OPTICAL DESIGN FOR EXTREMELY LARGE TELESCOPE

ADAPTIVE OPTICS SYSTEMS

by

Brian Jeffrey Bauman

A Dissertation Submitted to the Faculty of the

COMMITTEE ON OPTICAL SCIENCES (GRADUATE)

In Partial Fulfillment of the Requirements
For the Degree of
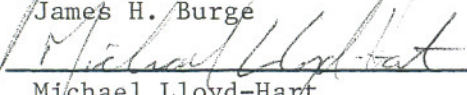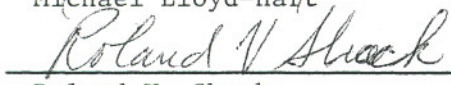
DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2 0 0 3
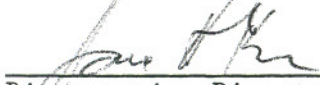
THE UNIVERSITY OF ARIZONA ®
GRADUATE COLLEGE

As members of the Final Examination Committee, we certify that we have

read the dissertation prepared by Brian Jeffrey Bauman

entitled Optical Design for Extremely Large Telescope Adaptive

Optics Systems

and recommend that it be accepted as fulfilling the dissertation

requirement for the Degree of Doctor of Philosophy

James H. Burge	Dec. 3, 2003 Date
Michael Lloyd-Hart	12/3/03 Date
Roland V. Shack	12/03/03 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   12/12/2003
                                                Date

STATEMENT BY AUTHOR

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

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgement 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 copyright holder.

SIGNED: _____

ACKNOWLEDGEMENTS

DEDICATION

One of the joys of an extended graduate student career is that when you finish, there are many people to thank in the dedication.

First, I give thanks to my advisor Jim Burge, who has been a constant source of encouragement and who would not let me fail. I thank my committee members, Roland Shack and Michael Lloyd-Hart for their dedication. I'm pleased that Professor Shack has received much well deserved recognition recently. When I speak to other OptSci graduates, they agree solemnly when I say, "The longer that I'm in this field, the more I appreciate what Roland Shack has done."

Many other OptSci members deserve special note: Kathy Creath, who I'm proud to call a friend and who provided valuable feedback at critical times; Jack Gaskill, who displayed great kindness when my mom became sick; and Didi Lawson, who took care of all of the administrative details during this rather long *in absentia* period.

I am honored to learn from terrific co-workers at Lawrence Livermore National Laboratory and in the adaptive optics community; the list is too long to be given here. But in particular, Scot Olivier and Don Gavel gave me a chance in the field, patiently answered questions, and didn't laugh when I told them my ideas. Elinor Gates, who worked with me countless nights at Lick Observatory, has been a valued friend.

My long journey has been enriched by the many friends that have walked along with me (even from afar), of which this is a woefully incomplete list: Elaine Barnett, Dawn Garcia-Nuñez, Nancy Narbut, Jonathan Howell, the Graf and Loeb families, and classmates Scott McNown, Fred Froehlich, Eric Mentzell, Phyllis Ryder, Andrew Lowman, Hope Queener, and Gene Campbell. Lastly, Rich Dekany was there at the beginning of grad school, provided the connection into adaptive optics 6 years ago to re-start my graduate career, and fittingly, was there at the end.

I dedicate this work to my family: to my brother-in-law Mike, my very favorite niece Julianne, and my very favorite nephew Maxwell; to all my cousins; to my in-laws—אבא ואמה ומוני ויניב ואודי ואלנית ומשפחה נגר: תודה! to all my aunties and uncles, who, quite simply, have always been there and who have stepped in where my parents once stood; to my stepfather, Dick Neill, who I think of like a father, because he treats me like a son; to my grandmother, Gertrude Bauman, and to the memories of my grandparents Barney Bauman, Max and Leah Fischel—I know that they are proud.

To my sister Vicki and my brother Daniel, who know just how hard it is;

To my wife Einat and my daughter Rebecca, who have been patient and have given the time, support, and encouragement that it took to finally finish;

And finally, I dedicate this to the memories of my father Harold and my mother June, who, by their presence, and who, by their absence, have shaped my life. Their efforts and their sacrifices enabled me to reach this day. Isaac Newton said, "If I have seen further, it is because I have stood on the shoulders of giants." I agree—my parents are those giants. I miss them more than I can express.

TABLE OF CONTENTS

TABLE OF CONTENTS - *Continued*

LIST OF FIGURES

LIST OF FIGURES - *Continued*

LIST OF FIGURES - *Continued*

LIST OF TABLES

ABSTRACT

Designing an adaptive optics (AO) system for extremely large telescopes (ELT's) will present new optical engineering challenges. Several of these challenges are addressed in this work, including first-order design of multi-conjugate adaptive optics (MCAO) systems, pyramid wavefront sensors (PWFS's), and laser guide star (LGS) spot elongation.

MCAO systems need to be designed in consideration of various constraints, including deformable mirror size and correction height. The $y,\overline{y}$ method of first-order optical design is a graphical technique that uses a plot with marginal and chief ray heights as coordinates; the optical system is represented as a segmented line. This method is shown to be a powerful tool in designing MCAO systems. From these analyses, important conclusions about configurations are derived.

PWFS's, which offer an alternative to Shack-Hartmann (SH) wavefront sensors (WFS's), are envisioned as the workhorse of layer-oriented adaptive optics. Current approaches use a 4-faceted glass pyramid to create a WFS analogous to a quad-cell SH WFS. PWFS's and SH WFS's are compared and some newly-considered similarities and PWFS advantages are presented. Techniques to extend PWFS's are offered: First, PWFS's can be extended to more pixels in the image by tiling pyramids contiguously. Second, pyramids, which are difficult to manufacture, can be replaced by less expensive lenslet

arrays. An approach is outlined to convert existing SH WFS's to PWFS's for easy evaluation of PWFS's. Also, a demonstration of PWFS's in sensing varying amounts of an aberration is presented.

For ELT's, the finite altitude and finite thickness of LGS's means that the LGS will appear elongated from the viewpoint of subapertures not directly under the telescope. Two techniques for dealing with LGS spot elongation in SH WFS's are presented. One method assumes that the laser will be pulsed and uses a segmented micro-electromechanical system (MEMS) to track the LGS light subaperture by subaperture as the light is returned from the upward-propagating laser pulse. A second method can be used if the laser is not pulsed. A lenslet array is described which creates "pixels" which are aligned with the axes of the elongated spot of each subaperture, without requiring special charge-coupled devices (CCD's).

1. INTRODUCTION

In man's insatiable quest to see ever-dimmer, ever more-distant, and ever-older objects in the universe, he has built ever-bigger telescopes at an astonishing clip—increasing the telescope diameter by about an order of magnitude every 100 years! This is 2 orders of magnitude per century in terms of collecting area!
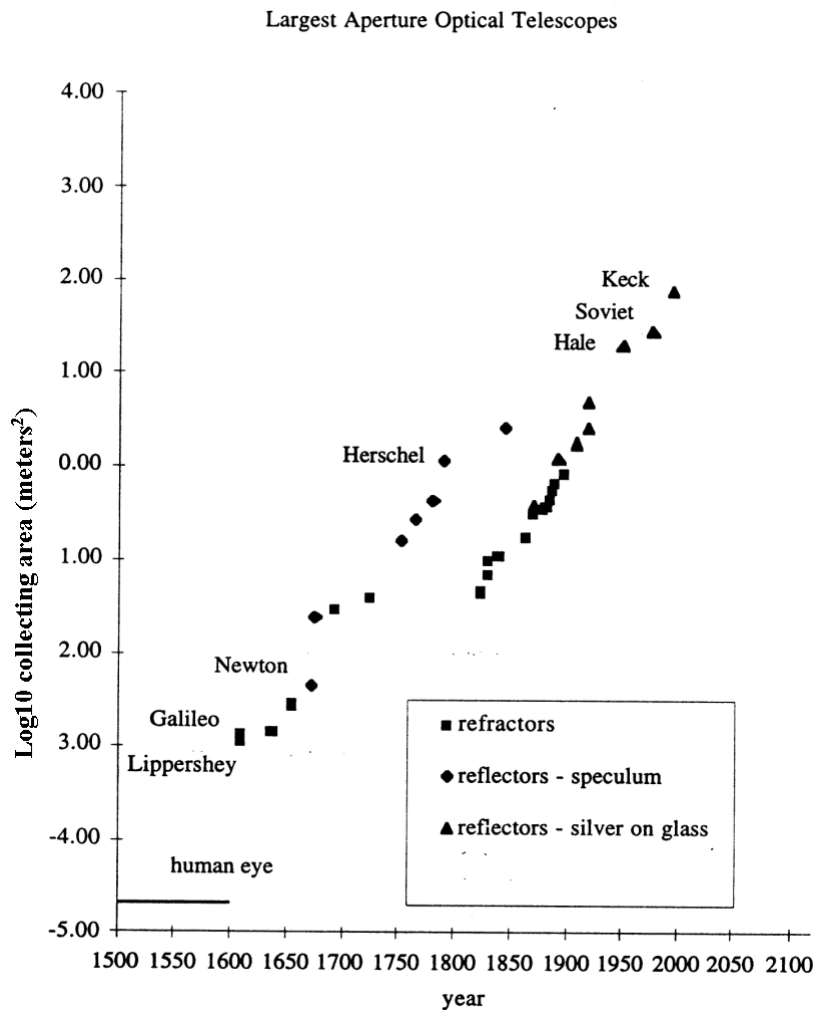
Figure 1- 1: Collecting area of several telescopes versus commission date. Telescope collecting areas have increased approximately 2 orders of magnitude each century. (Nelson)

As new, larger telescopes have been built, older telescopes have been threatened with obsolescence. However, the advent of adaptive optics (AO) has breathed new life into the smaller telescopes. AO systems correct for the deleterious effects of atmospheric turbulence that would otherwise degrade images (Babcock); see figure 1-2. With AO, telescopes achieve diffraction-limited images (full-width, half maximum (FWHM) of $\sim\lambda/D$), rather than seeing-limited ($\sim$0.5-1.0 arcsec) images. For a telescope such as the Lick 3 meter, the improvement from $\sim$1 arcsec FWHM to $\sim$0.14 arcsec at an observing wavelength of 2.2μ represents a dramatic increase in scientific throughput: approximately a factor of 50! At this point, current AO systems for medium and large telescopes are fairly mature—for example, the AO system on the Lick Observatory 3 meter telescope is on the telescope 30% of the time (Gavel), even with other potent instruments available to the observer.
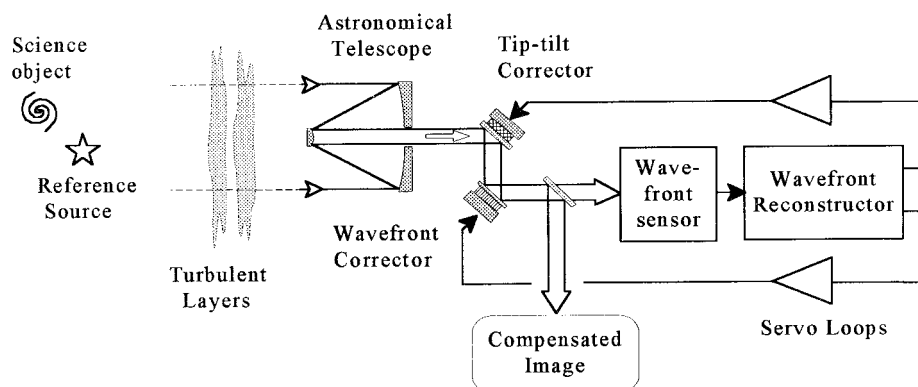
Figure 1- 2: A typical astronomical adaptive optics system (from Hardy).

The success of AO on smaller, existing telescopes has prompted planners of the next generation of telescopes, dubbed "Extremely Large Telescopes" (ELT's) to plan for a more integral implementation of AO, rather than as an "add-on". Indeed, the promise of "winning" at a D^4 rate (two powers of D for the collecting area, two powers of D for the improved resolution) is impossible to ignore. Thus, it is commonly accepted that the next generation of extremely large telescopes (ELT's) will use AO.

In the past and present, AO systems have been single-conjugate (SCAO) as shown in figure 2, i.e., they have one deformable mirror (DM) correcting at one height in the atmosphere according to wavefront sensor measurements taken from one reference source or "guide star" (GS). The guide star may be natural (NGS) or created by a laser (LGS). Typically, the deformable mirror (DM) is placed conjugate to the telescope primary (often called, "the ground layer"). This is done because most of the turbulence is located near the ground. The WFS detects the wavefront from the GS. This measured wavefront error is the integrated wavefront error over the entire height of the atmosphere.

If all of the turbulence were located at the ground layer and corrected with a DM conjugate to the ground layer, then we could have an arbitrarily large corrected field. But this is not what happens in reality—there are turbulence/refractive index changes at altitudes above the ground and this means that the required correction is different at different field angles; see figure 1-3. The angle over which the required correction is

relatively good (i.e., root-mean-square (rms) wavefront error < 1 rad^2) is termed the isoplanatic angle (Fried 1982). In other words, beyond the isoplanatic angle, the correction will be wrong because the light from that field angle will go through atmosphere other than that which was measured. Not surprisingly, then, the isoplanatic angle is fairly small (~10-20 arcseconds in diameter centered about the guide star, for an observing wavelength of 2.2μ). (California Association for Research in Astronomy (CARA))
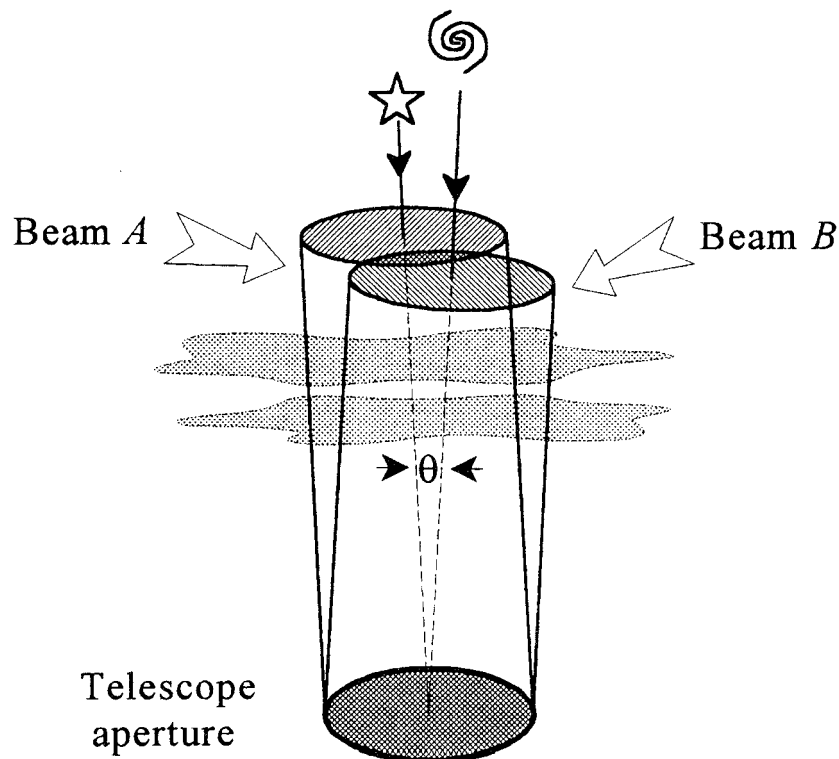
Figure 1- 3: Illustration of angular anisoplanatism. The wavefront measured within Beam A (the guide star) is not valid for Beam B (the science object) since the two volumes of atmosphere are different. Thus, a correction which is perfect for Beam A will not be perfect for Beam B. (from Hardy)

The isoplanatic angle can be calculated given statistical knowledge of the vertical distribution of the refractive index variations (Fried 1982).

$$\theta_0 = \left[2.914 k^2 (\sec \zeta)^{8/3} \int_h dh \, C_N^2(h) \, h^{5/3} \right]^{-3/5}$$

$$\text{where } k = \frac{2\pi}{\lambda}$$

$$\zeta = \text{zenith angle} \tag{1}$$

$$h = \text{altitude}$$

$$C_N^2(h) = \text{refractive index structure parameter}$$

Given the isoplanatic angle, the wavefront variance (in radians[2]) due to anisoplanatism can be calculated by equation 2 (Fried, Hardy):

$$\left\langle \sigma_\theta^2 \right\rangle = \left(\frac{\theta}{\theta_0} \right)^{5/3} \tag{2}$$

The limited isoplanatic angle has an adverse effect on system performance in two ways: first, the field that can be observed at any one moment with significant correction is limited by this isoplanatic angle; this limits the scientific "throughput" (measured, perhaps, in data points per second) of the AO system. The second effect is that the isoplanatic angle limits the fraction of the sky for which natural guide star (NGS) AO can be used. This effect comes about because we need a relatively bright guide star

(typically, brighter than ~14th magnitude), and stars this bright are, on average, much further apart than the isoplanatic angle. In fact, NGS SCAO can be used over only about 1% of the sky (CARA). These figures are for a science wavelength of 1.25μ and a target Strehl ratio of 3% under conditions with 0.5 arcsec seeing, θ_0=4 arcsec, and Greenwood frequency of 50Hz; the atmospheric parameters are measured at λ=0.55μ. Changing the science wavelength to 2.2μ improves the sky coverage by about a factor of 10-20. The complete set of assumptions is given in CARA.

This limitation is somewhat overcome by the use of laser guide stars (LGS's) (Foy, Max), which allow placement of a GS in an arbitrary field position on the sky, nominally above the atmosphere. In addition, a natural tip/tilt star of 17th-19th magnitude is required to stabilize the scene, but these are 1-2 orders of magnitude more common than appropriately-bright NGS's for high-order correction (CARA).   Furthermore, the isokinetic angle (i.e., that angle over which the tip/tilt component of the atmosphere is relatively constant) is considerably larger than the isoplanatic angle (~10 arcsec vs. ~60 arcsec). The end result is that the LGS AO system can be used over ~50% of the sky.

In order to expand the corrected field, Beckers proposed a multiconjugate adaptive optics (MCAO) technique; see figure 1-4 for two variations on MCAO to be discussed below. MCAO systems seek to correct over a larger field of view by correcting at multiple layers in the atmosphere with multiple DM's using wavefront measurements from multiple

guide stars. The appeal of MCAO is apparent: the scientific throughput of an MCAO system is approximately $\left(\theta_{0,MCAO}/\theta_{0,SCAO}\right)^2$. There is another potential improvement: that the PSF in an MCAO system may be more constant across the field of view than in an SCAO system. This makes deconvolving a scene much easier. It is also provides more opportunity to capture a well-corrected "PSF star" (i.e., a known "point source" whose image is used to deconvolve the image from the science target) within the same frame as the science target. Doing so eliminates the time-consuming and scientifically-ambiguous process of imaging the PSF star, then the science field, and then back to the PSF star. The strategy here is that if the PSF star's image stays constant, then the observer is more confident that the same PSF applies during the science image. But what happens if the PSF's image isn't constant? It is then difficult to know how to deconvolve the image. A larger field with a more consistent correction avoids this issue as well as improves the observing efficiency.
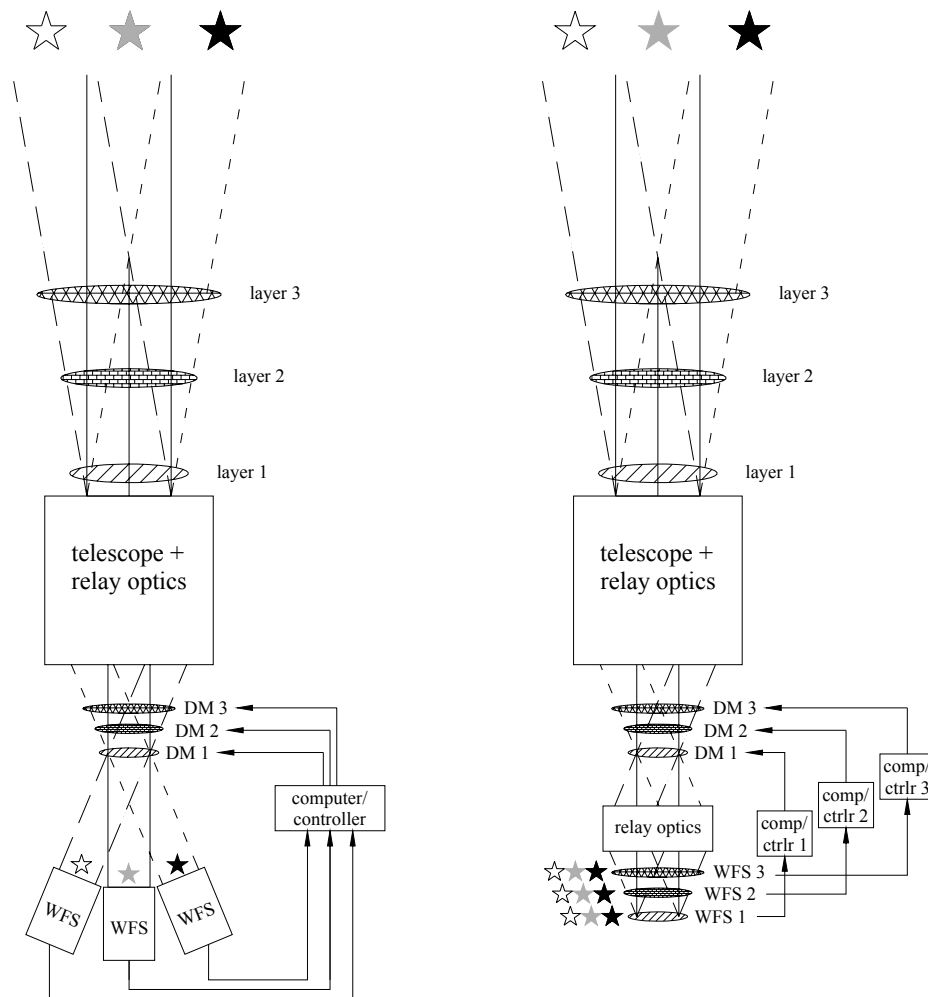
Figure 1- 4: Two methods of MCAO—star-oriented or tomographic MCAO (left) and layer-oriented MCAO (right). Both methods correct at multiple heights in the atmosphere using multiple DM's and multiple WFS's. Tomographic MCAO uses one WFS per guide star and brings all the information together in a computer, which issues commands to the collection of DM's. Layer-oriented AO combines the light from all guide stars, with one WFS per DM. In LOAO's purest form, each WFS controls its corresponding DM. Versions of LOAO have been suggested where the data from all three WFS's are brought together in one computer.

With this promise, an MCAO system will be commissioned for the first time (on the Gemini 8m telescope) shortly (Ellerbroek).   But the question of how to implement MCAO on ELT's is one with many open issues. We know that an MCAO system will require multiple DM's but it is not clear what the best way is to lay out such a system. We know that we do not currently have in hand DM's with sufficient numbers of actuators and sufficient size to implement MCAO. Where should we put our limited development money? The answer can be guided by an evaluation of the first-order optics of such a system.

Once we have DM's, what approach should we take to sensing wavefronts and controlling the DM's? Two approaches have been suggested: tomographic AO (a.k.a. "star-oriented AO") and layer-oriented AO (LOAO) (Ragazzoni 2000b). Conventional "tomographic" approaches to MCAO involve measuring wavefronts from a variety of angles (basically, one WFS per guide star[*]), reconstructing the three-dimensional aberration distribution function via tomography and assigning the resulting distribution function to the various DM's needed for MCAO. In contrast, a LOAO scheme would use a "confocal" technique to measure the wavefront aberration contribution from a designated atmospheric layer (integrating the light from the several guide stars together,

---

[*] although it may be possible to "slice-and-dice" multiple guide star wavefront measurements onto one physical camera; similar approaches can apply to LOAO as well.

so that there is one WFS per layer); the measured wavefront would then be used to control a DM conjugate to that designated layer. In this way, the control of an LOAO system would be via a few independent control loops. Since there are likely to be many fewer guide stars than layers, a LOAO system would yield fewer WFS's than a corresponding tomographic MCAO system. This is an important point – the smaller number of WFS's will result in important cost, complexity, and possibly performance improvements. These performance improvements from fewer WFS's derive from the concomitant reduction in pixels and hence reduction in total read noise.

Now, in an LOAO scheme, there is a question of how should the light from several GS's be sensed and combined. Ragazzoni (1996) has proposed using a pyramid WFS (figure 1-5). The PWFS, to be discussed in Chapter 3, is similar to the SH WFS, but reverses the order of aperture division and field division, and so produces 4 pupils on the CCD, each one corresponding to a field quadrant.

The PWFS is convenient for combining the light from several guide stars, as called for in LOAO. Such a beam-combining scheme is depicted in figure 1-6 (Ragazzoni 2002b).
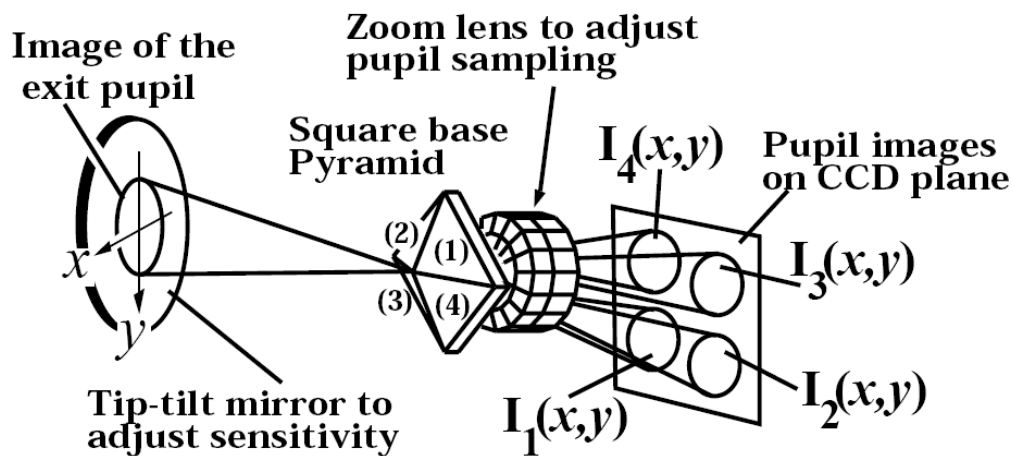
Figure 1- 5: Schematic of a pyramid wavefront sensor (from Esposito). A converging beam focuses on the apex of a 4-facet pyramid. The pyramid divides the light into 4 beams, one for each quadrant of the field, similar to a quad-cell. A subsequent field lens re-images each beam's pupil (I_1, I_2, I_3, I_4) onto a CCD. A tip-tilt mirror in the pupil plane enables a circular scan of the image around the apex of the pyramid; this enables adjustment of the sensitivity of the wavefront slope measurement.
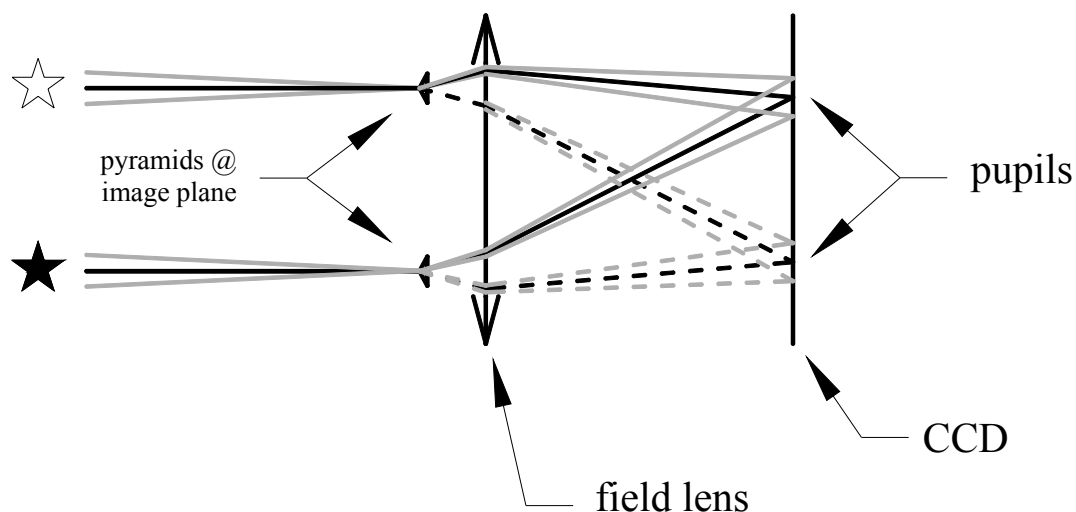
Figure 1- 6: Beam-combining scheme for layer-oriented AO. Each guide star is imaged onto its own pyramid that splits the light into 4 beams, as described in figure 1-5. A subsequent field lens common to all of the pyramids combines the beams and creates 4 pupils on the CCD. In the figure above, the black lines represent chief rays; the gray lines are marginal rays. The solid rays after the pyramids are directed to one pupil corresponding to one field quadrant/facet; the dashed rays are directed to another pupil corresponding to another field quadrant/facet.

While it has been suggested that LGS's may not be necessary for ELT's (Ragazzoni, 2000a), some groups are planning for LGS's for the same sky coverage reasons given above. The fact that ELT apertures are so large means that the LGS will appear from some subapertures to be quite elongated (~3.5 arcseconds for a center-launched LGS on a 30m telescope). While LGS spot elongation occurs on existing telescopes, the effect has been negligible because the apertures are not nearly as large as for ELT's.
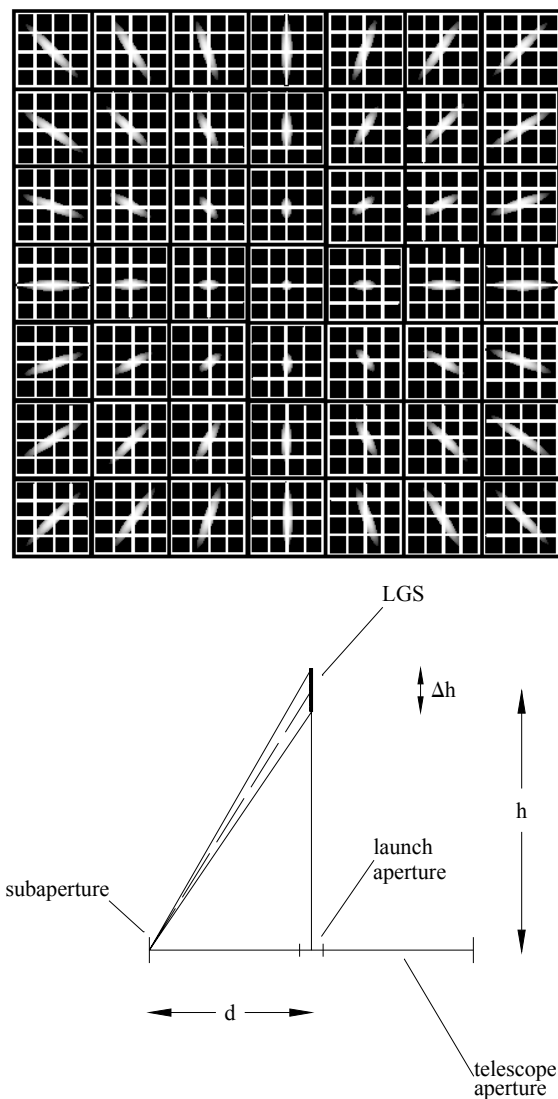
Figure 1- 7: (top) Schematic illustration of LGS spot elongation effect in a Shack-Hartmann WFS (from Goncharov). (bottom) Geometry of laser guide star showing spot elongation. The LGS, shown as a heavy line above the telescope, is located at an altitude of h with a depth of Δh. The elongation varies with the distance of the subaperture from the launch aperture (d); the left-most subaperture shown here. The angular elongation from the perspective of this subaperture is $\Delta\theta=(\Delta h/h)(d/h)=\Delta hd/h^2$.

The purpose of this work is to consider some of the optical design issues in AO systems

in ELT's. A complete discussion of all aspects of ELT AO system design is well beyond

the scope of this document, and well beyond the scope of any one person. The issues that

we will consider are some of the ones that concern an optical engineer, as opposed to astronomers or system designers. In particular, we will discuss the following issues: first-order optical design of MCAO systems, wavefront sensing of LGS's which appear elongated due to their finite depth, and pyramid wavefront sensors. These topics are introduced below.

