2014-06

Wavefront reconstruction and mirror surface optimization for adaptive optics

Axtell, Travis W.

Monterey, California: Naval Postgraduate School

NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

DISSERTATION

WAVEFRONT RECONSTRUCTION AND MIRROR SURFACE OPTIMIZATION FOR ADAPTIVE OPTICS

by

Travis W. Axtell

June 2014

Dissertation Supervisor: Roberto Cristi

Approved for public release; distribution is unlimited

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704–0188

1. AGENCY USE ONLY *(Leave Blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	06–06–2014	Dissertation 2009-09-28—2014-06-20

4. TITLE AND SUBTITLE

WAVEFRONT RECONSTRUCTION AND MIRROR SURFACE OPTIMIZATION FOR ADAPTIVE OPTICS

5. FUNDING NUMBERS

6. AUTHOR(S)

Travis W. Axtell

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Naval Postgraduate School
Monterey, CA 93943

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Department of the Navy

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. IRB Protocol Number: N/A

12a. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(maximum 200 words)*

The problem of wavefront reconstruction is important in high precision optical systems, such as astronomical telescopes, where it is used to estimate the distortion of the collected light caused by the atmosphere and corrected by adaptive optics. A generalized orthogonal wavelet wavefront reconstruction algorithm is presented in this research for use with gradient measurements from a Shack-Hartmann wavefront sensor. This algorithm can be implemented using a number of different wavelets for improved performance in the presence of noise. An extension of this algorithm is also presented to provide wavefront estimation in the presence of isolated branch points where the phase is undetermined. The wavefront is obtained by augmenting the wrapped observations with a filtered curl of the vector field. The wavefront estimation can then be used for surface control of a deformable mirror. A third contribution is in deformable mirror surface control. The control signals to a deformable mirror are computed that minimize the wavefront error using constrained optimization to ensure that the hardware actuator voltage limits are satisfied. A sequence of optimal solutions is used to verify the linear model of a deformable mirror. A multigrid approach to the optimization problem is shown to improve computation efficiency.

14. SUBJECT TERMS

Wavefront reconstruction, Adaptive optics, Wavelets, Atmospheric turbulence, Branch points, Mirror surface optimization, Space telescope, Segmented mirror

15. NUMBER OF PAGES 135

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UU

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2–89)
Prescribed by ANSI Std. 239–18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release; distribution is unlimited

WAVEFRONT RECONSTRUCTION AND MIRROR SURFACE OPTIMIZATION FOR ADAPTIVE OPTICS

Travis W. Axtell

B.S., Clemson University, 2007
M.S., Naval Postgraduate School, 2011

Submitted in partial fulfillment of the
requirements for the degree of

DOCTOR OF PHILOSOPHY IN ELECTRICAL ENGINEERING

from the

NAVAL POSTGRADUATE SCHOOL
June 2014

Author: Travis W. Axtell

Approved by:

Roberto Cristi Professor of Electrical and Computer Engineering Dissertation Advisor	Xiaoping Yun Professor of Electrical and Computer Engineering
Phillip Pace Professor of Electrical and Computer Engineering	Brij Agrawal Professor of Mechanical and Aeronautical Engineering
Jae-Jun Kim Professor of Mechanical and Aeronautical Engineering	Brett Borden Professor of Physics

Approved by: Clark Robertson
Chair, Department of Electrical and Computer Engineering

Approved by: Douglas Moses
Vice Provost for Academic Affairs

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

The problem of wavefront reconstruction is important in high precision optical systems, such as astronomical telescopes, where it is used to estimate the distortion of the collected light caused by the atmosphere and corrected by adaptive optics. A generalized orthogonal wavelet wavefront reconstruction algorithm is presented in this research for use with gradient measurements from a Shack-Hartmann wavefront sensor. This algorithm can be implemented using a number of different wavelets for improved performance in the presence of noise. An extension of this algorithm is also presented to provide wavefront estimation in the presence of isolated branch points where the phase is undetermined. The wavefront is obtained by augmenting the wrapped observations with a filtered curl of the vector field. The wavefront estimation can then be used for surface control of a deformable mirror. A third contribution is in deformable mirror surface control. The control signals to a deformable mirror are computed that minimize the wavefront error using constrained optimization to ensure that the hardware actuator voltage limits are satisfied. A sequence of optimal solutions is used to verify the linear model of a deformable mirror. A multigrid approach to the optimization problem is shown to improve computation efficiency.

THIS PAGE INTENTIONALLY LEFT BLANK

Table of Contents

List of Figures

List of Tables

THIS PAGE INTENTIONALLY LEFT BLANK

List of Acronyms and Abbreviations

1D	one-dimensional
2D	two-dimensional
3D	three-dimensional
AO	adaptive optics
DM	deformable mirror
DOD	Department of Defense
DSFT	discrete spatial Fourier transform
DSP	digital signal processing
DWT	discrete wavelet transform
FT	Fourier transform
IDWT	inverse discrete wavelet transform
LSI	linear shift-invariant
OPD	optical path difference
QMF	quadrature mirror filter
ROC	region of convergence
S-H WFS	Shack-Hartmann wavefront sensor
SMD	Segmented Mirror Demonstrator
SMT	Segmented Mirror Telescope

THIS PAGE INTENTIONALLY LEFT BLANK

Executive Summary

Optical systems, such as astronomical telescopes and laser weapons, are highly affected by perturbation from the propagating medium. The key problem addressed in this dissertation is how to estimate this perturbation of the optical waveform and compensate for its effects. In addition, optical wavefront reconstruction is used in many modern applications of engineering and science to provide insight into the performance of optical systems that can be used for correction and improvement.

Wavefront reconstruction can estimate the atmospheric distortion on the propagating wave of light. Adaptive optics (AO) uses information about the wavefront to correct the distortion and results in improved system performance. This technique can also detect optical manufacturing defects, thermal fluctuations of components, gravity sag of the components, and optical alignment of the components that also affect the wave. These defects are corrected with active optics. While AO corrections are usually small relative to the wavelength of light and rapidly changing, active optics corrections are large and slowly change.

Both AO and active optics systems commonly use a deformable mirror (DM) as the device to apply the correction. A DM has a reflective surface with actuators along the back structure that apply forces causing the mirror surface to adapt to a desired shape. A control law determines the individual voltages for each actuator that cause the mirror to take the desired shape. The common practice is that the mirror forms the conjugate of the wavefront, then the light reflecting off the mirror is a planar wavefront.

In this research we address the problems of phase reconstruction and control of the DM to compensate for phase distortions. The results are a number of algorithms which are efficient and scalable to systems with a larger number of sensor measurements and also robust in the presence of phase singularities due to high levels of atmospheric turbulence.

To address the research objectives, this dissertation has three contributions in AO. The first is wavefront reconstruction for a large number of sensor measurements. An algorithm for wavefront reconstruction is presented that uses orthogonal wavelet filters in a tree structure to estimate the relative phase of the wavefront from gradient measurements. The tree structure is implemented with two-dimensional quadrature mirror filters (QMFs), which yields

a computationally efficient approach. The measurements contain noise from the sensor and it is desirable for the algorithm to mitigate the impact of noise on the resulting wavefront. The noise filtering properties of this algorithm depend on the chosen wavelet filters used in the QMF. A modification for the Haar wavelet filters is presented that removes all dependencies on the boundary. The algorithm is designed to reconstruct a square wavefront of size $2^N \times 2^N$ for integer N. Although there are applications of AO systems with square apertures, an additional modification is proposed to handle realistic apertures with other shapes. The wavefront reconstruction algorithm is tested using simulated wavefront sensor data.

This algorithm has been designed for irrotational vector field gradient measurements, where there is no phase ambiguity and the phase is well defined at every point. This is, in general, the case of distortion generated by low atmospheric turbulence. Under more severe turbulence conditions, the intensity of the optical field might be zero at isolated points, thus causing phase uncertainties. In this case, the measured phase gradient becomes rotational and it is characterized by phase uncertainty, and branch points, which cause problems in all standard least-squares algorithms.

The second contribution is in wavefront reconstruction in the presence of significant atmospheric turbulence causing degradation of the optical beam. In this case, the measurements contain a rotational vector field component. An approach using a non-orthogonal decomposition to modify the rotational vector field to be irrotational is presented. The wavefront reconstruction algorithm operates on the irrotational measurements and estimates the phase that is consistent with the original measurements with rotational components. Our results show the wrapped phase measurements to make a comparison between the simulated phase and reconstructed wavefront phase. The algorithm is tested with simulated measurements of wavefronts. This approach can be applied to any wavefront reconstruction algorithm as the measurements are modified before the algorithm.

A third contribution is in DM surface control. The commands are calculated as the solution of a constrained optimization problem. The optimizer uses a linear influence function model for each actuator and determines a set of voltages that matches a desired surface shape. Although the model is linear, we expect nonlinear performance on laboratory hardware. An experiment performing mirror surface control using optimization is presented

where a sequence of optimal problems are used to verify the linear model of a DM. The hardware configuration uses a sensor which has many more measurement values than actuators. To decrease computation time, a multigrid approach to the optimization problem is used, which results in the same optimal solution and is 2.5 times faster. This research verified the control approach using mathematical optimization on laboratory hardware.

The research in this dissertation can be applied to a variety of AO systems. Wavefront reconstruction and mirror surface optimization are relevant to many military applications of precision optical systems. This research supports military capabilities in information dominance and directed energy weapon development.

THIS PAGE INTENTIONALLY LEFT BLANK

Acknowledgments

I am profoundly grateful for my advisor, Roberto Cristi. Without his guidance and immense patience, this work would not have been possible. I thank Professors Xiaoping Yun, Phillip Pace, Brij Agrawal, Jae-Jun Kim, and Brett Borden for their invaluable time and teaching. Each has qualities that I intend to promote throughout my career.

I also thank my colleagues in the Spacecraft Research and Design Center: Dr. John Bagnasco, Captain Alan Scott (Ret.), Dr. Lew DeSandre, Dr. Adam Yingling, Dr. Ty Martinez, Mr. Albert Jordan, Dr. Bautista Fernandez, Dr. Edwin Ahn, Dr. Gero Nootz, Commander J.J. Watson, Major Matt Allen, Lieutenant Lee Johnson, Ms. Eva Carrillo, Mr. John Hamilton, Mr. Mike Krol, Ms. Renee Oania, and Ms. Jackie Rosenkranz. Working with such a great group of people has been rewarding both on a professional and personal level.

I am fortunate to have Norris C. Mitchell as a mentor and friend. I thank him for all of his advice and encouragement throughout my years of government service.

My opportunity to be a student at Naval Postgraduate School was due to the generous support of the Office of the Secretary of Defense (SMART Scholarship Program) and my home organization, Space and Naval Warfare Systems Center Atlantic. Without these two organizations working together, I would not have been able to attend this school and follow my dreams.

I have been inspired by my family and friends, who have all offered me continued encouragement and their love during my education. I could not have achieved this milestone without them. I thank my parents for the strong sense of work ethic that they gave me.

This work is dedicated to Kayli and Amara. I am eager to see where your dreams take you.

THIS PAGE INTENTIONALLY LEFT BLANK

CHAPTER 1:
Introduction

Precision optical systems, such as high-resolution imaging or laser weapons, are highly affected by perturbation from the propagating medium. The key problem addressed in this dissertation is how to estimate this perturbation of the optical waveform and compensate for its effects. Wavefront reconstruction is used in many modern applications of engineering and science to provide insight into the performance of optical systems that can be used for correction and improvement. Wavefront reconstruction is particularly important in astronomy to estimate the distortion of the collected light caused by the atmosphere [1], [2]. In modern telescopes, this information is used to correct for the distortion using adaptive optics (AO) to reduce the distortions detected by wavefront reconstruction and gather high-quality observations of celestial objects. The AO corrections occur on a small time scale with updates faster than ten times a second. Wavefront reconstruction can also sense distortions from a variety of other sources, such as manufacturing defects, thermal fluctuations, gravity sag, and optical alignment. These corrections are performed by active optics [3] and occur on a longer time scale with corrections only once a second or slower.

An AO or active optics system contains three common components: a wavefront sensor, a computer, and a deformable mirror (DM) [1]. The computer uses measurements from the sensor to perform the wavefront reconstruction and then commands the DM actuators. The actuators cause forces along the back of the mirror structure and the mirror surface deflects to form the conjugate shape of the wavefront. The incoming light reflects off the DM surface and the wavefront becomes planar. The planar wavefronts improve the optical performance of the system, and the collected images have improved angular resolution to distinguish fine details on the distant object. The AO computer must estimate the wavefront in a short time period and calculate the actuator commands for the DM. If either procedure takes too long, the distortion may change, and the intended correction may reduce optical performance.

Another application of wavefront reconstruction is in laser weapon systems, which require AO for long horizontal distances to targets. As the laser beam propagates from the telescope

through the atmosphere to the target, the beam quality degrades from atmospheric effects and the weapon loses effectiveness. AO is used to apply a correction before the beam is emitted to reach the target and retain the high-energy laser performance.

Large telescopes are placed high on mountains that have excellent atmospheric "seeing" conditions. The *Fried parameter* (or *Fried coherence length*) quantifies the quality of the "seeing" and the best locations on Earth may have a Fried parameter value of 20 to 40 centimeters. If the telescope aperture diameter is larger then the Fried parameter, the telescope is limited by the atmospheric turbulence; incorporating AO into the telescope can overcome the limitation imposed by the atmosphere. Science missions require larger primary mirrors to increase the angular resolution and gather more light from the science objects of distant stars and their orbiting planets. Large monolithic mirrors are limited in maximum surface dimensions and may take years to correctly manufacture, polish, and test [4], [5]. This limitation has led to segmented mirror telescopes, where many smaller mirrors are combined to fill the large aperture [3]. These technologies are used in existing telescopes such as the Gran Telescopio CANARIAS (GTC) and the W. M. Keck observatory, as shown in Figure 1.1. A thorough historical perspective on segmented mirrors can be found in [6]. These telescopes are able to observe faint science objects using the highest fidelity in large-scale optics, active control, and sensor technology.

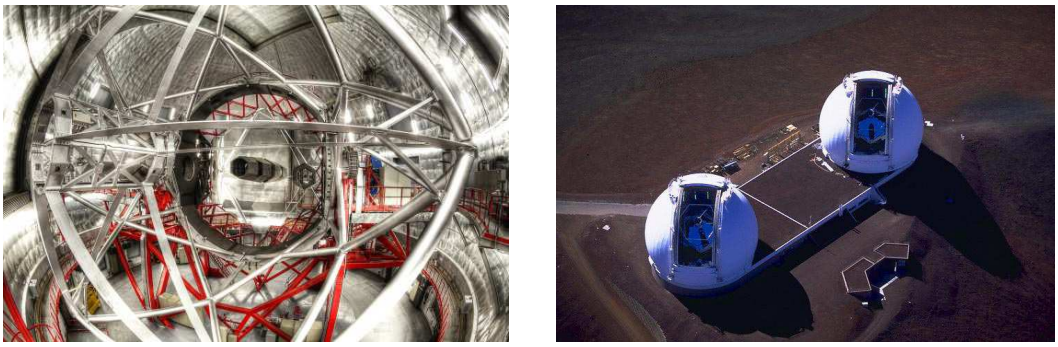


Figure 1.1. Left: The GTC primary mirror located in La Palma, Spain, is displayed (courtesy of GTC Digital, from [7]). Right: This is overhead view of the W. M. Keck telescopes at the observatory on Mauna Kea, Hawaii, is displayed (courtesy of NASA JPL and W. M. Keck Observatory, from [8]).

Future telescope designs incorporate technology advancements to expand the state-of-the-art science capabilities in celestial formations and fundamental physics. The Thirty Meter

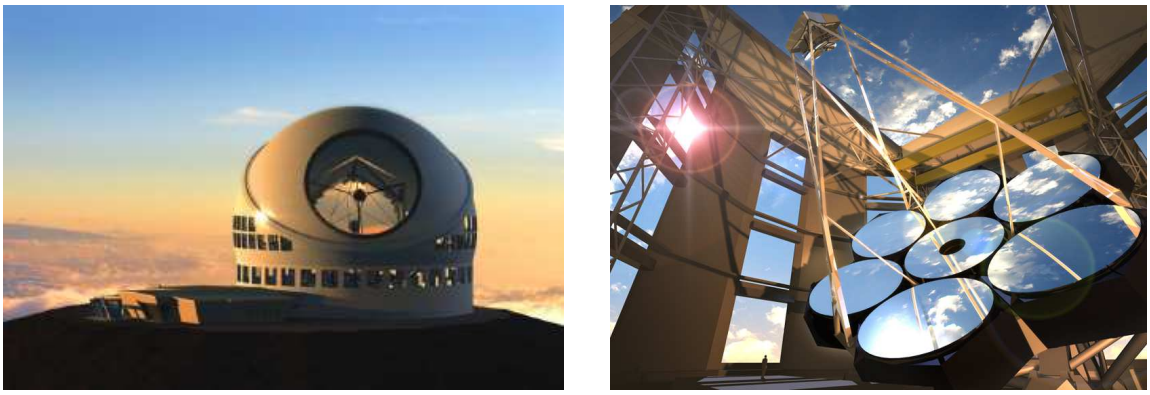


Figure 1.2. Left: Artist rendition of the TMT is displayed (courtesy of TMT Observatory Corporation, from [9]). Right: Artist rendition of the GMTO is displayed (courtesy of GMT Organization, from [10]).

Telescope (TMT) and Giant Magellan Telescope (GMT) will each be larger than 25 meters in diameter [9], [10]. Current artistic renderings of these observatories are shown in Figure 1.2. The designs require a combination of optics, structures, and controls to make a functional telescope.

Space-based telescope designs now incorporate segmented mirror technology. For example, the James Webb Space Telescope (JWST) consists of 18 mirror segments to form a 6.5 meter primary mirror [11]. A conceptual rendering of the JWST is shown in Figure 1.3. Current rocket launch vehicles and fairings can support a stowed telescope configuration that deploys the primary mirror on orbit. After deployment, the mirror segments are positioned precisely to form the primary mirror.

In addition to segmented mirror technology, space-based telescopes incorporate active optics for on-orbit conditions that cause mirror deformations. Active optics are desired by program managers for risk mitigation of manufacturing processes that result in an unnoticed defect until operation. For these reasons, active optics are highly desirable for space-based telescopes, as they balance performance risk and system cost.

To mature the segmented mirror with active optics technology for spacecraft, the Segmented Mirror Telescope (SMT) was built for the Segmented Mirror Demonstrator (SMD) program. Its lightweight, segmented primary mirror was constructed by sophisticated manufacturing processes to meet the size, weight and power (SWAP) requirements imposed on

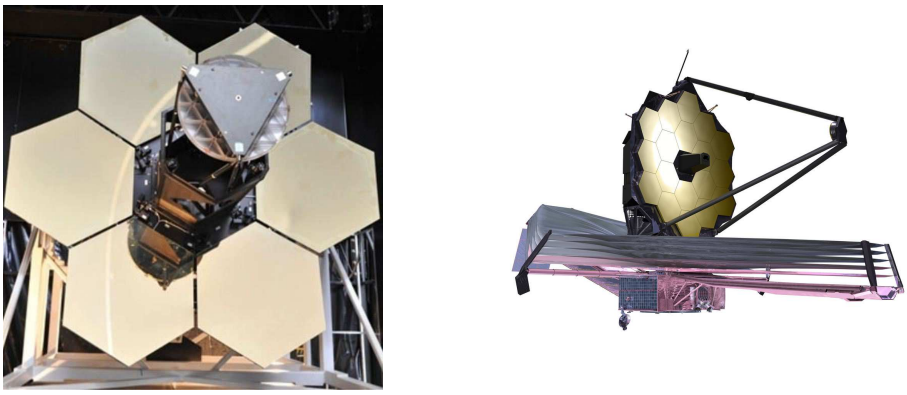


Figure 1.3. Left: The NPS laboratory of the Segmented Mirror Telescope testbed is shown. Right: Artist rendition of the James Webb Space Telescope is displayed (courtesy of NASA, from [13]).

satellites [12]. The mirrors were not of high enough quality for an operational mission. The manufacturing processes have matured since the SMD program finished and industrial methods are capable of producing excellent mirror surfaces. The SMT is now located at the Naval Postgraduate School (NPS) for research and shown in Figure 1.3.

Another important application in AO and segmented mirrors with active optics for the Department of Defense (DOD) is rapidly fielded imaging systems in support of military operations. Segmented mirrors with active optics technology can field an imaging spacecraft much faster than conventional imaging spacecraft by decreasing the payload manufacturing schedule. An individual segment can be produced in a few months time and multiple segments can be produced concurrently. Many mirrors can be manufactured and the mirrors with the best optical surfaces can be determined by wavefront reconstruction. Active optics reduce risk on the primary mirror fabrication by reducing polishing rework since the system can compensate for manufacturing defects. The most compelling justification for this technology is the cost savings from reduced schedule dedicated to the primary mirror.

1.1 Research Objectives

The first objective of this dissertation research is to develop a wavefront reconstruction algorithm that can be applied to large telescopes. A generalized orthogonal wavelet phase reconstruction algorithm is described. This approach is appropriate for AO systems with larger aperture size and increasing the number of sensor measurements. The increase in

number of measurements causes issues for the least-squares solvers which rely on matrix-vector multiplications. The approach presented in this dissertation is comparable with least-squares algorithms but has lower computational cost. The generalized orthogonal wavelet approach allows for filters that have improved noise-rejection properties compared to other existing algorithms. The phase reconstruction algorithm is tested using simulated wavefront sensor data.

The second objective of this dissertation research is to address wavefront reconstruction for significant wavefront distortion. Although many applications of phase reconstruction are used for weak disturbances on the wavefront, there are applications where the wavefront may experience significant distortion. Strong distortion of the wavefront can occur in long, horizontal paths through the atmosphere, such as in the maritime environment. Least-squares algorithms degrade with measurements from the significant distortion such that the wavefront estimate does not represent a meaningful quantity. The work presented in this dissertation allows for a correct wavefront to be reconstructed. The modification is tested with simulated measurements of wavefronts.

The final objective of the dissertation research is the control of a DM using wavefront sensor data and a mathematical optimization algorithm. The optimizer result selects the actuator commands to the DM that matches a desired mirror surface shape. The approach relies on a linear model of actuator influence and the validity of this control approach is investigated. Prior to this work, the technique has only been validated using computer simulation. This dissertation research verifies the control approach on laboratory hardware.

1.2 Dissertation Organization

In Chapter 2, the fundamentals of two-dimensional signal processing are presented. The signal representation notation and transforms that are used throughout the dissertation are explained. The wavelet phase reconstruction algorithm is presented in Chapter 3. The algorithm is an approach that recursively operates on gradient measurements to form an estimated wavefront. In Chapter 4, the branch point modification is presented for wavefront reconstruction algorithms. Mirror surface control using optimization is discussed in Chapter 5 and an experiment is conducted to verify the linear influence model of a DM. Concluding remarks and future research opportunities are given in Chapter 6.

THIS PAGE INTENTIONALLY LEFT BLANK

CHAPTER 2:
Two-Dimensional Signal Processing

A *signal* conveys information about a process of interest and usually is represented as a function of an independent variable, such as $f(t)$. Although in most applications the independent variable t is a scalar indicating time, in a more general setting, it is a vector $\mathbf{t} = [t_1, t_2, \ldots, t_n]$ indicating time, space, or a combination. This leads to the definition of *multidimensional signals* $f(t_1, t_2, \ldots, t_n)$. In this chapter, we present the fundamentals of multirate and multiresolution signal processing as applied to multidimensional signals. These concepts will be the basis of the phase reconstruction algorithms presented in subsequent chapters.

Although most physical signals evolve in a continuous domain of time and/or space, they are usually processed numerically in the discrete (or sampled) domain using digital signal processing (DSP). We denote this signal by $f(t_1, t_2, \ldots, t_M)$ in the continuous domain and $f[n_1, n_2, \ldots, n_M] = f(n_1\Delta t_1, n_2\Delta t_2, \ldots, n_M\Delta t_M)$ in the discrete domain, where each n_i is an integer for $i = 1, 2, \ldots, M$. The choice of variable name and brackets distinguish the difference between continuous and discrete signals. The terms $\Delta t_1, \Delta t_2, \ldots, \Delta t_M$ denote the sampling intervals in each respective dimension. Although general sampling of multidimensional signals is not a straightforward extension of one-dimensional (1D) sampling [14], in the case of simple sampling on a rectangular grid we can still use the well know sampling theorem. In this way, the sampling frequencies $F_i = 1/\Delta t_i$ are chosen to be larger than twice the maximum frequency associated to the respective variables.

Many important signals in imaging applications are represented as two-dimensional (2D) signals. In this dissertation, we are dealing with 2D signals that represent the phase of an optical field across a telescope aperture represented by two spatial coordinates, x and y. For simplicity, sometimes multidimensional signals are abbreviated with vector notation as $f[\mathbf{n}]$ for $\mathbf{n} = [n_1, n_2]$. In some places in this dissertation, a signal is referred to only by its symbolic letter name but no brackets for brevity. In these circumstances, a discrete signal is implied.

By the well-known sampling theorem, there is no loss of information if the signal is sam-

pled at a sufficiently fast rate F_s. Recovering the original band-limited continuous signal from the discrete signal can be done using Whittaker-Shannon interpolation

$$\begin{aligned} f(x,y) &= \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} f[n_1,n_2] \operatorname{sinc}\left(\frac{x-n_1\Delta x}{\Delta x}\right) \operatorname{sinc}\left(\frac{y-n_2\Delta y}{\Delta y}\right) \\ &= \left(\sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} f[n_1,n_2]\delta(x-n_1\Delta x, y-n_2\Delta y)\right) ** \operatorname{sinc}\left(\frac{x}{\Delta x}\right) \operatorname{sinc}\left(\frac{y}{\Delta y}\right) \end{aligned} \tag{2.1}$$

where the $**$ operator represents continuous time convolution, $\delta(t_1,t_2) \equiv \delta(t_1)\delta(t_2)$ is the 2D Dirac delta function, and $\operatorname{sinc}(t) \equiv \sin(\pi t)/(\pi t)$ that is defined for all $t \in \mathbb{R}$.

In the rest of the chapter we introduce the concept of frequency representation of 2D signals, followed by multiresolution decomposition. These signal processing concepts will play a major role in the phase reconstruction algorithms presented in this dissertation.

2.1 Two-Dimensional Fourier Transform

We define the continuous 2D *Fourier Transform* (FT) as

$$\begin{aligned} F(\kappa_x,\kappa_y) &= \mathrm{FT}\{f(x,y)\} = \\ &\equiv \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)e^{-j2\pi(\kappa_x x+\kappa_y y)}dxdy \end{aligned} \tag{2.2}$$

with κ_x, κ_y having the appropriate units to ensure that the quantities $\kappa_x x$ and $\kappa_y y$ are dimensionless.

Extension of the Fourier transform to the 2D sampled signal yields the 2D *Discrete Shift Fourier Transform* (DSFT) defined as

$$F(\boldsymbol{\omega}) \equiv \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} f[n_1,n_2]e^{-j(\omega_1 n_1+\omega_2 n_2)}. \tag{2.3}$$

The units of ω_1, ω_2 are dimensionless expressed in *radians* or *radians per sample*, which has a 2π periodicity in both dimensions of the DSFT; the two frequency variables are usually constrained to the intervals

$$-\pi \leq \omega_i < \pi \tag{2.4}$$

for $i = 1, 2$. All other values of ω are associated to aliased frequencies from the sampling process. If $f[n_1, n_2] = f(n_1\Delta x, n_2\Delta y)$, then ω_1 and ω_2 are related to κ_x, κ_y as

$$\omega_1 = 2\pi\kappa_x\Delta x, \quad \omega_2 = 2\pi\kappa_y\Delta y. \tag{2.5}$$

2.2 Two-Dimensional z-Transform and Filtering

In this dissertation, we numerically process discrete signals. In the 1D discrete time case, the z-transform is the tool which most conveniently describes linear shift-invariant (LSI) operations. It is defined as

$$F(z) = \mathcal{Z}\{f[n]\} \equiv \sum_{n=-\infty}^{\infty} f[n]z^{-n}, \quad z \in \text{ROC} \tag{2.6}$$

where z is a complex variable and is related to the frequency domain. In general, for signals of infinite length, the Region of Convergence (ROC) has to be defined. However when $f[n]$ has a finite interval of definition, then convergence is guaranteed for $0 < |z| < \infty$. When $z = e^{j\omega}$ and $z \in$ ROC, then (2.6) becomes the same as the DSFT. The inverse z-transform is given by

$$f[n] = \mathcal{Z}^{-1}\{F(z)\} \equiv \frac{1}{2\pi j}\oint_C F(z)z^{n-1}dz. \tag{2.7}$$

In the above definition, C is a closed, counter-clockwise directed contour that encircles the origin and is completely inside of the ROC.

The z-transform can be extended to the 2D case and its definition is given by

$$F(z_1, z_2) = \mathcal{Z}\{f[n_1, n_2]\} \equiv \sum_{n_1=-\infty}^{\infty}\sum_{n_2=-\infty}^{\infty} f[n_1, n_2]z_1^{-n_1}z_2^{-n_2}, \quad z_1, z_2 \in \text{ROC}. \tag{2.8}$$

One of the fundamental properties of the z-transform is that a shift in the domain of definition (time or space, accordingly), corresponds to an algebraic operation in the z-domain, as

$$\mathcal{Z}\{f[n_1 + L_1, n_2 + L_2]\} = z_1^{L_1}z_2^{L_2}F(z_1, z_2) \tag{2.9}$$

where L_1, L_2 are integers. By this respect, the variables z_1, z_2 can be viewed as shift

operators. This leads to the use of z_1, z_2 as *unit shift operators*, for which we indicate

$$z_1{}^{L_1} z_2{}^{L_2} f[n_1, n_2] = f[n_1 + L_1, n_2 + L_2]. \tag{2.10}$$

By this interpretation, we will make minimal use of the z-transform and rather use the shift operators.

Filtering is performed by using the convolution operation. For the 2D case, the convolution of f and g is stated as

$$f(x,y) * *g(x,y) \equiv \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(c_1, c_2) g(x - c_1, y - c_2) dc_1 dc_2 \tag{2.11}$$

for continuous signals and for the discrete signals is

$$f[n_1, n_2] * *g[n_1, n_2] \equiv \sum_{m_1=-\infty}^{\infty} \sum_{m_2=-\infty}^{\infty} f[m_1, m_2] g[n_1 - m_1, n_2 - m_2]. \tag{2.12}$$

Convolution expresses many physical processes and applications arise where one of the functions represents measured information and the other function represents a filtering kernel that modifies the measurements. In this dissertation, convolution is used to apply filters to the measured data of the AO system.

In order to efficiently represent convolution, throughout this dissertation we make frequent use of operator notation. Using the definition of the shift operator z_1, z_2, and the fact that $x[n_1 - m_1, n_2 - m_2] = z_1^{-m_1} z_2^{-m_2} x[n_1, n_2]$, we write the convolution of two signals as

$$y[n_1, n_2] = \sum_{m_1=-\infty}^{\infty} \sum_{m_2=-\infty}^{\infty} h[m_1, m_2] z_1^{-m_1} z_2^{-m_2} x[n_1, n_2]. \tag{2.13}$$

Then by defining the transfer function of the LSI system

$$H(z_1, z_2) \equiv \sum_{m_1=-\infty}^{\infty} \sum_{m_2=-\infty}^{\infty} h[m_1, m_2] z_1^{-m_1} z_2^{-m_2}, \tag{2.14}$$

we can write the convolution in operator form as

$$y[n_1, n_2] = H(z_1, z_2)x[n_1, n_2]. \tag{2.15}$$

This representation is completely equivalent to the convolution in (2.12) and to its z-transform correspondent in (2.8). We view (2.15) as a short-hand notation of the convolution operator, which, when necessary, can be extended to the corresponding z-transform with previous definitions of the correct ROC. Equation (2.15) is equivalent to the following sequence of operations:

$$\begin{aligned} X(z_1, z_2) &= \mathcal{Z}\{x[n_1, n_2]\} \\ Y(z_1, z_2) &= H(z_1, z_2)X(z_1, z_2) \\ y[n_1, n_2] &= \mathcal{Z}^{-1}\{Y(z_1, z_2)\}. \end{aligned} \tag{2.16}$$

For ease of notation, we skip the indices n_1, n_2 and the variables z_1, z_2 when they are apparent from the context. In this way, the following expressions are all equivalent:

$$\begin{aligned} y[n_1, n_2] &= H(z_1, z_2)x[n_1, n_2], \\ y &= H(z_1, z_2)x, \\ y &= Hx. \end{aligned} \tag{2.17}$$

Most of the operations in this dissertation are based on a combination of filtering and "resampling" operations, leading to a multirate signals, described next.

