INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

ACHROMATIC NULLING BEAM COMBINER FOR THE DETECTION OF EXTRASOLAR PLANETS

by

Rhonda Michelle Morgan

A Dissertation Submitted to the Faculty of the

COMMITTEE ON OPTICAL SCIENCES (GRADUATE)

In Partial Fulfillment of the Requirements
For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2001

THE UNIVERSITY OF ARIZONA ®
GRADUATE COLLEGE

As members of the Final Examination Committee, we certify that we have read the dissertation prepared by Rhonda Michelle Morgan

entitled Achromatic Nulling Beam Combiner for the Detection of Extrasolar Planets

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

James H. Burge	Nov. 19, 01 Date
James C. Wyant	11/19/01 Date
Neville J. Woolf	11-19-01 Date
	Date
	Date

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

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

Dissertation Director James H. Burge — Nov 30 2001 Date

STATEMENT BY THE AUTHOR

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

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: Rhonda Morgan

Existing instruments are unable to detect planets about stars other than the Sun but such detection would be important for the theory of origin of our solar system and in the search for extraterrestrial intelligence. Infrared offers an advantage of about 10^5 over visible light as regards the ratio of power received from star and planet. Infrared interferometry from Earth orbit would allow discrimination against the stellar infrared by the placement of an interference null on the star and a spinning infrared interferometer would modulate the planetary emission to permit extraction by synchronous detection from the background level.

R. Bracewell and R. MacPhie

Icarus, **38**, 136, 1979

ACKNOWLEDGMENTS

This work was supported by NASA Graduate Student Research Program award NGT5-50046 and grant # 961286 from Caltech's Jet Propulsion Laboratory, operated under contract to NASA. Any opinions, findings, and conclusions or recommendations expressed in this dissertation are those of the author and do not necessarily reflect the views of NASA.

This research would not have been possible without the assistance of many people, especially the professors, staff, and students of Steward Observatory and the Optical Sciences Center. I received problem solving advice from Jim Burge, Bill Hoffman, Roger Angel, Michael Lloyd-Hart, Don McCarthy, and Derek Sabatke. I benefited from the insights of Jim Breckenridge and Stuart Shaklan. Rigel Woida and Kameron Rausch assisted in lab. Matt Cheselka wrote software drivers, coded my algorithms and coaxed the computer into controlling the hardware of the experiment. Brian Duffy provided mechanical support. Doris Tucker, Kim Chapman, and Diana Dawson were invaluable for their administrative assistance at the Center for Astronomical Adaptive Optics. At OSC, Barbara Myers, my fellowship coordinator, and Did Lawson in the academic affairs office carried much of the burden of paperwork to satisfy the University.

My sanity would not have been possible without the help of my Tucson friends. Erin Sabatke is the best person I know to talk to when the going gets rough; she asks really good questions. Neil and I took many a late night walk to air research frustrations and contemplate the mysteries of truth. William and I discussed the travails of life. Matt and Laura refreshed me with home cooked meals and stimulating conversation. Kim cooked dinners that became a community of support. Tim saw poesis in my research.

I was strengthened by this prayer of the explorer Sir Frances Drake:

> Disturb us, Lord, when we are too well pleased with ourselves,
> when our dreams have come true because we have dreamed too little,
> when we arrive safely because we have sailed too close to the shore.
>
> Disturb us, Lord, when with the abundance of things we possess,
> we have lost our thirst for the waters of life;
> having fallen in love with life, we have ceased to dream of eternity;
> and in our efforts to build a new earth,
> we have allowed our vision of the new Heaven to dim.
>
> Disturb us, Lord, to dare more boldly,
> to venture on wider seas where storms will show your mastery;
> where losing sight of land, we shall find the stars.
> We ask you to push back the horizons of our hopes; and
> to push into the future in strength, courage, hope and love.

DEDICATION

I dedicate this work to my mentors and champions.

What you find in these pages and in me is
due to their generosity of knowledge, advice, and care.
In repayment, I offer my hand to assist and guide others on the path.

May all who sow generously, reap generously.

Ron Lipinski, my definitive mentor, taught me how to derive an equation and design an experiment. He is an altruistic and wise visionary that I am privileged to call friend. *Ad astra!*

Jeff Yu and **Stuart Shaklan** introduced me to interferometry and oversaw my bachelor's thesis. Stuart is still teaching me the subtleties of interferometry.

Keith Hege's warmth to his students humanizes research. He says that a PhD means that 1) you know what is important to do and 2) you know how to do it.

Marty Levine was the first at JPL to take me under his wings.

Sharon James tutored me through numerous science fairs to convey my research with clarity and impact.

Gene Jercenovich was the first to see past my clumsiness with numbers to an aptitude for math. His advanced teaching accelerated my research career.

Jean Moone - years later I realized how critical was her role behind the scenes.

Bethann Morgan taught me most about the journey of the heart.

Jan Morgan is the first and foremost of my Champions. She nurtured my interests and rejoiced in my accomplishments. Perhaps it started when she took me to see the space shuttle land at age 4, but it could not have continued without her support and participation.

Contents

List of Figures

List of Tables

ABSTRACT

Nulling stellar interferometry may enable the discovery of earth-like planets around other stars. In nulling mode, the zero order fringe is destructive and on axis, thus canceling light from a bright source and detecting dimer off-axis features. To create deep on-axis nulls, the phase must be shifted half a wave achromatically over a broad band. The phase shift is created by adding optical path thickness with dielectric plates. Plates of different materials can balance dispersion. The nulling solutions found for TPF (infrared) and for SIM (visible) are promising.

This dissertation describes the implementation of a nulling beam combiner test bed and presents data characterizing its performance. Although the implementation was limited so that a broad band null of 10^{-4} was not attained, the test bed revealed the extreme challenges of this technique and provided very valuable lessons that will enable future implementations to be successful and more precise.

The nulling beam combiner testbed was implemented in the laboratory as a Michelson interferometer with the goal of achieving a stabilized, 10^{-4} null over a spectral region from 600 nm to 800 nm. The beam combiner system has three tiers of control. Tier 1 controls phase achromaticity by tilting optical plates and is a static control loop. Tier 2 sweeps through the white light fringe and then searches for the null as the air path drifts over minutes. Tier 3 stabilizes the null with a 300 hertz servo loop.

A scheme for active control of the optical thicknesses was developed. The phase as a function of wavelength was measured by performing PSI on a spectrally dispersed fringe. The phase was fit to a model to solve for the optical thicknesses. The optical thicknesses were then adjusted to match the ideal thicknesses of an optimized solution. This process of measuring and adjusting the optical thicknesses is

performed iteratively to achromatize the phase.

The stabilizing servo loop sensed on a grey fringe at a short wavelength. At the shorter wavelength, the fringe was 90 degrees out of phase with the main pass band resulting in a grey fringe. The grey fringe intensity is more sensitive to OPD changes.

CHAPTER 1

INTRODUCTION

The heavens declare the glory of God; the skies proclaim the work of his hands. Day after day they pour forth speech; night after night they display knowledge. There is no speech or language where their voice is not heard. Their voice goes out into all the earth, their words to the ends of the world.

Psalms 19:1-4

1.1 The Call of the Heavens

Long have we gazed at the stars in wonder and awe. Standing before them we feel minuscule and also, somehow, important. We sense that somewhere in the voice of the stars is part of our story, if we have ears to hear it. We listen, we look, we search, we admire the heavens hoping to find our place. Where among the amazing, terrible, frightening, wonderful expanse of void and lights does this piece of work, this robust, fragile, small, daring race of men embrace our past and lean towards our future?

When man first made telescopes and gazed through them, we discovered that the heavens flaunt more than our mere eyes can see. Progressively greater telescopes revealed greater and greater beauty and complexity in the skies. Our telescopes have become marvels in their own right, triumphs of engineering and manufacturing. Ground based telescopes boast 8 and 10 meter diameters[1,2], sizes that require

adaptive optics to defeat atmospheric turbulence. Space borne telescopes avoid the atmosphere and probe the heavens from X-rays to the near-infrared. Our telescopes have collectively taught us of quasars, galaxies, nebulae, high energy jets and other esoterica. Recently, they have given us rumors of planets - planets around other stars.

In 1995 and 1996, two teams announced the first discovery of Jupiter mass planets by an indirect observational technique called radial velocity measurement[3,4]. In these initial discoveries, the most unusual planet was 51 Pegasi, a planet half the mass of Jupiter and with an orbital distance one eighth of Mercury's. It's orbital period was four days. It and the other newly discovered planets were completely unlike any planetary systems we expected to find. They defied our previous models of solar system formation - models that were based on a single data point: our own solar system. The discovery of more planets followed, many of them as bizarre and unexpected as 51 Peg. Now the tally is greater than 70, most of them large Jupiter mass planets. One is confirmed by transit observation[5]. A few are systems of multiple planets[6]. Only one multi-planet system, so far, is even slightly similar to our own solar system[7].

While theorists rushed to explain how such unusual star-planet relationships could form, observationalists relentlessly pursued the heavens for direct evidence of solar systems. In 1997, the brown dwarf Gliese 229B was photographed.[8] Next came pictures of dust disks. For a few sources we have a pictures of extensive dust disks. Upsilon Andromedae is the most extensive solar system, outside of our own, that we have discovered having indirect radial velocity measurements indicating three planets.

What excitement to finally discover evidence of planets! Both the direct and indirect evidence compels us to hope that planets are abundant. An abundance of

planets would make it worthwhile to build a new generation of observing instruments tailored for planet detection.

As we begin designing new instruments, we wonder: are there many planets like our own? What will it take to find small planets like the earth? What do we hope to find on them that we haven't found already?

1.2 Finding an Earth-like Planet

Finding a planet like the earth around a distant star is very difficult because such a planet has very small mass, reflects very little light, and is very close to its host star. Several direct and indirect techniques have the potential to discover its existence[9,10].

But we will not be satisfied to merely discover its existence. We know of the existence of 9 planets in our own solar system. Our sister planets we have scrutinized over this last century beginning with the Lowell and Mt. Wilson Observatories up through robotic probes culminating in the recent Mars Pathfinder landing, Galileo, and Cassini missions. We have acquired detailed photographs of their surfaces; we have radar maps, altimetry data, wind and temperature measurements. We have learned that Venus is hellish, that Jupiter's moons hold unexpected secrets, and that water may have carved Mars. But still we are not satisfied. In all these moons and planets, we have not found water in abundance. In all these alien landscapes, we have not found life.

As we set out to search for earth like planets, we set out also to search for life or for the conditions that could nurture it. Thus we are compelled to design an instrument to search for earth like planets that measures, spectroscopically, the composition of its atmosphere, with an emphasis on several key clues[11].

The clues are evident in the comparison of spectra from the three earth like

planets in our own solar system: Venus, Mars, and Earth[12]. The spectra show the composition of the atmospheres, not the surfaces. All three have carbon dioxide abundances in their atmospheres. Only earth shows methane, water, and ozone.

An infrared, coarse resolution spectra of newly discovered earth like planet could reveal several critical features[13]. The presence of carbon dioxide predicts that the planet has an atmosphere, as opposed to being a rock, like Mercury. Water in the spectra indicates plentiful water vapor in the atmosphere and possibly large bodies of standing water. (A manned expedition could survive on a planet with water. Such a discovery might drive us to serious development of technology needed to send an expedition.) Methane indicates the presence of organic molecules. Ozone indicates oxygen, which is difficult to maintain in abundance in an atmosphere except by photosynthesis. The discovery of ozone would be indirect evidence for the discovery of life.

The potential contents of a spectra make it a grail worth pursuing. How then do we gather spectra from an earth like planet?

As an initial design point, we use an analog of our own solar system. Imagine looking at an identical twin of our solar system that is 10 parsecs away. The earth would appear to be 0.1 arcseconds away from the sun. Resolving the earth from the sun would require a 100 m diameter telescope. So large a telescope is currently out of the question, but the same resolution can be achieved by interfering light from two smaller telescopes separated by 100 m.

Using interferometry, the planet can be resolved from the star, but it is hard to distinguish amid the glaring halo of the star. At the sun's brightest wavelength, which is 560 nm, the sun would be 10^9 times brighter than the earth. At the earth's brightest wavelength, 10 μm, the sun is only 10^6 times brighter than the earth. To block the glare of the star and detect the earth requires the observation system to

have a high dynamic range. Interferometry helps with this second major difficulty. Interferometry can produce a high dynamic range of detection and can be used to block the glare of the sun.

Interfering light results in the phenomenon of fringes: bands of alternating bright and dark. In stellar interferometry we can imagine the fringes projected onto the sky. Typically, the on-axis fringe is bright, and the star is on-axis, allowing for observation of the star. However, we could shift a dark fringe to be on-axis over the star. In this case, the star's light would cancel and the planetary companion, which is slightly off-axis, would come through a bright fringe. We term this mode of operation "nulling interferometry" as first suggested by Bracewell in 1975 [14].

1.3 Nulling Interferometry

Nulling interferometry is a mode of interferometry in which the main, on-axis fringe is a destructive fringe, also called a null fringe. For quasi-monochromatic fringes, this is an uninteresting case, because the coherence envelope is so long that many orders of fringes can have high visibility. For broad band interferometry, the coherence envelope is short; the visibility deteriorates quickly away from the zero order fringe. In regular mode, the central, zero order fringe is constructive and the $\pm$ 1/2 orders are destructive, but not of great depth because of the quick deterioration of visibility. In nulling mode, the central, zero order fringe is destructive with a great depth, much deeper than the $\pm$ 1/2 orders of regular mode. Pictures of white light fringes in regular and nulling mode are shown in Figure 1.1 to illustrate the quick decline of visibility; a profile of the fringes shows that the zero order null of nulling mode is substantially deeper than the null depth of the $\pm$ 1/2 orders of regular mode. For broad band interferometry, a higher contrast is achieved in nulling mode than in regular mode. High contrast translates into the high dynamic range detection

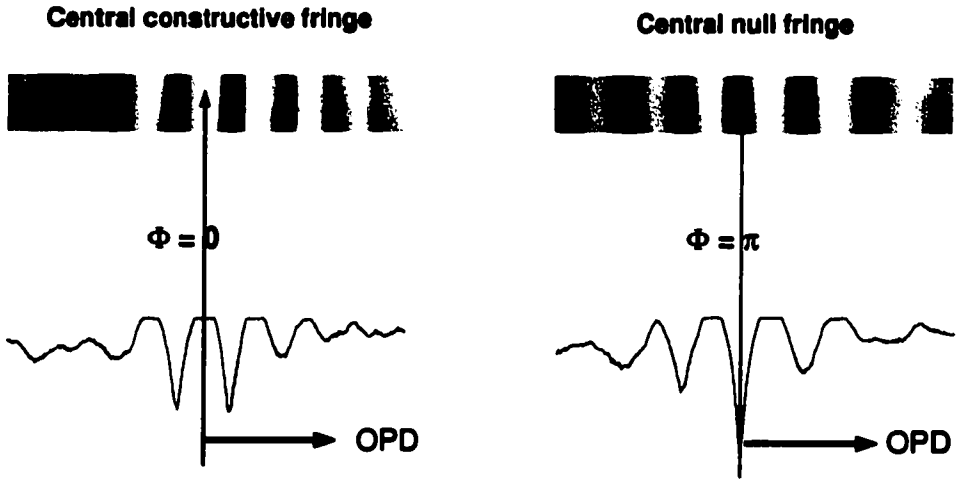


Figure 1.1 The left photograph shows normal white light interference fringes while the right photograph shows fringes in nulling mode. An arrow points to the central fringe, which is the zero OPD point. For the left image, the phase is zero at zero OPD. In nulling mode, the fringe phase is half a wave at the zero OPD point. The line plots are the logarithm of the profile of each fringe and show that the null depth of nulling mode is much greater than the null depth of the $\pm$ 1/2 orders of regular mode. The photographs were taken with a color digital camera.

needed for the planet detection.

Nulling mode fringes are the complimentary output to regular mode fringes. An interferometer has two outputs, one from each side of the beamsplitter. One output is complimentary to the other, so that one output is a bright fringe on axis while the other output is a null fringe on axis. The simple solution for nulling is to use the output with the null fringe, but the architecture of most beam combiners requires an additional beamsplitter to access the null which causes loss of photons. Loss of photons must be avoided because the planet is so dim and every photon is essential. Also, using the complimentary output allows for a phase difference between telescopes of exactly 180 degrees, but some interferometer arrays may use configurations requiring non-180 degree phase differences. Another possibility for a null is to use the output that is bright on-axis, but to put the first dark fringe

over the star. Unfortunately, at a broad bandwidth, the dark fringe visibility is too poor to provide adequate visibility. Several other approaches have been proposed for nulling interferometry, all of which must be achromatic.

Nulling interferometry for planet detection requires an achromatic phase shift. The phase shift must be achromatic, because observing must be done over a large bandwidth to collect enough photons to keep the telescopes of reasonably small size (4 m) and observing times per target reasonably short (weeks). It is difficult to create an achromatic phase shift because phase, which is some fraction of a wave, corresponds to a different physical optical path difference (OPD) for each wavelength. A half-wave of phase is 400 nm of OPD at 800 nm wavelength, but 200 nm of OPD at a wavelength of 400 nm.

Several techniques for producing achromatic phase shift have been proposed and are under development. M. Olliver, et al. exploit the Goi effect which says that 180 degrees of phase are added by going through focus[15]. Shao & Colavita[16] and Serabyn[17,18] use an out of plane reflection and roof mirrors to flip polarization in order to add a 180 degree phase shift. Angel, Burge, and Woolf suggested adding phase by adding optical thickness provided by dielectric plates[19]. Plates of two different materials balance dispersion so that the phase shift is achromatic; this technique is not restricted to phase increments of 180 degrees.

Several interferometer array configurations have been proposed using nulling interferometry. Some require a 180 degree phase shift to create the null, others require $\lambda/3$ or $\lambda/5$ phase shifts[20]. Optical path thickness is the only technique that can produce nulls at non-180 degrees and can be actively controlled to change phase in situ for on the fly array reconfiguring.

1.4 A Nulling Beam Combiner using Dielectric Plates

Details of the Terrestrial Planet Finder mission are currently under study, considering various concepts, configurations, and spectral regions. The mission, whatever it's final architecture, will take nearly fifteen years to develop, construct, and launch. The mission will be very demanding, very precise and require a great amount of technology development. The work at the University of Arizona has centered on using nulling interferometry for multiple apertures in the infrared, as proposed by Woolf and Angel[21].

The beam combiner for such an interferometer must produce a null of 10^{-6} from 7 to 20 μm. This level of null can theoretically be achieved using ZnSe and ZnS to achromatize the phase shift, as detailed in section 2.3.1. A 10^{-6} null requires that the amplitudes of the two beams must be equal to within 0.1% and the interference phase be achromatic to within 0.001 radians, which corresponds to 1 nm at 10 μm[22]. These requirements are extremely challenging and were thought overly ambitious for the first attempt at experimental implementation of this technique.

The challenges of a planet-detecting beam combiner were divided into several parts. A program to observe exo-zodiacal dust was conducted by Phil Hinz, who designed and built an infrared nulling beam combiner in a dewer for use on the 6.5 meter MMT telescope[23]. The combiner was designed to attain nulls of 10^{-4} because atmospheric seeing would limit a null to this level. Amplitude control was evaluated by Erin Sabatke through computer modeling of several control schemes[24]. I explored the issue of achromatic phase control.

I chose to evaluate phase control with a nulling beam combiner implemented in the visible where phase control could be separated from amplitude control. The visible beam combiner could demonstrate the 1 nm level path requirements of the infrared, planet-detecting beam combiner. A phase error of 1 nm in the visible

corresponds to a null of 10^{-4}. A 10^{-4} null requires amplitudes to be equal to $\pm$ 2%, which could be achievable in the visible without implementation of active amplitude control.

As an additional advantage for the visible, detectors and optical components are more abundant, less expensive, and easier to work with than in the infrared. However, the surface roughness of visible optical components is not at the 1 nm level ($\lambda/1000$). A single mode fiber is used as a spatial filter to eliminate the background irradiance due to scatter and surface irregularity.

For a laboratory beam combiner experiment, a wavefront must first be split into two beams, a task equal in challenge to the beam combiner. Consequently, a Michelson style interferometer was chosen as the beam combiner architecture because it both splits the wavefront and recombines it. As an additional advantage, the amplitudes are theoretically equal because A Michelson interferometer is double pass; each arm makes one reflection and one transmission through the beamsplitter. Plates of BK7 glass and fused silica are added to each arm to control the phase chromaticity.

The goal of the nulling beam combiner is to achieve a visible, broad band null of 10^{-4} stabilized for a second. This would validate in hardware that the phase plate technique can both achromatize and stabilize a null to the nm level, which is the level necessary in the infrared for planet detection. Further laboratory demonstration of a phase plate achromatized null for planet detection would require a beam combiner implemented in the infrared, most likely in a dewer, with the addition of amplitude control, to achieve a 10^{-6} null.

1.5 Outline of this Work

As my contribution in the quest for exo-earths, I wanted to demonstrate that the technology is feasible to the demanding levels of planet finding. Such a demonstration is beyond the scope of a single dissertation. I limit my work here to the development and demonstration of controlling phase. I develop a theoretical model for creating achromatized phase changes and for controlling the achromatization. I used the model as the basis for the design of a nulling beam combiner and then implemented the technique in hardware as a Michelson interferometer with computer control. Working with the hardware revealed the main difficulties of the technique and led to refinement of the control approach and aspects of the design.

The presentation of this dissertation follows this technique from theoretical model through experimental design to implementation and characterization. The theoretical model is developed and explored in Chapter 2. Several solutions are found for the cases of interest to several proposed missions: one solution is found for earth like planet detection in the 7 to 20 μm regime for the Terrestrial Planet Finder mission and another solution is found for a technology demonstration in the visible which could fly on the Space Interferometry Mission. Also developed is the theoretical model for control of the null: how to use the measured phase to adjust the null chromaticity and to solve for absolute phase. The control model is presented in its final, refined form.

The next chapter, Chapter 3 describes the experiment design in detail. The system is broken into three subsystems, one to measure the chromaticity of the phase, one to sweep through the fringe packet and locate the null, and one to stabilize the null in a servo loop. This chapter is similar to an engineering document, as it describes each subsystem and the resolution and sensitivities required of each of the experimental components.

Chapter 4 describes the characterization of the experiment hardware. Data is presented on the noise on the detectors, the lab environment, and the performance of each subsystem. Some subsystems did not match theoretical performance, which led to scrutinization of the hardware and the theory. The beamsplitter coating introduced an unanticipated systematic dispersion and undesired phase shift of the grey fringe. The final null results are not presented here but in the next chapter. Most of the data is taken not on an optimized null; to characterize the system behavior, the optical paths are adjusted through large ranges.

Chapter 5 shows the best null achieved in laser and white light. The hardware was limited to ultimately achieve a broad band null of 10^{-4}, but successfully demonstrated key components of the beam combiner and revealed the main challenges to any implementation of this technique. This chapter comments on the results and limitations of the experiment and suggests improvements. The lessons learned and discovered pitfalls are discussed as cautions for further implementations.

CHAPTER 2

PHASE PLATE THEORY

This chapter describes the theory of using dielectric plates to produce an achromatic phase shift. Nulling interferometry is a specific case of a π phase shift; the theory is presented for a general phase shift. In the first section, white light fringes are visualized and the analogy drawn to an achromatic lens. In the next section, the phase shift is considered as optical thickness is incrementally added starting with air path, then a single dielectric plate, and then a second dielectric plate. The phase is described by equations which are manipulated to solve for optimal thicknesses. The next section discusses the behavior as a function of bandwidth in the visible and the infrared. The following section presents the specific solutions for nulling in the visible for the Space Interferometry Mission and in the infrared for the Terrestrial Planet Finder. The final section proposes a scheme for extracting the optical thicknesses from the phase chromaticity. The measured optical thicknesses can be adjusted to meet the ideal solution. The findings of the chapter are summarized at the end.

2.1 Introduction

White light interference fringes can be thought of as the superposition of monochromatic fringes at many wavelengths. This is represented in Figure 2.1 as a three

dimensional surface of irradiance; the OPD is plotted longitudinally and the wavelength is plotted horizontally. The fringes are in phase at the zero OPD point and as OPD increases they become increasingly out of phase, decreasing coherence and subsequently visibility. At a particular OPD, the chromaticity of the irradiance can be evaluated by considering a slice of the surface in the wavelength direction.

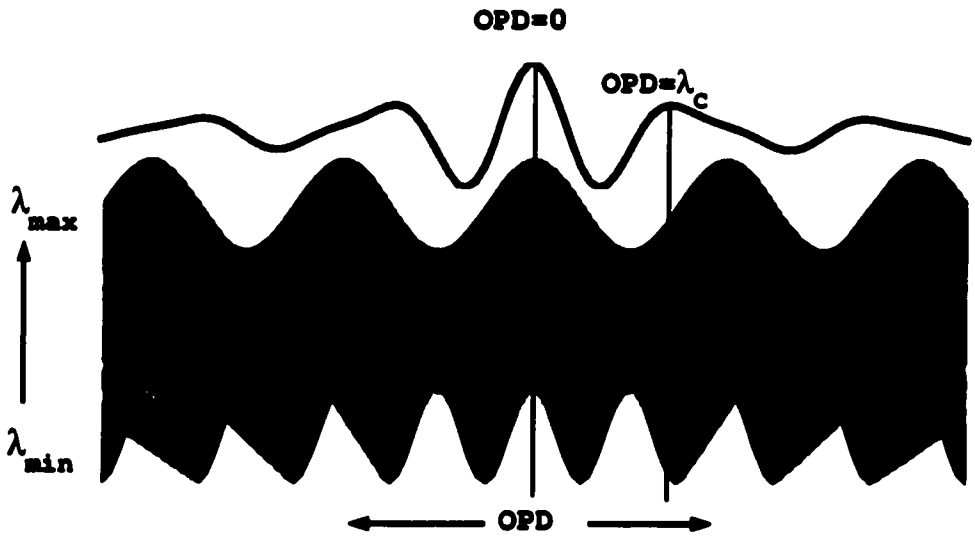


Figure 2.1 At the central fringe, the OPD is zero and fringes of all wavelengths are in phase.

Suppose a phase shift is desired. A phase shift can be achieved by adding OPD, but with the penalty of decreased visibility. A phase shift can also be created by adding OPD with a dispersive medium. The dispersion shifts each monochromatic fringe by a different amount corresponding to its wavelength. The total effect, after summing the differently shifted fringes, is to shift the coherence envelope (change the group delay) and, for chromatic OPD changes, to change the shape of the coherence envelope.

A phase shift can be achieved by adding OPD and dispersive OPD balanced in a manner similar to an achromatic lens. This analogy requires some detailed explanation. A singlet lens comes to an ideal focus for only wavelength; the other

wavelengths focus at different distances along the optical axis. This is similar to adding OPD to white light fringes: the desired phase shift is created at only a single wavelength, the other wavelengths are all shifted by different amounts. An achromatic lens uses two balanced glasses to bring two wavelengths to the same focal position. Similarly, OPD balanced by OPD through a dispersive material can produce the desired phase shift at two different wavelengths. An apochromatic lens uses three glasses to focus three wavelengths at the same position. Analogously, OPD balanced with two different dispersive materials could create the desired phase shift at three wavelengths. If even greater visibility is required, more dispersive materials can be added.

2.2 Creating a Phase Shift while Balancing Wavelength

Let us consider the dispersion balancing process for refractive materials in an interferometer. Assuming equal amplitudes between the two arms, the interferometer produces light having an irradiance described by

$$I = 2I_0(1 + \gamma \cos\phi), \tag{2.1}$$

where I_0 is the average irradiance of light, and ϕ is the phase difference between the two arms, and γ is the modulation. For this discussion, the modulation is assumed to be unity and the phase difference is dependent only on the difference in optical path between the two arms. An ideal Michelson white light interferometer has arms of exactly equal length and a compensator plate identical to the beamsplitter at the identical tilt angle so that the optical paths through the glass in each arm are equal and there is no dispersion. The output of the interferometer is a set of colored fringes in which the central fringe at the point of equal path lengths (zero OPD point) is a pure white constructive fringe, as shown in the photo of Figure 1.1. Moving away

from the central fringe, the fringes at the longer and shorter wavelengths become increasingly out of phase causing the visibility to decrease and the colors to smear.

