with an average line spacing of 20 µm, the pattern distortion gives a wavefront error of 0.005λ PV for the first diffraction order.

6.3.2 Example: error analysis of a phase etched CGH

The fabrication non-uniformities of a 5-inch phase etched CGH used for testing an off-axis parabola (OAP) were measured. The CGH was designed with a duty-cycle of 50% and an etching depth that gives 0.33λ for the transmitted phase depth. Phase depth of 0.33λ is an optimized parameter that gives good diffraction efficiencies with less

wavefront errors from fabrication non-uniformities [**Zhou 2007b**]. Table **6.1** shows that for 50% duty-cycle, the wavefront sensitivities to phase depth are the same for both the zero order and the first order. By subtracting the first order measurement from the zero order one, the wavefront errors due to phase depth can be cancelled [**Zhou 2007c**].

The CGH substrate error was measured using the zero order diffraction of the CGH. A 4D Technology PhaseCam 4010 interferometer with an F/7 transmission objective was used in a test setup similar to that shown in Figure **6.4**. The transmitted wavefront, shown in Figure **6.13** had an error of 13.6 nm rms. This error is mainly the substrate error, with some phase errors from the variation in duty-cycle, etching depth and amplitude. It will be shown later that the phase errors due to fabrication non-uniformity are small for this CGH.

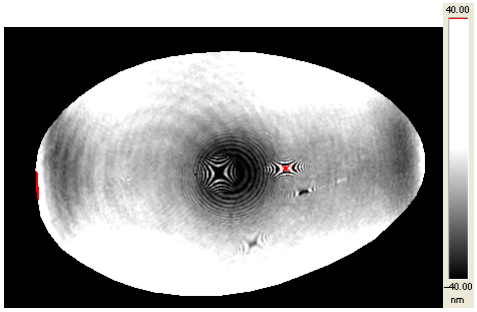


Figure **6.13**: Transmitted zero order wavefront of a 5 inch phase etched CGH, showing 13.6 nm rms of error due to the substrate.

The duty-cycle and etching depth variations were estimated using measured diffraction efficiencies up to the 11^{th} order. The CGH was illuminated with a 1 mm diameter collimated HeNe laser beam. A Newport low-power 918-UV detector and a 2930-C power meter were used to measure the power of each diffraction order. Five

points on the CGH were chosen to measure the fabrication non-uniformities. For each power measurement, 500 data points with an interval time of 20 ms were averaged by the power meter. The background light was subtracted from each measurement. A non-linear least-squares fit was applied to find the duty-cycles and etching depths of these five positions, shown in Table **6.3**. The RMS variations of duty-cycle and phase depth were determined to be 0.18% and 0.67%. Note that the RMS variation refers to the standard deviation of all the sample points, and RMS variation in percentage is the standard deviation divided by the nominal value.

Table **6.3**: The measured duty-cycle and etching depth of the five positions on the hologram.

	D	ϕ (waves)
Point 1	0.4897	0.3274
Point 2	0.4906	0.3315
Point 3	0.4897	0.3317
Point 4	0.4904	0.3326
Point 5	0.4881	0.3275
Average	0.4897	0.33014
RMS variation	0.0009	0.0022
RMS variation in percentage	0.18%	0.67%

For Point 1 on the CGH, the measured intensities for different orders and a fit is given in Figure **6.14**. The fitted curve is plotted using the fitted duty-cycle and etching depth. The model listed in Table **6.1** shows that the period of oscillations in intensity as a function of order is determined by the duty-cycle, while both duty-cycle and etching depth determine the amplitude of oscillations.

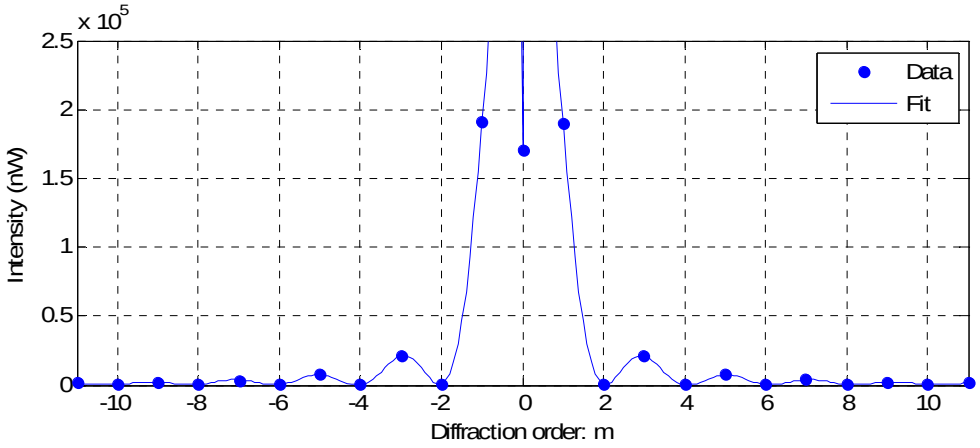


Figure **6.14**: The measured intensities for different diffraction orders and its fit.

The surface roughness of the etched and unetched areas was measured with a WYKO NT2000 interference microscope. Table **6.4** shows that the amplitude variation due to scattering is 0.0033%, which is small enough to be neglected.

Table **6.4**: Surface roughness measurement on a WYKO NT2000 interference microscope.

	Unetched area		Etched area		Ratio of amplitudes $\lvert A_0/A_1 \rvert$
	RMS roughness (nm)	I_{scat}	RMS roughness (nm)	I_{scat}	
Point 1	0.950	0.00002	2.188	0.00012	1.000048
Point 2	1.090	0.00003	1.540	0.00006	1.000015
Point 3	1.110	0.00003	1.840	0.00008	1.000027
Point 4	2.000	0.00010	2.663	0.00017	1.000038
Point 5	1.738	0.00007	2.169	0.00012	1.000021
				Average	1.000030
				RMS variation	0.000033
			RMS variation in percentage		0.000033

The first order wavefront of the CGH is used to test the OAP. Table **6.5** lists all the wavefront errors from the CGH. The estimated RSS wavefront error of the first order measurement is 13.6 nm. Duty-cycle non-uniformities have no impact on the wavefront error for the first order diffraction. Among all the fabrication errors, the substrate error is dominant. To achieve higher measurement accuracy, the substrate errors need to be characterized.

Table **6.5**: Wavefront errors from CGH fabrication non-uniformities without substrate calibration. Pattern distortion is not included.

Source of Errors	**Description**	**Measured variation**	**Sensitivities**	**RMS wavefront Errors at 1st order**
CGH substrate	Transmitted wavefront error	13.6 nm	1:1	13.6 nm
Etching depth	Effect on diffracted wavefront	0.67%	$\frac{\partial \Psi_{m=1}}{\partial \phi}$	0.700 nm
Duty-cycle	Effect on diffracted wavefront	0.18%	0	0
Amplitude ratio	Effect on diffracted wavefront	0.0033%	$\frac{\partial \Psi_{m=1}}{\partial (A_0 / A_1)}$	0.001 nm
Root-Sum-Squares Error:				**13.6 nm**

The CGH substrate errors can be removed by subtracting the zero order measurement from the OAP wavefront measurement, but the wavefront errors caused by the fabrication non-uniformities cannot be totally removed. Using the fabrication errors measured above, the residual wavefront errors can be determined by assessing both the zero order and the first order diffraction wavefronts, and then subtracting one from the other for each error source. Table **6.6** lists the residual wavefront errors for each error

source. The sensitivity functions can be found in Table **6.1**. The RSS wavefront error drops from 13.6 nm to 0.3 nm.