2.3 Fundamentals of Multirate DSP

In this section, we review the major definitions and results of multirate DSP, as applied to general multidimensional signals. Multirate DSP will be the basis of the proposed phase estimation technique, which processes observed phase differences into a tree-like structure at different scales and then recombines them into the final phase estimation.

In multirate DSP, there are two elementary resampling operations: *downsampling* and *upsampling*. By downsampling, we eliminate samples, which is equivalent to resampling the signal at a lower rate. By upsampling, we interpolate between samples, which is equivalent

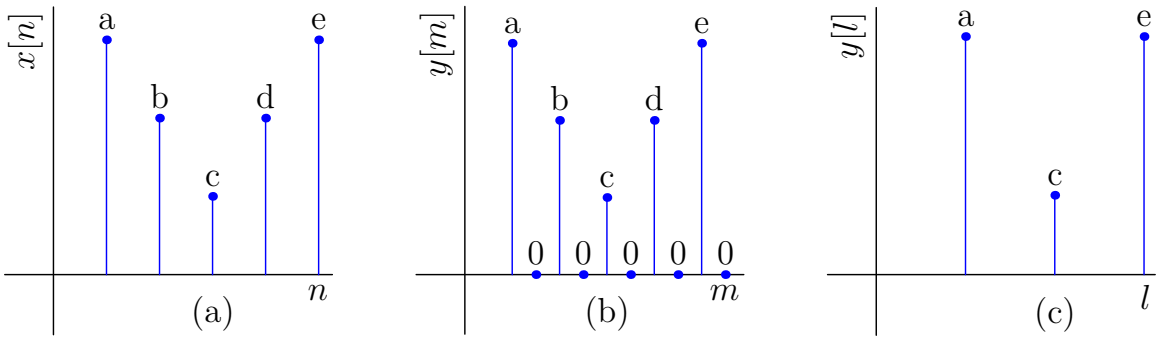


Figure 2.1. A signal is shown in (a) that is being upsampled in (b) and downsampled in (c). The index variable identifies the different sampling grids. Both examples use $L = 2$.

to resampling the signal at a higher rate. Whether we do this in the time or space domain is immaterial. Although we define these operations in one dimension, they are easily extendable to two or more dimensions.

Upsampling involves adding new samples between existing samples and all of the new samples are zero valued. The upsampling of the signal $x[n]$ as $y[n]$ is expressed as

$$y[n] = \begin{cases} x[n/L] & \text{if } n/L \in \mathbb{Z}, \\ 0 & \text{otherwise} \end{cases} \tag{2.18}$$

where L is the integer *upsampling factor* (two or larger). Adding zeros between data points is a fundamental signal processing tool to be used in conjunction with filtering. The length of $y[n]$ is now L times the length of $x[n]$. An example of upsampling a signal in Figure 2.1 (a) is shown in Figure 2.1 (b).

As previously stated, downsampling is the reduction of samples and is defined as

$$y[n] = x[Ln] \tag{2.19}$$

where L is the integer downsampling factor (two or larger). Whereas upsampling retains all original samples, downsampling loses information. An example of downsampling is shown in Figure 2.1 (c).

In this dissertation, the resampling is always performed using a factor of two on multidimensional signals, in which the operators are treated independently and operate only in

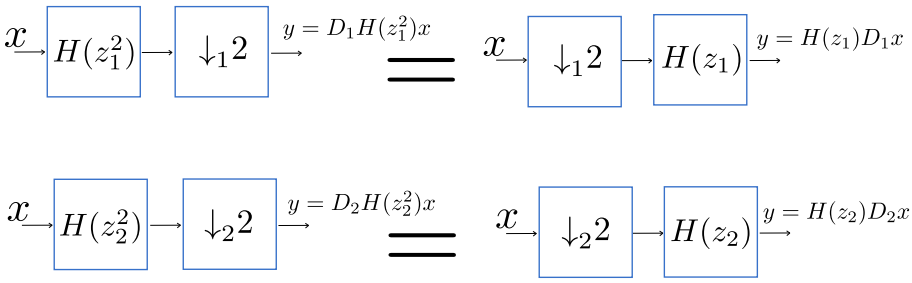


Figure 2.2. Noble identities in block diagram showing equivalence of (2.22) for both dimensions. The serial combination of two operations that both occur in the same dimension cannot change order without changing the filtering operation. The subscript on the downsampling arrow indicates the dimension and is then followed by the downsampling factor.

their dimension. To simplify the notation, the upsampling operator is defined as

$$\begin{aligned} y_1 = U_1 x \quad &\Leftrightarrow \quad \begin{cases} y_1[2n_1, n_2] = x[n_1, n_2] \\ y_1[2n_1+1, n_2] = 0 \end{cases} \\ y_2 = U_2 x \quad &\Leftrightarrow \quad \begin{cases} y_2[n_1, 2n_2] = x[n_1, n_2] \\ y_2[n_1, 2n_2+1] = 0 \end{cases} \end{aligned} \tag{2.20}$$

for each dimension. The downsampling operator is similarly defined as

$$\begin{aligned} y_1 = D_1 x \quad &\Leftrightarrow \quad y_1[n_1, n_2] = x[2n_1, n_2], \\ y_2 = D_2 x \quad &\Leftrightarrow \quad y_2[n_1, n_2] = x[n_1, 2n_2], \\ y_3 = D_1 D_2 x = D_2 D_1 x \quad &\Leftrightarrow \quad y_3[n_1, n_2] = x[2n_1, 2n_2]. \end{aligned} \tag{2.21}$$

Both operations of upsampling and downsampling are always combined with filtering operations, as shown in Figure 2.2. A full treatment is given in numerous well-known references (see [15], [16]), and not reproduced here for brevity. Just to give an idea, the reason we filter before downsampling in Figure 2.2 is that lowering the data rate causes aliasing, thus requiring an anti-aliasing filter. As far as upsampling, the interpolation of zeros is clearly unsatisfactory. However, the addition of a filter is equivalent to interpolating the signal by the impulse response of the filter. Again, it is always an advantage to represent all of these operations in terms of operators, which we can easily and efficiently manipulate. Operator notation is particularly important in the problem we have at hand, which, as we will see, is

based on resampling in two dimensions and can lead to a fair amount of complexity.

The Noble identities [15], [16] are used in multirate signal processing to describe the equivalence of swapping the ordering of downsampling-filtering to filtering-downsampling. The Noble identities correctly establish the relationship in instances such as

$$\begin{aligned} y_1 &= D_1 H(z_1^2)x = H(z_1)D_1 x, \\ y_2 &= D_2 H(z_2^2)x = H(z_2)D_2 x. \end{aligned} \tag{2.22}$$

The equivalence of (2.22) in block diagram form is shown in Figure 2.2. For example, let $H(z_1^2) = z_1^{-2L}$ then the left-hand expression of (2.22) is

$$y_1[m] = D_1 z_1^{-2L} x[n] = D_1 x[n-2L] = x[2m-2L] = x[2(m-L)] \tag{2.23}$$

and the right-hand expression is

$$y_1[m] = z_1^{-L} D_1 x[n] = z_1^{-L} x[2m] = x[2(m-L)], \tag{2.24}$$

which are the same expression and equivalent. Similar expressions could be stated for upsampling. Since these operators are linear, they can be applied to any linear combination, thus to any arbitrary filter $H(z)$.

2.4 Multiresolution Representation of a Signal in a Tree Structure

A wavefront phase reconstruction algorithm that is based on multiresolution decomposition is described in this dissertation. This concept is related to the wavelet transform, used in a variety of applications. The discrete wavelets are a particular class of filter kernels and they form a basis (rather than a Fourier transform basis or similar) for signal representation. We use an implementation of discrete, multiresolution wavelets that are orthogonal for the work presented herein.

This decomposition is used in a number of applications, such as signal compression. For example, in computer image storage, the JPEG2000 standard uses the discrete wavelet transform [17]. This compression is possible because most images are vastly oversampled

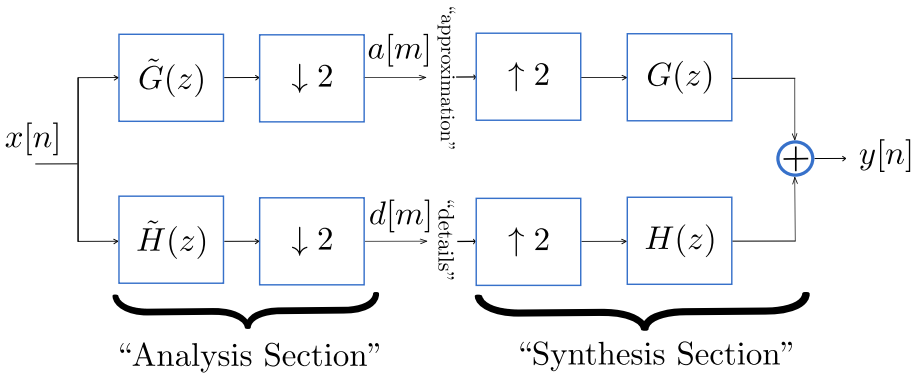


Figure 2.3. Tree structure of a quadrature mirror filter for single dimensional signal $x[n]$. Tree structures with the perfect reconstruction property result in $y[n]$ being equivalent to a shifted $x[n]$. The channel with $\tilde{G}(z)$ is a low-pass filter and the channel with $\tilde{H}(z)$ is a high-pass filter.

relative to their content and can be compactly expressed at coarse sampling. A compression algorithm can exploit the multiresolution signal structure for improved compression ratios at very little computational cost.

In this dissertation, the AO wavefront sensor has a relationship to the low and high pass filter nature of the wavelets. This relationship is exploited to apply the general orthogonal wavelets to reconstruct a signal of interest in AO by using a tree structure.

The multiresolution decomposition of a signal is recursively obtained as a tree structure, by successively dividing the signal into two components: low frequency (called the "approximation") and high frequency (called the "details"). Since each one of these components retain half the bandwidth of the signal, the sampling rate can be reduced by a factor of two, so that the overall data rate stays the same.

Multiresolution decomposition leads to the structure shown in Figure 2.3 called the *quadrature mirror filter* (QMF). In particular the "analysis section" performs the decomposition into the two channels (approximation and details), while the "synthesis section" recombines them into the original signal. The approximation and details are given by

$$\begin{aligned} a[m] &= D\tilde{G}(z)x[n], \\ d[m] &= D\tilde{H}(z)x[n]. \end{aligned} \tag{2.25}$$

It is important to notice two things:

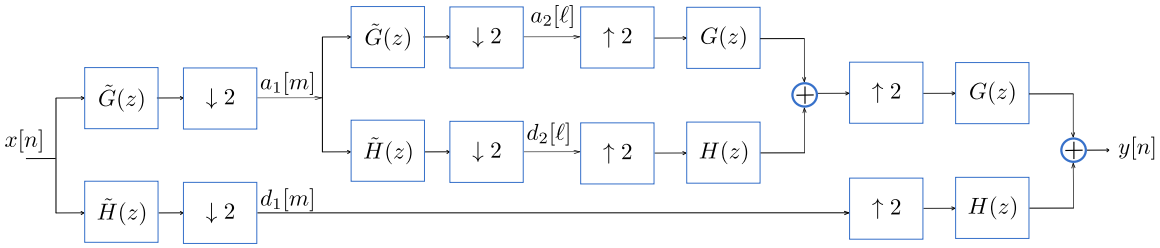


Figure 2.4. Tree structure of a QMF for single dimensional signal $x[n]$. The approximation signal $a_1[m]$ is processed through a second level tree structure which results in $a_2[\ell]$ and $d_2[\ell]$.

1. There is no change in overall data rate between the given signal $x[n]$ and the two components $a[m]$, $b[m]$;
2. It can be easily shown by standard arguments that if the two low-pass filters, $\tilde{G}(z)$, $G(z)$ and the two high-pass filters $\tilde{H}(z)$, $H(z)$ are ideal with bandwidth of $F_s/4$, then there is no aliasing and perfect reconstruction occurs. We can express the output in operator form as

$$y = \left(G(z)UD\tilde{G}(z) + H(z)UD\tilde{H}(z)\right)x \tag{2.26}$$

with U and D indicating upsampling and downsampling by two, respectively.

When the filters are not ideal, such as Finite Impulse Response (FIR), then perfect reconstruction can still occur, provided that the filters are specially chosen to satisfy some orthogonality conditions that preserve the perfect reconstruction property [16]. The filters that are used in a QMF system are specifically designed in a manner to cancel out aliasing effects due to non-ideal frequency response of the filters such that

$$G(z)UD\tilde{G}(z) + H(z)UD\tilde{H}(z) = z^{-L}. \tag{2.27}$$

The output $y[n]$ is a perfect reconstruction of $x[n]$ shifted by L, which can be thought of as a processing delay of the filters. This arrangement is ideal for lossless compression.

As stated at the beginning of the section, the QMF can be replicated in a tree structure as shown in Figure 2.4 where the low-frequency component (approximation) is further decomposed by a similar QMF; the left half of Figure 2.4 depicts the discrete wavelet transform (DWT) for two levels and the right half is the inverse discrete wavelet transform (IDWT). After performing the DWT, the original signal $x[n]$ is decomposed into $a_2[\ell]$,

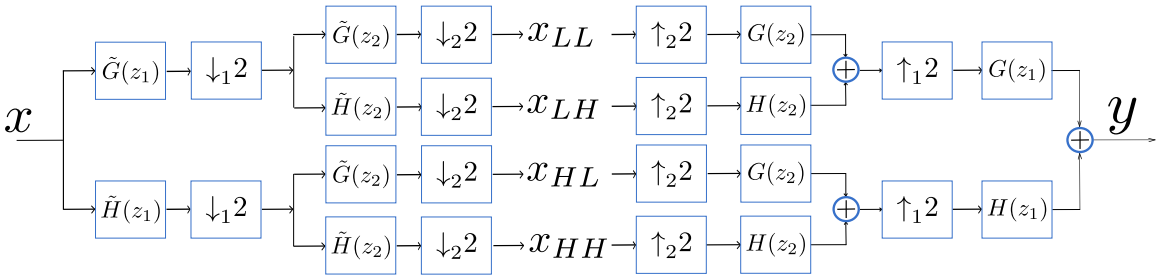


Figure 2.5. The 2D QMF breaks the original signal $x[n_1,n_2]$ into four components that can be recombined to form a copy of the signal as $y[n_1,n_2]$.

$d_2[\ell]$, and $d_1[m]$; this representation contains the same number of samples as $x[n]$ and each component contains content at different frequency ranges.

In this dissertation, we will make use of the extension of QMF to the 2D case. The 2D QMF is performed by decomposing each row of an $N \times N$ 2D signal $x[n_1,n_2]$, into the two components: a low pass and high pass, as $x_L[m_1,n_2]$, $x_H[m_1,n_2]$, with $m_1 = 0,1,\ldots,N/2-1$. Then decomposing again each of the two signals x_L, x_H into $x_{LL}[m_1,m_2]$, $x_{LH}[m_1,m_2]$ and $x_{HL}[m_1,m_2]$, $x_{HH}[m_1,m_2]$ again with the indices $m_i = 0,1,\ldots,N/2-1$. The result is the QMF structure in 2D as shown in Figure 2.5.

The four output channels can be described similar to the QMF terminology where x_{LL} is the approximation and x_{LH}, x_{HL} and x_{HH} are the details in the horizontal (LH), vertical (HL) and diagonal (HH) directions. Each of the channels is now one fourth of the original in size. This reduction in signal size scales with the dimensionality, and thus makes the tree structure an excellent tool for high-dimensional problems. To perform the 2D DWT, we take x_{LL} and recursively decompose the low-pass component through a 2D QMF. Likewise, the four components can be recombined into the original signal by the synthesis filters.

In Figure 2.6, each row of $x[n_1,n_2]$ is processed independently. The resulting signals are $x_L[m_1,n_2]$ and $x_H[m_1,n_2]$. Following this, the next step is to process each column, as shown in Figure 2.7. The four signals of the 2D QMF x_{LL}, x_{LH}, x_{HL}, and x_{HH} are constructed.

With large data sizes, it is appropriate to perform the 2D QMF recursively on the x_{LL} component. A second level QMF result is shown in Figure 2.8. The original signal is then represented with seven signals, each of which contains signal content at different frequency

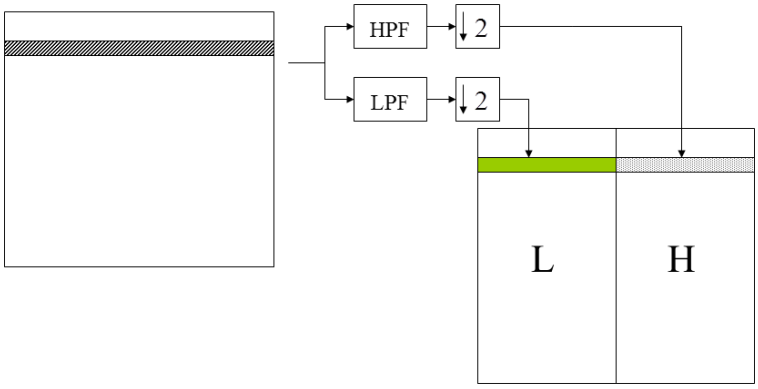


Figure 2.6. Each row of input data is passed through a low-pass and high-pass filter, and then downsampled. This process results in a row of data that is the same dimension as the original row. Figure is courtesy of Roberto Cristi, from [18].

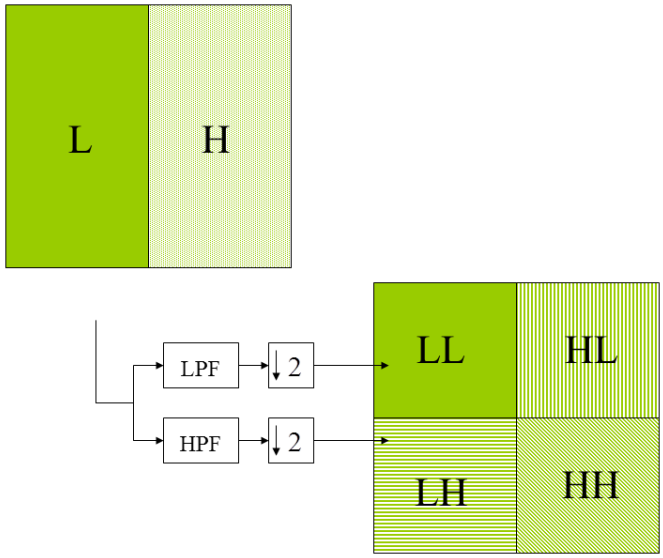


Figure 2.7. After each row of Figure 2.6 is processed, then each column is processed in the same manner. This process results in the four signal blocks. Figure is courtesy of Roberto Cristi, from [18].

ranges.

As an example of the usage of the 2D QMF, the decomposition of an image is shown in Figure 2.9. The approximation is a lower-resolution version of the original sized image. The details can be seen to be contain smaller features and a ghost-outline of objects can be seen.

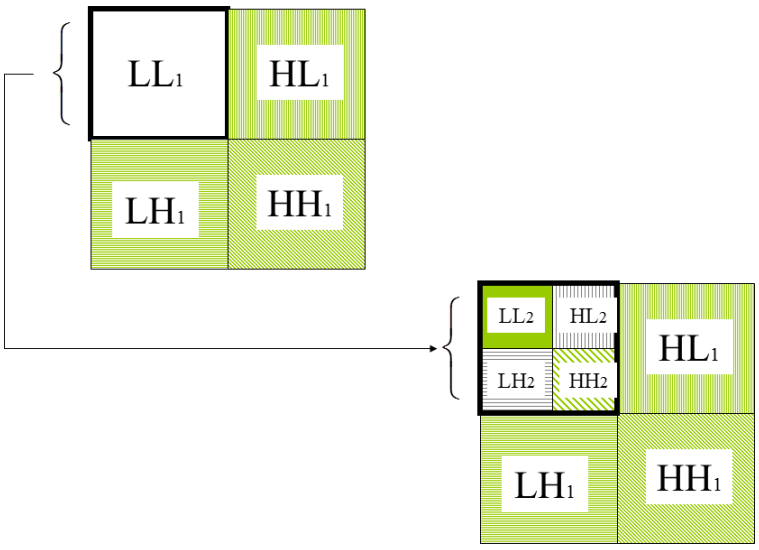


Figure 2.8. The 2D QMF can be processed recursively. Figure is courtesy of Roberto Cristi, from [18].

In this chapter, we discussed the signal processing fundamentals that are necessary to understand generalized orthogonal wavelet phase reconstruction. These concepts included the Fourier transform, the unit shift operator z, convolution, resampling, and multirate signal representation using quadrature mirror filters. With this knowledge, we can now explain the importance of phase and our algorithm to estimate it using wavelet filters.

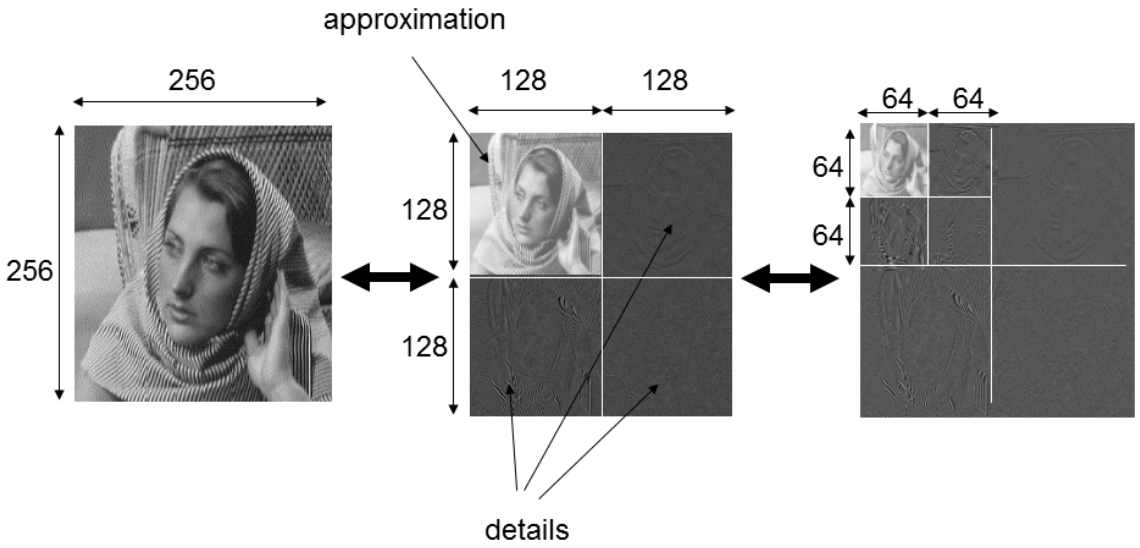


Figure 2.9. A two-level 2D QMF decomposition of a square image is shown. Figure is courtesy of Roberto Cristi, from [18].

CHAPTER 3:
Wavelet Wavefront Reconstruction

Telescopes image distant objects by collecting collimated light and focusing it onto a sensor. The sensor is usually a photographic camera, but can be replaced with a variety of optical instruments, e.g., a spectrometer. The telescope pupil and sensor are located in two different planes. The incoming light enters the *entrance pupil plane* and diffracts from the aperture, and exits at the *exit pupil plane* as shown in Figure 3.1. The *image plane* is where a sensor can be placed that would take measurements about the distant science object under study. In telescope imaging, these are the important planes that define the system capabilities and are shown in Figure 3.2. It is a standard convention to define a Cartesian coordinate frame with the z axis defined to be the optical axis perpendicular to the two planes. Consequently, any point in either plane is characterized by the x and y coordinates.

Telescopes collect collimated light, represented as waves propagating parallel to the optical axis, and are expressed using complex scalar notation as

$$u(x,y,z,t) = A(x,y,z)e^{j(kz-\omega t+\phi(x,y,z))} \tag{3.1}$$

where k is the wavenumber, ω is the angular frequency, $A(x,y,z)$ is the amplitude, and $\phi(x,y,z)$ is the phase shift of the wave. The wavenumber $k = 2\pi/\lambda$ is defined by the wavelength of the light, λ, and the angular frequency $\omega = 2\pi\nu$ is defined by the frequency of the light, ν. For a wave propagating in a three-dimensional (3D) space, a *wavefront* is a surface of the wave where the phase is constant. We call it a *plane wave* when the wavefront can be represented as a planar surface. Most light emanates from a point source and is described using spherical waves, but after propagating some large distance, these waves appear planar over a small solid angle. Thus, plane waves are ideal for telescope collection.

When the light wave is a plane wave propagating parallel to the optical axis, the field in (3.1) can be simplified to

$$u(x,y,z,t) = Ae^{j(kz-\omega t+\phi)} \tag{3.2}$$

where we also assume constant amplitude. At the pupil, which by convention is placed at

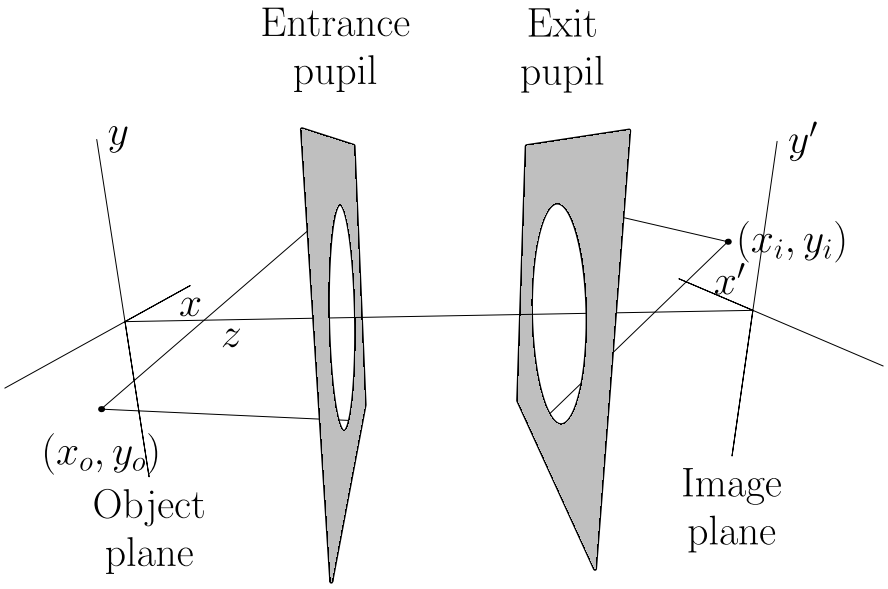


Figure 3.1. A simplified model of optical system analysis in which an image is formed of the object.

$z = 0$, the *pupil function* is defined as

$$P(x,y) \equiv \begin{cases} 1 & \text{inside the aperture} \\ 0 & \text{otherwise} \end{cases} \tag{3.3}$$

which can be used to describe the scalar field (3.2) directly inside the aperture as

$$u(x,y,0,t) = P(x,y)Ae^{j(-\omega t+\phi)} = U(x,y)e^{-j\omega t} \tag{3.4}$$

where the complex-valued phasor

$$U(x,y) = P(x,y)Ae^{j\phi} \tag{3.5}$$

represents the wave inside the pupil. Equation (3.5) is an idealization that the wave incident on the pupil is planar with a constant phase throughout the aperture.

From Fourier Optics [19], it is common knowledge that the optical field in the image plane is proportional to the Fourier transform of the field in the pupil plane indicated in (3.5). If the phase in the pupil plane is constant, the resulting image plane field is then propor-

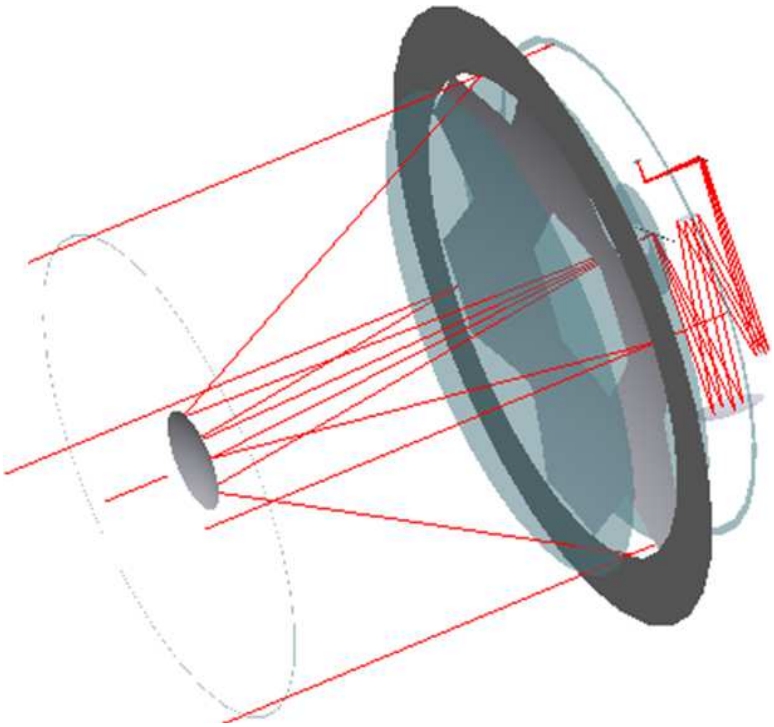


Figure 3.2. Light rays enter an aperture and reflect through the three-mirror anastigmat telescope. The entrance pupil plane is the large primary mirror; the image plane forms at the end of the telescope and is where the camera is placed.

tional to the Fourier transform of the pupil function, which indicates best possible system performance. However, because of atmospheric inhomogeneities in the *index of refraction*, the plane wave is disrupted and the phase can vary as a function of spatial coordinates as $\phi(x,y)$. Additional phase disturbances can occur due to the manufacturing tolerances of optics such as alignment, prescription defects, and vibration. The nonuniform phase across the aperture caused by the system optics is called *aberration* [20]. To consider the combination of these effects, we substitute $\phi(x,y)$ into (3.2) and determine the pupil field to be of the form

$$U(x,y) = P(x,y)Ae^{j\phi(x,y)}. \tag{3.6}$$

The effect of the phase in (3.6) when compared to (3.5) on the resulting image plane field causes diminished system performance.

Phase must be measured indirectly using sensors that can make scalar measurements of the irradiance, $I[n_1,n_2]$, which is proportional to the electric field amplitude squared [21].

Some sensors, like interferometers, use the superposition of two waves to determine a relative phase between the two waves [22]. Other sensors, such as a Shack-Hartmann, perform a centroid operation on irradiance measurements to determine a local wavefront gradient relative to the unperturbed centroid location [23].

Wavefront reconstruction is a mathematical formulation that uses sensor measurements and outputs an estimate of the wavefront surface, which can be used for optical characterization or feedback control. The goal of wavefront reconstruction algorithms is to estimate the wavefront $\phi(x,y)$ from measurements of its gradient vector $\nabla\phi(x,y)$, which is the focus of Chapters 3 and 4.

3.1 Wavefront Reconstruction in Adaptive Optics

The main goal of AO is to compensate for the phase distortion of the incoming optical field. Phase compensation is provided by a DM which properly adds optical path length to the incoming light, thus equalizing the phase as desired. As shown in Figure 3.3, the command to the DM is provided by a control system and a phase sensor; the DM makes a mirror shape to match the conjugate of the phase, thereby making the wavefront more planar. Closed-loop AO systems commonly estimate of the wavefront phase to command the controllable DM actuators [3], [24]. Most implemented systems reconstruct the phase at the phase points. The control of a DM will be discussed in a later chapter.

In some wavefront reconstruction applications, local phase differences are sensed, from which the whole relative phase has to be estimated. To this goal, a number of algorithms exist in the literature, mostly based on a least-squares solution. The AO community has developed several algorithms for wavefront phase reconstruction. Ideally the approaches would:

1. Be computationally efficient for a large number of data points;
2. Be robust to measurement noise;
3. Result in perfect reconstruction that is consistent with measurements and boundary conditions of the wavefront using noise-free wavefront sensor data.