Now assume there is no tilt to the fringes; the entire aperture of the interferometer is at the same phase. To create a phase offset ϕ, the simplest option is to increase the optical path in one arm by moving an end mirror. The phase difference is now:

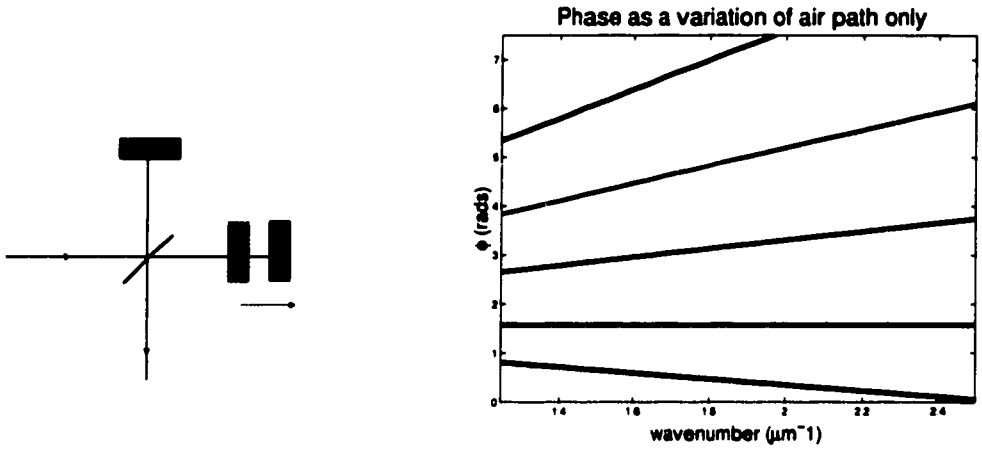
$$\phi(\lambda) = \frac{2\pi}{\lambda} s, \tag{2.2}$$

where s is the physical difference in path length through air or vacuum due to translating the mirror. In this paper we term s as "air spacing" recognizing that the spacing may be for a path in air or in a vacuum. The above phase term is strongly wavelength dependent; it is linear with wavenumber. The resulting phase chromaticity is shown in Figure 2.2 for various values of s in the visible region for an octave of bandwidth. These are the phase surfaces corresponding to the irradiance in Figure 2.1, found by moving along the phase axis and taking a slice of the surface across wavelength for a general phase shift. Note that the phase at $\pi/2$ is the zero OPD phase because of an achromatic 90 phase shift contributed by the beamsplitter. Any desired phase offset has a chromatic error.

The chromatic error may be minimized and balanced using dielectric materials. We can balance the $1/\lambda$ term by adding optical thickness to the other arm of the interferometer. We can't add air: that would result in the original zero OPD. We can add optical thickness by adding glass. The phase difference is then

$$\phi(\lambda) = \frac{2\pi}{\lambda}(s + n(\lambda)t), \tag{2.3}$$

where t is the glass thickness difference and $n(\lambda)$ is the index of refraction. The optical thickness is precisely balanced at one particular wavelength when $s = -n(\lambda_0)t$. The configuration and resulting chromaticity is illustrated in Figure 2.3. In the



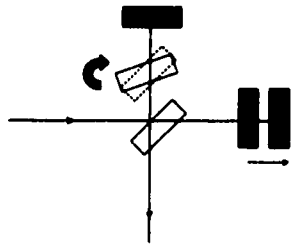
(a) (b)

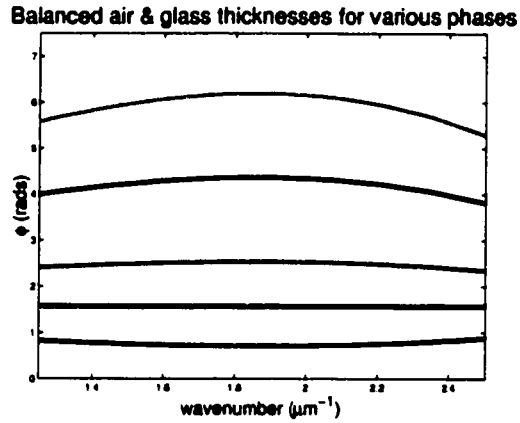
Figure 2.2 A phase shift can be created by translating one of the mirrors of the interferometer. This physical optical path difference corresponds to a different phase shift for each wavelength. The phase at various mirror spacings is plotted.

figure, the optical thickness is added by tilting the compensator plate so that the optical path through the compensator glass is longer. Equations and an illustration describing the optical path length as a function of tilt angle are in section 3.3.4. The selection of s and t may be optimized for minimal wavelength errors. The remaining phase error is due to the nonlinear dispersion of the glass.

We can balance this dispersion of the glass or dielectric by using a second dielectric.

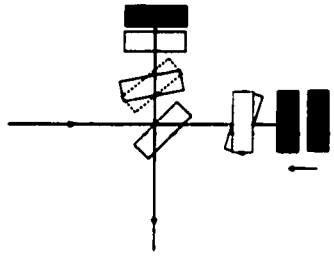
The two dielectrics are carefully selected to balance dispersions to minimize the residual wavelength effects, just as with the glass selection for an achromatic lens. The three thicknesses (the air spacing and the two glass compensators $t_1 n_1(\lambda)$ and $t_2 n_2(\lambda)$) must be balanced to achieve the desired phase and to minimize the phase

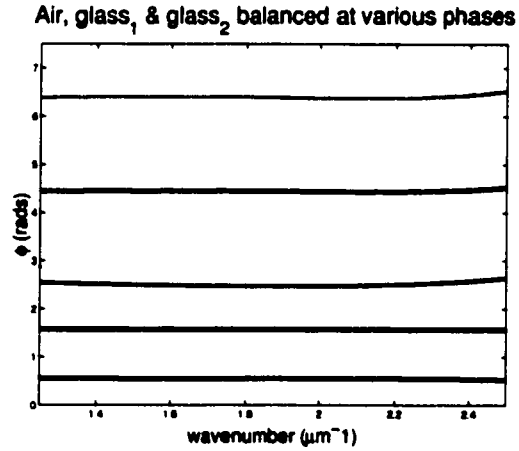




(a) (b)

Figure 2.3 A phase shift is introduced by varying both the mirror translation and the tilt of the compensator plate. The linear variation of phase with wavelength is balanced by the plate and mirror translation, leaving a residual error due only to the nonlinear dispersion of the glass plate.





(a) (b)

Figure 2.4 A second set of glass plates is added. A net phase shift is produced by balancing the tilts of the two types of plates and the mirror translation. Residual error with phase is now third-order.

error. The phase difference is

$$\phi(\lambda) = \frac{2\pi}{\lambda}\left[s + n_1(\lambda)t_1 + n_2(\lambda)t_2\right]. \tag{2.4}$$

A Taylor expansion of the phase as a function of λ provides a closed-form solution for 2 plates that fully corrects $\phi(\lambda)$ to second order:

$$\begin{aligned}
\phi(\lambda) &= \phi(\lambda_0) + (\lambda - \lambda_0)\phi'(\lambda_0) + \tfrac{1}{2!}(\lambda - \lambda_0)^2\phi''(\lambda_0) + \cdots \\
\phi(\lambda_0) &= \tfrac{2\pi}{\lambda_0}[s + n_1(\lambda_0)t_1 + n_2(\lambda_0)t_2] \\
\phi'(\lambda_0) &= \tfrac{-2\pi}{\lambda_0^2}[s + n_1(\lambda_0)t_1 + n_2(\lambda_0)t_2] + \tfrac{2\pi}{\lambda_0}[n_1'(\lambda_0)t_1 + n_2'(\lambda_0)t_2], \\
\phi''(\lambda_0) &= \tfrac{4\pi}{\lambda_0^3}[s + n_1(\lambda_0)t_1 + n_2(\lambda_0)t_2] + \tfrac{-4\pi}{\lambda_0^2}[n_1'(\lambda_0)t_1 + n_2'(\lambda_0)t_2] \\
&\quad + \tfrac{2\pi}{\lambda_0}[n_1''(\lambda_0)t_1 + n_2''(\lambda_0)t_2],
\end{aligned} \tag{2.5}$$

where the derivatives are with respect to λ. By setting the desired phase shift to ϕ_{offset} and the first and second derivatives of phase to zero, a solution for the two glass thicknesses and air spacing is found:

$$\begin{aligned}
t_1 &= -\frac{n_2''}{n_1''}t_2 \\
t_2 &= \tfrac{1}{2}\frac{1}{n_1' - n_2'\frac{n_1''}{n_2''}} \\
s &= \phi_{offset}\frac{\lambda_0}{2\pi} - n_1(\lambda_0)t_1 - n_2(\lambda_0)t_2.
\end{aligned} \tag{2.6}$$

Substituting these values into the next term of the expansion, the third derivative with respect to λ, gives the residual phase error:

$$\phi''(\lambda_0) = \frac{2\pi}{\lambda_0}[n_1'''(\lambda_0)t_1 + n_2'''(\lambda_0)t_2]. \tag{2.7}$$

This provides an adequate starting point, but an optimization routine can find a better solution. The residual phase error is now third order and rather small, as shown in Figure 2.4. The residual error due to imprecisely matched dispersion can be minimized by the appropriate selection of glasses.

The chromatic error can be reduced further by adding more plates of dielectric material. For multiple plates of dielectric material, the phase difference is

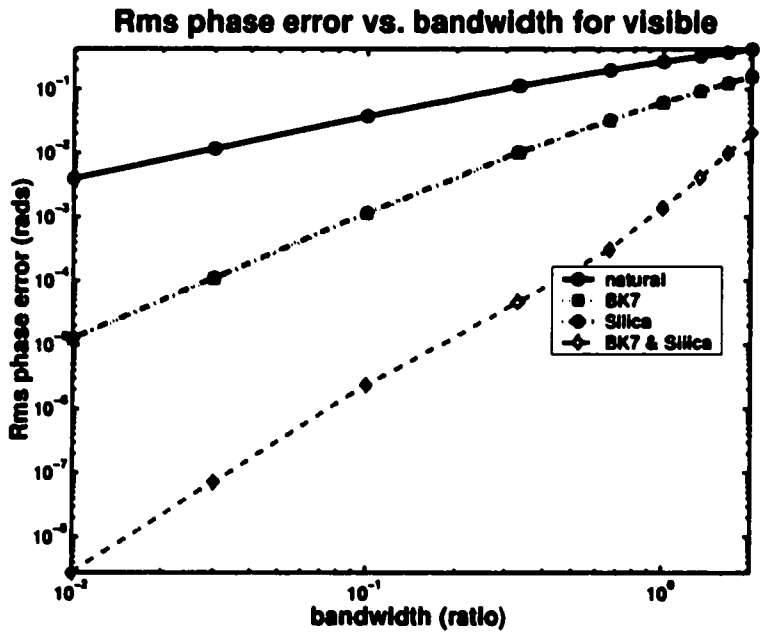
$$\phi = \frac{2\pi}{\lambda}\left(s + n_1(\lambda)t_1 + n_2(\lambda)t_2 + \cdots + n_i(\lambda)t_i\right). \tag{2.8}$$

2.2.1 Bandwidth Performance

To characterize this technique, we show the behavior of a two-beam nulling interferometer as bandwidth increases for a natural phase offset, an offset balanced with a single plate, and an offset balanced with two plates of glass. As the bandwidth increases, the $1/\lambda$ term dominates but is weighted by the dispersion of the material. Infrared materials, whose dispersion is lower than visible materials, should perform better at equivalent bandwidths. The rms phase error versus bandwidth is plotted in Figure 2.5 for a natural null (using only air spacing), for a single plate of material, and for two plates. We selected BK7 and silica for the visible and ZnSe and ZnS for the infrared. The bandwidth region is centered on 650 nm for the visible and 13 μm for the infrared. In the infrared, at two hundred percent bandwidth, a natural phase offset has an rms phase error of greater than 0.4 rads while two plates can yield nearly .001 rads. In the visible, two plates can yield .02 rads at two hundred percent bandwidth. The same rms phase as the infrared of .001 rads is achieved by two plates at a little less than one hundred percent bandwidth.

The chromatic phase error can be reduced by shifting the bandwidth region to where the material has lower dispersion. Optimum performance using a single plate occurs at the minima of dispersion for the material. BK7 has a dispersion minima at 1.3 μm, one of the shortest wavelengths for visible/NIR materials. Operating near 1.3 μm may provide adequate achromatic phase shifts using a single plate instead of two. In the infrared, few materials have dispersion minima at wavelengths longer than 10 μm.

(a)



(b)

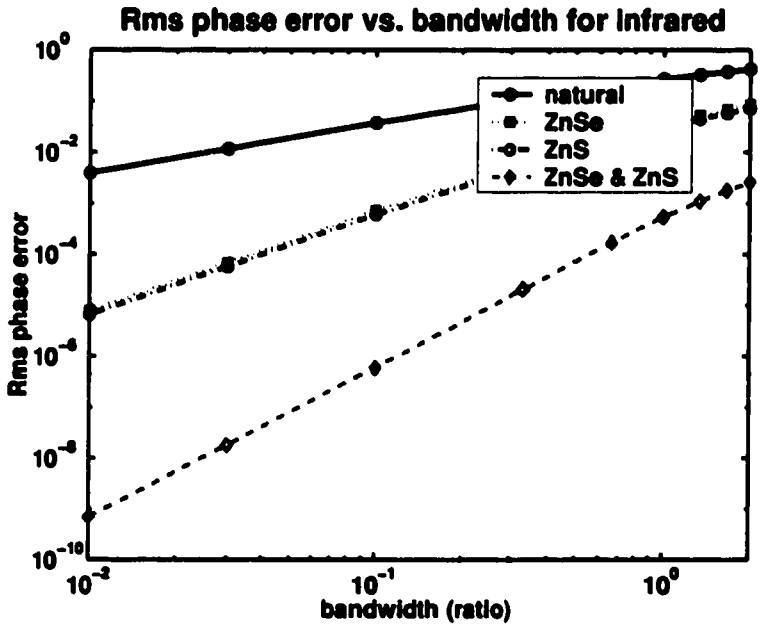


Figure 2.5 The rms phase error is plotted at various bandwidths for a phase shift produced by four optical path configurations: air space alone (natural shift), a single plate balanced by air OPD, a different single plate, and two plates balanced by air OPD. The central wavelength of the bandwidth region was 650 nm for the visible and 13 μm for the infrared.

2.3 Particular Solutions

2.3.1 Infrared Nulling

A very demanding application for achromatic phase shifts is nulling interferometry for high dynamic range, high resolution imaging. Nulling is a particular interferometry mode in which the fringes are shifted by half a wave so that the central fringe is completely destructive and cancels out the light of an on axis bright object, such as a star or an active galactic nuclei. In the infrared, a fringe contrast of 10^{-6} is desired from 7 to 20 μm for the Terrestrial Planet Finder (TPF) mission[9]. This requires a phase error of less than .001 rads over more than an octave of bandwidth for a half wave phase shift.

The available infrared materials that transmit out to 20 μm are limited. All combinations of the candidate materials were optimized for the deepest null. The materials and resulting null ratio (summed over wavelength) are listed in Table 1. The combination of ZnSe with ZnS met the requirements, producing a null of less than 10^{-6} over the entire band, as shown in Figure 2.6.

2.3.2 Visible Nulling

Nulling is planned to be demonstrated in the visible, to the level of a 10^{-4} null at 20% bandwidth[25]. Our solution is shown in Figure 2.7 and uses BK7 and fused silica. Both materials have very low dispersions, very similar second order dispersion, but different indices of refraction, as shown in Figure 2.8. These glasses easily meet the requirements for a 10^{-4} null at 20% bandwidth. In fact, the requirement could still be met using the region from 440 nm to 920 nm, a bandwidth of 70%.

For selecting glass pairs in the visible regime, an exhaustive evaluation of all pairs is possible. For expediency, we have developed a few rules of thumb for selecting glasses. First, a single plate of dielectric should be considered if the dispersion

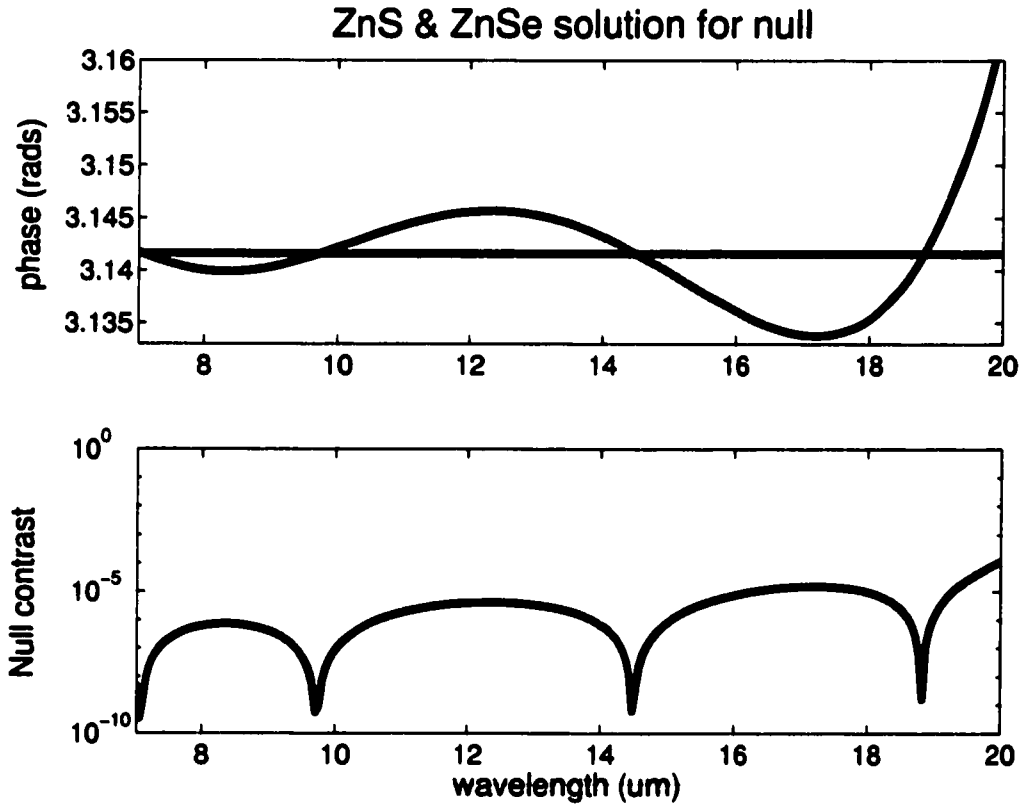


Figure 2.6 In the infrared region from 7 to 20 μm, the best pair of materials is ZnS/ZnSe. The top plot shows the phase variation with wavelength for a phase shift of π. The bottom plot shows the normalized irradiance of the null. This solution meets the TPF requirement of an rms null $< 10^{-6}$.

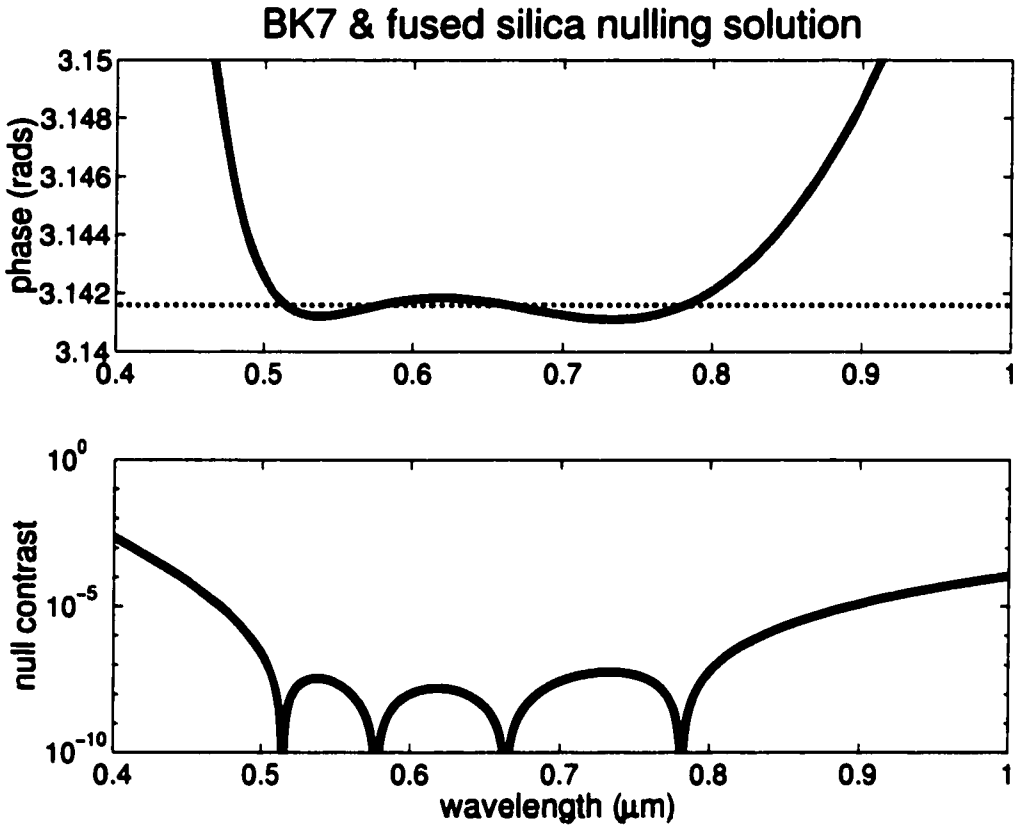


Figure 2.7 BK7 and fused silica can be balanced to produce a phase shift of π, creating a null across the visible spectrum. The top plot shows the phase variation with wavelength. The bottom plot shows the normalized irradiance of the null. This solution meets the SIM requirement of a null $< 10^{-4}$ over a 30% bandwidth in the visible. The null is $< 10^{-4}$ from 440 nm to 920 nm.

minima exists within the desired band pass. If the shift provided is not achromatic enough, then two materials should be considered. Most important for matching two glasses,the secondary dispersions, defined as $d^2n/d\lambda^2$, should be similar to allow the dispersion of the glasses to be well balanced. The indices of refraction should not be identical. The more different the refractive indices are, the larger the optical thicknesses t_1 and t_2 will be, and the looser the tolerances will be on the optical thicknesses.

2.4 Controlling Optical Thicknesses

The precision required for visible nulling is very demanding. As an example, consider two specific solutions for creating a π phase shift to yield a null of 10^{-4} over a 20% bandwidth. This null level can be reached using a single plate of BK7 of thickness 8.98 μm with an air spacing -13.77 μm; the tolerance on the glass thickness is 1.2 μm and the air path tolerance is 1 nm. A second solution uses dual plates, one of BK7 ($t_1 = 1.9153$ mm) and one of fused silica ($t_2 = -2.3453$ mm) and air thickness of 516.8 μm. The glass tolerances are 45 nm. These glass thicknesses are impractical from a manufacturing viewpoint. Instead, a plate of each material is put into each arm, one plate thicker than other by t_1 or t_2 as shown conceptually in Figure 2.9; one plate is used as the beamsplitter. One plate of each material can be thinned the appropriate amount, the tolerances are difficult to achieve through standard optical manufacturing. The glass OPD can be actively controlled by tilting the plate. Increasing the tilt of the plate increases the path the light travels through the plate and is quantified in section 3.3.4. In this way, the tolerances can be met by adding small amounts of optical thickness. Commercially available rotation stages can meet the tilt precision.

Active control of the optical thicknesses is necessary to meet the tolerances for

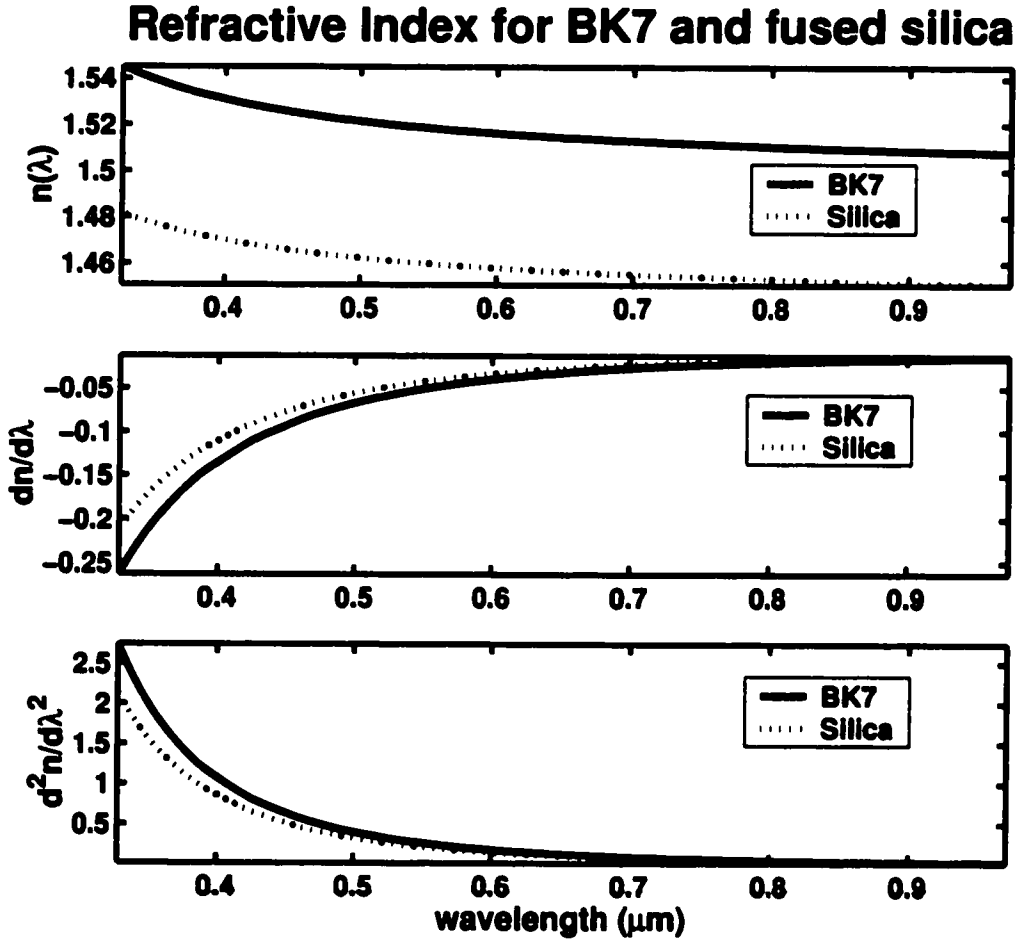


Figure 2.8 The index of refraction for BK7 and silica are plotted in the visible from 400 nm to 2 μm. The dispersions $dn/d\lambda$, shown in the middle plot, are very similar for both glasses. The secondary dispersion, $d^2/d\lambda^2$, in the bottom plot is also very similar for both glasses. These similarities in dispersion and secondary dispersion allow the two glass thicknesses to be well balanced with very little residual phase error.