Table **6.6**: Wavefront errors from CGH fabrication errors after subtracting the zero order measurement.

Source of Errors	Measured variation	Sensitivities	RMS wavefront Errors
Etching depth error	0.67%	$\frac{\partial \Psi_{m=1}}{\partial \phi} - \frac{\partial \Psi_{m=0}}{\partial \phi}$	0.056 nm
Duty-cycle error	0.18%	$-\frac{\partial \Psi_{m=0}}{\partial D}$	0.300 nm
Amplitudes Error	0.0033%	$\frac{\partial \Psi_{m=1}}{\partial (A_0 / A_1)} - \frac{\partial \Psi_{m=0}}{\partial (A_0 / A_1)}$	0.004 nm
Root-Sum-Squares Error:			**0.3 nm**

Note that the pattern distortion error is not included in the above analysis, because we did not measure the pattern distorion. Assuming the pattern distortion error is 0.1 μm PV for the 20 μm spacing and the wavelength is 632.8 nm, then the wavefront error caused by pattern distortion for the first order is 3.2 nm PV. Pattern distortion becomes the most important error after calibrating the CGH substrate error.

The error analysis of this phase CGH demonstrates that the CGH substrate is the primary error source. The wavefront errors caused by etching depth, duty-cycle and amplitude are determined to be small enough that we cannot measure them using the interferometer. These errors can be neglected in optical testing.

6.3.3 Demonstration of substrate calibration

To demonstrate that the substrate error can be calibrated by the zero order wavefront, a custom CGH with 30% duty-cycle and 0.35λ phase depth was fabricated, where the wavelength of interest is 632.8 nm. The CGH diameter is 15 mm, and it generates a spherical wavefront. A reference sphere is used to reflect the light back for both the zero order and first order measurements. The experimental setup is illustrated in Figure **6.15**. The zero order phase map was subtracted from the first order map to remove the CGH substrate error. The difference of these two phase maps should reflect only the fabrication errors in duty-cycle, etching depth, amplitude and pattern distortion. The transmission sphere does not introduce any errors in the measurement because they cancel out during the subtraction.

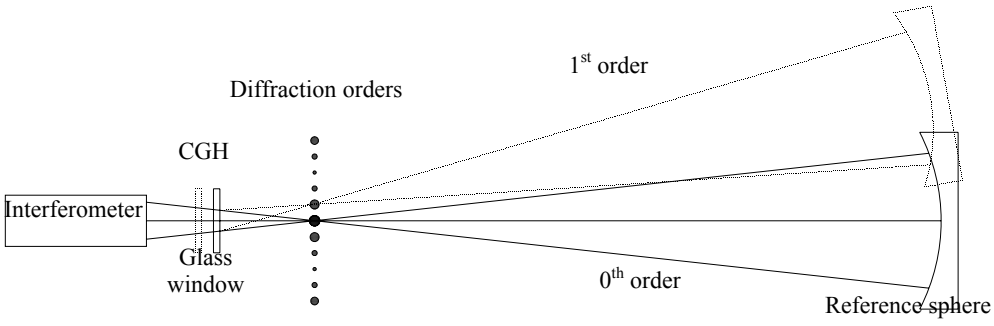


Figure **6.15**: Setup of the CGH substrate calibration. The first order measurement is subtracted from the zero order measurement to remove CGH substrate errors.

In order to increase the CGH substrate error, a glass window was placed between the interferometer and the CGH. Its wavefront distortion is equivalent to a CGH substrate error. Table **6.7** shows the measurement results of three configurations with

different amounts of equivalent substrate errors. Configuration 1 has no glass window in the setup, while configurations 2 and 3 have different window glasses. Up to about 15 nm substrate error is introduced in configuration 3. Tilt, power and astigmatism which come from system misalignment in the optical test, are removed from the measurement. For all cases, the difference of the first order and zero order wavefronts is about 1 nm rms. The wavefront errors from both fabrication non-uniformities and measurement noises contribute to this difference. This experiment shows that the substrate errors can be removed by subtracting the zero order wavefront from the first order one.

Table **6.7**: Measurements of three configurations with different amounts of equivalent substrate errors.

	RMS error: zero order (nm)	**RMS error: first order (nm)**	**Measured difference (nm)**
Configuration 1	1.677	1.648	0.714
Configuration 2	7.596	7.801	0.779
Configuration 3	16.82	17.43	1.291

6.4 Optimization of CGH designs to minimize the sensitivity to fabrication errors

Section **6.3** discussed methods to determine the fabrication uncertainties for a CGH, and provides a way to estimate the wavefront phase errors due to CGH fabrication errors. In this section, an optimal design for CGHs, based on the fabrication uncertainties, will be discussed. An optimally designed CGH generates a wavefront, which is the least sensitive to fabrication errors. The diffraction efficiency is also

optimized so that the light intensity returned to the interferometer is maximized. The optimal designs for both chrome-on-glass and phase etched CGHs are discussed in this section.

6.4.1 Chrome-on-glass CGH

A chrome-on-glass CGH used in transmission acts as a pure amplitude CGH. This is because there is no transmitted beam from the coated portion of the CGH, so the chrome coating either reflects or absorbs all the light; therefore, chrome thickness error does not affect the wavefront. CGH duty-cycle and amplitude errors also have no effect on wavefront for either the zero or other diffraction orders. The wavefront sensitivity function to duty-cycle and amplitude is zero for the zero and non-zero orders according to the equations listed in Table **6.1**.

The pure amplitude CGHs are not sensitive to variations in duty-cycle, chrome thickness or amplitude, leaving just the CGH substrate irregularities and pattern distortion errors. The overall wavefront accuracy using an amplitude CGH depends only on the pattern distortion errors, since substrate irregularities can be removed by calibration.

Since the fabrication errors on duty-cycle do not affect the wavefront phase for amplitude CGHs, a duty-cycle can be chosen to maximize the first order diffraction efficiency. A duty-cycle of 50% has the maximum diffraction efficiency of 10% at the first diffraction order, while the zero order has 25% diffraction efficiency. These results can be calculated from the equations in Table **6.1**.

When a chrome-on-glass CGH is used in reflection, A_0 and A_1 are both non-zero. The phase difference between the chrome and the glass must be considered. There is also a phase shift introduced between the interfaces of the metal and the air due to metal thickness and complex Fresnel reflections. For this case, both amplitude and phase effects must be considered when analyzing how the fabrication errors affect the wavefront phase. The following section discusses an example of phase CGHs, and provides an optimal design based on knowledge of the fabrication uncertainties.

6.4.2 Phase etched CGH

When testing a bare glass surface using chrome-on-glass CGHs in transmission, light passes through the CGH twice and only 0.04% ($10\% \times 10\% \times 4\% = 0.04\%$) of the light returns to the interferometer, which may not be enough in optical testing. Phase etched CGHs can have up to 40% diffraction efficiency at the first order, so the light back into the interferometer increases by a factor of 16. This makes phase CGHs useful for low-light-intensity optical testing.