In chapter 2, we will discuss the first-order design of a MCAO relay. It is a non-trivial task to figure out how to design an AO system that meet requirements of DM mirror sizes and altitudes at which they are placed, order of correction, and other constraints. It turns out that the Delano, or $y,\bar{y}$ method of first-order optical design is the perfect tool for considering these requirements. The $y,\bar{y}$ method is a graphical design technique that represents an optical system as a trace of connected line segments in a Cartesian plot. The (x,y) coordinates of the trace are given by, respectively, the chief and marginal ray heights at the various optical elements in the system. Figure 1-5 shows a sample optical system in the common side-view layout and as represented in a $y,\bar{y}$ diagram.
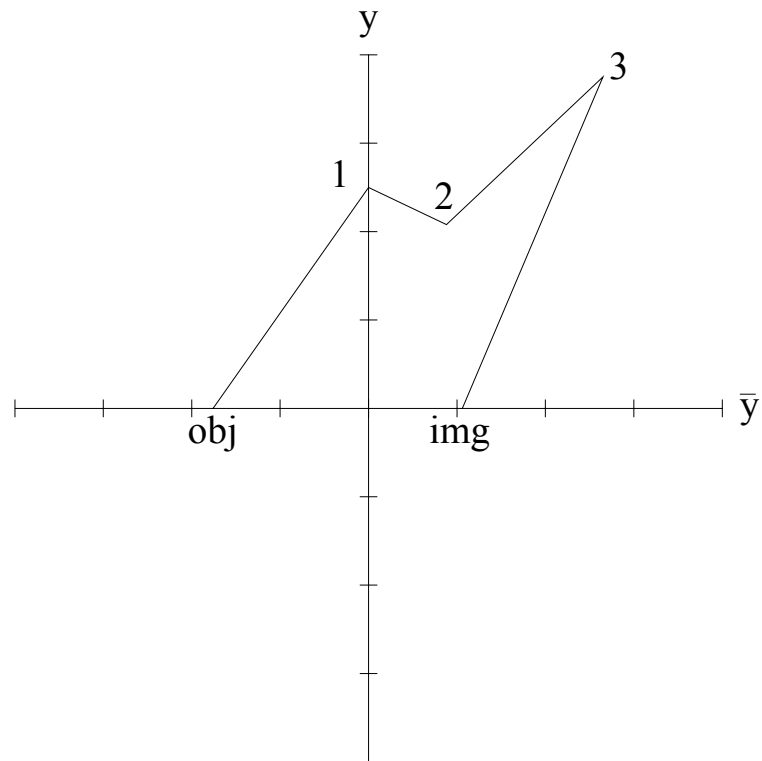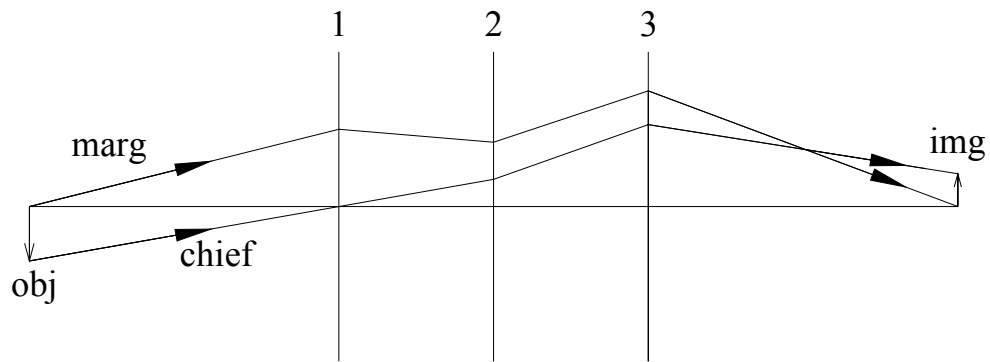
Figure 1- 8:   A sample optical system in a conventional side-view (top) and as represented in a y,$\bar{y}$ diagram (bottom).

In chapter 3, we will discuss the issues related to PWFS's which are proposed as the workhorse of LOAO (Ragazzoni 2000b). We will examine properties of PWFS's and extend the techniques to use multi-cell wavefront sensing. We will also propose a novel method for constructing a PWFS and point out how existing SH WFS's can be easily turned into PWFS's.

Chapter 4 will deal with methods to combat LGS elongation. LGS elongation imposes stiff penalties in laser power requirements and so is worth investigating (University of California and California Institute of Technology). A method will be proposed to reduce spot elongation for the case of pulsed lasers. A second method will be proposed for use with non-pulsed lasers.

For concreteness, we will deal with a specific ELT design. The lessons learned here can be extended to other currently contemplated ELT's.

The ELT design to be used here is a Ritchey-Cretien telescope with a 30m diameter, f/1.5 primary and an f/15 image located 15 m behind the primary (in an unfolded layout); this is the design used in the conceptual design of the California Extremely Large Telescope (CELT). The prescription is given below in Table 1-1. (University of California and California Institute of Technology). A layout of the telescope is given in figure 1-4.

	Radius of curvature (m)	Thickness (m)	Diameter	Conic constant
primary	-90	-39.545	30	-1.002716
secondary	-12.121	54.545	3.64	-1.524629
image				

Table 1- 1: Optical prescription of ELT to be used in this work.
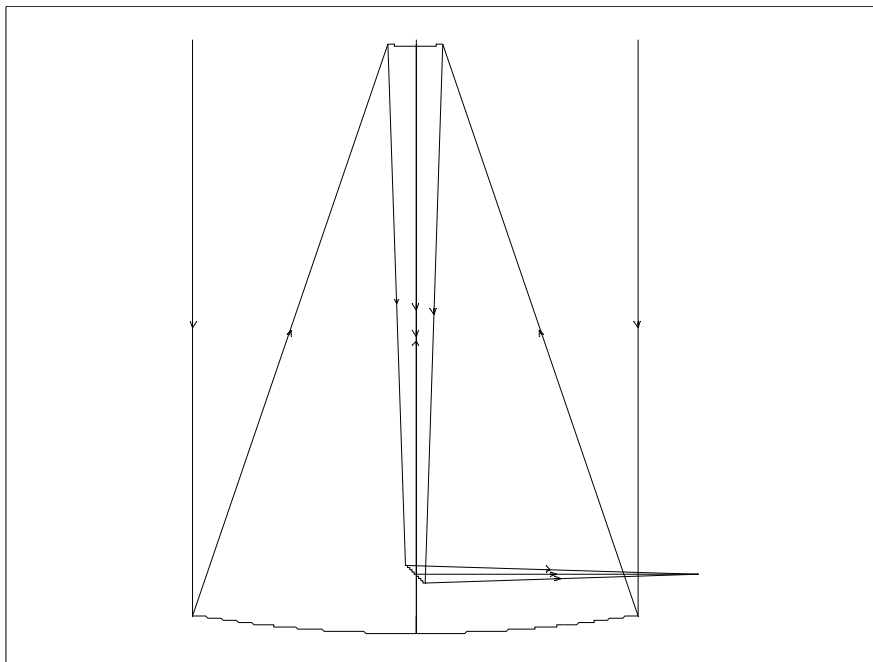
Figure 1- 9: Layout of ELT to be used in this work. The image is folded to the side of the telescope with a flat mirror. The prescription is given in the text.

The resulting first-order properties of this ELT are listed in Table 1-2 below:

Primary diameter	30m
Focal length @ cassegrain focus	450m
Front focal point	3.3km above telescope
Rear focal point	19m behind primary
Field-of-view	2 arcmin diameter

Table 1- 2: First-order properties of ELT to be used in this work.

The design of the CELT primary mirror calls for a large number (~1000) of hexagonal segments tiled together to create a 30m diameter mirror. For purposes of this work, we will neglect the non-circular nature of the aperture due to hexagonal segments.

2. THE USE OF THE Y,$\overline{\text{Y}}$ DIAGRAM IN MCAO DESIGN

2.1. Introduction

The optical design of an MCAO system logically begins with a first-order design. This allows us to identify key design issues to be answered and we will be able to state clearly the penalties for imposing various constraints upon the design that impact science return.

In this chapter, we will use the y,$\overline{y}$ design technique (Delano, Shack 1973, Shack 1991, Ditteon, Bauman 2001) for the MCAO relay design. As noted in Chapter 1, the y,$\overline{y}$ design technique is a graphical approach to first-order optical design; the method will be sketched in section 2.3 and its principles reviewed in section 2.5. While the y,$\overline{y}$ technique is not new, its application to MCAO as a superior/elegant design and communication method is new.

2.2. Choices in first-order design techniques

An optical design task begins with a set of requirements and constraints. In proceeding from these requirements to a first-order design, a designer traditionally follows one of the following approaches:

- Take an already-developed first-order design and adapt or extend it to the current application.

- Solve (either by hand, by commercial or homegrown mathematical code, or via commercial lens design code) sets of simultaneous equations where the variables are element powers and distances between elements, and the constraint equations may include desired focal length, magnification, and size constraints.

- Use an iterative, trial-and-error approach to satisfying the constraints and requirements.

The first approach is useful for "conventional" applications that are at most evolutionary from existing designs. The second and third approaches can be used for less conventional designs, but do not yield much insight into the problem or the trade-offs or opportunities therein.

2.3.   The y,$\overline{y}$ method

However, there is another design technique that uses the Delano, or "y,$\overline{y}$" diagram to represent the optical system. The y,$\overline{y}$ diagram is outlined below; the principles are reviewed in more detail in section 2.5.

For any optical system, we can trace the marginal and chief rays through the system, as shown in figure 2-1. We can then record the marginal ray and chief ray values (referred to as "y" and $\overline{y}$", and pronounced "y" and "y-bar", respectively, from here on) as shown in Table 2-1.
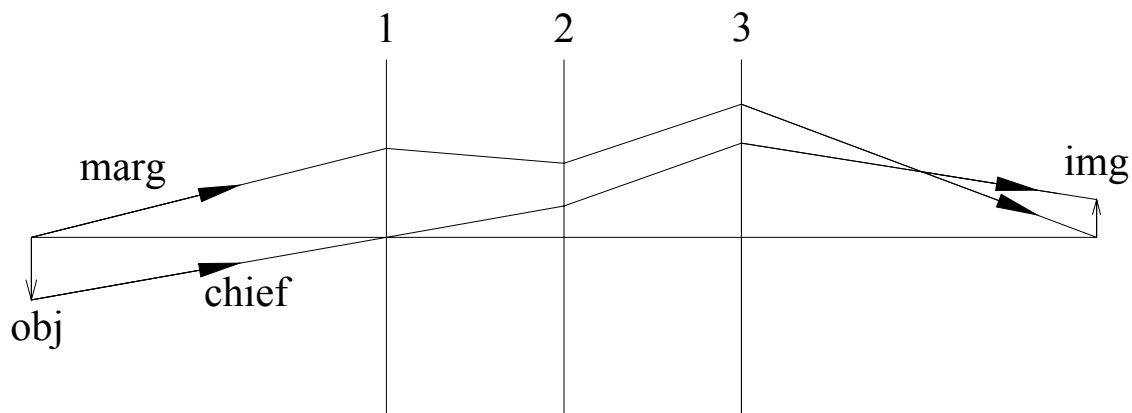
Figure 2- 1: Marginal and chief rays traced through an optical system with 3 optical elements (1, 2, 3)

	y (marginal ray height) (mm)	$\bar{y}$ (chief ray height) (mm)
Object	0	-17.6
Lens 1 (pupil)	25	0
Lens 2	20.8	8.8
Lens 3	37.5	26.5
Image	0	10.6

Table 2- 1: Marginal and chief ray heights at various surfaces in the optical system

Alternatively, we could display this same information as a series of connected line segments in a Cartesian plot with $\bar{y}$ on the horizontal axis and y on the vertical axis, as shown in figure 2-2. This plot, called the y,$\bar{y}$ diagram, is a graphical representation of the system.
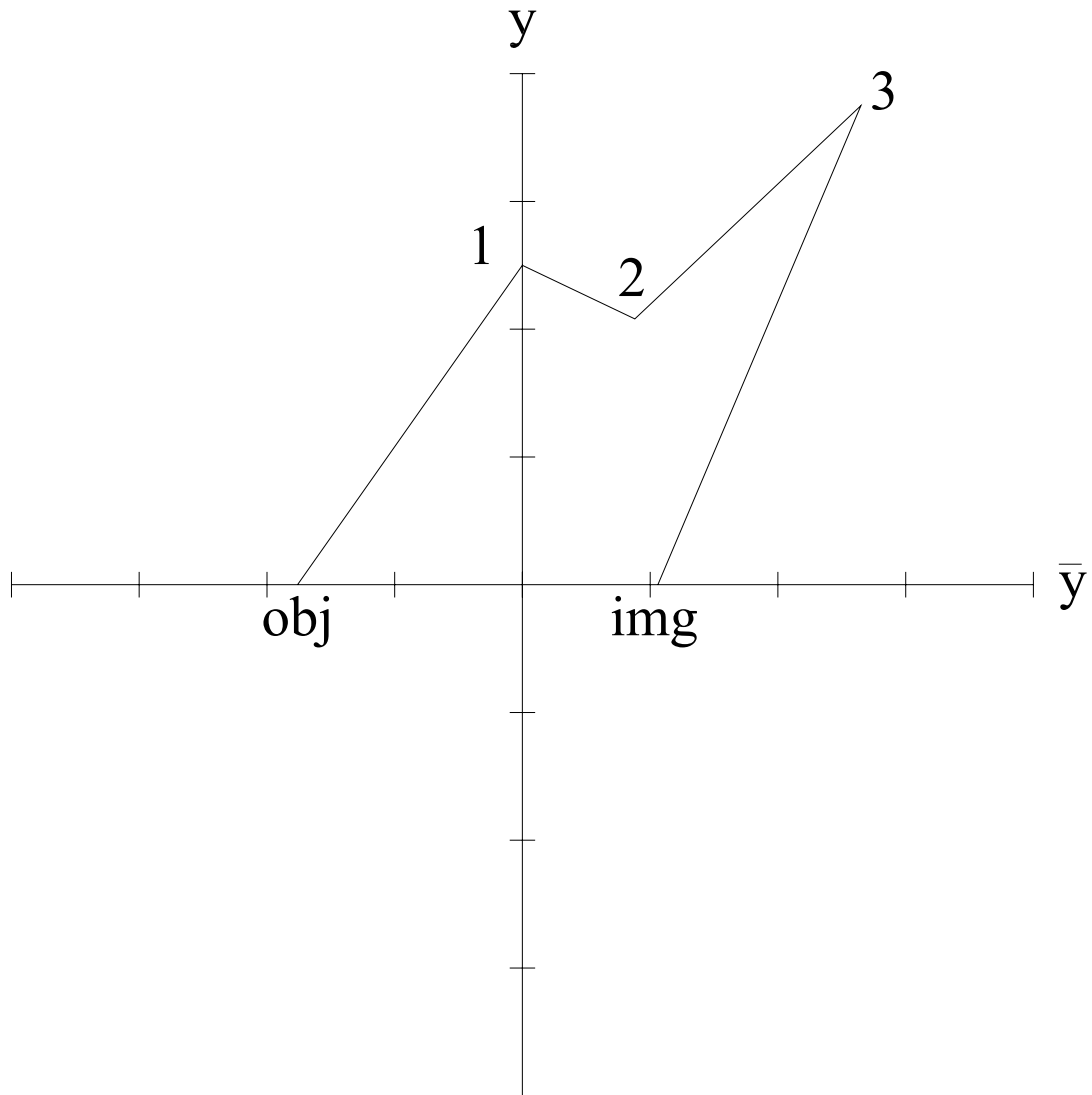
Figure 2- 2: y,$\bar{y}$ diagram of optical system represented in figure 2-1 and table 2-1. The y,$\bar{y}$ ray begins at the object and traces to the first optical element, labeled "1". Each kink in the y,$\bar{y}$ diagram represents a powered optical element. The y,$\bar{y}$ ray ends at the image point.

Comparing the $y, \bar{y}$ method to the previously-discussed first-order design techniques (extrapolation of existing design, solving simultaneous equations, and trial-and-error), the $y, \bar{y}$ method offers several technical benefits:

- Graphical representation: the first-order properties of a system and of optical elements can be represented graphically. This is a powerful aid to communicating to the non-optical engineer, who often considers optical design to be a black art.

- Physical realizability: The $y, \bar{y}$ method yields all and only physically-realizable solutions (i.e., "if and only if") as long as one simple graphical rule is obeyed. This is one of the rare cases where an optical engineer can answer with certainty that a design is or is not possible, i.e., the $y, \bar{y}$ method provides an "existence proof."

- Development of intuition and unorthodox solutions: The $y, \bar{y}$ method illuminates the tradeoffs to be considered in an optical design and points out solutions to a design task that do not use conventional optical blocks such as telescopes and afocal relays. Further, these benefits are gained while learning intuition for the design, as opposed to solving simultaneous equations or letting design codes run blindly to satisfy an underconstrained problem.

- Ease of use: No lens design software is required for $y, \bar{y}$ analysis. A spreadsheet or CAD program (my usual preference) makes the method very easy to use. It is also amenable to "white-board" discussions.

2.4.   y,$\overline{y}$ as communication tool

One of the strengths of the y,$\overline{y}$ method is that its graphical form and arguments are easily accessible and appealing to non-optics experts. The stakeholders of leading-edge telescope projects such as GSMT or CELT constitute a widely diverse group. The end-users of systems are astronomers who often have little background in optical design or in instrumentation; they often just want to know why their AO system must have so many optical surfaces that are stealing precious photons from them and adding deleterious infrared background photons. An unelaborated answer of "It can't be done" does not satiate the questioner—his response, whether stated or not, is likely to be, "Well, try harder."

At the other end of the spectrum are those that sponsor projects such as these – donors, governmental program managers, and congressional staff members. They want to know why the astronomers' concerns are not being adequately addressed, and they feel somewhat powerless in understanding issues or being able to contribute the resolution. A method such as y,$\overline{y}$ offers would-be bystanders a chance to participate.

In sum, the fact that y,$\overline{y}$ design follows easily understood rules of geometry is valuable in breaking down the barriers between users and engineers.

It is perhaps unusual to consider concerns that one might dismiss as being "sociological matters", but these very issues are part of the reality of large telescope programs. A thirty-meter telescope, along with its integral AO system, does not get built without a successful reckoning with the complex sociology of such a project. To the extent that technical issues can be aired and resolved in an inclusive manner, the project derives strength from its wide ownership.

Finally for optical engineers, the geometric nature of the $y,\bar{y}$ technique offers great power in grasping the complex first-order properties of an MCAO system. As will be seen, $y,\bar{y}$ is well suited for considering the multiple layers of correction in MCAO, the order of correction, sizes of optics (including deformable mirrors) and entrance/exit pupil considerations. These considerations will be crystallized into well-defined trade-offs. We will be able to state, for example, a requirement to correct atmospheric layers in order of increasing altitude will result in additional optical surfaces, equal in number to the number of DM's, whereas waiving this requirement allows us to have only one surface beyond the number of DM's. This is the kind of statement which brings the resolution of this question to front-burner status.

Just as important as any of these other points is the fact that a $y,\bar{y}$ discussion ends arguments. An ELT project can be delayed and made more costly by concerns that are not adequately addressed in the questioners' view. As an example, the end-users would

love to have an MCAO system which uses small (~300mm) DM's and which does not have any optics other than the DM's themselves. We will be able to state without question that this notion is simply not practical, and that no combination of money, design effort, or DM development will change that answer.

2.5.   Review of relevant $y, \overline{y}$ principles

We will now briefly review the common features and properties of $y, \overline{y}$ diagrams. Figure 2-3 provides an example $y, \overline{y}$ trace exhibiting many common features.
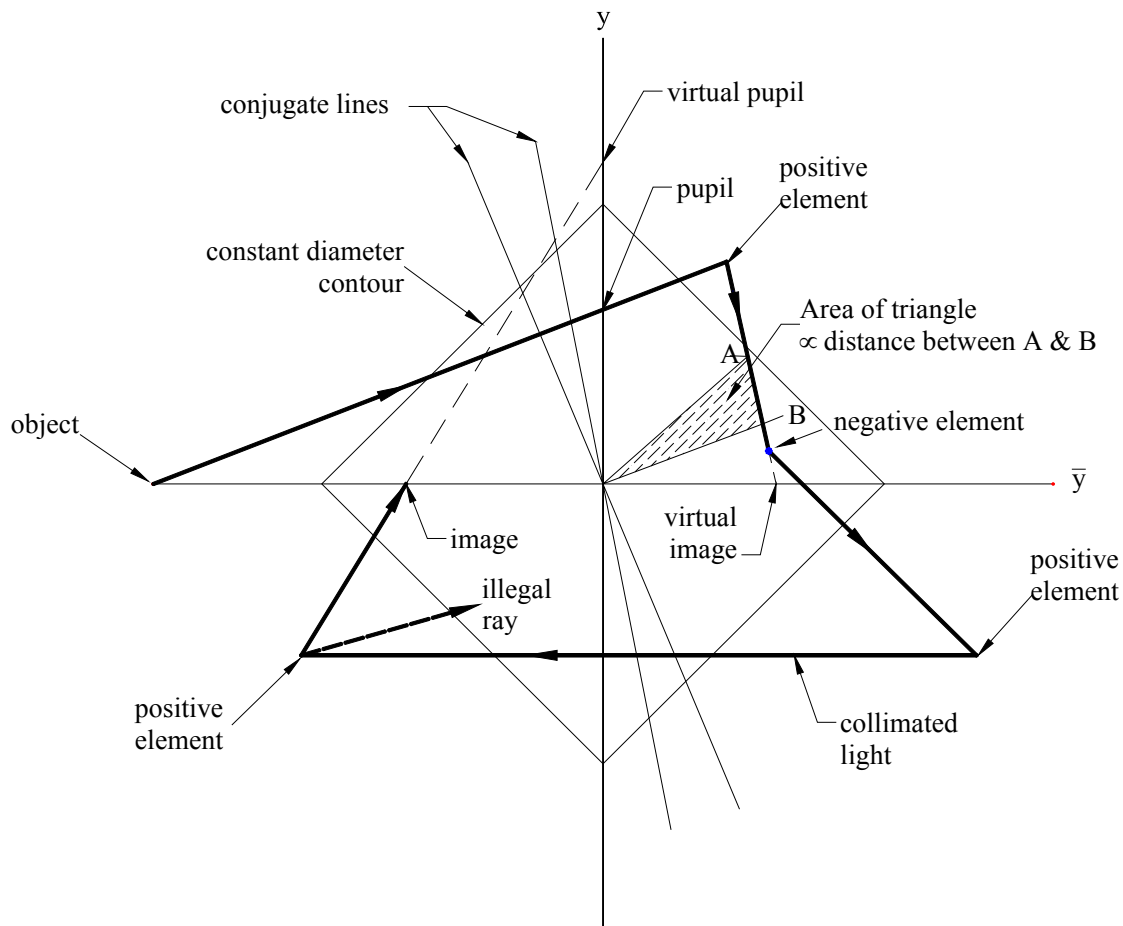
Figure 2- 3: Sample of a y,$\overline{y}$ diagram with several common features illustrated.

Collimated light:	For collimated light, the marginal ray height is constant, and so is represented as a horizontal line on the diagram.
Real object/image:	An image is formed where the marginal ray height is zero, i.e., where the y,$\overline{y}$ ray crosses the horizontal axis. The $\overline{y}$ coordinate at that point gives the height of the image. Note that we can interpret the horizontal axis (the $\overline{y}$ axis) in terms of an f/# for a given system since the image size is proportional to the f#.

Virtual object/image:	same as real image, but the y,$\overline{y}$ ray does not cross the horizontal (y) axis; rather the extension of this ray crosses the horizontal axis—this is the virtual image.
Real pupil:	For a pupil, the chief ray height is zero, so a pupil is located where the y, ray crosses the vertical (y) axis. The y coordinate at that point gives the radius of the pupil.
Virtual pupil:	Same as real pupil, but again the y,$\overline{y}$ ray in that optical space does not itself intersect the vertical ($\overline{y}$) axis; rather, the extension of the y,$\overline{y}$ ray does.
Element powers:	A positive element will cause the y,$\overline{y}$ ray to "kink" and bend the ray towards the origin. A negative element will bend the ray away from the origin.
Conjugate lines:	A line through the origin represents the locus of points that are conjugate to each other. This would seem reasonable given that the horizontal axis (which passes through the origin, of course) represents images, which are all conjugate to each other, and that the vertical axis represents pupils, which are all conjugate to each other.
Distances:	The distance between two points (A and B in figure 2-3) is proportional to the area of the triangle consisting of the two

points of interest, and the origin. The proportionality constant is the Lagrange Invariant:

$$\text{Distance between A\&B} = t_{AB} = 2\, nA_{\Delta OAB}\, /\, Ж,$$

$$\text{where } Ж = \text{Lagrange Invariant} = n(y\bar{u} - \bar{y}u)$$

$$A_{\Delta OAB} = \text{area of triangle OAB}$$

$$n = \text{index of refraction}$$

Diameter contours: The locus of points describing a constant optic diameter is represented by a diamond such that $2(|y| + |\bar{y}|) = $ diameter.

Focal length: The focal length of an element is obtained by finding the distance between the element and its focal point; see figure 2-4. The focal point is found by constructing a line through the origin that is parallel to the $y, \bar{y}$ ray in the space before the element (i.e., the object space of the element). The focal point is located where the image space $y, \bar{y}$ ray crosses this construction line. The focal length is then obtained by calculating the distance between the element and the focal point.

If the $y, \bar{y}$ ray "kinks" towards the origin at the lens, then the area is positive and the focal length is positive; this is in agreement

with the "element powers" section above.   If the y,$\bar{y}$ ray kinks

away from the origin at the lens, then the area is negative.
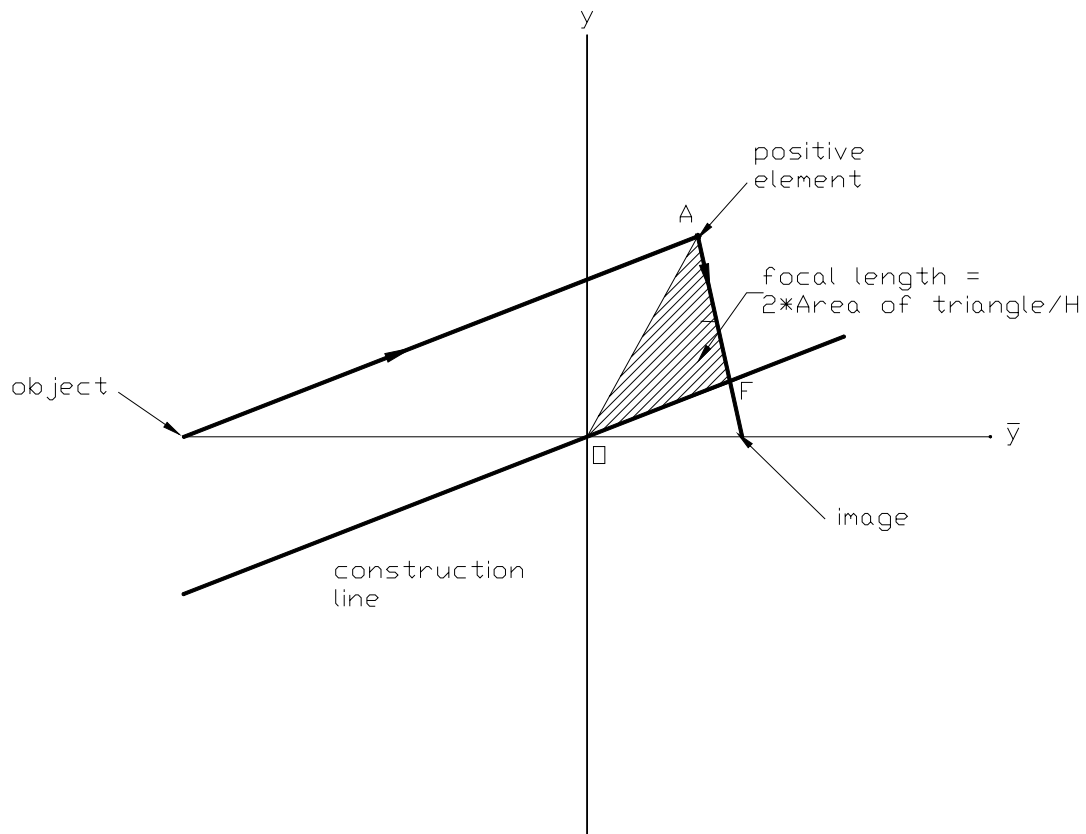
Figure 2- 4: Finding the focal length of an optical element. Construct a line through the origin parallel to the object space y,$\bar{y}$ ray. The image space y,$\bar{y}$ ray will intersect the construction line at a point F.   The focal length is 2 nA$_{\Delta OAF}$ / Ж.

Physical realizability: An optical system is physically realizable if and only if the y,$\bar{y}$ ray

always maintains the same sense of rotation (i.e., clockwise or

counterclockwise) about the origin. Physical realizability does not contemplate aberrations nor prohibit extremely fast f#'s.

Under the definition of the Lagrange Invariant used here, the $y,\bar{y}$ ray rotates clockwise around the origin if the Lagrange Invariant is positive and counterclockwise if the Lagrange Invariant is negative.

On occasion, we will need to display the $y,\bar{y}$ diagram on an anamorphic scale. This is convenient when the marginal ray heights are very large (many meters) and the chief ray heights (i.e., field angles) are very small, as with an ELT.

We state without proof that as long as the axes' labeling remains unchanged, there is no change in any $y,\bar{y}$ property, with a single exception: the "diamond" nature of the optic diameter contour line (i.e., a square rotated at 45° to the axes) is distorted to a rhombus where the angles are not 90°. Figure 2-5 demonstrates this.

2.6. $y,\bar{y}$ diagram of an extremely large telescope

For work presented later in this chapter, we will need the $y,\bar{y}$ diagram for our telescope. The specifications, presented in Chapter 1, that we will use will be:

- 30m diameter, f/1.5 primary (aperture stop)

- 4m diameter secondary

- f/15 cassegrain focus

- Field of view (FOV) is $\pm 300\mu rad \approx \pm 1 arcmin$ (the FOV of a Ritchey-Cretién telescope is much larger, but the AO system will use this smaller field)

In figure 2-5, we can see the resulting $y,\bar{y}$ diagram. As required, we have collimated space (object at infinity), so the $y,\bar{y}$ ray is horizontal. We know that the marginal ray height needs to be 15m (30m/2) and that at the primary, the chief ray height is zero (because the primary is the aperture stop). The $y,\bar{y}$ ray following the primary is aimed at $\bar{y}$ intercept of (1.5*30m)*(300µrad)=13.5mm. The secondary intercepts the $y,\bar{y}$ ray at the 4m diameter diamond contour, and redirects it towards the f/15 intercept ($\bar{y}$=135mm).
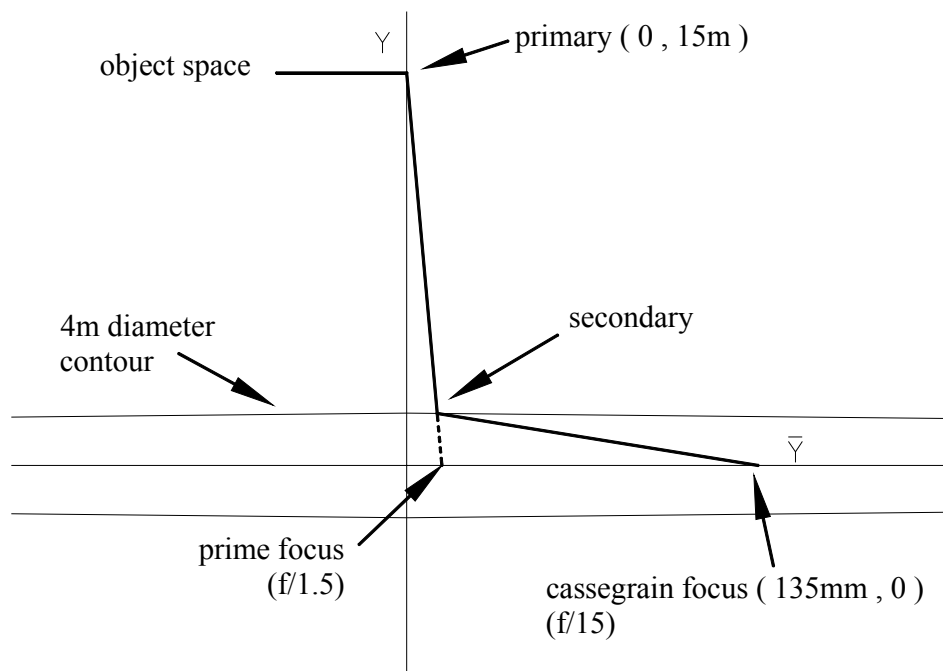
Figure 2- 5: y,$\overline{y}$ diagram of 30m, f/15 ELT with f/1.5 primary and 4m diameter secondary. Note that the scaling is anamorphic.