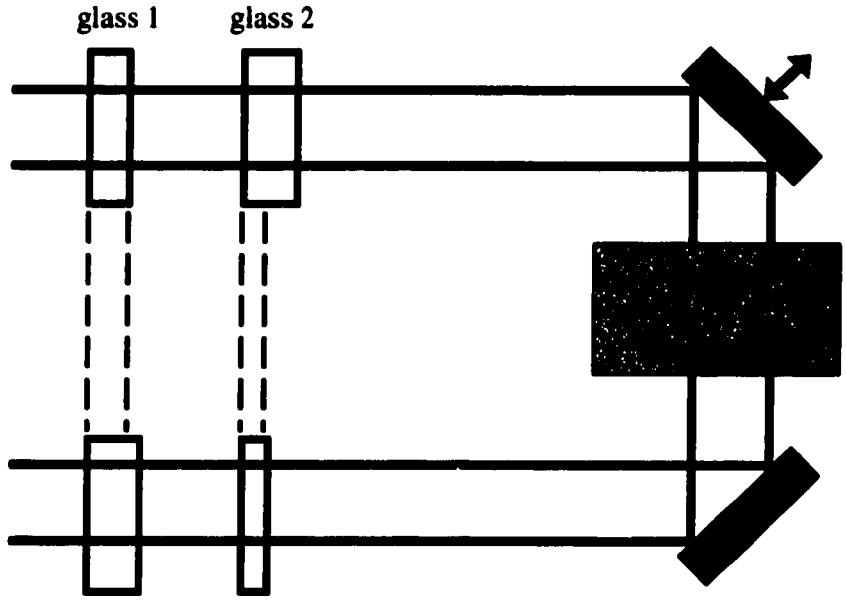


Figure 2.9 A phase shift is created by adding optical path thickness in glass to each beam of an interferometer. The glass thicknesses that are needed are too small to manufactured as a single plate ($\sim 10 - 100\mu$m). A plate of each glass type is placed in each arm so that the difference between the plate thicknesses is the desired optical path through glass.

visible nulling. The solution provides the ideal optical thicknesses, but in hardware, the optical thicknesses are not known exactly. The optical thicknesses must measured, or derived from a measurement of the system.

I suggest a scheme based on a linear systems viewpoint. The input, or the control variables, are the optical thicknesses $\vec{t} = [s\ t_1\ t_2]$. The output is the phase as a vector over wavelength $\vec{\phi_\lambda}$. We desire the phase to be flat (achromatic) at a specific value (the phase shift we are trying to produce, $\langle \vec{\phi} \rangle$). The phase plate system is described generally by

$$\vec{\phi} = \mathbf{H}\vec{t}, \tag{2.9}$$

where H is the system response function. We can measure $\vec{\phi_\lambda}$ and find $\vec{t}$ using the inverse of the system response function:

$$\vec{t} = \mathbf{H}^{-1}\vec{\phi}. \tag{2.10}$$

Once the actual $\vec{t}$ is known, the thicknesses can be adjusted until they are equal the ideal $\vec{t_0}$ given by the solution for the achromatized phase shift.

One might ask that as long as we are actively controlling the optical thicknesses, why bother solving for $\vec{t}$? A control loop could use an optimization routine to minimize the chromaticity of the phase, for example by minimimizing the first and second derivatives of the phase. Such a control loop would optimize the optical thicknesses to zero, where the total OPD is zero and there is absolutely no dispersion. This control scheme would not produce the desired, absolute phase shift. The control system could be constrained to the desired phase, but the solution space has many local minima due to the modulo 2π nature of the cosine of the interference term. A simplex hill-climbing routine would be insufficient. Rather than devote effort to an optimizer, I focus and utilize the characteristics of the phase plate system.

Next I turn to the question of how the system response function can be calculated. Since any measurement of the phase is likely to be a discrete measurement, equation 2.9 is discretized:

$$\phi[\lambda_i] = \mathbf{M}\, t_j, \tag{2.11}$$

where $\mathbf{M}$ is the system matrix, t_j is each of the optical thicknesses and $\phi[\lambda_i]$ is the phase measured at discrete wavelengths. The system matrix is derived by putting into matrix form the continuous equation 2.4 which describes the phase as a function of optical thickness and wavelength:

$$\begin{bmatrix} \varphi(\lambda_1) \\ \varphi(\lambda_2) \\ \vdots \\ \varphi(\lambda_i) \end{bmatrix} = 2\pi \begin{bmatrix} 1/\lambda_1 & n_1(\lambda_1)/\lambda_1 & n_2(\lambda_1)/\lambda_1 \\ 1/\lambda_2 & n_1(\lambda_2)/\lambda_2 & n_2(\lambda_2)/\lambda_2 \\ \vdots & & \\ 1/\lambda_i & n_1(\lambda_i)/\lambda_i & n_2(\lambda_i)/\lambda_i \end{bmatrix} \begin{bmatrix} s \\ t_1 \\ t_2 \end{bmatrix}. \tag{2.12}$$

For a single dielectric plate, the system is described by a similar equation:

$$\begin{bmatrix} \varphi(\lambda_1) \\ \varphi(\lambda_2) \\ \vdots \\ \varphi(\lambda_i) \end{bmatrix} = 2\pi \begin{bmatrix} 1/\lambda_1 & n(\lambda_1)/\lambda_1 \\ 1/\lambda_2 & n(\lambda_2)/\lambda_2 \\ \vdots & \\ 1/\lambda_i & n(\lambda_i)/\lambda_i \end{bmatrix} \begin{bmatrix} s \\ t_1 \end{bmatrix}. \tag{2.13}$$

The system matrix can be constructed using the discrete measurement wavelengths and the theoretical index of refraction of the dielectric plates. The index of refraction can be calculated with the Sellmeier formula for visible glasses[26] and other formulas for infrared materials found in Wolfe and Zisiss[27]. This is the theoretical system matrix and can be used by automated computer control to extract the optical thicknesses from the measured phase.

Instead of a theoretical matrix, the system matrix can be empirically derived from measurements to an experimental setup. Each column of the system matrix is acquired by measuring the phase at two different values of the input variable. The left most column is the difference between the phase at one value of s and

the phase measured at another value of s. The second column is found similar by measurements of phase for two different values of t_1 and the third column by measurements for different values of t_2. The empirical system matrix will include systematics and the actual refractive index of the dielectric materials.

To solve for the optical thicknesses for a measured $\phi(\lambda_i)$, the system matrix is inverted and then multiplied by the measured phase. The system matrix inverse is created using Moore-Penrose pseudo inverse[28]. The optical thicknesses are now found by

$$\begin{bmatrix} s \\ t_1 \\ t_2 \end{bmatrix} = \begin{bmatrix} \\ M^{-1} \\ \\ \end{bmatrix} \begin{bmatrix} \varphi(\lambda_1) \\ \varphi(\lambda_2) \\ \vdots \\ \varphi(\lambda_i) \end{bmatrix}. \tag{2.14}$$

By this matrix multiplication, the OPDs are extracted from the phase data. The OPDs can then be adjusted to more closely match the OPDs in the ideal solution. This process can be performed iteratively until the ideal solution is attained, and the phase is adequately achromatic.

2.4.1 Nonorgthogonality of Control Variables

A difficulty arises in controlling the achromaticity of the phase shift because the control variables are not orthogonal. The control variables are the basis set for the phase plate system. An ideal basis set is orthogonal (the dot product of any of the basis vectors are zero). However, the optical thicknesses are not an orthogonal basis set to map the phase; they are nearly colinear. This concept is illustrated in Figure 2.10 where $\widehat{s}$, $\widehat{t_1}$, and $\widehat{t_2}$ are basis vectors that point in nearly the same direction. If the basis set is orthogonalized, one vector is very long, one vector is very short, and one vector is very, very short.

This concept can be explained another way. To create an achromatic phase shift is difficult because the OPDs must be balanced to very high precision but

Table 2.1 Infrared materials were evaluated in pairs to balance a phase shift of half a wave to create a null fringe. The figure of merit listed in the table is the average of the normalized irradiance of the null fringe multiplied by 10^6 for a band width from 7 to 17 μm.

Material	CdSe	CdTe	CsBr	CsI	Ge	KBr	KCl	KRS5	ZnS	ZnSe
AgCl	82	29	11	0.9	3	49	15	4	15	75
CdSe		149	545	106	20	91	54	69	32	79
CdTe			91	186	10	40	16	397	3	6
CsBr				11	5	5	12	22	13	33
CsI					2	1.3	2	37	5	9
Ge						5	2	1.7	194	33
KBr							23	7	17	69
KCl								1.7	15	311
KRS5									1.1	1.7
ZnS										0.7

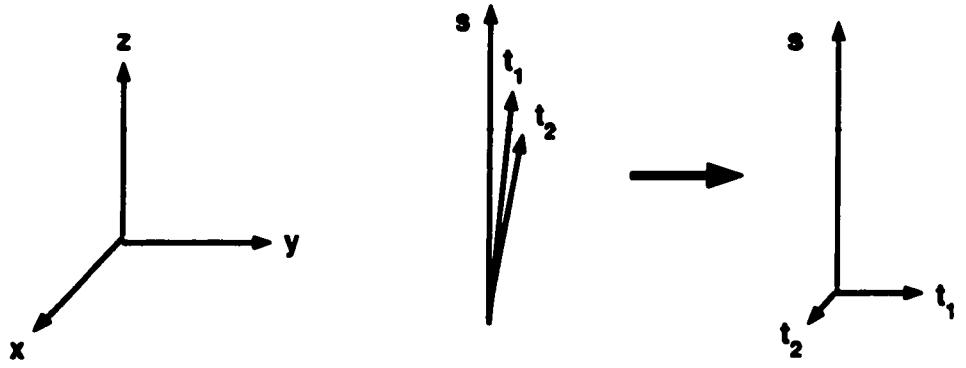


Figure 2.10 An ideal system has orthogonal eigenfunctions, as seen on the left as vectors. The input variables of the phase plate system are nearly colinear (middle). If they are transformed into orthogonal eigenfunctions, their vector lengths vary substantially (right).

also with large changes of OPD. From a systems view, the air and glass OPDs are the control variables and the degrees of freedom being controlled are the phase values at each wavelength of interest. The input variables are not orthogonal; they are nearly colinear. Adjusting only one input variable offsets the phase and adds slope that is linear with wavenumber. These first order effects must be balanced using another input variable until only the second order effects remain. When using two glasses, the second order effects must be balanced, requiring even greater precision. Consequently, the input variables must be balanced using large offsets to high precision.

This concept is quantified by examining the Singular Value Decomposition (SVD) of the system matrix. The SVD finds the eigenfunctions of the system matrix[28]. The eigenfunctions effect the system according the weight (eigenvalue) prescribed to them by the SVD weighting matrix. The weighting matrix of the SVD is a diagonal matrix whose elements w_{ii} describe the sensitivity of the system to each eigenfunction. The system is described with the SVD matrices as

$$[M] = \begin{bmatrix} u_{11} & \cdots & u_{i1} \\ u_{12} & \cdots & u_{i2} \\ \vdots & \ddots & \\ u_{1i} & \cdots & u_{ii} \end{bmatrix} \begin{bmatrix} w_{11} & 0 & 0 \\ 0 & w_{22} & 0 \\ 0 & 0 & w_{33} \\ \vdots & \vdots & \vdots \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_{11} & & \\ & \ddots & \\ & & v_{33} \end{bmatrix}. \tag{2.15}$$

In order to resolve the smallest eigenfunction, the data must have a dynamic range, at the least, calculated by divided by the smallest eigenvalue into the largest eigenvalue. The dynamic range is defined as the maximum possible value of the data divided by the smallest resolution element. For example, an 8-bit CCD has a full range of 256 possible values, and the smallest value is 1. The data, assuming no noise, has dynamic range of 256. In the presence of noise, the rms noise is considered the smallest resolution element of the data. For an rms pixel noise of 2 counts, the

dynamic range of the data is 256/2 = 128.

For the two-plate, ZnS/ZnSe solution for TPF from 7 to 20 μm, the ratio of largest to smallest to largest eigenvalues (w_{11}/w_{33} is 1.06x10^4. The required dynamic range of the phase data is 1.06x10^4. The same dynamic range is required for the two plate BK7/silica solution in the visible band pass of 600 nm to 750 nm.

A single plate solution for the visible bandwidth is also considered. A single plate of BK7 can provide a 10^{-4} null in the 600-800 nm passband. An SVD of that system matrix (equation 2.13) shows that a dynamic range of 2.06x10^3 is required. The phase data could be acquired with an SNR of 1 by using a 12 bit CCD with 2 bit read noise to acquire the phase as described in section 3.3.1.

2.5 Summary

A generalized phase shift can be produced achromatically over broad bandwidths by adding balanced optical thicknesses. Any desired phase shift can be generated. The challenging case of nulling interferometry demonstrates the potential of the technique. Specific solutions show that a null of 10^{-6} can be reached in the infrared over an octave and a null of 10^{-4} can be reached in the visible over a 20% bandwidth.

The tolerances on the optical thicknesses for these deep level nulls are very precise. The tolerances can be met only through active control of the optical thickness. The optical thicknesses can be extracted by a measurement of the phase as a function of wavelength. The phase measurement must have a high dynamic range to successfully extract all the optical thicknesses. Extracting the optical path thicknesses is difficult because, as control variables, the optical thicknesses are nearly colinear.

CHAPTER 3

DESCRIPTION OF THE EXPERIMENT

A beam combiner was implemented in the laboratory as a testbed for the phase plate technique of producing an achromatic phase shift. The testbed was designed with the goal of attaining and stabilizing a 10^{-4} null for a 20% bandwidth in the visible. This chapter describes the engineering of the beam combiner system. After an overview of the system, the specifications, tolerancing, and relevant simulations are described for each tier.

3.1 Top level requirement

The driving specification of the experiment is the goal to measure a null of 10^{-4} with a 20% bandwidth. Before considering how the null shapes the system specifications, one must first consider how to define the null. By a null of 10^{-4}, what is actually meant is a fringe contrast, where the minimum irradiance is 10^{-4} of the maximum irradiance. In characterizing fringe contrast, Michelson defined the visibility[29] as

$$\mathcal{V} = \frac{I_{max} - I_{min}}{I_{max} + I_{min}}, \tag{3.1}$$

where $\mathcal{V}$ is the visibility and the minimum and maximum irradiances are I_{min} and I_{max}. Using Michelson's definition the fringe visibility may range from zero to

one,and a 10^{-4} fringe contrast has a visibility of .9999. This being an inconvenient metric, we adopt the null contrast ratio proposed by Serabyn [30]:

$$N = \frac{I_{min}}{I_{max}}. \tag{3.2}$$

Because white light fringes decrease in contrast away from the zero OPD point, the null ratio N will also decrease. For the purposes of this experiment, the null ratio of a white light fringe will use the current null irradiance for I_{min} and the larger irradiance value of the neighboring peaks for I_{min}.

The degradation of the contrast of white light fringes is greatly dependent on the spectral bandwidth of the source. In this chapter, the source spectrum is approximated as a rect function whose width is described by the bandwidth:

$$BW = \frac{\lambda_{max} - \lambda_{min}}{\frac{1}{2}(\lambda_{max} + \lambda_{min})}. \tag{3.3}$$

A bandwidth of 100% corresponds to an octave in frequency. Two octaves correspond to 400% bandwidth. The tolerances are calculated for the spectral range from 600 nm to 800 nm which is a 28% bandwidth. This is larger than the required 20% bandwidth to allow for the slope in the cutoff of the pass band filters. Occasionally, performance extended to 100% bandwidth is also calculated.

3.2 System Overview

The goal for the Terrestrial Planet Finder (TPF) mission is a null of 10^{-6} from 7 to 20 μm. The null will have to be stabilized in the presence of vibrations to allow for long integrations on the exo-solar system. This null level requires the amplitude of each arm to be equal within 0.1% and the phase to be shifted by a half wave $\pm$.001 radians, which is about 1 nm of OPD at 10 μm. Both amplitude and phase

requirements are exceedingly difficult to meet and were thought overly ambitious for the first attempt at experimental implementation of this technique.

A beam combiner implemented for visible wavelengths could demonstrate the 1 nm level OPD requirements of the infrared, planet-detecting beam combiner. A phase error of 1 nm in the visible corresponds to a null of 10^{-4}. A 10^{-4} null requires amplitudes to be equal to 4%, which could be achievable in the visible without implementation of active amplitude control. In the visible, phase control can be separated from amplitude control.

As an additional advantage for the visible, detectors and optical components are more abundant, less expensive, and easier to work with than in the infrared. However, the surface roughness of visible optical components is not at the 1 nm level ($\lambda/1000$). A single mode fiber is used as a spatial filter to eliminate the background irradiance due to scatter and surface irregularity.

The laboratory demonstration is implemented as a Michelson style interferometer as in figure 3.1. The Michelson interferometer first splits the wavefront and then recombines it. Because a Michelson is a double pass interferometer, the two arms have equal amplitudes, and the phase control issue can be addressed without simultaneously controlling beam amplitudes.

A Michelson interferometer consists of two identical mirrors and a beamsplitter with matching compensator plate. Only a few modifications turns this into a nulling interferometer: a PZT controls one of the mirrors, a rotary stage adjusts the tilt angle of the compensator plate, and the output is spatially filtered by a single mode fiber.

A beam combiner could be designed in a single pass configuration. A single pass architecture is a possible design for an observational instrument that combines the beams from two telescopes. The wavefront was already been split by the spatial

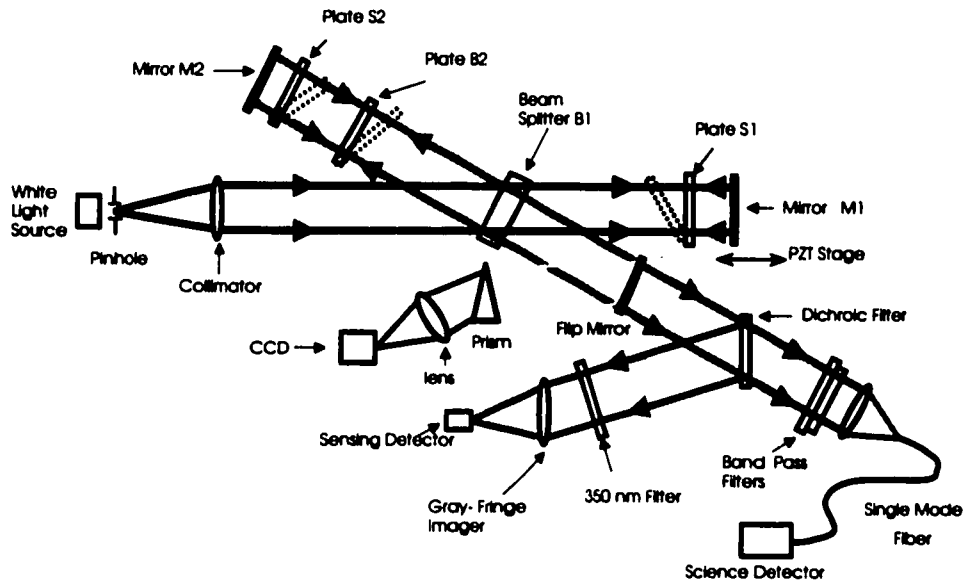


Figure 3.1 Experimental layout of the white light nulling interferometer. The output is spatially filter by a single mode fiber and detected by a PMT. A photodetector senses the grey fringe for the stabilizing servo loop. The branch with the prism and CCD measures $\phi(\lambda)$

sampling by the telescopes. In a single pass beam combiner, one beam is transmitted through the beamsplitter and the other is reflected by the beamsplitter. The beam splitter contributes a 90 degree phase shift to the transmitted beam[31]. The 90 degree phase shift is explained by conservation of energy. Assuming a completely symmetric beam combiner, when one output is a bright fringe, the other fringe is a dark fringe. When one output is grey, the other output is grey. Both outputs can be grey only if the beamsplitter imparts a 90 degree phase shift upon transmission. In a single pass architecture, a null fringe requires only an additional 90 degree phase shift to be contributed by the dielectric plates. In a double pass configuration, the phase plates must provide the full 180 degrees phase shift; all the optical thickness are twice what they would be in a single pass implementation.

In designing the experiment, one must consider not only the tolerances, but also how the null can be measured to the desired precision. Two different aspects of the null, the phase and the contrast, are measured. This provides a check of the data so that we may be convinced that the correct phase shift has been reached.

The system is composed of three tiers of control. The first tier tilts the plates to perform the achromatization. The second tier sweeps through the white light fringe to record the peak irradiance and to locate the null. The third tier is a fast servo loop that stabilizes the null. Each tier operates over different time scales. Tier 1 controls chromaticity by tilting optical plates and is a static control loop. Tier 2 searches for the null as the air path drifts over minutes. Tier 3 stabilizes the null at a 300 hertz rate.

The first tier achromatizes the null. First the chromaticity of the phase must be measured. The output of the interferometer is directed by fold mirror to a dispersing prism and then imaged onto a CCD. The dispersed fringe is then phase shifted in four discrete steps. The four frames are reduced to give the phase at

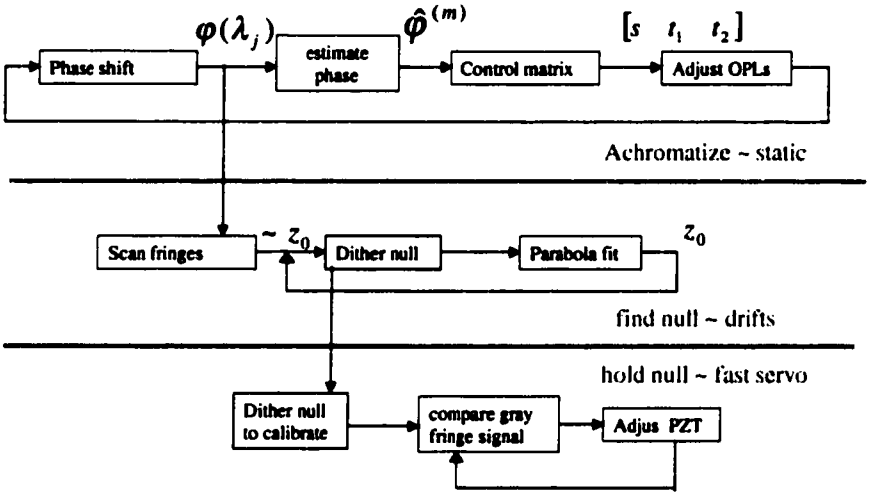


Figure 3.2 Control of the null is divided into three tiers. The first tier is a static control loop to achromatize the null. The second tier locates the center of the null as it drifts over seconds. The third tier is a fast servo loop to stabilize the null.

each pixel in the CCD. Since each pixel corresponds to a wavelength, the phase as a function of wavelength is now known. Using the matrix multiplication described in the previous chapter, the current optical path thicknesses of the two glasses and the air spacing are extracted from the phase data. The thicknesses are then changed to more closely match the "ideal thicknesses" of a previously calculated optimal solution. This process is iterated until the phase is adequately achromatic.

Then the experiment proceeds to the second tier. The second tier locates the precise center of the null. The fold mirror is removed so that the output of the interferometer is coupled into a single mode fiber for spatial filtering and then detected by a PMT. The white light fringe packet is recorded while scanned in OPD using a PZT translation stage. The highest contrast fringe is selected and then the null of that fringe is scrutinized by dithering. The dithered scans of the bottom of the null are averaged. The average null is fit to a parabola to determine the center of the null. The dither-and-fit iterates until the center of the null is found to the required

precision.

Once the center of the null is found, it is stabilized. The third tier stabilizes the null in a simple servo loop. The detector for the servo loop is not the PMT; it is a photodetector. The output of the interferometer, before entering the single mode fiber, is reflected by a dichroic beamsplitter to the photodiode. Wavelengths shorter than 450 nm are reflected to the photodiode while the longer wavelengths of the main passband, 600 to 800 nm, are transmitted through the dichroic beamsplitter and couple into the single mode fiber. At the shorter wavelength, the fringe phase is different than the phase of the main band pass. The dichroic split wavelength is chosen so the fringe at the shorter wavelength has a quarter wave phase difference from a the main band pass so that the shorter wavelength fringe is grey when the main band pass fringe is a null. The OPD is stabilized on the grey fringe detected by the photodiode. The grey fringe provides the greatest sensitivity to path change. The null fringe is transmitted through the single mode fiber and is incident on the PMT and is recorded as the science channel during the stabilization.

3.3 Tier 1 - Achromatize the Null

Tier 1 achromatizes the phase. Tier 1 must measure the phase as a function of wavelength, then use that measure to better balance the glass and air path lengths. Tier 1 iterates in a closed loop to achromatize the phase, but the control loop is static, not dynamic.

Computer control of the phase is necessary. From the resolutions stated in the previous chapter, two plates could not be balanced by hand. A single plate and air spacing, having looser tolerances, could possibly be balanced manually using a motorized rotary stage to adjust the plate tilt and a PZT translation stage to adjust air spacing.

3.3.1 Measuring Phase by Phase Shifting Interferometry

In order to control the phase chromaticity, one must measure the phase chromaticity. What must be measured is $\phi(\lambda)$, phase as a function of wavelength. A similar measurement is made in white light fringe tracking. There are several currently used techniques: 2D FFTs of channeled fringes, inverse Fourier transforms of coherence packets, and Colavita's method of modulating dispersed fringes[32]. I use a method similar to Colavita's modulation of dispersed fringes. Modulating the fringe allows the phase to be computed in a technique referred to as phase shifting interferometry (PSI)[33].

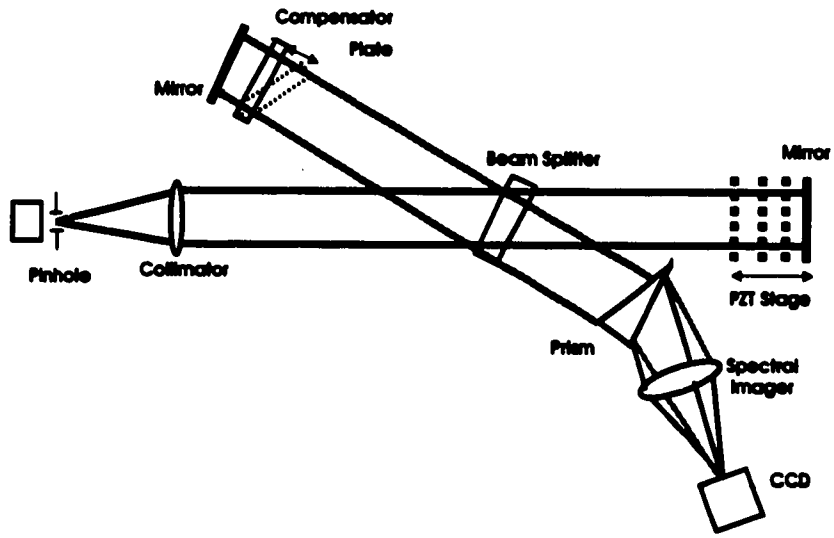


Figure 3.3 The phase as a function of wavelength is measured using PSI on a dispersed fringe. The output of the interferometer is dispersed by a prism and imaged onto a CCD. The OPD is stepped in $\lambda/4$ increments over a wave. The sequence of four frames is reduced to produce the phase at each pixel. Each pixel corresponds to a wavelength.

The light coming out of the interferometer is aligned such that there are no tilt fringes. The phase variation with λ is measured by dispersing the light onto a CCD and using phase shifting interferometry (PSI). The output is dispersed by a

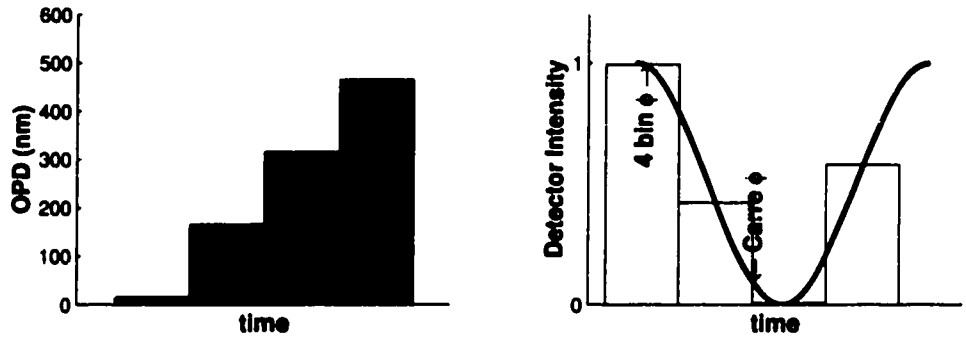


Figure 3.4 In phase shifting interferometry, the OPD is modulated in $\lambda/4$ increments (left graph). The resulting detector signal samples the cosine fringe at four locations. The 4 bucket algorithm calculates the phase of the cosine fringe at the location of the first sample. The Carré algorithm calculates the phase between the second and third samples.

prism and imaged onto a CCD, as illustrated in Figure 3.3. Only a single row of pixels is needed because there is no tilt to the interference fringe. The calibration of this subsystem is described in Section 4.2.1. The phase of the dispersed fringe is modulated over a wave by dithering a mirror with a PZT while a sequence of images is taken. The images are processed with Carré's algorithm for reducing phase shifting interferometry data. The resulting data is the phase of the interferometer at each pixel, where each pixel corresponds to a wavelength. This is an unusual form of phase shifting interferometry because each pixel sees a different phase shift. A 90 degree phase step at 600 nm results in a 120 degree step at 800 nm.