For phase etched CGHs, the zero and non-zero order diffractions have different sensitivities to duty-cycle, phase depth, and amplitude. When we measure the CGH using the zero order diffraction and then subtract the result from the non-zero order surface measurement, there are residual wavefront errors left from the fabrication non-uniformities in duty-cycle, etching depth, and amplitude. Furthermore, all these fabrication errors are coupled together to affect the wavefront phases for both zero and

non-zero orders of diffraction. The parametric model introduced in Section **6.2.1** relates each fabrication error to the wavefront performance. To evaluate the overall effect of all the fabrication errors, the wavefront errors from each fabrication error are combined by a root-sum-square (RSS).

To determine wavefront errors caused by etching depth, duty-cycle, and amplitude variation, the PV variations in these parameters are assumed to be 2%, 1% and 0.5%, respectively. Additionally, both A_1 and A_0 are assumed to unity. The PV wavefront errors from CGH fabrication non-uniformities after subtracting the zero order wavefront can be calculated using the parametric model. It is shown in Table **6.8**.

Table **6.8**: Wavefront errors from fabrication variations after subtracting the zero order wavefront for substrate calibration.

Source of Errors	Fabrication variation	Wavefront error
Etching depth error	2%	$(\frac{\partial \Psi_{m=1}}{\partial \phi} - \frac{\partial \Psi_{m=0}}{\partial \phi}) \cdot \left(\frac{\Delta \phi}{\phi} \right) \cdot \phi$
Duty-cycle error	1%	$-\frac{1}{2\pi} \frac{\partial \Psi_{m=0}}{\partial D} \cdot \left(\frac{\Delta D}{D} \right) \cdot D$
Amplitude Error	0.5%	$\frac{1}{2\pi} (\frac{\partial \Psi_{m=1}}{\partial A_1} - \frac{\partial \Psi_{m=0}}{\partial A_1}) \cdot \left(\frac{\Delta A_1}{A_1} \right) \cdot A_1$

The plots in Figure **6.16** give the wavefront errors caused by the variations assumed above for different duty-cycles and etching depths. Figure **6.16 (a)** shows that the wavefront sensitivities to etching depth are the same for the zero and first orders at 50% duty-cycle. Therefore, the wavefront error is zero at 50% duty-cycle after subtracting the zero order wavefront from the first order. The variation of duty-cycle

only affects the wavefront for the zero order diffraction. Figure **6.16 (d)** shows the RSS wavefront errors caused by the three fabrication error sources. It shows that when the phase depth is close to 0.5λ, the wavefront error increases dramatically because the diffraction efficiency drops quickly at the zero order diffraction. At 50% duty-cycle and 0.5λ phase depth, the wavefront phase is extremely sensitive to duty-cycle and etching depth at the zero order because there is also no light from that order. When designing a CGH, a phase depth close to 0.5λ should be avoided unless the duty-cycle is away from 50%.

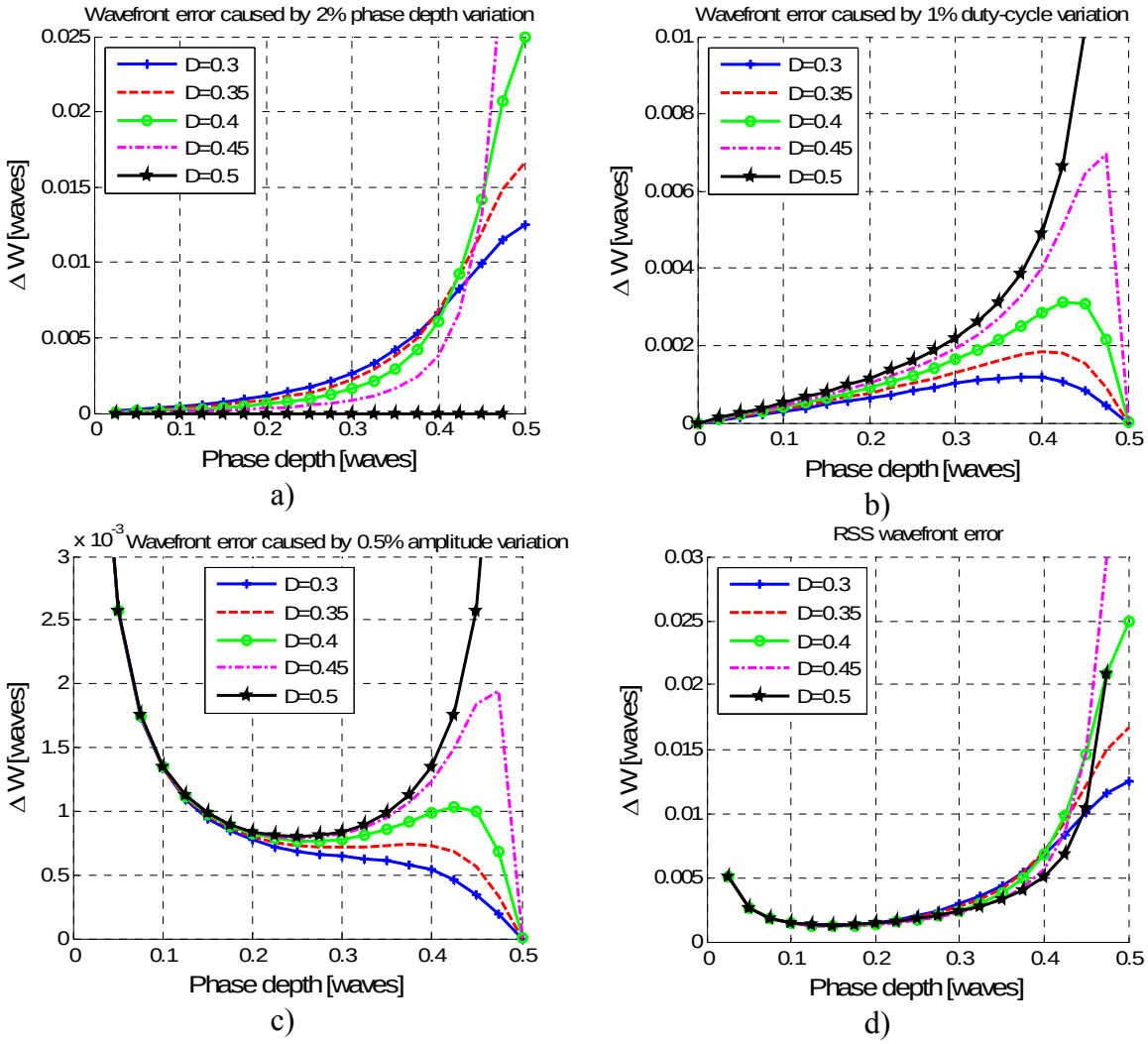


Figure **6.16**: a) Wavefront error caused by 2% phase depth variation; b) Wavefront error caused by 1% duty-cycle variation; c) Wavefront error caused by 0.5% amplitude variation; d) RSS wavefront error caused by 2% phase depth variation, 1% duty-cycle variation and 0.5% amplitude variation.

Assuming that the first order of diffraction is used for testing the surface, Figure **6.17 (a)** shows that the diffraction efficiency monotonically increases with duty-cycle up to 50% and phase depth up to 0.5λ. Larger phase depths have higher diffraction

efficiency but larger wavefront errors. The maximum diffraction efficiency of the first order occurs at 0.5λ phase depth.