The first property is essential for future telescopes that increase the number of actuators or sensor measurements seeking diffraction-limited system performance. The second property

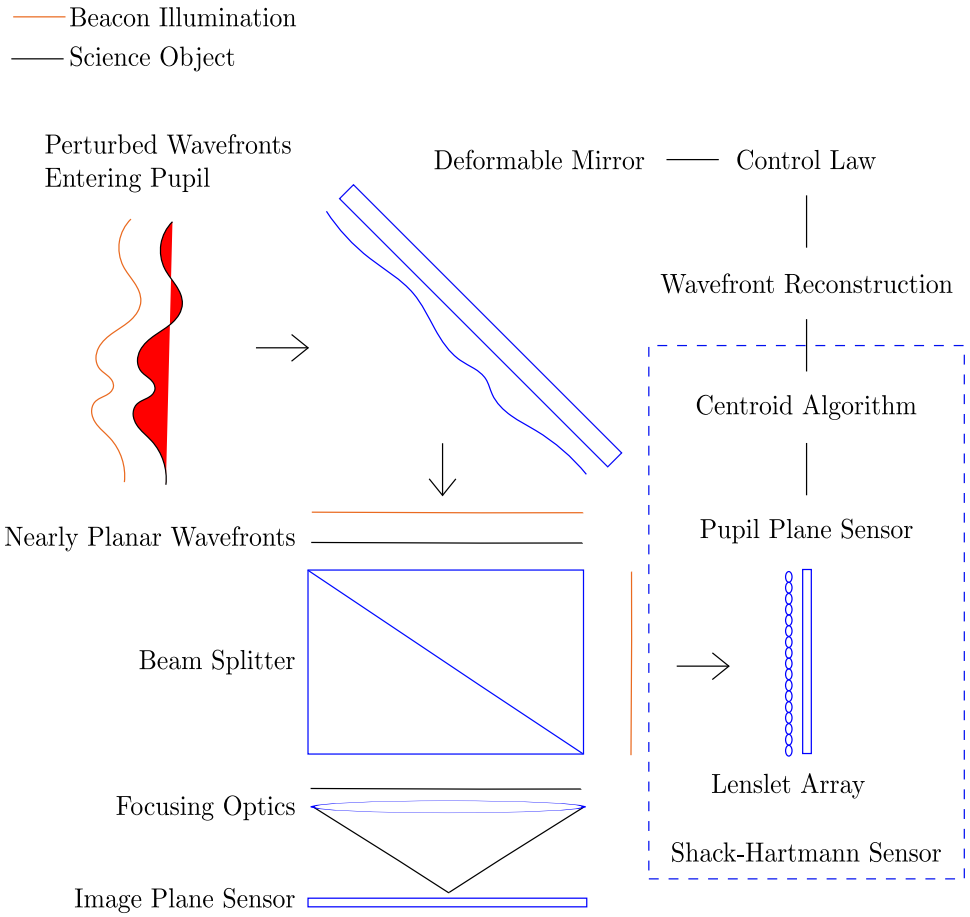


Figure 3.3. Light enters an AO system aperture from a beacon illuminator and science object. It is assumed that the beacon illumination and science object experience the same distortion. The filled red area on the incoming light represents the undesirable distortion on the wavefront that is corrected by the DM.

depends on the mathematical operations performed by the algorithm and in general depends on the statistics of the measurements, including noise correlations. The final property was discussed previously by Southwell [25] and Poyneer [26] and the difficulty arises from the pupil geometry interaction with the algorithm.

The original wavefront reconstruction techniques used matrix-vector multiplies [25,27–29] with computational complexity of $O(n^2)$ or higher with n the total number of samples of the sensed gradient. Later, a Fourier transform technique was proposed by Freischlad [30] with further refinement by Poyneer [26] that has $O(n\log_2 n)$ complexity. The computational cost of the algorithm is limited by the speed of the implementation of the fast Fourier transform

for the change of basis. The papers [26, 30, 31] treat the wavefront sensor measurements as a filtering operation, which is similar to the concept of this dissertation, except that their filtering operations occur on the global data set and solve the entire phase surface at once. During the same time period as Poyneer, Gilles [32] produced a multigrid preconditioned conjugate-gradient method with the same computational complexity.

The first accurate wavefront reconstruction algorithm with complexity $O(n)$ was performed by MacMartin, who developed a multi-resolution hierarchic reconstructor [33]. His work addressed the issues of controlling an actuator by using only local measurements adjacent to the actuator (rather than using all measurements). This "zonal" control can suffer from degradation of "global" modes and his work improved the performance by combining local and global estimates in a hierarchy. To significantly decrease computation, he downsampled by factors of 4 or more and results show that the larger downsampling factor decreases the relative performance and increases the noise multiplier (or noise propagator of [26]) of the reconstructor. In his work, MacMartin averaged gradient measurements, solved actuator estimates at a coarse resolution, and solved for piston at the coarsest level using interpolation and spatial averaging. Our work focuses on the global estimate only and uses a downsampling factor of 2, which gives improved performance of the second property.

Additional $O(n)$ complexity work followed, such as that of Gilles [34]. The work of Gilles was a direct application of a multi-grid solver of a minimum-variance reconstructor based on a sparse approximation of the wavefront inverse covariance matrix. Vogel also improved sparse matrix methods [35] and a multigrid least-squares algorithm [36]. Another minimum-variance solver that followed Vogel is the Fractal Iterative Method (FrIM) [37], [38], which performs a change of basis (using a fractal preconditioner). Minimum-variance reconstruction is an excellent choice, as it is optimal in the sense of maximizing the Strehl ratio [39].

In the last few years, several new wavefront reconstruction algorithms have been proposed. Rosensteiner has produced the Cumulative Reconstructor (CuRe), which is a direct integration reconstructor [40]. More recently, de Visser has shown the SABRE algorithm using B-spline basis functions [41].

An approach based on the wavelet representation was first applied to wavefront reconstruc-

tion by Dowla [42]. This original work did not fully exploit the features of the discrete wavelet transform in a 2D QMF; rather it was an approximation of two iterations of the low-pass filters and downsampling. Peter Hampton and colleagues developed an algorithm that used the complete discrete wavelet transform in a 2D QMF and was able to perform reconstruction using the Haar wavelet [43–45]. This dissertation extends their work to allow for the use of wavelets with the orthogonal property.

In this dissertation, we will further describe the benefits of the wavelet phase reconstruction. The wavelet technique offers a computational efficiency of $O(n)$ using Finite Impulse Response filters. Hampton's derivation uses the Haar wavelet and then uses either a Poisson smoother or recommends a de-noiser as a follow-on step [45]. Our approach in this dissertation is made to be robust to noise using wavelets that have a longer basis length, yielding a smoother reconstruction without a follow-on step. Using noise-free data, the approach also yields an exact reconstruction with a zero mean of the two-dimensional data. We also extend on Hampton's work to provide for a solution that requires no boundary conditions for the Haar wavelet on a square aperture where the side dimensions are a power of 2.

The wavelets approach is based on a multi-resolution analysis solution type. As a consequence of how the discrete wavelet transform is employed, the grid size must be a power of 2. In a Cartesian coordinate frame, wavefront values are iteratively reconstructed first on a 2×2 grid, then 4×4, then 8×8, and expanded as a power of two in size for each iteration. Iteration in this context means that the two matrix dimensions of processed data are doubled each cycle, not that the entire data is processed repeatedly. However, there is no preconditioner or approximation required. The solution algorithm constructs the data for each iteration entirely using the slope measurements provided from the Shack-Hartmann wavefront sensor.

The outline of the remainder of the chapter is as follows: the sensor and model that the wavelet phase reconstruction uses is described in Section 3.2. The wavefront reconstruction 2D QMF signals for each iteration of the analysis section is developed in Section 3.3. The composition of the wavefront estimate using the QMF synthesis filters is shown in Section 3.4. Discussion and simulated wavefront reconstruction examples for several cases are given in Section 3.5. Additional steps to improve performance for data from a telescope with obscurations are given in Section 3.6. The chapter is summarized in Section 3.7.

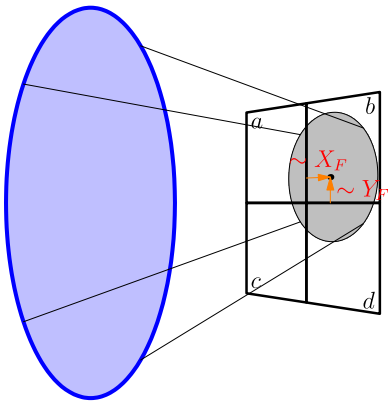


Figure 3.4. A single lenslet is shown. In this example, a centroid algorithm uses the four irradiance measurements at the pixels a through d to determine the center of the beam, which is proportional to the two slope measurements, X_F and Y_F.

3.2 Shack-Hartmann Sensor Geometry

A wavefront reconstruction algorithm is designed to work for a particular sensor geometry. The algorithm presented here is designed to work with sensor data from a Shack-Hartmann wavefront sensor (S-H WFS), which measures local gradients of the wavefront. The sensor has an array of small lenses, or lenslets, and is placed in a relayed pupil plane of the telescope. These lenslets focus the incoming light onto a focal plane detector (which is a standard camera of sufficiently high resolution); the spatial sampling rate of the pupil plane wavefront is determined by the grid of lenslets.

The two local gradient measurements are performed by a centroid algorithm using the measured irradiance at the focal plane pixels, as shown in Figure 3.4. Most S-H WFS treat the centroid operation internally and only provide the two measurement values for each lenslet. The S-H WFS is calibrated so that with a planar wavefront, the centroid of each lenslet yields zero-valued gradients.

This sensor is modeled mathematically by the Fried geometry [27]. In particular, each lenslet is associated to four phase points in the rectangular grid. The model places measurements that are centered on each lenslet so that every lenslet has the two measurements of the vertical and horizontal components of the gradient, denoted as X_F and Y_F, respectively.

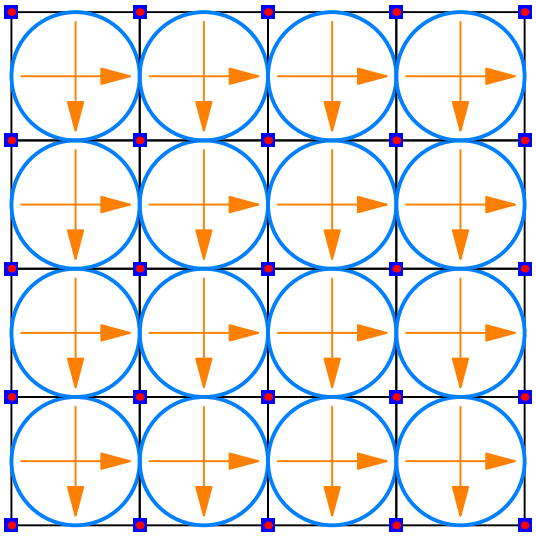


Figure 3.5. An array of 4×4 lenslets shown as blue circles. The red dots at the corners of the grid are the phase point locations that are reconstructed from the orange-arrow slope values.

The phase points are the locations where the wavefront reconstruction obtains phase values. An array of 4×4 lenslets is shown in Figure 3.5. These 16 lenslets are surrounded by 25 phase points and provide 32 slope measurements. The Fried model of a Shack-Hartmann sensor is shown in Figure 3.6 and defines the relationship the slope measurements to the phase point values as

$$\begin{aligned} X_F[n_1,n_2] &\equiv \frac{1}{2}(\phi[n_1-1,n_2]-\phi[n_1-1,n_2-1]+\phi[n_1,n_2]-\phi[n_1,n_2-1]), \\ Y_F[n_1,n_2] &\equiv \frac{1}{2}(\phi[n_1,n_2-1]-\phi[n_1-1,n_2-1]+\phi[n_1,n_2]-\phi[n_1-1,n_2]). \end{aligned} \tag{3.7}$$

Equation (3.7) was rewritten from its original form in [27] to appear in causal form. Close observation of (3.7) reveals that the slope measurements can be written as a separable filtering operation on the phase points. In order to rewrite (3.7) in operator form, we define two filter functions:

$$\begin{aligned} g(z) &\equiv \frac{1}{\sqrt{2}}\left(1+z^{-1}\right), \\ g(-z) &\equiv \frac{1}{\sqrt{2}}\left(1-z^{-1}\right), \end{aligned} \tag{3.8}$$

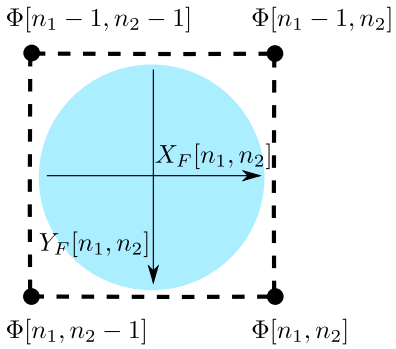


Figure 3.6. The Fried geometry relationship between phase points and their measured slopes lattice for a single lenslet. A Shack-Hartmann sensor will have an array of these lenslets.

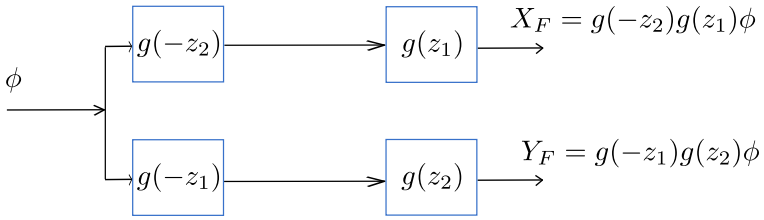


Figure 3.7. The block diagram relationship between the phase points and the Fried geometry.

where the coefficient is chosen that splits the $1/2$ in (3.7) evenly. The geometry involves one low-pass filter and one high-pass filter and is stated in operator form as

$$\begin{aligned} X_F &= g(-z_2)g(z_1)\boldsymbol{\phi}, \\ Y_F &= g(-z_1)g(z_2)\boldsymbol{\phi}. \end{aligned} \tag{3.9}$$

Figure 3.7 depicts the block diagram relationship of the Fried geometry to the wavefront. The form of (3.9) was first stated in [43]. However, as we will see in Section 3.3.1, there is not a direct solution to solve for $\boldsymbol{\phi}$ based on (3.9) alone.

The Haar wavelet is the most simple and was introduced in 1909 [46]. As it turns out, its two filters (low pass and high pass) are simple first-order filters which are identical to the functions in (3.8) that were used for the operator form. Thus, the operator notation of (3.9) is the connection of the Haar wavelet to the Fried geometry.

In order to reconstruct a smooth, continuous $\hat{\phi}[\mathbf{n}]$ from the slope measurements, the tree structure 2D QMF is employed. The signals X_F and Y_F can be processed using multirate

DSP to provide the necessary ϕ_{LL}, ϕ_{LH}, ϕ_{HL}, and ϕ_{HH} signals for the 2D QMF.

In a more general setting, it can be shown that the filters in Figure 2.5 that are associated to the orthogonal wavelet families can be factored [47] as follows

$$\begin{aligned} G(z) &= g(z)G_0(z), \\ H(z) &= g(-z)H_0(z), \\ \tilde{G}(z) &= g(z)\tilde{G}_0(z), \\ \tilde{H}(z) &= g(-z)\tilde{H}_0(z) \end{aligned} \tag{3.10}$$

where $G_0(z)$, $H_0(z)$, $\tilde{G}_0(z)$, $\tilde{H}_0(z)$ are transfer functions of FIR filters.

In the particular case of the Haar wavelet, $G_0(z) = H_0(z) = \tilde{G}_0(z) = 1; \tilde{H}_0(z) = -1$. The factorization in (3.10) is at the basis of the phase reconstruction algorithm presented next.

3.3 Analysis 2D QMF Stage

The phase is reconstructed by the proposed algorithm in two stages: analysis and synthesis. In the analysis, we determine the wavelet coefficients of the phase at different resolutions, based on measured gradients. These coefficients are then used by the synthesis to reconstruct the phase. The challenge is in creating the analysis signals from the measured gradients, as the synthesis section is standard. In this section, we develop the mathematical relationship between slope measurements and the QMF signals at each iteration.

3.3.1 Iteration for Level 1

From observation of the 2D QMF in Figure 2.5, we can write equations which relate the phase points to the four channels of the 2D QMF, where we omit indices n_1, n_2 for convenience, as

$$\begin{aligned} \phi^1_{LL} &= D_2 D_1 \tilde{G}(z_2)\tilde{G}(z_1)\phi, \\ \phi^1_{LH} &= D_2 D_1 \tilde{H}(z_2)\tilde{G}(z_1)\phi, \\ \phi^1_{HL} &= D_2 D_1 \tilde{G}(z_2)\tilde{H}(z_1)\phi, \\ \phi^1_{HH} &= D_2 D_1 \tilde{H}(z_2)\tilde{H}(z_1)\phi. \end{aligned} \tag{3.11}$$

We have swapped the order of the operators in (3.11) for a convenience of notation and use the superscript 1 to denote the first iteration, not an exponent. Expanding (3.11) using the

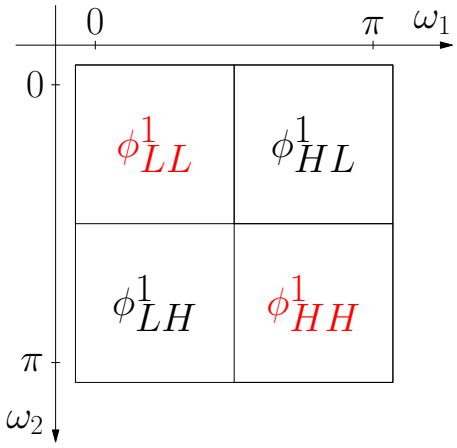


Figure 3.8. The 2D QMF is performed for the first iteration. At this point, half of the frequency domain content is directly computable (in black). The LL and HH sections must be properly iterated.

factoring of (3.10), we can make substitutions from (3.9) that result in

$$\begin{aligned}
\phi^1_{LH} &= D_2 D_1 \tilde{H}_0(z_2)\tilde{G}_0(z_1)\left(g(-z_2)g(z_1)\phi\right) \\
&= D_2 D_1 \tilde{H}_0(z_2)\tilde{G}_0(z_1)X_F, \\
\phi^1_{HL} &= D_2 D_1 \tilde{G}_0(z_2)\tilde{H}_0(z_1)\left(g(z_2)g(-z_1)\phi\right) \\
&= D_2 D_1 \tilde{G}_0(z_2)\tilde{H}_0(z_1)Y_F.
\end{aligned} \tag{3.12}$$

The remaining two quantities, ϕ^1_{LL} and ϕ^1_{HH}, do not have a simple relationship to the slope measurements X_F and Y_F, because the filters of Equation (3.11) are both low-pass or high-pass, whereas the slope definitions require one of each. The quantities are shown in the frequency domain in Figure 3.8 with L and H corresponding to low and high frequencies, respectively.

As seen in (3.12) the terms LH and HL have the proper mix of low-pass $g(z)$ and high-pass $g(-z)$ filters to generate the measured gradients X_F and Y_F. The terms LL and HH do not contain a low-pass and high-pass filters, which indicates the need to further decimate these signals as shown in the next subsection.

3.3.2 Iteration for Level 2

The arguments presented in what follows are based on two fundamental operations:

1. the Noble identities, Equation (2.22) in Section 2.3, recalled here for convenience:

$$\begin{aligned} y_1 &= D_1 H(z_1^2)x = H(z_1)D_1 x, \\ y_2 &= D_2 H(z_2^2)x = H(z_2)D_2 x. \\ y_3 &= D_1 H(z_2)x = H(z_2)D_1 x, \\ y_4 &= D_2 H(z_1)x = H(z_1)D_2 x. \end{aligned} \tag{3.13}$$

2. geometric sum identities, listed in (3.14), which are properties of the Haar high-pass filter:

$$g(-z^N) = \left(\sum_{\ell=0}^{N-1} z^{-\ell} \right) g(-z) \qquad \forall N \geq 1 \tag{3.14a}$$

$$= \left(\sum_{\ell=0}^{\frac{N}{2}-1} z^{-2\ell} \right) \sqrt{2} g(z) g(-z), \qquad \text{if } N \text{ is even.} \tag{3.14b}$$

Equation (3.14) reveals that high-order complexity filters can be implemented as a delayed summation of first-order filters and its complete proof is shown in the Appendix. Equation (3.14a) is used when a high-pass filter is needed, whereas (3.14b) is used when a low-pass filter is needed.

In this algorithm, when there is an unknown quantity, we apply the 2D DWT and break apart the unknown quantity into 4 new channels. We now explore the second iteration because we have two unknown quantities from the previous section; we must do this for both the ϕ_{LL}^1 and ϕ_{HH}^1 channels. First we will only consider ϕ_{LL}^1; we start by writing out the expressions for the four channels in the analysis:

$$\begin{aligned} \phi_{LL/L}^2 &= D_2 D_1 \tilde{G}(z_2)\tilde{G}(z_1)\phi_{LL}^1 \\ &= D_2 D_1 \tilde{G}(z_2)\tilde{G}(z_1) D_2 D_1 \tilde{G}(z_2)\tilde{G}(z_1)\phi, \\ \phi_{LH/L}^2 &= D_2 D_1 \tilde{H}(z_2)\tilde{G}(z_1)\phi_{LL}^1 \\ &= D_2 D_1 \tilde{H}(z_2)\tilde{G}(z_1) D_2 D_1 \tilde{G}(z_2)\tilde{G}(z_1)\phi, \end{aligned}$$

$$
\begin{aligned}
\phi^2_{HL/L} &= D_2 D_1 \tilde{G}(z_2)\tilde{H}(z_1)\phi^1_{LL} \\
&= D_2 D_1 \tilde{G}(z_2)\tilde{H}(z_1) D_2 D_1 \tilde{G}(z_2)\tilde{G}(z_1)\phi, \\
\phi^2_{HH/L} &= D_2 D_1 \tilde{H}(z_2)\tilde{H}(z_1)\phi^1_{LL} \\
&= D_2 D_1 \tilde{H}(z_2)\tilde{H}(z_1) D_2 D_1 \tilde{G}(z_2)\tilde{G}(z_1)\phi.
\end{aligned} \tag{3.15}
$$

We are using the superscript 2 for the second iteration and add the /L subscript for the ϕ^1_{LL} data. There is additionally a /H analysis for the ϕ^1_{HH} data. For brevity, we will only state the HH results at the end of this section since their development follows the same analysis, but we emphasize the resulting expression is not exactly the same. Examining Equation (3.15) reveals that while $\phi^2_{LL/L}$ is only comprised of low-pass filters, the remaining three channels have a combination of low-pass and high-pass filters. If we factor the filters as in (3.10), we will again find substitutions of (3.9) with the measured slope data. For $\phi^2_{LH/L}$ we obtain

$$
\begin{aligned}
\phi^2_{LH/L} =& D_2 D_1 \tilde{H}_0(z_2) g(-z_2)\tilde{G}(z_1) \\
& D_2 D_1 \tilde{G}(z_2)\tilde{G}_0(z_1) g(z_1)\phi \\
=& D_2 D_1 \tilde{H}_0(z_2)\tilde{G}(z_1) \\
& D_2 D_1 \tilde{G}(z_2)\tilde{G}_0(z_1)\left(g(-z_2^2) g(z_1)\phi\right) \\
=& D_2 D_1 \tilde{H}_0(z_2)\tilde{G}(z_1) \\
& D_2 D_1 \tilde{G}(z_2)\tilde{G}_0(z_1) g(z_2)\left(\sqrt{2} X_F\right).
\end{aligned} \tag{3.16}
$$

Using the same procedure, we can also solve for $\phi^2_{HL/L}$ as

$$
\begin{aligned}
\phi^2_{HL/L} =& D_2 D_1 \tilde{G}(z_2)\tilde{H}_0(z_1) g(-z_1) \\
& D_2 D_1 \tilde{G}_0(z_2) g(z_2)\tilde{G}(z_1)\phi \\
=& D_2 D_1 \tilde{G}(z_2)\tilde{H}_0(z_1) \\
& D_2 D_1 \tilde{G}_0(z_2)\tilde{G}(z_1)\left(g(-z_1^2) g(z_2)\phi\right) \\
=& D_2 D_1 \tilde{G}(z_2)\tilde{H}_0(z_1) \\
& D_2 D_1 \tilde{G}_0(z_2)\tilde{G}(z_1) g(z_1)\left(\sqrt{2} Y_F\right).
\end{aligned} \tag{3.17}
$$

The final channel, $\phi^2_{HH/L}$, yields two possible simplifications

$$\begin{aligned}
\phi^2_{HH/L} =& D_2 D_1 \tilde{H}_0(z_2) g(-z_2) \tilde{H}(z_1) \\
& D_2 D_1 \tilde{G}(z_2) \tilde{G}_0(z_1) g(z_1) \phi \\
=& D_2 D_1 \tilde{H}_0(z_2) \tilde{H}(z_1) \\
& D_2 D_1 \tilde{G}(z_2) \tilde{G}_0(z_1) \left(g(-z_2^2) g(z_1) \phi \right) \\
& \text{or similarly} \\
=& D_2 D_1 \tilde{H}_0(z_1) \tilde{H}(z_2) \\
& D_2 D_1 \tilde{G}(z_2) \tilde{G}_0(z_1) \left(g(-z_1^2) g(z_2) \phi \right).
\end{aligned} \tag{3.18}$$

The two possible substitutions arise from the flexibility of having two high-pass filters. Either simplification is exact when using noise-free data. Rather than choose one definition over the other, we take an average

$$\begin{aligned}
\phi^2_{HH/L} = & \frac{1}{2} D_2 D_1 \tilde{H}_0(z_2) \tilde{H}(z_1) \\
& D_2 D_1 \tilde{G}(z_2) \tilde{G}_0(z_1) g(z_2) \left(\sqrt{2} X_F \right) \\
& + \frac{1}{2} D_2 D_1 \tilde{H}(z_2) \tilde{H}_0(z_1) \\
& D_2 D_1 \tilde{G}_0(z_2) \tilde{G}(z_1) g(z_1) \left(\sqrt{2} Y_F \right).
\end{aligned} \tag{3.19}$$

The averaging of (3.19) allows for some robustness against noise in the slope measurements at very little computational cost. The $1/2$ coefficient simply assumes additive white Gaussian noise. Noise correlation statistics analysis may provide a better coefficient. Using

the same process for the HH data, we now state the results as

$$\begin{aligned}
\phi^2_{LH/H} &= D_2D_1\tilde{H}_0(z_2)\tilde{G}(z_1) \\
&\qquad D_2D_1\tilde{H}(z_2)\tilde{H}_0(z_1)g(-z_2)\left(\sqrt{2}Y_F\right), \\
\phi^2_{HL/H} &= D_2D_1\tilde{G}(z_2)\tilde{H}_0(z_1) \\
&\qquad D_2D_1\tilde{H}(z_1)\tilde{H}_0(z_2)g(-z_1)\left(\sqrt{2}X_F\right), \\
\phi^2_{HH/H} &= \frac{1}{2}D_2D_1\tilde{H}_0(z_2)\tilde{H}(z_1) \\
&\qquad D_2D_1\tilde{H}(z_2)\tilde{H}_0(z_1)g(-z_2)\left(\sqrt{2}Y_F\right) \\
&\quad + \frac{1}{2}D_2D_1\tilde{H}(z_2)\tilde{H}_0(z_1) \\
&\qquad D_2D_1\tilde{H}_0(z_2)\tilde{H}(z_1)g(-z_1)\left(\sqrt{2}X_F\right).
\end{aligned} \tag{3.20}$$

We have now completed the derivation of the second iteration and show it in Figure 3.9. While the equations look complex on paper, actual implementations are straightforward and efficient in processing performance. Every expression is simply a serial grouping of the filter-filter-downsample-downsample block.

3.3.3 Further Iterations

We are able to generalize the formulation for additional iterations. After the second iteration, there are two unknown quantities $\phi^2_{LL/L}$ and $\phi^2_{LL/H}$. We now seek to generalize the formulas for each level $k \geq 2$. Until the final iteration, there will still be two unknown quantities. We first write out the formulas

$$\begin{aligned}
\phi^k_{LH/L} &= \left(D_2D_1\tilde{H}(z_2)\tilde{G}(z_1)\right)\left(D_2D_1\tilde{G}(z_2)\tilde{G}(z_1)\right)^{k-2} \\
&\qquad \left(D_2D_1\tilde{G}(z_2)\tilde{G}(z_1)\right)\phi, \\
\phi^k_{HL/L} &= \left(D_2D_1\tilde{G}(z_2)\tilde{H}(z_1)\right)\left(D_2D_1\tilde{G}(z_2)\tilde{G}(z_1)\right)^{k-2} \\
&\qquad \left(D_2D_1\tilde{G}(z_2)\tilde{G}(z_1)\right)\phi, \\
\phi^k_{HH/L} &= \left(D_2D_1\tilde{H}(z_2)\tilde{H}(z_1)\right)\left(D_2D_1\tilde{G}(z_2)\tilde{G}(z_1)\right)^{k-2} \\
&\qquad \left(D_2D_1\tilde{G}(z_2)\tilde{G}(z_1)\right)\phi.
\end{aligned} \tag{3.21}$$

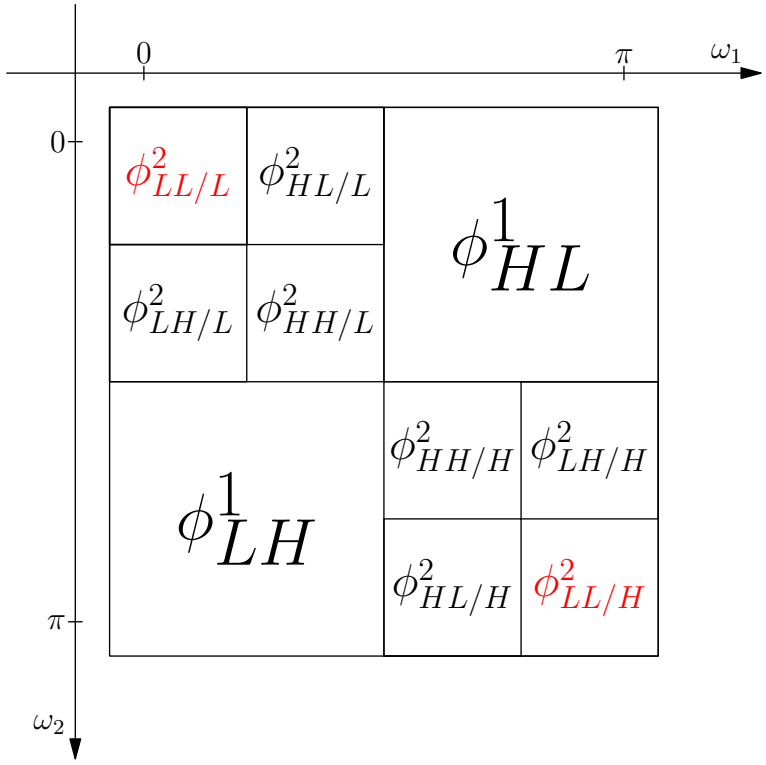


Figure 3.9. The 2D QMF diagram shows the channels at the second iteration. The upper left and lower right are each divided into four channels. Only two blocks remain unknown (in red) that require further decomposition.