In the commonly used 4-bucket method of phase shifting interferometry, the phase is incremented by quarter wave increments, and the intensity of light read at each increment, as illustrated in Figure 3.4. The sequence of four intensity data

points is reduced to yield the phase of light:

$$\begin{aligned} I_1(x) &= I_0\left[1+\gamma\cos\left(\varphi(x)\right)\right] \\ I_2(x) &= I_0\left[1+\gamma\cos\left(\varphi(x)+\tfrac{1}{2}\pi\right)\right] \\ I_3(x) &= I_0\left[1+\gamma\cos\left(\varphi(x)+\pi\right)\right] \\ I_4(x) &= I_0\left[1+\gamma\cos\left(\varphi(x)+\tfrac{3}{2}\pi\right)\right] \end{aligned} \tag{3.4}$$

where I_0 is the average intensity, γ is the interference modulation, and φ is the unknown phase. The phase step is in increments of 90 degrees, $\delta = \pi/2$. φ is the phase at the first data point of the sequence. The phase on a single pixel detector is given by:

$$\varphi = \arctan\left(\frac{I_4 - I_2}{I_1 - I_3}\right). \tag{3.5}$$

The four-bucket technique is used because of the simplicity of its phase calculation. There are a multitude of techniques and reduction algorithms that use various phase increments δ.

Typically, PSI is performed with a monochromatic source and an imaging detector, such as a CCD, and the physical step ΔOPD corresponds to the same phase δ for each pixel. For this experiment, the phase is modulated by 4 steps of the same physical distance. This physical step is a different fraction of a wave, e.g. different phase δ, for each wavelength. Each wavelength corresponds to a different pixel, so the step increment ΔOPD corresponds to a different phase δ for each pixel. The Carré algorithm is well suited to reducing this data because the phase shift δ is calculated for each pixel; δ does not have to be known or calibrated as with many of the PSI reduction algorithms. And the phase step can be different for each pixel. Colavita, et al. have also performed PSI on dispersed fringes, but use a 4 bin method and achieve the same δ by using a read-out time tailored for each pixel.

The Carré algorithm uses a sequence of four data points to calculate the phase increment between each data point[34]. The same increment must be used for each of the four samples. The four data points have phase $\varphi - 3\delta/2$, $\varphi - \delta/2$, $\varphi + \delta/2$, and $\varphi + 3\delta/2$. The computed phase is not the phase at the first sample but rather the phase halfway between the second and third sample points. The intensity of the four data points are

$$\begin{aligned} I_1(x) &= I_0\left[1 + \gamma\cos\left(\varphi(x) - \tfrac{3}{2}\delta\right)\right] \\ I_2(x) &= I_0\left[1 + \gamma\cos\left(\varphi(x) - \tfrac{1}{2}\delta\right)\right] \\ I_3(x) &= I_0\left[1 + \gamma\cos\left(\varphi(x) + \tfrac{1}{2}\delta\right)\right] \\ I_4(x) &= I_0\left[1 + \gamma\cos\left(\varphi(x) + \tfrac{3}{2}\delta\right)\right] \end{aligned} \tag{3.6}$$

where I_0 is the average intensity and γ is the modulation of the interference. The phase at a particular pixel is calculated to be

$$\varphi = \arctan\left\{\frac{\sqrt{|[(I_1 - I_4) + (I_2 - I_3)]\,[3(I_2 - I_3) - (I_1 - I_2)]|}}{|(I_2 + I_3) - (I_1 + I_4)|}\right\}. \tag{3.7}$$

Due to the arctan, the phase is between 0 and $\pi/2$ and has ambiguities. The ambiguities are clarified depending on the sign of the quantities:

$$(I_2 - I_3) = \left[2I_0\gamma\sin\tfrac{1}{2}\delta\right]\sin\varphi \tag{3.8}$$

$$(I_2 + I_3) - (I_1 + I_4) = \left[2I_0\gamma\cos\tfrac{1}{2}\delta\sin^2\tfrac{1}{2}\delta\right]\cos\varphi. \tag{3.9}$$

These quantities are proportional to $\sin\varphi$ and $\cos\varphi$. The correct quadrant of the phase is determined by the sign of these quantities according to the table 3.1.

The data reduction, including phase ambiguity correction, produces the phase modulo 2π. The phase is then unwrapped so that it is continuous. The unwrapping algorithm starts at the left of the phase data and proceeds sequentially through the phase data looking for 2π jumps and removing them by adding 2π to the remainder

Table 3.1 The arctan of the Carré produces a phase ambiguity. The adjustment to the phase is determined by the sign of the $\sin\varphi$ and $\sin\varphi$ terms of equations 3.9.

$\sin\varphi$	$\cos\varphi$	adjusted phase
+	+	φ
+	-	$\pi - \varphi$
-	-	$\pi + \varphi$
-	+	$2\pi - \varphi$
0	$\forall$	2π
+	0	$\pi/2$
-	0	$3\pi/2$

of the phase data. The final calculated, unwrapped phase is not known absolutely but only known relative to the first phase datum, which is modulo 2π. This leaves an ambiguity of 2πm, m is an integer, in the absolute value of the phase.

The unwrapping process also places limits on the noise. Noisy phase data can lead to false unwrapping for jumps in the data of slightly under π. Phase noise of 3 radians rms always produce unwrapping jumps; jumps are usually avoided for rms noise less than 1 radian.

The phase shifting algorithm, including automated phase ambiguity removal and unwrapping, was simulated in Matlab in order to test the implementation of the Carré algorithm and to determine it's sensitivity to noise. The phase shifting data reduction algorithm was ultimately implemented in C on the computer that controlled the lab experiment so that that phase could be calculated in real time. The C implementation was tested with the same simulated data and yielded the same results, verifying that the code worked properly.

One phase shifting simulation is shown in figure 3.5. First the theoretical phase was generated using the optical thicknesses optimized for a null and then adding a 0.875 waves of air spacing. To the theoretical phase, I added 150 nm OPD for

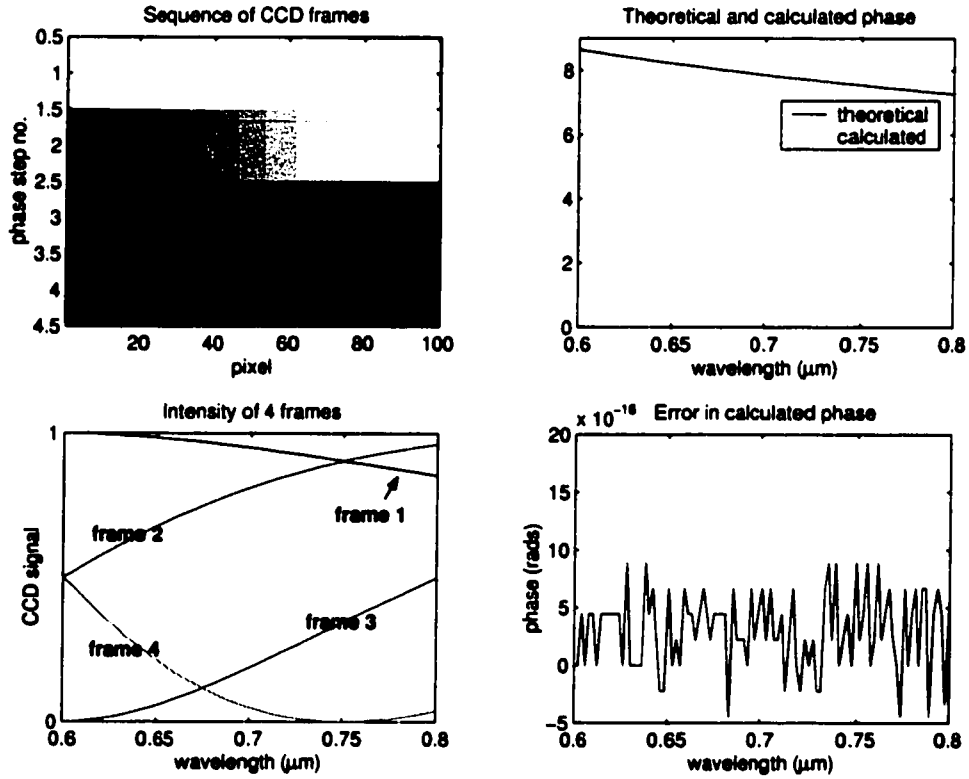


Figure 3.5 The Carré phase shifting technique was simulated. The dispersed fringe is imaged on to a row of the CCD. The OPD is stepped by 150 nm ($\lambda/4$ for $\lambda = 600$ nm) to generate a sequence of four frames(top left). The intensity of the four frames (bottom left) are reduced to by the Carré algorithm into the calculated phase. The theoretical phase is 2π greater than the calculated phase because of unwrapping effects (top right). Removing the 2π difference, the error between the theoretical and calculated phase is the floating point error of the simulation (lower right).

each step increment and took the cosine of these stepped phases to generate the intensities seen in the upper and lower left plots. The phase steps are set for $\lambda/4$ at $\lambda = 600$ nm so that the pixel signal for 600 nm corresponds to the detector signal shown in figure 3.4. The phase step at the 800 nm end of the pass band is 0.33 waves. The four intensity frames were reduced according to the Carré algorithm, the phase ambiguities removed, and the modulo 2π phase unwrapped to produce the calculated phase shown in the upper right plot. The calculated phase does not exactly match the theoretical phase; the calculated phase is 2π less than the theoretical phase. The left most calculated phase datum always lies between 0 and 2π, illustrating that the absolute phase is not calculated. The calculated phase is always an integer multiple of 2π times the absolute, true phase. The error between the true phase and the calculated phase, removing the integer multiple of 2π, is within the floating precision of the computation. Thus the Carré phase shifting technique is implemented correctly.

The simulation of phase shifting was used to explore the effects of noise on the phase calculation as detailed in the next section.

3.3.2 PSI Noise

Noise in the calculated phase will come from two main sources: noise in the pixels of the CCD and air turbulence. The pixel noise effects the intensity of each frame and is independent for each pixel. Pixel noise may be reduced by averaging. The air turbulence moves the phase of each frame; it appears as noise on the air path term of the phase. Air turbulence produces a phase offset with a slight chromatic dependence.

Pixel Noise

The pixel noise was added to the phase shifting simulation. A single pixel was used to restrict the simulation to a single wavelength and a phase step size of a quarter wave. This purposefully excludes the effect of noise on non-quarter wave step sizes which is evaluated below. The simulated intensities of the 4 bins was normalized and then noise with a normal distribution was added to each bin. The variance of the noise was increased to evaluate its effect on the calculated, or estimated, phase. The error was defined as the difference between the estimated phase and the ideal, simulated phase modulo 2π. The results are shown in figure 3.6.

The simulation used rms pixel noise from 0.001 to 0.1 of the normalized intensity. The rms error of the phase exhibits a power law dependence on the rms pixel noise. The exponent of the power law is 1.02. An 8-bit CCD has 256 possible levels measured as analog-digital-units (ADU). For an 8-bit CCD, the rms pixel noise of 0.1 corresponds to 2.6 ADU and results in 0.02 radians of rms phase error which corresponds to 2 nm OPD for a wavelength of 600 nm.

In the simulation, the mean of the phase error should be zero for normal noise, but this was not the case, as seen in the top plot. The mean off the phase error becomes significantly nonzero at an rms normalized intensity of 0.03 which corresponds to 7.8 ADU for an 8-bit CCD. When the noise became a large fraction of the intensity, the phase reduction process was effected, producing a mean error in the phase. Noise higher than rms 0.3 of normalized intensity, which corresponds to 77 ADUs, causes 2π errors in the phase unwrapping.

Pixel noise was also examined for dependence on the size of the step increment. The previous simulations of pixel noise were repeated for different step sizes. At 600 nm, the step size is 0.25 waves (90 degrees), but at 800 nm the step size is 0.33 waves (120 degrees). The results are shown in Figure 3.7. In the right plot, the rms

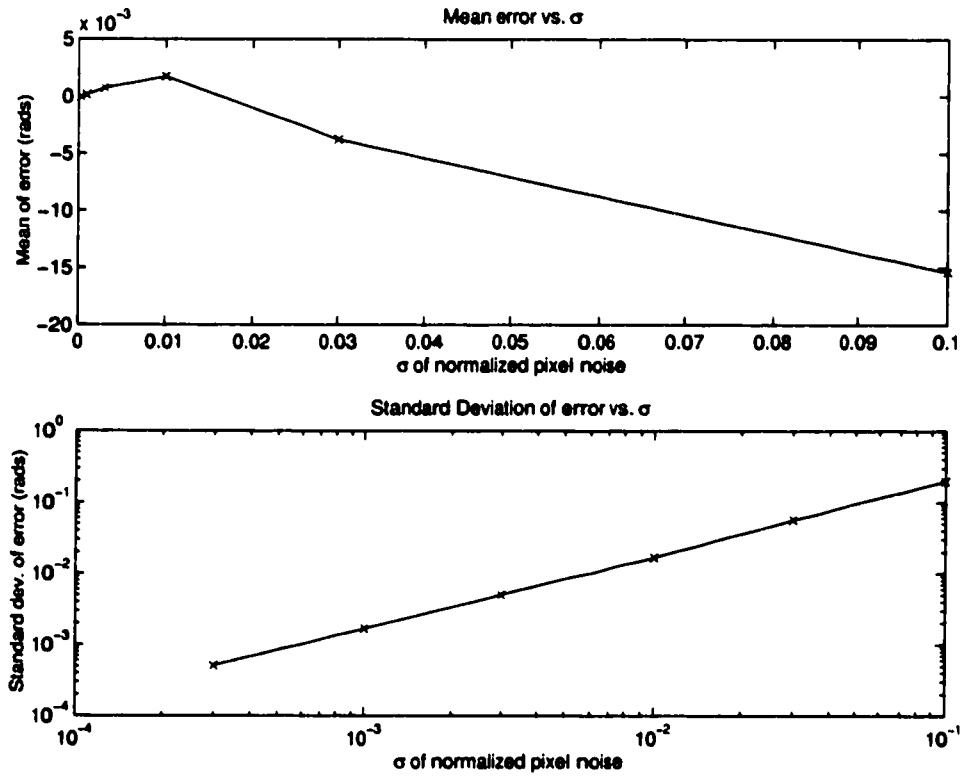


Figure 3.6 For a single pixel, normal noise was added with increasing variance and the effect on the calculated phase is considered. The error is between the estimated phase and the theoretical phase. The top plot shows the mean, which becomes biased as the noise approaches 10% of the intensity level. The bottom plot shows that the standard deviation of the error exhibits a power law with exponent 1.

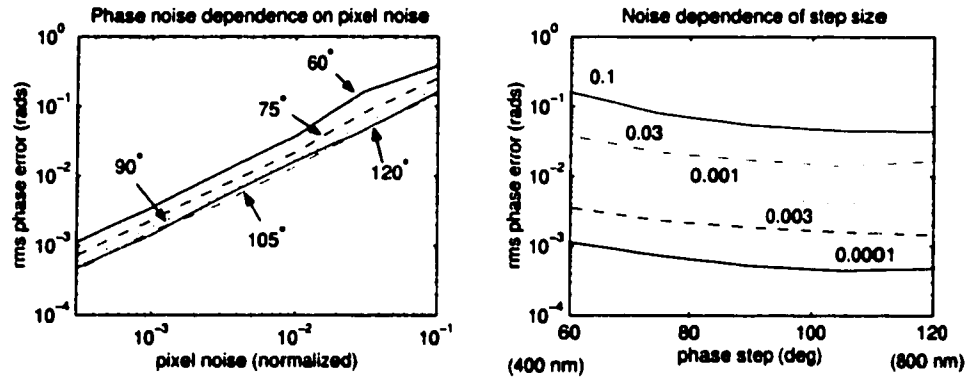


Figure 3.7 In phase shifting, the OPD is incremented by 150 nm, which is $\lambda/4$ at 600 nm. This increment at $\lambda = 400$ nm is 60 degrees and at 800 nm the increment is 120 degrees. The Carré phase reduction is more sensitive to noise for non-90 degree steps. The rms phase error due to various levels of pixel noise are plotted as a function of step size (right). The same information is presented in the left plot as rms phase error as a function of normalized pixel noise for several step sizes.

phase error is plotted as a function of step size for pixel noise from 0.1 to 0.0001, normalized. The region from 90 degrees to 120 degrees is fairly flat with only a 20% peak-to-valley variation in the rms phase error. As the step size moves away from 90 degrees, the rms phase error is more strongly effected by the pixel noise. The effect is more extreme at 60 degrees where the rms phase error is nearly double that at 105 degrees. For phase steps smaller than 45 degrees and larger than 180 degrees, aliasing makes the phase reduction unreliable.

Air Turbulence Noise

The second main source of noise is air turbulence. Air motion adds random path to the phase:

$$\phi = \tfrac{2\pi}{\lambda}\left(s + \Delta s + t_1 n_1(\lambda) + t_2 n_2(\lambda)\right), \tag{3.10}$$

where Δs is the additional air path due to turbulence. This additional air path was modeled as normally distributed noise added to the phase and produced a different

phase offset for each frame of data. The noise in the offset of each frame effects the estimated phase in its average value, its slope, and its curvature. The phase error is the difference between the theoretical phase and the phase calculated from the noisy frames. The rms of the phase error describes the effect of the air motion, which dominated by a phase offset. In addition to producing a phase offset, the air turbulence produces a chromatic effect. The chromatic effect is best described by the rms of the curvature (the second derivative) of the phase error. Figure 3.8 shows both these metrics as a function of the rms noise of the air path.

An rms air path noise of 0.1 radians, which is typical of controlled laboratory conditions, produces an rms phase error of 0.008 rads which corresponds to about a thousandth of a wave and almost 1 nm OPD at a wavelength of 600 nm. The curvature of the phase will have an rms error of 1. That means that if the phase error were fit to a second order polynomial, the quadratic coefficient would be off by one hundred percent.

Noise effects on optical thickness estimation

Noise in the phase will produce noise in the estimation of the system parameters s, t_1, t_2. Simulations of the noise in estimation were performed. Noise was added to the absolute phase, then the system parameters solved for using the matrix multiplication. Figure 3.9 shows that for a single phase plate, for a good fit to the necessary precision, and convergence of the solution, .1 or .3 rads rms is required.

A difficulty arises because the estimated phase is known only relatively, not absolutely. The absolute phase is the estimated phase plus some integer multiple of 2π. This is detailed in section 4.2.3 and demonstrated with data.

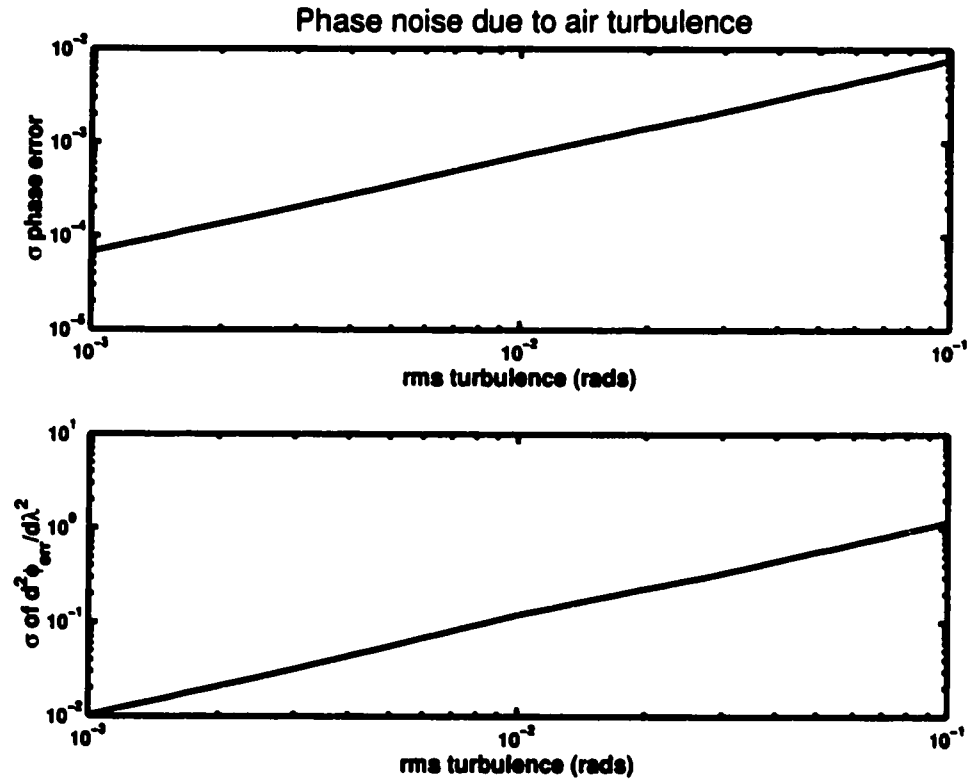


Figure 3.8 Air turbulence adds a different amount of path to each frame of a PSI measurement. The phase is effected not only in offset but also in curvature. The top plot shows the rms phase error dependence on the rms path noise due to air motion. The second plot shows the effect of air motion on the curvature of the difference between the theoretical phase and the estimated, noisy phase.

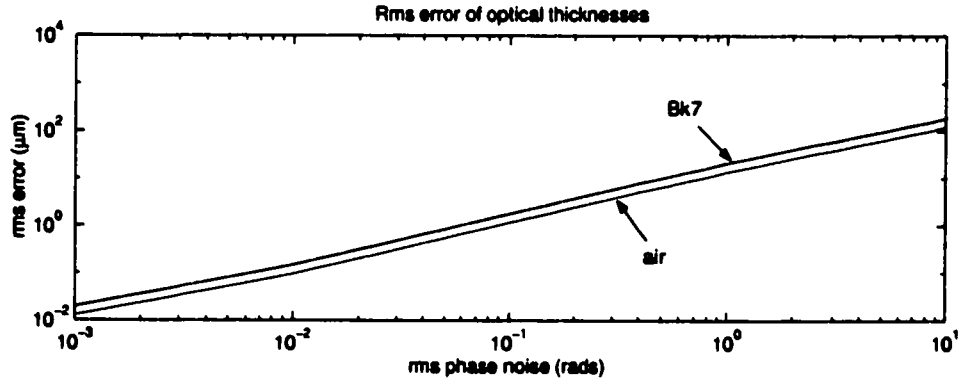


Figure 3.9 Noise was added to the phase. The matrix multiplication of equation 2.14 gives the optical thicknesses. The error in the estimated optical is plotted for a single glass plate.

3.3.3 Controlling the OPDs

The phase measurement, data fit, and changing the glass thicknesses is performed in a closed loop. The final element of the control loop changes the plate thicknesses and thereby the phase. The estimated system parameters are subtracted from the ideal solution parameters to create the error signal of the control loop. Proportional and derivative dampening factors are applied, generating the desired system parameters. The new parameters are translated into the appropriate signal to the actuators. The actuators change the tilt of the glass plates and the position of the end mirror in order to change the phase. Then the loop is begun again with the measurement of the new phase.

3.3.4 Plate Tilt

The optical path length through glass is modified by changing the tilt of the glass plate. The optical path through a glass plate is

$$l = \frac{w}{\cos\theta_t}, \tag{3.11}$$

where l is the length of the optical path, w is the width of the plate, and θ_t is the angle of light transmitted inside the plate, as illustrated in Figure 3.10. The optical path length (OPL) in glass that are required for this experiment are so small that single plate cannot be manufactured with the necessary optical quality to the required thickness. Consequently, two plates are used, one in each arm. The OPD is the difference between the thickness of the two plates. Assuming both plates have the same thickness, the desired OPD is achieved by tilting one of the plates. In this case, the OPL of one arm is w while the OPL of the other arm is l; the OPD is $l - w$. The desired OPD t is equal to $l - w$:

$$t = l - w. \tag{3.12}$$

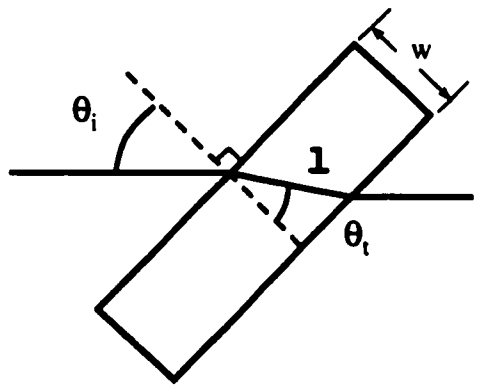


Figure 3.10 The optical path length through a tilted plate increases with the cosine of the tilt angle. This figure illustrates the notation used in equation 3.11.

The angle of the light path inside plate that will give an OPD of t found using equation 3.11:

$$\theta_t = \arccos \frac{t}{t+w}. \tag{3.13}$$

Snell's Law is used to convert the angle of light inside the plate to the tilt angle of the plate[35]:

$$\theta_i = \arcsin \left[\mathrm{n} \sin \left(\arccos \frac{t}{t+w} \right) \right], \tag{3.14}$$

where n is the index of refraction of the dielectric plate. In this way, the physical tilt of the plate is calculated that gives the desired OPD t in glass.

The tolerance on the tilt angle θ_i for an allowed tolerance on the OPD t is considered. An error in the OPD is expressed as

$$\begin{aligned} t(1+\Delta) &= l_\Delta - w \\ l_\Delta &= w + t(1+\Delta) \end{aligned} \tag{3.15}$$

where Δt is the error. This translates into an error in the tilt of the plate:

$$\begin{aligned}
\cos(\theta_t + \Delta\theta_t) &= \frac{w}{w + t(1+\Delta)} \\
\theta_t + \Delta\theta_t &= \arccos\left[\frac{w}{w + t(1+\Delta)}\right] \qquad (3.16) \\
\theta_i + \Delta\theta_i &= \arcsin[n\sin(\theta_t + \Delta\theta_t)] \qquad (3.17) \\
\Delta\theta_i &= (\theta_i + \Delta\theta_i) - \theta_i \\
\Delta\theta_i &= \arcsin[n\sin(\theta_t + \Delta\theta_t)] - \arcsin[n\sin\theta_t], \qquad (3.18)
\end{aligned}$$

where $\Delta\theta_i$ is the error in the plate tilt for an OPD error of Δt. The error can also be thought of as the tolerance of the tilt angle for a particular tolerance of OPD. The tilt angle tolerance depends on the tilt angle of the plate, as seen in the bottom plot of Figure 3.11. The tolerance of the tilt angle must be less than the resolution of the rotation stage.

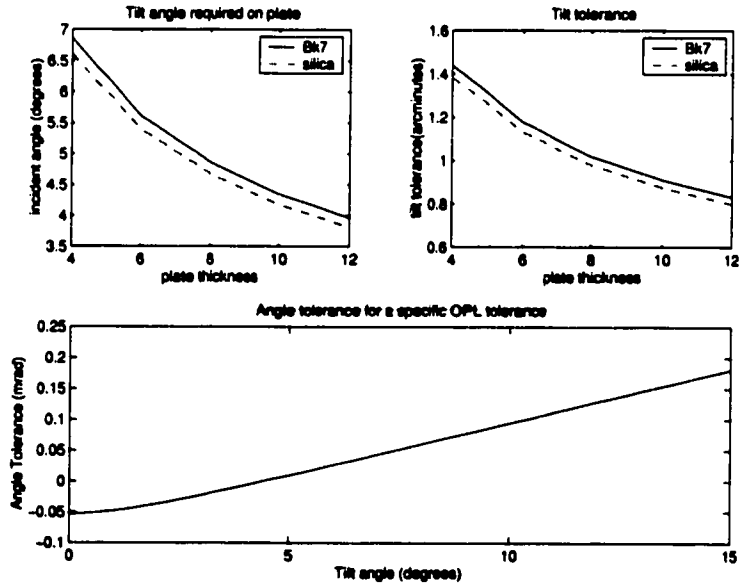


Figure 3.11 OPL is produced by tilting the glass plate. For a target OPL, the plate angle depends on thickness of the plate (upper left). For a given OPD tolerance, the tilt angle tolerance also depends on the plate thickness (upper right). The tolerance on the plate tilt angle also depends on the tilt angle (bottom).