Figure **6.17 (b)** shows the relationship between the first order diffraction efficiency and the RSS wavefront error for the case of 2% phase depth error, 1% duty-cycle error and 0.5% amplitude error. The RSS wavefront error is plotted on a log scale. Each curve represents a different duty-cycle. The points on each curve represent different phase depths. It can be shown that when the phase depth is less than 0.4λ, the diffraction efficiency increases rapidly as the phase depth increases, while the RSS wavefront error does not change much. When the phase depth is greater than 0.4λ, the RSS wavefront error increases significantly, but the diffraction efficiency remains almost unchanged. Using the previously assumed CGH fabrication errors, a duty-cycle of 50% and phase depths between 0.3λ and 0.4λ are reasonable designs, because the diffraction efficiencies are greater than 25% and the wavefront errors are relatively small.

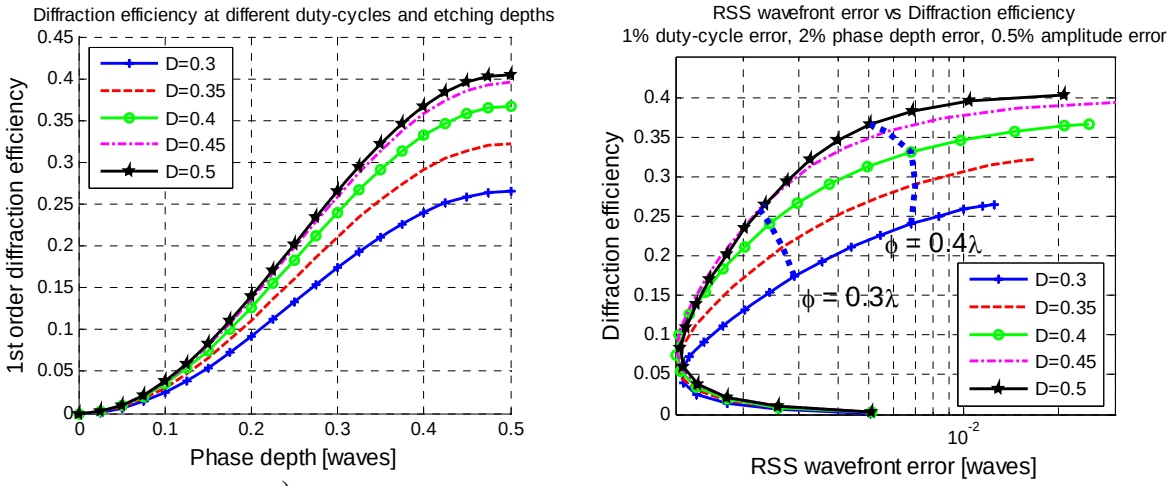


Figure **6.17**: a) Diffraction efficiency at different duty-cycles and phase depths; b) Relationship between RSS wavefront error and the first order diffraction efficiency (Each point on the curve represents different etching depth).

When designing a CGH, it is important to know the magnitude of fabrication uncertainties during the CGH writing process. Based on those fabrication uncertainties, a plot like Figure **6.17 (b)**, showing the relationship between diffraction efficiency and RSS wavefront error, can be obtained. Depending on the application, there is a tradeoff between diffraction efficiency and wavefront errors. This plot helps optical engineers to determine the CGH parameters that are optimal for their applications.

The amplitude errors due to scatter are usually small in a CGH. Phase depth and duty-cycle are typically the main errors. Consider a case where duty-cycle errors dominate, with 5% duty-cycle variation and 1% phase depth variation. The relationships between the RSS wavefront error and diffraction efficiency are shown in Figure **6.18 (a)**. For this case, CGHs with a larger duty-cycle have larger wavefront errors. This can be seen in Figure **6.16 (b)**, where the larger duty-cycles always have more wavefront errors.

Therefore, when the duty-cycle error is dominant in CGHs, a duty-cycle of 40% with a phase depth of 0.5 λ is preferable. With 40% duty-cycle, the zero order measurement can be made accurately with a phase depth of 0.5λ. However, all the curves in Figure **6.18** are very closely spaced, which means that the wavefront error and diffraction efficiency are not very sensitive to duty-cycle as long as the design is not close to 50% duty-cycle and 0.5 λ phase depth.

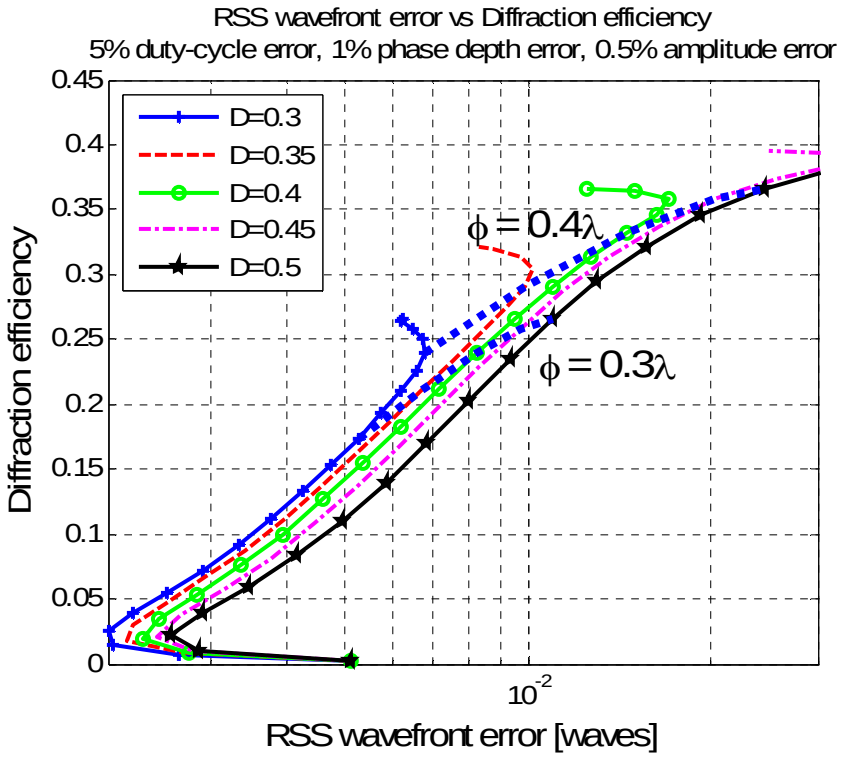


Figure **6.18**: RSS wavefront error vs. the first order diffraction efficiency (duty-cycle error dominant). Each point on the curve represents different phase depth.

For the case where phase depth error dominates, consider 1% duty-cycle variation and 5% phase depth variation. CGHs with 50% duty-cycle have less wavefront error and higher diffraction efficiency. This is because the difference of wavefront errors caused

by phase depth between the zero order and first order is zero for 50% duty-cycle, as shown in Figure **6.16** (a). The relationships between the RSS wavefront error and diffraction efficiency are shown in Figure **6.19**, which shows that a duty-cycle of 50% with a phase depth ranging from 0.3λ to 0.4λ is preferable.