The intent of the exponential notation is that the operations inside occur $k-2$ times. We choose to express these equations as three groups since the left group will be factored to shift a filter to the right group. Again we will factor the $\tilde{H}(z)$ on the left side, then move the $g(-z)$ to the right. As $g(-z)$ is swapped position with the downsampling operators, the Noble identities will apply, resulting in $g(-z^{2^{k-1}})$. Then the high-order filter will be simplified to a delayed summation of the first-order filter. The relationship to the slope measurements can then be made. The end result for the LL data is

$$\phi^k_{LH/L} = \left(D_2 D_1 \tilde{H}_0(z_2)\tilde{G}(z_1)\right)\left(D_2 D_1 \tilde{G}(z_2)\tilde{G}(z_1)\right)^{k-2} \left(D_2 D_1 \tilde{G}(z_2)\tilde{G}_0(z_1)\right)\left(\left(\sum_{\ell=0}^{2^{k-1}-1} z_2^{-\ell}\right) X_F\right), \tag{3.22}$$

$$\phi^k_{HL/L} = \left(D_2D_1\tilde{G}(z_2)\tilde{H}_0(z_1)\right)\left(D_2D_1\tilde{G}(z_2)\tilde{G}(z_1)\right)^{k-2} \left(D_2D_1\tilde{G}_0(z_2)\tilde{G}(z_1)\right)\left(\left(\sum_{\ell=0}^{2^{k-1}-1} {z_1}^{-\ell}\right)Y_F\right)$$

with a combination summation for the HH/L channel

$$\begin{aligned}\phi^k_{HH/L} = &\frac{1}{2}\left(D_2D_1\tilde{H}_0(z_2)\tilde{H}(z_1)\right)\left(D_2D_1\tilde{G}(z_2)\tilde{G}(z_1)\right)^{k-2} \left(D_2D_1\tilde{G}(z_2)\tilde{G}_0(z_1)\right)\left(\left(\sum_{\ell=0}^{2^{k-1}-1} {z_2}^{-\ell}\right)X_F\right) \\ &+\frac{1}{2}\left(D_2D_1\tilde{H}(z_2)\tilde{H}_0(z_1)\right)\left(D_2D_1\tilde{G}(z_2)\tilde{G}(z_1)\right)^{k-2} \left(D_2D_1\tilde{G}_0(z_2)\tilde{G}(z_1)\right)\left(\left(\sum_{\ell=0}^{2^{k-1}-1} {z_1}^{-\ell}\right)Y_F\right).\end{aligned} \tag{3.23}$$

The HH data is again developed through the same manner and results in definitions with some slight differences

$$\begin{aligned}\phi^k_{LH/H} = &\left(D_2D_1\tilde{H}_0(z_2)\tilde{G}(z_1)\right)\left(D_2D_1\tilde{G}(z_2)\tilde{G}(z_1)\right)^{k-2} \left(D_2D_1\tilde{H}(z_2)\tilde{H}_0(z_1)\right)\left(\left(\sum_{\ell=0}^{2^{k-2}-1} z_2^{-2\ell}\right)\sqrt{2}g(-z_2)Y_F\right), \\ \phi^k_{HL/H} = &\left(D_2D_1\tilde{G}(z_2)\tilde{H}_0(z_1)\right)\left(D_2D_1\tilde{G}(z_2)\tilde{G}(z_1)\right)^{k-2} \left(D_2D_1\tilde{H}_0(z_2)\tilde{H}(z_1)\right)\left(\left(\sum_{\ell=0}^{2^{k-2}-1} z_1^{-2\ell}\right)\sqrt{2}g(-z_1)X_F\right),\end{aligned} \tag{3.24}$$

and the final channel, HH/H, is again defined as a combination

$$\begin{aligned}\phi^k_{HH/H} = &\frac{1}{2}\left(D_2D_1\tilde{H}_0(z_2)\tilde{H}(z_1)\right)\left(D_2D_1\tilde{G}(z_2)\tilde{G}(z_1)\right)^{k-2}\\ &\left(D_2D_1\tilde{H}(z_2)\tilde{H}_0(z_1)\right)\left(\left(\sum_{\ell=0}^{2^{k-2}-1} z_2^{-2\ell}\right)\sqrt{2}g(-z_2)Y_F\right)\\ &+\frac{1}{2}\left(D_2D_1\tilde{H}(z_2)\tilde{H}_0(z_1)\right)\left(D_2D_1\tilde{G}(z_2)\tilde{G}(z_1)\right)^{k-2}\\ &\left(D_2D_1\tilde{H}_0(z_2)\tilde{H}(z_1)\right)\left(\left(\sum_{\ell=0}^{2^{k-2}-1} z_1^{-2\ell}\right)\sqrt{2}g(-z_1)X_F\right).\end{aligned} \tag{3.25}$$

The summations represent either a zero-padded shift or a circular shift of the data, and the choice should match the preferred implementation of how the sequences are treated for boundary conditions.

By developing this level k implementation, we are able to scale the algorithm for any $2^N \times 2^N$ sized data quickly. This algorithm is possible due to the high-pass filter simplifications of (3.14). We are able to construct the multirate 2D QMF signals for both the LL and HH channels using the measured quantities X_F and Y_F.

3.3.4 Defining the Values for Unsensed Modes

At the final iteration, we have two sets of 2×2 matrices. Each one is of the form

$$\begin{bmatrix}\phi_{LL/L} & \phi_{HL/L}\\ \phi_{LH/L} & \phi_{HH/L}\end{bmatrix}, \qquad \begin{bmatrix}\phi_{HH/H} & \phi_{LH/H}\\ \phi_{HL/H} & \phi_{LL/H}\end{bmatrix} \tag{3.26}$$

and no further downsampling is possible because each entry is a 1×1. The upper-left scalar value of the $\phi_{LL/L}$ channel and lower-right scalar value of the $\phi_{LL/H}$ channel have gone undetermined for all prior iterations. Each value represents undetected modes of the Fried geometry: the piston and "waffle" modes. The piston represents the mean of the entire $\phi[\mathbf{n}]$ data set. Since the S-H WFS only measures differences between phase points and not absolute values, the mean value cannot be known and a separate sensor is required for measuring piston. We can assign $\phi_{LL/L} = 0$ and accept that we are within a constant value of the actual piston of the wavefront. The "waffle mode" represents a nuisance checkerboard pattern

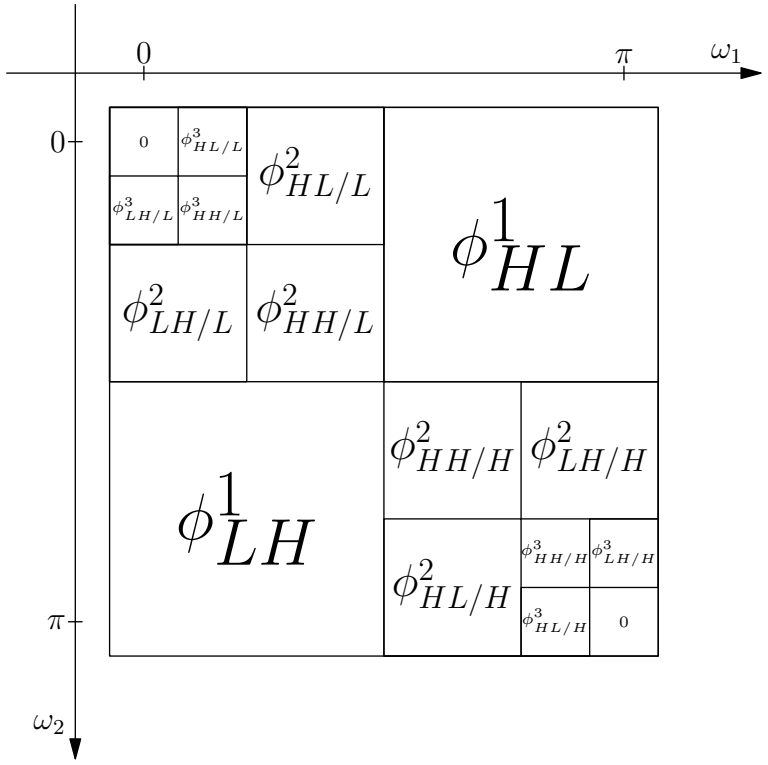


Figure 3.10. The 2D QMF is performed iteratively on the upper left and lower right until scalar values remain. This example figure has three levels until only scalar values remain.

(same amplitude but alternating sign between adjacent values) along the phase points with a mean of zero; it is the highest frequency component in κ_x, κ_y and we can set $\phi_{LL/H} = 0$ by assumption. We show the completed 2D QMF analysis section structure in Figure 3.10, where all values are known.

3.4 Synthesis 2D QMF Stage

Up until this section, all of the previous algorithm steps have been used to iteratively create the four channel blocks of the analysis section. We separated each unknown channel into four sub-channels. While we did not have the direct information for each channel, we were able to substitute for it using the measurements that were available.

The analysis section is now complete and must now take the four channel blocks and perform the inverse discrete wavelet transform as shown in Figure 3.11. The result can be

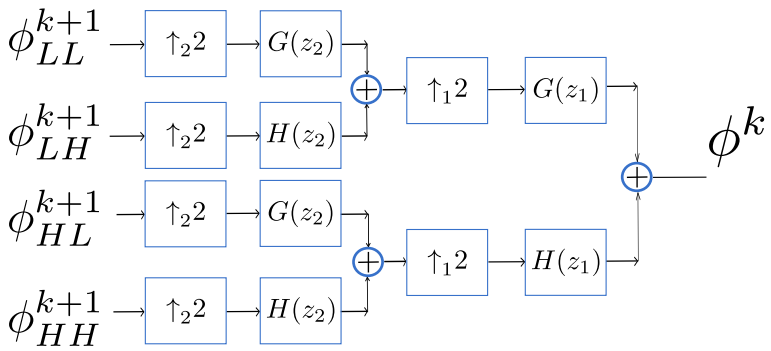


Figure 3.11. Synthesis section of the 2D QMF is shown.

expressed in operator form as

$$\begin{aligned}\phi^k[\mathbf{n}] =& G(z_1)U_1\left(G(z_2)U_2\phi_{LL}^{k+1}[\mathbf{m}]+H(z_2)U_2\phi_{LH}^{k+1}[\mathbf{m}]\right)\\ &+H(z_1)U_1\left(G(z_2)U_2\phi_{HL}^{k+1}[\mathbf{m}]+H(z_2)U_2\phi_{HH}^{k+1}[\mathbf{m}]\right).\end{aligned} \tag{3.27}$$

Equation (3.27) is the standard form of the synthesis section of a 2D QMF and has not been modified to fit the algorithm. For a given index k, both channels $\phi_{LL/L}^k$ and $\phi_{HH/H}^k$ must be processed from their respective QMF signals. The index progresses from the largest iteration index, $k+1=k_{\max}$ and decrements down to $k=1$. Thus, we recursively perform this until we have no more four channel blocks; once $k=1$, the four blocks ϕ_{LL}^1, ϕ_{LH}^1, ϕ_{HL}^1, and ϕ_{HH}^1 can be processed through the synthesis filters one final time that results in the estimated solution of the wavefront phase surface, $\phi[\mathbf{n}]$.

3.5 Discussion

3.5.1 Resampling Haar Wavelet

As previously discussed, the choice of wavelets has an effect on the phase reconstruction and its sensitivity to noise and boundary conditions. Higher-order wavelets (such as the Daubechies family) yield filters with longer impulse response and better filtering capabilities. The drawback of filters with a longer impulse response is clearly the fact that they have longer transient responses, thus extending the effects of the boundary conditions.

On the other hand, the Haar wavelet, which yields the simplest first-order filters at all the stages, if properly implemented, is completely independent of the boundary conditions,

provided the data matrix is square with dimensions as power of two. By adding a positive shift at the analysis network (see Figure 2.3), the four filters become

$$\begin{aligned} \tilde{G}(z) &= zg(z), \qquad & \tilde{H}(z) &= zg(-z) \\ G(z) &= g(z), \qquad & H(z) &= -g(-z) \end{aligned} \tag{3.28}$$

using $g(z)$ defined in (3.8). With this choice of filters, the "approximation" and "details" signals $a[m]$ and $d[m]$ in Figure 2.3 can be related to the input $x[n]$ as

$$\begin{aligned} a[m] &= \frac{1}{\sqrt{2}}\left(x[2m+1]+x[2m]\right) \\ d[m] &= \frac{1}{\sqrt{2}}\left(x[2m+1]-x[2m]\right), \end{aligned} \tag{3.29}$$

and the output becomes

$$\begin{aligned} y[2m] &= \frac{1}{\sqrt{2}}\left(a[m]-d[m]\right) = x[2m] \\ y[2m+1] &= \frac{1}{\sqrt{2}}\left(a[m]+d[m]\right) = x[2m+1]. \end{aligned} \tag{3.30}$$

This choice of filters yields perfect reconstruction $y[n] = x[n]$, $n = 0,\ldots,N-1$, provided the data length N is even, regardless of boundary conditions. In other words, in the case of the Haar wavelet, the effects of boundary conditions get discarded by the resampling operations. If the data length is a power of 2, this will be true for all resolution levels in the decomposition.

3.5.2 Effects of Filter Selection on Noise

The major contribution of this work compared to Hampton's derivation [43] is the extension to more general QMF filters. The Fried model is not always an accurate reconstruction of the wavefront since it only relates neighboring sample points as seen in (3.7). In this section we look at the Daubechies family of filters, namely db$n$, with n positive integer and their effects on the phase reconstruction. In particular, db1 corresponds to the Haar wavelet we presented. The number n corresponds to different filter lengths (or equivalently polynomial lengths). The number of filter coefficients in $\tilde{G}(z)$, $\tilde{H}(z)$, $G(z)$, and $H(z)$ are all twice the

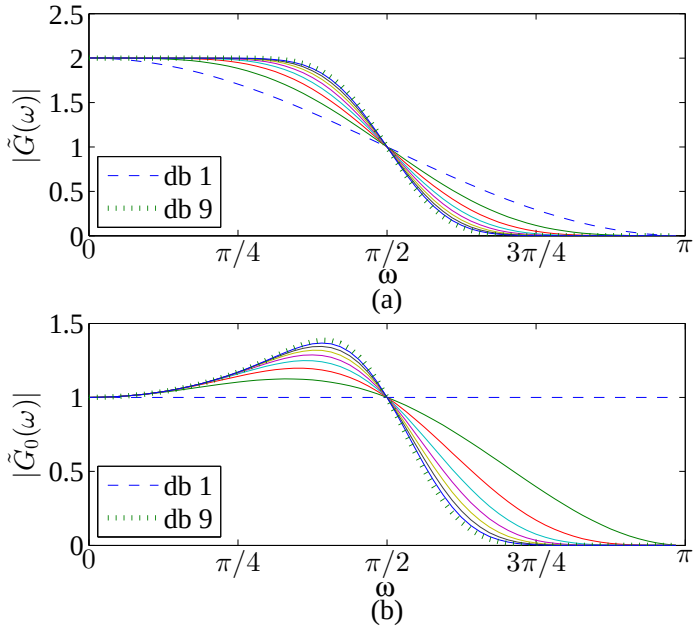


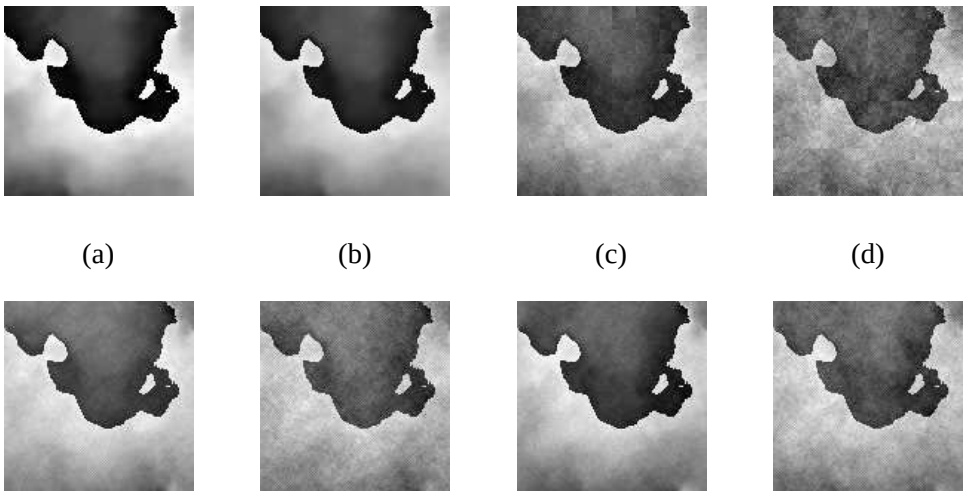
Figure 3.12. (a) The digital frequency response of $\tilde{G}(z)$ for the Daubechies family is illustrated. The frequency response for $\tilde{H}(z)$ would be mirrored at $\pi/2$. (b) The digital frequency response of $\tilde{G}_0(z)$ for the Daubechies family is illustrated. The frequency response for $\tilde{H}_0(z)$ would be mirrored at $\pi/2$. The filtering improvement can be readily seen.

Daubechies number. For example, Daubechies 3 uses filters of length 6.

These factored polynomials of the wavelet factoring offer the same characteristics as the other approaches while also fitting into the wavelet technique. Longer polynomials can smooth the noise, but the choice of filter length should be considered against the dimension length of the data. Having a large filter length but using fewer data does not yield better results.

The frequency response for the Daubechies family is shown in Figure 3.12. These filters are also known as "max-flat" since they are smooth at low and high frequencies. As a consequence, as the order increases the filters become and more selective and provide better attenuation of the aliased components of the noise. The noise has impact on each of the four channels (LL, LH, HL, and HH). This improved noise rejection implies that the noise will be diminished on a subset of the channels.

As AO systems increase in size and therefore also the density of actuators and sensors,



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 3.13. (a) The original 128×128 wavefront; the remaining images are all reconstructed (b) with the Haar wavelet, result is the same as would be for [43]; (c) 10 dB SNR Haar wavelet; (d) 3dB SNR Haar wavelet; (e) 10 dB SNR with the Daubechies 3 wavelet; (f) 10 dB SNR with Daubechies 3 wavelet; (g) 10 dB SNR with Daubechies 9 wavelet; and (h) 3 dB SNR with Daubechies 9 wavelet.

the larger filters are less sensitive to noise than the Haar wavelet and are an appropriate choice. For example, in Figure 3.13, we show the results of several reconstructions using the Daubechies family wavelets with 10 dB and 3 dB signal-to-noise ratio (SNR).

Using only the Haar wavelet, the resulting reconstruction contains a 2×2 checkerboard pattern. The pattern is also apparent at larger resolution in Figure 3.13 (c) and (d). With the longer wavelet lengths, we are able to have a smoothing effect on the result, as shown in Figure 3.13 (e) through (h). In the paper by Hampton et al. [45], they perform smoothing using an iterative Poisson solver which relies on having a previously reconstructed wavefront as an estimate and gradient data that is independent from the estimate. In our work, we are able to provide the smoothing effect from longer filter lengths; which can be seen in comparing Figure 3.13 (e) to (g) or (f) to (h).

3.6 Telescope Apertures

The wavefront reconstruction algorithm presented in Sections 3.3 to 3.5 can be directly applied to data from a telescope with a non-square aperture and other features such as

segmented primary mirror and central obscurations from the secondary mirror and support structure. There will, however, be errors near the boundary edges where the Fried model is incorrect. For improved performance, this section explains how to correct for errors at the boundary for masked data stored within a $2^N \times 2^N$ matrix; however, the boundary correction adds computational complexity, but still quite a bit lower than the standard least-squares approach. The results presented here are a full theoretical explanation, and reductions in operations would be used in an actual real-time implementation.

We begin by defining the mask, or window function, as

$$w[\mathbf{n}] = \begin{cases} 0 & \text{outside aperture} \\ 1 & \text{inside aperture} \end{cases} \tag{3.31}$$

where $\mathbf{n} = [n_1, n_2]$. As we did before, we define the Fried gradient operator as

$$\nabla_F(z_1, z_2) = \begin{bmatrix} g(z_1)g(-z_2) \\ g(-z_1)g(z_2) \end{bmatrix} \tag{3.32}$$

which calculates the values X_F and Y_F for (3.9).

With these two definitions, we can now define two sets of indices for the boundary and inside the aperture as

$$\begin{aligned} \mathscr{B} &= \Big\{ \mathbf{n} \quad | \quad \|\nabla_F w[\mathbf{n}]\| \neq 0 \Big\}, \\ \mathscr{W} &= \Big\{ \mathbf{n} \quad | \quad \|\nabla_F w[\mathbf{n}]\| = 0 \quad \text{and} \quad w[\mathbf{n}] = 1 \Big\}. \end{aligned} \tag{3.33}$$

For simplicity, let us define the entire reconstruction phase reconstruction algorithm presented in Sections 3.3 and 3.4 as an operator $\mathcal{H}$ such that

$$\phi[\mathbf{n}] = \mathcal{H}\left(\nabla_F \phi[\mathbf{n}]\right), \tag{3.34}$$

provided the mean and high frequency modes are both zero. Since all operations in the algorithm are linear, the entire algorithm is linear. By having this property, we can then

write the expression

$$\begin{aligned}\nabla_F\left(w[\mathbf{n}]\phi[\mathbf{n}]\right) &= w[\mathbf{n}]\nabla_F\phi[\mathbf{n}] + \left[\nabla_F\left(w[\mathbf{n}]\phi[\mathbf{n}]\right) - w[\mathbf{n}]\nabla_F\phi[\mathbf{n}]\right] \\ &= w[\mathbf{n}]\begin{bmatrix} X_F[\mathbf{n}] \\ Y_F[\mathbf{n}] \end{bmatrix} + E[\mathbf{n}]\end{aligned} \tag{3.35}$$

where the first term on the right shows the gradient X_F, Y_F masked by the aperture $w[\mathbf{n}]$ and we define the error, $E[\mathbf{n}] \equiv \nabla_F\left(w[\mathbf{n}]\phi[\mathbf{n}]\right) - w[\mathbf{n}]\nabla_F\phi[\mathbf{n}]$. Since it can be easily seen that $E[\mathbf{n}]$ is identically zero outside the boundary, as

$$E[\mathbf{n}] = 0 \quad \Leftrightarrow \quad \mathbf{n} \notin \mathscr{B}, \tag{3.36}$$

the error can be written as

$$E[\mathbf{n}] = \sum_{\boldsymbol{\ell}\in\mathscr{B}} \begin{bmatrix} \tilde{X}_{\boldsymbol{\ell}} \\ \tilde{Y}_{\boldsymbol{\ell}} \end{bmatrix} \delta\left[\mathbf{n} - \boldsymbol{\ell}\right] \tag{3.37}$$

where $\delta[\mathbf{n}]$ is the two-dimensional Kronecker delta function. The terms $\tilde{X}_{\boldsymbol{\ell}}$ and $\tilde{Y}_{\boldsymbol{\ell}}$ are the corrections we need to apply at the boundaries of the aperture. We can solve for $\tilde{X}_{\boldsymbol{\ell}}$ and $\tilde{Y}_{\boldsymbol{\ell}}$ by performing the Haar reconstruction operator to both sides of (3.35), which results in

$$\begin{aligned} w[\mathbf{n}]\phi[\mathbf{n}] = \mathcal{H}&\left(w[\mathbf{n}]\begin{bmatrix} X_F[\mathbf{n}] \\ Y_F[\mathbf{n}] \end{bmatrix}\right) \\ &+ \sum_{\boldsymbol{\ell}\in\mathscr{B}} \tilde{X}_{\boldsymbol{\ell}}\mathcal{H}\left(\begin{bmatrix} \delta\left[\mathbf{n}-\boldsymbol{\ell}\right] \\ 0 \end{bmatrix}\right) \\ &+ \sum_{\boldsymbol{\ell}\in\mathscr{B}} \tilde{Y}_{\boldsymbol{\ell}}\mathcal{H}\left(\begin{bmatrix} 0 \\ \delta\left[\mathbf{n}-\boldsymbol{\ell}\right] \end{bmatrix}\right) \quad \forall \mathbf{n}. \end{aligned} \tag{3.38}$$

By taking the Fried gradient operator of both sides of (3.38) we obtain, for $\mathbf{n} \in \mathscr{W}$

$$\begin{bmatrix} X_F[\mathbf{n}] \\ Y_F[\mathbf{n}] \end{bmatrix} = \nabla_F\mathcal{H}\left(w[\mathbf{n}]\begin{bmatrix} X_F[\mathbf{n}] \\ Y_F[\mathbf{n}] \end{bmatrix}\right) + \sum_{\boldsymbol{\ell}\in\mathscr{B}} \Gamma_X\tilde{X}_{\boldsymbol{\ell}} + \sum_{\boldsymbol{\ell}\in\mathscr{B}} \Gamma_Y\tilde{Y}_{\boldsymbol{\ell}}. \tag{3.39}$$

Using linearity, we define the impulse responses as

$$\begin{aligned}\Gamma_X[\mathbf{n},\boldsymbol{\ell}] &= \nabla_F \mathcal{H}\left(\begin{bmatrix}\delta[\mathbf{n}-\boldsymbol{\ell}]\\ 0\end{bmatrix}\right) \in \mathbb{R}^{2\times 1},\\ \Gamma_Y[\mathbf{n},\boldsymbol{\ell}] &= \nabla_F \mathcal{H}\left(\begin{bmatrix}0\\ \delta[\mathbf{n}-\boldsymbol{\ell}]\end{bmatrix}\right) \in \mathbb{R}^{2\times 1}\end{aligned} \tag{3.40}$$

for $\mathbf{n} \in \mathscr{W}, \boldsymbol{\ell} \in \mathscr{B}$. These definitions can be precomputed, and have no dependence on the slope measurements.

Using the impulse responses $\Gamma_X[\mathbf{n},\boldsymbol{\ell}]$ and $\Gamma_Y[\mathbf{n},\boldsymbol{\ell}]$, for $\mathbf{n} \in \mathscr{W}$ and $\boldsymbol{\ell} \in \mathscr{B}$, we can solve for $\tilde{X}_{\boldsymbol{\ell}}$ and $\tilde{Y}_{\boldsymbol{\ell}}$ in (3.38) from a set of linear equations as

$$\begin{bmatrix}X_F[\mathbf{n}]\\ Y_F[\mathbf{n}]\end{bmatrix} - \nabla_F \mathcal{H}\left(w[\mathbf{n}]\begin{bmatrix}X_F[\mathbf{n}]\\ Y_F[\mathbf{n}]\end{bmatrix}\right) = \sum_{\boldsymbol{\ell}\in\mathscr{B}} \Gamma_X[\mathbf{n},\boldsymbol{\ell}]\tilde{X}_{\boldsymbol{\ell}} + \sum_{\boldsymbol{\ell}\in\mathscr{B}} \Gamma_Y[\mathbf{n},\boldsymbol{\ell}]\tilde{Y}_{\boldsymbol{\ell}} \tag{3.41}$$

for $\mathbf{n} \in \mathscr{W}$. The left-hand side is known from the measured gradients and (3.41) yields $n_{\mathscr{W}} = 2|\mathscr{W}|$ equations in $n_{\mathscr{B}} = 2|\mathscr{B}|$ unknowns, with $|\mathscr{W}|$ and $|\mathscr{B}|$ the number of sample points inside the aperture and on the boundary respectively. It can easily be seen that (3.41) can be written in matrix form as

$$\underline{z}_{\mathscr{W}} = \boldsymbol{\Gamma}\underline{\tilde{z}}_{\mathscr{B}} \tag{3.42}$$

with $\underline{z}_{\mathscr{W}}$ and $\underline{\tilde{z}}_{\mathscr{B}}$ are the corresponding $n_{\mathscr{W}} \times 1$ and $n_{\mathscr{B}} \times 1$ vectors on the left and right hand sides of (3.41), and $\boldsymbol{\Gamma}$ the corresponding $n_{\mathscr{W}} \times n_{\mathscr{B}}$ matrix. Equation (3.42) is underdetermined and therefore $\underline{\tilde{z}}_{\mathscr{B}} = \left(\boldsymbol{\Gamma}^{\mathrm{T}}\boldsymbol{\Gamma}\right)^{-1}\boldsymbol{\Gamma}^{\mathrm{T}}\underline{z}_{\mathscr{W}}$ is solved by least-squares that can be reduced in operations due to many zero-value eigenvalues [21].

To give an idea of the dimensionality, for a 64×64 data matrix containing a circular aperture with radius $\rho = 29$, the matrix $\boldsymbol{\Gamma} \in \mathbb{R}^{4976\times 456}$. However, since all gradients on the boundary yield redundant information, it turns out that the number of unknowns can be considerably reduced. In this example, the matrix $\boldsymbol{\Gamma}^{\mathrm{T}}\boldsymbol{\Gamma}$ can be decomposed as

$$\boldsymbol{\Gamma}^{\mathrm{T}}\boldsymbol{\Gamma} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^{\mathrm{T}} \tag{3.43}$$

where $\lambda_i = \mathrm{diag}(\boldsymbol{\Lambda})$, are the eigenvalues as shown in Figure 3.14. The first 283 eigenvalues

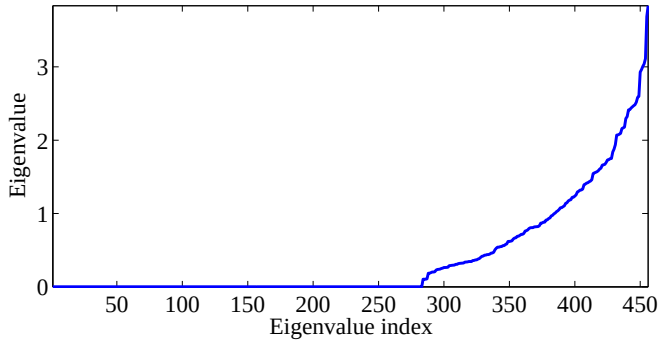


Figure 3.14. The eigenvalues of the 64×64 circular aperture in monotonic order are shown.

are zero and only $M = 173$ eigenvalues are not zero. As a consequence, we can factor

$$\mathbf{\Gamma}^{\mathrm{T}}\mathbf{\Gamma} = \overline{\mathbf{U}}\,\overline{\mathbf{\Lambda}}\,\overline{\mathbf{U}}^{\mathrm{T}} \tag{3.44}$$

with $\overline{\mathbf{\Lambda}} = \mathrm{diag}(\lambda_1, \ldots, \lambda_M)$, $\lambda_i > 0$ and $\overline{\mathbf{U}} \in \mathbb{R}^{n_{\mathscr{B}} \times M}$ orthonormal. Based on this decomposition, the corrective term $\underline{\tilde{z}}_{\mathscr{B}}$ can be solved as

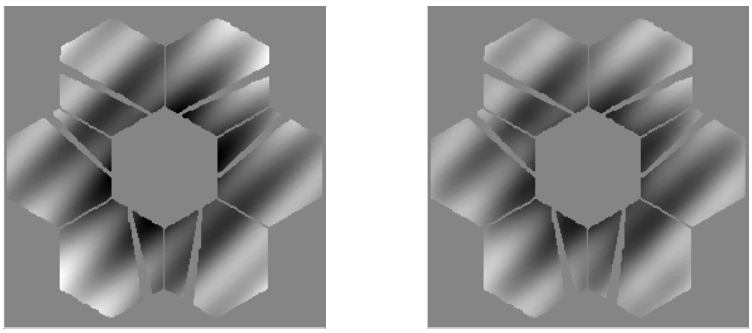
$$\begin{aligned} \underline{\alpha} &= \overline{\mathbf{\Lambda}}^{-1}\overline{\mathbf{U}}^{\mathrm{T}}\overline{\mathbf{\Gamma}}^{\mathrm{T}}\underline{z}_{\mathscr{W}} \\ \underline{\tilde{z}}_{\mathscr{B}} &= \overline{\mathbf{U}}\underline{\alpha} \end{aligned} \tag{3.45}$$

where the matrices $\overline{\mathbf{\Lambda}}^{-1}\overline{\mathbf{U}}^{\mathrm{T}}\overline{\mathbf{\Gamma}}^{\mathrm{T}}$ and $\overline{\mathbf{U}}$ are $M \times n_{\mathscr{W}}$ and $n_{\mathscr{B}} \times M$ respectively. In the example, they would be 173×4976 and 456×173 respectively. All of these matrices are precomputed, since they depend on the aperture only.

Having solved for $\underline{\tilde{z}}_{\mathscr{B}}$, we now can use $\tilde{X}_{\boldsymbol{\ell}}$ and $\tilde{Y}_{\boldsymbol{\ell}}$ in (3.41), so that we have the corrected gradients of $w[\mathbf{n}]\phi[\mathbf{n}]$ as

$$\begin{aligned} X[\mathbf{n}] &= X_F[\mathbf{n}] + \sum_{\boldsymbol{\ell} \in \mathscr{B}} \tilde{X}_{\boldsymbol{\ell}}\delta[\mathbf{n} - \boldsymbol{\ell}], \\ Y[\mathbf{n}] &= Y_F[\mathbf{n}] + \sum_{\boldsymbol{\ell} \in \mathscr{B}} \tilde{Y}_{\boldsymbol{\ell}}\delta[\mathbf{n} - \boldsymbol{\ell}]. \end{aligned} \tag{3.46}$$

Using the corrected gradients X and Y, we run the wavefront reconstruction algorithm (with the option of a different wavelet for better results) to obtain the wavefront phase estimate.



(a) (b)

Figure 3.15. (a) The original 256×256 wavefront with a telescope mask applied; (b) the reconstructed wavefront using the Daubechies 3 wavelet.

In the previous section, we provide example results using a square aperture. We now consider a realistic segmented mirror telescope scenario where there is an outer edge that is non-square and central obscuration by a secondary mirror and its support structure. We simulated this by generating data on a square aperture and then using zero value entries outside of the telescope aperture mask.

In Figure 3.15, the algorithm is applied to simulated data for a notional segmented telescope system. We do not use any boundary correction or modification of the measured wavefront data and the result is still successful in reconstructing the wavefront. In Figure 3.16, we plot the 256 pixels across a row for the original wavefront in comparison with two reconstructions using the Daubechies 3 and Daubechies 9 wavelets. The reconstruction has errors near the boundary edges. Since the Daubechies 3 wavelet is shorter in filter length, it is able to converge to the actual wavefront values closer to the edge than the Daubechies 9. The Daubechies 9 wavelet also has more smoothing than the Daubechies 3 due to increased filter length, and its result has less error in reconstruction when far from the edges such that they have no influence.

In Figure 3.17, the corrected reconstructed wavefront can be compared to the original wavefront and the wavefront reconstructed with the algorithm of Section 3.3 only. The improved performance near the boundary edge is apparent. In addition, the correction also estimates the wavefront hidden underneath the structural support of the secondary mirror.