In the two plate solution, the thicknesses were 1.9 mm for BK7 and 2.3 mm for fused silica. The angle needed to generate these OPDs, for a 6 mm thick plate is rather large, so the plates should be thinned. At large angles, polarization splitting will degrade the null. To keep polarization splitting small enough, one must keep the tilt of the plate to less than 8 degrees. The only way to do this to thin one of the plates. Figure 3.11 shows the angle as a function of plate thickness, assuming a thinning of 2.5 mm. An 8 degree angle requires a 8 mm thick plate.

Actuator adjustment of OPD

The desired change in OPD is translated into the appropriate actuator signal. For actuators, I used picomotors, which are finely threaded screws that are pushed in inch-worm style steps by a PZT. The picomotors are not strictly linear; they may have hysteresis and considerable backlash.

The actuators provide coarse OPD control. The resolution of the picomotor driven translation stage is less than 30 nm. This is too coarse to achieve the final null. The final null will be located and stabilized using a flexure translation stage driven by a PZT. The PZT will have a total travel of 10 μm and a resolution approaching 1 nm.

In a single plate solution of the null, the thickness t is 8.998 μm and the tolerance is 1.188 μm. This OPL is achieved with a plate tilt angle of 4.753 degrees. At this angle, the tilt angle tolerance is 5.3 mrad. The resolution of the picomotor driven rotation stage is 2 mrad, so the OPL tolerance can be met.

Practical limits must be set on the amount of travel allowed for each actuator during an iteration. The broad band white light fringe has a very narrow coherence envelope. Dispersed into 10 nm bins, the coherence length is 10 μm. The OPD must not be changed by more than 10 μm, otherwise the fringe visibility would decrease so dramatically that the PSI could not be performed. The implication for

the picomotor actuators that the rotation stage must not be rotated by more than 30 counts liberally (15 counts conservatively) and the translation stage cannot move the mirror by more than 400 counts liberally (100 counts conservatively).

Limits are set on the amount of travel but not on the direction of travel. The sign of physical direction can either increase or decrease the OPD. The interferometer only measures optical path *difference* between each arm, not which arm has greater path. The change in OPD will change sign when passing through the zero OPD point for one physical sign direction. The iteration of the control system should converge to the correct direction.

The position of the actuators is calculated from the values of the system parameters that are found using equation 2.14. Physically, two solutions exist: the ideal, $[s_0\ t_0]$, and the negative of the ideal, $[-s_0\ -t_0]$. To constrain and thus stabilize the control loop, only the positive solution is allowed. Section 4.2.4 presents data for increments s and t.

3.4 Tier 2 - Locate the Null

The first tier varies the glass thickness to achromatize the null. Once it is finished, the precise location of the null must be determined to within 1 nm. The null can drift due to air turbulence slowly, on a time scale of minutes. Tier 2 relocates the drifted null without the encumbrance of phase shifting tilting plates. Additionally, Tier 2 scans and records the entire white light fringe. Data is presented in Section 4.3.

The white light fringe is scanned. This provides a quick estimate of the null depth as a check of the Tier 1 achromatization and allows selection of the correct null. The PZT is scanned through 10 fringes (about 20 μm) while the fringe irradiance is read out on the science channel, which is a PMT. The data must be synchronized with

PZT scan position to determine absolute, not relative, position. The fringe scan is inspected and the deepest fringe selected as the null. The PZT location at the deepest fringe is used as the starting point of the fine resolution null search.

The PZT is sent to the null position and then dithered back and forth over a fraction of a wave while the PMT is read. Initially, the fraction of the wave can be coarse, up to half a wave, and gradually narrowed as the null location is refined. The dither data is averaged and fit by least squares to a parabola. A parabola is an adequate approximation of the bottom of the cosine fringe. The center of the parabola is set as the calibrated null position. The uncertainty of the central position is calculated. When the uncertainty is 1 nm, the control passes to Tier 3.

3.5 Tier 3 - Grey Fringe Control

Tier 3 is a fast servo loop that stabilizes the null. As an observing instrument, a beam combiner would have to stabilize the null for minutes or even hours to allow for sufficient integration. However, for this experiment, one second is considered an adequate demonstration of stabilization.

Most of the drift or change in the system comes from high frequency influences: turbulence and vibration. We have taken steps to isolate the experiment from the lab environment to minimize those effects. The characterization of the lab environment is described in section 4.1.

The null is stabilized in a simple servo loop. The loop locks onto the intensity of the fringe and then adjusts the OPD with a PZT translation stage to maintain that intensity. The bottom of a null fringe provides a poor control signal because has (hopefully) very low light levels and presents a challenging SNR on which to control. Also, the bottom of a cosine fringe is not very sensitive to changes in OPD; the slope of the change in intensity with change in OPD is at a minimum, $\Delta I/\Delta OPD = 0$.

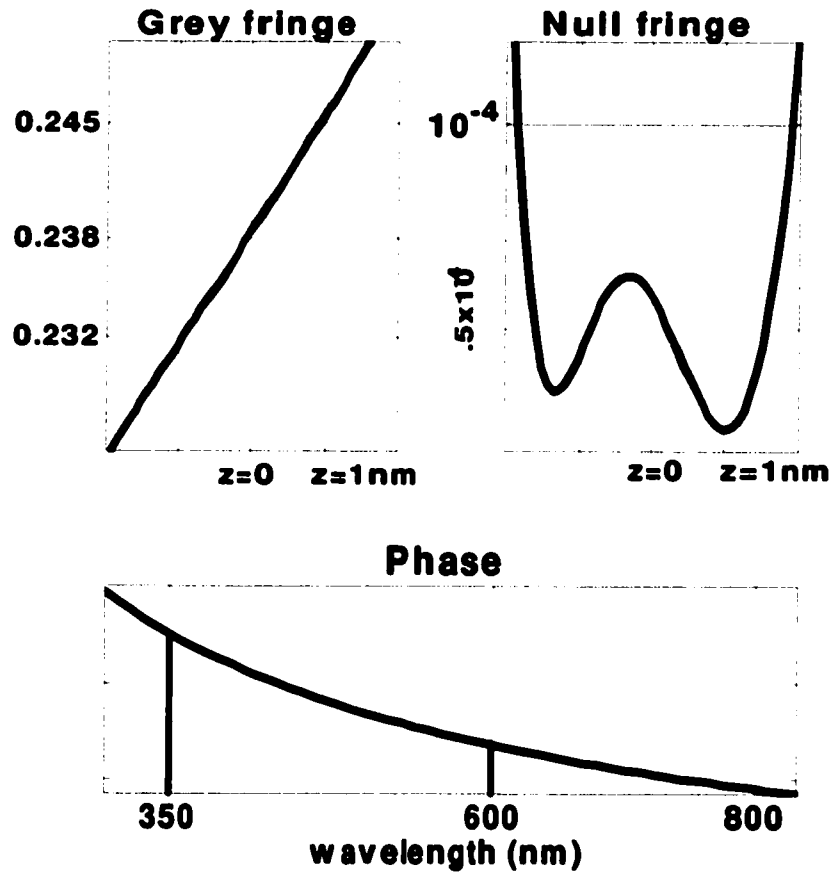


Figure 3.12 The grey fringe is more sensitive to OPD changes as seen by comparing the top two plots of normalized fringe intensity over. The right plot shows the null fringe where 1 nm of OPD will give a 10^{-4} increment of intensity. The left plot shows the grey part of the fringe where 1 nm of OPD produces an increase of 0.013 in the normalized intensity. The grey fringe is a quarter wave phase different from the null fringe. The quarter wave phase difference is achieved at a shorter wavelength. The bottom plot shows the phase shift created by the dielectric plates. At the main band pass of 600 - 800 nm, the phase shift is π the fringe is at a null. At 350 nm, the phase shift is $3\pi/2$ so the fringe is grey.

The ideal control signal comes from the fringe where the slope is greatest. The slope of the fringe is greatest a quarter of a wave away from the null at the grey fringe. The intensity over 2 nm of OPD is shown in the top two plots of Figure 3.12 for a null and a grey fringe.

To sense the grey fringe, a shorter wavelength is used. The shorter wavelength is chosen so that its phase is a quarter wave shifted from null. The bottom plot of Figure 3.12 shows the phase due to the dielectric plates over a broad bandwidth. The beamsplitter is assumed to be ideal with no dispersion and no contribution to the phase. At the main band pass of 600 - 800 nm, the phase shift is π so the fringe is at a null. At 350 nm, the phase shift is $3\pi/2$ so the fringe is grey. The fringe was measured simultaneously at the shorter wavelength and the main pass band. The results are presented and discussed in section 4.4.2.

The detector for the servo loop is not the PMT; it is a photodetector. The output of the interferometer, before entering the single mode fiber, is reflected by a dichroic beamsplitter to the photodiode. Wavelengths shorter than 450 nm are reflected to the photodiode while the longer wavelengths of the main passband, 600 to 800 nm, are transmitted through the dichroic beamsplitter and couple into the single mode fiber. At the shorter wavelength, the fringe phase is grey when the main band pass fringe is a null. The OPD is stabilized on the grey fringe intensity detected by the photodiode. The null fringe is transmitted through the single mode fiber to the PMT and is recorded during the stabilization. The PMT is the science channel and the photodiode is the sensing channel. The optical source is a battery powered incandescent source which is very stable, so intensity fluctuations, which could cause errors, should not be present.

When control initially proceeds to Tier 3, Tier 2 passes to Tier 3 the location of the null. Tier 3 performs a dither centered on the null location while reading

the photodiode signal of the grey fringe intensity. The dither is 1 nm on either side of the null. It sets the maximum and minimum intensities that are allowed for the null stabilization. The dithered signal is averaged to give the target signal for stabilization. Once the dither is complete, the servo loop locks onto the target signal for one second. The servo loop can be a simple proportional - derivative control loop.

Chapter 4

Characterization of the Beam Combiner System

The beam combiner test bed described in Chapter 3 was implemented in the laboratory. The phase measurement by PSI in real time performed successfully. Tier 2 scanned the fringe to measure the contrast and show the symmetry of the fringe. Tier 2 also successfully located the bottom of a null by fitting to a parabola. The Tier 3 servo loop stabilized a fringe for a second. This chapter describes the implementation and presents data characterizing its performance. Although the implementation was limited so that a broad band null of 10^{-4} was not attained, the test bed provided very valuable lessons for employing this technique and lessons in dealing with hardware that will enable future implementations to be more precise. The lessons learned are detailed throughout this chapter and discussed in ensemble in Chapter 5.

Phase shifting on a dispersed fringe was demonstrated in real time, and a range of step sizes were used to test the sensitivity of noise. The pixel noise of the CCD was measured and found to limit the data fit for the system parameters. The data fit modified to solve for the absolute phase as one of the system parameters. A

systematic dispersion was discovered in the system that limited the achromaticity of the null and is attributed to the anti-reflective coating. Tier 3 servo control was implemented in software at a rate of 300 Hz and stabilized the OPD to 6 nm peak-to-valley. The short wavelength fringe was not 90 degrees out of phase with the main band pass fringe due to the systematic dispersion, but the servo control did stabilize the OPD on the short wavelength fringe intensity while the null fringe data was collected on the science channel.

Figure 4.1 The interferometer was constructed on a small, floating optical bench on top of a larger, floating optical table to minimize vibration effects. A krypton lamp bulb was the source and the output was filtered through a single mode fiber. Most optics are 0.5 inches in diameter. The interferometer was constructed with a small beam and short arms to minimize atmospheric effects.

The nulling beam combiner was implemented as a Michelson interferometer. The

compensator plate was mounted on a rotary stage and one mirror mount on a PZT translation stage mounted on top of a picomotor driven translation stage. The rotary stage and both translation stages were computer controlled. The same computer gathered data from the three detectors: a CCD, a photodiode, and a PMT. The electronics interface and control was provided through custom software written in C on a linux computer system.

For the interferometer, pictured in Figure 4.1, the source was an incandescent krypton bulb. The light beam was spatially filtered by a 5 μm diameter pinhole and then collimated to a 4 mm diameter beam. The optics were 0.5 inch diameter. The two end mirrors had matching coatings of protected silver which provides greater than 98% reflectivity in order to ensure that the amplitude of the light in each arm was equal to within 4%. The beamsplitter and compensator plate were 6 mm thick Bk7 plates with MgF anti-reflective (AR) coatings. The beamsplitter coating was CVI's standard dielectric broad band 50/50 beam splitter coating for 15 degree incident angle.

4.1 Characterizing the lab environment

The Michelson, nulling interferometer was used to characterize the lab environment for mechanical vibrations, acoustic effects, and air turbulence. Air disturbances were minimized by constructing a small interferometer with short arms and enclosing it in a box made of acoustic dampening material. The environment was characterized with and without the environmental isolation measures to verify the measures were adequate and that a null of suitable depth could be attained in the environment.

4.1.1 Air motion

Air motion, if fast, can threaten the phase shifting measurements by adding a random phase on top of each frame. Also, the air motion must not be faster than the

transition time between control tiers. Slow drifts require the use of tier 2 to track the center of the null.

Air motion was measured using the nulling interferometer with white light. The output of the interferometer was coupled through a single mode fiber to a calibrated power meter using a silicon photodiode. No band pass filters were used; the fringe coherence envelope was determined only by the spectral response of the photodiode and the spectral emission of the tungsten filament bulb. Sampling was at 40 Hz. The interferometer was vibration isolated by two floating optical tables to ensure that air turbulence was the dominate effect.

First, the OPD was scanned forward and backward through several fringes to determine maximum and minimum fringe values in order to convert an irradiance value into a phase. The fringe contrast is used to place the measured irradiance in a normalized value, that is one in which the highest contrast fringe has irradiance values from -1 to $+1$. The phase is then found by taking the arccosine of the normalized irradiance:

$$I_{normalized} = \frac{I_{msr} - I_{min}}{(I_{max} - I_{min})/2} - 1 \tag{4.1}$$

$$\varphi = \arccos(I_{normalized}), \tag{4.2}$$

where I_{msr} is the irradiance data point, I_{max} and I_{min} are the maximum and minimum irradiance values of the scan through the fringe packet, and φ is the computed phase of the fringe. These equations are derived from the general expression of interference given in equation 2.1.

The path was adjusted so that the darkest null fringe was placed on the detector. Then the system was allowed to drift with the box open, and with the box closed while data was collected. The greatest sensitivity occurs on the grey fringe where $\Delta I/\Delta OPD$ is maximum. Only data from this grey fringe level of irradiance was used.

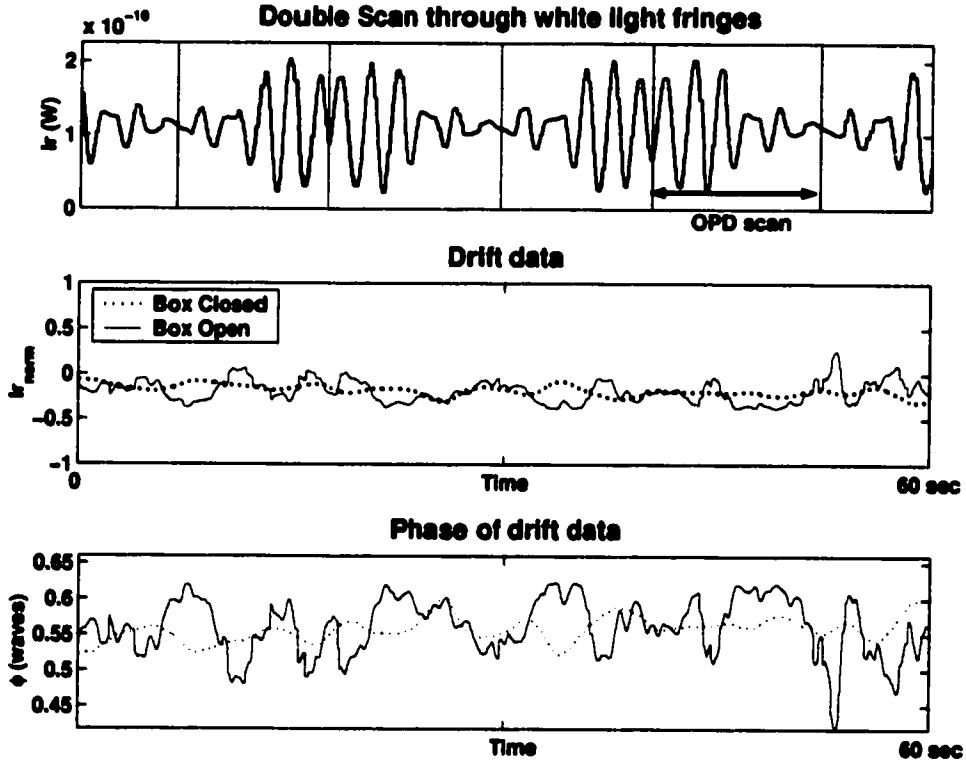


Figure 4.2 The interferometer was enclosed in a box to minimize atmospheric effects. The top plot shows scans forward and backward through the white light fringe packet. The fringe contrast of the deepest null fringe was used to normalize the irradiance of the data. Data was taken over a minute with the box sealed (dashed line) and open (solid line). The normalized irradiance was converted to phase, as shown in the lower plot. The box keeps the OPD wander to a under a tenth of a wave.

For full exposure to the atmospheric turbulence of the lab environment, largely due to air conditioning, the highest frequency OPD fluctuations appear to be at 1 Hz. This will not limit the servo loop's ability to stabilize the null. This will not limit the > 10 Hz frame rate for the phase shifting. For phase shifting, the data frames are taken at about 10 Hz, so no significant phase error expected for each frame. For averaging over 10 frames or more, a phase error of as much as 0.2 waves could be introduced. The consequences of this error is discussed in the section 3.3.1.

The enclosing box eliminates the high frequency air turbulence, after a settling time of about 15 minutes. The data shows a time scale of about 10 seconds for twentieth wave fluctuations. This could impact averaged sets of phase shifting data by adding a twentieth wave phase error over the set. This is a much smaller error than with the open box and should be acceptable.

4.1.2 Vibration

The greatest threat from vibration is high frequency excitations that could limit the servo loop. The servo loop operates at about 1 kHz, so variations less than 100 Hz could be compensated.

The vibration environment was measured by a Michelson interferometer using a laser source. The output of the interferometer was coupled through a single mode fiber to a silicon photodetector. The photodetector signal was digitized by a 12-bit analog to digital converter.

A simple HeNe laser was used as the source of the interferometer. A laser source provides the same contrast for each fringe. This is important if the vibration is larger than a fringe and the intensity variation of fringe motion is calibrated to path change (dI/dz). The greatest sensitivity to path change occurs at the gray portion of the fringe, where the slope of the fringe cosine is greatest. The fringe is position so that the interferometer is resting on the grey fringe when the vibration measurements are

taken. The data was accumulated while the experiment was not disturbed, allowing it to detect the vibrations of its environment. The interferometer was sealed inside a box to minimized air turbulence.

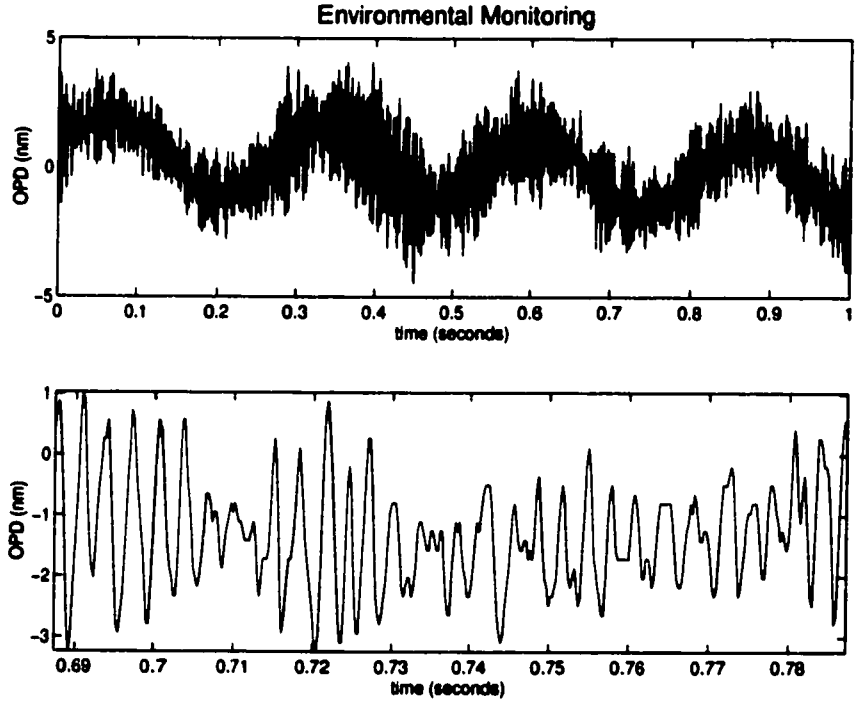


Figure 4.3 The interferometer output was monitored to observe the vibration environment. The top plot shows data taken over 1 second with slow variations of 5 nm. The lower plot is a shorter time sampling and shows 3 nm fluctuations at about 350 Hz.

The measured irradiance of the data was converted into phase using equation 4.2. The phase was converted to OPD by multiplying by the wavelength of the HeNe laser, 633 nm:

$$z = \varphi \frac{\lambda}{2\pi}, \tag{4.3}$$

where φ is the phase in radians and λ is the wavelength of the laser source in nanometers, and z is the OPD in nanometers. The data shown in Figure 4.3 as nanometers of OPD.

The data shows that the high frequency vibrations have an amplitude of about 3 nm. The standard deviation of the data is 1.5 nm, which is sufficient for a 10^{-4} null. The highest frequency vibrations are at about 350 Hz and should be controllable to the 1 nm level by the 1 kHz servo loop. As an added precaution, a second layer of vibration isolation was added by placing a small floating breadboard on top of the floating optical benches.

4.1.3 Acoustic

Acoustic effects were measured in the same way as the vibration effects with a Michelson interferometer and laser source. Acoustic noise has the potential to excite vibrations in the optical mounts; these are vibrations not transmitted through the floor and optical bench. These measurements reveal the behavior of the optical mounts and components used to make the measurement. The initial results showed that some mounts were very susceptible to vibration: those mounts were replaced. The data presented here is the final acoustic measurements after the mounts had been upgraded.

With the box open, the interferometer is susceptible to acoustic noise causing OPD fluctuations as great as 30 nm (a tenth wave) at frequencies as high as 400 Hz. This may be manageable to the servo loop, but as a precaution acoustic dampening material was added to the experiment enclosure. Even so, the most precise nulling data was acquired during off hours when vibration and acoustic noise were at a minimum.

4.2 Tier 1 - Phase shifting the dispersed fringe

The first step on the road to achieving a null is to balance the dispersion of the system. This is done by dispersing the fringe and evaluating it in spectral channels. Michelson used dispersed tilt fringes naming them "channeled spectra" in order to

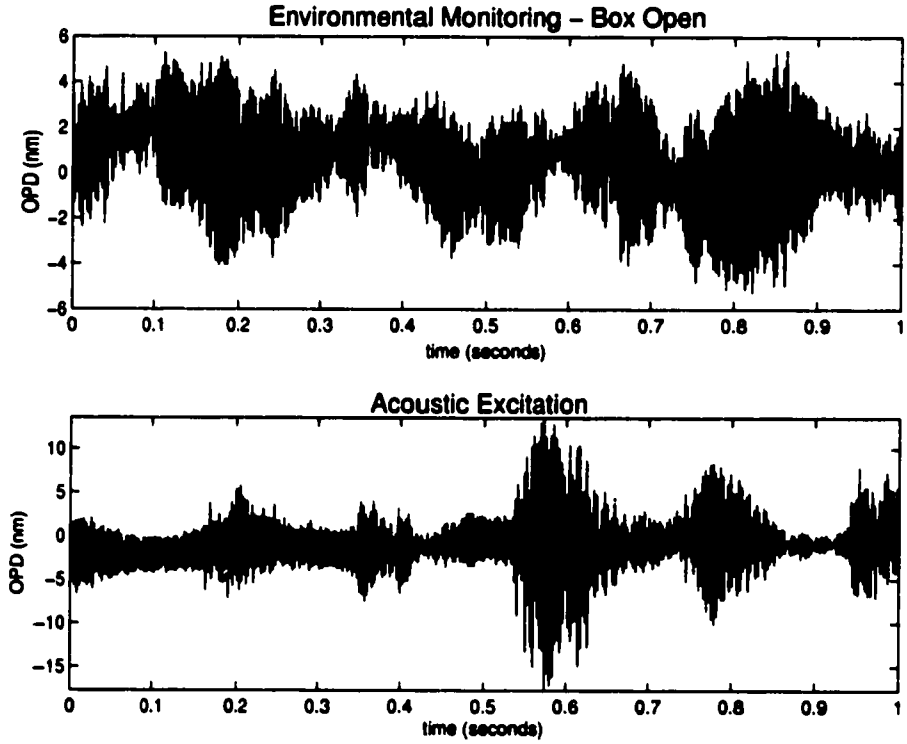


Figure 4.4 Environmental noise in the OPD when the box is open (top) and excited by talking and clapping (bottom) results in fluctuations as large as 30 nm at 400 Hz.

adjust a pair of glass wedges to dispersion balance his stellar interferometer [29]. In this experiment, all tilt is removed from the fringes. The phase of each spectral channel is determined through phase shifting, as discussed in section 3.3.1.

4.2.1 Imaging the dispersed fringe

CCD noise

The imaging detector for the dispersed fringe is an 8 bit CCD. The CCD is controlled in exposure, gain, and bias to maximize the signal to noise ratio for each data set. Noise in the CCD was measured as a function of these three variables.

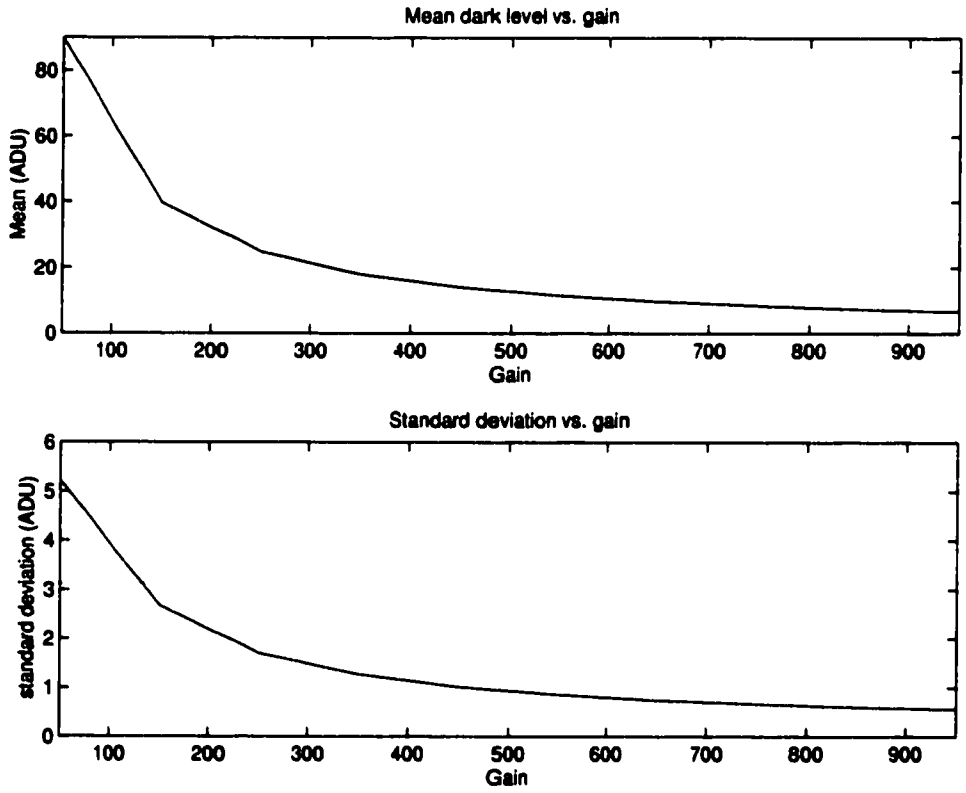


Figure 4.5 The CCD gain varies from 1, the most gain, to 999, the least gain. The highest gain yields the highest background noise and the highest standard deviation of pixel noise.