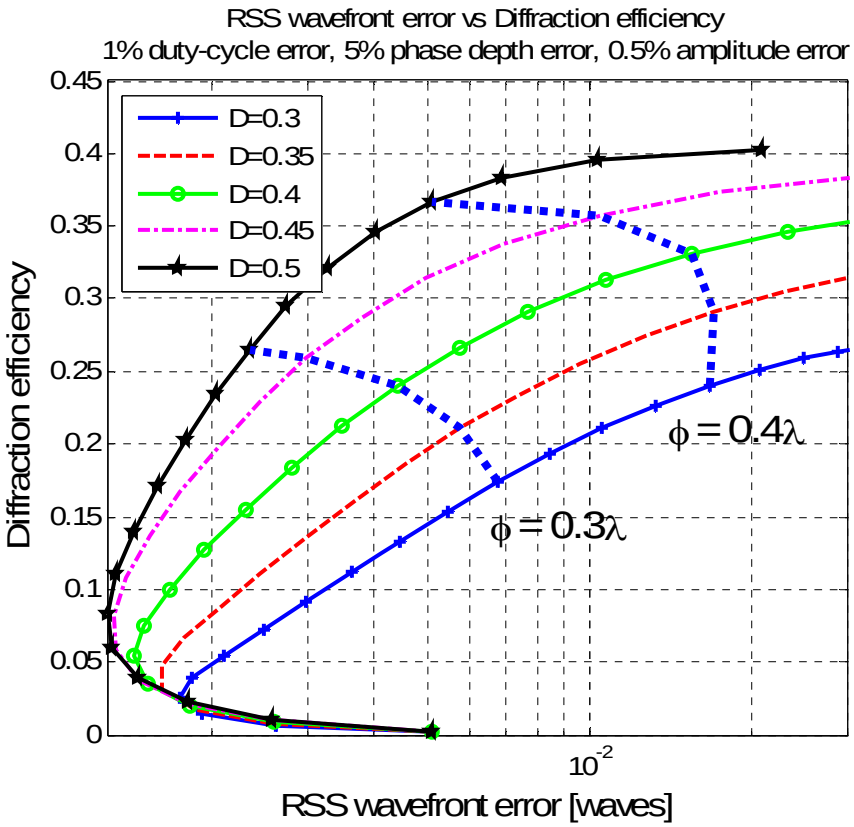


Figure **6.19**: RSS wavefront error vs. the first order diffraction efficiency (phase depth error dominant). Each point on the curve represents different etching depth.

In conclusion, for designing a phase etched CGH when the magnitude of fabrication errors are known, the duty-cycle and phase depth can be optimized to achieve best performances. If the fabrication errors are unknown, then a 50% duty-cycle and 0.35λ phase depth should be chosen. In this way, the CGH will not suffer large

wavefront error and low diffraction efficiency, no matter what fabrication error is dominant.

6.4.3 Example of estimating the total errors using a CGH in optical testing

The optimal design method and calibration of CGHs have been discussed in the previous sections. In this section, a typical example of estimating the total wavefront errors using a CGH to test some optic is provided.

According to current CGH fabrication techniques, the likely errors in duty-cycle, phase depth and amplitude are 0.5% rms, 1% rms, and 0.1% rms, respectively. If we design a CGH with 50% duty-cycle and 0.35λ phase depth, the diffraction efficiency will be about 30%. Assuming that the CGH is used at the red He-Ne wavelength (632.8 nm), the RMS wavefront errors from variations in duty-cycle, phase depth and amplitude will be 0.99 nm, 0 nm and 0.12 nm. The wavefront error from the phase depth is zero, because the wavefront sensitivities to phase depth are the same for both the zero order and the first order. However, the nominal duty-cycle may not be exactly 50% due to fabrication errors. Assuming that the actual measured duty-cycle is 49%, the wavefront sensitivities to phase depth are no longer the same for the zero order and the first order, which will cause 0.11 nm wavefront errors. To distinguish this error from the 0 nm errors from phase depth, we call it the 2^{nd} order phase depth error.

The wavefront errors from pattern distortion will be 0.63 nm rms, if the average line spacing of the CGH is 20 μm and the pattern distortion is 0.02 μm rms.

The measurement uncertainties in both CGH substrate calibration measurement and surface measurement also contribute to the total wavefront errors. The RMS wavefront errors from each error source are listed in Table **6.9**. If all error sources are characterized using the methods described in this chapter, the accuracy of using a CGH to measure a surface is about 2.2 nm.

Table **6.9**: Estimate of total RMS wavefront errors of using a CGH.

Sources of error	RMS Uncertainties	RMS Wavefront errors
CGH errors		
Duty-cycle error	0.50%	0.99 nm
1st order etching depth error	1.00%	0.00 nm
2nd order etching depth error	1.00%	0.11 nm
Amplitude error	0.10%	0.12 nm
Substrate error		0.00 nm
Pattern distortion	0.02 um	0.63 nm
Measurement errors		
Uncertainty in CGH substrate calibration		1.60 nm
Uncertainty in test surface measurement		1.00 nm
Root-Sum-Squares error		2.23 nm

CHAPTER 7

CONCLUSION

Random errors and systematic errors in the interferometric surface measurements are studied from the perspective of an end user of commercial phase shifting interferometers. A method of estimating the random errors in a single measurement or an average map was provided, which is useful in determining the number of measurements required to achieve the desired accuracy based on the measurement environment. Systematic errors include geometric errors and errors due to diffraction effects. In this dissertation, we concentrated on the study of diffraction effects in interferometry. Fresnel diffraction and Talbot imaging theories were used to quantify the phase smoothing effect and edge diffraction. We described the diffraction effects in terms of Fresnel number, which can simplify the complex geometric propagation to a single parameter. An ASAP simulation was run to verify our analysis. A remaining area of future work is to conduct an experiment and demonstrate the diffraction effects with an interferometer to validate our analysis.

This dissertation reviewed several interferometer calibration techniques, also called absolute tests. We used the random ball test as an example to study the remaining errors after the interferometer calibration. Three aspects of the remaining errors were studieds: random error, geometric error, and error due to diffraction effects. This provided an example of estimating the errors in other interferometric calibration methods.

This dissertation also addressed the interferometer phase modulation function. That is useful information and can be used to retrieve some middle spatial frequency phase information. An example of obtaining the phase modulation function for an IR interferometer and using the Wiener filter on the experiment data was given. The fundamental difficulty of using the Wiener filter is to estimate the power spectrum of noise completely. Developing a comprehensive technique for assessing noise and using it to develop optimal filtering for interferometer data should remain as future work.

We present a smoothing filter method that can be used in map registration and subtraction. This method requires smoothing the map to be subtracted to a certain resolution first and then subtracting the smoothed map from the test. This can reduce the sensitivity to registration error and improve the overall measurement accuracy.

This dissertation also provides the wavefront sensitivities to CGH manufacturing errors, and develops a parametric model that can be used for optimizing the CGH design to minimize the wavefront sensitivity to fabrication errors while giving good diffraction efficiency. This parametric model is a very useful tool for engineers in determining the design parameters of the CGH.

APPENDIX A

DERIVATION OF EFFECTIVE PROPAGATION DISTANCE FOR WAVEFRONTS WITH DIFFERENT SIGNS OF CURVATURE

A converging wavefront starting with radius of curvature R_1, diameter $2a_1$, and ripples with period p_1, propagates to a position where it has radius of curvature R_2 with a negative sign. To convert this propagation to equivalent propagation in collimated space, we evaluate the case where a lens with arbitrary focal length f is used to convert the light into collimated light. Then we use geometric imaging relationships to calculate equivalent dimensions.