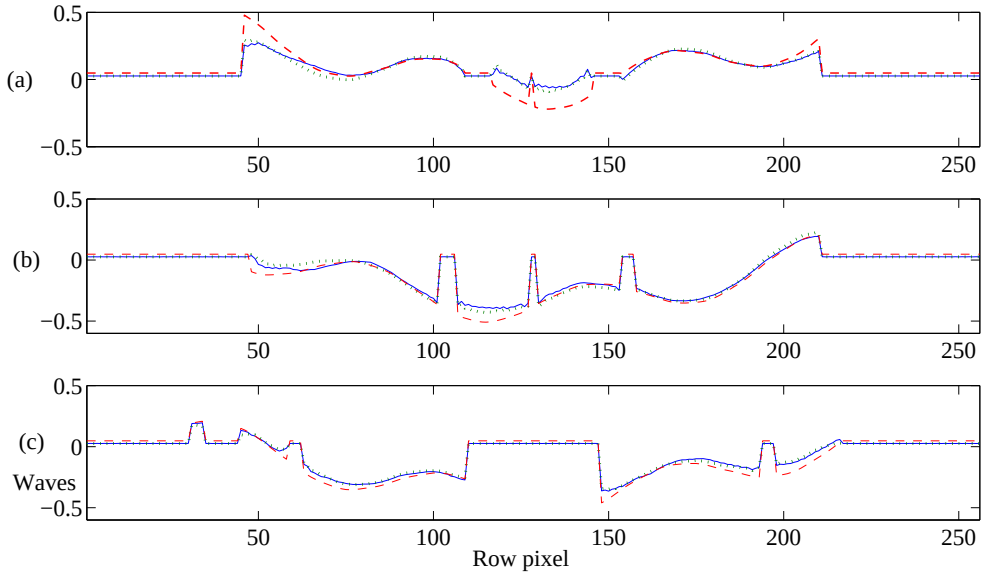


Figure 3.16. Rows (a) 220, (b) 177, and (c) 90 from Figure 3.15 are shown. In each plot, the dashed line shows the original wavefront compared against the reconstructed wavefront shown by solid line for Daubechies 3 and by dotted line for Daubechies 9.

3.7 Summary

In this chapter, we presented a wavelet phase reconstruction algorithm to estimate the phase $\hat{\phi}[\mathbf{n}]$. This quantity is important for AO systems to achieve their goal of improved imaging of science objects. With an estimate of the phase, a DM can be commanded to compensate for atmospheric turbulence which degrades the performance of telescopes.

The algorithm presented in this chapter relies on the premise that the measurements X_F and Y_F are the gradients of the phase function $\phi[\mathbf{n}]$. In the next chapter, we discuss when this is not the case and provide a means to estimate $\hat{\phi}[\mathbf{n}]$ even when the measurements do not satisfy this condition.

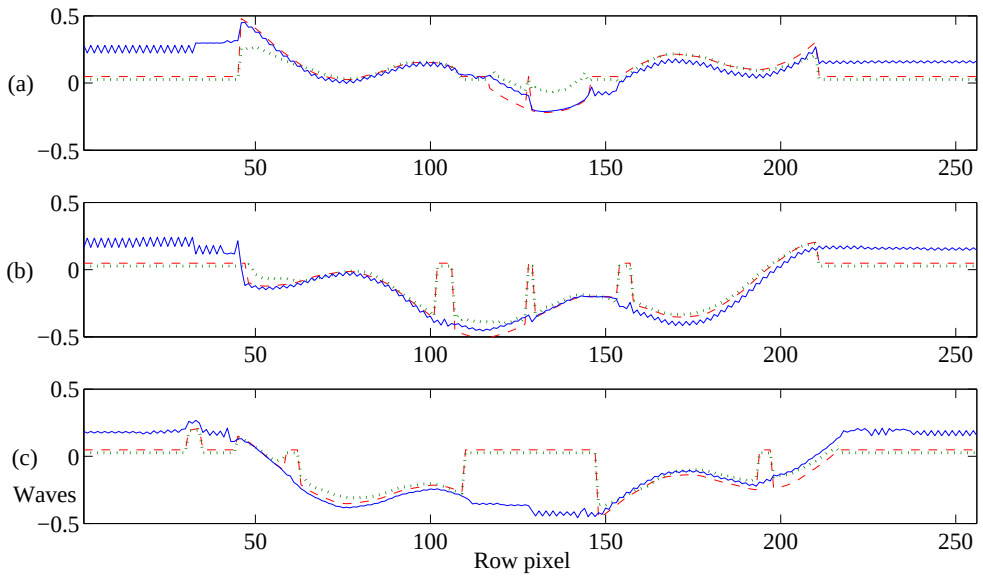


Figure 3.17. Three rows from Figure 3.15 are shown after applying the boundary correction. In each plot, the dashed line shows the original wavefront compared against the reconstructed wavefront shown by dotted line for Daubechies 3 and by solid line for Daubechies 3 that has the boundary correction applied.

THIS PAGE INTENTIONALLY LEFT BLANK

CHAPTER 4:
Reconstruction When Branch Points Are Present

In this chapter, we extend the algorithm of Chapter 3 to more general rotational vector fields, where the irradiance may be zero at isolated points, thus making the phase undefined. The phase can have discontinuities, as the wavefront may be severely degraded.

Strong atmospheric turbulence causes *scintillation*, which is the fluctuation of irradiance. Regions of the telescope pupil can have non-uniform amplitude $A(x,y)$, or *apodization* of the light wave [20]. In some regions, the amplitude decreases significantly and nulls form. A particularly difficult problem can occur for wavefront sensors that take gradient measurements of the scalar $\phi(x,y)$ field near these nulls, as the irradiance measurement to estimate phase will have poor SNR. This issue is referred to as the branch point problem of adaptive optics [48].

The *Rytov variance* for a plane wave [49] is commonly used as a metric of the strength of the scintillation and is defined as

$$\sigma_R^2 = 1.23 C_n^2 k^{7/6} L^{11/6} \tag{4.1}$$

where C_n^2 is the refractive index structure parameter, k is the wavenumber, L is the path length through the atmosphere. Some AO literature uses the measurements of the *Rytov number* σ_χ^2 for the same purpose. The relationship between the two quantities can approximated by $\sigma_R^2 \approx 4\sigma_\chi^2$; for further details see Section 8.2 in [49] and Equation (2.111) in [50]. Although the literature for atmospheric beam propagation [49] establishes the weak scintillation regime by $\sigma_R^2 < 1$, Barchers et al. [51] states that branch points form around a Rytov number of $\sigma_\chi^2 = 0.08$ ($\sigma_R^2 \approx 0.32$), and the presence of branch points becomes significant at $\sigma_\chi^2 = 0.2$ ($\sigma_R^2 \approx 0.8$) such that a Shack-Hartmann wavefront sensor has reduced performance.

In this chapter, we develop a modification to the slope measurements that improves the quality of the reconstruction. This modification is independent of the reconstruction algorithm. In the next section, we introduce branch points. Phase wrapping is discussed

in detail in Section 4.2. In Section 4.3, the vector field decomposition is explained. Section 4.4 is the description and comparison of Fried gradients and wrapped Fried gradients. Section 4.5 is one of the main results of this dissertation and has the least-squares phase estimation from the wrapped Fried gradients. In Section 4.6, we provide several examples of the algorithm. Finally, in Section 4.7, the conclusions are presented.

4.1 Effects of Branch Points on Phase Reconstruction

The branch point phenomenon was first observed by Nye and Berry in 1974 [52]. The original work in phase reconstruction when branch points are present was done by Fried and Vaughn [53]. Their analysis shows that branch points appear in areas of low irradiance. Phase discontinuities form between branch points along *branch cuts*, which prevent path-dependent phase unwrapping from traversing the cuts. Branch cuts are placed between branch points of opposite signs or between a branch point and an edge boundary of the surface. The path of a branch cut is not unique and they can be placed along areas of low irradiance by taking irradiance (SNR) into account in the reconstructor. The issue is how to pair branch points for branch cuts, which is not a trivial problem. Regions of the wavefront may not be completely surrounded by branch cuts, as this prevents phase unwrapping. Short branch cut lengths are desirable for AO systems with a continuous DM that cannot form shapes with discontinuities. The branch cut selection algorithm must also be computationally tractable for AO feedback control. Many branch cut selection algorithms were developed for Synthetic Aperture Radar [54], and can be applied to AO.

Fried [48] continued the analysis of branch points in wavefront reconstruction. He noted that if branch points are present, then the measured phase difference is not a gradient of a scalar field. Instead, the measured phase differences have two components: a gradient of a scalar potential and a curl of the vector potential. The rotational curl component leads to "hidden phase," as the least-squares reconstruction can only recover the irrotational scalar potential. In order to recover the original phase, the curl component must be zero valued. The vector potential is determined as the solution of a Poisson equation, which requires the locations of the branch points. This analysis leads to a closed-form solution of the hidden phase contribution from each branch point. The "total phase" is then the sum of the least-squares reconstructed phase and the hidden phase.

Tyler [55] performed an analysis of branch points by using the Fourier transform of the vector field. He decomposed the observed gradient phase field into irrotational and rotational components. Both components are orthogonal to each other and the rotational component is shown to be hidden in the null space of the standard least squares reconstructor. The "slope discrepancy" was determined to be the difference of the gradient field of the least-squares reconstructed phase and the original measurements. Tyler defines the slope discrepancy as a matrix operator on the gradient measurements that can be used to reconstruct the correction to the least-squares wavefront reconstruction. The correction procedurally follows the least-squares reconstructor.

In our approach, which is presented in this chapter and in [56], we interpret the 2D vector field problem in terms of linear algebra and vector spaces. We begin with the Helmholtz decomposition of the measured vector field in terms of orthogonal subspaces of the gradient and curl components. Our approach is different from [48], [55] by changing to a non-orthogonal decomposition and using least-squares. We can show that a family of wavefronts can be generated that match the measured gradients and our solutions can be adapted to the desired branch cuts.

4.2 Phase Wrapping

A least-squares reconstructor cannot produce the discontinuities of the wavefront caused by the presence of branch points, and incorrectly estimates the phase values across the wavefront surface. Although a smooth phase function cannot be determined, there is a family or ensemble of phase functions that all have the same gradient measurements and different algorithms may result in different phase functions using the same measurements for this reason. Unwrapped phase functions are also members of the ensemble.

There are several phase quantities to note. The true phase is $\phi[n_1, n_2]$. This quantity can never be known exactly, but it can be estimated as a relative phase. The only guaranteed relationship between true phase and the estimated phase is that the wrapped quantities are equivalent, as

$$\mathcal{W}\{\phi[\mathbf{n}]\} = \mathcal{W}\{\hat{\phi}[\mathbf{n}] + C\}, \tag{4.2}$$

where C is a constant. Wrapped phase is always guaranteed to be contained in a 2π range; for our work here we prefer the definition $-\pi \leq \mathcal{W}\{\phi[\mathbf{n}]\} < \pi$. The difficulty imposed by

the wrapping is that it is a nonlinear operator of the form

$$\mathcal{W}\{\phi[\mathbf{n}]\} = \phi[\mathbf{n}] + 2\pi k[\mathbf{n}] \tag{4.3}$$

where $k[\mathbf{n}]$ are integer values that ensure $-\pi \leq \phi[\mathbf{n}] + 2\pi k[\mathbf{n}] < \pi$. It is impossible to know the original true phase from the wrapped values; only relative phase values can be known.

Unwrapping algorithms can remove any sharp discontinuities by finding values for $k[\mathbf{n}]$ that result in a continuous surface for the relative phase. Unwrapped phase is not restricted to a particular range, but it is desirable that the surface is continuous and there are no sharp transitions between adjacent phase points. Thus, neighboring phase points are kept to less than a 2π difference to avoid any ambiguities. This definition means that the constraint for unwrapped phase is on its difference as

$$\begin{aligned} |\phi[n_1, n_2] - \phi[n_1 - 1, n_2]| &< 2\pi \\ |\phi[n_1, n_2] - \phi[n_1, n_2 - 1]| &< 2\pi \end{aligned} \tag{4.4}$$

Increasing the spatial sampling rate is one way to avoid a 2π or larger difference between two adjacent phase values (or the limit imposed by the sensor).

In general, the estimated phase output from the wavefront reconstruction is not guaranteed to be either a wrapped or unwrapped phase. In many cases, the estimated wavefront is smooth and can be considered the unwrapped phase. However, there are instances where this condition is not satisfied by the wavefront reconstruction algorithm, and the result requires a phase unwrapping algorithm to yield a continuous wavefront surface. While some reconstruction algorithms do perform unwrapping (such as the exponential reconstructor [57–59]), we treat reconstruction and unwrapping as two separate operations. In this dissertation, we are only presenting a novel reconstruction algorithm and not an unwrapping algorithm.

4.3 Vector Field Decompositions

The Helmholtz decomposition states that a vector field in the three dimensional space is represented by a gradient component and a rotational component as

$$\psi = \nabla\phi + \nabla \times \mathbf{v} \tag{4.5}$$

where $\nabla = [\nabla_x, \nabla_y, \nabla_z]^T$ represents the gradient operator, $\phi(x,y,z) \in \mathbb{R}$ is the scalar potential and $\mathbf{v}(x,y,z) \in \mathbb{R}^3$ is the vector potential, with the outer product $\nabla\times$ defining the curl of the vector. As we can observe in (4.5), vector fields are composed of a gradient of function and a curl of a vector field. The gradient of a function is an irrotational (curl-free) vector field; whereas the curl of the vector field is a rotational (divergence-free) vector field. In particular, in the case of a two dimensional vector field in the x,y plane

$$\psi(x,y) = \begin{bmatrix} \psi_x(x,y) \\ \psi_y(x,y) \end{bmatrix} \tag{4.6}$$

where we assume the component along the z axis to be identically zero, the scalar potential is $\phi(x,y)$ and the vector potential is along the z axis as $\mathbf{v} = [0,0,v(x,y)]^T$. These assumptions lead to a simple expression of the decomposition (4.6) in matrix form as

$$\begin{bmatrix} \psi_x(x,y) \\ \psi_y(x,y) \end{bmatrix} = \begin{bmatrix} \nabla_x & \nabla_y \\ \nabla_y & -\nabla_x \end{bmatrix} \begin{bmatrix} \phi(x,y) \\ v(x,y) \end{bmatrix}. \tag{4.7}$$

In the Fourier domain, equation (4.7) relates complex vectors as

$$\begin{bmatrix} \Psi_x(\boldsymbol{\kappa}) \\ \Psi_y(\boldsymbol{\kappa}) \end{bmatrix} = \begin{bmatrix} j2\pi\kappa_x & j2\pi\kappa_y \\ j2\pi\kappa_y & -j2\pi\kappa_x \end{bmatrix} \begin{bmatrix} \Phi(\boldsymbol{\kappa}) \\ V(\boldsymbol{\kappa}) \end{bmatrix} \tag{4.8}$$

with $\Phi(\boldsymbol{\kappa})$ and $V(\boldsymbol{\kappa})$ the 2D Fourier transforms of $\phi(x,y)$ and $v(x,y)$ respectively, and $\boldsymbol{\kappa} = [\kappa_x, \kappa_y]$. Equation (4.8) corresponds to the decomposition of the complex vector $\Psi(\boldsymbol{\kappa})$ in terms of the orthogonal reference frame defined by the two columns of the matrix in

(4.8). Simple matrix inversion yields the two potentials (scalar and vector) computed as

$$\left((j2\pi\kappa_x)^2+(j2\pi\kappa_y)^2\right)\begin{bmatrix}\phi(\boldsymbol{\kappa})\\ V(\boldsymbol{\kappa})\end{bmatrix}=\begin{bmatrix}j2\pi\kappa_x & j2\pi\kappa_y\\ j2\pi\kappa y & -j2\pi\kappa_x\end{bmatrix}\begin{bmatrix}\Psi_x(\boldsymbol{\kappa})\\ \Psi_y(\boldsymbol{\kappa})\end{bmatrix}. \tag{4.9}$$

The problem of phase reconstruction is to recover the overall phase we call $\phi_0(x,y)$ from observed gradients $\psi(x,y)$. This problem is typical in adaptive optics [24] or Interferometric Synthetic Aperture Radars [60], where local phase differences are observed directly. When the vector field is irrotational, i.e., the curl component $v(x,y)$ is absent, the overall phase $\phi_0(x,y)$ is the same as the scalar potential $\phi(x,y)$ since by definition,

$$\psi(x,y)=\nabla\phi(x,y). \tag{4.10}$$

When the field is irrotational, we have seen that the integral of the measured gradient is independent of the path of integration. In practice, we do not integrate along a path, due to noise and the fact that we want to use all the data we have rather than only the measurements along the path.

Algorithms designed for this reconstruction are based on a matrix representation of phase differences [26–28] as

$$\mathrm{vec}\,(\psi[:,:])=\mathbf{\Gamma}\mathrm{vec}\,(\phi[:,:]) \tag{4.11}$$

with $\mathrm{vec}(\cdot)$ representing matrix-to-vector reshaping of sampled gradients ψ and potential ϕ, and $\mathbf{\Gamma}$ a matrix of appropriate dimensions with approximately twice the number of rows then columns. Solving for ϕ in (4.11) yields an overdetermined set of equations solved by least squares to estimate $\hat{\phi}$ as

$$\mathrm{vec}(\hat{\phi}[:,:])=\left(\mathbf{\Gamma}^T\mathbf{\Gamma}\right)^{-1}\mathbf{\Gamma}^T\mathrm{vec}(\psi[:,:]). \tag{4.12}$$

Although in the applications of interest, the observation vector $\psi(x,y)$ is made of phase gradients, the presence of singularities and the fact that all phase values are wrapped to lie within the interval $[-\pi,\pi)$, makes the vector potential $v(x,y)$ to be nonzero. As a consequence, the phase $\hat{\phi}(x,y)$ to be estimated is not the same as the scalar potential $\phi(x,y)$. The substitution of (4.5) into (4.12) shows that the vector potential information is lost in the orthogonal frame. Therefore the computation in (4.12) yields a least squares approxima-

tion for the scalar potential and cannot account for the orthogonal component (the "hidden phase" in multiple references such as [48] and [55]) associated to the null space of the matrix $\mathbf{\Gamma}$ in equation (4.11).

A solution to this problem, proposed in this dissertation, is to use a different, non-orthogonal reference frame. This frame will be shown in the next section to be well suited to the computation of phase data in the presence of phase wrapping and singularities such as branch points. In order to see this, we replace the representation in (4.7) with a non-orthogonal frame as follows

$$\begin{bmatrix} \psi_x(x,y) \\ \psi_y(x,y) \end{bmatrix} = \begin{bmatrix} \nabla_x & 0 \\ \nabla_y & -\nabla_y \end{bmatrix} \begin{bmatrix} \phi_0(x,y) \\ \overline{c}(x,y) \end{bmatrix} \tag{4.13}$$

or, equivalently,

$$\begin{bmatrix} \psi_x(x,y) \\ \psi_y(x,y) \end{bmatrix} = \begin{bmatrix} \nabla_x & \nabla_x \\ \nabla_y & 0 \end{bmatrix} \begin{bmatrix} \phi_1(x,y) \\ \overline{c}(x,y) \end{bmatrix}. \tag{4.14}$$

In this non-orthogonal frame, where the two basis vectors in the frequency domain are given by

$$e_1 = \begin{bmatrix} j2\pi\kappa_x \\ j2\pi\kappa_y \end{bmatrix}, \quad e_2 = \begin{bmatrix} 0 \\ -j2\pi\kappa_y \end{bmatrix} \text{ or } e_2 = \begin{bmatrix} j2\pi\kappa_x \\ 0 \end{bmatrix}, \tag{4.15}$$

the two components $\phi_0(x,y)$ or $\phi_1(x,y)$ and $\overline{c}(x,y)$ are given by the following:

Lemma. Let $c(x,y)$ be the curl of the vector field $\psi(x,y)$, i.e.,

$$c(x,y) \equiv \nabla_y \psi_x(x,y) - \nabla_x \psi_y(x,y) \tag{4.16}$$

and let $\phi(x,y)$, $v(x,y)$ be the scalar and vector potentials as in (4.7). Note that the sign convention of our definition of curl in (4.16) is opposite of the standard mathematics definition; we do this to simplify some signs in subsequent results.

Also define $\overline{c}(x,y)$ and $\overline{v}(x,y)$ in differential equation form such that

$$\begin{aligned} c(x,y) &\equiv \nabla_x \nabla_y \overline{c}(x,y), \\ v(x,y) &\equiv \nabla_x \nabla_y \overline{v}(x,y) \end{aligned} \tag{4.17}$$

with the boundary condition $w(x,y)$. The integral form of (4.17) is

$$\overline{c}(x,y) \equiv \int_0^y \int_0^x c(\lambda_1,\lambda_2) d\lambda_1 d\lambda_2 + w(x,y). \tag{4.18}$$

Then the vector field $\psi(x,y)$ can be expressed as in (4.13) or (4.14) with

$$\begin{aligned} \phi_0(x,y) &= \phi(x,y) + {\nabla_y}^2 \overline{v}(x,y) + w_y(y), \\ \phi_1(x,y) &= \phi(x,y) - {\nabla_x}^2 \overline{v}(x,y) + w_x(x) \end{aligned} \tag{4.19}$$

with $w_x(x)$ and $w_y(y)$ depending on boundary conditions.

Proof. From the bottom equation in (4.9) and the definition of the curl $c(x,y)$ in (4.16), we obtain

$$\left({\nabla_x}^2 + {\nabla_y}^2\right) v(x,y) = c(x,y). \tag{4.20}$$

Substitution of $v(x,y),\ c(x,y)$ with $\overline{v}(x,y),\ \overline{c}(x,y)$ as in (4.17) yields

$$\left({\nabla_x}^2 + {\nabla_y}^2\right) \overline{v}(x,y) = \overline{c}(x,y) + w(x,y) \tag{4.21}$$

with $w(x,y)$ such that

$$\nabla_x \nabla_y w(x,y) = 0. \tag{4.22}$$

As shown in the Appendix, Equation (4.22) implies that we can write $w(x,y)$ in the form

$$w(x,y) = w_x(x) - w_y(y). \tag{4.23}$$

Substituting $v(x,y)$ in (4.7) with $\nabla_x \nabla_y \overline{v}(x,y)$, we obtain

$$\begin{bmatrix} \psi_x(x,y) \\ \psi_y(x,y) \end{bmatrix} = \begin{bmatrix} \nabla_x \\ \nabla_y \end{bmatrix} \phi(x,y) + \begin{bmatrix} \nabla_x {\nabla_y}^2 \overline{v}(x,y) \\ \nabla_y {\nabla_x}^2 (-\overline{v}(x,y)) \end{bmatrix}. \tag{4.24}$$

Combine (4.24) with (4.21) and we can rewrite it as

$$\begin{bmatrix} \psi_x(x,y) \\ \psi_y(x,y) \end{bmatrix} = \begin{bmatrix} \nabla_x \\ \nabla_y \end{bmatrix} \left(\phi(x,y) + {\nabla_y}^2 \overline{v}(x,y) + w_y(x,y)\right) - \begin{bmatrix} 0 \\ \nabla_y \end{bmatrix} \overline{c}(x,y). \tag{4.25}$$

Call $\phi_0(x,y) = \phi(x,y) + {\nabla_y}^2 \overline{v}(x,y) + w_y(y)$ and the lemma is proven. This relationship shows that $\phi_0(x,y)$ contains information from the scalar and vector potentials. The model in (4.14) follows immediately by adding the column vector $[\nabla_x, \nabla_y]^T \overline{c}(x,y)$ to both sides of (4.13) which yields

$$\phi_1(x,y) = \phi_0(x,y) - \overline{c}(x,y) \tag{4.26}$$

and proves the result. ■

The significance of this result is that, in certain cases, as addressed in the next section, the term $\phi_0(x,y)$ is the total phase and it can be computed from equation (4.13) as

$$\begin{bmatrix} \psi_x(x,y) \\ \psi_y(x,y) \end{bmatrix} + \begin{bmatrix} 0 \\ \nabla_y \overline{c}(x,y) \end{bmatrix} = \begin{bmatrix} \nabla_x \\ \nabla_y \end{bmatrix} \phi_0(x,y) \tag{4.27}$$

using standard techniques. From this result, a rotational field in $\psi(x,y)$ can be made irrotational by combining it with its own curl. In this case, $\phi_0(x,y)$, $\phi_1(x,y)$ or any combination thereof, becomes a possible scalar potential function. In the example below, where phase wrapping causes the phase gradient to become rotational, it is shown that the scalar potential $\phi_0(x,y)$ coincides with the actual phase $\theta(x,y)$, which is sensed by the wrapped gradients.

Based on the fact that any signal is equivalent to its own convolution with the impulse, as $c(x,y) = c(x,y) ** \delta(x)\delta(y)$, we can write

$$\begin{aligned} \overline{c}(x,y) &= c(x,y) ** u(x)u(y), \\ \nabla_x \overline{c}(x,y) &= c(x,y) * u(y), \\ \nabla_y \overline{c}(x,y) &= c(x,y) * u(x), \end{aligned} \tag{4.28}$$

with $u(\cdot)$ the unit step function and the "star" operations indicating 2D convolution with $u(x)u(y)$ and 1D convolutions with $u(x)$ and $u(y)$ respectively. For clarity, we note that the first line of (4.28) is equivalent to (4.18).

The following example illustrates the results in the Lemma presented above.

Example. Consider the phase

$$\theta(x,y) = \text{phase}(x+jy), \tag{4.29}$$

with the phase wrapped to the interval $[0, 2\pi)$. Simple differentiation yields the *wrapped* gradient of θ and its Fourier transform, as computed in [55]

$$\begin{aligned} \psi_x(x,y) = -\frac{y}{x^2+y^2} \quad &\Leftrightarrow \quad \Psi_x(\kappa_x,\kappa_y) = \; j\frac{\kappa_y}{{\kappa_x}^2+{\kappa_y}^2}, \\ \psi_y(x,y) = \; \frac{x}{x^2+y^2} \quad &\Leftrightarrow \quad \Psi_y(\kappa_x,\kappa_y) = -j\frac{\kappa_x}{{\kappa_x}^2+{\kappa_y}^2}. \end{aligned} \tag{4.30}$$

The actual *unwrapped* gradient of $\theta(x,y)$ has to take the discontinuity at $y=0$, $x>0$ into account as

$$\begin{bmatrix} \nabla_x \\ \nabla_y \end{bmatrix} \theta(x,y) = \begin{bmatrix} \psi_x(x,y) \\ \psi_y(x,y) \end{bmatrix} + \begin{bmatrix} 0 \\ 2\pi\delta(y)u(x) \end{bmatrix}. \tag{4.31}$$

From substituting (4.30) into (4.9), we can easily solve for the scalar potential $\phi(\boldsymbol{\kappa})$, vector potential $V(\boldsymbol{\kappa})$ and the curl $C(\boldsymbol{\kappa})$ from (4.16) as

$$\begin{aligned} \phi(\boldsymbol{\kappa}) &= 0, \\ V(\boldsymbol{\kappa}) &= \frac{1}{2\pi({\kappa_x}^2+{\kappa_y}^2)}, \\ C(\boldsymbol{\kappa}) &= -2\pi. \end{aligned} \tag{4.32}$$

Now we can verify that equation (4.13) holds. We can substitute from (4.17) to solve

$$\nabla_x {\nabla_y}^2 \overline{v}(x,y) = \nabla_y v(x,y) = \text{IFT}\left\{ \frac{j\kappa_y}{{\kappa_x}^2+{\kappa_y}^2} \right\}. \tag{4.33}$$

From (4.30), the right-hand side of the above equation is $\nabla_x \theta(x,y)$, and therefore

$${\nabla_y}^2 \overline{v}(x,y) = \theta(x,y) + w_y(y) \tag{4.34}$$

with $w_y(y)$ accounting for boundary conditions. Substitution into (4.19) and using the fact

that the scalar potential $\phi(x,y)$ is zero, we obtain

$$\phi_0(x,y) = \theta(x,y) + w_y(y). \tag{4.35}$$

The quantity $\phi_0(x,y)$ is the total phase we want to reconstruct, on the right-hand side of (4.27). Substituting for $\bar{c}$ on the left-hand side of (4.27), we obtain

$$\begin{bmatrix} \psi_x(x,y) \\ \psi_y(x,y) \end{bmatrix} + \begin{bmatrix} 0 \\ 2\pi u(x)\delta(y) \end{bmatrix} = \begin{bmatrix} \nabla_x \\ \nabla_y \end{bmatrix} \phi_0(x,y). \tag{4.36}$$

The left-hand side is $\nabla\theta(x,y)$ from (4.31), which implies

$$\theta(x,y) = \phi_0(x,y) + C \tag{4.37}$$

where C is a constant. Equation (4.37) is consistent with (4.35) and $w_y(y) = C$.

In the next section, where we address the sampled data implementation, we actually prove analytically that $\phi_0[n_1,n_2]$ and the phase sensed by the wrapped gradients differ by integer multiples of 2π, thus yielding the same wrapped values. In other words the "hidden phase" is included in ϕ_0, and therefore no "slope discrepancy" is in the gradients of ϕ_0.

4.4 Fried Geometry

In the sampled data case, we extend the concepts introduced in the previous sections by defining the gradient operators on the basis of the Fried geometry.

To this extent, given the sampled phase $\phi_0[n_1,n_2] = \phi_0(n_1\Delta x, n_2\Delta y)$ we define the gradients in the two directions as

$$\begin{aligned} \psi_1[n_1,n_2] &= \frac{1}{2}\left(\phi_0[n_1+1,n_2+1] + \phi_0[n_1,n_2+1]\right) - \frac{1}{2}\left(\phi_0[n_1+1,n_2] + \phi_0[n_1,n_2]\right), \\ \psi_2[n_1,n_2] &= \frac{1}{2}\left(\phi_0[n_1+1,n_2+1] + \phi_0[n_1+1,n_2]\right) - \frac{1}{2}\left(\phi_0[n_1,n_2+1] + \phi_0[n_1,n_2]\right). \end{aligned} \tag{4.38}$$

Using the shift operators z_1, z_2, we can write (4.38) in a more compact form

$$\begin{aligned}\psi_1[n_1,n_2] &= \frac{1}{2}(z_1+1)(z_2-1)\,\phi_0[n_1,n_2],\\ \psi_2[n_1,n_2] &= \frac{1}{2}(z_1-1)(z_2+1)\,\phi_0[n_1,n_2]\end{aligned} \tag{4.39}$$

from which we define the discrete gradient operators as

$$\begin{aligned}\nabla_1(z_1,z_2) &\equiv \frac{1}{2}(z_1+1)(z_2-1),\\ \nabla_2(z_1,z_2) &\equiv \frac{1}{2}(z_1-1)(z_2+1).\end{aligned} \tag{4.40}$$

Substituting $z_1 = e^{j\omega_1}$ and $z_2 = e^{j\omega_2}$ into (4.40), we obtain the discrete frequency response of the operators as

$$\begin{aligned}\nabla_1(\boldsymbol{\omega}) &= 2e^{-j\frac{\omega_1+\omega_2}{2}}\cos\left(\frac{\omega_1}{2}\right)\sin\left(\frac{\omega_2}{2}\right),\\ \nabla_2(\boldsymbol{\omega}) &= 2e^{-j\frac{\omega_1+\omega_2}{2}}\sin\left(\frac{\omega_1}{2}\right)\cos\left(\frac{\omega_2}{2}\right).\end{aligned} \tag{4.41}$$

It is easy to see that both $\nabla_1(\boldsymbol{\omega})$ and $\nabla_2(\boldsymbol{\omega})$ are zero when $\boldsymbol{\omega} = [0,0]$ ("piston" mode) and $\boldsymbol{\omega} = [\pi,\pi]$ ("waffle" mode). As a consequence,

$$\nabla x[\mathbf{n}] = 0 \;\;\Rightarrow\;\; x[\mathbf{n}] = C_0 + C_1(-1)^{n_1+n_2} \tag{4.42}$$

for some constants C_0, C_1 that depend on the boundary conditions.

It is well known that Fried derivatives are good models for Shack-Hartmann sensors, which measure local phase gradients. However, Barchers demonstrated that the Fried geometry performance degrades in high scintillation when compared to Hudgin geometry [51]. When branch points causing phase wrapping are present, it is imperative to properly embed the wrapping operation within the Fried gradients computations. The development of the wrapped Fried gradient presented here is sufficient for reconstructing the high turbulence wavefront properly.