2.7. Application of y,$\overline{y}$ principles in MCAO design

2.7.1. Introduction

In this section, we will apply the principles of y,$\overline{y}$ to MCAO design. In the rest of section 2.7, we will implement various constraints into the y,$\overline{y}$ diagram. Then in sections 2.8 and 2.9, we will engage in a "design study" where we change various assumptions and evaluate the consequences. In section 2.8, we will perform the design study assuming no requirement on the order of correction of DM's; in section 2.9, we will assume that there is an order-of-correction requirement, from lower altitudes to higher altitudes. We will discuss this order-of-correction issue in section 2.7.6.

2.7.2.   DM diameters

Selection of DM sizes is subject to opposing design pressures: on one hand, optical performance requirements (i.e., aberrations) tend to push DM's larger so that the field angles can be correspondingly smaller; and on the other hand, the desire for smaller optical elements and shorter AO relays advocate smaller DM's. In particular, there is pressure to shrink DM's to the point where MEMS could be used. For a DM with 100 actuators across (30cm actuator spacing on a 30m telescope pupil) and 300μ actuator spacing (a commercially-available actuator spacing), the result is a 30mm DM. A quick first-order analysis indicates the possible problem. If y=15m and $\bar{u}$=300 μrad at the telescope aperture in object space, then at a 30mm diameter pupil, $\bar{u}$=300mrad=17°! This immediately raises red flags for optical designers, who know that an optical design with $\pm$17° fields with diffraction-limited performance is not a trivial matter. In addition, packaging becomes an issue; this will be discussed in section 2.8.

In appendix A a short analysis is performed which finds a minimum DM size in MCAO under a certain set of assumptions. This minimum is about 300mm in diameter and we shall use this figure later in this chapter. The family of MCAO designs using small DM's such as these is an interesting region of design space.

In this chapter, we are not considering the use of adaptive secondaries, which obviously provides another family of solutions. The methodology of this chapter, though, can be

applied equally well to derive a set of conclusions about MCAO design with adaptive secondaries.

This constraint of DM size is easily incorporated into a $y,\bar{y}$ diagram—it is simply a diamond contour such that $2(|y| + |\bar{y}|)$=DM diameter. See figure 2-6.

2.7.3. Conjugate lines for various altitudes

For MCAO systems, we will have DM's conjugate to several different heights, and these DM's will have certain size constraints upon them. We've seen how the DM size is implemented; what about the correction heights?
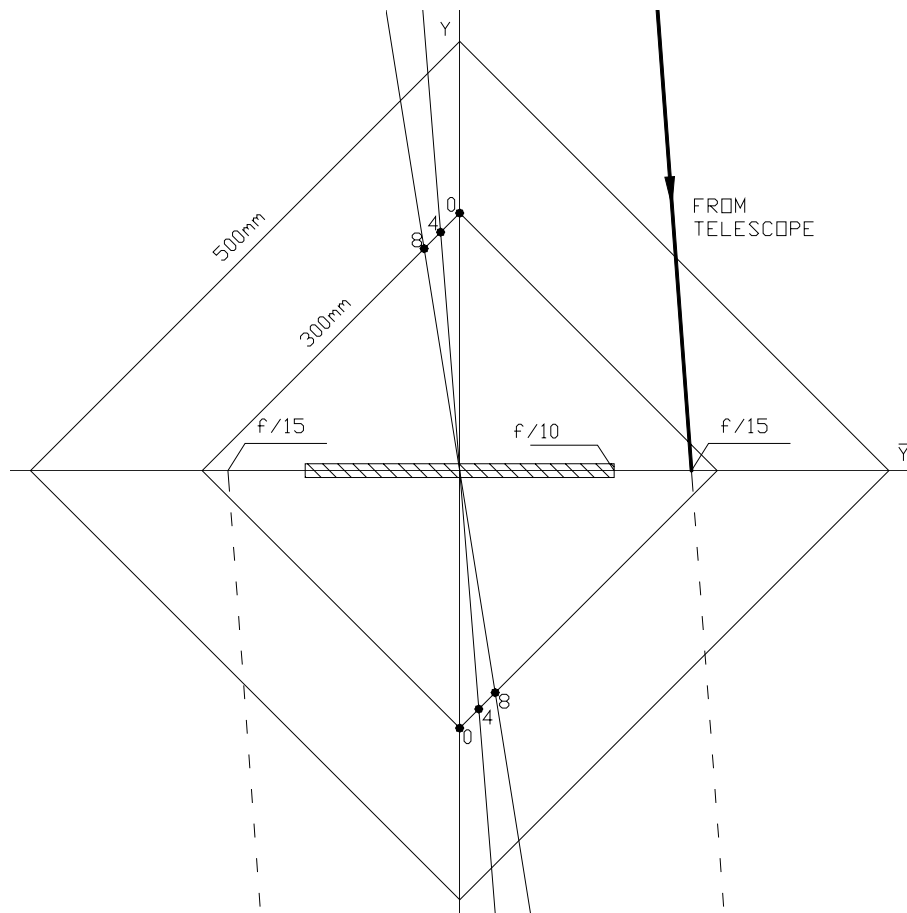
Figure 2- 6: y,$\overline{y}$ diagram of ELT with "strawman" requirements. The vertical and steeply downward-sloping lines through the origin are the conjugate lines for altitudes 0km, 4km, and 8km. Points labeled 0, 4, and 8 refer to DM's which are 300mm and located conjugate to 0km, 4km, and 8km altitudes respectively; note that there are two such points for each altitude. The solid y,$\overline{y}$ ray is from the telescope (see figure 2-5); the dashed extension of this ray and the similar dashed line on the left side of the diagram represents y,$\overline{y}$ rays with the same image and pupil properties as the telescope. A "keep-out" area corresponding to faster than f/10 is indicated.

Again, y,$\overline{y}$ provides an easy implementation of this constraint. As seen earlier in this chapter, a line through the y,$\overline{y}$ origin is the locus of points conjugate to one another.

Where the $y,\overline{y}$ ray intersects this conjugate line, that point is conjugate to all other intersections of the $y,\overline{y}$ ray and that line.

Thus, the various correction altitudes are implemented as lines through the y origin, with each height possessing a different slope. The line is most easily found by finding the (y, $\overline{y}$) values for a given height and drawing a line though the origin at the line. For a 30m diameter telescope and a 300μrad field, the marginal ray height in object space will be (telescope diameter)/2 since the light is collimated at this point. The chief ray height is given by $\overline{y}$= (altitude)*(field angle). See figure 2-6 for examples of conjugate lines with heights of 0, 4, and 8 km.

These heights have been chosen in a somewhat arbitrary manner, without regard to, say, a specific telescope site or a specific set of system performance requirements. Different sites or performance requirements may well require different sets of correction heights. These heights have been chosen as a set of plausible heights that illustrate the $y,\overline{y}$ method.

2.7.4. F-numbers

We want to have some minimum f# just so that we don't stray into unrealistic regimes. To start, we will use a minimum of f/10. This is implemented in $y,\overline{y}$ as a "keep-out"

region along the horizontal axis. Not only should the y,$\overline{y}$ ray not cross this keep-out area, but also the extensions of each y,$\overline{y}$ segment need to keep-out of this area.

2.7.5. Exit pupil and image properties

We will initially assume that the exit pupil and image size at the output of the MCAO system must match that of the telescope itself. It may well be that the MCAO system will have a dedicated science camera that can be designed to the output space requirements of the MCAO system rather than the other way around. Nevertheless, we start with this constraint so that we can see its effect on the design.

2.7.6. The "Hardy conjecture"

Hardy has conjectured that in an MCAO system that the correction must be done from lower altitudes to higher altitudes. For example, in an MCAO system with DM's conjugate to 0, 4, and 8km, the correction must occur in this order. In other words, aberrations must be undone in a LIFO (last in, first out) manner or the correction will be less than optimal. While this point is uncontested, the question is to what degree does violating this Hardy conjecture degrade the image. Flicker has performed a numerical simulation which models the Gemini telescope for a specific model of the atmosphere and field of view. His analysis suggests that the order of correction does not greatly degrade the images for most IR wavelengths, but that the effect on visible wavelengths

may become significant. Since this is an analysis for a specific set of conditions, it is not a definitive analysis. In short, in response to the question, "Does the order of correction matter?", the answer is "maybe", depending on the wavelength and on the conditions.

Since the order of correction may matter, we will analyze both cases: where the correction is done in the "correct" order, and where the correction is done in the "wrong" order. In section 2.8, we will design assuming that we need not comply with the Hardy conjecture (the "wrong" order). Later, in section 2.9, we will design so as to satisfy the Hardy conjecture (the "right" order), and look at the implications. Having performed both designs, we see that the Hardy conjecture is a key design issue—correcting the atmosphere in the right order will cost a number of additional optical surfaces relative to the correcting the atmosphere in the wrong order.

2.7.7. "Strawman" MCAO relay requirements

For concreteness, we will impose a set of "strawman" requirements on our MCAO design. These constraints are given in the table below:

Object space pupil diameter and field angle	30m diameter, $\pm$ 300µrad field angle
Telescope design	F/15 with f/1.5 primary
DM correction altitudes	0,4,8 km
Order of correction	any
DM diameters (ignoring elongation due to angle of incidence; to be considered)	300 mm
Image and pupil output after AO system	Same as for telescope
Minimum f/#	f/10

Table 2- 2: "Strawman" requirements for an MCAO system

For use in calculating distances later, the resulting Lagrange Invariant is (30m/2)*(300µrad)=4.5mm.

2.8. y-$\overline{y}$ MCAO design with Hardy conjecture "off"

In this section, we will consider the AO relay design assuming that we need not comply with the Hardy conjecture.

The constraints of section 2.7.7 are implemented in figure 2-7 below. The solid ray on the right side of the drawing indicates the y,$\overline{y}$ ray from the telescope to the Cassegrain focus (f/15 at cassegrain image, f/1.5 primary, 4m diameter secondary). Only a portion

of the $y,\bar{y}$ ray from the telescope is shown for clarity. The dashed lines indicate other optical spaces that share the same image and pupil properties as the telescope itself.

The 300mm diameter diamond contour is shown (as well as some other diameters); the 0, 4, and 8km conjugate lines are the vertical and steeply downward-sloping lines. The intersections of the conjugate lines and the diamond contour represent 300mm diameter DM's which are conjugate to the various heights. Thus, any design that meets our requirements must have a $y,\bar{y}$ ray which intersects these points. Note that for each conjugate height there is a pair of points of intersection. The $y,\bar{y}$ trace must intersect one of each pair of points.

The horizontal axis intercept represents the height of the image; thus, we can interpret the horizontal axis in terms of f/#'s. The vertical axis intercept (y axis) indicates the pupil size. Thus for any design to match the pupil and image size requirements, the $y,\bar{y}$ ray must finish along one of the rays in the diagram.

In the remainder of this section, we will evaluate the benefits of modifying or relaxing these constraints. The next several diagrams show $y,\bar{y}$ traces that meet variations on the above requirements. For reference, several of these $y,\bar{y}$ diagrams will be accompanied by a traditional, side-view layout. The layout will be rendered as a transmissive system (i.e., all optics are shown as being transmissive), for clarity and convenience.
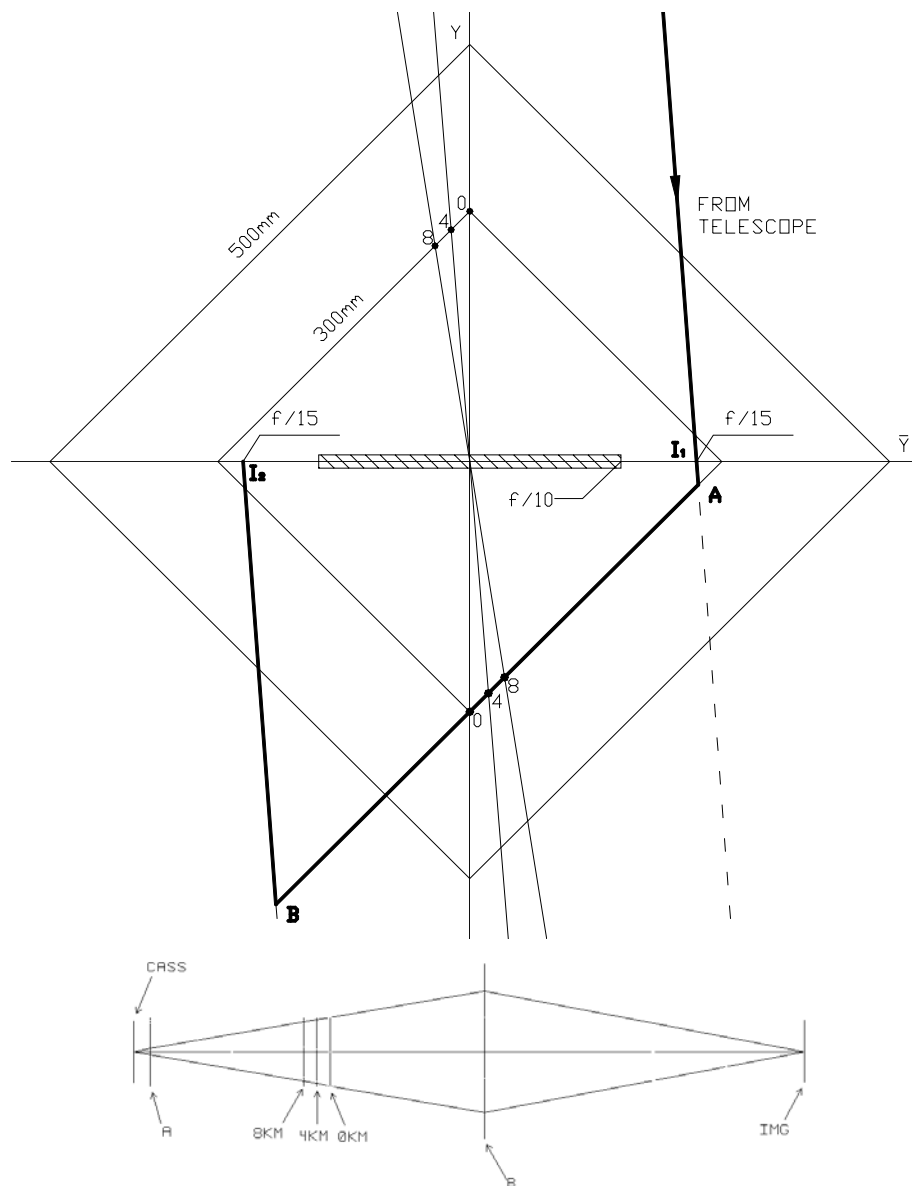
Figure 2- 7: y,$\overline{y}$ diagram showing a design with two powered relay optics (A and B).

The 3 DM's are located at the bullets labeled 0, 4, and 8 in the lower half of the y,$\overline{y}$ diagram. The exit pupil and image size properties of the telescope have been retained. A is 300mm in diameter, B is 760mm in diameter. For scale, the distance between A and B is 8.4m and the overall length from I_1 to I_2 is 18.4m. For reference, a traditional side-view layout, rendered with transmissive elements, is shown at the bottom of the figure. The scale is anamorphic.

Figure 2-7 shows the simplest solution for the requirements stated. It is apparent that 2-powered optics are required (from here, "relay optic" will refer to a powered optical element other than a DM). Starting with the $y,\bar{y}$ ray from the telescope, we see that the ray will not intersect the required points; thus, an optic is needed to bend the $y,\bar{y}$ ray towards the DM's. Since the DM bullets are collinear, the logical place for this relay optic is at the intersection of the $y,\bar{y}$ ray from the telescope and the aforementioned line of DM bullets; this yields point A.

Since the ray from A to the DM bullets will not intersect the $\bar{y}$ axis (i.e., this space does not have a real image), a second relay optic is required to bend the $y,\bar{y}$ ray back towards a real image on the $\bar{y}$ axis. This relay optic is logically placed at the intersection of the $y,\bar{y}$ ray extending from A and the ray through I_2 on the left side of the diagram that yields the appropriate output space image/pupil characteristics; this is point B.

Note that the DM's are not in collimated space. This is permissible, as the notion of DM conjugate to heights in the atmosphere does not require any particular state of vergence. Most AO systems do not choose this option, but the $y,\bar{y}$ diagram makes explicit this possibility.

Figure 2-8 shows a similar solution but with the DM requirements relaxed slightly so as to allow the DM's to be in collimated space. A and B are now more equally sized. A

visual comparison of the areas of polygons I_1ABI_2 in figures 2-7 and 2-8 shows that the overall length is approximately the same.
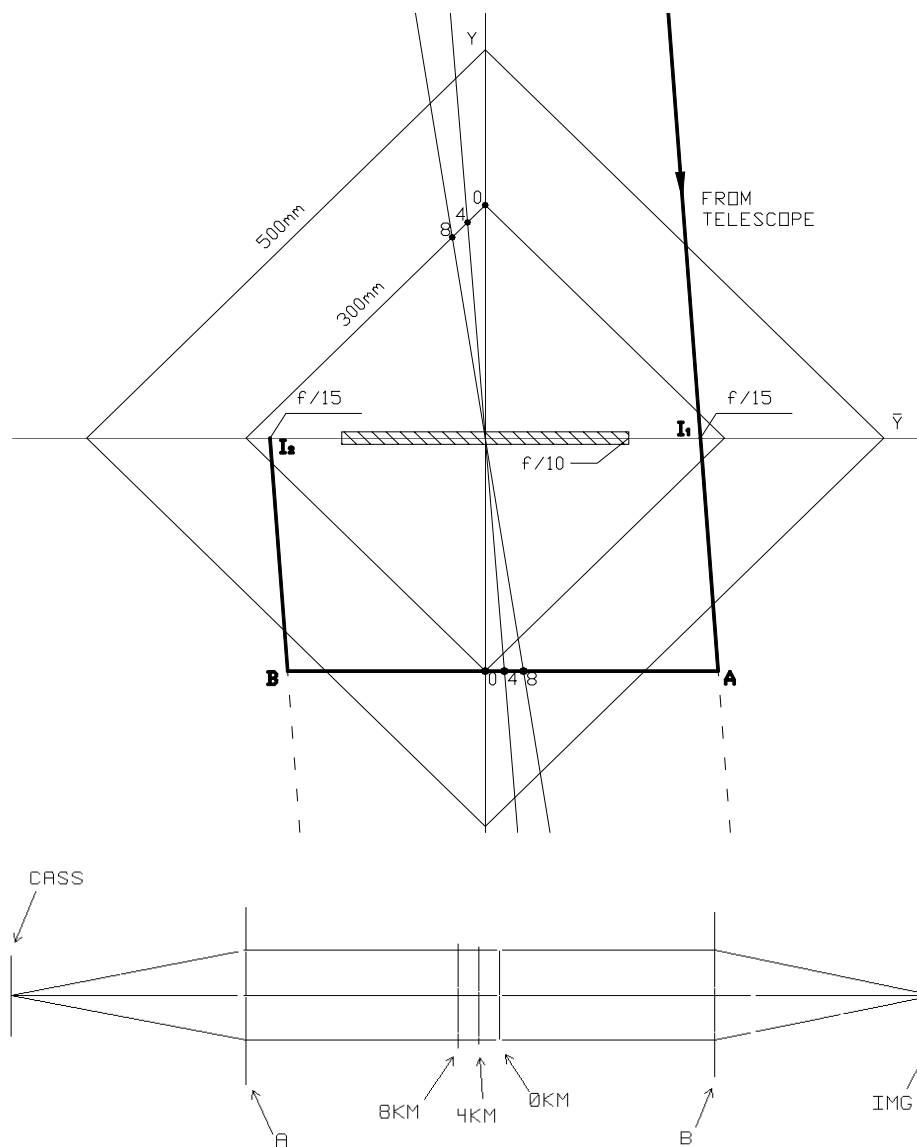
Figure 2- 8: Another solution, but with the DM size requirements relaxed slightly (the DM's range from 300mm to 350mm). This allows the first powered relay optic to be a bit further from the cassegrain focus, at the cost of a larger size. Note that the DM's are now in collimated space and the overall length is unchanged from figure 2-7.

One might consider removing the exit pupil requirement. Since the AO system will have much higher resolution than any seeing-limited set-up, one might argue that the science camera will be designed specifically for the MCAO system, and therefore the exit pupil can be relaxed to something more convenient.

Two implementations are shown in figure 2-9. The solid ray shows some length being cut out relative to figure 2-8, while keeping the same f#; the relay optic "B" from figure 2-8 has been shifted to B' in figure 2-9. The dashed ray shows a solution where one optical surface has been removed (and the system shortened as well) by making the 0km DM a powered surface. In addition, we have removed the collimation constraint, so as to shorten to the minimum possible—replacing "A" with "A'''" in figure 2-9; the $y,\bar{y}$ ray would then follow $I_1A'B''I_2$. This design would have an overall length of 9.7m.

It is well worth noting that adding DM's to this "non-Hardy" layout would not increase the length of the system appreciably, nor add any relay optics.
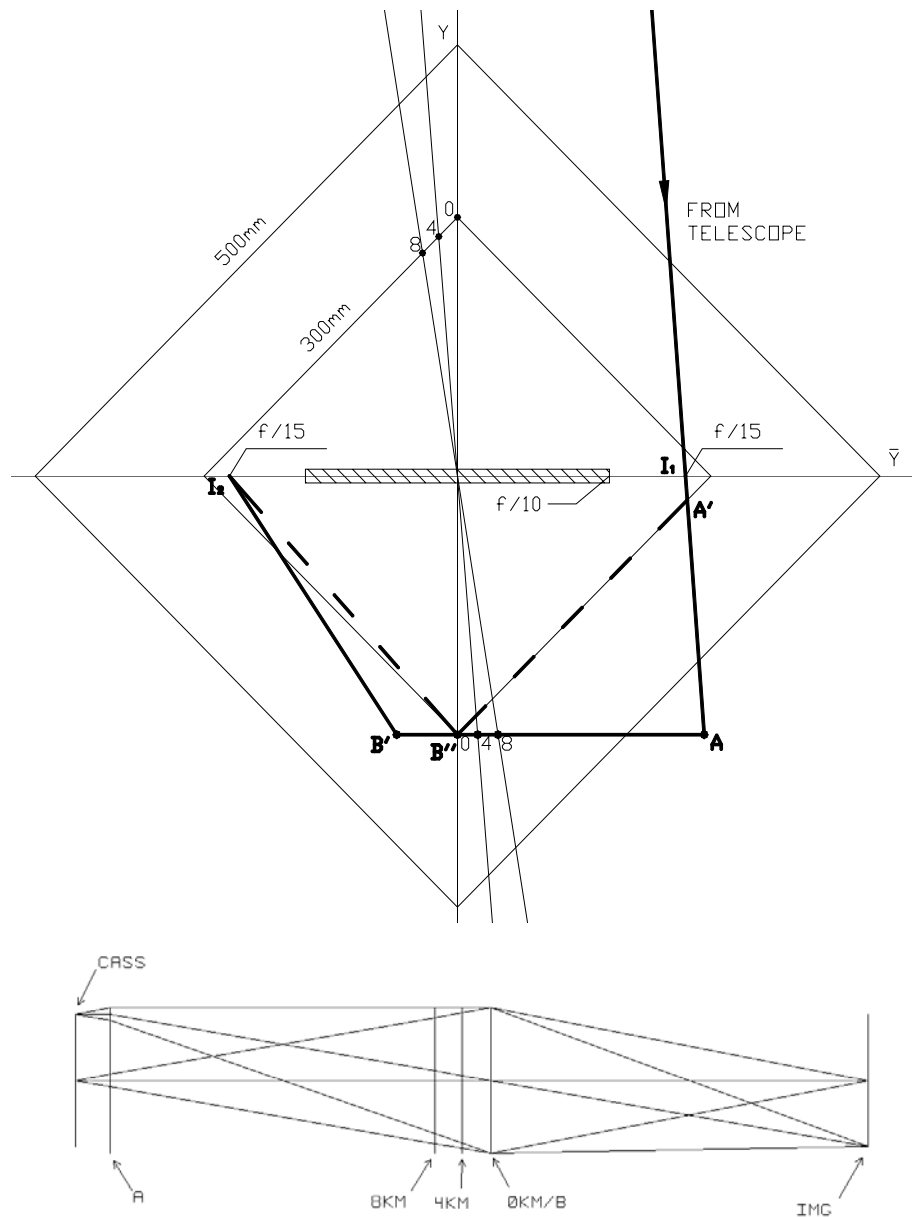
Figure 2- 9: Similar to figure 2-8, except that the exit pupil requirements have been lifted. Two solutions are shown here: the solid ray, with a powered relay optic at B', is somewhat shorter than the solution in figure 2-8. The dashed ray indicates a solution where a relay optic has been eliminated by allowing the 0km DM to be a powered optical surface and some length has been trimmed by moving A to A'; the overall length of this solution from I_1 to I_2 is 9.7m.

The dashed line solution in figure 2-9 requires only one relay optic, but at the cost of making one DM powered. Is it possible to have only one relay optic and flat DM's? The answer is shown in figure 2-10. We see that lifting the exit pupil constraint has allowed a solution that requires only one relay optic in addition to the DM's. Figure 2-10 has the same number of surfaces as the dashed solution in figure 2-9.— the tradeoff is powered DM's in collimated light versus flat DM's in non-collimated light and a large relay optic.
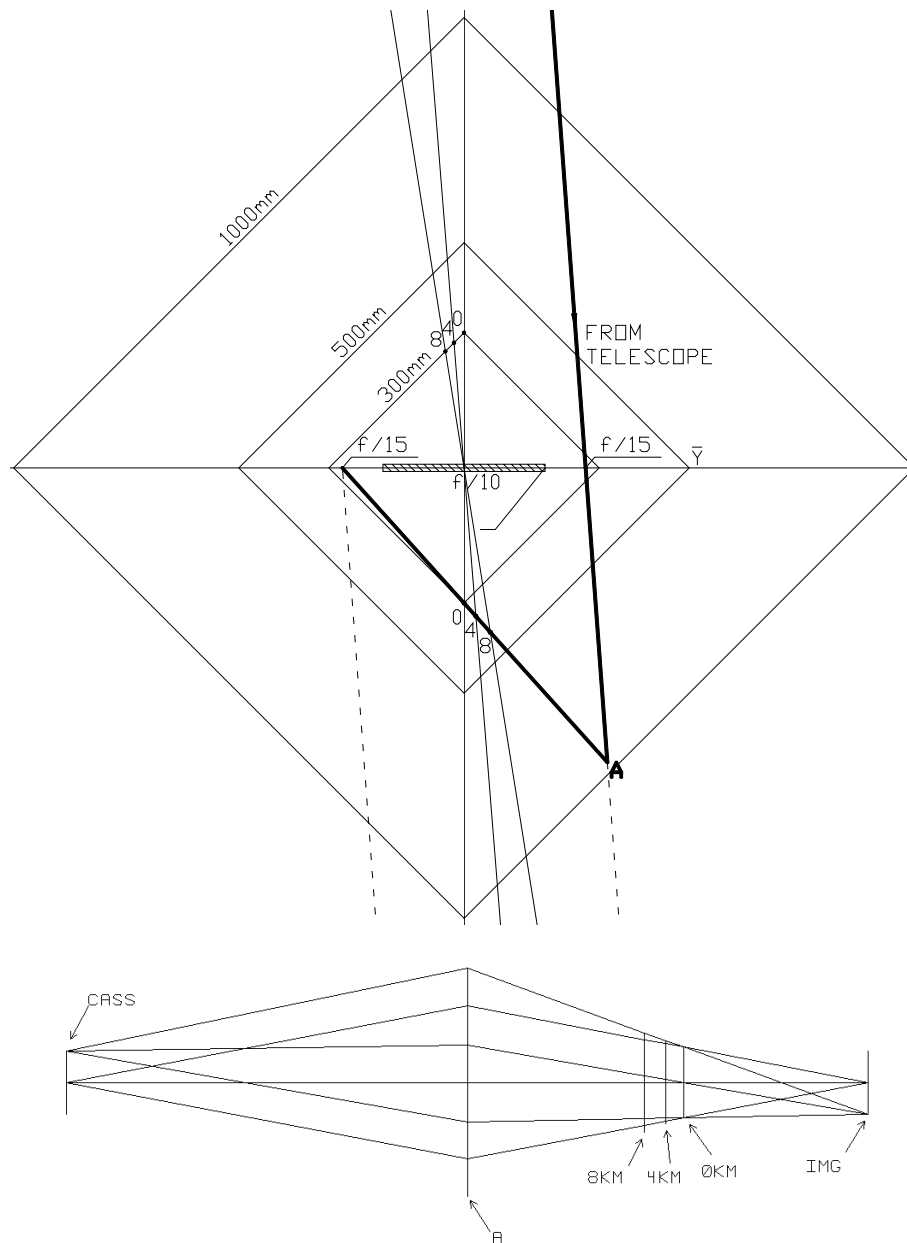
Figure 2- 10: Another solution where the exit pupil constraint has been removed and the DM size constraint has been relaxed. This solution requires only one powered relay optic, although it is large—nearly 1m in diameter. The DM's range in size from 300 to 425mm in diameter. Note scale has been changed for space reasons.