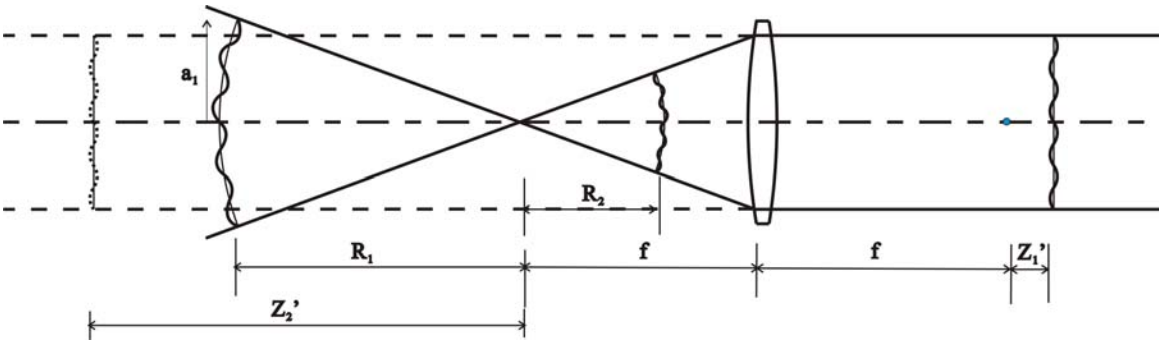


Fig. **A.1:** Propagation in converging space is converted to equivalent propagation in collimated space when the two wavefronts have opposite signs.

In the collimated space, the equivalent propagation distance is

$$\Delta Z' = Z_2' - Z_1' = f^2\left(\frac{1}{R_2} - \frac{1}{R_1}\right), \qquad \textbf{(A.1)}$$

and the phase ripple has a period of

$$p' = \frac{f}{R_1} p_1 . \qquad \textbf{(A.2)}$$

The attenuation of the phase ripple in the equivalent collimated space becomes

$$W' = W \cdot \cos\left(\frac{\pi\lambda \cdot \Delta Z'}{p'^2} \right) = W \cdot \cos\left(\frac{\pi\lambda R_1 \cdot (R_1 - R_2)}{R_2 \cdot p_1^2} \right) = W \cdot \cos\left(\frac{\pi\lambda L_e}{p_1^2} \right), \qquad \textbf{(A.3)}$$

where L_e is the effective propagation distance. The effective propagation distance in this case is the same as the wavefronts with the same sign of curvature. Therefore, all the equations derived in Chapter 2 also work when the wavefronts have opposite signs.

APPENDIX B

ZERNIKE POLYNOMIALS

Zernike polynomials are composed of a product of a function with radial dependence and one with azimuthal dependence. The radial function is $R_n^m(r)$ and the azimuthal function is given by either $\cos(m\theta)$ or $\sin(m\theta)$. The Zernike polynomials are ordered by an index j such that an even value for j corresponds to a polynomial with angular dependence $\cos(m\theta)$ and an odd value for j corresponds to $\sin(m\theta)$:

$$Z_j(r,\theta) = \begin{cases} \sqrt{n+1}R_n^m(r)\sqrt{2}\cos(m\theta) & \text{even } j, m \neq 0 \\ \sqrt{n+1}R_n^m(r)\sqrt{2}\sin(m\theta) & \text{odd } j, m \neq 0 \\ \sqrt{n+1}R_n^0(r) & m = 0 \end{cases}, \qquad \textbf{(B.1)}$$

where the radial function is

$$R_n^m(r) = \sum_{s=0}^{(n-m)/2} \frac{(-1)^s (n-s)!}{s![(n+m)/2-s]![(n-m)/2-s]!} r^{n-2s}. \qquad \textbf{(B.2)}$$

j is the general index of Zernike polynomials,

n is the power of the radial coordinate r,

m is the multiplication factor of the angular coordinate θ, and

n and m have the following relations: $m \leq n$ and $(n - m)$ is even.

The first 37 Zernike polynomials using Noll's numbering is listed below.

j	*n, m*	*Zernikes*	*j*	*n, m*	*Zernikes*
1	*n*=0 *m*=0	$Z_1 = 1$	20	*n*=5 *m*=5	$Z_{20} = \sqrt{12} r^5 \cos 5\theta$
2	*n*=1 *m*=1	$Z_2 = 2r\cos\theta$	21	*n*=5 *m*=5	$Z_{21} = \sqrt{12} r^5 \sin 5\theta$
3	*n*=1 *m*=1	$Z_3 = 2r\sin\theta$	22	*n*=6 *m*=0	$Z_{22} = \sqrt{7}(20r^6 - 30r^4 + 12r^2 - 1)$
4	*n*=2 *m*=0	$Z_4 = \sqrt{3}(2r^2 - 1)$	23	*n*=6 *m*=2	$Z_{23} = \sqrt{14}(15r^6 - 20r^4 + 6r^2)\sin 2\theta$
5	*n*=2 *m*=2	$Z_5 = \sqrt{6} r^2 \sin 2\theta$	24	*n*=6 *m*=2	$Z_{24} = \sqrt{14}(15r^6 - 20r^4 + 6r^2)\cos 2\theta$
6	*n*=2 *m*=2	$Z_6 = \sqrt{6} r^2 \cos 2\theta$	25	*n*=6 *m*=4	$Z_{25} = \sqrt{14}(6r^6 - 5r^4)\sin 4\theta$
7	*n*=3 *m*=1	$Z_7 = \sqrt{8}(3r^3 - 2r)\sin\theta$	26	*n*=6 *m*=4	$Z_{26} = \sqrt{14}(6r^6 - 5r^4)\cos 4\theta$
8	*n*=3 *m*=1	$Z_8 = \sqrt{8}(3r^3 - 2r)\cos\theta$	27	*n*=6 *m*=6	$Z_{27} = \sqrt{14} r^6 \sin 6\theta$
9	*n*=3 *m*=3	$Z_9 = \sqrt{8} r^3 \sin 3\theta$	28	*n*=6 *m*=6	$Z_{28} = \sqrt{14} r^6 \cos 6\theta$
10	*n*=3 *m*=3	$Z_{10} = \sqrt{8} r^3 \cos 3\theta$	29	*n*=7 *m*=1	$Z_{29} = \sqrt{16}(35r^7 - 60r^5 + 30r^3 - 4r)\sin\theta$
11	*n*=4 *m*=0	$Z_{11} = \sqrt{5}(6r^4 - 6r^2 + 1)$	30	*n*=7 *m*=1	$Z_{30} = \sqrt{16}(35r^7 - 60r^5 + 30r^3 - 4r)\cos\theta$
12	*n*=4 *m*=2	$Z_{12} = \sqrt{10}(4r^4 - 3r^2)\cos 2\theta$	31	*n*=7 *m*=3	$Z_{31} = \sqrt{16}(21r^7 - 30r^5 + 10r^3)\sin 3\theta$
13	*n*=4 *m*=2	$Z_{13} = \sqrt{10}(4r^4 - 3r^2)\sin 2\theta$	32	*n*=7 *m*=3	$Z_{32} = \sqrt{16}(21r^7 - 30r^5 + 10r^3)\cos 3\theta$
14	*n*=4 *m*=4	$Z_{14} = \sqrt{10} r^4 \cos 4\theta$	33	*n*=7 *m*=5	$Z_{33} = \sqrt{16}(7r^7 - 6r^5)\sin 5\theta$
15	*n*=4 *m*=4	$Z_{15} = \sqrt{10} r^4 \sin 4\theta$	34	*n*=7 *m*=5	$Z_{34} = \sqrt{16}(7r^7 - 6r^5)\cos 5\theta$
16	*n*=5 *m*=1	$Z_{16} = \sqrt{12}(10r^5 - 12r^3 + 3r)\cos\theta$	35	*n*=7 *m*=7	$Z_{35} = \sqrt{16} r^7 \sin 7\theta$
17	*n*=5 *m*=1	$Z_{17} = \sqrt{12}(10r^5 - 12r^3 + 3r)\sin\theta$	36	*n*=7 *m*=7	$Z_{36} = \sqrt{16} r^7 \cos 7\theta$
18	*n*=5 *m*=3	$Z_{18} = \sqrt{12}(5r^5 - 4r^3)\cos 3\theta$	37	*n*=8 *m*=0	$Z_{37} = \sqrt{9}(70r^8 - 140r^6 + 90r^4 - 20r^2 + 1)$
19	*n*=5 *m=3*	$Z_{19} = \sqrt{12}(5r^5 - 4r^3)\sin 3\theta$			

APPENDIX C

SPECIAL FUNCTIONS

In this appendix, we list a number of functions used in this dissertation. They are all real-valued, one-dimensional functions of the real independent variable x. In our notation, x_0 is a real constant that determines the position of the function along the x-axis, and the real constant b is a scaling factor.