From Figure 4.1, the block diagram approach of the Fried geometry can be seen. In 4.1 (a), the standard Fried gradients are shown in terms of transfer functions in z_1 and z_2. The first blocks (z_1-1) and (z_2-1) provide for phase differences in the vertical and horizontal

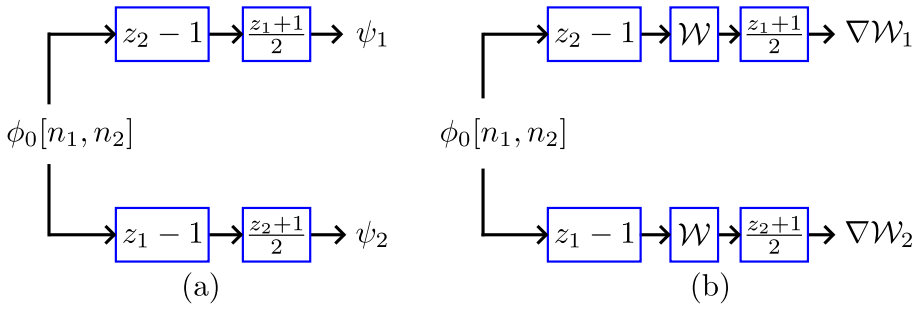


Figure 4.1. The block diagram form (a) the traditional Fried gradient and (b) the wrapped Fried gradient.

directions, while the second blocks $(z_2+1)/2$ and $(z_1+1)/2$ provides simple averaging (half the sum) in the opposite directions (horizontal and vertical).

When phase wrapping is present, the first blocks, which model phase difference measurements, must be augmented with the phase wrapping operation $\mathcal{W}$. Wrapping guarantees that phase differences multiples of 2π are not sensed.

Call $\nabla\mathcal{W}$ the wrapped Fried derivative gradient in Figure 4.1 (b) and $\boldsymbol{\psi}[\mathbf{n}]$, the sensed wrapped gradients, as

$$\boldsymbol{\psi}[n_1,n_2] = \begin{bmatrix} \nabla\mathcal{W}_1 \\ \nabla\mathcal{W}_2 \end{bmatrix} \boldsymbol{\phi}_0[n_1,n_2]. \tag{4.43}$$

From the definition of the wrapping operator $\mathcal{W}$ and the factor $1/2$ in the averaging second block we can relate the gradient and wrapped gradient by

$$\nabla\mathcal{W}\boldsymbol{\phi}_0[\mathbf{n}] = \nabla\boldsymbol{\phi}_0[\mathbf{n}] + \pi\ell[\mathbf{n}] \tag{4.44}$$

with $\ell[\mathbf{n}]$ having integer values only.

The wrapped Fried gradients are the basis of the phase estimation presented in the next section.

4.5 Least Squares Phase Estimation from Wrapped Fried Gradients

Along the same lines as in the previous section, the vector field ψ can be represented in terms of scalar and potential functions ϕ and v as

$$\begin{bmatrix} \psi_1[n_1,n_2] \\ \psi_2[n_1,n_2] \end{bmatrix} = \begin{bmatrix} \nabla_1 & \nabla_2 \\ \nabla_2 & -\nabla_1 \end{bmatrix} \begin{bmatrix} \phi[n_1,n_2] \\ v[n_1,n_2] \end{bmatrix}. \tag{4.45}$$

Equation (4.45) is the same as for the continuous case in the previous section, and in the frequency domain this just becomes a matrix-vector operation

$$\begin{bmatrix} \Psi_1(\boldsymbol{\omega}) \\ \Psi_2(\boldsymbol{\omega}) \end{bmatrix} = e^{-j\frac{\omega_1+\omega_2}{2}} \begin{bmatrix} c_1 s_2 & s_1 c_2 \\ s_1 c_2 & -c_1 s_2 \end{bmatrix} \begin{bmatrix} \Phi(\boldsymbol{\omega}) \\ V(\boldsymbol{\omega}) \end{bmatrix} \tag{4.46}$$

with $c_i = \cos(\omega_i/2)$ and $s_i = \sin(\omega_i/2)$ for $i = 1,2$. We can easily verify that the two columns of the matrix represent two orthogonal vectors.

The result of the previous section can then be extended to the sampled data case by defining the curl of the vector field as

$$c[n_1,n_2] \equiv \nabla_2 \psi_1[n_1,n_2] - \nabla_1 \psi_2[n_1,n_2]. \tag{4.47}$$

with ∇_1, ∇_2 the standard unwrapped Fried derivatives. Then the vector field ψ can be expressed as

$$\begin{bmatrix} \psi_1[n_1,n_2] \\ \psi_2[n_1,n_2] \end{bmatrix} = \begin{bmatrix} \nabla_1 & 0 \\ \nabla_2 & -\nabla_2 \end{bmatrix} \begin{bmatrix} \phi_0[n_1,n_2] \\ \overline{c}[n_1,n_2] \end{bmatrix} \tag{4.48}$$

with $\overline{c}$ defined as

$$c[n_1,n_2] \equiv \nabla_1 \nabla_2 \overline{c}[n_1,n_2]. \tag{4.49}$$

In order to obtain an expression for $\overline{c}$, first notice that any signal can be represented in convolution (double convolution in the 2D case) form

$$c[n_1,n_2] = c[n_1,n_2] ** \delta[n_1]\delta[n_2] \tag{4.50}$$

with $\delta[n]$ the discrete impulse, being $\delta[0]=1$ and $\delta[n]=0$ for all $n\neq 0$. In equation (4.49), the product of the two operators can be expressed as

$$\begin{aligned}\nabla_1(z)\nabla_2(z) &= \frac{1}{4}(z_1-1)(z_1+1)(z_2-1)(z_2+1)\\ &= \frac{1}{4}\left({z_1}^2-1\right)\left({z_2}^2-1\right).\end{aligned} \tag{4.51}$$

Now if we define the sequence

$$\overline{u}[n] \equiv (1+(-1)^n)\,u[n-2], \tag{4.52}$$

plotted in Figure 4.2, we can verify that $\overline{u}[n+2]-\overline{u}[n]=2\delta[n]$, and therefore

$$\delta[n_1]\delta[n_2] = \nabla_1(z)\nabla_2(z)\overline{u}[n_1]\overline{u}[n_2] \tag{4.53}$$

so that $\overline{c}$ can be written as

$$\overline{c}[n_1,n_2] = c[n_1,n_2] \ast\ast\, \overline{u}[n_1]\overline{u}[n_2]. \tag{4.54}$$

With these premises, we can state the main result of this research.

Main Result. Let $\psi[n_1,n_2]$ be the vector field of the wrapped Fried gradient of the phase $\phi_0[n_1,n_2]$, defined as

$$\begin{bmatrix}\psi_1[\mathbf{n}]\\ \psi_2[\mathbf{n}]\end{bmatrix} \equiv \begin{bmatrix}\nabla\mathcal{W}_1\\ \nabla\mathcal{W}_2\end{bmatrix}\phi_0[\mathbf{n}]. \tag{4.55}$$

Also let its curl, $c[\mathbf{n}]$, be such that

$$c[\mathbf{n}] = \pi\ell[\mathbf{n}] \tag{4.56}$$

i.e., it assumes values only integer multiples of π.

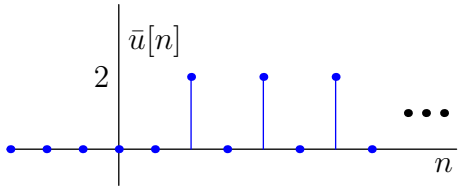


Figure 4.2. The function $\bar{u}[\mathbf{n}]$ is used to create $\bar{c}[\mathbf{n}]$ from $c[\mathbf{n}]$.

Let $\hat{\phi}[\mathbf{n}]$ be such that

$$\begin{bmatrix} \psi_1[\mathbf{n}] \\ \psi_2[\mathbf{n}] \end{bmatrix} + \begin{bmatrix} 0 \\ c[\mathbf{n}] \ast\ast q[\mathbf{n}] \end{bmatrix} = \begin{bmatrix} \nabla_1 \\ \nabla_2 \end{bmatrix} \hat{\phi}[\mathbf{n}] \tag{4.57}$$

with $\ast\ast$ denoting 2D convolution and $q[\mathbf{n}]$ defined as

$$q[n_1, n_2] \equiv \frac{1}{2}(z_1 - 1)(z_2 + 1)\bar{u}[n_1, n_2]. \tag{4.58}$$

Then there exist constants C_0, C_1 for which

$$\phi_0[\mathbf{n}] = \hat{\phi}[\mathbf{n}] + C_0 + C_1(-1)^{n_1+n_2} + 2\pi\ell[\mathbf{n}] \tag{4.59}$$

with the rightmost term a sequence assuming only integer multiples of 2π.

Proof. By exactly the same arguments as in the previous section, equation (4.57) holds for some $\hat{\phi}[\mathbf{n}]$. What we need to show is that the estimated phase $\hat{\phi}[\mathbf{n}]$ and the original phase $\phi_0[\mathbf{n}]$ are the same apart from integer multiples of 2π, a piston mode C_0 and a "waffle" mode $C_1(-1)^{n_1+n_2}$.

The argument is based on the fact that the curl sequence $c[n_1, n_2]$ assumes values which are all integer multiples of π. Also it is a simple exercise to verify that the sequence $q[\mathbf{n}] = 0, \pm 2$ for all n. As a consequence, for all $\mathbf{n}$,

$$c[\mathbf{n}] \ast\ast q[\mathbf{n}] = 2\pi\ell[\mathbf{n}]. \tag{4.60}$$

where, again $\ell[\mathbf{n}]$ denotes a sequence of integer values. Recall that the relation between Fried and wrapped Fried gradients in (4.44) and the observed wrapped phase gradient $\psi[\mathbf{n}]$. Then by substitution into (4.57) we obtain that the sequence

$$\begin{bmatrix} \nabla_1 \\ \nabla_2 \end{bmatrix} \left(\phi_0[\mathbf{n}] - \hat{\phi}[\mathbf{n}]\right) = \pi \begin{bmatrix} \ell_1[\mathbf{n}] \\ \ell_2[\mathbf{n}] \end{bmatrix}, \tag{4.61}$$

i.e., it assumes values which are integer multiples of π for all $\mathbf{n}$. Since for any sequence

$$\ell[\mathbf{n}] = \ell[\mathbf{n}] ** \delta[n_1]\delta[n_2] \tag{4.62}$$

and

$$\begin{aligned} \delta[n_1]\delta[n_2] &= \nabla_1 \left(2(-1)^{n_1-1}u[n_1-1]u[n_2-1]\right) \\ &= \nabla_2 \left(2(-1)^{n_2-1}u[n_1-1]u[n_2-1]\right), \end{aligned} \tag{4.63}$$

we have that

$$\begin{aligned} \ell[\mathbf{n}] &= 2\nabla_1\ell[\mathbf{n}] \\ \ell[\mathbf{n}] &= 2\nabla_2\ell[\mathbf{n}]. \end{aligned} \tag{4.64}$$

In other words, a sequence of integers is the Fried derivative of a sequence of even integers. Finally combine equations (4.61) and (4.64) to obtain

$$\nabla\left(\phi_0[\mathbf{n}] - \hat{\phi}[\mathbf{n}] + 2\pi\ell[\mathbf{n}]\right) = 0, \tag{4.65}$$

which proves the result. ■

Estimation of $\phi_0[\mathbf{n}]$ based on sensed Fried gradients $\psi[\mathbf{n}]$ is then carried out by computing $\hat{\phi}[\mathbf{n}]$ from equation (4.57), using either standard least squares or (as is shown in the next section) the multi-resolution algorithm presented by the authors in [61].

Then, from the result in equation (4.59), we obtain

$$\phi_0[\mathbf{n}] = \mathcal{W}\left(\hat{\phi}[\mathbf{n}] + C_0\right) \tag{4.66}$$

with C_0 a constant determined by a reference value. The "waffle" term is usually neglected since the data is assumed not to contain this term.

The algorithm for Fried geometry can be summarized as a procedural list:

1. Collect the sensor measurements $\psi_1[n_1,n_2]$ and $\psi_2[n_1,n_2]$.
2. Compute the curl
$$c[n_1,n_2] = \nabla_2\psi_1[n_1,n_2] - \nabla_1\psi_2[n_1,n_2]. \tag{4.67}$$

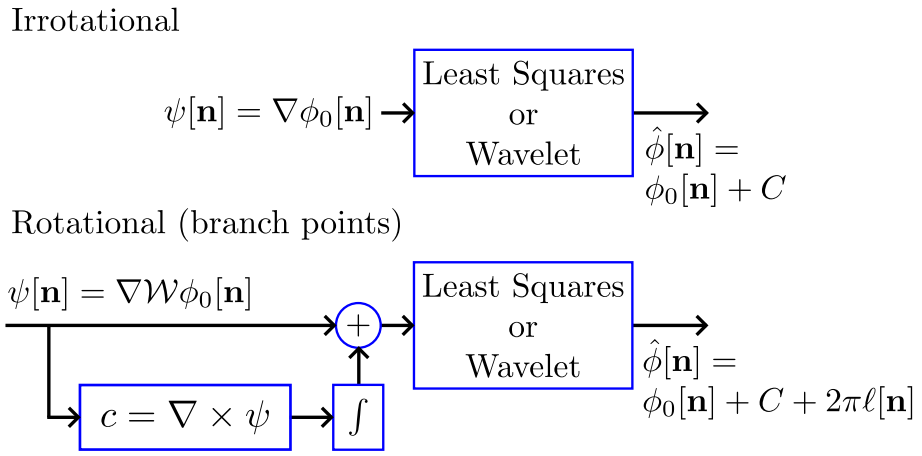


Figure 4.3. The block diagram compares the traditional least-squares approach to the new form that is capable of handling branch points. When the curl is equal to zero, the output is exactly the same for both forms.

3. Compute the quantity

$$\nabla_2\bar{c}[n_1,n_2] = c[n_1,n_2] ** q[n_1,n_2]. \tag{4.68}$$

4. Modify the measurement

$$\psi_{2,new}[n_1,n_2] = \psi_2[n_1,n_2] + \nabla_2\bar{c}[n_1,n_2]. \tag{4.69}$$

5. Use $\psi_1[n_1,n_2]$ and $\psi_{2,new}[n_1,n_2]$ in the standard least-squares reconstructor to solve for $\hat{\phi}[n_1,n_2]$

$$\begin{bmatrix} \psi_1[n_1,n_2] \\ \psi_{2,new}[n_1,n_2] \end{bmatrix} = \begin{bmatrix} \nabla_1 \\ \nabla_2 \end{bmatrix} \hat{\phi}[n_1,n_2]. \tag{4.70}$$

The comparison of this algorithm with the traditional approach is given in Figure 4.3.

4.6 Application to Phase Estimation

The algorithm presented in Section 4.5 has been applied to a number of phase signals both geometric and simulated wavefront phase functions.

In the following examples when noise is present, the Gaussian white noise is added to the

phase difference measurement quantities as

$$\begin{bmatrix} \psi_{x,noisy}[n_1,n_2] \\ \psi_{y,noisy}[n_1,n_2] \end{bmatrix} = \begin{bmatrix} \psi_x[n_1,n_2] \\ \psi_y[n_1,n_2] \end{bmatrix} + \begin{bmatrix} \alpha n_x[n_1,n_2] \\ \alpha n_y[n_1,n_2] \end{bmatrix}, \tag{4.71}$$

where α is chosen to ensure the desired SNR for simulation. The noise source models the uncertainty in the centroid operation of the S-H WFS. Unless stated otherwise, amplitude is not used in the reconstruction and no noise added to the amplitude.

In addition, we provide a comparison with the proprietary SPhase algorithm in AOTools and WaveProp, developed by the Optical Sciences Company [62], [63]. SPhase uses amplitude and phase information for wavefront reconstruction and its goal is to place the branch cuts in areas of low irradiance for a continuous DM. SPhase also performs phase unwrapping. Thus, the goals of SPhase are different than the algorithm presented here.

4.6.1 Example 1: Geometric Signal

Let $s = x + jy$ and define $\phi_0[\mathbf{n}]$ as samples of the phase

$$\phi(x,y) = \text{phase}(s) \tag{4.72}$$

with sampling intervals $\delta_x = \delta_y = 0.01$, the number of data points $N = 256 \times 256$ and a shift by $\delta_x/2$ and $\delta_y/2$ so that the singular point $x = y = 0$ is not in the sampling grid.

In Figure 4.4, the 3D plot of the wrapped phase $\mathcal{W}\phi_0[\mathbf{n}]$ is shown and in Figure 4.5 the wrapped estimated phase $\mathcal{W}\hat{\phi}[\mathbf{n}]$ is shown. The reconstruction is an exact match. The significance of this is observable from Figures 4.6 and 4.7, where the large discontinuity is not apparent. The lack of discontinuity in the measurements is the importance of the wrapped Fried gradient model, otherwise the ridge would be in the gradient data that is the input to the algorithm.

In this particular example, because the discontinuity is along the same dimension as our non-orthogonal correction, the result is exactly the same as the input. If the input had a discontinuity at a different angle relative to the origin (which would still result in the same wrapped measurements), the resulting output would still be the same as the one shown in Figure 4.5.

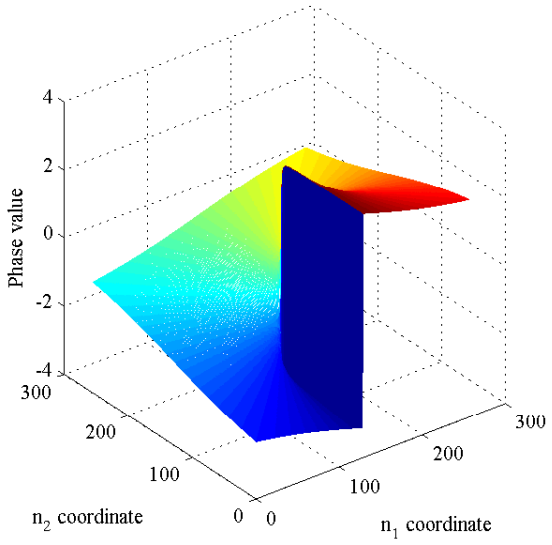


Figure 4.4. The original $\phi[\mathbf{n}]$ phase data for example 1 is shown.

4.6.2 Example 2: Geometric Signal

Similarly, in Figure 4.8 the samples of the phase for the function

$$\phi(x,y) = \text{phase}\left(\frac{(s-b_1)(s-b_2)}{(s-a)}\right), \tag{4.73}$$

with $b_1 = 0.5150 - j0.26$, $b_2 = 0.0050 + j0.26$ and $a = 0.005 + j0.005$ are shown. Two estimates are shown: without noise in Figure 4.9 and with noise added to the observations (with 40dB SNR) in Figure 4.10.

The comparison between Figures 4.8 and 4.9 along the top center shows two different boundaries of maximum (red) and minimum phase (blue). In this example, we show that C_0 from (4.66) is set to a constant that causes a slightly different wrapping than Figure 4.8. The gradient measurements are the same and we show that the discontinuity can be positioned.

With this example, we are able to know the amplitude and phase. If we run SPhase with the phase, but set the uniform amplitude to be unity, SPhase chooses a simple branch cut scheme of connecting the two closest branch points to one another, and the third (closer to the bottom) branch point straight to the bottom edge. Our algorithm connects the lower

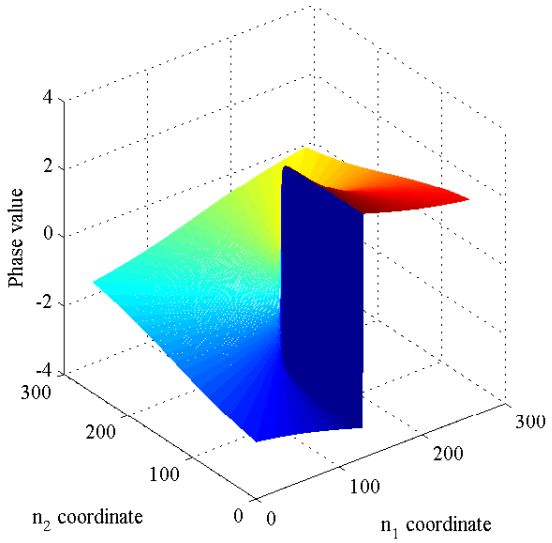


Figure 4.5. The reconstructed $\mathcal{W}\hat{\phi}[\mathbf{n}]$ phase data for example 1 is shown.

branch point to the right edge (due to the non-orthogonal basis). Wrapping the phase for either result has the same output.

Using both the amplitude and phase information from (4.73), SPhase creates a slightly more complicated branch cut between the upper two branch points that takes advantage of the lower irradiance path between these singularities.

4.6.3 Example 3: High Turbulence Phase Signal

WaveProp was used to generate the algorithm input data for this example. We tried the algorithm under a variety of operation conditions, but only present the highest turbulence results here as other cases also were successful. WaveProp simulated a 1.0 meter diameter circular aperture in a 2048×2048 E-field grid. The simulation used $\lambda = 1\mu$m through a 4 km horizontal path. The atmospheric effects were assumed to be a constant turbulence through 5 phase screens. The C_n^2 value is 7×10^{-15} with a calculated Rytov number of 0.3051.

The phase signal is shown in Figure 4.11, with estimates in Figure 4.12 (no noise) and Figure 4.13 (noise with 40dB SNR). For the noiseless case, the location of the detected

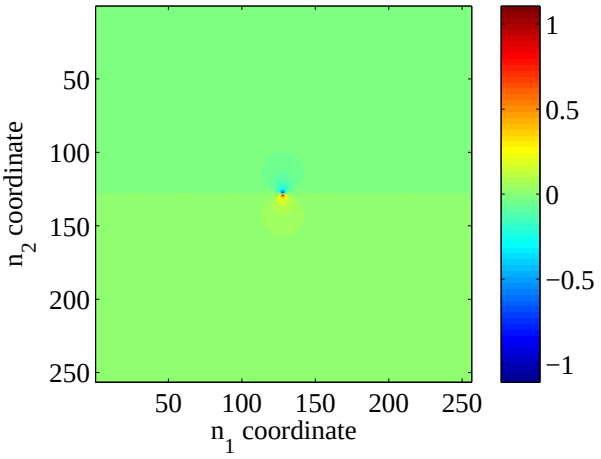


Figure 4.6. The wrapped gradient $\psi_1[\mathbf{n}]$ data for example 1 is shown. The large discontinuity seen in Figure 4.4 is not apparent in the wrapped measured gradient.

branch points are shown in Figure 4.15, while the branch points of the original phase as determined by WaveProp are shown in Figure 4.14.

SPhase only works on this example when the correct (original) amplitude is also supplied to its input. Setting a constant amplitude results in a signal of little interest (even the wrapped output did not match the original data). The wrapped output of SPhase (using the WaveProp amplitude) is identical to the output of our algorithm. Thus, we can say that the amplitude information is important in the SPhase algorithm, whereas the amplitude is not used by our algorithm proposed here.

4.6.4 Example 4: Double Spiral

Our last example is the double spiral shear from [54]. Although this dataset, shown in Figure 4.16, is used to test unwrapping, we decided to include it here. Ghiglia states that this example has failed in unwrapping when there is noise on the measurements for all unwrapping algorithms covered by their book. The actual spiral data has one arm ascending (with, a positive n_1 gradient) and the other spiral arm descending with a negative n_1 gradient of the same magnitude.

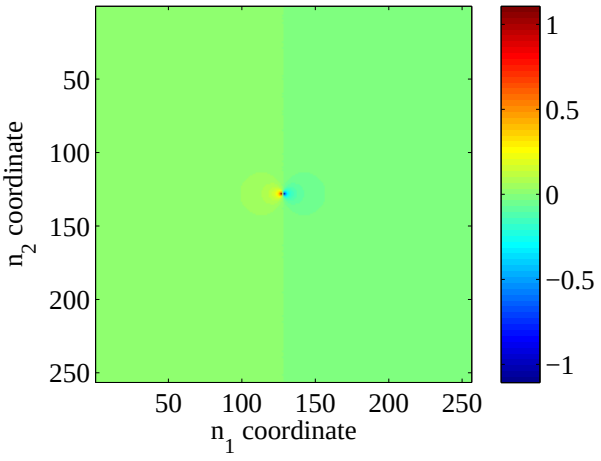


Figure 4.7. The wrapped gradient $\psi_2[\mathbf{n}]$ data for example 1 is shown. The large discontinuity seen in Figure 4.4 is not apparent in the wrapped measured gradient. The correction term proposed in this dissertation will be added to $\psi_2[\mathbf{n}]$ to create the large discontinuity.

Our algorithm results in Figure 4.17 for no noise, and in Figure 4.18 for 40 dB SNR. In the case of noise, the noise can potentially cause the phase value to wrap and the horizontal bar pattern can form. However, in the no noise case, the reconstruction is exact. The determined branch point locations match Figure 3.10 in [54].

Since SPhase also includes unwrapping, it has difficulty on this data set. While its output does show the double spiral pattern, the spiral arms are flat areas. The boundary pixels between the spiral arms often do not fully resolve correctly and have discontinuities. The wrapped output of SPhase is not a good match to the original surface. One spiral arm takes on zero value for all pixels, and the other spiral arm has areas that are close to $\pm\pi$. The boundary pixels of the spirals in the wrapped output also have discontinuities.

4.7 Summary

In this research, we addressed the problem of estimating a phase signal based on observation of wrapped local variations. This approach is based on a particular representation of the vector field in terms of a non-orthogonal basis which seems to be better suited than the

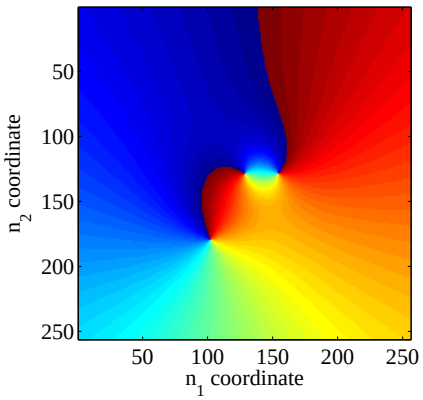


Figure 4.8. The original $\phi[\mathbf{n}]$ plotted for example 2 is shown.

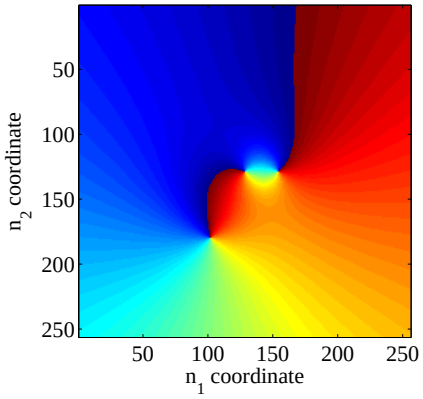


Figure 4.9. The estimated phase $\mathcal{W}\hat{\phi}[\mathbf{n}]$ plotted for example 2.

standard orthogonal basis associated to scalar and potential field.

It was shown that by correcting the observed gradient with a filtered curl, the overall phase (including what has been called the "hidden phase") is estimated by standard least-squares solver. A number of computer simulations support what has been stated based on mathematical analysis. A comparison with SPhase shows that our algorithm results in the same wrapped phase measurements, which is expected since the algorithms output different phase functions of the ensemble of wavefront surfaces that have the same gradient measurements. The examples show that the wrapped ϕ_0 is equal to the wrapped total phase.

This approach is able to efficiently determine a wavefront surface that is a member of the

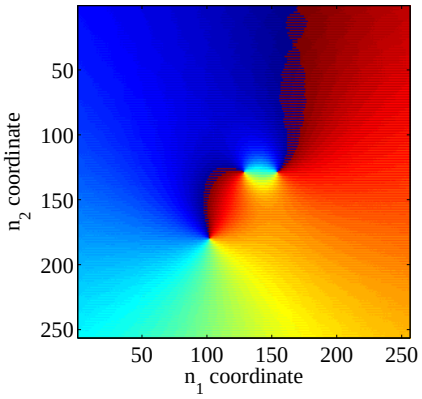


Figure 4.10. The estimated phase $\mathcal{W}\hat{\phi}[\mathbf{n}]$ with 40 dB SNR for example 2 is shown. Pixels with values close to $-\pi$ or π may wrap due to the noise.

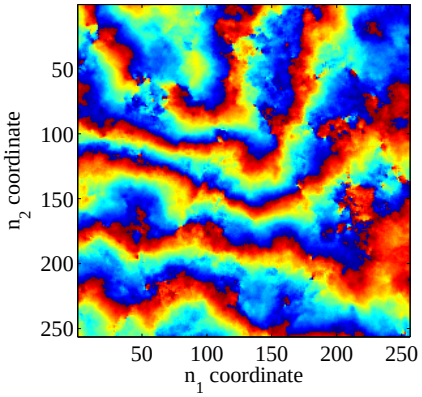


Figure 4.11. A high turbulence wavefront phase $\phi[\mathbf{n}]$ created using WaveProp for example 3.

ensemble of wavefronts that all have the same gradient measurements. The approach is as computationally efficient as the least-squares or equivalent reconstructor chosen. The approach does not unwrap the phase, as we leave that as a follow on step to the output of our algorithm presented here.

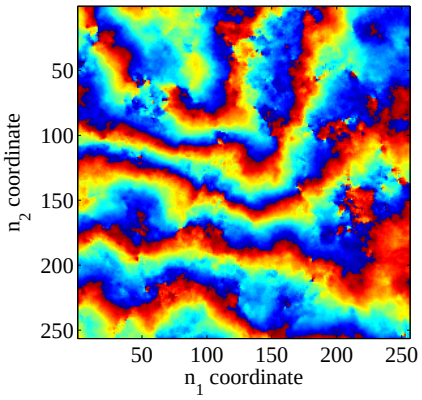


Figure 4.12. The estimated wavefront $\mathcal{W}\hat{\phi}[\mathbf{n}]$ is reconstructed for example 3.

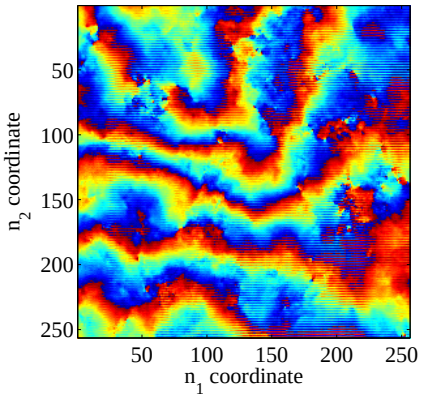


Figure 4.13. The estimated wavefront $\mathcal{W}\hat{\phi}[\mathbf{n}]$ is reconstructed with 40 dB SNR.

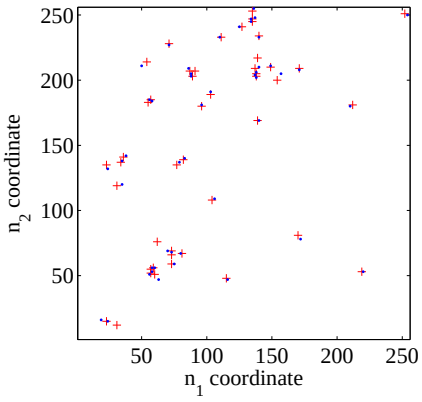


Figure 4.14. The known branch points for example 3 with no noise are shown. Locations were determined by WaveProp. Positive branch points are indicated with a red plus while negative branch points are indicated with a blue dot.

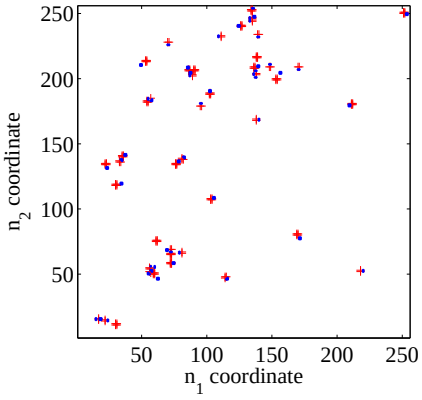


Figure 4.15. The detected branch points for example 3 with no noise are shown. Positive branch points are indicated with a red plus while negative branch points are indicated with a blue dot. Because the Fried geometry averages neighboring values, the locations on this plot are quadrupled compared to Figure 4.14.

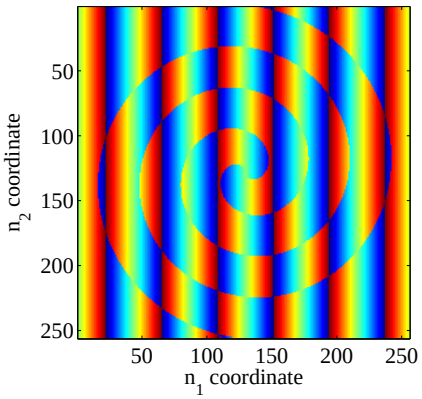


Figure 4.16. The spiral dataset $\phi[\mathbf{n}]$ from [54]. This dataset is known to be difficult to process correctly.