Having removed one relay optic, could we remove the other? This is a possibility by allowing a powered DM, as seen in figure 2-11; note that the $y,\bar{y}$ ray kinks at the 8km DM position. However, the DM is quite large (>3.5m) and the whole AO relay has become impractically long (>200m). Given our set of assumptions, there is no way to escape these long distances and large diameters—there must be an optical element at the intersection of the telescope $y,\bar{y}$ ray and the 8km conjugate line. Only shifting the 8km DM to a much higher altitude will reduce the lengths/diameters, as seen in figure 2-12.
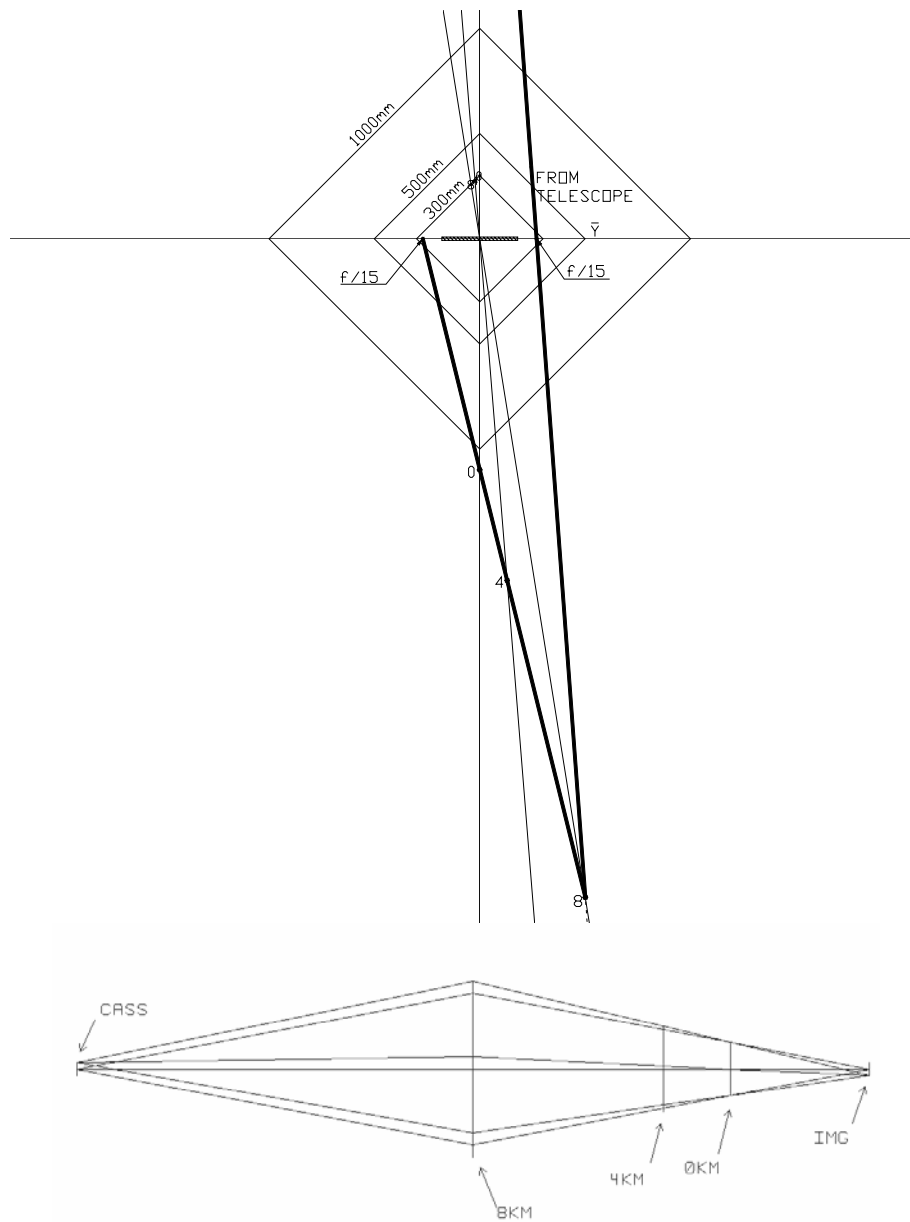
Figure 2- 11: y,$\overline{y}$ trace of system with no optics other than the DM's. The y,$\overline{y}$ ray from the telescope intersects the 8km line far from the origin, so the diameter is large. The 8km, 4km, and 0km DM's are 3.6m, 1.75m, and 1.1m in diameter, respectively. The distance from the cassegrain focus to the final f/15 image is over 94m!! Since this is the only way to avoid adding a relay optic to the system, we conclude that for practicality, at least one relay optic will be necessary.
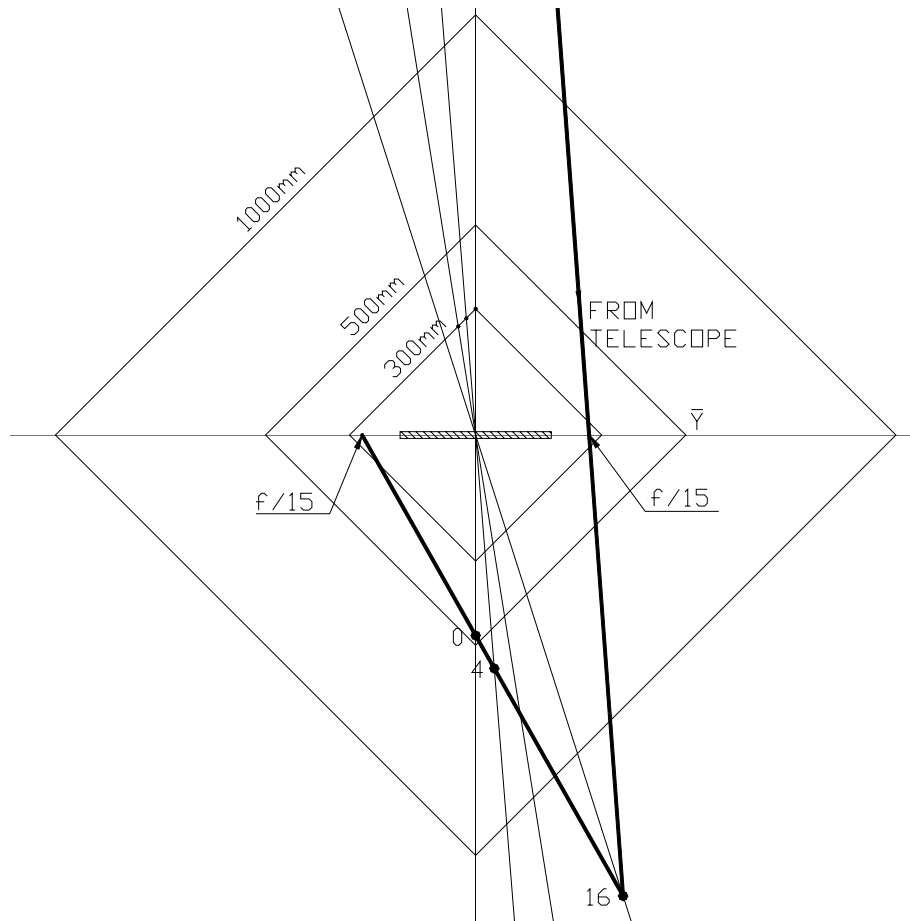
Figure 2- 12: Similar to figure 2-11, but solution has been shortened and DM's made smaller by using a higher correction height (16km instead of 8km). The 16km DM is 1.45m in diameter and the overall length from the cassegrain image to the final image plane is 32.7m.

Conclusions

From the above examples, we draw the following conclusions about first-order AO relay

design without the Hardy order-of-correction constraint:

- At least one relay optic is inevitable without excessive length/diameters or much

  higher correction heights.

- No more than 2 relay optics are necessary.

- A requirement that DM's are in collimated space results in either one additional relay optic or a powered DM.

- Under the 2 relay optic case, some length (4.5m) can be reduced by relaxing the exit pupil requirements.

- Adding DM's would not add length to the system, nor any optics.

In this next section, we will consider the impact of the Hardy conjecture, i.e., requiring that we correct aberrations in a LIFO manner.

2.9. y,$\overline{y}$ MCAO design with Hardy conjecture "on"

We will now perform a y,$\overline{y}$ design of an MCAO system where we comply with the Hardy conjecture, i.e., the DM's must occur in the order 0, 4, 8km.

First, we establish that the rules of the y,$\overline{y}$ diagram do not allow the design to be as "direct" as in the previous section. In other words, we will not be able to put the 0, 4, 8km DM's in that order, in the same optical space, without additional relay optics. As an introduction, Figure 2-13 shows that trying to "pick up" the 4km DM after the 0km DM would require the y,$\overline{y}$ ray to switch its sense of rotation about the origin, and so is forbidden. Figure 2-14 shows a "legal" way of picking up the DM's in Hardy's order. Figure 2-15 shows the 0 and 4km DM's in the same optical space, but only at the price of

unreasonably fast f#'s. Thus, we conclude that we will not be able to place multiple

DM's in the same optical space; that is, a relay optic will be required between DM's.
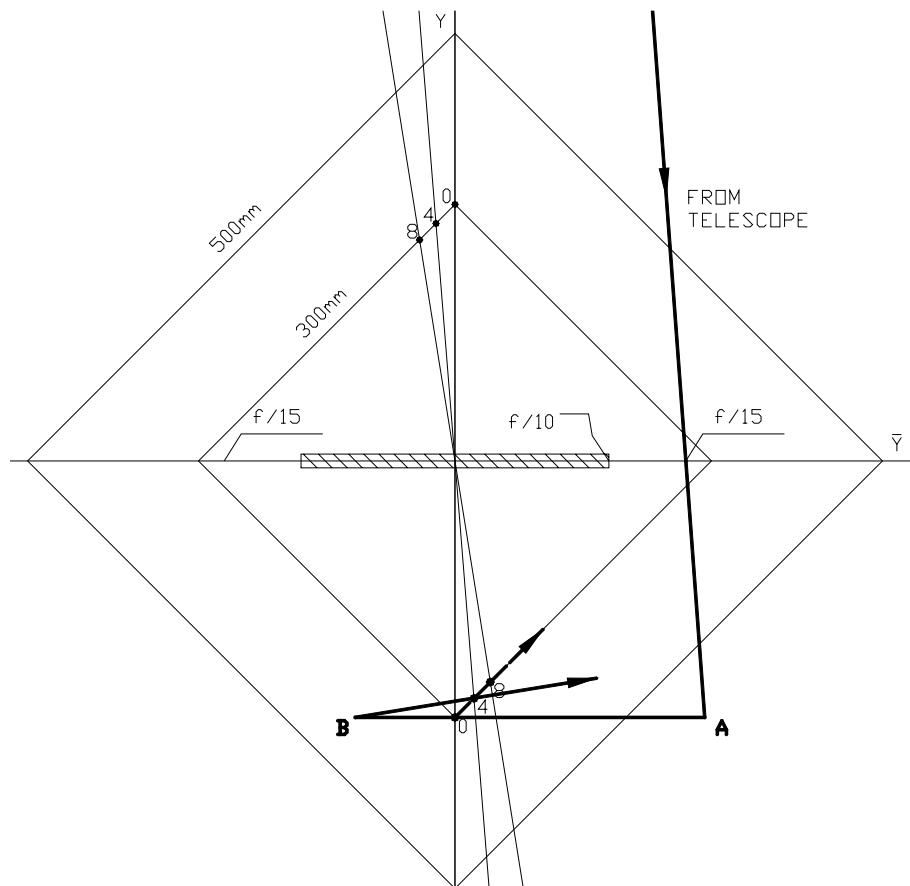
Figure 2- 13: Example of violating "sense of rotation" rule. The y,$\overline{y}$ ray from A to B "picks up" the 0km DM, but the next segment (extending from B) that would pick up the 4km DM has the opposite sense of rotation about the origin, and so is not physically realizable. The y,$\overline{y}$ ray from the telescope and from A to B has a clockwise rotation about the origin, whereas the ray from B to the 4km DM would be counter-clockwise. We can also see that any ray from the 0km DM to the 4km DM on the lower half of the diagram has counterclockwise rotation about the origin (ray from 0 to 4).
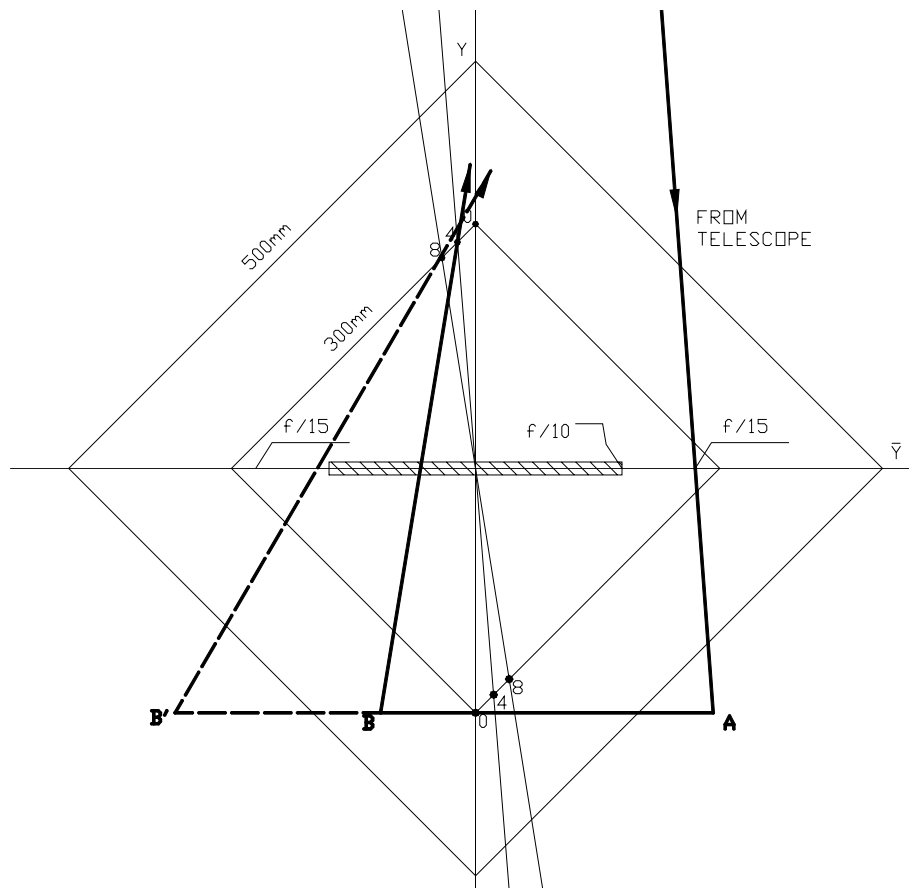
Figure 2- 14: Example of y,ȳ trace that does not violate the sense of rotation rule. After the 0km DM is picked up in the space between A&B, the 4km DM is picked up in the space following B. Note that the f# is quite fast (f/2.5) in this space. The beam can be slowed down, as shown by the dashed line, but at the price of greater diameter at optic B', and more space (the area of the triangle B-B'-4km).
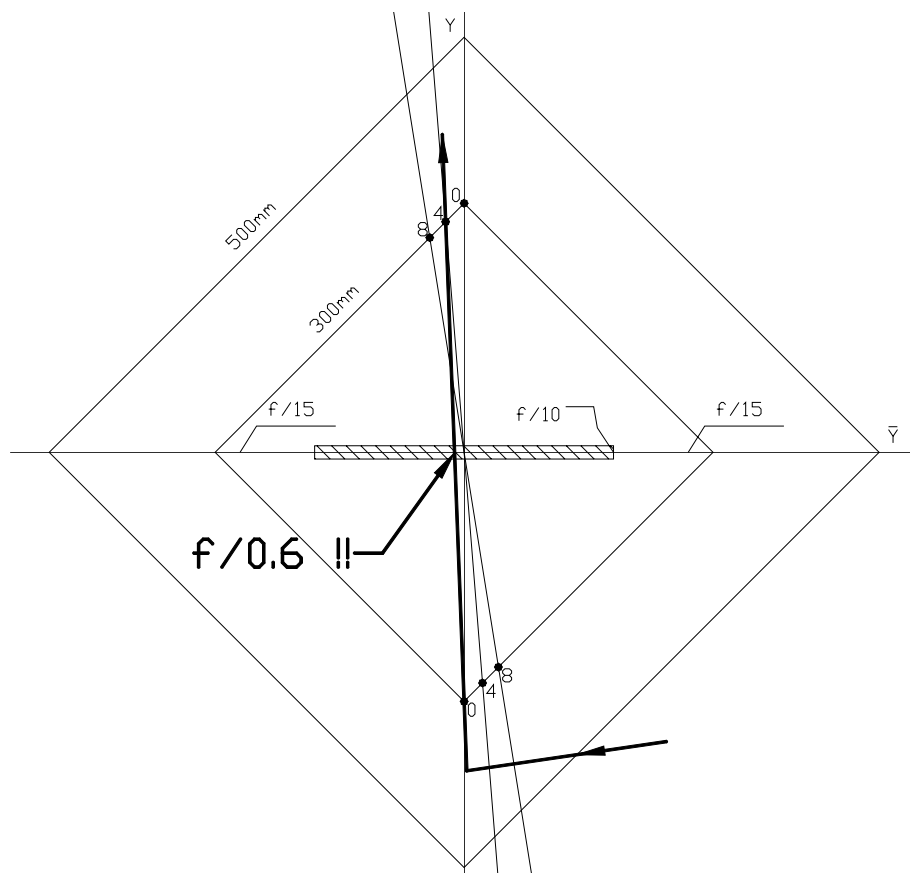
Figure 2- 15: This figure shows a y,$\overline{y}$ ray picking up the 0km DM and then picking up the 4km DM on the other side in one "pass", i.e., in one optical space.   While theoretically "legal", this results in an f/0.6 beam which is impractical. Even if the y,$\overline{y}$ ray proceeds from the 0km DM to the highest altitude DM (8km in this case), we see that the f# is still 1.2! We conclude that it is impractical to pick up two or more DM's in the same optical space according to the Hardy conjecture, i.e., lower altitudes first.

We have seen that we need at least one relay optic between DM's under the Hardy

assumption.   From the section 2.7, we have seen that we will also need a relay optic

between the telescope and the first DM; otherwise we will have a very large DM and a

very long relay (figure 2-11).   Thus, for a 3 flat-DM design, the minimum number of

optics is three: one between the telescope and the 0km DM, one between the 0km and

4km DM's, and one between the 4km and 8km DM's. How can we find a design that meets the minimum number of optics?

As a starting point for a design, let us assume the following constraints on our design: 1) flat DM's; 2) nowhere is the beam faster than f/10; and 3) each optical space is either collimated or telecentric. This third requirement is fairly restrictive, but adopting it initially is illustrative.   Figure 2-16 shows the step-by-step development of the design. The resulting design (figure 2-17) is informative: we are forced to a solution that has *six* optical surface other than the DM's, with 2 optics between each pair of DM's!
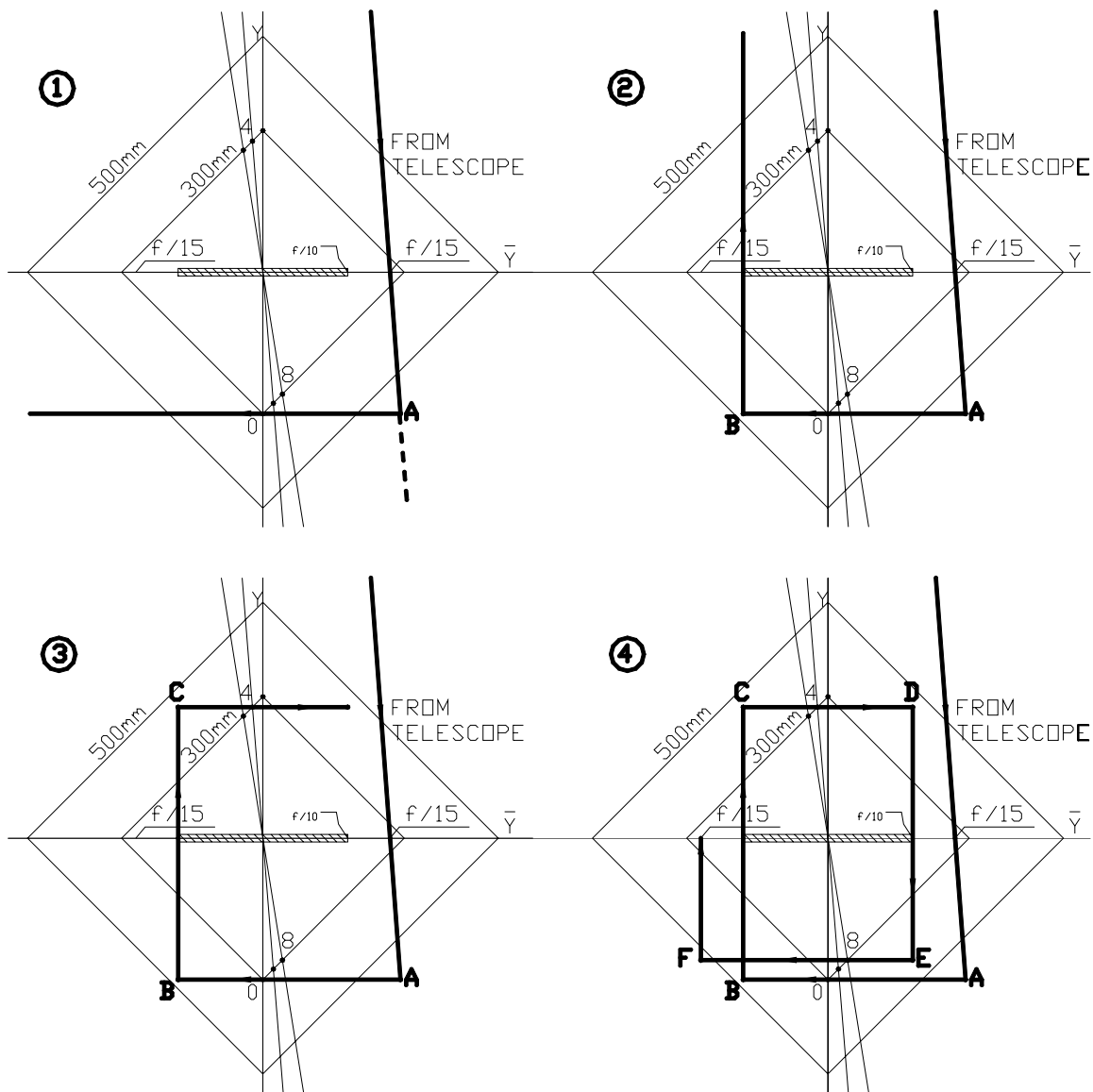
Figure 2- 16: Step-by-step development of result in figure 2-17. In the upper left drawing, we start with an extrapolation of the telescope $y,\overline{y}$ line. We know that the next $y,\overline{y}$ ray must intersect the 0km DM and that the light is to be collimated; thus, the $y,\overline{y}$ ray through the 0km line must be horizontal. Where the two lines intersect, we will have an element ("A"). We know that the next segment of the $y,\overline{y}$ ray must be vertical (telecentricity), and that it must not cross the f/10 keep-out area, so we construct a vertical line through the f/10 point; where this vertical line crosses the horizontal line through the 0km DM is the location of the next relay optic ("B"). This same process is continued in the third drawing (lower left). The final drawing is the finished result.
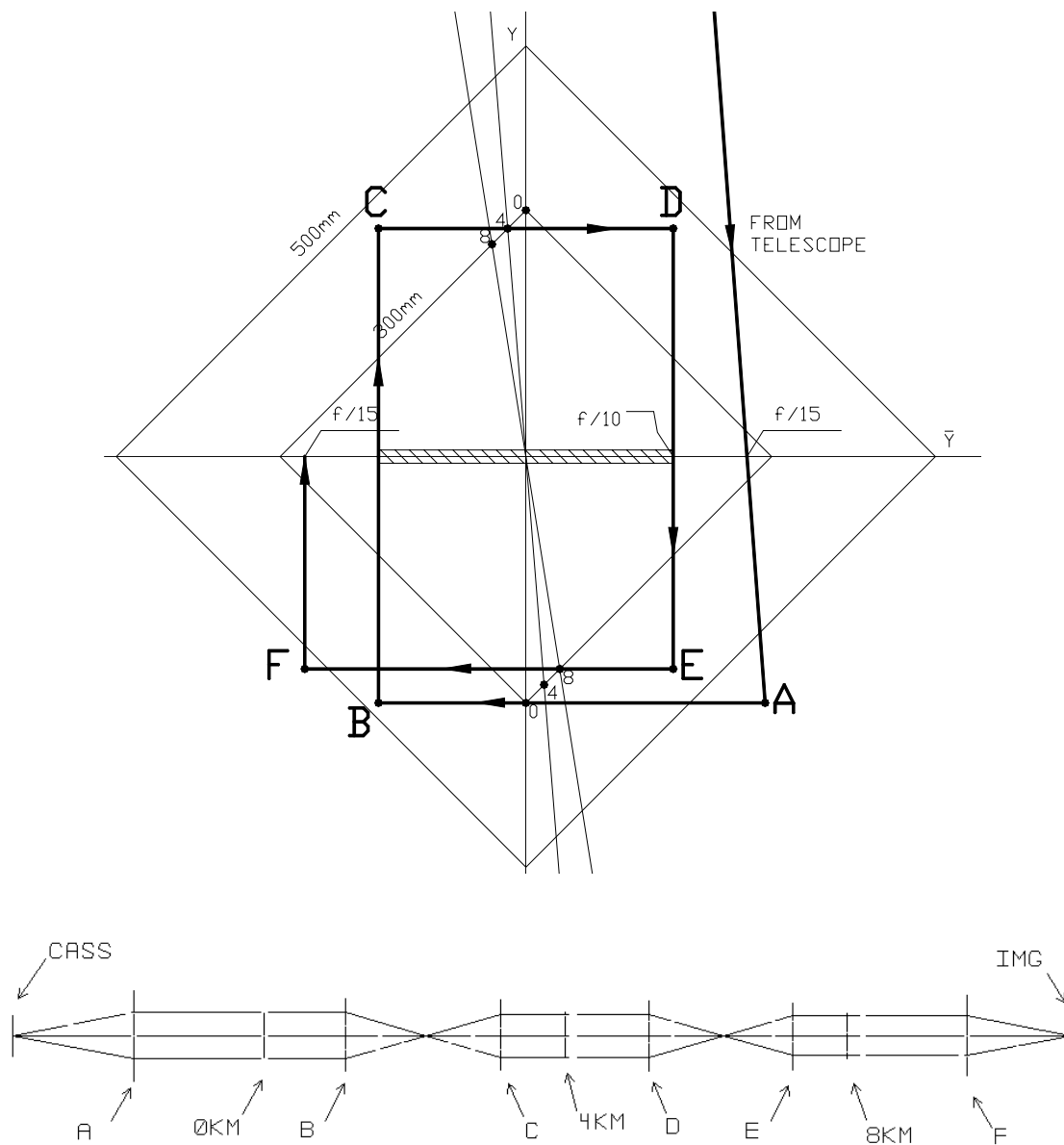
Figure 2- 17: y,$\overline{y}$ trace showing a relay that complies with the Hardy conjecture with slower f#'s. Additional constraints are 1) flat DM's; 2) nowhere is the beam faster than f/10; and 3) each optical space is either collimated or telecentric. Note that 6 optical surfaces besides the DM's are required! Graphically, see that each DM requires two other powered surfaces. The overall length is 39.2m.

Now we'll try to reduce the number of surfaces required. Let us see if using powered DM's while retaining the other constraints reduces the relay optic count.

Figure 2-18 shows the result. Clearly, the very fast f# required (f/1.4) violates our f# requirement. The telecentricity requirement (vertical $y,\bar{y}$ rays) heavily constrains the design.

In figure 2-19, we lift the telecentricity requirement (alternating spaces are collimated) and find that we can reduce the relay optic count from 6 to 4. A process similar to that in figure 2-16 is used, with the exception that in order to find "B", we construct a line that connects the left edge of the f/10 keep-out zone and the upper 4km DM point. The intersection of this line with the horizontal line passing through 0km locates "B". A similar process locates "C". The last relay optic, "D", is somewhat arbitrarily placed. "D" needs only to take the horizontal $y,\bar{y}$ ray after the 8km DM point and redirect it to the image plane. Another option for "D" would be to place such that the exit pupil of the telescope (or equivalently, entrance pupil of the AO relay) is re-created, as indicated by the dashed line in figure 2-6.
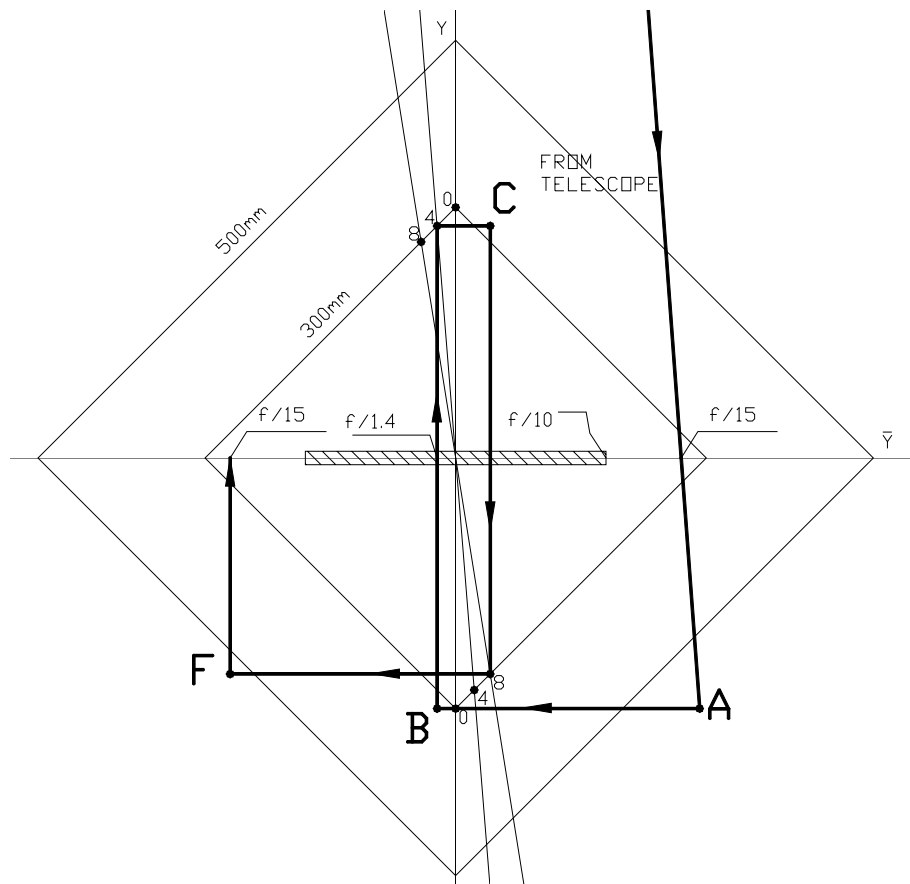
Figure 2- 18: Same as figure 2-17, but with powered DM's allowed. This solution violates the f# requirement, similar to figure 2-15. If feasible, this design would have saved 2 surfaces (7 total) relative to figure 2-17.
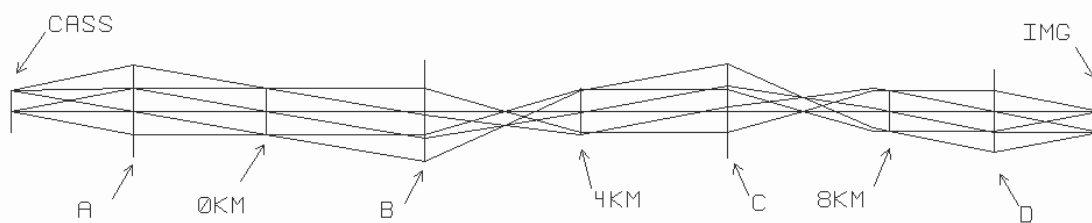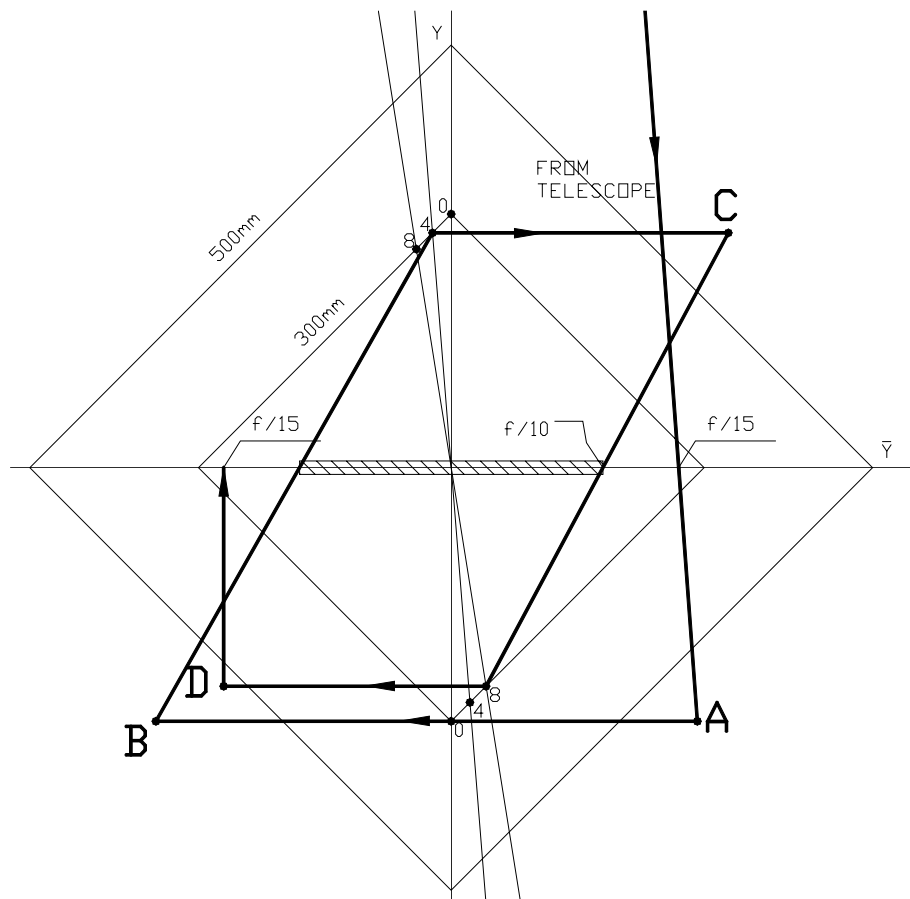
Figure 2- 19: Same as figure 2-17, but with telecentricity and flat DM requirements removed; the 4km and 8km DM's are powered. The result is one relay optic per DM, plus one to focus the light after the 8km DM. There is a cost in mirror diameters at B and C compared to figure 2-17, but little change in overall length (compare polygon origin-0km-B-C-4km in figure 2-17; similarly for origin-4km-D-E-8km).

Now, as figure 2-19 shows, the requirement of alternating collimated spaces keeps us from achieving the minimum of three optics in this design. Figure 2-20 shows a solution that result after lifting the constraint of alternating collimated spaces. This is similar to the design of figure 2-19, but with the y,$\bar{y}$ ray proceeding directly from the 8km DM to the image. In figure 2-21, we see the shortest possible solution given the f/10 speed constraints. The overall length is still almost 21m, significantly longer than the minimum non-Hardy layout of figure 2-9 (9.7m).