The rectangle function is defined by

$$\operatorname{rect}\left(\frac{x-x_0}{b}\right)=\begin{cases}0, & \left|\dfrac{x-x_0}{b}\right|>\dfrac{1}{2}\\ \dfrac{1}{2}, & \left|\dfrac{x-x_0}{b}\right|=\dfrac{1}{2}\\ 1, & \left|\dfrac{x-x_0}{b}\right|<\dfrac{1}{2}\end{cases}. \tag{C.1}$$

The sinc function is defined by

$$\operatorname{sinc}\left(\frac{x-x_0}{b}\right)=\frac{\sin\pi\left(\dfrac{x-x_0}{b}\right)}{\pi\left(\dfrac{x-x_0}{b}\right)}. \tag{C.2}$$

The comb function is defined by

$$\operatorname{comb}\left(\frac{x-x_0}{b}\right)=|b|\sum_{n=-\infty}^{\infty}\delta\left(x-x_0-nb\right). \tag{C.3}$$

REFERENCES

Ai, C. and J. C. Wyant, "Effect of spurious reflection on phase shift interferometry," Appl. Opt. Vol.27, No.14, 3039-3045 (1988)

Ai, C. and J. C. Wyant, "Absolute testing of flats by using even and odd functions," Appl. Opt. Vol.32, 4698-4705 (1993)

Arnold, S. M., "How to test an asphere with a computer-generated hologram," in Holographic Optics: Optically and Computer Generated, I.N. Cindrich and S.H. Lee, Editors, Proc. SPIE 1052, 191-197 (1989)

Arnold, S. M., "An interferometer for aspheric testing: calibration and error compensation," Precision Interferometric Metrology, Proc. ASPE 5, 83-86 (1992)

Arnold, S. M., "Verification and certification of CGH aspheric nulls," Proc. SPIE 2536, 117-126 (1995)

Augustyn, W., "Versatility of a microprocessor-based interferometric data reduction system," Proc. Soc. Photo-Opt. Instrum. Eng., Vol.192 (1979)

Barrett, H. H., K. J. Myers, "*Foundation of image science*" Wiley Series in Pure and Applied Optics (2003)

Brock, N., J. Hayes, B. Kimbrough, J. Millerd, M. North-Mirris, M. Novak and J. C. Wyant, "Dynamic interferometry," Proc. SPIE 5875, 58750F (2005) Brophy, C. P., "Effect of intensity error correlation on the computed phase of phase-shifting interferometry," J. Opt. Soc. Am. A, 7, 537 (1990)

Bruning, J. H., "Fringe scanning interferometers," in Optical Shop Testing, D. Malacara, Ed., Wiley, New York (1978)

Burge, J. H., "Advanced techniques for measuring primary mirrors for astronomical telescopes," Ph.D. Dissertation, University of Arizona, Tucson (1993)

Chang, Y., "Diffraction wavefront analysis of computer-generated holograms," Ph.D. dissertation, University of Arizona, Tucson (1999a)

Chang, Y. and J. H. Burge, "Error analysis for CGH optical testing," Proc. SPIE 3782, 358-366 (1999b)

Chang, Y., P. Zhou, and J. H. Burge, "Analysis of phase sensitivity for binary computer-generated holograms," Appl. Opt. Vol.45, No.18, 4223-4234 (2006)

Chu, J., Q. Wang, J. P. Lehan, G. Gao, and U. Griesmann, "Measuring the Phase Transfer Function of a Phase-Shifting Interferometer", Proc. SPIE 7064, 70640C (2008)

Creath, K., "Phase-Measurement Interferometry Techniques," in Progress in Optics. Vol. XXVI, E. Wolf, Ed., Elservier Science Publishers, Amsterdam, 349-393 (1988)

Creath, K. and J. C. Wyant, "Holographic and speckle tests," in Optical Shop Testing, D. Malacara, Editor (Wiley, New York) 599-651(1992)

de Groot, P. and L. Deck, "Numerical simulations of vibration in phase-shifting interferometry," Appl. Opt., 35, 2172 (1996)

de Groot, P. and X. Colonna, "Interpreting interferometric height measurements using the instrument transfer function", Proceedings of the 6th International workshop on automatic processing of fringe patterns, Springer, Berlin, Germany, 30-37 (2006)

Evans, C. and R. N., Kestner, "Test optics error removal," Applied Optics, Vol.37, No.7, 1015-1021 (1996)

Fercher, A. F., "Computer-generated holograms for testing optical elements: error analysis and error compensation," Opt. Acta 23, 347-365 (1976)

Frieden, B. R., "Statistical, Optics, and Data Testing: A Problem Solving Approach," 3rd Edition, Springer-Verlag (2001)

Fritz, B. S., "Absolute calibration of an optical flat," Opt. Eng. 23, 379-383 (1984)

Gardner, N. and A. Davies, "Retrace error evaluation on a figure-measuring interferometer," Proc. SPIE 5869, 58690V (2005)

Gaskill, J., "*Linear System, Fourier Transform, and Optics,*" John Wiley & Son, Inc., N.Y. (1978)

Gaylord, T. K. and M. G. Moharam, "Analysis and applications of optical diffraction by gratings," Proc. IEEE 73, 864-937 (1985)

Goodman, J. W. "*Introduction to Fourier Optics*" Roberts & Company (2004)

Griesmann, U., Q. Wang, J. Soons, and R. Carakos, "A simple ball averager for reference sphere calibrations," Proc. SPIE 5869, 58690S (2005)

Griesmann, U., Q. Wang, and J. Soons, "Three-flat tests including mounting-induced deformations," Optical Engineering 46(9), 093601 (2007)

Han, S., E. Novak, and M. Schurig, "Ritchey-Common Test used fro Measurement of Astronomical Optic," Proc. SPIE Vol.4842, 270-273 (2003)

Harris, J. S., "The Universal Fizeau Interferometer," Ph.D. Thesis, U. Reading, U.K. (1971)

Hopkins, J. S., "Applied Optics at Reading," Opt. Laser Technol. 2, 158-163 (1971)

Ishii, Y., J. Chen, and K. Murata, "Digital phase-measuring interferometry with a tunable laser diode," Opt. Lett., 12, 233 (1987)

Jensen, A. E., "Absolute Calibration Method for Laser Twyman-Green Wave-Front Testing Interferometers," J. Opt. Soc. Am. 63, 1313A (1973)