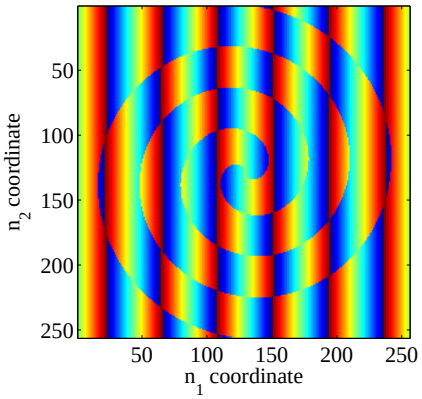


Figure 4.17. The estimated phase $\mathcal{W}\hat{\phi}[\mathbf{n}]$ reconstructed for the spiral dataset.

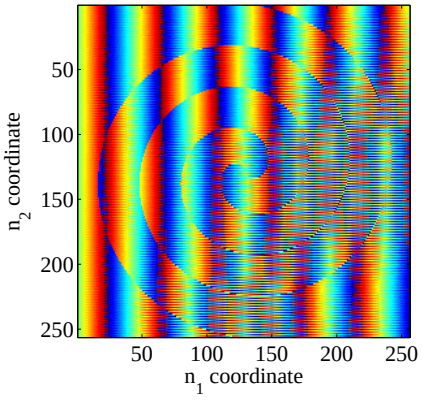


Figure 4.18. The estimated phase $\mathcal{W}\hat{\phi}[\mathbf{n}]$ reconstructed for spiral dataset with 40 dB SNR.

CHAPTER 5:
Mirror Surface Control Using Optimization

As mentioned in the introductory chapter, the goal of the AO system is to compensate for phase distortion. To achieve this goal, the phase distortion determined by the wavefront reconstruction, presented in the previous two chapters, has to be properly compensated. In this chapter we address the problem of setting the control actuators of a DM to compensate for the phase distortions detected by the wavefront reconstructor.

The phase distortion detected by the phase reconstruction algorithms presented in the previous two chapters, is at the basis of using mathematical optimization for mirror surface control that was conducted in the SMT laboratory. Although the original mirror control software determines actuator settings by using mathematical optimization routines, these routines were never validated on the hardware. The DM behavior has nonlinear characteristics; however, the original optimal control problem uses a linear model approximation of the actual hardware performance. We sought to determine whether control using the linear model was valid, the range of operation where the linear assumption is true, and what the resulting performance levels were as measured in root-mean-square (rms) error of the wavefront surface.

SMT is an active optics system with surface parallel actuators. Applying a voltage across the actuator causes the mirror surface to change. The goal of this technology is to allow larger variances in manufacturing tolerances. Deviations from the intended optical prescription are removed by the actuators during operation. This design saves money by reducing costly manufacturing rework.

The primary mirror of the SMT has six segments that are hexagonal-shaped mirrors, each having 156 controllable actuators. Although in the previous chapters a S-H WFS was used to estimate the wavefront, the SMT primary mirror is measured using an interferometer sensor placed in front of the telescope as shown in Figure 5.1. The sensor choice allows for high resolution sampling of the mirror surface that would not be available with the laboratory S-H WFS. A Stewart platform is used to position the interferometer along the optical axis of the primary mirror. The interferometer and null corrector are mounted to remove the

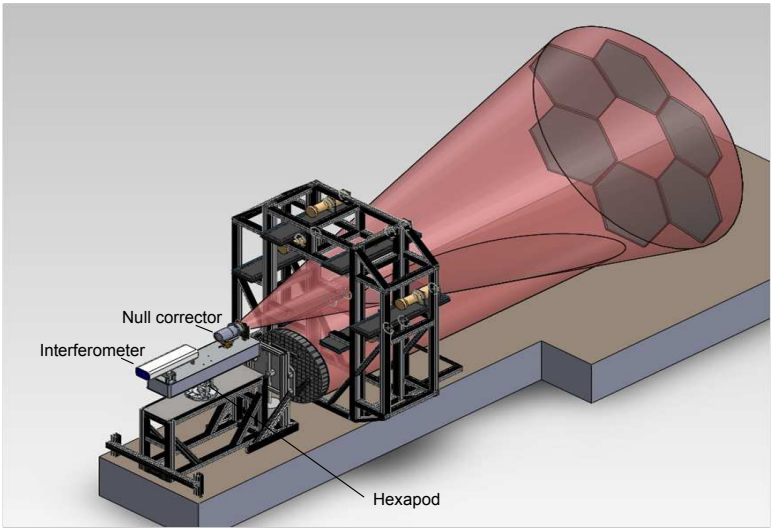


Figure 5.1. The hardware arrangement in the laboratory that was used for influence functions and measuring surface curvature. Image is courtesy of John Bagnasco.

spherical aberration at the primary mirror center of curvature. Each interferometer sample measures the mirror surface height in units of wavelength.

In the next section, we describe the original optimization problem developed by the SMT manufacturers to control the mirror surface. In Section 5.2, we explain a hardware test algorithm which solves a series of optimization problems and each result can be used to control the mirror surface. In Section 5.3, we modify the algorithm to use a multigrid approach for the optimization problems and reduce computation time. We summarize the chapter in Section 5.5.

5.1 Original Optimization Problem

The original SMT control software solves a constrained optimization problem for each primary mirror segment to determine the actuator voltages. The constrained optimization problem is

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & J = \tfrac{1}{2}\,\|\mathbf{C}\mathbf{x}-\mathbf{d}\|_2^2 \\ \text{subject to} \quad & \mathbf{A}\mathbf{x} \leq \mathbf{b} \end{aligned} \tag{5.1}$$

where the vector $\mathbf{x}$ contains the actuator voltages that minimize the cost function J, $\mathbf{C}$ is the linear influence function matrix, and the $\mathbf{d}$ vector contains the desired mirror shape. The matrix $\mathbf{A}$ and vector $\mathbf{b}$ model the hardware voltage limits.

The cost function is minimized when the actuators change the mirror surface to match the desired mirror surface shape. When the mirror forms the conjugate shape of the wavefront distortion as measured by the interferometer, the wavefront rms error decreases. The notation $\|\cdot\|_2$ in the cost function identifies the L_2 norm [21], or the Euclidean magnitude of a vector.

The vector $\mathbf{x}$ consists of the $M = 156$ actuator voltage biases from the nominal operating voltages. After solving the optimization problem, the software commands the actuator voltages.

The linear influence matrix $\mathbf{C}$ has as many rows as the number of measured samples, P, and as many columns as the number of actuators, M. The influence function for actuator m is defined as

$$f_m[n_1, n_2] = \frac{\mathrm{OPD}_m(V_1) - \mathrm{OPD}_m(V_2)}{V_1 - V_2} \tag{5.2}$$

where $\mathrm{OPD}_m(V)$ is a discrete 2D optical path difference (OPD) for the applied voltage V to actuator m. Each OPD is calculated from multiple interferograms collected by an interferometer. Forming each influence function requires 2 measurements, one with positive actuator voltage bias V_1 and the other with a negative voltage bias V_2. Both V_1 and V_2 are chosen to have the same distance from the nominal voltage and represent the range over which the actuators are assumed to have linear operating characteristics. Further details on measuring influence functions are found in [6] and simulated influence functions can be created by using integrated optomechanical analysis [64].

Although the interferometer collects approximately $1,000 \times 1,000$ samples per OPD, a segment only covers a fraction of the area. Each segment has a *mask* that identifies which samples display the segment surface. As a result of the optical configuration, a particular segment has $P \approx 60,000$ samples that overlap the measurement. Each masked $f_m[n_1, n_2]$ is formed into a column vector

$$\mathbf{c}_m = \mathrm{vector}(\mathrm{masked}(f_m)) \tag{5.3}$$

where $\mathrm{masked}(\cdot)$ keeps the samples of the measured mirror surface and discards all others.

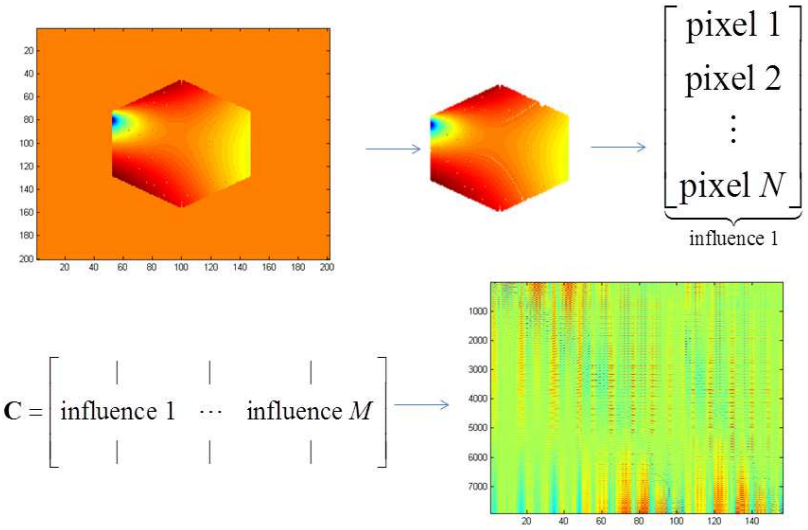


Figure 5.2. The **C** matrix processing is shown for a single simulated influence function.

After all of the influence functions are collected the **C** matrix is defined as

$$\mathbf{C} = \begin{bmatrix} | & | & | & | \\ \mathbf{c}_1 & \mathbf{c}_2 & \cdots & \mathbf{c}_M \\ | & | & | & | \end{bmatrix}. \tag{5.4}$$

The process of creating **C** is shown in Figure 5.2 using simulated data with approximately 8,000 samples on the surface.

The vector **d** is the desired wavefront as measured by the interferometer and is of the same dimensions as a single column of **C**. Immediately before running the optimization, these data are collected. The optical system has alignment issues that cause large piston, tip and tilt aberration. These modes are removed from the data, as they are not of interest to correct in this study.

Table 5.1. The dimensions of the quantities used in the optimization problem where $M = 156$ and $P \approx 60,000$.

Variable	Dimensions
$\mathbf{A}$	$2M \times M$
$\mathbf{b}$	$2M \times 1$
$\mathbf{C}$	$P \times M$
$\mathbf{d}$	$P \times 1$
$\mathbf{x}$	$M \times 1$
lower	$M \times 1$
upper	$M \times 1$

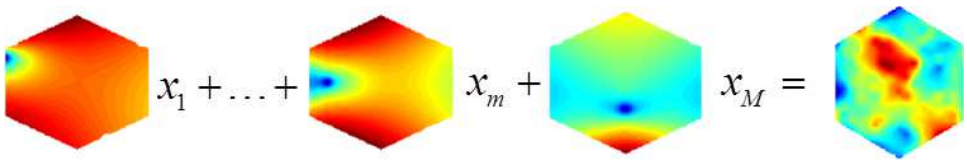


Figure 5.3. The optimization problem combines the SMT influence functions to create the desired mirror surface shape by treating each x_i as a scaling factor for the influence function.

The constraints $\mathbf{Ax} \leq \mathbf{b}$ can be used to form a convex hull [65] using $\mathbf{A}$ and $\mathbf{b}$ of the form

$$\mathbf{A} = \begin{bmatrix} \mathbf{I} \\ -\mathbf{I} \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} \mathbf{upper} \\ \mathbf{lower} \end{bmatrix}. \tag{5.5}$$

The set of $\mathbf{x}$ values that satisfy the constraints is called the *feasible set* and the optimization solution must be contained in this set. The dimension of $\mathbf{A}$ is $2M \times M$ and dimensionality of $\mathbf{b}$ is $2M \times 1$, resulting in twice as many constraints as actuators. We use **lower** and **upper** to signify bounds on each actuator. The implementation used by the SMT developers was a constant vector that constrained each actuator to the same range used to calculate influence functions in (5.2).

Table 5.1 summarizes the dimensions of the optimization problem. Since $M \ll P$, the optimization problem is a vastly overdetermined linear system of equations. As shown in Figure 5.3, each actuator is commanded to achieve the desired surface using the obtained optimal solution $\mathbf{x}^*$.

To solve the optimization problem of (5.1), the control software uses the `LSQLIN` optimizer

function in the MATLAB Optimization toolkit. The algorithm details are documented in [66], [67]; however, one significant limitation of the optimizer function is that it only returns the final solution and the sub-optimal solutions of the trajectory are not available by design. For the general use-case scenario of `LSQLIN`, intermediate sub-optimal solutions are of little interest. Cohan and Miller state that control using an optimization problem solution of a modeled segmented mirror matches to within 7% of NASTRAN finite element model prediction [68]; however, we sought to test this on hardware and to determine if the hardware response is similar to the linear influence function approximation. To do so, we need more than just the final optimization solution.

The linear influence function model assumes that the actuators operate independently and their combination follows the principle of superposition. In the actual hardware, we expect that the system is not truly linear and that actuators have nonlinear coupling. With only a single optimization solution to compare against the hardware, we cannot collect significant data to make a determination as to whether the linear system model accurately represents the hardware characteristics.

In order to produce more than one solution to test, we must create a *trajectory*, or sequence of solutions. In the next section, we present our developed technique to create a trajectory of actuator voltages to command the DM. We can determine the linear region from the trajectory.

5.2 Trajectory Creation Algorithm

We developed an algorithm to generate a trajectory from the solutions of a sequence of optimization problems with varying constraints. The solutions can be used to compare the linear influence function model against the hardware performance. We wanted to generate a series of solutions with a "small" step from the previous iteration along the trajectory to the final result $\mathbf{x}^*$. We define "small" steps as a combination of two constraints: a norm constraint and a moving boundary constraint. First, the L_2 norm of the actuator change in value is forced to be less than or equal to a chosen constant. Second, each actuator movement is individually constrained from the previous iteration value. Since each individual actuator has a boundary constraint, placing a constraint on the L_2 norm limits the number of actuators that are on the boundary constraint. This combination of

constraints allows all actuators to change value, but also have a new solution that is "local" to the previous iteration. The algorithm form to create a trajectory is

$$\begin{aligned}
&\underset{j}{\text{iterate}} \\
&\qquad \underset{\mathbf{x}_j}{\text{minimize}} \quad J = \left\|\mathbf{C}\mathbf{x}_j - \mathbf{d}\right\|_2 \\
&\qquad \text{subject to} \quad \mathbf{lower} \leq \mathbf{x}_j \leq \mathbf{upper} \\
&\qquad\qquad\qquad \left\|\mathbf{x}_j - \mathbf{x}_{j-1}\right\|_2 \leq \alpha \\
&\qquad\qquad\qquad -\beta \leq \mathbf{x}_j - \mathbf{x}_{j-1} \leq \beta \\
&\text{stop condition} \quad \left\|\mathbf{x}_j - \mathbf{x}_{j-1}\right\|_2 \leq \varepsilon
\end{aligned} \tag{5.6}$$

where the first constraint enforces the hardware voltage limits and we add two new constraints. We use the threshold values α and β to limit the change of values between iterations and ε to determine when the trajectory has "settled" and the algorithm is finished. We initialize the problem with $j = 1$ and $\mathbf{x}_0 = \text{zeros}(M, 1)$ and the iterations cease when the stop condition has been satisfied.

The form of the new constraints only affect the change per iteration and do not affect the final optimal solution $\mathbf{x}^*$ for our original problem; the final iteration solution is equivalent to the solution of the optimization problem (5.1) to within numerical rounding.

To give an idea of how the combination of constraints work, an example trajectory is shown as a projection into the 2D space of actuators x_1 and x_2 in Figure 5.4. For this example, each vector $\mathbf{x}_i$ contains variables x_1 to x_k ($k > 2$). The norm constraint prevents all of the variables from moving beyond a "small ball" per iteration. The radius of the norm constraint when projected in 2D space of x_1 and x_2 is a function of the orthogonal variables x_3 to x_k. Each moving boundary constraint ensures that every actuator does not change by more than a threshold value which forms a box shape. For solutions $\mathbf{x}_1$ and $\mathbf{x}_2$, the boundary constraint was the active constraint that restricted the solution, whereas solution $\mathbf{x}_3$ was restricted by the norm constraint.

The norm constraint is not compatible with the `LSQLIN` solver, which is designed for equality and inequality constraints. For this reason, we use SeDuMi [69] as our solver for the optimization problem. Although we could use the SeDuMi routines directly, we instead choose to use YALMIP [70] to construct the optimization problem. YALMIP translates

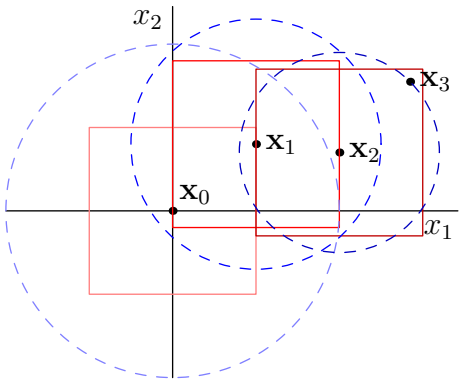


Figure 5.4. An example trajectory of $\mathbf{x}_0$ to $\mathbf{x}_3$ is shown as a 2D projection. The blue circles indicate the α norm constraint and its radius is a function of variables x_3 to x_k (which are not shown). The red squares show the convex β bounding box constraint.

optimization problems into forms understood by solvers and can simplify implementation.

For each run of the optimization problem, SeDuMi returns a new solution that has a lower cost function result. We then continue to iterate until the norm difference between $\mathbf{x}_i$ and $\mathbf{x}_{i-1}$ is less than threshold ε, which indicates that SeDuMi has settled near the optimal solution $\mathbf{x}^*$. With the combination of constraints, the correction begins with low spatial frequencies. As the iterations progress, finer detail is possible in the result.

A drawback to this approach is that the computation time for SeDuMi is much longer than `LSQLIN` since many optimization problems are being solved. The computation time was not an important factor to consider as this control software was not intended for real-time feedback. However, as a means to decrease the processing time, we implemented the optimization problem on multiple grids.

5.3 Multigrid Optimization Problem

The system of equations is overdetermined since $P \gg M$ by an factor of 400, and the optimization algorithm must perform many linear algebra computations.

Adjacent OPD measurements are similar in value, which led to the idea of using the low-pass filters of the Discrete Wavelet Transform to consolidate measurements. Since we are using only the low-pass filters, this is considered a multigrid approach [71]. In Chapter 2, we discussed multiresolution analysis of signals. The significant difference between

multiresolution and multigrid is that multigrid methods discard the high-frequency content whereas multiresolution methods do not [72].

In Section 5.1, we described the construction of the $\mathbf{C}$ matrix and $\mathbf{d}$ vector. For our multigrid technique, we create $\mathbf{C}$ at the multiple grids. We can use operator notation to express the resizing of influence functions in (5.2) as

$$f_m^i[m_1, m_2] = (D_1 D_2 g(z_1) g(z_2))^{i_{\max} - i} f_m^0[n_1, n_2] \tag{5.7}$$

where we use the same operators and $g(z)$ definition from Chapter 2, i is the grid number, and $i_{\max}$ is the number of grids. The mask must also be redefined for each grid, which we can do by downsampling without filtering. At each downsampled resolution, the number of masked samples decreases by a factor of 4 and the masked$(\cdot)$ function is redefined. Each column of $\mathbf{C}$ is still of the form

$$\mathbf{c}_m^i = \text{vector}(\text{masked}(f_m^i)) \tag{5.8}$$

so that the $\mathbf{C}$ matrix at grid i becomes

$$\mathbf{C}_i = \begin{bmatrix} | & | & | & | \\ \mathbf{c}_1^i & \mathbf{c}_2^i & \cdots & \mathbf{c}_M^i \\ | & | & | & | \end{bmatrix}. \tag{5.9}$$

The trajectory creation algorithm begins at the coarsest grid. At this grid, the optimizer is iteratively run until the stop condition is satisfied. The solution is used as the initial guess on the next grid. This process continues until the optimizer is run at the highest resolution

and its stop condition is satisfied. We write this in algorithm form as

$$
\begin{array}{lll}
\underset{i}{\text{iterate}} & & \\
 & \underset{j}{\text{iterate}} & \\
 & & \underset{\mathbf{x}_j}{\text{minimize}} \quad J = \left\|\mathbf{C}_i\mathbf{x}_j - \mathbf{d}_i\right\|_2 \\
 & & \text{subject to} \quad \mathbf{lower} \leq \mathbf{x}_j \leq \mathbf{upper} \\
 & & \qquad\qquad \left\|\mathbf{x}_j - \mathbf{x}_{j-1}\right\|_2 \leq \alpha \\
 & & \qquad\qquad -\beta \leq \mathbf{x}_j - \mathbf{x}_{j-1} \leq \beta \\
 & \text{stop condition} & \left\|\mathbf{x}_j - \mathbf{x}_{j-1}\right\|_2 \leq \varepsilon
\end{array}
\tag{5.10}
$$

where $\mathbf{d}_i$ is the matching data for grid i.

5.4 Experimental Results

To perform a hardware test, a modification was made to the optical arrangement to add another DM into the beam path of the interferometer, which is shown in Figure 5.5. This DM has 140 actuators that act locally on the mirror surface. This additional DM was added because the primary mirror actuators do not have sufficient actuation range for the magnitude of wavefront error. Rather than command the primary mirror actuators, its actuators are set to manufacturer voltages and we use the optimization problem to determine actuator voltages for the small DM only. The optical configuration has $M = 55$ actuators near the primary mirror segment under study; all other actuators were kept at their nominal flat-mirror position. The allowable range of actuator values for the control software to the DM is $0 \leq \mathbf{x} \leq 100$, which uses percentages and not voltages.

For the optimization problem, we set $\alpha = 10$, $\beta = 1$, and $\varepsilon = 10^{-4}$. These were chosen arbitrarily and give a sufficient trajectory with the desired "small" actuator movement.

For the multigrid algorithm, the approximate number of masked samples for each grid is given in Table 5.2. We stopped at the fifth iteration since the next level would result in fewer equations than unknowns ($P < M$).

We show the process of (5.10) visually in Figure 5.6. The DM multigrid trajectory first solves optimization problems on the coarsest grid. Once a solution has been determined

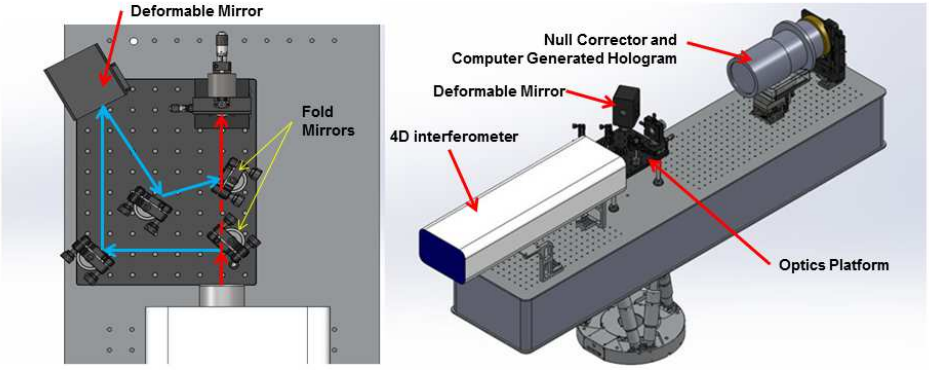


Figure 5.5. Left: The DM table is shown with an overhead view of the optics platform. The laser follows the blue arrows. Right: A view of the optics table that labels the components. Images are courtesy of John Bagnasco.

Table 5.2. The number of samples used for each grid of the multiple resolutions.

Grid level	Number of samples
1	$P \approx 60,000$
2	$P \approx 14,000$
3	$P \approx 3,500$
4	$P \approx 800$
5	$P \approx 190$

for a grid, the result is used as the initial condition on the next higher grid.

The multigrid trajectory algorithm was run on the same data as the single grid trajectory. A comparison of the cost function as a function of time is shown in Figure 5.7 for each approach. Although both approaches have the same final optimal solution, the multigrid algorithm finished approximately 2.5 times faster. The rate of decrease of the cost function is thirty times steeper for the initial two multigrid solutions compared to the initial two single grid solutions (see the tick marks in Figure 5.7).

We expect that for each iteration, the cost function will decrease. The number of iterations at each level i is shown in Table 5.3. In Figure 5.7, we see that while the single grid cost function monotonically decreases, this is not the case for the multigrid. The cost function increase is understood by examining Figure 5.8. The cost function is evaluated at each res-

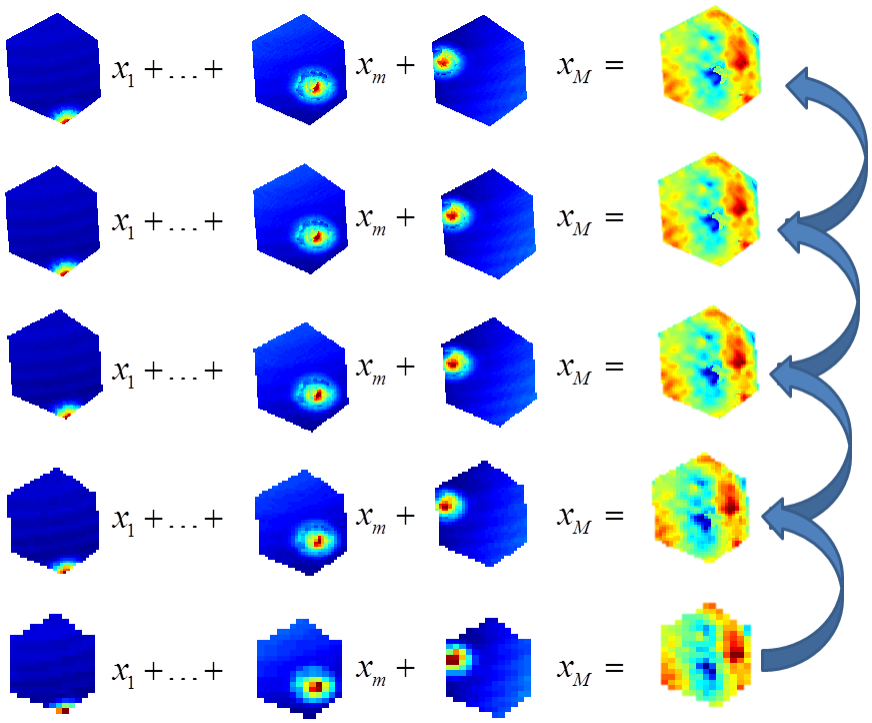


Figure 5.6. The DM multigrid trajectory first solves optimization problems on the coarsest grid and works towards the highest resolution.

olution using $\mathbf{C}_i$ and $\mathbf{d}_i$. We see that each segment of the plot is monotonically decreasing. When switching to a higher resolution, the cost function has an increase. Thus at each grid level, the optimizer is doing the best available solution for the influence functions at that grid and cannot correct for higher resolution distortion. The rate of cost function decrease at the coarse grid is significant compared to using the full grid, which implies that the mirror is effectively controlled at the coarse grid. There is not a substantial improvement to control the mirror at the full resolution. With much higher actuator density ($M \gg 55$), we would expect that the mirror surface would improve at higher resolution control.

Another feature to note in Figure 5.8 is the distances between the tick marks which indicate an iteration of the optimization problem. At the high resolution, the tick marks are broadly spaced because of the time required to perform the large-scale linear algebra. The multigrid approach solves each iteration much faster, though it has to perform more iterations to arrive at the same solution.

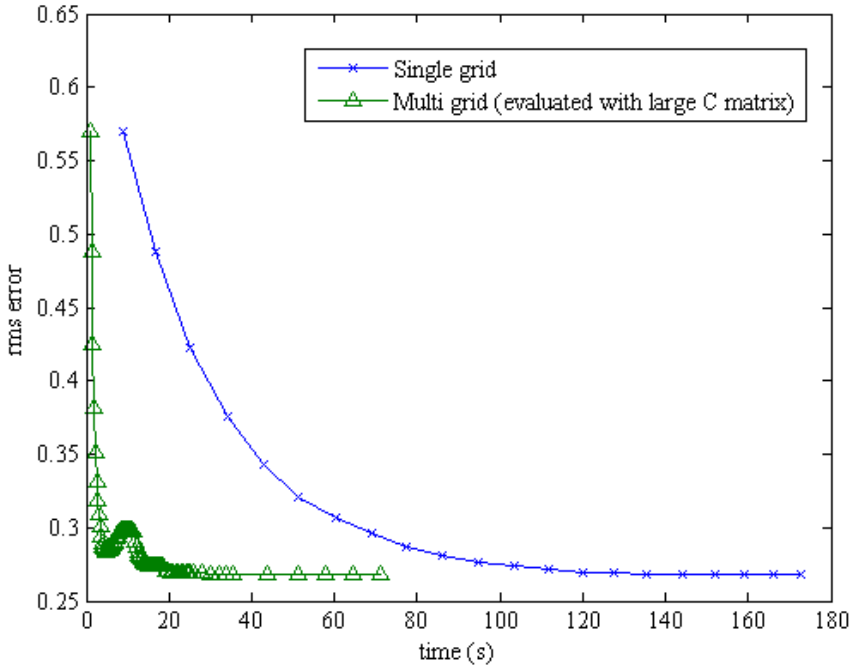


Figure 5.7. The cost function evaluated at the highest resolution and plotted against the processing time of the trajectory creation algorithm for single grid and a 5-level multigrid.

Table 5.3. The number of iterations for multigrid and single grid algorithms.

	Grid level	Number of iterations
Multigrid	1	42
	2	21
	3	13
	4	5
	5	5
Single grid		22

The hardware response shown in Figure 5.9 is compared against the linear model prediction using the multigrid solutions. Up until approximately 5 optimizer iterations, the agreement between the linear model and hardware response is very close. After about 10 optimizer iterations, the model does not accurately predict the hardware response, which is due to the expected nonlinear behavior of the mirror [73], [74]. Despite this, the mirror surface does not see a significant change in the wavefront rms error as it levels off. However, while we can see that the hardware response has small oscillations in the wavefront rms error, the

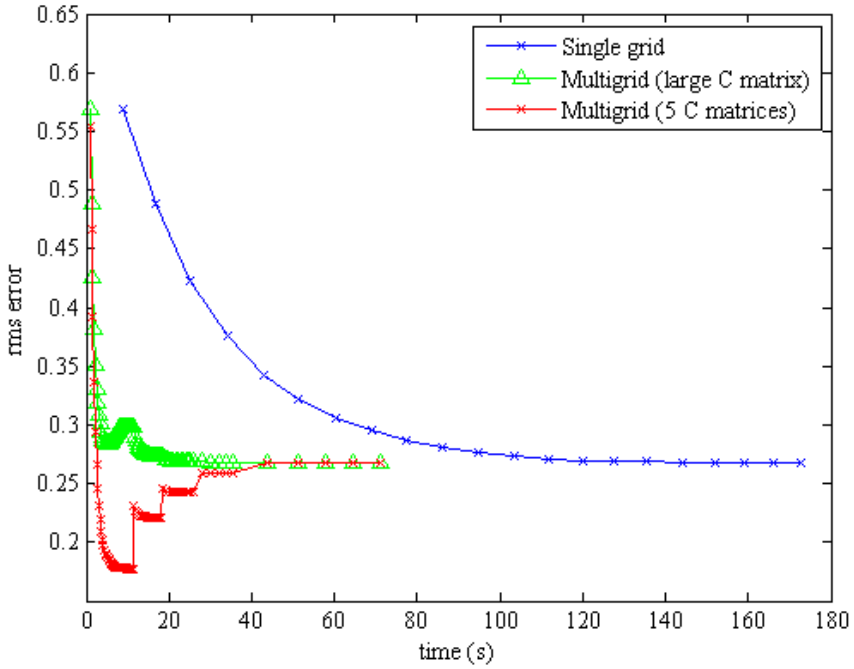


Figure 5.8. The cost function evaluated at the same resolution as the iterative solution and plotted against time.

overall wavefront rms error is fairly constant. The changes in the wavefront rms error can be attributed to nonlinearities in the mirror surface and measurement noise in the interferometer. The optical alignment is moving due to mechanical vibration. Each interferometer measurement takes a duration of time and the motion has an effect on the result. In some cases, small areas of the mirror surface experience a different phase unwrapping result from the interferometer software. All of these effects attribute to the oscillations in wavefront error rms.

In Figure 5.10, we show the mirror surface results from the final iteration of each grid. In each case, the rms wavefront error is consistent despite the changing the actuator settings. Although the DM has nonlinearities, the performance did not degrade due to them.