In studying figure 2-21, we see that if we were to add more DM's to the system, we would need another "half-circuit" around the origin of the y,$\bar{y}$ diagram, which would add one relay optic per DM, plus about 5.5m of path length.
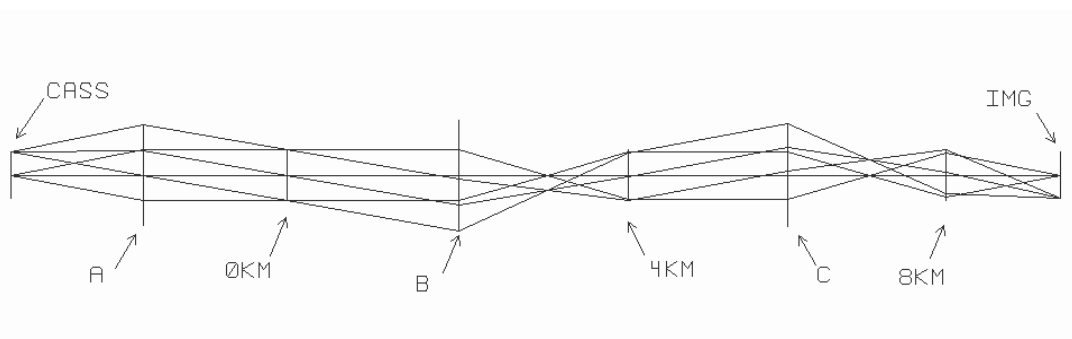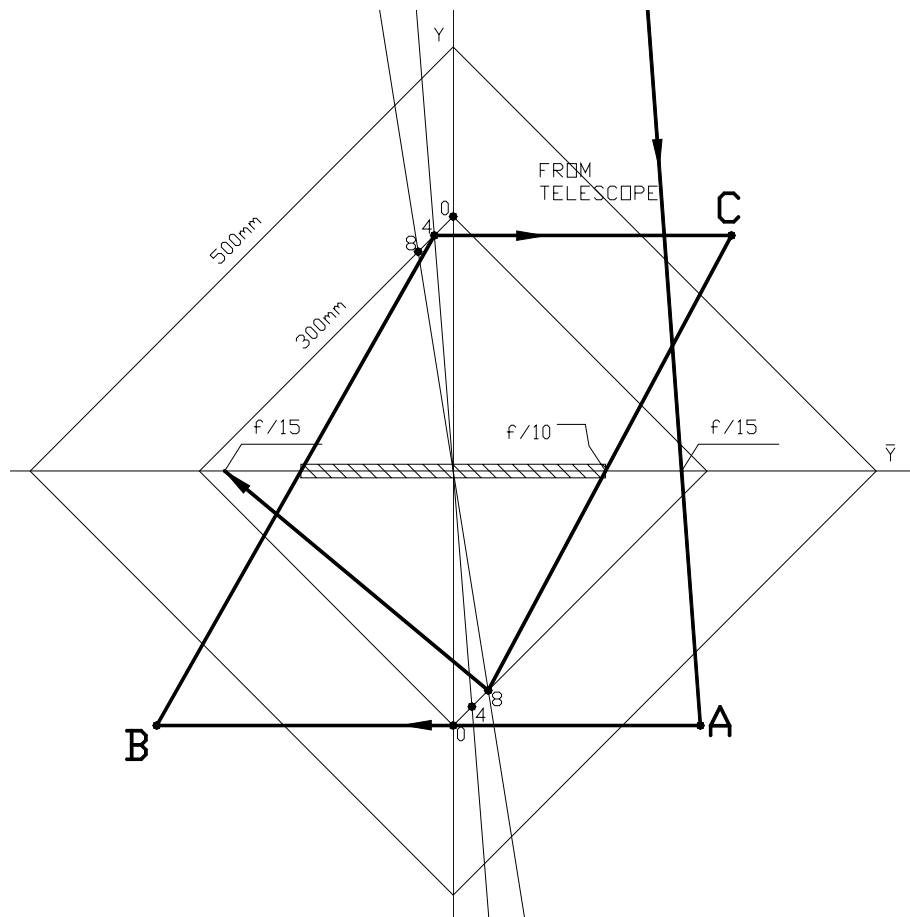
Figure 2- 20: A design using the minimum number of optics possible under the Hardy order-of-correction assumption. There are 3 optics in addition to the 3 DM's. Overall length is 35.6m.
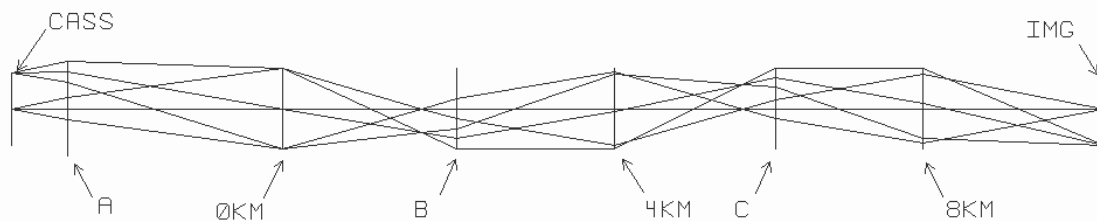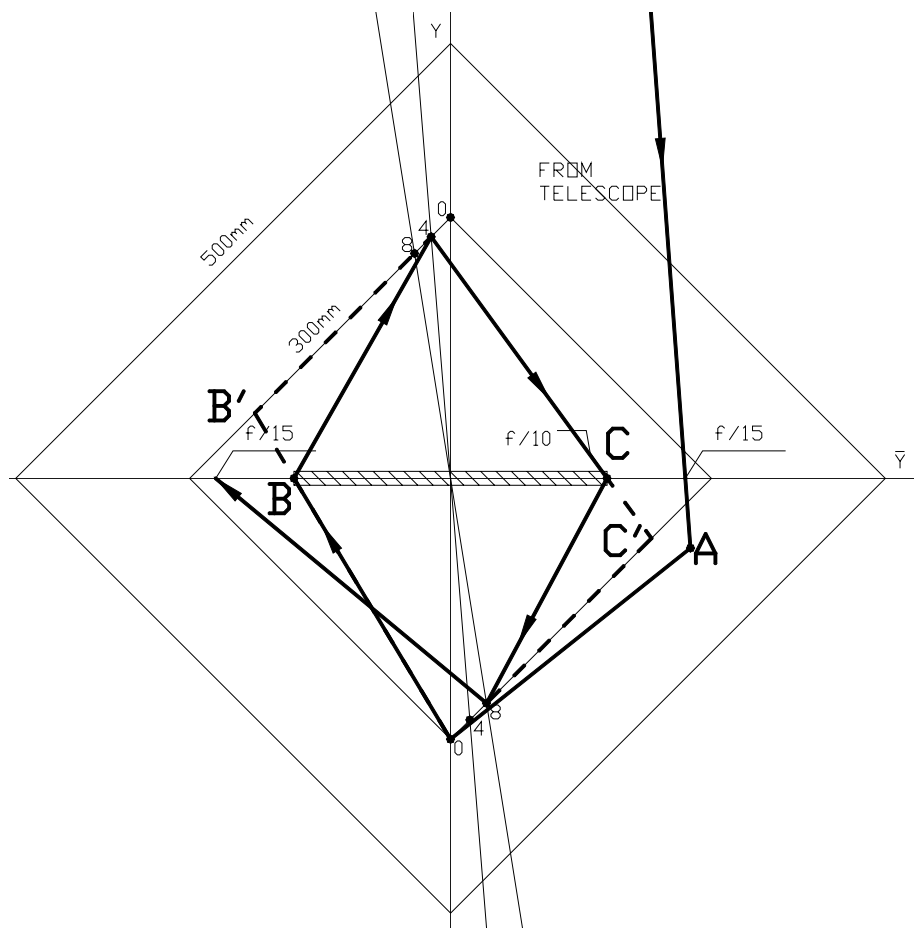
Figure 2- 21: Minimum length solution for minimum number of optics under Hardy constraints. The fact that B and C are at image planes may be problematic since any dust on these optics would appear in focus. This may be alleviated at a minor cost in length as shown by moving the relay optics away from the image plane, illustrated by the dashed line above. Some length could be reduced by moving "A" closer to the telescope focus; this position for "A" is used for clarity in the drawing. The overall lengths of these layouts are 20.8m and 23.4m.

In figures 2-22 and 2-23, we try to retain the minimum relay optic count (3), but using only flat DM's. We see that doing so is possible, but at the cost of space and relay optic diameters. Figure 2-23, which has eased the minimum f# constraint to f/7.5 is still 67.6m long!
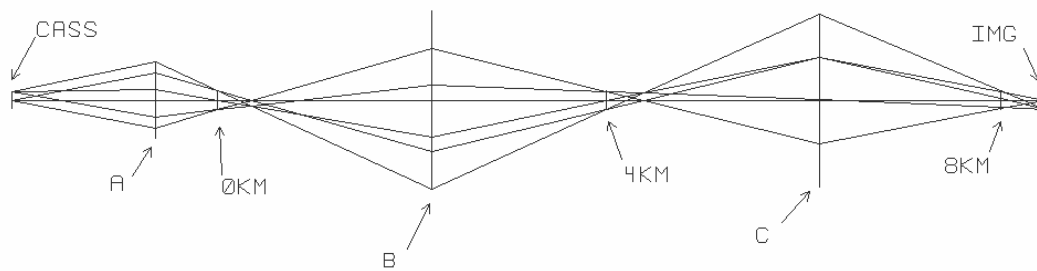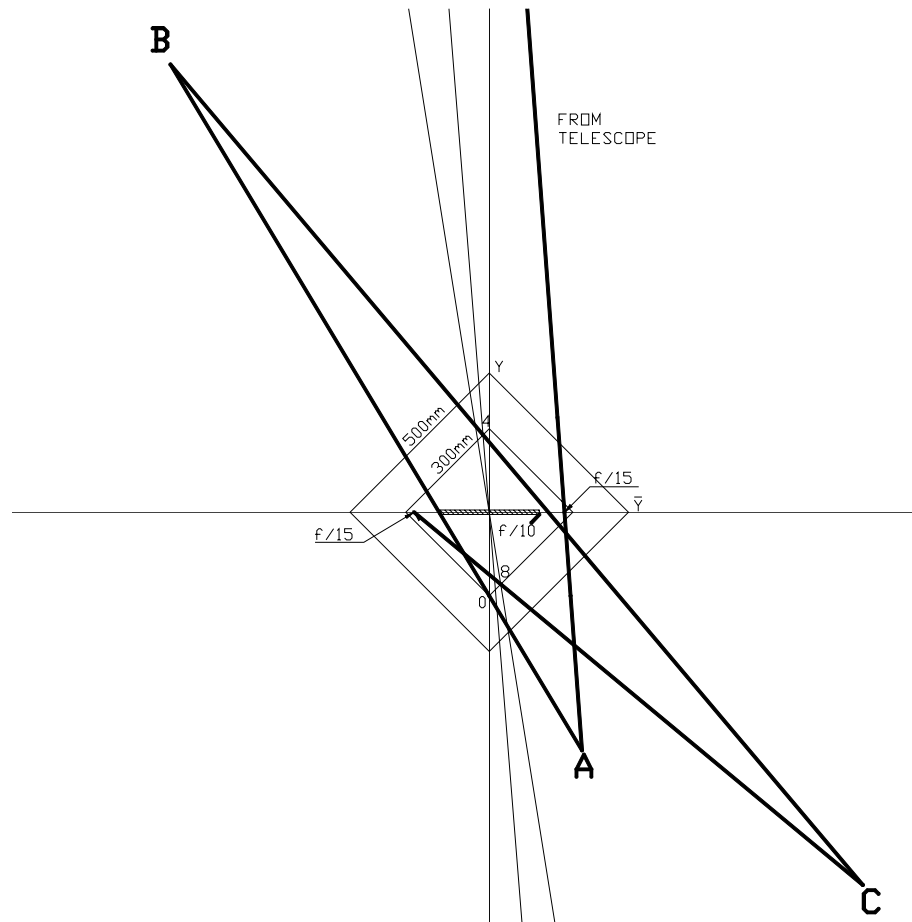
Figure 2- 22: A solution using flat DM's and the minimum number of optics (A, B, and C). The diameters of the optics are large (2.7m). The overall length is 92.3m.
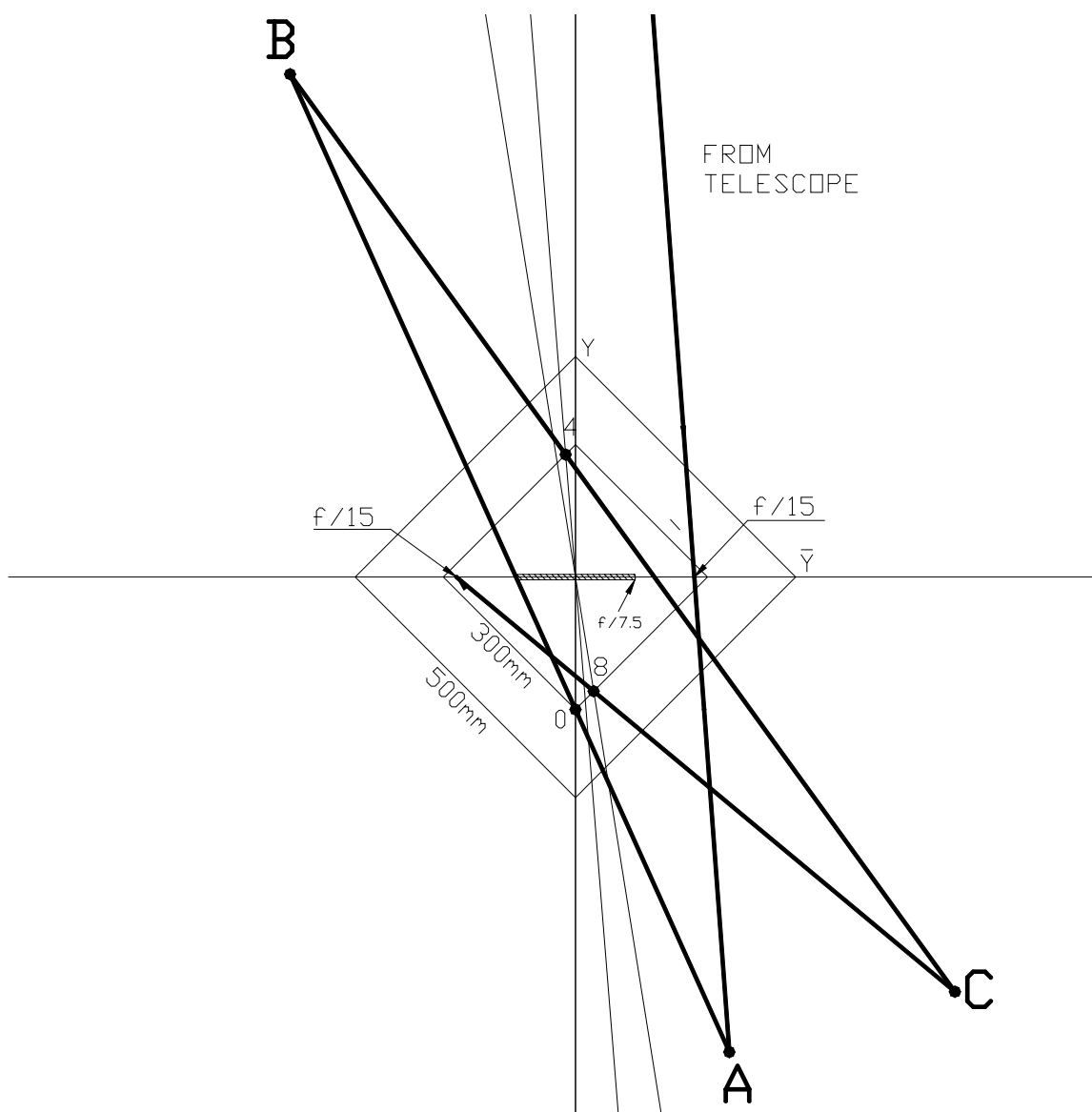
Figure 2- 23: Another "Hardy" solution with flat DM's and minimum number of optics. The relay optic diameters are quite a bit smaller (1.8m) than in figure 2-22; the minimum f# constraint has been changed from f/10 to f/7.5 to allow this. The overall length is still long—67.6m.

The following summarizes the last section which considered MCAO designs using the Hardy conjecture as a constraint.

Conclusions

In the case of the Hardy conjecture being applied, we conclude:

- It is not possible to put 2 DM's in the same optical space without very fast f#'s

- DM's must be separated by at least one relay optic.

- It is possible to have a flat DM solution with the minimum number of elements, but it is quite long with large relay optic diameters.

- Requiring only collimated or telecentric spaces results in 2 optics per DM.

- The minimum length solution for a 3 DM system is about 21m.

- Using powered DM's cuts the overall path length approximately in half (compare figures 2-17 and 2-19 versus figure 2-21)

- Adding more DM's to the system will cost at least one relay optic per DM added plus an additional 5.5m of path length.

Comparing the results of sections 2.8 and 2.9, we can reach some important additional conclusions:

- The AO relay design will be much simpler and much shorter if we can conclude that we need not obey the Hardy conjecture for order-of-correction.

- Using powered DM's can result in a significant savings in space but not in numbers of optics.

3. PYRAMID WAVEFRONT SENSORS

3.1. Introduction

Most current astronomical adaptive optics instruments use Shack-Hartmann (SH) WFS's (Shack 1971) to measure the wavefront aberrations from guide stars. In this scheme, illustrated in figure 3-1, the converging beam is collimated by a lens that images the pupil of the system (usually the telescope primary) onto a lenslet array. The lenslet array
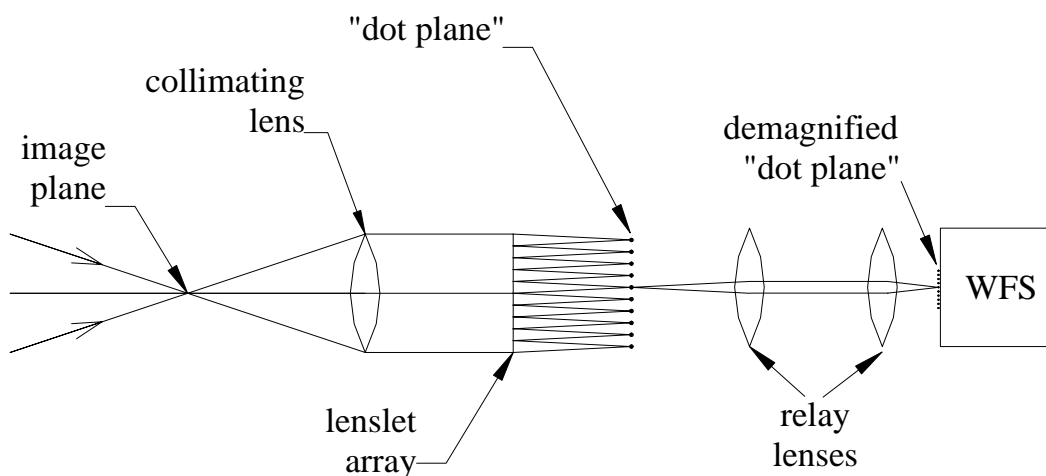
Figure 3- 1: Conventional Shack-Hartmann (SH) wavefront sensor. The collimating lens collimates the light from the guide star. The lenslet array is commonly placed at the pupil of the system, i.e., conjugate to the ground layer, but can be placed conjugate to other heights, as in an LOAO scheme. Some systems have additional relay optics between the focal plane of the lenslets and the CCD, as depicted here. For clarity, only one dot's rays are relayed to the CCD in this illustration.

divides the pupil into subapertures and produces images of the guide star at the focal points of the several lenslets. A CCD is then located (perhaps after some relay optics) at the image plane.

In some configurations, the image of the guide star is steered so that the image is centered at or near the junction of 4 (2x2) neighboring pixels (a "crosshair"), and the pixels are used as a quad-cell to find the centroid of the spot, thereby determining the local slope of the beam over that subaperture. In other configurations, each subaperture may have more than 2x2 pixels per subaperture. In this case, a centroid is computed via center-of-mass algorithm (Hardy), iterative region-of-interest centroiding (Hofer, Williams), or matched-filter (cross-correlation) algorithm (Poyneer, et al.).

pyramid — field lens

pupils with CCD
pixels demarking
subapertures

incoming
beam

CCD at
pupil plane
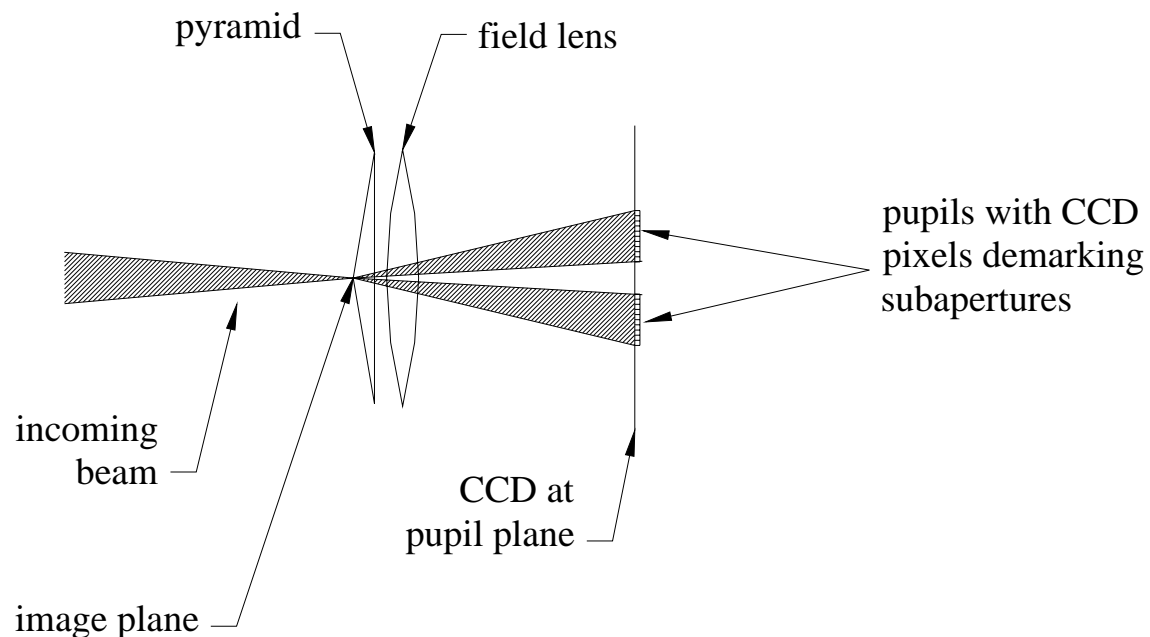
image plane

Figure 3- 2: A pyramid wavefront sensor. The pyramid (or lenslet array in the proposed approach) is located at the image plane and the spot is positioned over the apex of the pyramid, dividing up the field into 4 quadrants. A subsequent lens re-images the pupil (or other conjugate height) onto the CCD. The CCD pixels divide the pupil into subapertures.

In 1996, Ragazzoni invented the notion of pyramidal wavefront sensing which can be considered a variant of the Foucault knife-edge test. His idea was to place a 4-faceted pyramid at the image plane, with the apex of the pyramid located at the guide star image; see figure 3-2. In Ragazzoni's original formulation, the pyramid divides the beam into four field quadrants; a subsequent field lens images the pupil for each beam onto a CCD. Each subaperture is defined by a pixel on the CCD.

The PWFS can be thought of as a Foucault knife-edge test (Foucault) (figure 3-3), where two orthogonal knife-edges are implemented simultaneously. It has been noted that the

knife-edge test is excellent as a qualitative test, but is not as convenient for quantitative measurement (Malacara). This issue is eased somewhat by scanning the knife-edge in the focal plane. In a similar way, Ragazzoni's original formulation calls for oscillating the guide star image around the apex of the pyramid where the size of the oscillation depends on the size of the image. When the aberrations are small, small oscillations are enough to generate significant modulation in the subaperture intensities; when the aberrations are large, the oscillations are large so that adequate modulation is achieved.
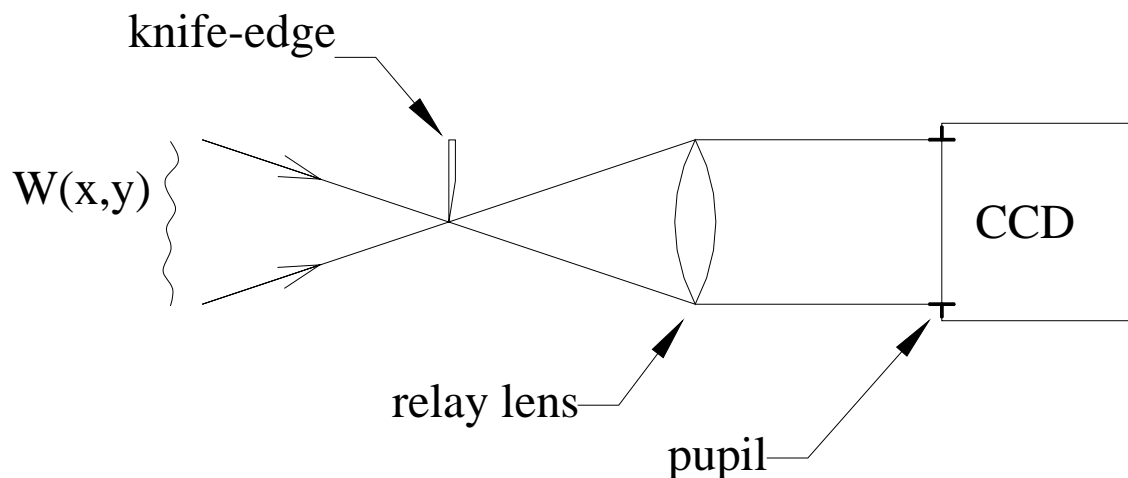
Figure 3- 3: Layout of a Foucault knife-edge test. The wavefront to be measured, W(x,y), converges onto a knife-edge in the focal plane. A relay lens images the pupil onto the CCD.

3.2. Comparison of PWFS's and SH WFS's

We now offer some insights into the comparison of SH WFS's and pyramid WFS's.

Upon examination, one observes that PWFS's are very similar to SH WFS's. Whereas a SH WFS (figure 3-1) first divides the pupil into subapertures (via a lenslet array placed at the pupil), then divides the field into quadrants (via the pixel boundaries on the WFS CCD), a PWFS performs these operation in reverse order: dividing the field into quadrants (via the pyramid, or other technique, as will be discussed), then dividing the subsequent pupil(s) into subapertures via the pixel boundaries of the CCD (figure 3-2). The pupils are created by a lens that follows the pyramid. Thus, in a geometrical optics sense, the information is exactly the same, just organized differently as shown in figure 3-
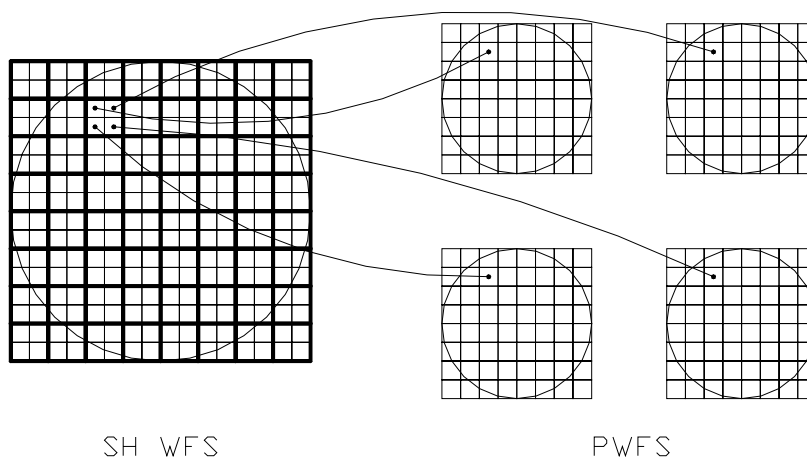
Figure 3- 4: Organization of SH wavefront data (left) versus pyramid wavefront data (right). The circle indicates the beam footprint on the WFS. The heavily-weighted squares on the left indicate the various subapertures (8x8 grid of subapertures). Each subaperture has 4 pixels (a quad cell). In a pyramid wavefront sensing scheme, each pixel represents a subaperture; the 4 images of the pupil correspond to the quadrants of the quad cell.

However, Ragazzoni's notion of an oscillating element is not peculiar to PWFS's. In the same way that Esposito suggested using a tip/tilt mirror in order to oscillate the guide star image around the apex of the pyramid, one could use a tip/tilt mirror to steer the guide star image around in a SH WFS. In a SH WFS, the image would be steered around the crosshair of pixels rather than about the vertex of a pyramid. In this characteristic, then, SH WFS's and PWFS's do not differ.

3.3.   Advantages of PWFS's

There are two very important differences between PWFS's and SH WFS's. First, as Ragazzoni notes, the sensitivity of the WFS depends linearly on the size of the spot incident on the image-plane quad-cell (e.g., the pixel crosshairs in a SH WFS). In a SH WFS, the size of this spot is limited by the diffraction limit of the subaperture, i.e., λ/d, where λ is the wavefront sensing wavelength, d is the subaperture size; in practice, $d \approx r_0$, where r_0 is Fried's atmospheric correlation length at the science wavelength (Fried 1966). In a PWFS, the size of the spot is ultimately limited by λ/D, where D is the diameter of the entire telescope aperture (Ragazzoni 1996). The size of the spot will decrease from $\lambda/r_0 \approx \lambda/d$ as the wavefront correction improves, approaching λ/D. For an AO system with 100 subapertures across (such as that proposed for a 30m ELT), this is a 100x smaller spot, with a corresponding improvement in centroid measurement. This λ/D limit may be realized in a highly-corrected, high-Strehl ratio regime, such as that required by Extreme AO. Extreme AO endeavors to achieve very high Strehl ratios so that planets

around other stars can be directly imaged—this requires being able to detect objects that are $\sim 10^7 - 10^8$ dimmer than the parent star (Angel 1994, Macintosh).

The remaining discussion in this section represents a newly considered advantage of the PWFS. The second important difference between SH WFS's and PWFS's is in the mechanism used to detect the spot's position. In a SH WFS, each subaperture's spot is divided by the pixel boundaries. A recent report (van Dam) indicates that the pixel boundaries are in fact, quite indistinct (figure 3-5) – a photon landing just to one side of a pixel boundary would have equal chances of arriving in either of the two bordering pixels; this phenomenon is called "charge diffusion". This is the worst case: the information (signal) of these photons is lost, yet all of the noise (Poisson statistics) is retained—it would be better if the photon were lost entirely. van Dam has quantified this effect for the Lincoln Labs 64x64 CCD used for the Keck and Lick WFS's: the charge diffusion effect is equivalent to convolving the guide star image with a Gaussian with a 1/2 pixel FWHM, which is comparable to the size of the guide star itself.
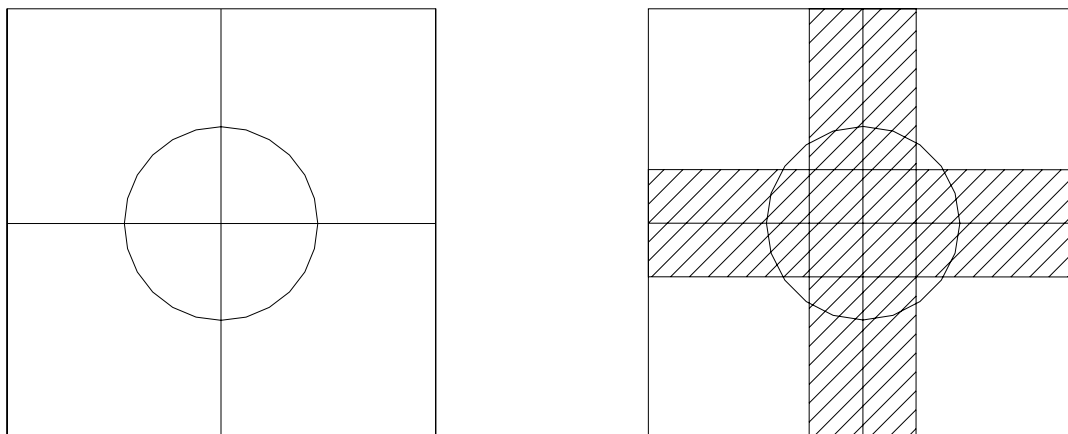
Figure 3- 5: A spot incident on the junction of 4 pixels. The figure at left shows the idealized representation of the pixel boundaries. However, van Dam has shown that the pixel boundaries are actually quite indistinct, as shown at right, so that a large percentage of the light strikes ambiguous regions (shaded areas).

In contrast, a PWFS is divided at the field quadrant boundaries (the edges of the pyramid in the Ragazzoni scheme), then divided into subapertures by the pixel boundaries. Thus, assuming sharp pyramid edges, the field is divided sharply, while the subapertures are divided in a blurry manner. Intuitively, we can expect a performance gain in the PWFS scheme: the sharp division in the field produces high sensitivity measurements of the centroid that we need, while the blurriness at the subaperture boundary correctly reflects the fact that we are somewhat indifferent as to which subaperture a boundary photon should be attributed.