Kallioniemi I., Saarinen, J., Blomstedt, K. and Turunen, J., "Polygon approximation of the fringes of diffractive elements," Appl. Opt. 36, 7217-7223 (1997)

Koliopoulos, C., "Interferometric optical phase measurement techniques," Ph.D. dissertation, University of Arizona, Tucson (1981)

Liu, D., Y. Yang, C. Tian, Y. Luo and L. Wang, "Practical methods for retrace error correction in nonnull aspheric testing," OPTICS EXPRESS Vol.17, No.9, 7025-7035 (2009)

Loomis, J. S., "Computer-generated holography and optical testing," in Recent Advances in Holography, T.C. Lee and P.N. Tamura, Editors, Proc. SPIE 2115, 59-69 (1980)

Lohmann, A. W. and D. P., Paris, "Binary Fraunhofer holograms, generated by computer," Appl. Opt. 6, 1739-1748 (1967)

Malacara, D., "*Optical Shop Testing*," Wiley series in pure and applied optics (1991)

Martin, H. M., J. H. Burge, S. M. Miller, B. K. Smith, R. Zehnder, C. Zhao, "Manufacture of a 1.7 m prototype of the GMT primary mirror segments", Proc. SPIE 6273, 62730G (2006)

Martin, H., "Specifications for the 3.75 m fold sphere for the GMT principle test," Steward observatory internal report (2008)

Millerd, J., B. Neal, J. Hays, M. North-Morris, M. Novak and J. C. Wyant, "Pixelated phase-mask dtnamic interferometer," Proc. SPIE 5531 304-314 (2004)

Millerd, J., N. Brock, J. Hayes, B. Kimbrough, M. Novak, M. North-Morris, and J. C. Wyant, "Modern approaches in phase measuring metrology," Proc. SPIE 5856, 14-22 (2005)

Murphy, P. E., T. G. Brown, and D. T. Moore, "Measurement and calibration of interferometric imaging aberrations,",Appied Optics Vol.39, No.34, 6421-6429 (2000)

Noll, R. J., "Zernike polynomials and atmospheric turbulence," J. Opt. Soc. Am., Vol.66, No.3, 207-211 (1976)

Novak., E., C. Ai, and J. C. Wyant, "Transfer function characterization of laser Fizeau interferometer for high spatial frequency phase measurements," Proc. SPIE Vol.3134, 114-121 (1997)

Ono, A. and J. C.Wyant, , "Plotting errors measurement of CGH using an improved interferometric method," Appl. Opt. 23, 3905-3910 (1984)

Ricks, D. W. and L.V. Chizek, "Light scattering from binary optics," Proc. SPIE 1211, 24-37 (1990)

Saxer, C. and K. Freischlad, "Interference microscope for sub Angstrom surface roughness measurement," Proc. SPIE 5144, 37-45 (2003)

Schulz G. and J. Schwider, "Interferometric testing of smooth surfaces," in Progress in Optics (North-Holland, Amsterdam), Vol.13, 93-167 (1976)

Selberg,.L. A., "Interferometer accuracy and precision", Proc. SPIE 1400, 24-32 (1990)

Siegel, N., "Algorithms for data registration and subtraction in optical testing," Master Thesis, University of Arizona, Tucson (2005)

Su, P., J. H. Burge, R. A. Sprowl, and J. Sasian, "Maximum likelihood estimation as a general method of combining sub-aperture data for interferometric testing," Proc. SPIE Vol.6342, 63421X (2006)

Sykora, D., "A model for cavity induced errors with wavefront slope in high accuracy spherical Fizeau metrology", OSA OF&T, OWB7 (2008)

Swanson, G. J., "Binary optics technology: theoretical limits on the diffraction efficiency of multilevel diffractive optical elements," technical report (Lincoln Library, Massachusetts Institute of Technology, 1991)

Takeda, M., "Temporal versus spatial carrier techniques for heterodyne interferometry," in Proceedings, 14th congress of the international commission of optics: Henri Arsenault, ed., Proc. SPIE 813, 329-330 (1987)

Tatsuno, K. and Y. Tsunoda, "Diode laser direct modulation heterodyne interferometer," Appl. Opt., 26, 37 (1987)

Totonto, M., "Review of Various MTF Measurement Methods Comparison of Bar Target Variance and Modulation," MS Thesis, University of Arizona (1995)

Parks, R. E., "Removal of test optics errors," in Advances in Optical Metrology I, N. Balasubramanian and J. C. Wyant, eds., Proc. SPIE 153, 56-63 (1978)

Parks, R., C. J. Evans and L. Shao, "Calibration of interferometer transmission spheres," in Optical Fabrication and Testing Workshop OSA Technical Digest Series 12, 80-83 (1998)

Parks, R. E., "A practical implementation of the random ball test," in Frontiers in Optics, OSA Technical Digest, OFMC12 (2006)

Womack, K. H., J. A. Jonas, C. Koliopoulos, K. L. Underwood, J. C. Wyant, J. S. Loomis, and C. R. Hayslett, "Microprocessor-based instrument for analyzing video interferograms," Proc. Soc. Photo-Opt. Instrum. Eng., Vol.192 (1979)

Wyant, J. C. and V. P. Bennett, "Using computer-generated holograms to test aspheric wavefronts," Appl. Opt. 11, 2833-2839 (1972)

Wyant, J. C., P. K., O'Neill and A. J., MacGovern, "Interferometric method of measuring plotter distortion," Appl. Opt. 13, 1549-1551(1974a)

Wyant, J. C. and P. K. O'Neil, "Computer-generated hologram, null lens test of aspheric wavefronts," Appl. Opt. 13, 2762-2765 (1974b)

Wyant, J. C., "Use of an ac heterodyne lateral shear interferometer with real-time wavefront correction systems," Appl. Opt., 14, 2622 (1975)

Wyant, J. C., Optical testing classnotes, University of Arizona, Tucson (1998)

Zhao, C. and J. H. Burge, "Imaging aberrations from null correctors," Proc. SPIE 6723, 67230L (2007a)

Zhao, C. and J. H. Burge, "Orthonomal Vector Polynomials in a Unit Circle, Part I: Basis Set Derived from Gradients of Zernike Polynomials," Optics Express, Vol.15, No.26 (2007b)

Zhao, C. and J. H. Burge, "Orthonomal Vector Polynomials in a Unit Circle, Part II: Completing the Basis Set," OPTICS EXPRESS, Vol.16, No.9 (2008)

Zhao, C. and J. H. Burge, "Orthonormal Vector Polynomials in a Unit Circle, Application: Fitting mapping distortions in a null test," Proc. SPIE Vol.7426, 74260V (2009a)

Zhao, C. and J. H. Burge, "Estimate of wavefront error introduced by encoding of computer generated holograms," CLEO-Pacific Rim (2009b)

Zhou, P. and J. H. Burge, "Coupling of surface roughness to the performance of computer-generated holograms," Appl. Opt. Vol.46, No.26, 6572-6576 (2007a)

Zhou, P. and J. H. Burge, "Optimal design of computer-generated holograms to minimize sensitivity to fabrication errors", Optics Express, Vol.15, No.23, 15410-15417 (2007b)

Zhou, P. and J. H. Burge, "Fabrication error analysis and experimental demonstration for computer-generated holograms," Appl. Opt. Vol.46, No.5 657-663 (2007c)

Zhou, P. and J. H. Burge, "Limits for interferometer calibration using the random ball test," Proc. SPIE Vol.7426 74260U (2009)