The data was taken with no light on the CCD. The CCD image was averaged to calculate the mean background level for each gain setting. The standard deviation

of the CCD image was calculated for each gain setting.

Spectral calibration

The spectral channels were calibrated to correspond to pixels. The spectral calibration sources were mercury vapor and argon lamps as well as passband filters.

Recalibration had to be done each time any alignment in the system was changed. The nonlinear part of dispersion was changed if the prism tilt was changed or if the focus was adjusted. The spectra was simply shifted when the pinhole was translated or the fold mirror angle was changed. The data presented here is typical and illustrates the calibration process.

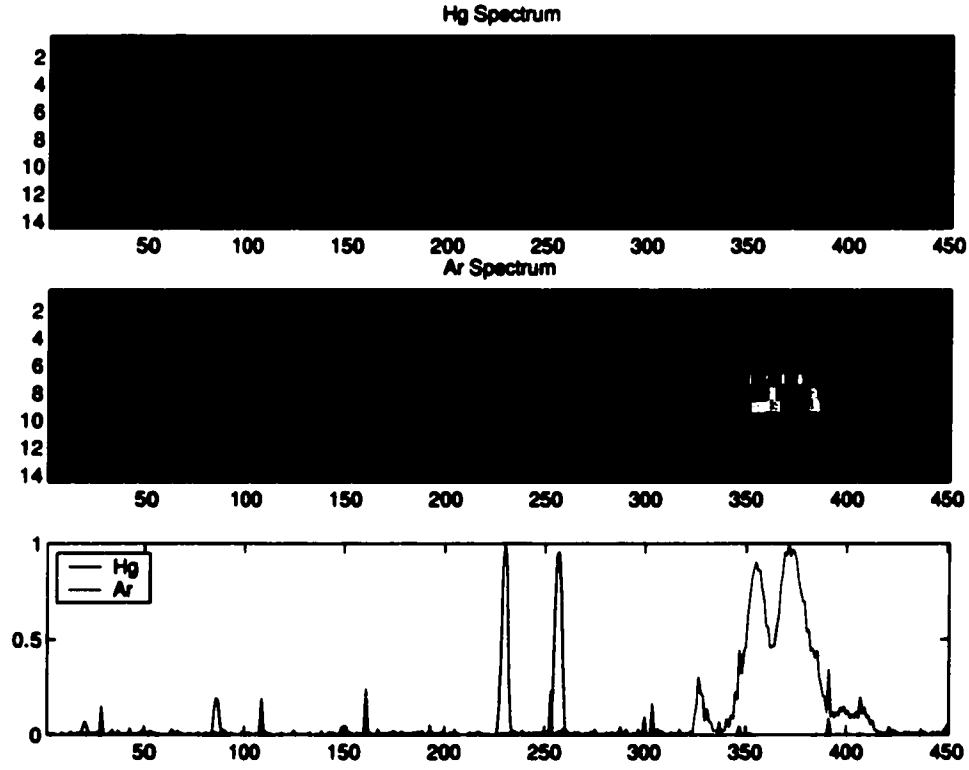


Figure 4.6 Mercury and Argon vapor lamps were used as calibration sources. The two strongest Mercury lines are saturated to bring out two faint lines (top). The Argon spectra is not fully resolved in the second image. The average spectral profiles are plotted (bottom).

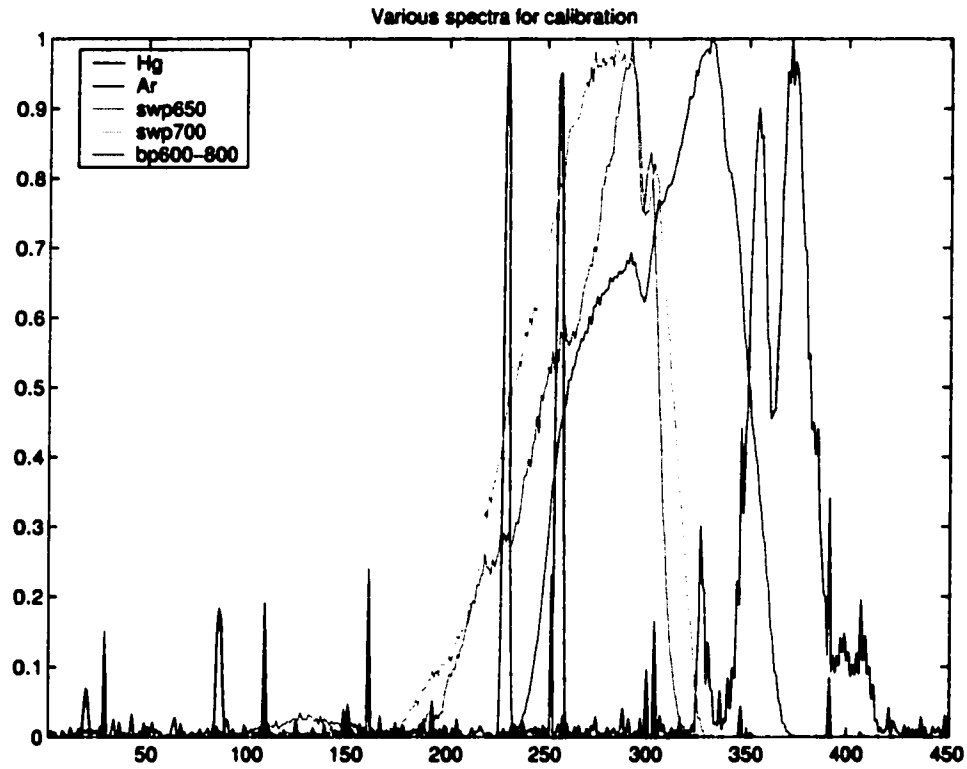


Figure 4.7 In addition to Mercury and Argon vapor lamps, a white light source was used with bandpass filters at 600 nm, 650 nm, 700 nm, and 800 nm.

Each of the spectral sources are imaged onto the CCD. The centroids of the spectral lines are found and the crossover points of the bandpasses determined. This is the raw data of wavelengths corresponding to pixels. The raw data is fit in a least squares sense with a second order polynomial.

Fitting the data has some uncertainty. The Argon lines are not fully resolved (which would require a 1 nm spectral resolution), so the centroids are assigned an assumed wavelength. Levels of saturation can effect the centroids. The polynomial fit was performed for each wavelength extreme, and the best fit chosen based on the polynomical behavior in the blue region.

The spectral resolution on the CCD at 500 nm is 1 nm/pixel and is 2 nm/pixel at 700 nm for the fit shown in Figure 4.8.

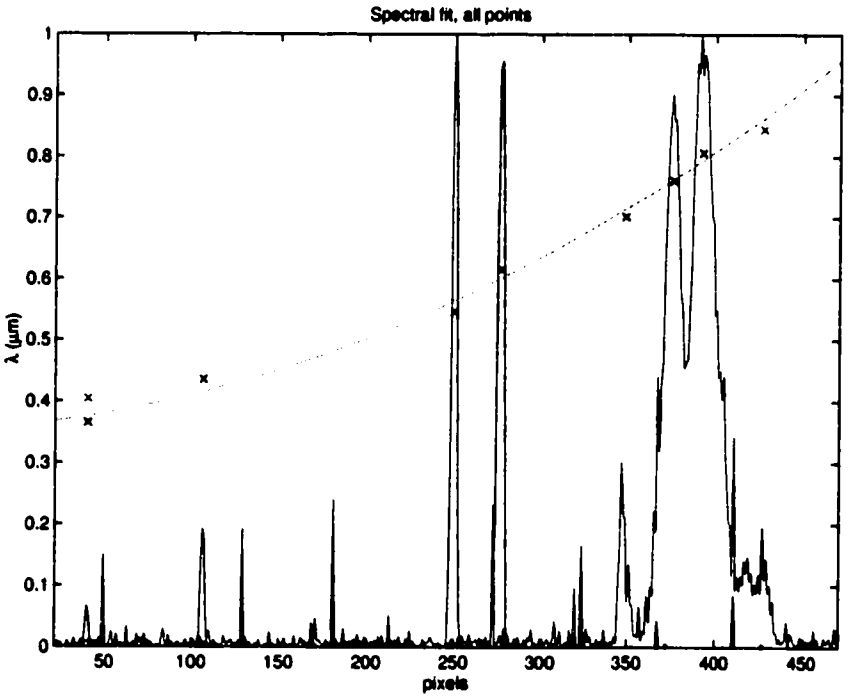


Figure 4.8 The Hg and Ar spectral lines were centroided and assigned wavelengths.The x's mark the centroid locations and possible wavelength assignments. Centroids were fitted to a second order polynomial (the superposed curve).

4.2.2 Phase shifting

The phase of the dispersed fringe was successfully measured in real time. The phase shifting measurement was controlled by setting the initial voltage, the final voltage, and the number of loops through the bins for averaging. There were always 4 steps for the phase shifting algorithm. The typical values for the other parameters were an initial voltage of 0 Volts which is the center of PZT travel, a final voltage of 1.8 V which corresponds to a full wave OPD at 600 nm wavelength, and 10 loops through the PSI process for averaging.

The averaging was performed to reduce pixel noise. The sequence of 4 frames was looped several times. Then the frames from the first step are average, the frames from the second increment are averaged, and so on for the third and fourth increment. Then the averaged frames from each increment are processed by the Carré algorithm.

The phase step is different for each wavelength. The Carre algorithm is rugged against this, but there is some susceptibility to noise as the phase step deviates from 90 degrees. The phases measured in figure 4.9 were fit to a third order polynomial and the residual was used to calculate the standard deviation. This is the rms error over wavelength listed in Table 4.1.

Table 4.1 The same phase was measured four times using different phase step increments. The band pass was 500 - 800 nm. The standard deviation over wavelength is listed for the four step sizes.

increment	at 600 nm	rms error
66 nm	0.11 waves	0.104 ADU
99 nm	0.165 waves	0.041 ADU
149 nm	0.25 waves	0.0356 ADU
182 nm	0.3 waves	0.0372 ADU

Pixel noise also effects the phase shifting reduction. Typical CCD pixel noise,

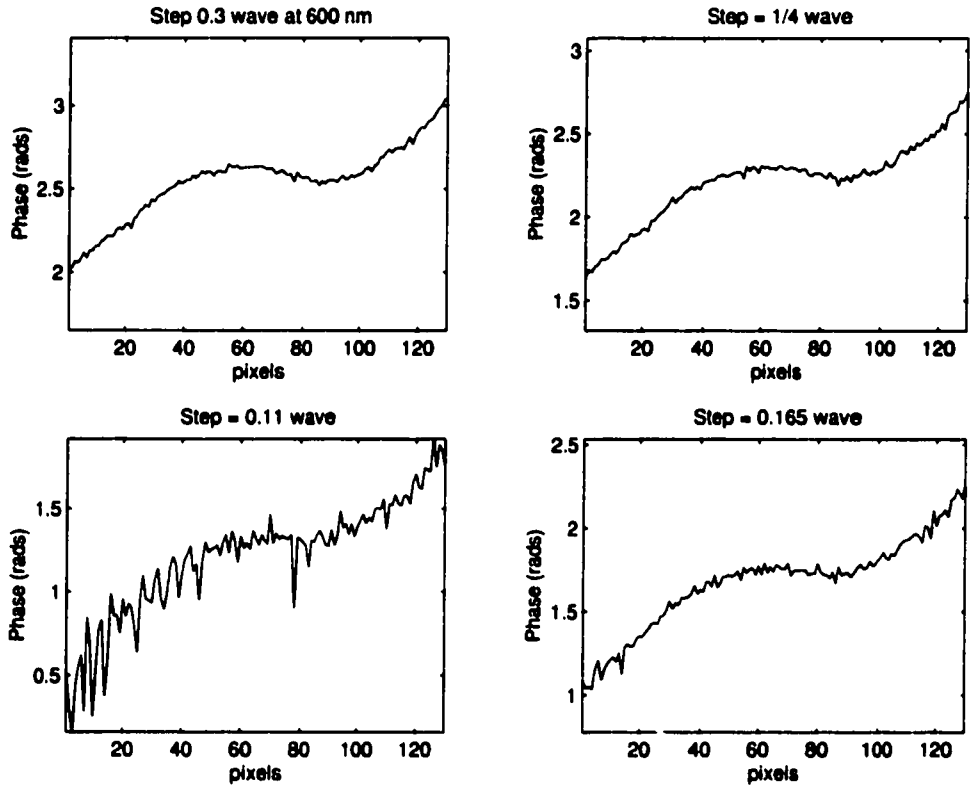


Figure 4.9 The same phase was measured four times using different phase step increments. The Carré algorithm performs with increasing noise as the phase step moves away from a quarter wave.

measured above, was 2 ADU rms. The pixel noise was modeled in section 3.3.2. According to that simulation, a pixel noise of 2 ADU, which is normalized to 0.008, should produce an rms phase error of 0.02 rads. The data of Figure 4.9 shows an rms phase error on the order of 0.02 rads rms.

Once the phase as a function of wavelength, $\varphi(\lambda)$ is measured, the optical path thicknesses can be extracted. Using the phase to find the optical thicknesses is discussed in the next section.

4.2.3 Data Fit for Absolute Phase

The system parameters, the optical thicknesses, are found by multiplying the estimated phase by the system matrix, equation 2.14. This is equivalent to using a model of the system (equation 2.4 to fit several parameters $[s\ t_1\ t_2]$ to the phase data $\varphi(\lambda)$. However, there is a problem with this data fit. The phase, as described in section 3.3.1, is not calculated absolutely, but with it's first datum between 0 and 2π. The true phase is the estimated phase plus some integer multiple of 2π:

$$\varphi_0 = \tilde{\varphi} + m2\pi, \tag{4.4}$$

where φ_0 is the absolute phase, $\tilde{\varphi}$ is the phase estimated from the PSI process, and m is an integer. Now m must be included as a parameter in the fit to the data.

The integer m is found by optimization. The estimated phase $\tilde{\varphi}$, with an added $m2\pi$, is fit to the data for various values of m. The correct value of m is taken to be the one that produces the least χ^2 error of the fit.

The data fit for absolute phase was tested on real data. The data uses only a single plate of BK7, so the system matrix is described by equation 2.13. Four data sets were taken, each with about 7.5 μm more air path than the previous data set. The estimated phases were fit to $[s\ t]$ for various values of m. Figure 4.10 shows the data and the fit error. For various values of m, the fitted values for $[s\ t]$ are

plotted for each of the data sets (lower right). The χ^2 error of the fit as a function of m for each data set is shown in the lower left plot. The m with the lowest χ^2 error was chosen to generated the adjusted, absolute phases according to equation 4.4 (upper left). For 7.5 μm of air path between each data set, the phases at 600 nm wavelength should be separated by 78.5 radians. The separation are close, but not exactly 78.5, and the final data set has a greater separation than the rest. This could be explained by the fact that the final data set was noisier in estimated phase than the others. The noise of all the data sets is seen in the upper right plot which is the difference between the absolute phase φ_0 of the upper left plot and an theoretical phase generated using the values $[s\ t]$ from the fit. The peak to valley error in the absolute phase is about a fifth of a wave.

The estimated values of $[s\ t]$ generated by this fit were disappointing. They are listed in Table 4.2. In the sequence of data, the plate tilt was not changed, so the value for t should not have changed, but the fitted value of t increases by as much as 57 μm between data sets. The value of s incremented between each data set should be about 7.5 μm. The table shows that s increments from 17 to 93 μm between data sets. These values are not reliable for controlling the OPDs towards an achromatic null.

Table 4.2 These are the values of the system parameters $[s\ t]$ given by best fit to the phase data plotted in Figure 4.10. OPD in air of 7.5 μm was added to each successive data set. The OPD in glass was not altered, even though the data fit calculates it.

actuale air path (μm)	0	+7.5	+15	+22.5
estimated s (μm)	-782	-765	-676	583
estimated t (μm)	507	500	445	389
estimated Δs		17	89	93
estimated Δt		7	55	57

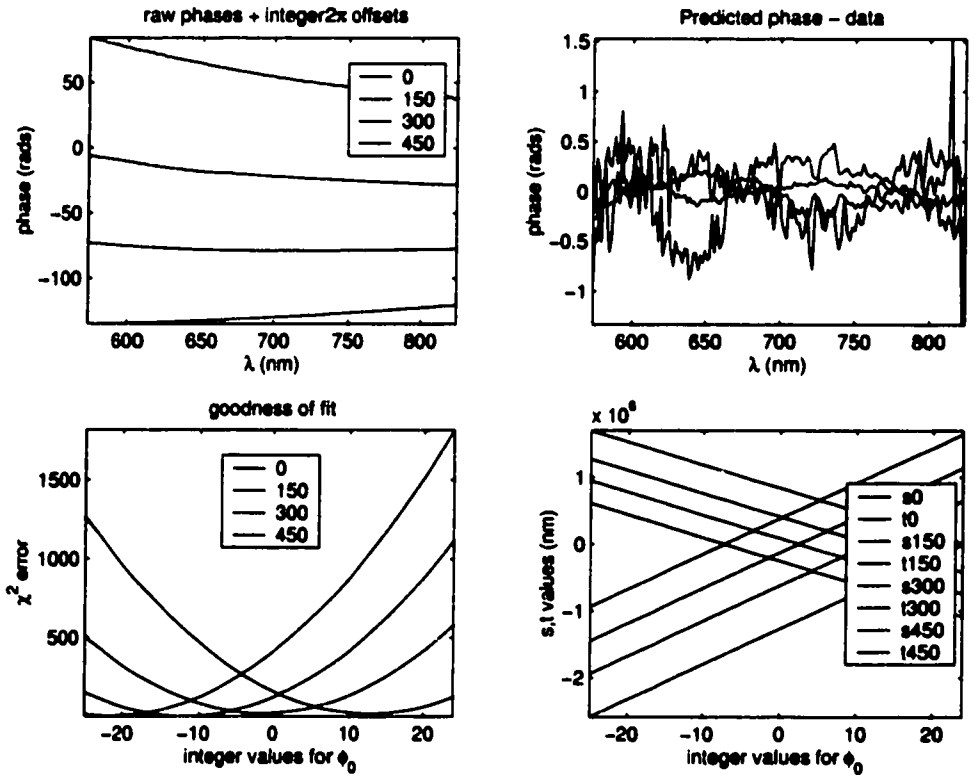


Figure 4.10 PSI phase measurements were taken for 4 data sets, adding about 7.5 μm of air path to each set. The parameters $[s\ t]$ were fit to the phase plus $m2\pi$ for various integer values of m (bottom right). The χ^2 error was calculated for each value of m (bottom left). The m with the least χ^2 error was used to generate the absolute phase (upper left). The upper right plot shows the difference between the absolute phase and an ideal phase using the fitted values for $[s\ t]$.

4.2.4 Changing Optical Thicknesses to Balance Dispersion

The OPD in air, s, and the OPD in glass, t, were varied and then $\varphi(\lambda)$ was measured to verify that the actual behavior matched the theoretical model. Increasing the OPD has the zeroth order effect of changing the absolute phase. The absolute phase can't be directly measured with this system, but can be fitted as discussed in section 4.2.3. I do not do that here, but instead look at the first and second order effects.

The first order effect of changing the OPD is to change the slope of the $\varphi(\lambda)$. The OPD in air was incremented by 320 nm for each data set using the fine precision of the PZT. The relative phase generated by the PSI reduction was converted by hand into an absolute phase, because the change in OPD was known so the appropriate $m2\pi$ could be deduced by examining the data. The data, shown in Figure 4.11 illustrates that the zeroth order effect of changing OPD is to change the phase offset and the first order effect is to change the slope of the phase.

Next, the OPD in glass was varied. The first order effect has the same result as changing the OPD in air, so it was not measured. The second order effect should change the curvature of the $\varphi(\lambda)$. The OPD in glass, t, was incremented by a large amount, and then s was changed so the OPDs balanced each other to minimize the peak to valley extent of the phase. The curvature of the phase should have become greater as more dispersion was added by increasing the OPD in glass, but it did not, as seen by the data plotted in Figure 4.12. The chromaticity of the phase was dominated by a systematic. The systematic turned out to be the dispersion of the dielectric coatings.

4.2.5 Beamsplitter Coating

The data described above showed that the phase was limited in its achromaticity. Large amounts of glass OPD was added, but the effect was barely discernible; some

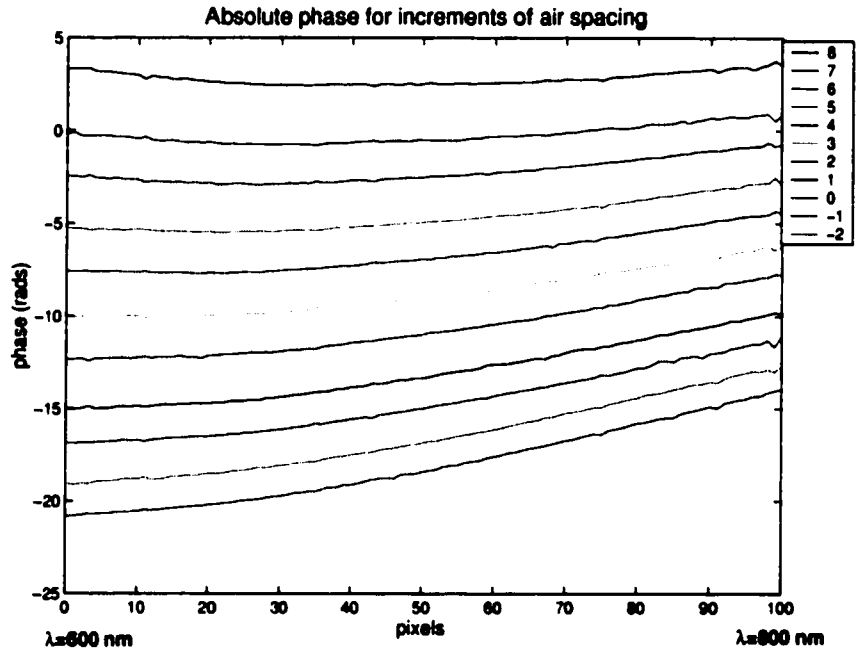


Figure 4.11 The phase was measured for 320 nm increments of OPD in air.

other systematic was dominating with an achromaticity of 1.5 radians peak to valley from 500 to 800 nm. A realignment of the interferometer reduced this systematic dispersion, to such a low amount that a third order form was observed in the residual which is shown in Figure 4.13. This indicated that a second dispersive material was in the optical path: a dielectric coating.

There are two dielectric coatings on the optics: one for the beamsplitter, another for anti-reflection (AR). The beam splitter coating is the standard broad band coating by CVI for the visible spectrum and consists of a stack of dielectric materials. The anti-reflective coating used on the compensator plate and beamsplitter for this experiment was a single layer of MgFl. Both coatings exhibit dispersive behavior that is dependent on angle of incidence. The coating dispersion must be balanced with the glass dispersion and the air path.

The beamsplitter coating should not be the contributor because the beam com-

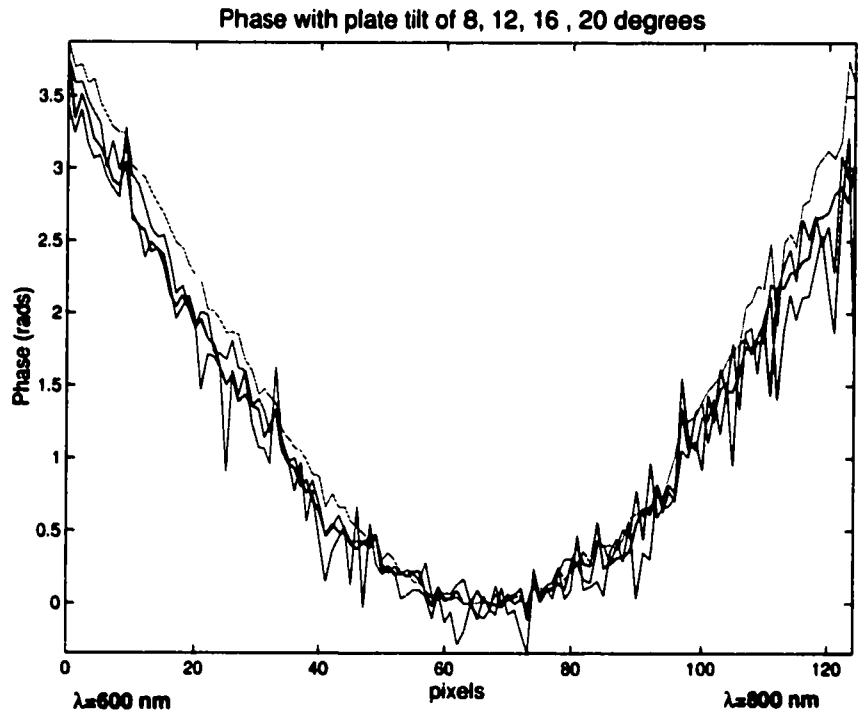


Figure 4.12 The phase was measured for 4 degree increments of the tilt of the glass plate. The curvature of the phase should increase with each tilt angle, but it does not.

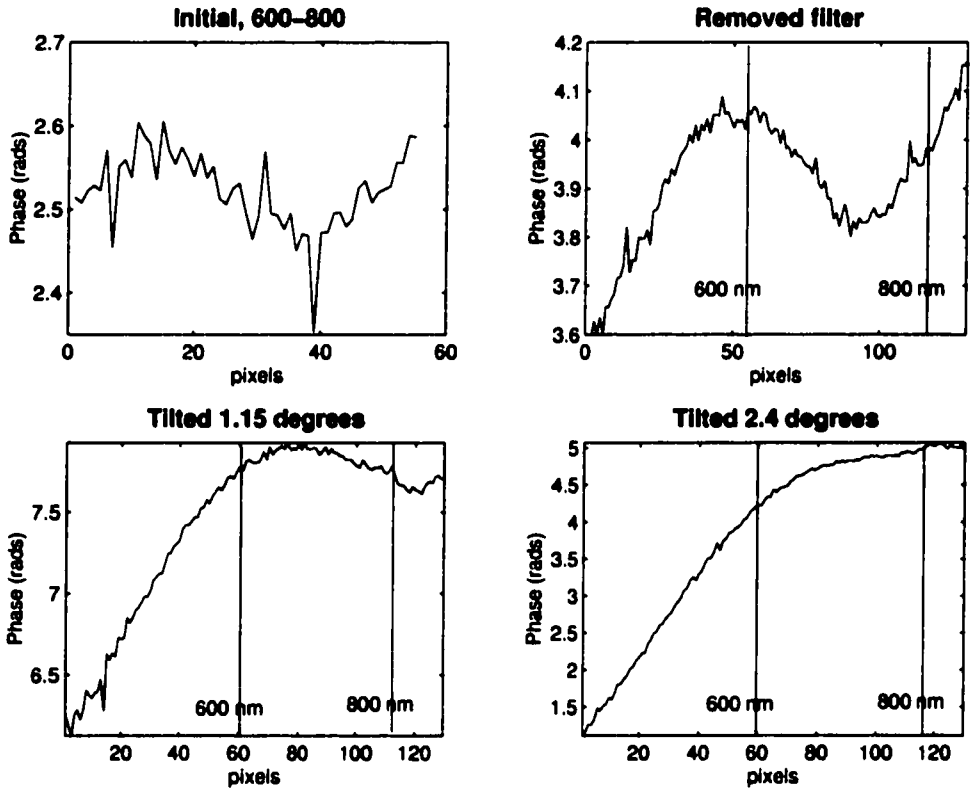


Figure 4.13 With the 600 - 800 nm band pass filter in place, the plate was tilted until the phase achromaticity was minimized (top left). The band pass filter was removed and the plate tilted to reoptimize the achromaticity (top right). Then the plate was tilted 1.15 degrees (bottom left) and 2.4 degrees (bottom right).

biner is double pass; each arm of the interferometer undergoes one reflection and one transmission through the beamsplitter coating. The beamsplitter is AR coated on one side and the compensator plate is AR coated on both sides. One arm of the interferometer transmits through the AR coating twice while the other arm, the one with the compensator plate, transmits four times through the coating. The optical path difference between the two arms is two transmissions through the AR coating.