The effect of reducing the effective spot size on a quad cell (in this case, by increasing the sharpness of our field-division knife-edge) is well understood. What is the effect,

though, of blurring the subapertures? Intuitively, one would expect that the blurring of

the subapertures will make little difference in measuring low spatial frequency

components in the pupil plane, but would make a greater difference in measuring high

spatial frequency components.

We can understand the difference between the sharp and blurry demarcation of

subapertures by setting aside the actual measurement of the wavefront and considering

only the action of the subaperture boundaries. The approach here will be to view the

pixelation of the pupil (subapertures) in the same manner that one would view pixelation

of an image.

The following discussion uses the function and Fourier transform definitions found in

Gaskill, and is cast in one dimension for clarity; the analysis is easily extended to two

dimensions.   In analyzing spatial frequency performance (i.e., modulation transfer

function (MTF)) of imaging optical systems, one considers various MTF components,

cascading (multiplying) them together to form a system MTF (Lloyd). Examples of

component MTF's are diffraction MTF, electronics MTF, and detector pixel MTF. The

pixel MTF is the frequency response due to the finite size of the pixel in the image plane,

i.e. due to the integration of the image over the extent of the pixel. For example, a typical

pixel function in the spatial domain would be $(1/a)\,\text{rect}(x/a)$ where "a" is the pixel size;

the scaling factor 1/a is chosen so that the area under the rect function is unity. "a" can

be expressed in terms of object space (e.g., arc-sec) or image space (e.g., mm) quantities. The pixel MTF is just the Fourier transform of the spatial pixel function, i.e., $\mathrm{MTF}(\xi) = \mathrm{sinc}(a\xi)$, where ξ is a spatial frequency coordinate. As one would expect, the MTF is unity at zero spatial frequency and rolls off at higher frequencies, becoming zero at $\xi = 1/a$.

Analogous to the pixel in the image plane is the subaperture in the pupil plane—both image pixel and subaperture describe an extent over which the image or pupil, respectively, is integrated. Thus, in the same way that the pixel MTF describes the effect of finite pixel size in the image plane, a subaperture MTF describes the effect of finite subaperture size in the pupil plane. In a SH WFS, this subaperture function is similar to the pixel function mentioned above: $(1/d)\mathrm{rect}(x/d)$ where d is the subaperture size. For the PWFS, the subaperture boundaries are defined by the detector pixel response function, i.e., by a function that has rolled-off edges compared to a rect function. A reasonable approximation for the pixel response function is shown in figure 3-6, and is described by the function

$$f_{\mathrm{diffusion}}(x) = \frac{1}{ab}\left[\mathrm{rect}(x/a) * \mathrm{rect}(x/b)\right], \tag{1}$$

where "a" is the nominal pixel size, and "b" represents the pixel blurring size. The component rect functions that are convolved to make the pixel response function are illustrated in figure 3-7.
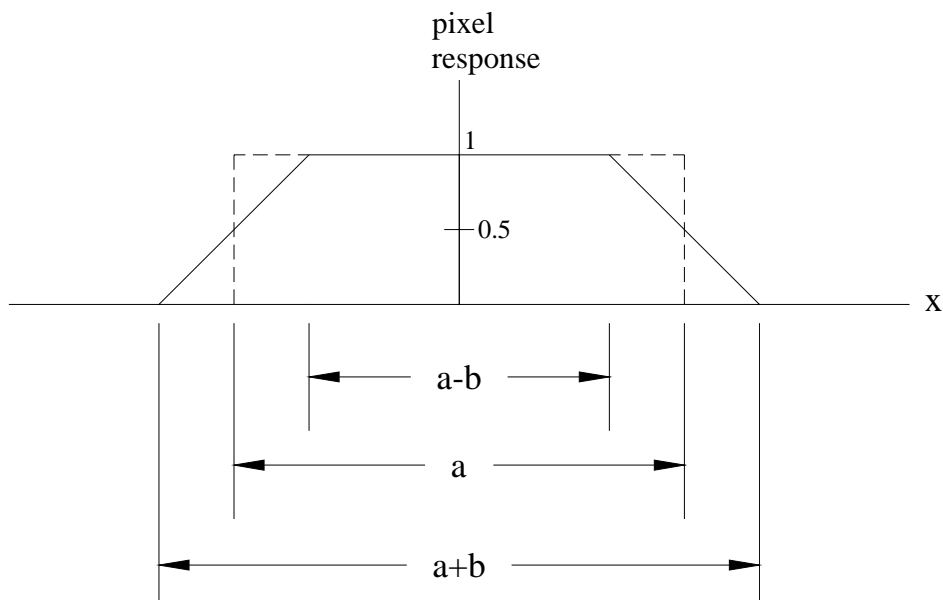
Figure 3- 6: Pixel response function. The nominal width of the pixel is "a".   The dashed line indicates the ideal pixel function; the solid line indicates the model pixel function due to charge diffusion, which extends the pixel response to width "a+b".   A linear interpolation is assumed.
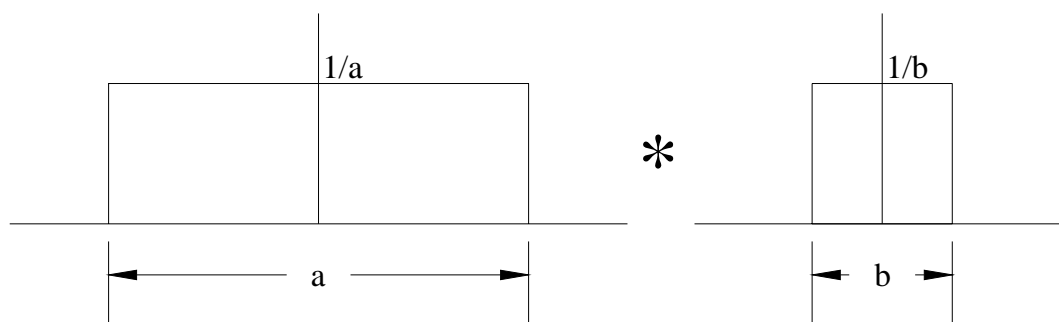
Figure 3- 7: Rect functions that are convolved to yield pixel response function in figure 3-6. "a" is the nominal pixel size and "b" is the pixel blur size.

Taking the Fourier transform of $f_{\text{diffusion}}(x)$,

$$\text{MTF}_{\text{diffusion}}\left(\xi\right) = \mathcal{F}\left(\text{f}_{\text{diffusion}}(x)\right)$$
$$= \mathcal{F}\left(\frac{1}{ab}\text{rect}(x/a)*\text{rect}(x/b)\right) \quad\quad (2)$$
$$= \text{sinc}(a\xi)\,\text{sinc}(b\xi)$$

This MTF is compared to the MTF of the ideal pixel case in figure 3-8.

In the spatial frequency domain of the pupil, the subaperture function imposes a low-pass filter on the measured wavefront. At low frequencies, there is little difference between the ideal and the charge-diffused MTF's, so the blurring of the subaperture has little effect. At large frequencies, the high-frequency components are attenuated. Although there is some loss in sensitivity in the high-frequency components, this can be an advantage because it suppresses the spatial frequency components that are above the Nyquist frequency which would otherwise be aliased.
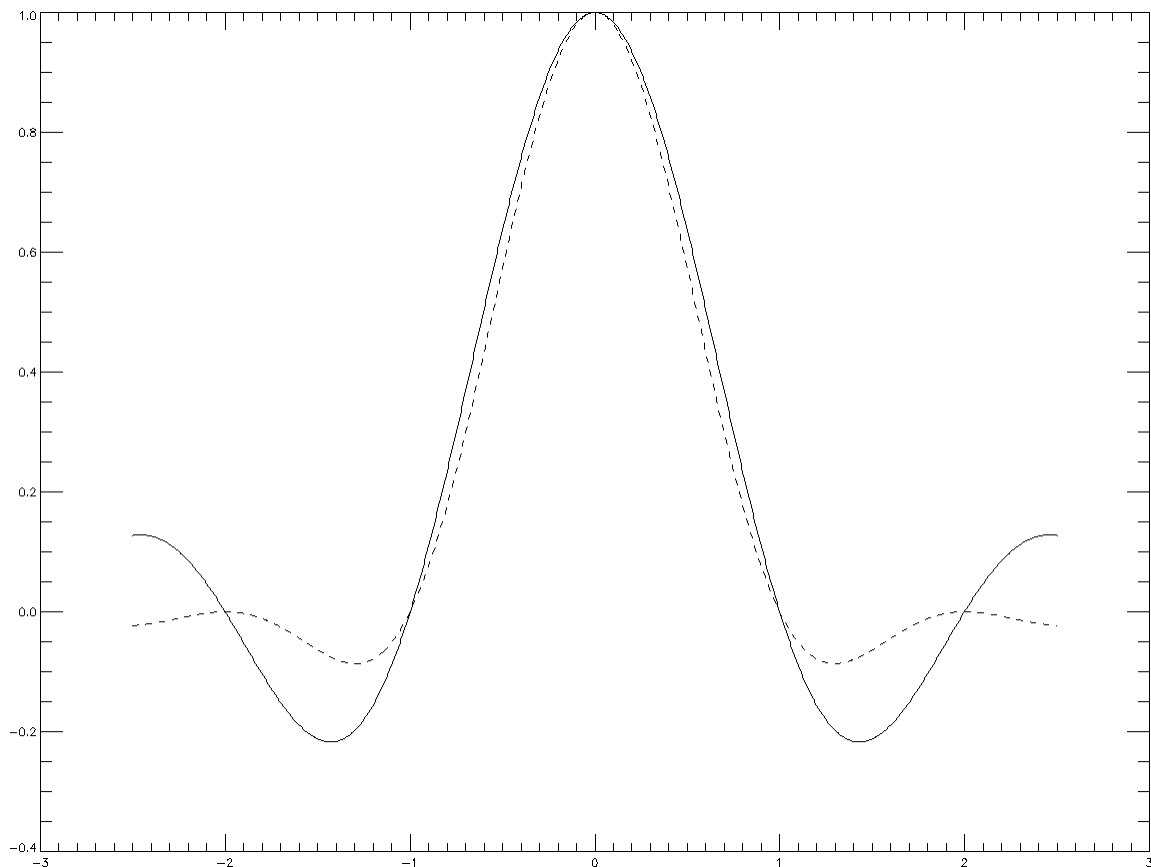
Figure 3- 8: MTF of ideal pixel (solid line) and of charge-diffused pixel (dashed). The horizontal axis is spatial frequency in units of 1/2a, where a is the nominal pixel size. Frequencies beyond 1/2a are aliased and so the suppression of these frequencies by the charge-diffused pixel is desirable. The response at spatial frequencies lower than 1/a is only slightly different between the two pixel functions.

3.4. Extension of Pyramid WFS to larger arrays

As mentioned earlier, in some configurations the PWFS uses a circular scan of the image

to increase the ability of the WFS to measure larger slopes. In SH WFS's, this can be

achieved by using a multicell (i.e., greater than 2x2 pixels per subaperture) and

calculating the centroid via the aforementioned center-of-mass algorithm, iterative region-of-interest centroiding algorithm, or matched-filter (cross-correlation) algorithm.

Ragazzoni's method uses a four-faceted prism to divide the field into quadrants. As previously stated, this is analogous to using a 2x2 matrix of pixels (a quad cell) in a SH WFS.

Using a novel technique to be described next, the notion of a larger number of field pixels (larger than 2x2) can be implemented in a PWFS.

Implementing a multicell wavefront measurement can be done in PWFS's with an array of pyramids (or an array of lenslets, to be seen later in this chapter); see figure 3-9. The length of the facet of the pyramid sets the plate scale, e.g., pyramids with 2 mm wide facets at an image plane with a plate scale of 0.5 arcsec/mm produces a wavefront sensor with a 1.0 arcsec/facet plate scale. Each pyramid is then followed by a lenslet to produce the set of 4 pupils corresponding to that pyramid's 2x2 pixel field. Figure 3-9 shows the resulting system and Figure 3-10 compares the organization of the data between SH and PWFS forms of the 4x4 multi-cell arrays.

Thus with the availability of this new technique, both PWFS's and SH WFS's can use quad-cell or multi-cell techniques, and so we need not choose SH WFS's or PWFS's because of the centroiding techniques that are available.

The size of the facets is an important design parameter, the selection of which depends on the desired number of subapertures across the pupil. Section 3.8 provides some guidance on the facet size.
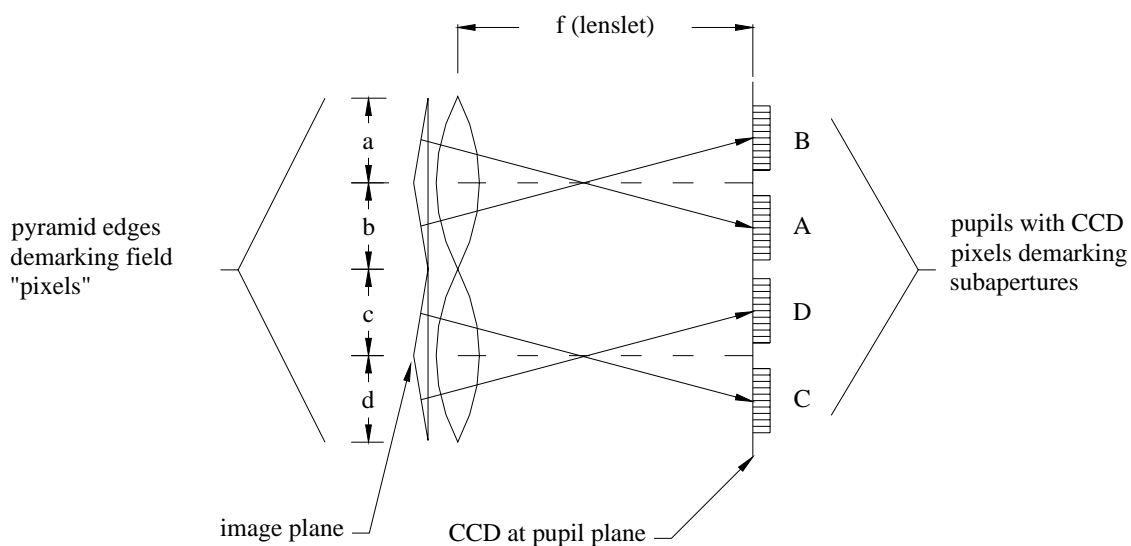
Figure 3- 9: Side view of PWFS in a 4x4 configuration. Each facet of the pyramid is a pixel (a,b,c,d) in the image plane. Light from a pixel in the image plane is mapped onto a corresponding pupil (A, B, C, D). The incident spot spans multiple pyramids at the image plane.
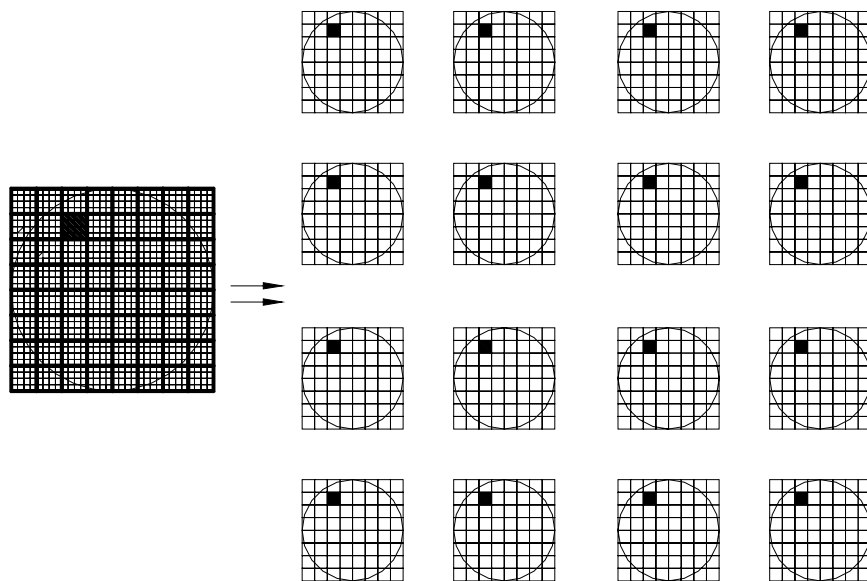
Figure 3- 10: Organization of data for SH WFS (left) and PWFS (right). The WFS has 8x8 subapertures, with 4x4 field pixels. The pixels for one subaperture are shaded.

It may be noted that the multicell PWFS bears resemblance to the wire test in optical metrology. The wire test (Malacara), which is a relative of the Foucault knife-edge test, uses a wire placed in the image plane (instead of a knife-edge) and a screen or detector at the pupil plane. Looking at the transmission cross-section of the focal-plane mask (i.e., wire or a single facet of a multicell pyramid), the similarities become more obvious (figure 3-10); the pyramid transmission curve is the complement of that of the wire. Of course, analogous to the comparison of the pyramid and the knife-edge, the pyramid facet yields a simultaneous measurement of the wire in two orientations at the image plane.
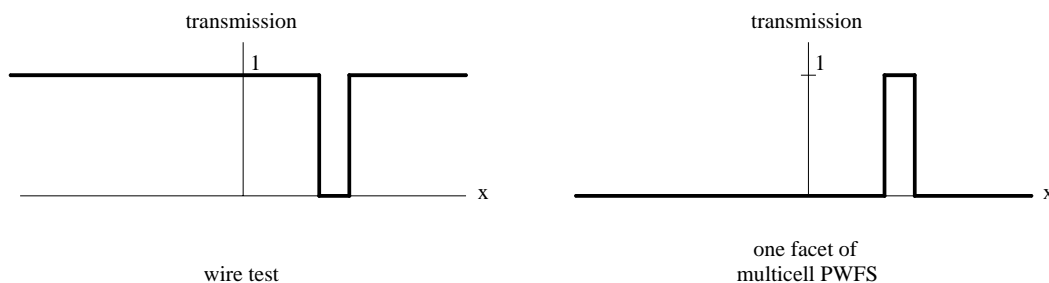
Figure 3- 11: One-dimensional transmission profile in the focal plane: wire test (left) and one facet in a multicell PWFS (right).

3.5. Lenslet-based PWFS's: a novel approach to constructing a PWFS

While the concept of PWFS, as proposed by Ragazzoni, shows promise, the implementations have been awkward, costly, and difficult, involving precision manufacture of glass facets and assembly of the facets into small (~12.7 mm) pyramids

(Ghigo 2001). The specification on pyramid edges quality can be eased somewhat by slowing down the light converging onto the pyramid to f/45; the pyramid edges need only to be good to a fraction (~10% or 5µ) of the spot size.

In order to minimize the number of pixels needed on the CCD, the various pupils are positioned close together. This requires that the apex angle is very shallow: $\alpha \approx$(NA of incoming beam)/(n-1)=(1/90)/(n-1)=1.3°; see figure 3-12. but one would need to contend with an even larger focal plane.
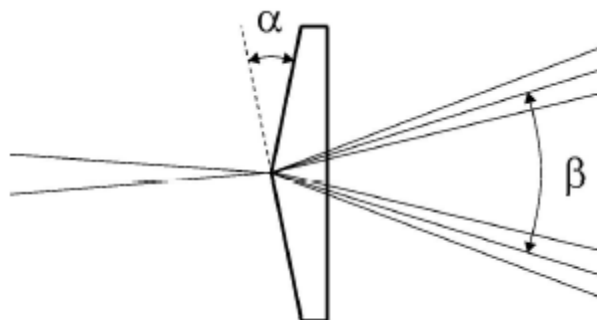
Figure 3- 12: The pyramid for a PWFS. The apex angle of the pyramid is α, and the beams leading to the pupils are separated by β; angles are exaggerated. (from Diolaiti)

Ragazzoni has found that it is very difficult to maintain pristine edges on the pyramid as the succeeding facets of the pyramid are polished—it is just too easy to mar or knock off the tip or edges of the pyramid; a lithographic technique has not yet solved the problem (Ghigo 2003). A design with two opposing pyramids (Diolaiti) produces more

manageable apex angles (~a few degrees; see figure 3-13), but it is still a significant

development effort.
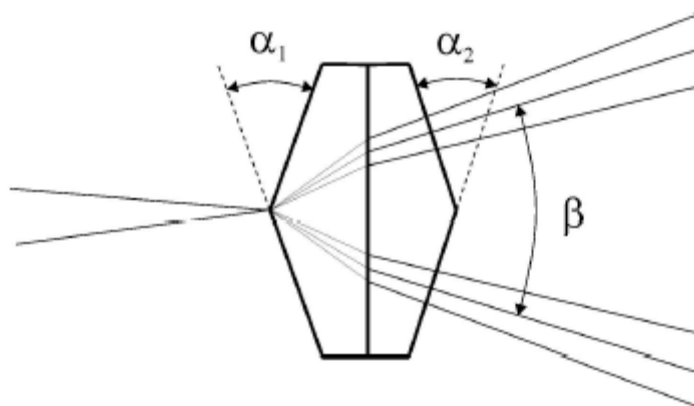
Figure 3- 13: Pyramid for PWFS made with two opposing pyramids. Using opposing pyramids allows the apex angles (α_1, α_2) to be larger, which are easier to fabricate. Note that only the left pyramid requires good quality facet edges since the beam is far from the facet edges in the right pyramid. β is the angle by which the pupils are separated. (from Diolaiti)

A novel approach to making pyramid wavefront sensors is suggested here. The key

insight is that the Ragazzoni method of a pyramid followed by a field lens is equivalent to

a lenslet array, where the field lens and lenslet have the same focal length. In an

elementary manner, this equivalency is shown in figure 3-14.
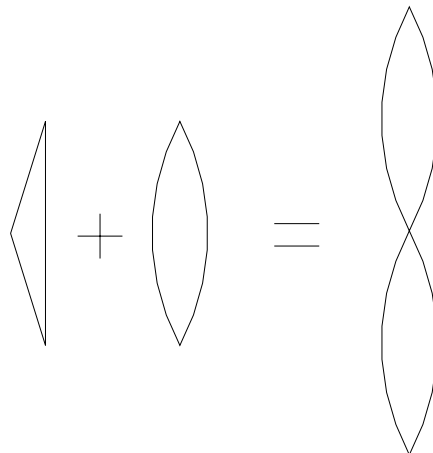
Figure 3- 14: Simple view of equivalency of a pyramid + field lens and a lenslet array. In this relation, the pyramid is actually the "negative" of the pyramid shown (i.e., a concave pyramid), but is shown in this manner for ease of illustration.

The layout is shown in figure 3-15. The converging beam is incident on the intersection of 4 lenslets; the pupil is assumed to be at infinity (telecentric) for convenience, but this is not required. The quad-cell formed by the lenslets splits the light into four beams and each lenslet forms a pupil at which the WFS CCD is located; in this side view, only two of the four beams are shown.

This approach has the benefit of using lenslet arrays, which are inexpensive, commercially available in a variety of configurations, easily customizable, and robust. Lenslet arrays are commonly formed by a molding process (compression molding on acrylic, or epoxy replication onto glass) as well as microlithographic techniques used by integrated circuit fabrication.
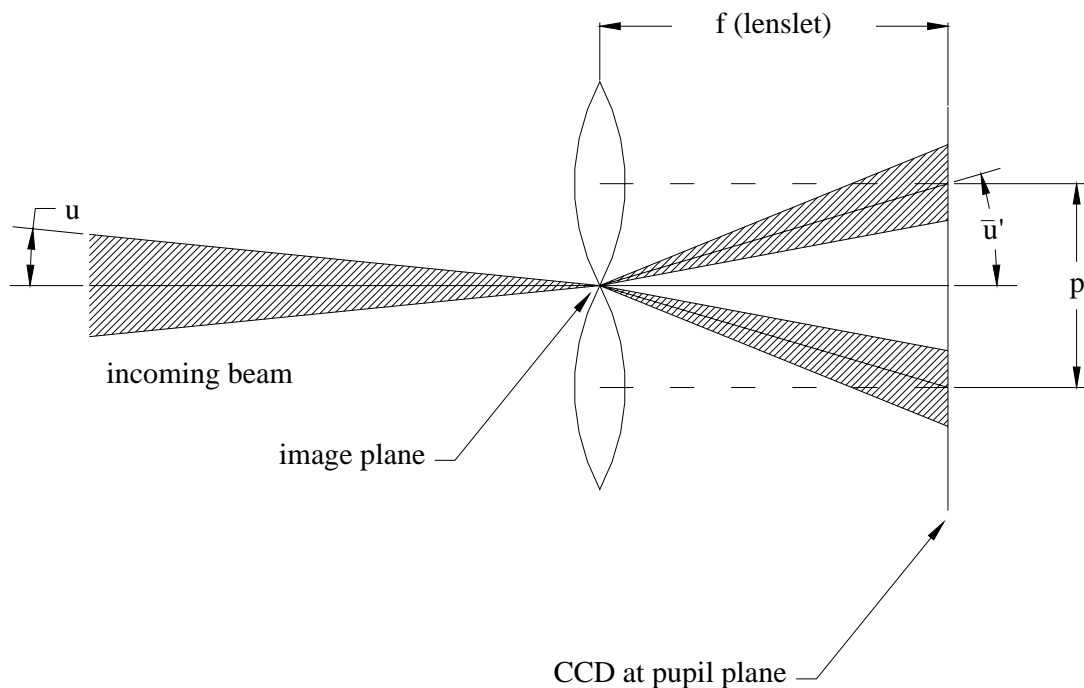
Figure 3- 15: A lenslet-based pyramid wavefront sensor, with the incoming beam to be measured centered on crosshairs, i.e., at the junction of the lenslets. The lenslet edges divide the field into quadrants and so separate the beam into 4 parts corresponding to the field quadrants; 2 are shown here. Each lenslet forms a corresponding pupil at the CCD. The pixels on the CCD divide the pupil into subapertures. u (= ½(f#)$_{beam}$) is the numerical aperture of the incoming beam (n=1 assumed), $\bar{u}$' indicates the angle through which each quadrant beam is steered, p is the pitch between lenslets, and f is the focal length of the lenslet. The beam is assumed telecentric in the space before the lenslet array. The biconvex shape of the lenslet is for illustration only.

The two key design parameters are, of course, the lenslet focal length and pitch.

Referring to figure 3-15, we do not want to have the pupils overlapped—this requires that

the lenslet steers the beam by more than u, i.e., $\bar{u}$'>u. Since u=1/2(f#)$_{beam}$ and

$\bar{u}$'=1/2(f#)$_{lenslet}$, we require

$$\left(f\#\right)_{lenslet} > \left(f\#\right)_{beam} \tag{3}$$

for pupil separation. We also need to have the pupil be the correct diameter, as set by our CCD characteristics. The necessary pupil size on the CCD is

Pupil diameter = (# subapertures across diameter) * (CCD pixel size)

In order to achieve this pupil diameter, we need

$$\frac{f_{lenslet}}{f\#_{beam}} = \text{pupil diameter}$$

So,

$f_{lenslet} = (f\#)_{beam}$ * (# subapertures across diameter) * (CCD pixel size)

Of course, relay optics may be imposed between the lenslet array and the CCD to adapt the size of one to the other, similar to what is done for current SH WFS's.

A sample design will be presented in section 3.7.

Now, the lenslets are clearly being used here in a manner that was not originally intended. SH WFS's placed relatively mild requirements on the lenslet arrays: lenslet edges only defined subapertures rather than performing the crucial knife-edge test. Thus,

current commercially available lenslet arrays may not be adequate for the task—we will need to test them to find out. This testing is planned for early 2004.

The question of tolerancing a PWFS has been examined in Diolaiti. Since a lenslet is just the combination of the pyramid facet and subsequent field lens, the tolerancing approach taken in Diolaiti applies here as well.

3.6. Lenslet-based multi-cell PWFS's

Earlier, we saw how Ragazzoni's pyramid WFS can be extended so as to use configurations other than quad-cell centroiding, such as 4x4 pixels. A new technique, to be described next, shows how the lenslet approach to PWFS's can be extended easily to these configurations.

Figure 3-16 shows the necessary configuration. Each lenslet is analogous to a pixel in a multi-cell SH array. As before, each lenslet creates a pupil which is divided by subapertures by the CCD pixels. This configuration is easy to achieve since lenslets are usually available in large arrays; in fact, one could measure wavefronts from several guide stars on one CCD with this approach.

The organization of the resulting data is similar to that shown in figure 3-10. The selection of lenslet size is similar to that in the multi-cell PWFS; see section 3.8.
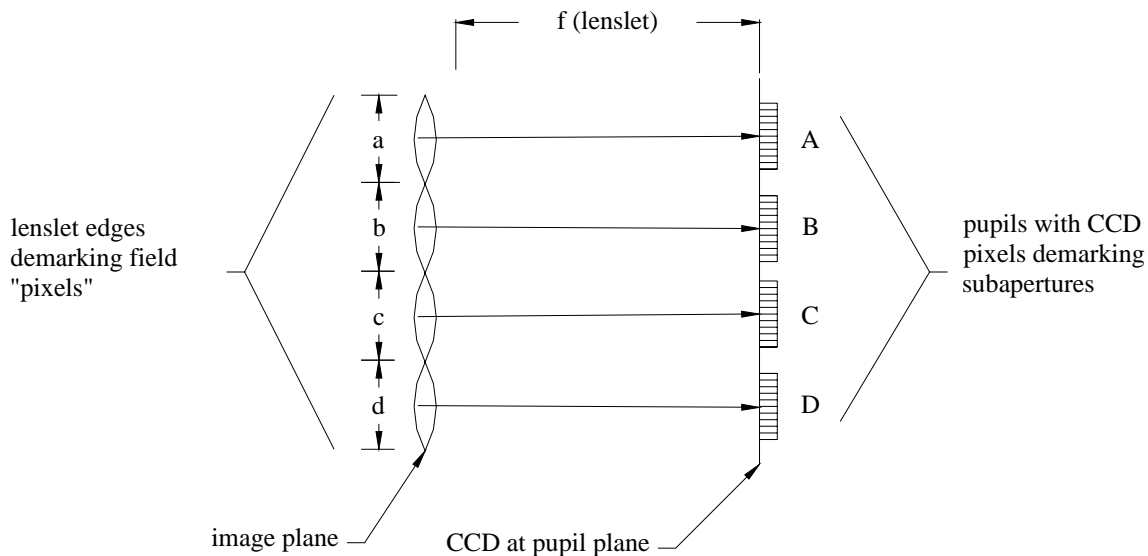
Figure 3- 16: Lenslet-based PWFS in a 4x4 configuration. The image of the guide star at the lenslet plane would span multiple lenslets (a,b,c,d). Each lenslet creates a pupil (A,B,C,D), which is divided into subapertures via CCD pixels. The optical space before the lenslets is telecentric in this illustration, although it is not required.

3.7.   Adaptation of SH WFS's to lenslet-based PWFS's

The following is a novel technique to easily adapt existing S-H WFS's to lenslet-based PWFS's. The adaptation may not be perfect, initially, but it can be tried out without major anguish in order to test whether a more permanent change is worthwhile.