For our optical configuration, we verified the performance of the `LSQLIN` solution by generating a trajectory of solutions to test because we did not see the wavefront error increase as the iterations progressed. The linear range is shown to be about ± 10 percentage points

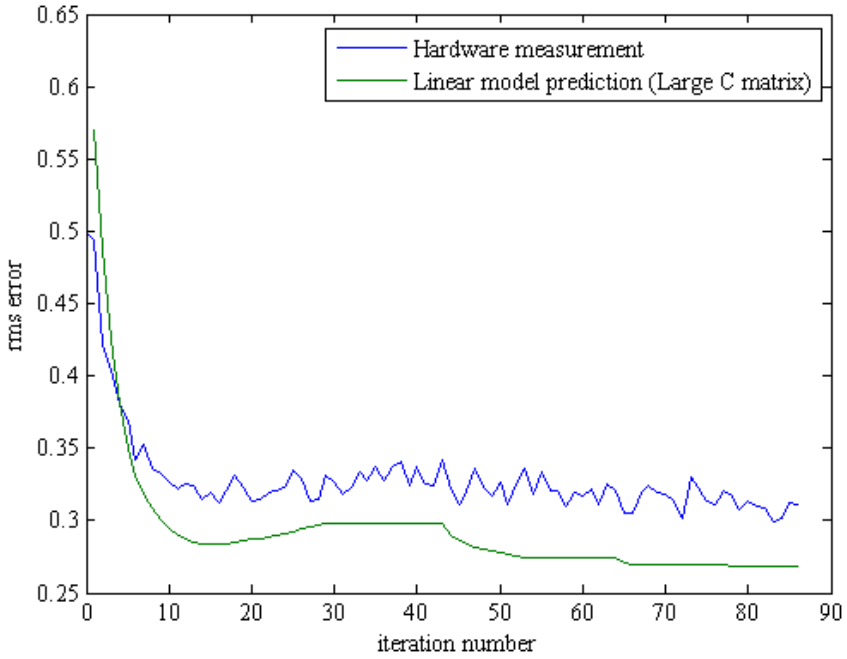


Figure 5.9. The hardware measurement is compared against the multigrid solution.

of the actuators nominal setting. Past this range, the linear system model and hardware response differ due to the nonlinearities. The difference did not have a significant impact on the wavefront rms error. We also showed that the rms error was not influenced by using the full resolution interferometer image. For our system, the same level of wavefront error performance was possible from using a coarse grid of measurements.

5.5 Summary

In this chapter, we sought to experimentally test whether an optimization problem effectively controlled a DM. We sought to understand the hardware response of using optimization for mirror surface control and developed a trajectory creation algorithm to generate solutions to test on the hardware. To improve the computational efficiency, we used a multigrid approach and the optimizer solves the trajectory algorithm on several grids. The multigrid approach resulted in the same optimal solution at reduced computational cost when compared to the original trajectory creation algorithm. We verified the performance of a linear influence function model and determined the valid linear range of the DM. We

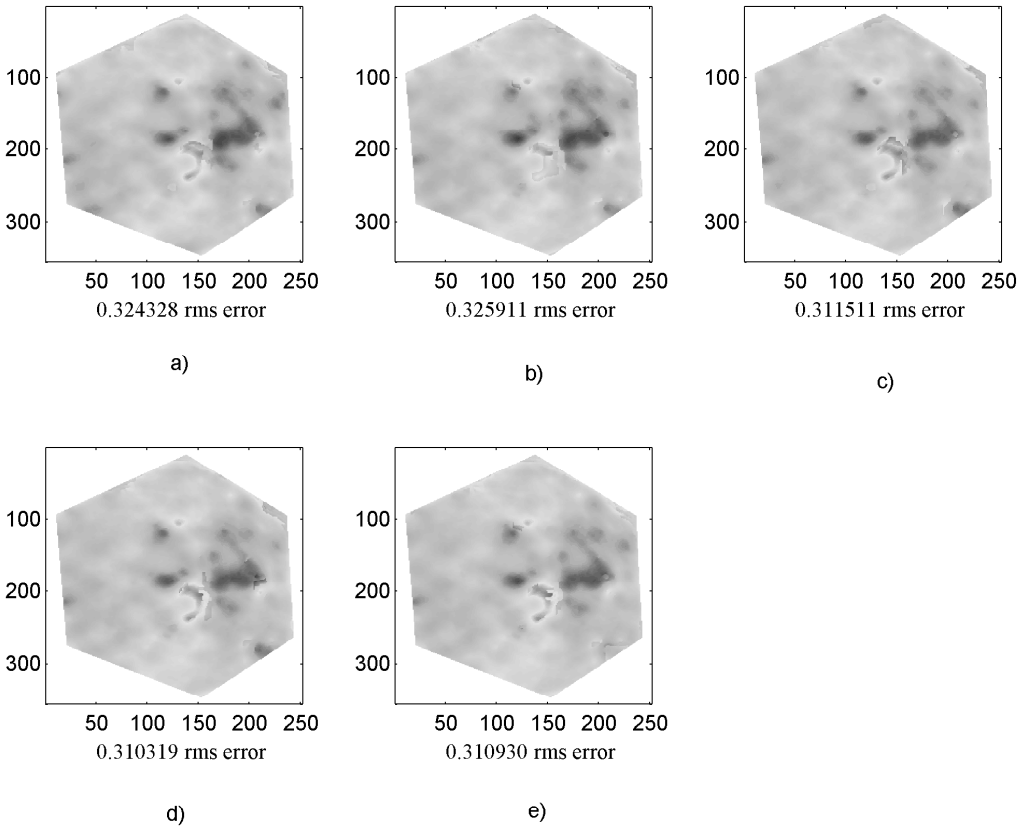


Figure 5.10. The hardware measurement of the segment at the final solution is shown for a) the coarsest grid through e) the finest grid.

would expect further reductions in the wavefront rms error by using a DM with higher actuator density.

CHAPTER 6:
Conclusion

6.1 Main Contributions

This dissertation presented three contributions to AO. The first is a wavelet approach to wavefront reconstruction. The algorithm uses local gradient measurements from a Shack-Hartmann wavefront sensor to estimate the wavefront phase. The 2D QMF tree structure and Noble identities are used to decompose the wavefront into the spatial frequency components. The inverse discrete wavelet transform is performed using these components to estimate the phase. Our method allows for the use of orthogonal wavelet filters and using longer filter lengths has improved noise-rejection performance for the estimation compared to the Haar wavelet used by Hampton et al. [43], [44]. He later proposed a Poisson solver to improve the noise performance [45]. For the high SNR data case, a modification was shown to make the resulting wavefront estimate not depend on the boundary condition.

This algorithm has been designed for irrotational vector fields, where there is no phase ambiguity and the phase is well defined at every point. This is, in general, the case of distortion generated by low atmospheric turbulence. Under more severe turbulence conditions, the intensity of the optical field might be zero at isolated points, thus causing phase uncertainties. In this case, the measured phase gradient becomes rotational and it is characterized by phase uncertainty, and branch points, which cause problems in all standard least-squares algorithms.

The second contribution adapts the proposed algorithm to work when branch points are present from significant atmospheric turbulence. An analysis of vector spaces shows that the branch points cause the rotational components in the measured gradients. An approach using a non-orthogonal decomposition to modify the rotational vector field to be irrotational is presented. The wavefront reconstruction algorithm operates on the irrotational measurements and estimates the phase that is consistent with the original measurements with rotational components. Our results show the wrapped phase to make a comparison between the simulated phase and reconstructed phase. This approach can be applied to any wavefront reconstruction algorithm as the measurements are modified before the algorithm.

A third contribution is made in segmented mirror with active optics. The control signals to a DM are computed that minimize the wavefront error using constrained optimization to ensure that the hardware actuator voltage limits are satisfied. The optimization problem uses a linear influence function model for the actuators to determine the voltages that form the desired mirror surface. The best results in terms of efficiency and convergence are obtained with a multigrid approximation of the influence functions.

An experiment verified the performance of the mirror surface control using constrained optimization. A trajectory creation algorithm solved a sequence of optimization problems which forms a sequence of actuator values. Each solution has small changes from the previous voltages and results in the same final optimal solution. The trajectory provides insight into the validity of the linear influence function model when compared to the hardware response. An OPD is collected for each iteration and used to evaluate the residual wavefront error. The multigrid approach is shown to have a rapid decrease in the cost function when compared to the single grid solution and results in the same final solution as the single grid 2.5 times faster. Using a large amount of interferometer measurements did not significantly change the residual wavefront error, as the residual wavefront error for the optimal solution at the coarsest grid was comparable to the residual wavefront error for the fine grid. For the linear influence function model, the number of OPD measurements only needs to be on the order of the number of DM actuators. Analysis determined the range of linear operation for the DM, which for the particular DM used in the experiment, is small over the entire range of operation. Mirror surface control using optimization of a linear influence model has similar performance to other tested linear control methods [75] to decrease the wavefront error of a segmented mirror. For large corrections, a nonlinear controller can obtain better performance for the DM [76].

6.2 Future Work

The wavelet phase reconstruction algorithm was applied to a Cartesian lattice. Other lattices, such as the hexagonal lattice, are also used in AO. Sensors use hexagon shaped lenslets which can be packed tightly together and use all of the collected light for measurements. The wavefront reconstruction algorithm can be extended to work on measurement data taken in this lattice.

Any orthogonal wavelets can be used for the phase reconstruction algorithm. One may develop specialized filter functions for the purpose of wavefront reconstruction that uses the statistics of the atmosphere to determine filter coefficients. Choosing wavelet coefficients may also be done in an adaptive manner to improve performance.

In the study of the algorithm performance with noisy measurements, only Gaussian noise was used. This noise model is appropriate for the the case of AO sensor centroid measurements. For using this algorithm on RADAR data, a noise analysis should be conducted with other noise models.

The wavefront reconstruction algorithm and branch point modification were studied extensively in computer simulation. The phase reconstruction can be studied in the laboratory to understand the estimation error for different wavelets given simulated atmospheric conditions. The branch point modification can be studied experimentally to verify performance for laboratory or atmospheric branch points.

The wavefront reconstruction algorithm operates on a set of measurements independently from previous estimations of the wavefront. Successive estimations of the phase can be smoothed or blended to analyze AO performance and understand how the wavefront evolves temporally. The branch cuts can significantly change paths for each estimation due to noise but this is undesirable for hardware performance. Future work may use previous estimations try to smooth the evolution of branch cut placement.

Following wavefront reconstruction, phase unwrapping may be necessary if the estimated phase contains large discontinuities. Some wavefront reconstruction algorithms (such as the complex exponential reconstructor [58], [59]) also perform phase unwrapping. Additional work may be done to analyze which phase unwrapping algorithm works best with the wavelet wavefront reconstruction algorithm.

For the segmented mirror with active optics contribution, the analysis of the mirror surface optimization can be extended by using a DM with higher actuator density. Our conclusions showed a small linear range of operation over which the linear approximation model can be used. Work can be done to extend the cost function to the nonlinear influence case. One possible approach is to redefine the influence function matrix to vary as a function of

actuator settings as $\mathbf{C}(\mathbf{x})$. After each iteration of the optimization problem, new influence functions can be generated for actuators to create a new linear influence matrix. Each optimization problem is treated linearly and the same approach can be applied with the new $\mathbf{C}$ matrix on the next iteration of the optimization problem.

We converted the optimization problem to modal basis by transforming $\mathbf{C}$ and $\mathbf{d}$, which changes the number of rows for each. For the Zernike polynomial basis, the result was not satisfactory due to having fewer equations (number of polynomials used) than number of actuators. Most Zernike decompositions only use ~ 50 polynomials, so for high actuator density, this basis is undesirable as there are more unknowns than equations. We also explored other techniques to decrease the number of rows of $\mathbf{C}$ and $\mathbf{d}$, such as compressed sensing. Random sets of samples (at least as many as number of actuators) were kept for each OPD. The estimated wavefront surface was usually a good match for compressed sensing but would miss features of the wavefront that occurred away from the random samples. The best performance of a reduced system of equations we observed was from the multigrid approach, but any technique that can exploit the vastly oversampled wavefront may provide further reduction in wavefront error.

One nonlinear control approach from the manufacturer is a lookup table to determine the actuator settings for a desired mirror surface. The use of a lookup table in a mirror surface control optimization problem has not been studied.

In the iterative optimization problem, only one value for ε was used at every grid level. Changing this value to be larger for coarser grids may reduce the number of iterations and result in faster convergence to the optimal solution.

The cost function can be adjusted for multiple deformable mirrors to be installed in the AO system. The optimization problem can then choose the actuator commands for each mirror. Adjusting the constraints can result in a variety of techniques to split the necessary correction among the multiple mirrors.

APPENDIX: Proofs

High-Order Wavelet Simplification Proof

This proof shows how the results of (3.14) are determined. We start with the definition

$$g(-z^N) \triangleq \frac{1-z^{-N}}{\sqrt{2}} \tag{A.1}$$

and the definition of the geometric series

$$\sum_{\ell=0}^{N-1} z^{-\ell} \triangleq \frac{1-z^{-N}}{1-z^{-1}}. \tag{A.2}$$

We immediately observe that Eqs. (A.1) and (A.2) can be combined:

$$g(-z^N) = \frac{1-z^{-N}}{\sqrt{2}} = \frac{\left(\sum_{\ell=0}^{N-1} z^{-\ell}\right)\left(1-z^{-1}\right)}{\sqrt{2}} = \left(\sum_{\ell=0}^{N-1} z^{-\ell}\right) g(-z). \tag{A.3}$$

We have now shown the first result. The second result takes some manipulation similar to the concept of polyphase decomposition where we split the sequence up into an even and odd component. We proceed from the result of (A.3) in

$$\begin{aligned}\left(\sum_{\ell=0}^{N-1} z^{-\ell}\right) g(-z) &= \left(\sum_{\ell=0}^{\frac{N}{2}-1} z^{-2\ell} + z^{-2\ell-1}\right) g(-z) \\ &= \left((1+z^{-1}) \sum_{\ell=0}^{\frac{N}{2}-1} z^{-2\ell}\right) g(-z) \\ &= \left(\sum_{\ell=0}^{\frac{N}{2}-1} z^{-2\ell}\right) \sqrt{2} g(z) g(-z). \quad \blacksquare\end{aligned} \tag{A.4}$$

The first simplification is the realization that the sum of the two sequences can be factored to $1+z^{-1}$. The final factoring swaps with the Haar scaling function and needs a $\sqrt{2}$ to cancel the denominator.

Branch Point Boundary Condition Proof

Let $w(x,y)$ be such that

$$\nabla_x \nabla_y w(x,y) = 0 \tag{A.5}$$

for all x, y real. Then $w(x,y)$ can be written as

$$w(x,y) = a(x) + b(y). \tag{A.6}$$

To show this, define

$$g(x,y) \equiv \nabla_x w(x,y) \tag{A.7}$$

Then $\nabla_y g(x,y) = 0$ and therefore,

$$g(x,y) = g(x,0), \tag{A.8}$$

i.e., independent of y. Substitute in (A.7) to obtain

$$w(x,y) = w(0,y) + \int_0^x g(\lambda,0) d\lambda, \tag{A.9}$$

which shows the result. ■

References

[1] R. K. Tyson, *Introduction to Adaptive Optics*. Bellingham, WA: SPIE Press, 2000.

[2] R. K. Tyson, *Principles of Adaptive Optics*, 2nd ed. Boston: Academic Press, 1998.

[3] P.-Y. Bely, Ed., *The Design and Construction of Large Optical Telescopes*. New York: Springer, 2003.

[4] R. S. Capers and E. Lipton, "Hubble error: Time, money and millionths of an inch," *The Academy of Management Executive*, vol. 7, no. 4, pp. 41–57, 1993. [Online]. Available: http://www.jstor.org/stable/4165153

[5] University of Arizona. (2014). "The University of Arizona Stewart Observatory Mirror Laboratory" [Online]. Available: http://mirrorlab.as.arizona.edu

[6] A. Yingling, "Integrated optics, structures, and controls of segmented mirror telescopes," Ph.D. dissertation, Naval Postgraduate School, Monterey, CA, Sept. 2012.

[7] GRANTECAN S.A. (2014). "The Gran Telescopio CANARIAS" [Online]. Available: http://www.gtcdigital.net

[8] NASA. (2014). "The Keck Interferometer" [Online]. Available: http://www.nasa.gov/centers/jpl/missions/keck.html

[9] TMT Observatory Corporation. (2014). "The Thirty Meter Telescope" [Online]. Available: http://www.tmt.org

[10] GMT Consortium. (2014). "The Giant Magellan Telescope Observatory" [Online]. Available: http://www.gmto.org

[11] J. P. Gardner, J. C. Mather, M. Clampin, R. Doyon, M. A. Greenhouse, H. B. Hammel, J. B. Hutchings, P. Jakobsen, S. J. Lilly, K. S. Long, J. I. Lunine, M. J. McCaughrean, M. Mountain, J. Nella, G. H. Rieke, M. J. Rieke, H.-W. Rix, E. P. Smith, G. Sonneborn, M. Stiavelli, H. S. Stockman, R. A. Windhorst, and G. S. Wright, "The James Webb space telescope," *Space Science Reviews*, vol. 123, no. 4, pp. 485–606, 2006. [Online]. Available: http://dx.doi.org/10.1007/s11214-006-8315-7

[12] J. R. Wertz, D. F. Everett, and J. J. Puschell, *Space Mission Engineering: The New SMAD*. Hawthorne, CA: Microcosm Press, 2011.

[13] NASA. (2014). "The James Webb Space Telescope" [Online]. Available: http://www.jwst.nasa.gov

[14] D. E. Dudgeon and R. M. Mersereau, *Multidimensional Digital Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1984.

[15] R. Cristi, *Modern Digital Signal Processing*. Pacific Grove, CA: Thomson/Brooks/Cole, 2004.

[16] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*. Englewood Cliffs, NJ: Prentice Hall, 1993.

[17] D. S. Taubman, *JPEG2000: Image Compression Fundamentals, Standards, and Practice*. Boston: Kluwer Academic, 2002.

[18] R. Cristi. (2014). Class Lecture, Topic: "Quadrature Mirror Filters" [Online]. Available: http://faculty.nps.edu/rcristi/EC4480/weekly.htm

[19] J. W. Goodman, *Introduction to Fourier Optics*. Englewood, CO: Roberts & Co., 2005.

[20] V. N. Mahajan, *Optical Imaging and Aberrations: Part II*. Bellingham, WA: SPIE Optical Engineering Press, 1998.

[21] H. H. Barrett and K. J. Myers, *Foundations of Image Science*. Hoboken, NJ: Wiley-Interscience, 2004.

[22] D. Malacara, Ed., *Optical Shop Testing*, 3rd ed. Hoboken, NJ: Wiley-Interscience, 2007.

[23] H. Campbell and A. Greenaway, "Wavefront sensing: from historical roots to the state-of-the-art," in *Astronomy with High Contrast Imaging III: Instrumental Techniques, Modeling and Data Processing*, vol. 22, 2006, pp. 165–185. [Online]. Available: http://dx.doi.org/10.1051/eas:2006131

[24] F. Roddier, *Adaptive Optics in Astronomy*. Cambridge: Cambridge University Press, 1999.

[25] W. H. Southwell, "Wave-front estimation from wave-front slope measurements," *Journal of the Optical Society of America*, vol. 70, no. 8, pp. 998–1006, Aug. 1980. [Online]. Available: http://www.opticsinfobase.org/abstract.cfm?URI=josa-70-8-998

[26] L. A. Poyneer, D. T. Gavel, and J. M. Brase, "Fast wave-front reconstruction in large adaptive optics systems with use of the Fourier transform," *Journal of the Optical Society of America A*, vol. 19, no. 10, pp. 2100–2111, Oct. 2002. [Online]. Available: http://josaa.osa.org/abstract.cfm?URI=josaa-19-10-2100

[27] D. L. Fried, "Least-square fitting a wave-front distortion estimate to an array of phase-difference measurements," *Journal of the Optical Society of America*, vol. 67, no. 3, pp. 370—375, Mar. 1977. [Online]. Available: http://www.opticsinfobase.org/abstract.cfm?URI=josa-67-3-370

[28] R. H. Hudgin, "Wave-front reconstruction for compensated imaging," *Journal of the Optical Society of America*, vol. 67, no. 3, pp. 375—378, Mar. 1977. [Online]. Available: http://www.opticsinfobase.org/abstract.cfm?URI=josa-67-3-375

[29] J. Herrmann, "Least-squares wave front errors of minimum norm," *Journal of the Optical Society of America*, vol. 70, no. 1, pp. 28–35, 1980.

[30] K. Freischlad, "Wavefront reconstruction from noisy slope or difference data using the discrete Fourier transform," in *Proceedings of SPIE*, vol. 551, Arlington, VA, 1985, pp. 74–80.

[31] L. A. Poyneer, "Advanced techniques for Fourier transform wavefront reconstruction," in *Proceedings of SPIE*, vol. 4839. Waikoloa, HI: SPIE, 2003, pp. 1023–1034. [Online]. Available: http://dx.doi.org/10.1117/12.459461

[32] L. Gilles, C. R. Vogel, and B. L. Ellerbroek, "Multigrid preconditioned conjugate-gradient method for large-scale wave-front reconstruction," *Journal of the Optical Society of America A*, vol. 19, no. 9, pp. 1817–1822, Sept. 2002. [Online]. Available: http://josaa.osa.org/abstract.cfm?URI=josaa-19-9-1817

[33] D. G. MacMartin, "Local, hierarchic, and iterative reconstructors for adaptive optics," *Journal of the Optical Society of America A*, vol. 20, no. 6, pp. 1084–1093, Jun. 2003. [Online]. Available: http://josaa.osa.org/abstract.cfm?URI=josaa-20-6-1084

[34] L. Gilles, "Sparse minimum-variance open-loop reconstructors for extreme adaptive optics: order N multigrid versus preordered Cholesky factorization," in *Proceedings of SPIE*, vol. 5169, Dec. 2003, pp. 201–205. [Online]. Available: http://dx.doi.org/10.1117/12.506893

[35] C. R. Vogel, "Sparse matrix methods for wavefront reconstruction, revisited," in *Proceedings of SPIE*, vol. 5490, Oct. 2004, pp. 1327–1335. [Online]. Available: http://dx.doi.org/10.1117/12.548689

[36] C. R. Vogel and Q. Yang, "Multigrid algorithm for least-squares wavefront reconstruction," *Applied Optics*, vol. 45, no. 4, pp. 705–715, Feb. 2006. [Online]. Available: http://ao.osa.org/abstract.cfm?URI=ao-45-4-705

[37] C. Béchet, M. Tallon, and E. Thiébaut, "FRIM: minimum-variance reconstructor with a fractal iterative method," in *Proceedings of SPIE*, vol. 6272, Orlando, FL, Jun. 2006, pp. 62 722U–62 722U. [Online]. Available: http://dx.doi.org/10.1117/12.672197

[38] E. Thiébaut and M. Tallon, "Fast minimum variance wavefront reconstruction for extremely large telescopes," *Journal of the Optical Society of America A*, vol. 27, no. 5, pp. 1046–1059, May 2010. [Online]. Available: http://josaa.osa.org/abstract.cfm?URI=josaa-27-5-1046

[39] J. Herrmann, "Phase variance and Strehl ratio in adaptive optics," *Journal of the Optical Society of America A*, vol. 9, no. 12, pp. 2257–2258, Dec. 1992. [Online]. Available: http://josaa.osa.org/abstract.cfm?URI=josaa-9-12-2257

[40] M. Rosensteiner, "Cumulative reconstructor: Fast wavefront reconstruction algorithm for extremely large telescopes," *Journal of the Optical Society of America A*, vol. 28, no. 10, pp. 2132–2138, Oct. 2011. [Online]. Available: http://josaa.osa.org/abstract.cfm?URI=josaa-28-10-2132

[41] C. C. de Visser and M. Verhaegen, "Wavefront reconstruction in adaptive optics systems using nonlinear multivariate splines," *Journal of the Optical Society of America A*, vol. 30, no. 1, pp. 82–95, Jan. 2013. [Online]. Available: http://josaa.osa.org/abstract.cfm?URI=josaa-30-1-82

[42] F. U. Dowla, "Fast Fourier and wavelet transforms for wavefront reconstruction in adaptive optics," in *Proceedings of SPIE*, San Diego, CA, 2000, pp. 118–127. [Online]. Available: http://spie.org/Publications/Proceedings/Paper/10.1117/12.407495

[43] P. J. Hampton, P. Agathoklis, and C. Bradley, "A new wave-front reconstruction method for adaptive optics systems using wavelets," *IEEE Journal of Selected Topics in Signal Processing*, vol. 2, no. 5, pp. 781–792, Oct. 2008. [Online]. Available: http://dx.doi.org/10.1109/JSTSP.2008.2006386

[44] P. J. Hampton, "Robust order N wavelet filterbanks to perform 2-D numerical integration directly from partial difference or gradient measurements," Ph.D. dissertation, Dept. of Electrical and Computer Engineering, University of Victoria, 2009.

[45] P. J. Hampton, P. Agathoklis, R. Conan, and C. Bradley, "Closed-loop control of a woofer-tweeter adaptive optics system using wavelet-based phase reconstruction," *Journal of the Optical Society of America A: Optics and Image Science, and Vision*, vol. 27, no. 11, pp. A145–A156, 2010, compendex. [Online]. Available: http://dx.doi.org/10.1364/JOSAA.27.00A145

[46] A. Haar, "Zur theorie der orthogonalen funktionensysteme (On the theory of orthogonal function systems)," Ph.D. dissertation, University of Göttingen, July 1909.

[47] G. Strang and T. Nguyen, *Wavelets and Filter Banks*. Wellesley, MA: Wellesley-Cambridge Press, 1996.

[48] D. L. Fried, "Branch point problem in adaptive optics," *Journal of the Optical Society of America A*, vol. 15, no. 10, pp. 2759–2768, 1998. [Online]. Available: http://www.opticsinfobase.org/abstract.cfm?URI=JOSAA-15-10-2759

[49] L. C. Andrews, *Laser Beam Propagation through Random Media*, 2nd ed. Bellingham, WA: SPIE Press, 2005.

[50] R. J. Sasiela, *Electromagnetic Wave Propagation in Turbulence: Evaluation and Application of Mellin Transforms*, 2nd ed. Bellingham, WA: SPIE, 2007.

[51] J. D. Barchers, D. L. Fried, and D. J. Link, "Evaluation of the performance of Hartmann sensors in strong scintillation," *Applied Optics*, vol. 41, no. 6, pp. 1012–1021, Feb. 2002. [Online]. Available: https://www.opticsinfobase.org/ao/abstract.cfm?uri=ao-41-6-1012

[52] J. F. Nye and M. V. Berry, "Dislocations in wave trains," *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, vol. 336, no. 1605, pp. 165–190, Jan. 1974. [Online]. Available: http://rspa.royalsocietypublishing.org/content/336/1605/165

[53] D. L. Fried and J. L. Vaughn, "Branch cuts in the phase function," *Applied Optics*, vol. 31, no. 15, pp. 2865—2882, May 1992. [Online]. Available: https://www.opticsinfobase.org/ao/abstract.cfm?uri=ao-31-15-2865

[54] D. C. Ghiglia and M. D. Pritt, *Two-Dimensional Phase Unwrapping: Theory, Algorithms, and Software*. New York: Wiley, 1998.

[55] G. A. Tyler, "Reconstruction and assessment of the least-squares and slope discrepancy components of the phase," *Journal of the Optical Society of America A*, vol. 17, no. 10, pp. 1828–1839, Oct. 2000. [Online]. Available: https://www.opticsinfobase.org/josaa/abstract.cfm?URI=josaa-17-10-1828

[56] R. Cristi and T. W. Axtell, "Least-squares phase estimation with wrapped measurements and branch points," *Journal of the Optical Society of America A*, vol. 30, no. 11, pp. 2225–2236, Nov. 2013. [Online]. Available: http://josaa.osa.org/abstract.cfm?URI=josaa-30-11-2225

[57] D. L. Fried and J. L. Vaughn, "Branch cuts in the phase function," *Applied Optics*, vol. 31, no. 15, pp. 2865–82, May 1992. [Online]. Available: https://www.opticsinfobase.org/ao/abstract.cfm?uri=ao-31-15-2865

[58] D. L. Fried, "Adaptive optics wave function reconstruction and phase unwrapping when branch points are present," *Optics Communications*, vol. 200, no. 1–6, pp. 43–72, Dec. 2001.

[59] S. C. Moran, "Comparison of a noise-variance weighted complex exponential reconstructor with traditional reconstructors in the presence of deep turbulence," M.S. thesis, Dept. of Mechanical and Aerospace Engineering, Naval Postgraduate School, Monterey, CA, June 2013.

[60] R. Olsen, *Remote Sensing from Air and Space*. Bellingham, WA: SPIE Press, 2007. [Online]. Available: http://spie.org/Publications/Book/868232

[61] T. W. Axtell and R. Cristi, "Generalized orthogonal wavelet phase reconstruction," *Journal of the Optical Society of America A*, vol. 30, no. 5, pp. 859–870, 2013. [Online]. Available: http://josaa.osa.org/abstract.cfm?URI=josaa-30-5-859

[62] T. J. Brennan and P. Roberts, "AOTools: The adaptive optics toolbox for use with MATLAB," 2010.

[63] T. J. Brennan, P. Roberts, and D. Mann, "WaveProp: A wave optics simulation system for use with MATLAB," 2010.

[64] K. Doyle, V. Genberg, and G. Michels, *Integrated Optomechanical Analysis*, 2nd ed. Bellingham, WA: SPIE Press, 2012.

[65] S. P. Boyd, *Convex Optimization*. Cambridge: Cambridge University Press, 2004.

[66] MATLAB, *version 7.10.0 (R2013b)*. Natick, MA: The MathWorks Inc., 2013.

[67] T. F. Coleman and Y. Li, "A reflective Newton method for minimizing a quadratic function subject to bounds on some of the variables," *SIAM J. on Optimization*, vol. 6, no. 4, pp. 1040–1058, Apr. 1996. [Online]. Available: http://dx.doi.org/10.1137/S1052623494240456

[68] L. E. Cohan, "Integrated modeling for design of lightweight, active mirrors," *Optical Engineering*, vol. 50, no. 6, p. 063003, Jun. 2011. [Online]. Available: http://spie.org/Publications/Journal/10.1117/1.3592520

[69] J. F. Sturm, "Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones," *Optimization Methods and Software*, vol. 11, no. 1-4, pp. 625–653, 1999. [Online]. Available: http://www.tandfonline.com/doi/abs/10.1080/10556789908805766

[70] J. Löfberg, "YALMIP: A toolbox for modeling and optimization in MATLAB," in *Proceedings of CACSD*, Taipei, Taiwan, 2004. [Online]. Available: http://users.isy.liu.se/johanl/yalmip

[71] W. L. Briggs, V. E. Henson, and S. F. McCormick, *A Multigrid Tutorial*. Philadelphia, PA: SIAM, 2000.

[72] W. L. Briggs and V. E. Henson, "Wavelets and multigrid," *SIAM Journal on Scientific Computing*, vol. 14, no. 2, pp. 506–510, Mar. 1993. [Online]. Available: http://dx.doi.org/10.1137/0914031

[73] Y. Zhou and T. Bifano, "Adaptive optics using a MEMS deformable mirror," in *Proceedings of SPIE*, vol. 6018, 2005, pp. 601 817–1–601 817–7. [Online]. Available: http://dx.doi.org/10.1117/12.669376

[74] D. Guzmán, F. J. De Cos Juez, R. Myers, F. Sánchez Lasheras, L. K. Young, and A. Guesalaga, "Deformable mirror models for open-loop adaptive optics using non-parametric estimation techniques," in *Proceedings of SPIE*, vol. 7736, Jul. 2010, pp. 77 361C–77 361C. [Online]. Available: http://dx.doi.org/10.1117/12.856894

[75] J. J. Watson, "Correcting surface figure error in imaging satellites using a deformable mirror," M.S. thesis, Dept. of Mechanical and Aerospace Engineering, Naval Postgraduate School, Monterey, CA, Dec. 2013.

[76] J. B. Stewart, A. Diouf, Y. Zhou, and T. G. Bifano, "Open-loop control of a MEMS deformable mirror for large-amplitude wavefront control," *Journal of the Optical Society of America A*, vol. 24, no. 12, pp. 3827–3833, Dec. 2007. [Online]. Available: http://josaa.osa.org/abstract.cfm?URI=josaa-24-12-3827

THIS PAGE INTENTIONALLY LEFT BLANK

Initial Distribution List

1. Defense Technical Information Center
   Ft. Belvoir, Virginia
2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California