The left plot of Figure 4.14 shows the theoretical plot for the phase on transmission of the AR coating (data provided by CVI). The beam transmits twice through this coating producing a systematic, dispersive OPD. The dispersion of the coating can be balanced with the dispersion of the compensator glass plate. The right plot shows the theoretical best achromaticity that can be achieved by balancing the dispersion of the AR coating with s and t.

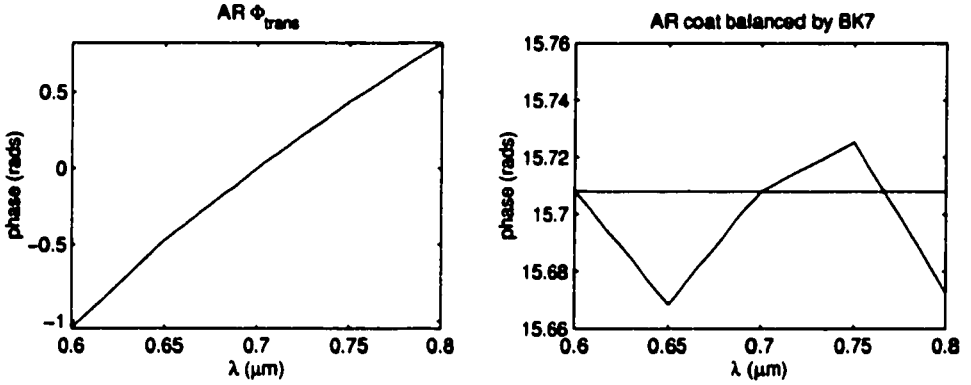


Figure 4.14 The left plot shows the theoretical OPD due to double pass transmission through the AR coating. The AR coating OPD can be balanced by the BK7 OPD and the air OPD to produce the achromatized phase shown in the right plot.

As the glass thickness is increased, the dispersion and thus the chromaticity of the phase also increases. The measured data follows the model for the balancing of dispersion by the AR coating. The best possible null that can be achieved in the presence of the AR coating is $1.1\text{x}10^{-3}$. For a null of this level, the OPD in

glass t must be -74.5 μm and the tolerance Δt is 60 nm. The BK7 plate must have a tilt angle of 11.19 degrees (using the thinned plate) and the tilt tolerance is 6.8×10^{-5} radians, which is 14 arcseconds. The resolution of the rotation stage is 40 arcseconds. Assuming that the OPD error is due to half the resolution, the theoretically best null that could be achieved by balancing the BK7 plate with the AR coating is a null of 0.14. Figure 4.15 shows the optimum null and the null that can be attained with the tilt resolution of the rotation stage.

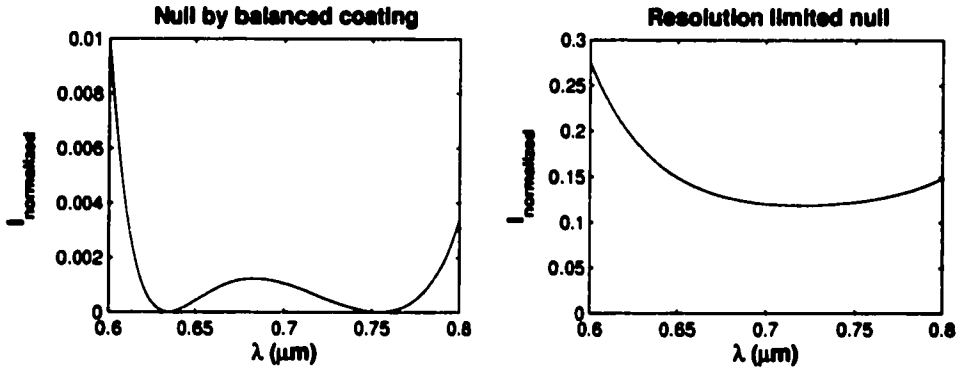


Figure 4.15 The AR coating is a systematic dispersion that can be balanced by BK7. The optimum null (left plot) has a mean intensity of 1.3E-3. The resolution for adjusting BK7 OPD limits the null to 0.14 (right plot).

Measuring the Phase of the Beamsplitter Coating

For a single pass beam combiner architecture, the design of the beamsplitter coating is critical. The coating should be tested to be compared to the theoretical phase. The phase of the coating must be measured. However, most coating facilities only have the capability to measure the intensity of transmission and reflection through a coated optic, not the phase. The phase could be measured by the beam combiner test bed of this experiment with a simple modification. The output of the Michelson interferometer that exits toward the light source has an OPD that is twice the difference between the reflection and transmission through the beamsplitter. A beamsplitter can pick off the output beam and direct it into the set of optics for

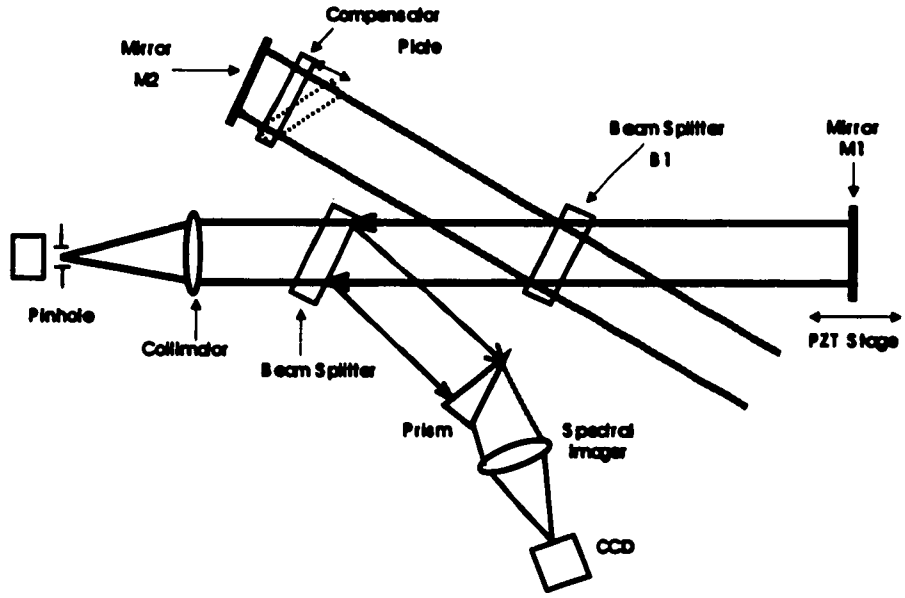


Figure 4.16 The beamsplitter coating can be characterizing by evaluating the output of the interferometer that returns towards the source. This output has the difference between reflection and transmission. The output is sampled by a beamsplitter in common path and then dispersed and phase shifted. The test layout uses the interferometer with compensator control and OPD control for phase shifting.

performing PSI on a dispersed fringe. The concept is illustrated in Figure 4.16. It is hard to isolate the coating to measure it, because the compensator plate is always balancing the dispersion. The measurement is actually a mixture of the coating, the compensator, and the OPD in air.

4.3 Tier 2 - Locating the null

After tier 1 is done, the plate is not changed in tilt anymore. All the variation of phase will be introduced and controlled by moving the mirror on the PZT translation stage. The PZT sweeps through the fringe packet and tracks the null as the OPD

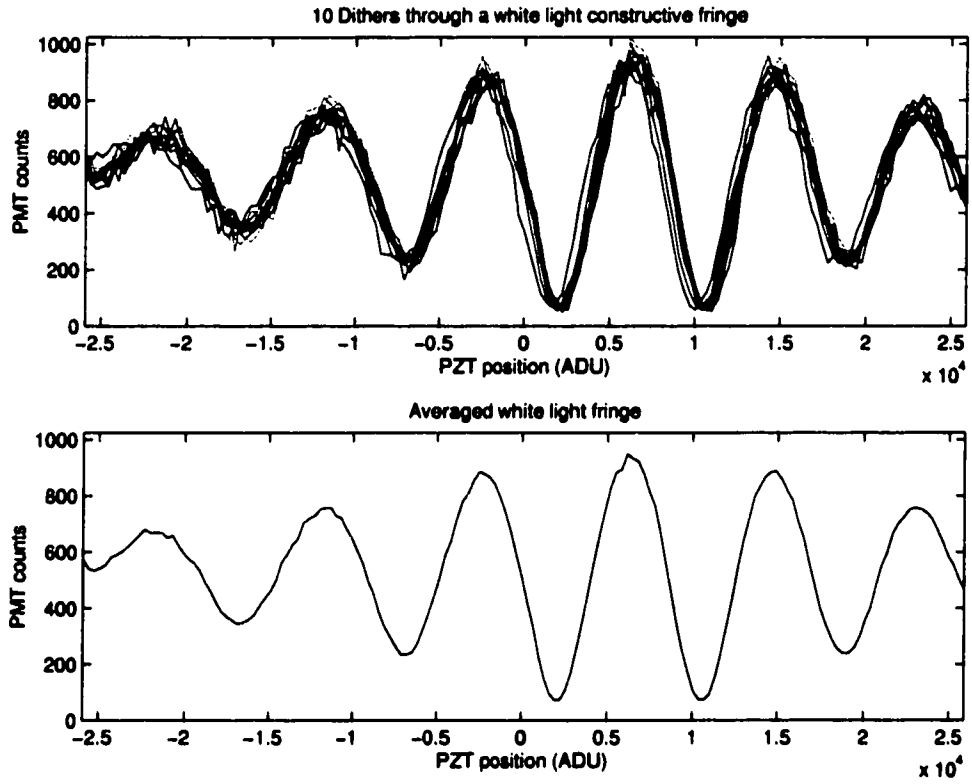


Figure 4.17 A full scan is performed on white light fringe packet. No band pass filters are used, so the fringe coherence envelope is set only by the PMT response and source spectrum. The fringe is symmetric about a constructive peak fringe. The top plot shows ten dithers through the fringe and the bottom plot is their average.

fluctuates on the order of minutes due to air motion. Also, the phase shifting algorithm calculates the null for the phase between the two middle bins, but the position of the mirror after phase shifting is the first bin. So, we have to find the null.

First, the PZT scans through the white light fringe packet. The fringe can be checked for symmetry and the correct order null can be chosen. The wrong order null has worse contrast and will not be as deep. The PZT scans a large range to read the peak fringe; the largest scan available is 20V which gives 6.4 micrometers of OPD or about 10 waves of fringes.

A PZT has hysteresis of typically 10 to 15%. Hysteresis effects a long scan of length 10 μm by half a μm. For a one half wave scan, hysteresis introduces an error of a tenth wave. In dithering over a 60th of a wave, the error is 1 nm. For dithering, a saw tooth wave, not a triangle wave, so that the PZT is scanned always scanned in the same direction. This eliminates hysteresis effects on a scan in the reverse direction. However, the PZT may have some error on returning to the same location at the restart of each scan.

Tier 2 then locates the center of the null. Looking at the broad scan of the white light fringe, the user selects the deepest null and enters its position as the starting origin for the search for the center of the null. The PZT dithers repeatedly on either side of the origin. The dithers are averaged. A parabola is fitted to the average by least squares. The averaged fringe and the parabola are plotted on screen. If the fit is not sufficiently accurate, the user runs the dither-and-fit process again. Usually, the center of the null is found within 4 iterations.

The process for finding the null works for any step size, but the width of the scan must be less than half a wave, otherwise the parabolic fit is inaccurate. The coarsest step size is used to scan half a wave, then the step size is refined and the scan width narrowed. The resolution of the step limits the depth of null that can be measured. A resolution of less than 3 nm is required to guarantee that the full null depth was measured.

As an example, high contrast fringes were achieved using a HeNe laser source. The PMT served as the detector and the counting electronics integrated 30 msec per sample. The null and peak regions of the fringe were sampled every 3 nm for 91 samples as shown in Figure 4.18. This gives adequate sampling of the bottom of the null which has an average intensity of 27 counts. The peak has an intensity of 18,700 counts. The null contrast ratio is 9.1×10^{-4}.

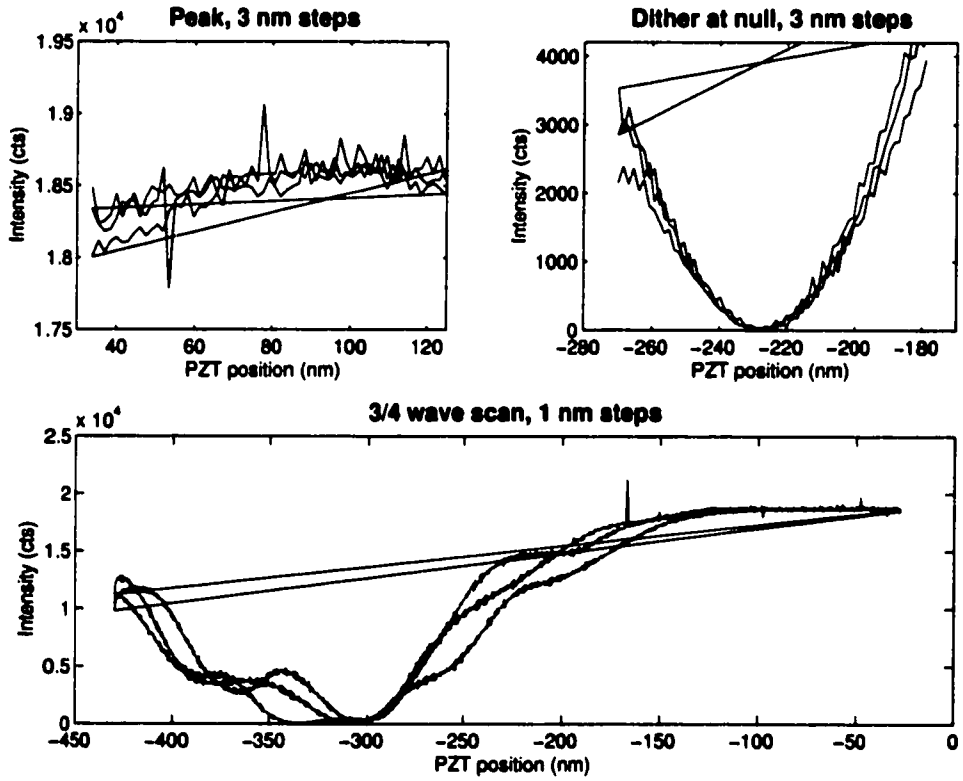


Figure 4.18 A laser was used to produce high contrast fringes. The PZT was scanned over 270 nm of OPD in 3 nm increments to finely sample the peak (top left) and the null (top right). The fringe was scanned with 1 nm step sizes and disturbed by air motion (bottom).

For extremely fine sampling, 3/4 of the fringe was scanned with 1 nm steps. The bottom plot of Figure 4.18 shows that air motion disturbed the three scans. The scans produce a poor and misrepresentative average. The time to perform the three scans was 400 samples * 30 msec/sample * 3 scans = 36 seconds. As discussed in the environmental characterization section, the air motion will vary on the order of seconds. To ensure the average is not limited by air motion, the total scan time should be restricted to less than 1 second.

4.4 Tier 3 - Stabilizing the null

Tier 3 stabilizes the null in the presence of vibrations and air turbulence. Data of the null is accumulated on the science channel, which is a PMT. The sensing channel uses a photodetector with 12 bit analog to digital (A/D) converter to read the irradiance of the grey fringe. A servo control loop locks onto the desired grey fringe light level. The control loop moves the PZT as necessary to maintain the grey fringe light level.

4.4.1 PZT translation stage calibration

The PZT translation stage is a low voltage PZT that is operated with a -10 to 10 Volt, 16 bit digital to analog (D/A) converter. The manufacture's specifications indicate that 1 Volt should correspond to about 165 nm of translation. The correspondence between voltage and translation was examined more carefully.

A HeNe laser source was used to create interference fringes for the PZT calibration measurement. The PZT was dithered through a full wave 10 times at 60 steps per wave. The fringe intensity was measured by the PMT. The intensity was normalized and converted to OPD according to equation 2.1. The measurement and derived OPD as a function of PZT position are shown in Figure 4.19. A best fit line provided the measured correspondence of PZT driving signal to physical translation.

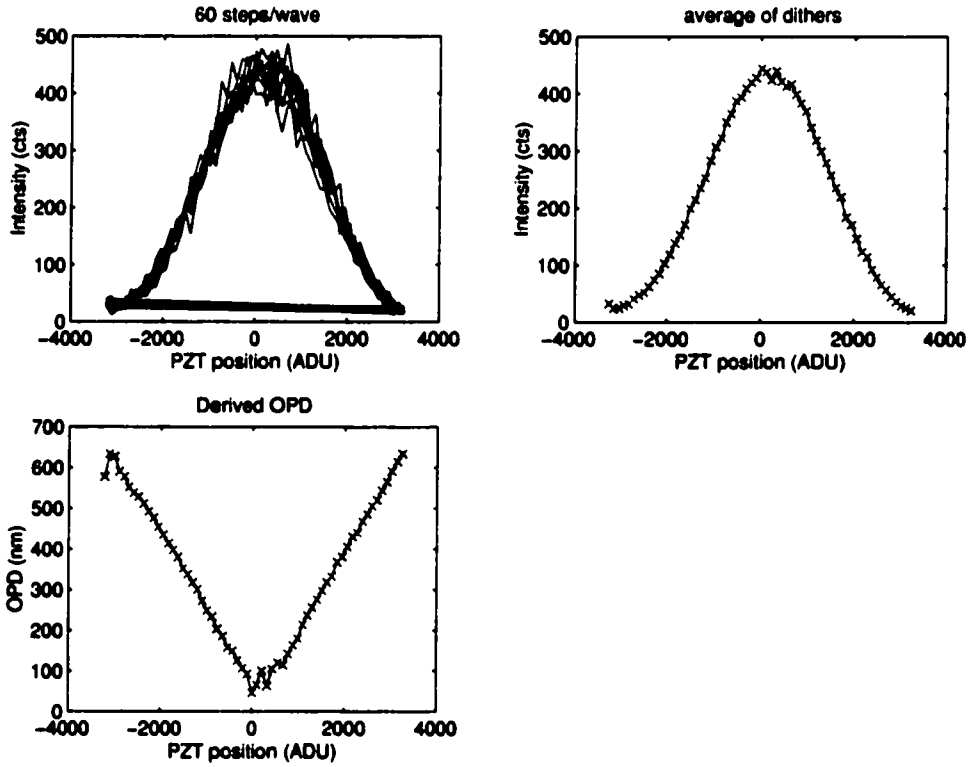


Figure 4.19 The PZT was dithered over a full wave of OPD with a 60 steps per wave (top left). The dithers were averaged (top right). The averaged intensity was normalized and converted into phase that was translated into OPD (bottom).

The actual correspondence is 1 Volt produces a mean of 185 nm of translation, with standard deviation .8 nm. The smallest resolution expected with the 16 bit D/A driver is 0.3 nm, which is below the standard deviation of this calibration. Achieving 1 nm resolution with the PZT stage is the best expected performance.

4.4.2 Gray fringe phase

A dichroic beamsplitter directs light of wavelengths shorter than 450 nm to the photodetector. The interference fringe at this wavelength should be a quarter wave out of phase with the 600-800 nm light in the science channel. To calibrate phase of the blue light, data was collected on both the science and sensing channels while the white light fringe was scanned by applying a sawtooth wave to the PZT from a function generator.

The plots show that the blue light is not a quarter wave out of phase as expected. The phase difference between the sensing and science channels varies with OPD. At the highest contrast OPD, the phase difference is not exactly a quarter wave. This is due to the beamsplitter coating. The theoretical solution shown in Chapter 3 predicted the light at 400 nm to be a quarter wave out of phase, thus providing an ideal grey fringe. That solution failed to account for the beamsplitter coating. The beamsplitter coating adds dispersion, and consequently the 400 nm light is not usable as a grey fringe. More care must be taken with the beamsplitter coating design to yield the correct phase at a shorter wavelength to provide grey fringe control.

4.4.3 Photodetector noise

The photodetector produced a signal from 0 to 7.5 Volts. That signal was digitized by a 12 bit A/D converter with 4 on board gain settings. The noise for each of these signal levels is listed in the table below.

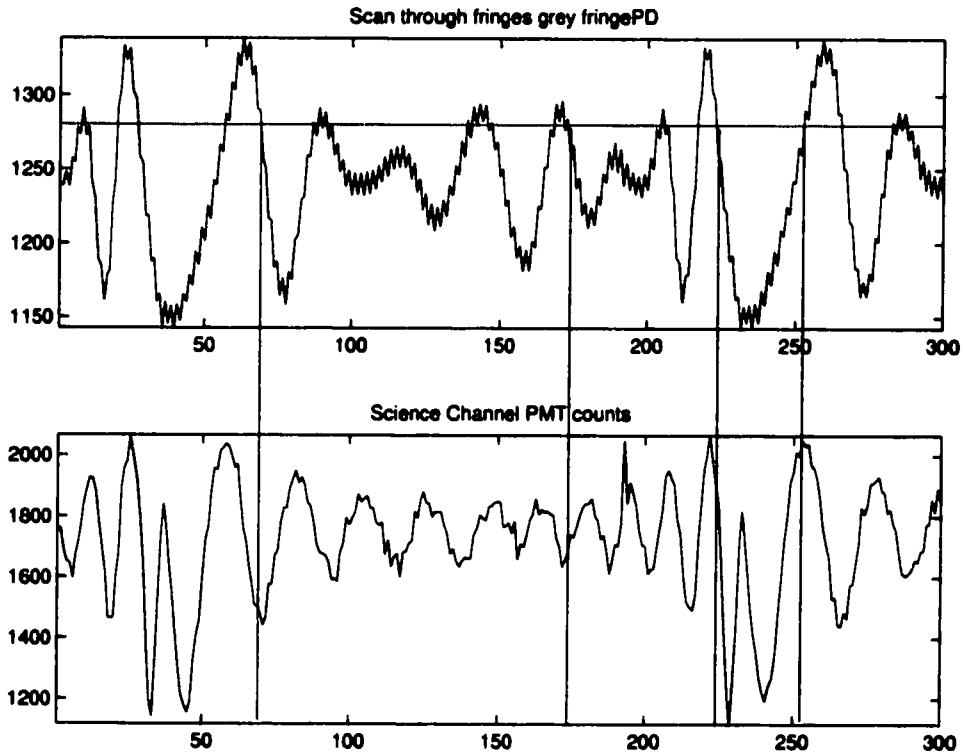


Figure 4.20 The sensing channel detects light short of 450 nm, which is ideally a quarter wave out of phase with the science channel. To find the actual phase difference, the white light fringe packet was scanned while data was taken on the grey fringe, sensing channel (top plot) and the broad band, PMT science channel (bottom plot). The phase difference varies with OPD.

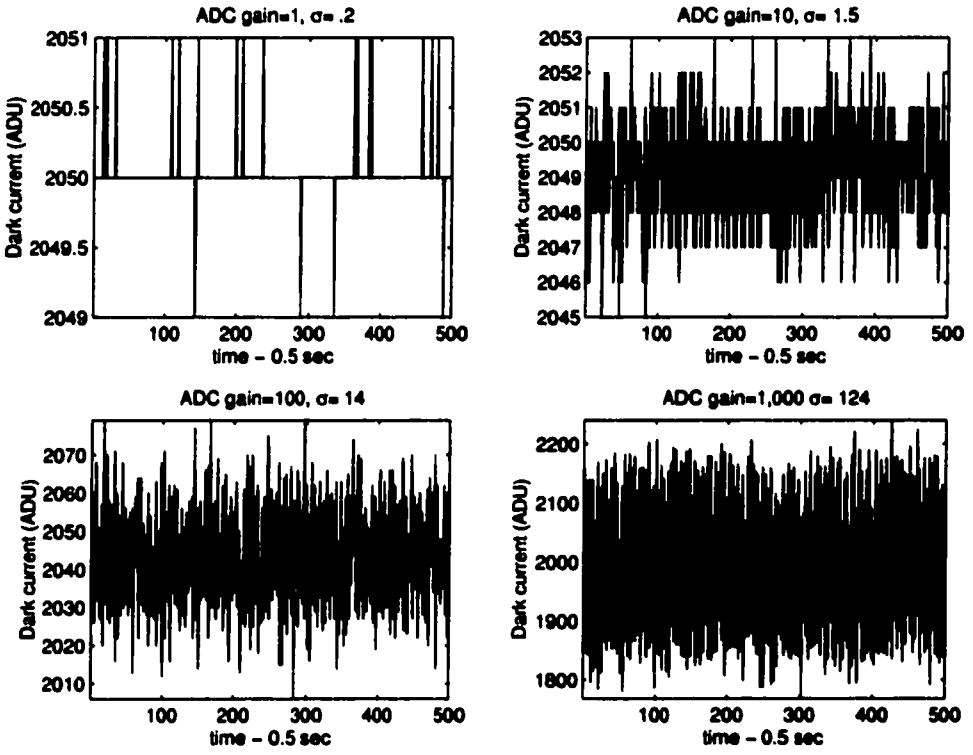


Figure 4.21 The photodiode signal is digitized by the analog to digital converter. The four A/D gain settings allow a small signal to be used, but with increased noise. The plots show the noise at each gain setting for dark current from the detector. A detector signal of 0 Volts corresponds to 2048 ADU.

Table 4.3 The 12-bit analog to digital converter digitizes the signal from the photodiode over 4096 analog-digital units (ADU). Each of the four gain settings has a different dark current noise.

gain	voltage range	standard deviation
1	10 Volts	0.2 ADU
10	1 Volt	1.5 ADU
100	100 mV	14 ADU
1,000	10 mV	124 ADU

The light reflected from the dichroic was very faint, in the 10 mV regime at peak fringe. At half value, the grey level, the signal is 5 mV requiring the gain setting of 1000. The signal was 1280 ADU. The noise for the 1000 gain has a standard deviation of 124. The signal to noise ratio was 10.3.

4.4.4 PMT noise

A PMT was used for the science channel because the light level is expected to be very low at the null and limited by coupling into the single mode fiber. Measuring the null level, which is (hopefully) 10^{-4} of the continuum intensity, requires photon counting. However, most PMTs have a peak spectral response between 400 and 600 nm.

A new hybrid PMT by Perkin Elmer called a Channel Photomultiplier (CPM) featured a peak spectral response between 600 and 700 nm, which is more suited to this experiment. The detector was part number C-962P and had a quantum efficiency of 0.4% at 800 nm. The CPM is composed of a high voltage photocathode (maximum bias 3 kV) followed by a solid state avalanche channel with a 300 V bias. After exiting the avalanche channel, the photoelectrons are incident on the anode. The gain is controlled by varying the high voltage bias. The CPM was a module that included the detector and detection electronics.

The module electronics produce pulses. The pulses are captured by an off-the-shelf counting board in the computer. The PMT integrates for 1 ms. The dark count is 100 counts per second. The counting board can accumulate counts for longer periods. The noise on the counting board is simple digitization noise in the smallest bit resulting in a 2 photon dark current noise peak-to-valley.

4.4.5 Servo Performance

The servo control loop is a simple proportional-derivative control loop. The servo is most sensitive to path changes when the fringes are steep, having high contrast. The damping coefficients depend on the sensitivity, which depends on the contrast. A laser source provided high contrast fringes. The damping coefficients were adjusted until the loop reached stability. The coefficients were 0.1 for both the derivative and the proportional loops.