For simplicity, consider a SH WFS which is telecentric in the optical space before the collimating lens of the SH WFS; the telecentricity assumption is nearly true in existing AO systems. A lens collimates the beam, and a lenslet is placed at the pupil plane; see figure 3-17. The lenslets create images (dots) located one focal length downstream of the

lenslet array. Often, there are relay optics which image the dot plane onto the CCD; the relay optics' function is to match the lenslet pitch to the pixel pitch, since in general, the dimensions of off-the-shelf lenslet arrays and the dimensions of off-the-shelf CCD's do not match.
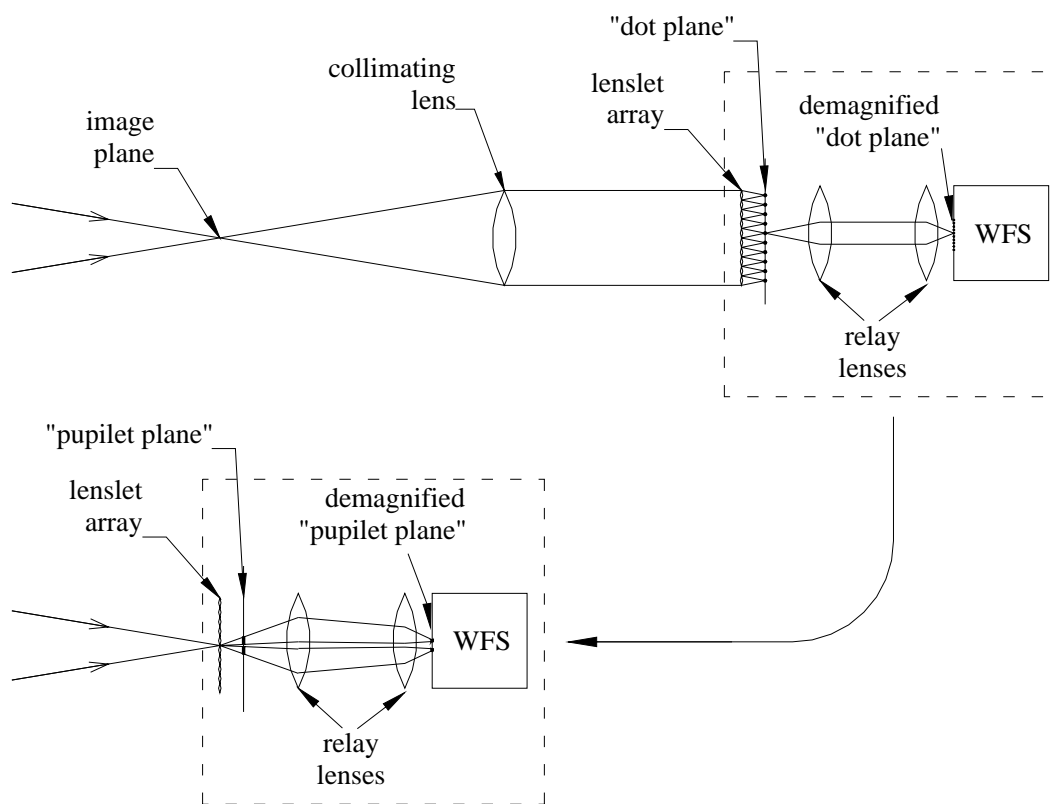
Figure 3- 17: Conversion of a SH WFS to a PWFS. Top figure: Converging light from the left comes to a focus and is then collimated by a collimating lens. The collimating lens creates a pupil downstream, where the lenslet array is placed. The lenslets produce a series of images or dots at the focal plane of the lenslets or "dot plane". Subsequent relay optics scale the dots as appropriate for the WFS CCD. For clarity, light from only one dot is shown after the dot plane. Bottom figure: to convert SH WFS to PWFS, remove collimating lens and translate lenslet array, relay optics, and WFS CCD upstream until the lenslet array is at the focus of the incoming beam. The lenslets now produce "pupilets" at the lenslet focal plane, i.e., where the dots were in the top figure. Thus, the relay optics will relay the pupilets to the WFS CCD.

To convert a SH WFS with a telecentric pupil into a PWFS, simply remove the collimating lens in front of the lenslet array and move the remainder of the WFS (lenslet array, subsequent optics, and CCD) upstream so that the lenslet array is at the focus. Since the pupil is at infinity relative to the lenslets, the lenslets each will now image the pupil one focal length away, i.e., at the same plane that their dots were in the SH scheme. Thus, any subsequent relay optics will produce pupils on the CCD.

Now, assuming the same number of subapertures and plate scale (to be discussed shortly), the data is the same as in the SH WFS, just sliced-and-diced differently, as previously shown in figure 3-10. A simple routine can be inserted into the control software to re-map the data from the PWFS format to the SH WFS format, and the rest of the software can remain the same. It is possible that the control matrix will need to be modified, since one may expect that the sensitivity of the WFS to wavefront errors will be different.

In general, this procedure will not yield the same number of subapertures or the same plate scale on the WFS. It will be shown shortly that a change in lenslet array may be enough to match the original number of subapertures and plate scale. A change in lenslet array can be a practical approach for some AO systems (such as the Keck and

Lick AO systems) since these systems already have a mechanism for interchanging lenslet arrays.

As an example of this adaptation, consider the AO system on the Lick Observatory 3 meter telescope (Bauman 1999). In this system, an f/28 beam is collimated by a 100mm focal length lens into a 3.57mm diameter pupil which is then incident on a lenslet array of 500μ wide lenslets with focal length of 10.7mm. The pupil is ~7 lenslets wide, but is spread out over 8 lenslets by aligning the pupil so that the edge subapertures are half-illuminated. These Hartmann spots, which are nominally 500μ apart, are demagnified by a factor of ~3.5 to approximately 144μ apart at the CCD plane, which is 6 pixels (each pixel is 24μ).

The plate scale at the f/28 focal plane is 2.5 arcsec/mm, and the plate scale at the WFS CCD is 1.9 arcsec/pixel. The fact that the f# of the lenslets (10.7mm/0.5mm=f/21.4) is smaller than the f# of the incoming beam indicates that when we convert the WFS into a PWFS, that the pupils will not overlap (equation 3). Let us evaluate the other details of the adaptation.

With the lenslet array at the f/28 image plane, the 500μ wide lenslets become the field angle pixels. Thus a 500μ wide lenslet subtends (2.5arcsec/mm)(0.5mm)=1.25 arcsec, which is different from our original plate scale (1.9 arcsec), but is still a reasonable plate

scale with ~1 arcsec seeing, which is typical for Lick. If it were necessary, we could match the original plate scale by choosing 750µ pitch lenslets ((2.5arcsec/mm)(0.75mm)=1.9 arcsec).

The lenslets will produce a pupil ~10.7mm from the lenslets, which will be 10.7mm/28=382µ in diameter. After the demagnification of 3.5x by the relay, we have a pupil of 111µ, which is 4.6 pixels. Thus, this PWFS will have ~5 subapertures rather 7 subapertures across the beam diameter. This is a noticeable difference in the order of wavefront that can be corrected, but it is good enough to run a simple trial. If we wanted to match the pupil size of the SH WFS, we would use lenslets with a longer focal length. Using the equations given in earlier, we conclude that a focal length of 16.5mm is appropriate.

The SH WFS in the Lick AO system is located on a stage that translates along the optical axis; the primary purpose of the stage is to allow the WFS to be used with LGS's. The stage does not have the travel necessary to move the lenslet array/relay optics/CCD upstream enough (100 mm) to bring the lenslet to the focal plane. The entire stage could be moved on the optical bench, but for purposes of a PWFS trial, it would be sufficient to move the image plane by defocusing the telescope and then refocusing the science camera, which does have sufficient travel. This offers a minimally invasive test.

3.8. Diffraction analysis of PWFS measurements

For a SH WFS, in cases where the subapertures are small enough so that the wavefront over a subaperture can be approximated as a tilt, it is fairly straightforward to interpret the data in a SH WFS and to reconstruct the wavefront. For a PWFS, it is somewhat less obvious. In the limit of geometric optics, i.e., large aberrations compared to the wavelength, it is apparent that SH WFS's and PWFS's produce the same answer—that they perform the same functions (division of aperture into subapertures; division of field into quadrants) but in opposite order.

In the case of a diffraction-limited or partially-corrected beam, the interpretation is less obvious for the PWFS. As mentioned earlier, a facet of a quad-cell PWFS can be thought of as a Foucault knife-edge test where a knife-edge has been implemented along both x and y axes simultaneously. Although the Foucault knife-edge test was initially a qualitative measure, later theory was developed to handle analysis of diffraction-limited wavefronts under the Foucault test. (Gaviola, Linfoot, Barakat, Katzoff, Wilson) These analyses make clear that one can measure and evaluate quantitatively a diffraction-limited wavefront using a Foucault knife-edge (or with a pyramid WFS, which gives both knife-edge orientations at once). Linfoot's work describes the forward analysis, i.e., it predicts the pattern that results from a given wavefront error. Katzoff's work solves the reverse analysis for small aberrations, i.e., given a pattern from the Foucault test, it deduces the wavefront error present.

A brief derivation of the equations describing the diffraction theory of the focal-plane mask test (e.g., Foucault knife-edge test, PWFS) is presented next. The derivation parallels similar ones in Linfoot, Barakat, Wilson, and Feeney. Figure 3-18 illustrates the layout of a focal-plane mask test.
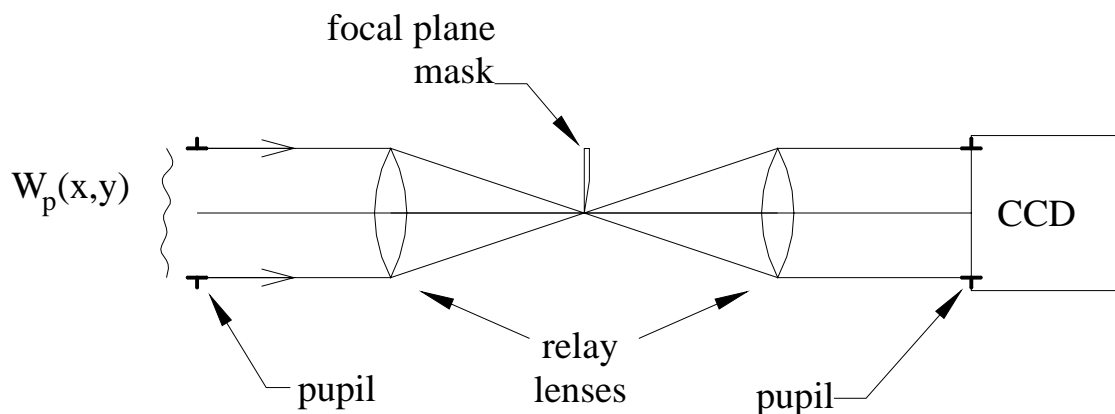
Figure 3- 18: Layout of focal-plane mask test. $W_p(x,y)$ is the wavefront to be measured, which is located at a pupil. A lens focuses the light onto a focal-plane mask and a second lens relays a pupil onto a CCD. A knife-edge mask is shown, but any focal-plane mask can be implemented. For convenience, the relay is assumed to be 1:1; the focal length of each lens is f.

The electric field at the pupil, W, is defined by

$$W_p(x,y) = A(x,y)\exp\big(2\pi j\phi(x,y)\big),$$
$$\text{where } A(x,y) = \begin{cases} 1, \text{inside aperture} \\ 0, \text{outside aperture} \end{cases} \tag{4}$$
$$\phi(x,y) = \text{ wavefront phase in waves}$$

The phase function $\phi(x, y)$ is what we seek to measure. The symbol U will be used for the electric field; a subscript, "i" or "p" denotes that the field is taken at an image plane or pupil plane, respectively.

At the image plane just before the focal-plane mask (e.g., knife-edge), the electric field is just the Fourier transform of the pupil function:

$$U_i^- (\xi, \eta) = \mathcal{F}\left(W_p (x, y)\right) \qquad (5)$$

The "-" superscript indicates that the electric field is taken just before the pyramid; a "+" superscript will indicate that the field is taken just after the focal-plane mask. ξ and η represent the coordinates in the image plane.

A mask function $K(\xi, \eta)$ is now applied in the focal plane, which multiplies the electric field just before the mask:

$$\begin{aligned} U_i^+ (\xi, \eta) &= U_i^- (\xi, \eta) \cdot K(\xi, \eta) \\ &= \mathcal{F}\left(W_p (x, y)\right) \cdot K(\xi, \eta) \end{aligned} \qquad (6)$$

For a knife-edge,

$$\begin{aligned} K(\xi, \eta) &= \text{step}(\xi) \cdot 1(\eta), \\ \text{where step}(\xi) &\equiv \begin{cases} 0, & \xi < 0 \\ 1, & \xi > 0 \end{cases} \\ 1(\eta) &= 1 \text{ for all } \eta \end{aligned} \qquad (7)$$

The function $1(\cdot)$ is a bookkeeping tool so that variable dependencies are not lost.

For one pyramid facet in a PWFS,

$$K(\xi,\eta) = \text{step}(\xi)\cdot\text{step}(\eta) \tag{8}$$

After the image plane mask, a subsequent lens creates a pupil, and so inverse Fourier transforms the electric field:

$$
\begin{aligned}
U_p(x,y) &= \mathcal{F}^{-1}\left(U_i^+(\xi,\eta)\right) \\
&= \mathcal{F}^{-1}\left(\mathcal{F}\left(W(x,y)\right)\cdot K(\xi,\eta)\right) \\
&= W(x,y)*\mathcal{F}^{-1}\left(K(\xi,\eta)\right)
\end{aligned} \tag{9}
$$

As a check, if there is no image-plane mask, then

$$
\begin{aligned}
K(\xi,\eta) &= 1(\xi)\cdot1(\eta) \\
&\Downarrow \\
\mathcal{F}^{-1}\left(K(\xi,\eta)\right) &= \delta(x)*\delta(y)
\end{aligned}
$$

where $\delta(x)$ is the Dirac delta function. After convolution, we have

$$
\begin{aligned}
U_p(x,y) &= W_p(x,y)*\delta(x)*\delta(y) \\
&= W_p(x,y)
\end{aligned}
$$

which is what we would expect.

For the case of a vertical knife-edge in the focal plane, $K(\xi,\eta) = \text{step}(\xi) \cdot 1(\eta)$. Taking the inverse Fourier transform of $K(\xi,\eta)$, we have

$$\mathcal{F}^{-1}\left(\text{step}(\xi) \cdot 1(\eta)\right) = \left(\tfrac{1}{2}\delta(x) - \frac{1}{2\pi j}\frac{1}{x}\right) * \delta(y) \tag{10}$$

and so the final pupil plane electric field is

$$\begin{aligned}
U_p(x, y) &= W(x, y) * \left(\tfrac{1}{2}\delta(x) - \frac{1}{2\pi j}\frac{1}{x}\right) * \delta(y) \\
&= W(x, y) * \left(\tfrac{1}{2}\delta(x) - \frac{1}{2\pi j}\frac{1}{x}\right)
\end{aligned} \tag{11}$$

At this point, Linfoot, Wilson, and Feeney explicitly write the convolution integral:

$$U_p(x, y) = \frac{1}{2}W(x, y) - \frac{i}{2\pi}\int_{-\infty}^{\infty}\frac{W(x', y)}{x' - x}dx' \tag{12}$$

For a given point (x,y) in the pupil, the integral is taken along the horizontal line through (x,y). Using the fact that W(x,y) is zero outside the pupil, we come to the equation most often cited:

$$U_p(x, y) = \frac{1}{2}W(x, y) - \frac{i}{2\pi}\int_{-B(y)}^{B(y)}\frac{W(x', y)}{x' - x}dx' \tag{13}$$

where B(y) represents the x-coordinate at the edge of the pupil at a vertical coordinate of y; see figure 3-19.
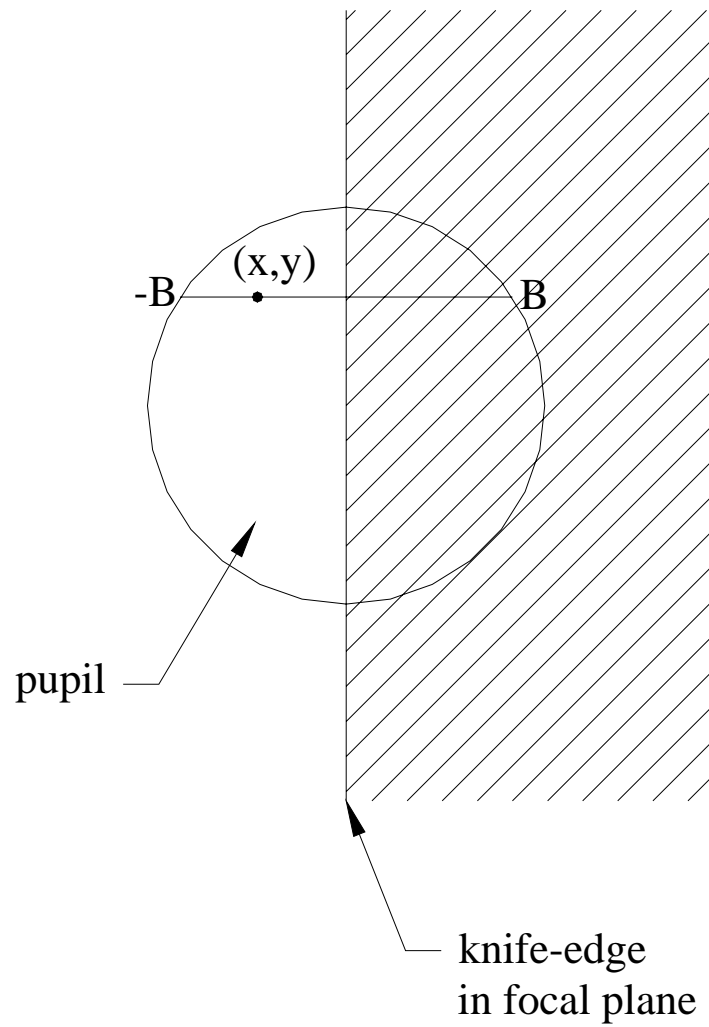
Figure 3- 19: Geometry of line integral in equation 13.

Facet/lenslet size considerations in multi-cell PWFS's

Equation 9 provides the basis for deciding what size the field pixels in a multi-cell array should be. Suppose that we have a pupil function W(x,y) with diameter D and that we wish to have n subapertures across the pupil, i.e., the desired resolution is the pupil plane is D/n. For ease of use, we want the function that is convolved with W(x,y) in equation 9 to have width $\lesssim$ D/n; this means that $\mathcal{F}^{-1}\left(K\left(\xi,\eta\right)\right)$ has width ~D/n. This in turn yields that $K\left(\xi,\eta\right)$ should have a width $\gtrsim$ n (λf/D), i.e., n times the diffraction spot size due to the full aperture; this is the same as the diffraction spot size due to a subaperture.

As an example, suppose the facet/lenslet in the focal plane is a square with side $a\left(\dfrac{\lambda f}{D}\right)$; "a" represents the width of the facet/lenslet in units of the diffraction spot size. Then,

$$K\left(\xi,\eta\right)=\left(\frac{1}{a\,\lambda f/D}\right)^{2}\mathrm{rect}\left(\frac{\xi}{a\,\lambda f/D}\right)\mathrm{rect}\left(\frac{\eta}{a\,\lambda f/D}\right)$$

and so,

$$\mathcal{F}^{-1}\left(K\left(\xi,\eta\right)\right)=\mathrm{sinc}\left(\frac{x}{D/a}\right)\mathrm{sinc}\left(\frac{y}{D/a}\right)$$

This is the function that is convolved with W(x,y)—it is a sinc function with a width which is 1/a times the pupil size. If D/a is the size of a subaperture, then we will get the same blurring of subapertures given in section 3.3, which is generally acceptable. But if a is much smaller than the number of subapertures, then the sinc function becomes broad

compared to the subaperture size and the slope for a given subaperture becomes more difficult to discern because of "pollution" from the other nearby subapertures. Thus, if the PWFS is to be used with a large number of subapertures, then the facets/lenslets must be large (many diffraction spot sizes wide). For a beam with large aberrations (large enough to span several pyramids/lenslets), then the multicell approach can be used. But for a (nearly-) diffraction-limited spot, either the spot will be near the junction of four facets/lenslets, in which case, it is essentially the previous quad-cell arrangement; or the spot will be away from junctions of facets/lenslets, in which case no knife-edge or modulation is applied, and so no information results. Thus, the multicell approach is best applied to cases of measuring large aberrations.

It is worth noting that the above discussion illuminates the trade-off between resolution in the pupil and resolution in the image plane (i.e., resolution in measuring slope). Having the spot span more field pixels improves the slope measurement, but at the cost of reducing the number of subapertures (to avoid confounding the wavefront measurement). In the limit of several facets/pyramids spanning a diffraction-limited spot, the sinc function which is convolved with the pupil becomes comparable to (or larger than) the size of the pupil, and so the information about subapertures becomes heavily confounded.

Interpretation of the knife-edge convolution equation (equation 13) in terms of the derivative of the pupil function

One may wonder, looking at equation 13, where exactly the measurement of the wavefront slope takes place. This is most easily seen by returning to equation 11 and working backwards: what function should the pupil function be convolved with in order to generate the derivative of the wavefront? The function that produces a derivative after convolution is the derivative of a delta function, $\delta^{(1)}(x)$ (Gaskill). Checking that convolving with $\delta^{(1)}(x)$ produces the correct result:

$$
\begin{aligned}
U_p(x, y) &= W(x, y) * \delta^{(1)}(x) \\
&= \left[A(x, y)\exp\left(j\varphi(x, y) \right) \right] * \delta^{(1)}(x) \\
&= \frac{\partial}{\partial x}\left(A(x, y)\exp\left(j\varphi(x, y) \right) \right)
\end{aligned}
\tag{14}
$$

A(x,y) is constant (=1 inside the aperture), except at the edge of the aperture, so we can pull A(x,y) out from the derivative for pupil points away from the aperture boundary.

Continuing,

$$
\begin{aligned}
U_p(x, y) &= \frac{\partial}{\partial x}\left[A(x, y)\exp\left(j\varphi(x, y) \right) \right] \\
&= A(x, y)\frac{\partial}{\partial x}\exp\left(j\varphi(x, y) \right) \\
&= jA(x, y)\,\exp\left(j\varphi(x, y) \right)\left\{ \frac{\partial\varphi(x, y)}{\partial x} \right\}
\end{aligned}
\tag{15}
$$

Thus, $U_p(x,y)$ is proportional to the wavefront slope (the factor in braces); all other factors have magnitude 1. That is, the wavefront slope has been manifested as amplitude. This confirms that convolving $W(x,y)$ with $\delta^{(1)}(x)$ is the desired result.

Now the question of identifying the slope measurement in equation 13 can be answered. If $W(x,y)$ were convolved with $\delta^{(1)}(x)$, we would get a derivative of the wavefront (wavefront slope). But in a knife-edge test, $W(x,y)$ is not convolved with $\delta^{(1)}(x)$; it is convolved with $1/x$, which can be recognized as a poor man's approximation to $\delta^{(1)}(x)$; see figure 3-20. In other words, to the (limited) extent that $1/x$ approximates $\delta^{(1)}(x)$, the knife-edge test returns a measurement of the slope of the wavefront in the direction perpendicular to the knife-edge.

f(x)

$\delta^{(1)}(x)$

$\frac{1}{x}$

x
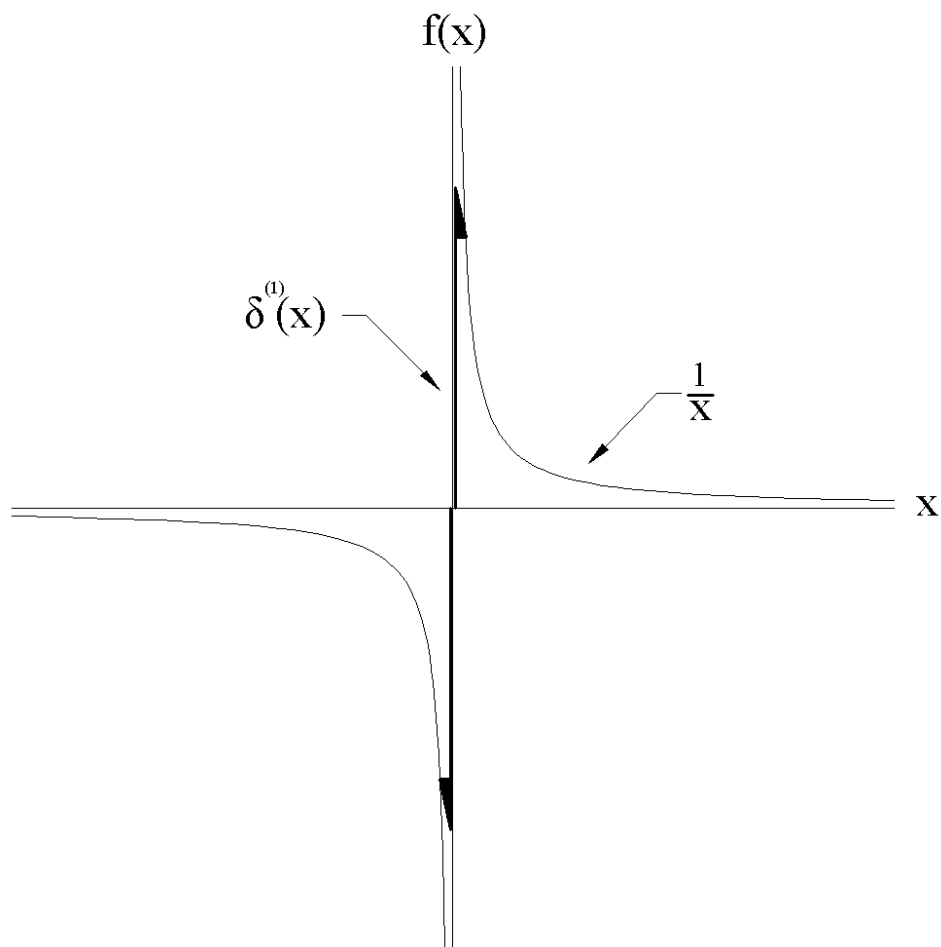
Figure 3- 20: Comparison of f(x)=1/x versus f(x)= $\delta^{(1)}(x)$. 1/x can be seen as an approximation to $\delta^{(1)}(x)$.

As a side note, the derivative of A(x,y) is infinite at the pupil boundary, and so one would expect infinite (or practically, very large) intensities at the edge of the aperture. The large intensity on the pupil boundary is, in fact, observed in the Foucault knife-edge test (known as a Rayleigh diffraction ring), as one might expect given the comparison of 1/x and $\delta^{(1)}(x)$.

The next logical question is what mask should be introduced at the focal plane in order to produce this desirable convolution with $\delta^{(1)}(x)$? This question was answered (Sprague, Horwitz 1976, Horwitz 1994) by taking the inverse Fourier transform of $\delta^{(1)}(x)$ in order to find the required focal-plane mask, $T(\xi)$:

$$
\begin{aligned}
T(\xi) = \mathcal{F}^{-1}\left(\delta^{(1)}(x)\right) = -j2\pi\xi \\
\propto (\xi)
\end{aligned}
\tag{16}
$$

This is a linear gradient transmission mask as shown in figure 3-21. The negative amplitude for x<0 can be implemented with a $\lambda/2$ phase plate. Of course, the transmission cannot increase indefinitely; it is limited to a maximum of unity, but as long as the focal plane spot were much smaller than the spatial extent of the transmission mask, this would not be a serious limitation.
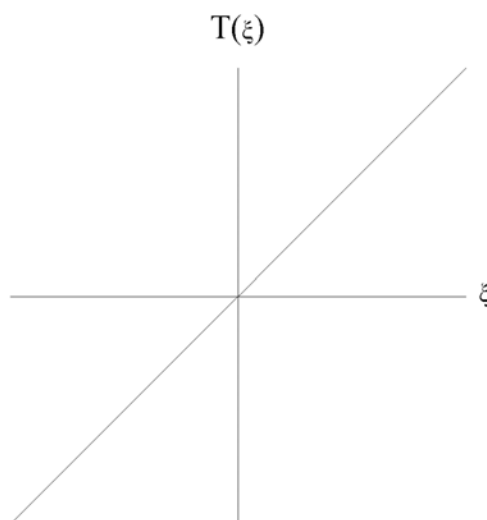
Figure 3- 21: Transmission versus position for linear gradient transmission mask described in equation 16.

To avoid using the λ/2 phase plate, one could add a constant term to the transmission mask, which would add a δ(x) term to the convolution in equation 14. This yields a constant intensity pedestal added to the pupil $U_p(x, y)$.

3.9.   Example of a PWFS measuring an aberration

In this section, a numerical example of a PWFS measuring an aberration (coma) is given. The present simulation uses the middle line of equation 9, implemented with fast Fourier transforms. The pupil intensity is calculated separately for each quadrant of the PWFS and all four pupils are displayed. The amount of coma varies from 1λ to 17λ peak-to-valley (P-V). Figure 3-22 displays the results.

Two items are of note in figure 3-22. First, one may observe bright pupil edges nearest and furthest from the apex of the pyramid. This may be seen as the PWFS analogue of the Rayleigh diffraction ring seen in the Foucault knife-edge test. Second, one may observe fringing in the pupils for aberrations larger than about 5 waves P-V. Certainly, the fringing complicates wavefront reconstruction and would defeat a zonal adaptive optics control matrix, which relies upon a linear (or at least monotonic) relationship between the wavefront slope and the centroid calculated from the pupil intensities. An explicit wavefront reconstruction would be necessary before applying such a control

matrix. Ragazzoni deals with the problem by scanning the guide star image around the pyramid apex so that the fringing is blurred out.

1λ coma

5λ coma
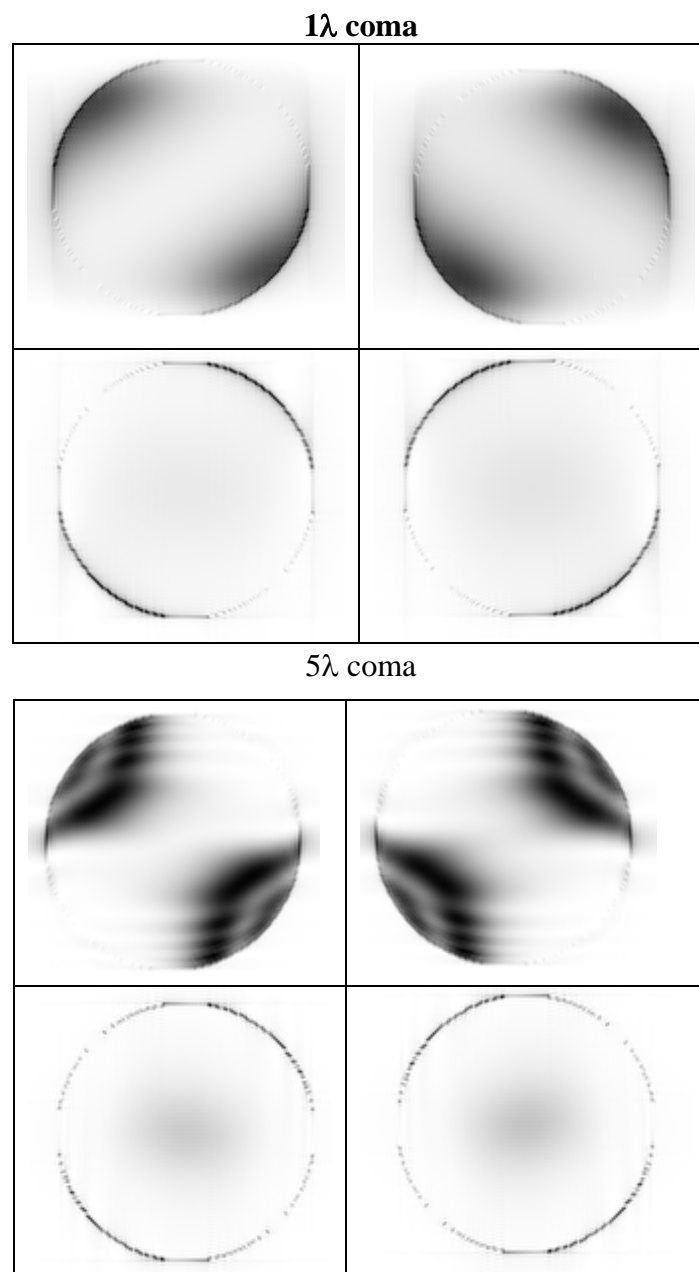
Figure 3- 22: PWFS output from wavefront with 1 wave and 5 waves (this page), and 10 waves and 17 waves P-V of coma (next page).

10λ coma
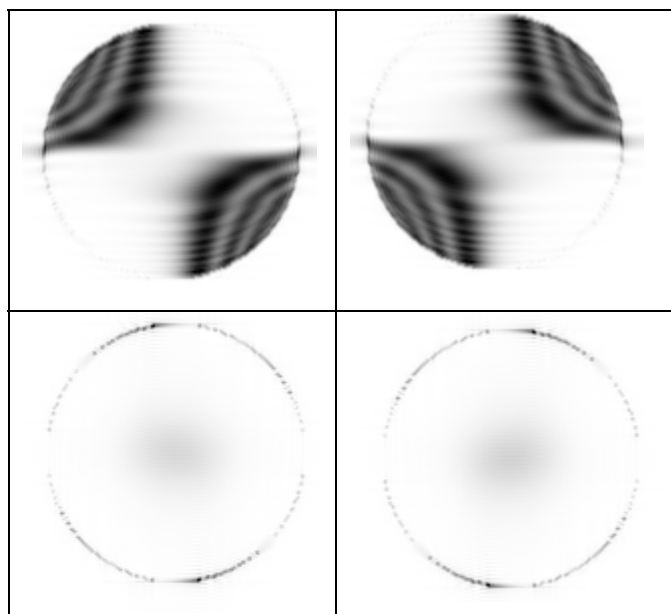
17λ coma
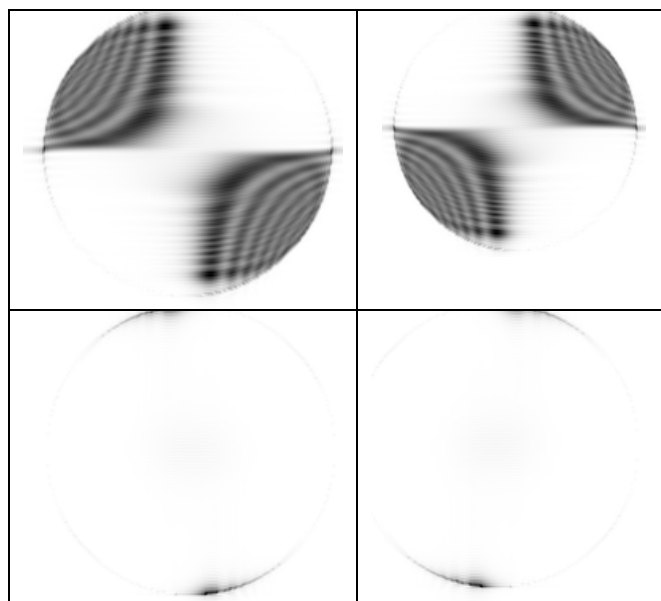
Figure 3-22, continued.