The servo loop was tested first on an idealized environment with isolation measures. The air was allowed to settle within the enclosed box. The servo loop was actively controlled the PZT for one second. The PZT wandered slowly over a quarter wave OPD to stabilize the power level of the photodiode, as shown in Figure 4.22. The photodiode acted as the sensing and the science channel since the laser source, being monochromatic, would have the same phase on each channel.

Next the servo loop was tested on more challenging air turbulence. While the servo loop was active for one second, the box was open and air was heated by a hand held over one arm of the interferometer. The PZT traveled more than a half wave to stabilize the photodiode signal.

The servo loop was also tested with a white light source. The PZT traveled about 80 nm in a gentle down slope that followed the the photodiode signal, as shown in Figure 4.23. The science channel shows a corresponding rise. The servo loop could have kept the science channel at a more constant level if the damping

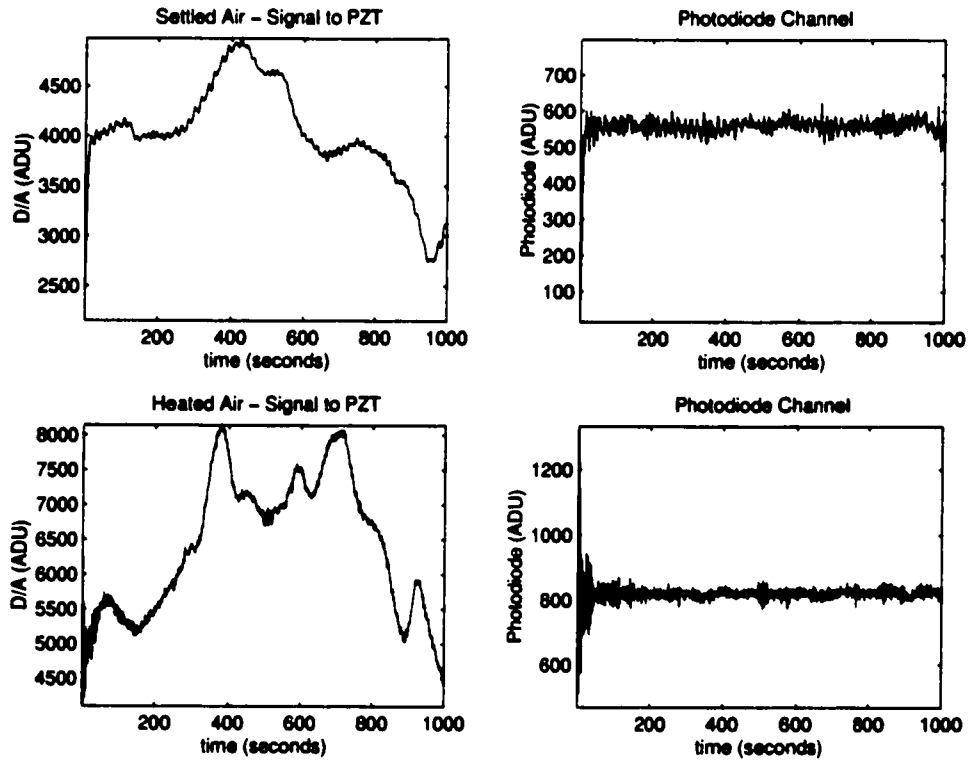


Figure 4.22 A laser provides a strong signal to the sensing channel. The control signal sent to the PZT (top left plot) reveals the environmental noise that was overcome to provide the stable signal on the photodiode (top right plot). The bottom two plots show the control signal and the stabilized signal in the presence of air turbulence.

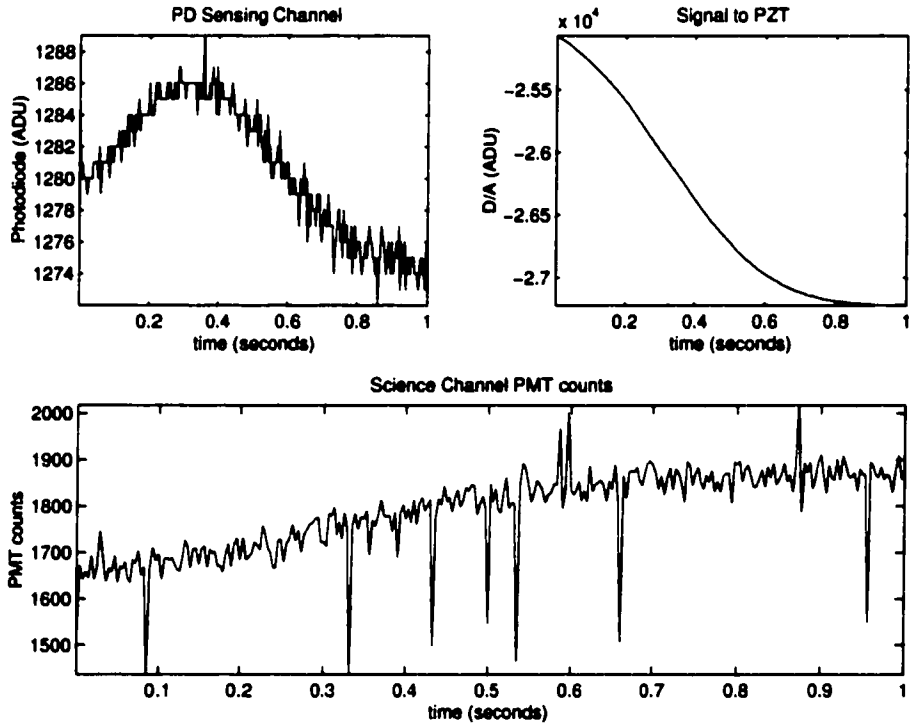


Figure 4.23 The photodiode provides the sensing signal (top left). The servo loop sends a signal to the PZT (top right) to stabilize the sensing signal. The data is accumulated on the science channel (bottom).

coefficients hadn't been as strong.

The speed of the servo limited could be very fast, because the counting electronics are very fast. The A/D converter can operate at 200 kHz, which is the slowest element of the servo. However, the software included a sleep factor that limited the servo speed to 300 Hz.

CHAPTER 5

CONCLUSION

5.1 Measurements of the Null

The driving requirement of the beam combiner was to generate a 10^{-4} null. This level null was achieved initially with a monochromatic source to demonstrate that the required precision of optical alignment could be met. Then a white light source was used to study the achromatization process. Due to hardware limitations, a 10^{-4} null was not achieved broad band. The achromatized null was measured at various bandwidths.

5.1.1 Laser Null

The deepest null achieved with a monochromatic source is $1.6\text{x}10^{-4}$. This null demonstrates that a null of the order magnitude of 10^{-4} is possible with the alignment of the optics. This result was gained early in the experiment as a proof of concept before the servo loop was functional, consequently the null is controlled by hand and detected by a silicon photodiode attached to a calibrated power meter.

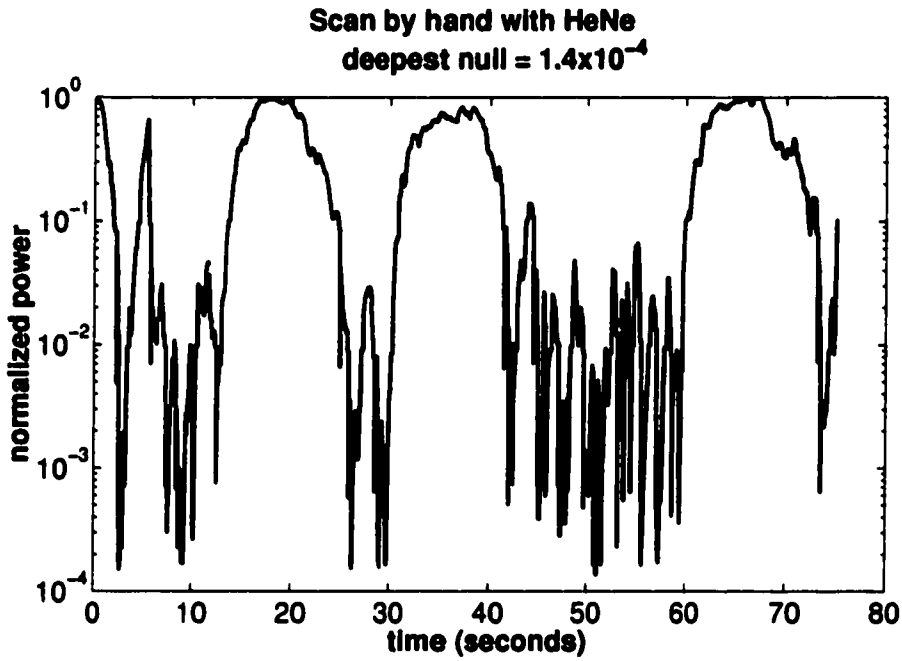


Figure 5.1 A monochromatic source avoids chromaticity problems to verify the optical alignment can generate a 10^{-4} level null. The interferometer output is coupled through a single mode fiber to a photodiode on a calibrated power meter. Sampling is 30 Hz. The fringe is swept by hand to yield a $1.6\text{x}10^{-4}$ null.

5.1.2 White light fringe sweeps

Scans through nulls for various bandwidths were measured. The null was found by manually adjusting the plate tilt until the fringe was symmetric about the null. The fringe was then visually examined to verify that the central null was black and that the fringe colors were symmetric about the null.

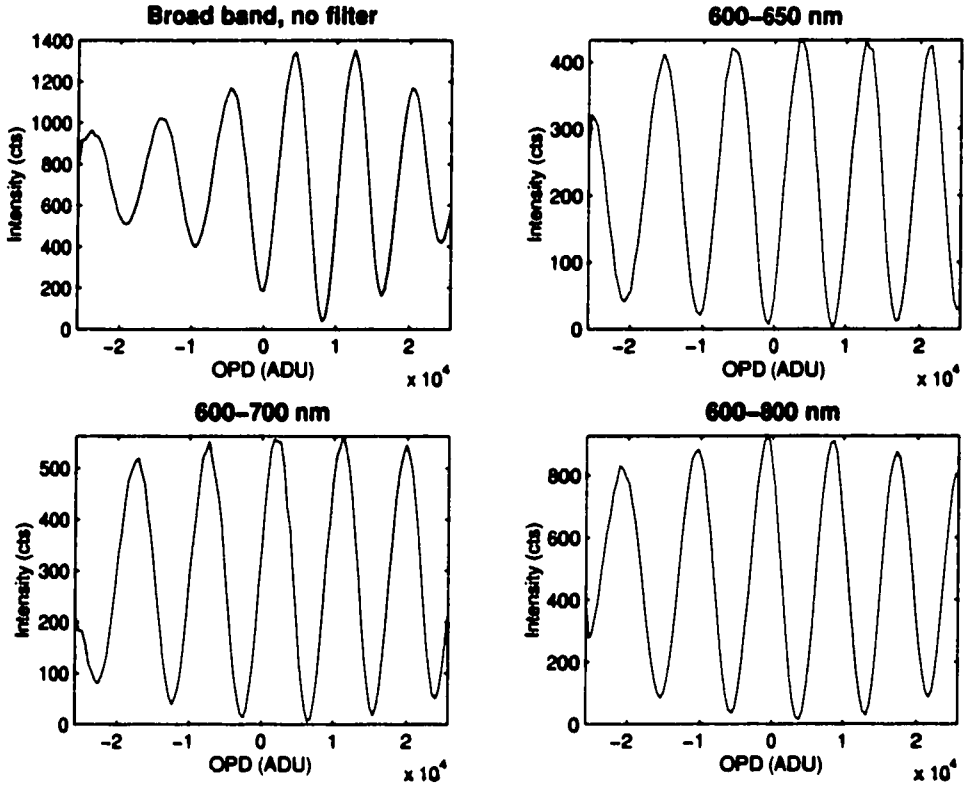


Figure 5.2 The plate was adjusted to produce a fringe that was symmetric about the null. The nulled fringe was scanned over 8 waves of OPD for various band passes: 600-650 nm (upper right), 600-700 nm (lower left), 600-800 nm (lower right). In the upper left plot, no filter was used. The coherence envelope was set by the source spectrum and the spectral response of the CPM.

The fringe intensity was acquired on the CPM while the PZT scanned the OPD repeatedly through 4.8 μm. The repeated scans were averaged to produce the nulled, broad band fringes shown in Figure 5.2. The OPD step resolution was 40 nm which could, in a worst sampling case, limit a monochromatic null to 0.033. The integration

time for each sample was 30 msec for 600-800 and 600-700 nm, 20 msec for 600-650 nm, and 10 msec for the broadest band pass. The null contrast ratio for each band pass is given in Table 5.1.2.

Table 5.1 The nulled, broad band fringes were measured for various band widths. The null contrast ratios were calculated from the data presented in Figure 5.2.

band pass	null ratio
no filter	0.028
600 - 800 nm	0.0069
600 - 700 nm	0.009
600 - 650 nm	0.015

5.2 Conclusions

Optical thickness of dielectric materials can be balanced to produce achromatic phase shifts. A half wave phase shift in the visible using BK7 glass and fused silica could produce of null of 10^{-4} in the visible spectral region of 600 nm to 800 nm. In the infrared region from 7 to 20 μm, Zinc Selenide and Zinc Sulfide could produce a null of 10^{-6}. The infrared null is at a level necessary for direct detection and spectroscopy of earthlike planets around other stars. This nulling technique may be a critical component of a future mission to search for small extrasolar planets and perform high resolution, high dynamic range astronomy.

These precision nulling solutions require optical thicknesses and tolerances that are prohibitive to manufacture. The optical thickness is introduced and actively controlled by using a differential thickness. One plate of a particular material is placed in each arm of the beam combiner; the differential thickness the optical path difference (OPD) through that material. The tolerances on the OPDs are achieved by tilting the dielectric plates. Tilting the plates provides active control of the phase

and the phase chromaticity.

A method for active control of the optical path thicknesses was developed. The phase as a function of wavelength is measured. The phase is fit to a model to determine the current optical thicknesses. The eigenfunctions of the model are nearly colinear and consequently the dynamic range required of the phase data is very high. For a single plate of BK7, the required dynamic range is $2x10^3$. For a dual plate solution using BK7 and fused silica, the required dynamic range is 10^4. The dynamic range required for a dual plate solution in the infrared using ZnSe and ZnS is the same. The best matched pairs of dielectric materials are well matched because they have very similar first and second order dispersion curves. Similar dispersion curves produces system eigenfunctions that are more colinear and makes the dynamic range more demanding.

The phase is measured using phase shift interferometry (PSI) techniques. The fringe is spectrally dispersed and imaged on a CCD. The OPD in air is incremented in 4 equal steps while a sequence of frames are captured. The sequence of 4 frames are reduced by the Carré algorithm. The Carré algorithm allows for OPD increments that are non-90 degrees and not known. The physical OPD step translates into a different increment of phase for each wavelength. The PSI process was modeled to determine the effect on the rms phase error by pixel noise and air movement. To meet the dynamic range of $2x10^3$ for a single plate solution, the rms phase error must be less than 0.013 radians. For an rms phase error of 0.01 radians, the pixel noise must be less than 0.005 of the normalized intensity (1.3 counts for an 8-bit CCD) and the air turbulence must be less than 0.1 radians, which is about a sixtieth of a wave and 10 nm at a wavelength of 600 nm.

The phase measured by the PSI process is known relatively, not absolutely. The absolute phase is an integer multiple of 2π added on to the relative phase. The

relative phase cannot be used in the fit to solve for the optical thicknesses; only an absolute phase will give the true optical thicknesses. The absolute phase must be included as one of the parameters of the fit.

A beam combiner was implemented in hardware as a laboratory testbed. The beam combiner testbed performed several unique feats. The first of which was phase shift on a dispersed fringe and reduce the data with the Carré algorithm. The sensitivity of this algorithm was explored with respect to environmental factors and hardware noise of the system. The phase shifting operation was successfully implemented in the hardware in real time to measure the phase as a function of wavelength. The air motion noise was about a hundredth wave rms over a second. This should provide low enough noise in the measured phase to successfully solve for the optical thicknesses. The pixel noise was typically 2 counts rms which is 0.008 of the normalized, maximum signal level. This level of pixel noise translates into an rms phase error of 0.04 radians; the error is too large to solve for the optical thicknesses.

The behavior of the chromaticity of the phase was studied by varying the OPD in air and the OPD in glass. A systematic dispersion was observed. The dispersion was attributed to dielectric coating. The dispersion of the coating, if known, can be included in the system model and can be balanced by the other dielectric materials in the system. The dynamic range required of the phase data then becomes on the order of 10^4, the same as for a dual plate solution. The coating dispersion, if balanced by BK7 dispersion, would limit the null to 1.1×10^{-3}. To reach this null, the BK7 OPD must be 56.7 μm with a tolerance of 60 nm. The OPD tolerance translates into a tolerance of plate tilt angle of 14 arcseconds. The rotation stage only had a resolution of 40 arcseconds. The 40 arcsecond tilt resolution could limit the null to 0.14.

The white light fringe was detected by a PMT while the PZT scanned through the fringes. The precise center of the null was located by dithering the OPD around the bottom of the null, averaging the dithers, and fitting the average to a parabola. The center of the null was found in less than four iterations as long as the dither was less than half a wave. A monochromatic fringe was measured this way to have a null contrast ratio of 9.1×10^{-4}. The time to perform all dithers of a measurement should be limited to less than a second to avoid air motion.

The OPD was stabilized to 6 nm peak-to-valley using a 300 Hz servo loop implemented in software. Grey fringe control was demonstrated whereby the servo locks onto the signal at a short wavelength of light while a second detector acquires the fringe intensity of the main pass band (600 - 800nm). The fringe at the shorter wavelength of light is supposed to be a quarter wave out of phase with the main pass band so that it is a grey fringe when the main pass band is a null. Controlling on the grey portion of the fringe gives greater sensitivity to OPD change than controlling at the bottom of the cosine fringe. Due to the dielectric coating, the shorter wavelength fringe did not have the ideal quarter wave phase shift. The servo was still able to stabilize on the intensity signal of the shorter wavelength fringe. Due to the detector and amplifier, the signal to noise ratio at the shorter wavelength was very low and consequently controlling at the shorter wavelength did not provide a better sensitivity to changes in OPD.

5.3 Recommendations for Future Implementations

The work on this test bed revealed the main challenges of this technique and developed viable solutions. This implementation was limited by hardware. I am confident that another engineering iteration using better hardware could reach a 10^{-4} null over a broad band. The hardware upgrades necessary to reach the goal of this experiment

are available at the current state of the art.

Another possible future implementation is a nulling beam combiner in the infrared to demonstrate a 10^{-6} null. Such an infrared testbed could be directly built from the lessons of this visible testbed without finalizing the 10^{-4} visible null. However, the infrared beam combiner for such a deep null will require active amplitude control, which should perhaps first be developed in its own testbed.

The next implementation of a nulling beam combiner with phase plates is a programmatic issue. An infrared testbed will become more relevant as the Terrestrial Plant Finder program advances to the technology development stage. An infrared nulling beam combiner will likely be one aspect of its technology development program. Is there further cause for a visible combiner? Currently, plans for a nulling demonstration on SIM have been dropped. A visible beam combiner could be built as an observing instrument, but atmospheric seeing limits the null so severely that interesting astronomical science could not be performed. A nulling beam combiner will likely be part of a possible precursor mission to TPF; the beam combiner requirements will be dictated by the goals of the mission and could be designed using the lessons gained from this work.

I recommend several pragmatic modifications and precautions for future implementations.

Of primary importance is the dispersion contributed by any dielectric coatings. If possible, I recommend avoiding coatings all together. If non-common path coatings are necessary, they should be empirically measured. The empirical coating phase should be verified with the theoretical phase of the coating. The empirical coating phase should be added to the system model. The dynamic range required of the phase data can be determined from the SVD of the system model matrix. A new solution of plate thicknesses and tolerances must be generated to balance the coating.

Examine the behavior of the shifted phase at the short wavelength to ensure it is usable for grey fringe control. If the beam combiner is implemented as a single pass architecture, extreme care must be put into the design of the coating to minimize its dispersion and control its behavior at the grey fringe sensing wavelength.

Another very important consideration is the detector used in measuring the phase of the dispersed fringe. The detector must have enough dynamic range to allow for solving for the optical thicknesses. For a dual plate solution, the required dynamic range of 10^4 could be met with a 16-bit CCD with 6.5 counts of pixel noise. If the dynamic range is not met, the optical thicknesses cannot be fit to the phase data, and the phase cannot be achromatized in an automated fashion. Balancing three nearly colinear OPDs is intractable manually.

Determining the absolute phase is a difficult part of the data fit. A scheme could be implemented using the PSI phase measurement and scans through the broad white light fringe to verify the absolute phase.

The detector used for grey fringe control should be tailored to the short, grey fringe wavelength. This test bed used a photodiode detector that has a peak spectral response near 800 nm. A PMT, which typically has peak spectral response between 400 and 500 nm, would have been more suited and given a higher signal to noise ratio.

The control of the beam combiner was hierarchical; each tier completed its work before control switched to the next tier. However, a nested control structure may be more helpful. The servo, as the inner most control loop, can run continuously to stabilize the phase as the phase shifting data frames are acquired or as the bottom of the null is dithered. It may permit longer averaging. And it may be necessary for a beam combiner constructed for an environment with more noise, such as an observing environment.

A practical consideration is alignment. To achieve the final broad band null, the interferometer must be aligned with enough precision to achieve a monochromatic null of the same depth. This testbed did not employ a convenient nor repeatable method for switching between the alignment source and the broad band source. Consequently, the optical alignment of the system was limited. I recommend using a fiber for source injection. A fiber y-junction can deliver light from the alignment laser and the broad band source to a location that does not have to be altered to switch between the sources. I also recommend using as an alignment source an amplitude stabilized laser. I used a HeNe laser that fluctuated 10% in amplitude even after an hour and a half warm up period.

REFERENCES

[1] S. West and et al., "Towards first light for the 6.5 m mmt telescope," in *Optical telescopes of Today and Tomorrow*, A. Ardeberg, ed., *Proc. SPIE* **2871**, 1996.

[2] J. Hill and P. Salinari, "The large binocular telescope project," in *Telescope Structures, enclosures, controls, assembly, integration, validation, and commissioning*, T. Sebring and T. Anderson, eds., *Proc. SPIE* **4004**, 2000.

[3] M. Mayor and D. Queloz, "A jupiter-mass companion to a solar-type star," *Nature* **378**, p. 355, 1995.

[4] G. Marcy and R. Butler, "A planetary companion to 70 virginis," *Ap. J. Lett.* **464**, p. L147, 1996.

[5] G. Marcy, "A transiting 51 peg-like planet," *Ap. J. Ltrs.* **529**, p. L41, 2000.

[6] D. Fischer, G. Marcy, R. Butler, S. Vogt, S. Frink, and K. Apps, "Planetary companions to hd 12661, hd 92788, hd 38529 and variations in keplerian residuals of extrasolar planets," *Ap. J.* **551**, pp. 1107–1118, 2001.

[7] D. Fischer, G. Marcy, R. Butler, G. Laughlin, and S. Vogt, "A second planet orbiting 47-uma," *Ap. J.* **10 Jan**, p. in press, 2002.

[8] T. Nakajima and et al., "Gliese229b," *Nature* **378**, p. 463, 1995.

[9] C. Beichman, N. Woolf, and C. Lindensmith, *The Terrestrial Planet Finder (TPF): A NASA Origins Program to Search for Habitable Planets*, NASA JPL, Pasadena, 1999.

[10] N. Woolf and J. Angel, "Astronomical searches for earth-like planets and signs of life," *Annu. Rev. Astron. and Astrophys.* **36**, pp. 507–37, 1998.

[11] J. Angel, A. Cheng, and N. Woolf, "A space telescope for infrared spectroscopy of earthlike planets," *Nature* **322**, p. 341, 1986.

[12] R. Hanel, B. Jennings, and R. Samuelson, *Exploration of the Solar System by Infrared Remote Sensing*, Cambridge University Press, Cambridge, 1975.

[13] J. Angel and N. Woolf, "An imaging nulling interferometer to study extrasolar planets," *Ap. J.* **475**, pp. 373–377, 1997.

[14] R. Bracewell, "Detecting nonsolar planets by spinning infrared interferometer," *Nature* **274**, p. 780, 1978.

[15] M. Ollivier and et al., "Nulling interferometry for the darwin mission: Laboratory demonstration experiment," in *Optical and IR Interferometry from Ground and Space*, S. Unwin and R. Stachnick, eds., *ASP Conf. Ser.* **194**, p. 443, 1999.

[16] M. Shao and M. Colavita, "Long baseline optical and infrared stellar interferometry," *Annu. Rev. Astron. and Astrophys.* **30**, p. 457, 1992.

[17] E. Serabyn, "Nanometer-level path length control scheme for nulling interferometry," *Appl. Opt.* **38**, p. 4213, 1999.

[18] E. Serabyn, "Deep nulling of visible laser light," *Appl. Opt.* **38**, p. 7128, 1999.

[19] J.R.P.Angel, J. Burge, and N. Woolf, "Detection and spectroscopy of exoplanets like earth," in *Optical Telescopes of Today and Tomorrow*, A. L. Ardebarg, ed., *Proc. SPIE* **2871**, pp. 516–519, 1996.

[20] B. Menesson and J.-M. Mariotti, "Array configurations for a space infrared nulling interferometer dedicated to the search for earthlike extrasolar planets," *Icarus* **128**, pp. 202–212, 1997.

[21] N. Woolf and J.R.P.Angel, "Planet finder options 1: New linear nulling array configurations," in *Beyond the Solar System and the Next Generation of Space Missions*, D. Soderblom, ed., *ASP Conf. Ser.* **119**, pp. 285–293, 1997.

[22] N. Woolf, J.R.P.Angel, and J. Burge, "Planet finder options 3: Focal plane instrumentation," in *Infrared Space Iterferometry: Astrophysics & the Study of Earth-like Planets*, C. Eiroa and et al., eds., pp. 295–307, Kluwer Academic Publishers, 1997.

[23] P. Hinz, *Comparison of amplitude control methods for Nulling Interferometry Applications.* PhD thesis, University of Arizona, Tucson, 2001.

[24] E. Sabatke, "Nulling interferometry for studying other planetary systems: Techniques and observations," tech. rep., Center for Astronomical Adaptive Optics, University of Arizona, Tucson, 1999.

[25] D. Peterson and M. Shao, *Space Interferometry Mission (SIM): Taking Measure of the Universe*, NASA JPL, Pasadena, 1999.

[26] W. Smith, *Modern Optical Engineering*, McGraw Hill, Boston, 1990.

[27] W. Wolfe and G. Zissis, *The Infrared Handbook*, Environmental Research Institute of Michigan, Michigan, 1993.

[28] H. Barrett and K. Myers, *Fundamentals of Image Science*, Wiley and Sons, in press, 2002.

[29] A. A. Michelson, "Measurement by light-waves," *Am. J. of Science* **39**, 1890.

[30] E. Serabyn, "Nulling interferometry: symmetry requirements and experimental results," in *Interferometry in Optical Astronomy*, P. Lena and A. Quirrenbach, eds., *Proc. SPIE* **4006**, 2000.

[31] M. Mansuripur, "Reciprocity in classical linear optics," *Opt. and Phot. News* **10Jul**, p. 53, 1998.

[32] M. Colavita and et al., "Aseps-0 testbed interferometer," in *Amplitude and Intensity spatial interferometry 2*, J. Breckenridge, ed., *Proc. SPIE* **2200**, p. 89, 1997.

[33] J. Grievenkamp and J. Bruning, "Phase shifting interferometry," in *Optical Shop Testing*, D. Malacara, ed., pp. 501–598, John Wiley & Sons, Inc, 1992.

[34] K. Creath, "Phase shifting speckle interferometry," *Appl. Opt.* **24**, p. 3053, 1985.

[35] M. Born and E. Wolfe, *Fundamentals of Optics*, Whiley and Sons, New Jersey, 1989.