It is apparent that PWFS's have the potential to measure larger aberrations; however, doing so requires dealing with the fringing in the pupils. Current analytic Foucault knife-edge analysis techniques do not support the deduction of larger wavefront errors, although one could in principle use a forward analysis and trial-and-error approach to deduce the wavefronts.

Clearly, more work is required here for PWFS's to be useful in measuring larger aberrations without modulation; this is reserved for future work.

4. LASER GUIDE STAR SPOT ELONGATION

4.1   Introduction

Most currently contemplated ELT AO systems plan to use some form of laser guide star in operation in order to increase the AO systems' sky coverage. A complete treatment of the optical engineering required to make laser guide stars work on ELT's is beyond the scope of this work—indeed, this topic would fill several volumes! Instead, we will focus on one important, "high-leverage" aspect of LGS's for ELT's: LGS spot elongation. LGS elongation results from the following combination of facts:

- LGS's are located at a finite height above telescope

- LGS's have a finite "depth"

- Subapertures of the wavefront sensors are in general not directly underneath the LGS, i.e., the subapertures are separated from the LGS's launch aperture

Figure 4-1, previously shown in Chapter 1 and reproduced here for convenience, shows an example of spot elongation on a SH WFS.
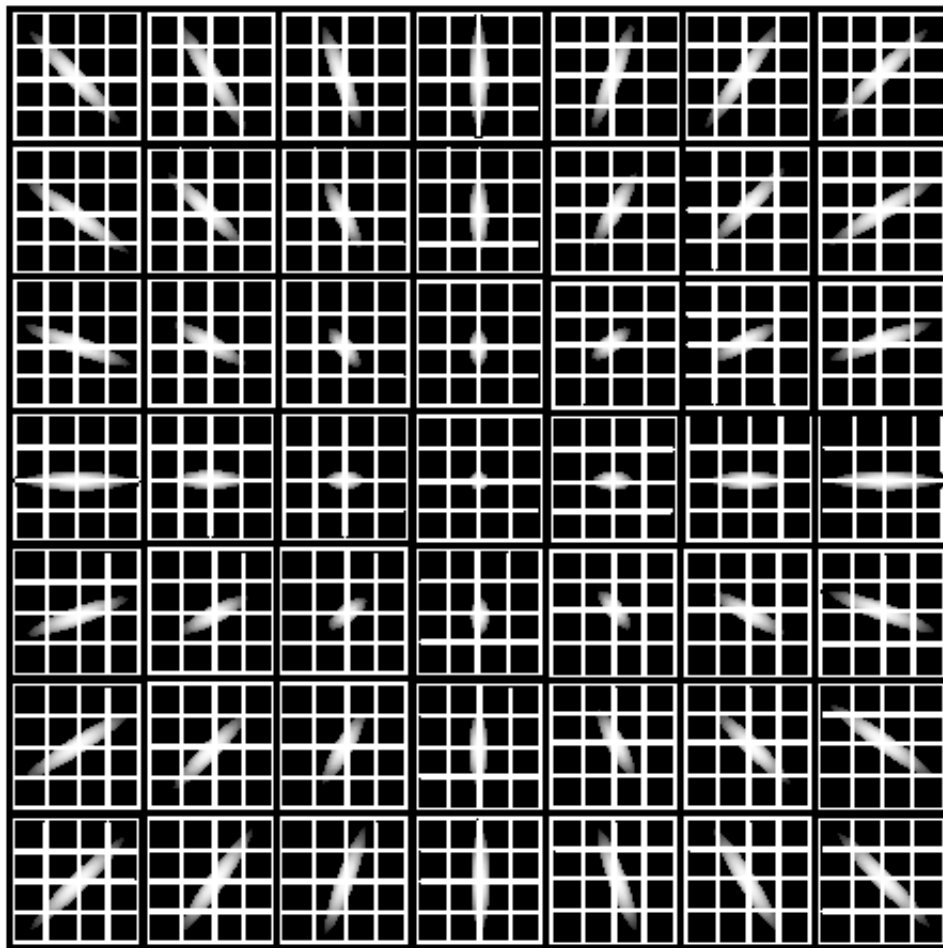
Figure 4- 1: Schematic illustration of LGS spot elongation effect in a Shack-Hartmann WFS. (From Goncharov)

4.2 Sodium LGS's and Rayleigh LGS's

As explained earlier, LGS's act as a source nominally above the atmosphere. The light from the LGS propagates downward to the telescope, propagating through the atmosphere that we wish to measure. LGS's are of two types: "Sodium" LGS's and "Rayleigh" LGS's.

Sodium LGS's

By serendipity, it turns out that there is a layer of sodium located approximately 90km above the earth, in the mesosphere, believed to be formed from meteors breaking up in the upper atmosphere. This layer is approximately 10km thick. The properties of the sodium layer vary considerably: sodium may be deposited anywhere between 80-110km high and sodium densities can vary by approximately a factor of 3 according to the time of year. The fact that the sodium layer is at a fixed height above the earth leads to an important point: that the apparent height of the sodium layer (as well as the apparent depth of the sodium layer) is proportional to the secant of the zenith angle (figure 4-2).
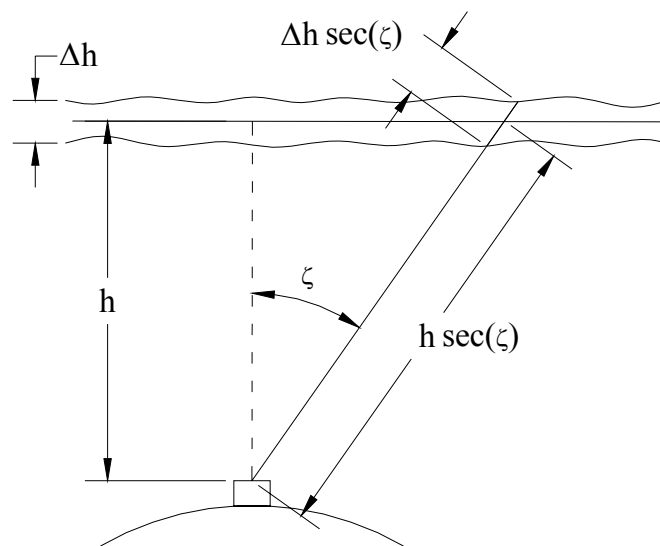
Figure 4- 2: Sodium guide star range and thickness variation with zenith angle. The sodium layer has a nominal height h with a depth of Δh. As the zenith angle increases, the sodium guide star becomes more distant and thicker proportional to the secant of the zenith angle.

The sodium LGS is created by projecting a laser, which is tuned to the D_2 line of the sodium atom (~589nm), into the atmosphere. The laser light excites the sodium and upon decay, 589 nm light is re-emitted. Currently, current facility-grade sodium LGS's are made with dye lasers, although other techniques are under investigation. "Sum-frequency" techniques mix two lasers in a crystal with refractive index non-linearities (Pennington). The two laser wavelengths are chosen so that the sum of their frequencies equals the frequency corresponding to a wavelength of 589nm.

There are a few important issues with sodium LGS's. First, the power has been limited: dye lasers that are in current use are limited to about 15-20 watts, which produces an apparent 9th magnitude LGS. As we will see, spot elongation will reduce the apparent brightness of the LGS. Second, the current approaches do not yet support a pulse format compatible with sodium LGS pulse tracking, which will be discussed later in this chapter.

Rayleigh LGS's

Rayleigh LGS's are, as the name suggests, created by Rayleigh scatter from a laser projected into the atmosphere. Since Rayleigh scatter is produced at every height in the atmosphere, Rayleigh LGS's are pulsed and range-gated on the WFS camera so that the WFS only sees light from a given altitude (or more accurately, a range of altitudes). Rayleigh LGS's are often made at short wavelengths (e.g., ultraviolet) since Rayleigh

scatter efficiency is proportional to λ^{-4}. The efficiency is an issue because we want to measure light from as high as possible in the atmosphere so that we measure as much of the desired atmosphere as possible (described further in the next paragraph). Of course, the atmospheric density (and so scattering particle density) decreases exponentially with increasing altitude (with decay parameter h_0); the collected return flux is proportional to $1/h^2$. Thus, the return photon flux decreases as $\dfrac{1}{h^2}\exp\left(-\dfrac{h}{h_0}\right)$ with increasing LGS height. In selecting an altitude for the Rayleigh LGS, the decreasing photon flux with altitude provides an opposing design pressure to the improved geometry of measurement with altitude.

Referring to figure 4-3, light from a LGS at a finite height will not probe the atmosphere above the LGS, nor will it probe the atmosphere outside the cone defined by the LGS and the aperture of the telescope. The error resulting from these incomplete measurements is referred to as focal anisoplanatism (a.k.a., "cone effect") (Fried 1994). In general, focal anisoplanatism becomes worse with lower LGS altitude, so that sodium LGS's (at ~90km altitude) suffer less focal anisoplanatism than Rayleigh LGS's (~10-40km).
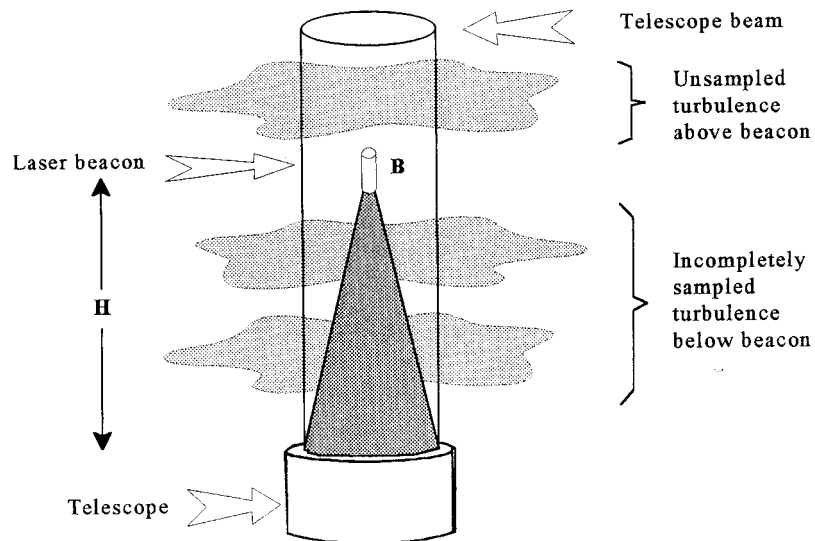
Figure 4- 3: Focal anisoplanatism. (from Hardy)

While sodium LGS's vary in height and depth according to zenith angle, Rayleigh LGS's need not suffer these effects, although one might imagine that at large zenith angles, it might be desirable to position the LGS further away so that it remains above the bulk of the atmosphere.

4.3 Geometry of LGS spot elongation

The geometry that leads to LGS spot elongation is shown in figure 4-4 below. h is the average altitude of the LGS, Δh is the depth of the LGS, and d is the distance between the launch aperture and the subaperture under consideration. Although the LGS is shown directly above the telescope, this is not required.
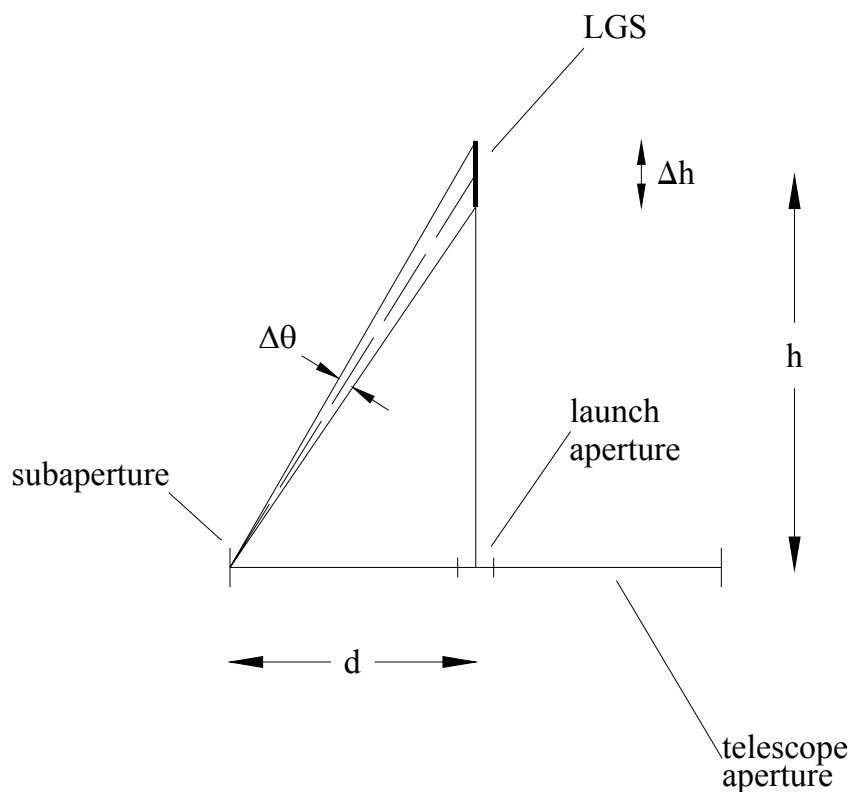
Figure 4- 4: Geometry of laser guide star showing spot elongation. The LGS, shown as a heavy line above the telescope, is located at an altitude of h with a depth of Δh. The elongation varies with the distance of the subaperture from the launch aperture (d); the left-most subaperture shown here. The angular elongation from the perspective of this subaperture is $\Delta\theta = (\Delta h/h)(d/h) = \Delta hd/h^2$.

In general, the elongation increases with the depth of the LGS and with the distance of the subaperture from the launch aperture of the LGS; the elongation decreases with LGS height.

For sodium LGS's on small telescopes, this elongation is benign—it is negligible for the side-launched Lick telescope with $d_{max} \approx 4$ meters, which yields a maximum elongation of

1 arcsec, which is significantly smaller than the ~1.5-2 arcsec blur size of the LGS (the laser is blurred by atmospheric seeing on the way up to the sodium layer and on the way back to the telescope).

For larger telescopes (say, 30m), the elongation becomes problematic. Assuming that the LGS launch aperture is behind the secondary to minimize the elongation, d_{max} will be 15m. As a result, we will have for CELT an elongation of (15m/90km)(10km/90km) = 4 arcsec, which is a factor of 4 greater than the Lick example. Rayleigh LGS's also exhibit large elongation due to their low altitude. For a 30m telescope, a 10km thick LGS at an average altitude of 25km will have 250μrad≈52 arcsec of elongation!

What are the implications of this spot elongation? We examine the effects below.

Our wavefront sensor uses the spots to estimate the local wavefront slope. The rms error in the estimate of the spot centroid is proportional to the spot size (Hardy):

$$\sigma_1 = \frac{3\pi}{16}\left(\frac{\theta}{SNR}\right) \qquad (1)$$

where σ_1 = rms error of centroid estimate in one axis

θ = FWHM of spot (Gaussian profile assumed)

SNR = signal-to-noise ratio of measurement

Thus, we will have a much poorer estimate of the centroid in the long axis; the estimate of the centroid along the short axis is basically unaffected.

The poorer estimate in the long axis, in principle, could be overcome with more photons from the guide star in order to improve the SNR. Assuming that we are signal-photon noise limited (i.e., the Poisson noise from the signal photons is much larger than background photon noise and detector read noise), SNR=(number of signal photons)$^{1/2}$. Thus, an additional factor of 16 in power will recover the loss due to the elongation. For sodium guide stars, which have been reluctant in producing a surfeit of power, this is not a desirable solution.

We might also wonder whether the elongation might make the centroid estimation more sensitive to the actual vertical distribution of the backscattering aerosol. For example, we will get very different estimates of the long-axis centroid when the top of the sodium layer is relatively bright than when the bottom of the sodium layer, 4 arcsec away for an edge subaperture, is relatively bright. This turns out not to be a problem, however. The shift of the centroid is proportional to the spot elongation, which is proportional to the distance of the subaperture from the launch aperture; this, not surprisingly, is just what focus looks like on a SH WFS. Thus, as the vertical distribution of the backscattering aerosol changes, we will merely measure a change in focus. For a sodium LGS, we already ignore the focus term since the sodium layer height is unknown; we find the

focus from an auxiliary natural guide star (which can be dim, since we are measuring low-order modes, and so the subapertures can be large) (CARA). Now, in the case of range-gated LGS's, we could, in principle, know the LGS height a priori, so we need not ignore the focus term from the WFS measurement. Using a NGS for focus, in any case, is not an imposing requirement, if needed.

An additional consideration is that if we have an elongated spot for each subaperture, this will greatly increase the number of pixels necessary for the WFS. Thus, the centroid estimation will become noisier due to the read noise of the additional pixels; this is only a problem if the photon count is so low that the Poisson noise of the signal photons does not dominate the detector read noise. An additional, very practical, consideration is simply that we will need wavefront sensors with larger number of pixels, which certainly increases cost and complexity.

4.4 Dynamic refocusing

One might try to solve the spot elongation problem by using a pulse format for the laser and range-gating over a smaller depth. The problem is that LGS WFS's are generally light-starved, so limiting the elongation by "throwing away" photons (say, 75% to reduce the elongation from 4 arcsec to 1 arcsec) is not appetizing.

4.4.1 Resonating discrete mirror

As an alternative to range-gating over a small depth, Angel, et al. have described and developed a method for dynamically refocusing the light as the LGS "pulse" returns from different heights in the atmosphere ((Angel 2000, Lloyd-Hart & Georges, Georges). This method involves moving a mirror axially via a resonating mechanical structure so that the image is stationary even as the object is moving. In designing the optical system, the required displacement of the mirror is plotted versus time; this plot is not linear since the depth of desired return layer is a sizable fraction of the altitude of the average altitude of the return layer. The required oscillator motion is determined by varying its amplitude, frequency, and starting point to match the desired motion via a least squares fit. This process is captured in figures 4-5 and 4-6 below. An optical pick-up from the mirror is then used to time the launch of the Rayleigh beacon pulses.
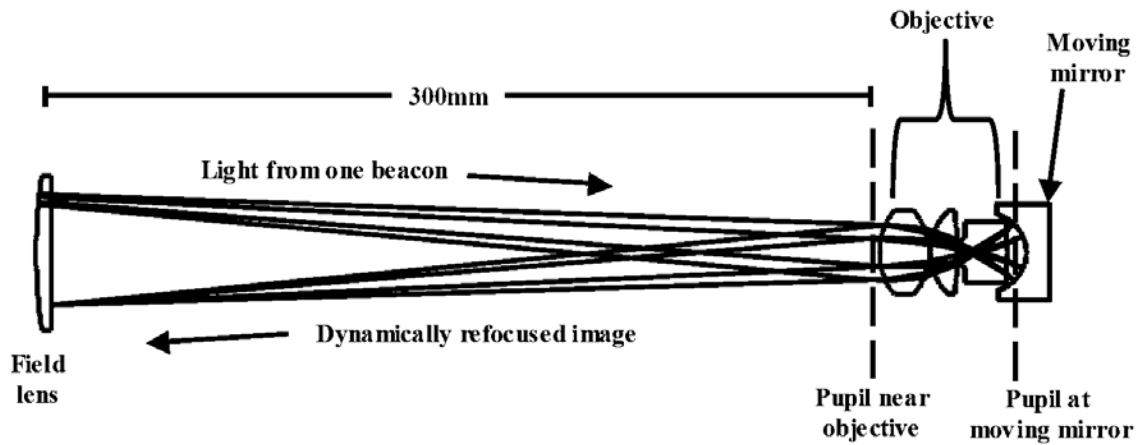
Figure 4- 5: Layout of dynamic refocusing unit. Light from the telescope (top left) passes through a field lens which forms a pupil 2mm away from the objective at right. The objective speeds up the beam to f/0.7 so that the resonating mirror needs to move only a short distance (~1mm) to keep the laser pulse's focus stationary following the mirror. The light, with stationary focus, is returned to the AO system through the objective and field lens. (Georges)
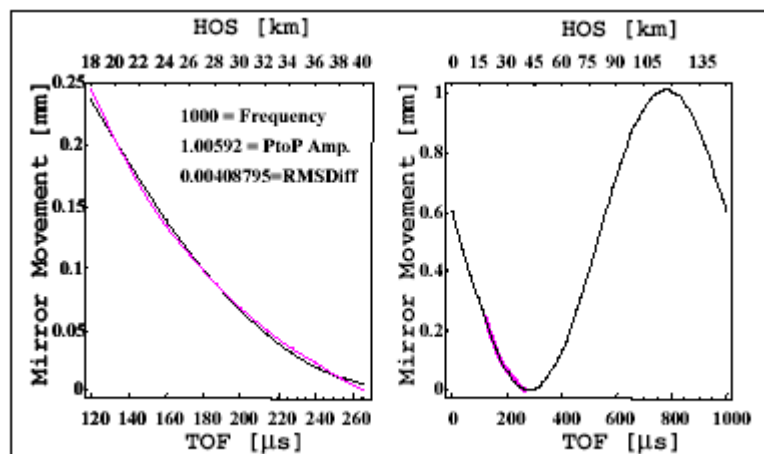
Figure 4- 6: Resonator mirror motion with LGS pulse timing. The resonating mirror moves in a sinusoidal pattern (right). A portion of the sinusoid is selected (left) such that the mirror's motion in this region is well-matched to the mirror motion required to track the focus of the LGS pulse. In this case, the sinusoid has a peak-to-peak amplitude of 1mm and frequency of 1 kHz. The selected portion of the sinusoid has a motion of 250μ over a duration of 140μs. TOF=time-of-flight; HOS=height of source (Georges)

A prototype for Rayleigh LGS's has been built and tested with encouraging results. Lloyd-Hart, et al. have proposed extending this method to sodium LGS's. This method is more complex since the distance and thickness of the sodium layer changes with zenith angle. By altering the amplitude of the resonator and the phase between the resonator and the launch pulse, one can adjust the mirror motion to keep the LGS light "in focus" (Lloyd-Hart, 2003).

This approach is promising, although there are still questions to be answered about altitude-dependent field aberrations arising from the optical relay.

4.4.2   Segmented micro-electromechanical system (MEMS) mirrors

A novel method for dynamically refocusing LGS light is proposed here. This method takes advantage of the fact that a SH WFS divides the LGS light into subapertures, so there is no penalty for using a segmented MEMS ahead of the SH WFS, as long as the segments boundaries correspond to the subaperture boundaries; i.e., relative piston between segments/subapertures is permissible.

One can use as the refocusing element a segmented MEMS array, ideally with one segment per subaperture and tip/tilt control over each segment; see figure 4-7. The desired power at any particular moment would be put on the MEMS array in a Fresnel

lens-like manner, i.e., the tilt at each subaperture would match the local tilt for the desired power, but the surface would not be continuous. This approach would not rely on "phase-wrapping" techniques, but if desired, phase-wrapping could be achieved by adding a piston degree-of-freedom to each segment.

Another way to view this approach is that from the point-of-view of a subaperture, the return light from an LGS pulse has varying (but known) amounts of tilt with change in altitude. Each segment would correct the tilt at any given moment for its corresponding subaperture. The spot elongation, then, is corrected on a subaperture-by-subaperture basis.

In order to reduce stroke requirements, the MEMS would be flat when the WFS is looking at light from the central altitude of the LGS; see figure 4-8. Any quasi-static focus (due to changes in the sodium layer height or changes in zenith angle) can be handled by translating the WFS or a discrete optic axially.
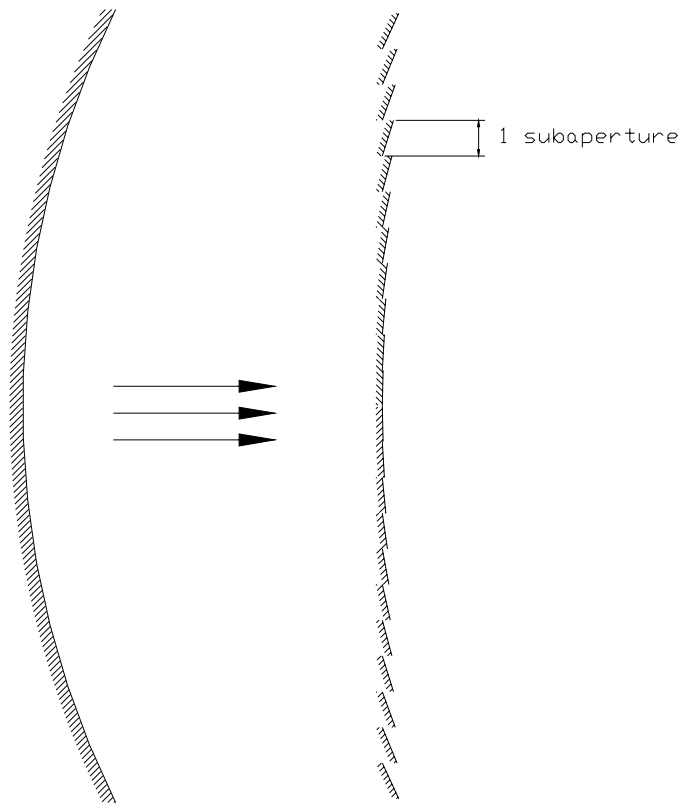
Figure 4- 7: Schematic of a segmented MEMS used for dynamic refocusing. The segmented MEMS at right is equivalent (over each subaperture) to the continuous mirror shape at left. Amplitudes are exaggerated.

143

bottom of   middle of   top of
  LGS   LGS   LGS
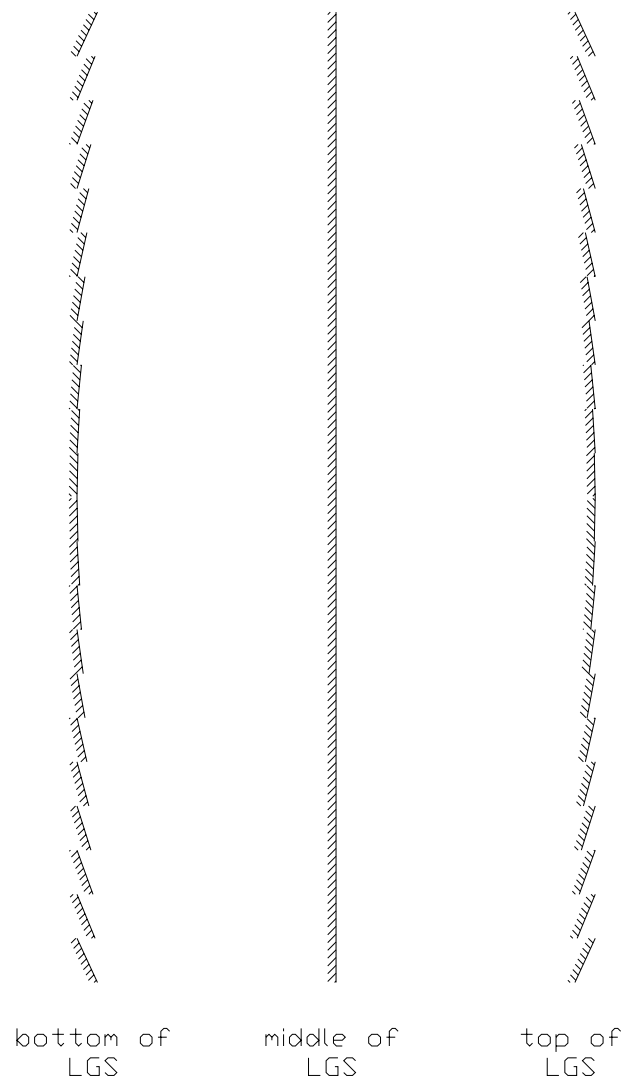
Figure 4- 8: Shape of segmented MEMS during tracking of a LGS pulse. Amplitudes are exaggerated.

There are two key questions to resolve for this technique: Do MEMS currently (or will soon) have the stroke necessary to accomplish the dynamic refocusing task; and are MEMS fast enough for the job?

To find the stroke necessary, we note that the largest stroke will be required by the edge subaperture. The stroke will be the change in local tilt required (i.e., the spot elongation) multiplied by the width of a subaperture. For a 0.30 m subaperture and a 4 arcsec $\approx$ 19μrad elongation, this corresponds to 5.7μ of OPD or 2.9μ of actuator motion.

The time required for this actuation is twice the time that it takes for a laser pulse to travel the "depth" of the LGS. For a sodium LGS, the depth is ~10km at zenith, so the time of flight, t, is

$$t = \frac{\Delta h}{c} = \frac{10km}{3x10^6 m/s} = 30\mu s \tag{2}$$

The actuation needs to take place, then, in ~60 μs. Since the vendors below quote actuation speed in terms of frequency of sinusoidal motion, it's useful to convert these frequencies to a characteristic actuation time. Since we are interested in the actuation in one direction only, we should use a characteristic actuation time that is half of the sinuosoidal period. For example, a 7kHz frequency corresponds to a period of 140μs, or a characteristic actuation time of 70μs. Of course, this characteristic actuation time does not tell the whole story, but it is a useful measure for determining feasibility.

Ideally, we would like this MEMS to operate open-loop, i.e., the segments go to where they are commanded to go; otherwise, we will need closed-loop control of the MEMS, which is an undesirable complication. The error due to open-loop operation should be much less than the error allotted for wavefront measurement in the error budget. The CELT error budget calls for a 40 nm measurement error (University of California and California Institute of Technology), so an open-loop rms error of 5-10nm would not take very much away from the budget.

We can now give a list of requirements for such a MEMS and compare them to commercially available MEMS from vendors Boston Micromachines (Perreault, Cornelissen) and Iris AO (Doble):

	Focus-tracking requirement	Boston Micromachines	Iris AO
MEMS type	Segmented with tip/tilt or piston/tip/tilt control	Segmented with tip/tilt control	Segmented with tip/tilt/piston control
Number of segments	~100 segments across diameter	32 x 32	7 across diameter
Stroke (full-range)	2.8μ	2-2.8μ	0.6° tilt = 6μ across 500μ segment
Speed	Full stroke in 60μs	7 kHz (70μs) in air, up to 70kHz (7μs) in vacuum	2-3 kHz (170-250μs) at full 6μ stroke
Open-loop accuracy	<5-10 nm	2 nm	2 nm

Table 4- 1: Requirements and current performance of segmented MEMS for dynamic refocusing. See text for discussion of speed parameters.

The actuator count is a little short currently, but efforts to reach 100x100 arrays are already underway. The speed and stroke parameters are at or near the specifications; commercial pressure in developing these devices for telecommunications and other AO uses will tend to push MEMS to higher speeds and stroke.

In addition to dynamically correcting for focus, one could use the MEMS to adjust the calibration of the WFS in the AO system; this would take the place of using WFS reference centroids that are deviated from the centers of their quad-cells. Furthermore, one could use the MEMS to correct for LGS-distance-dependent aberrations through the telescope and AO optical relay. In principle, this could be done quickly, correcting for

the LGS-distance-dependent aberrations due to the depth of the sodium layer, and/or slowly, to correct for zenith angle dependencies.

Figure 4-9 shows how the segmented MEMS might be integrated into the SH WFS leg of an AO system. The difference between this layout and that shown in Figure 3-1 is that there is an additional relay that is necessary so that both the segmented MEMS and the lenslet array can be located at pupils. Assuming that the segmented MEMS pixels are approximately 300μ wide (the typical value for vendor Boston Micromachines) and that we have 100 segments across the diameter of the beam, the MEMS size would be ~30mm. With an f/15 beam entering the WFS, the focal length of the collimating lens would be 30mm*15=450mm, which is a reasonable length to work with.
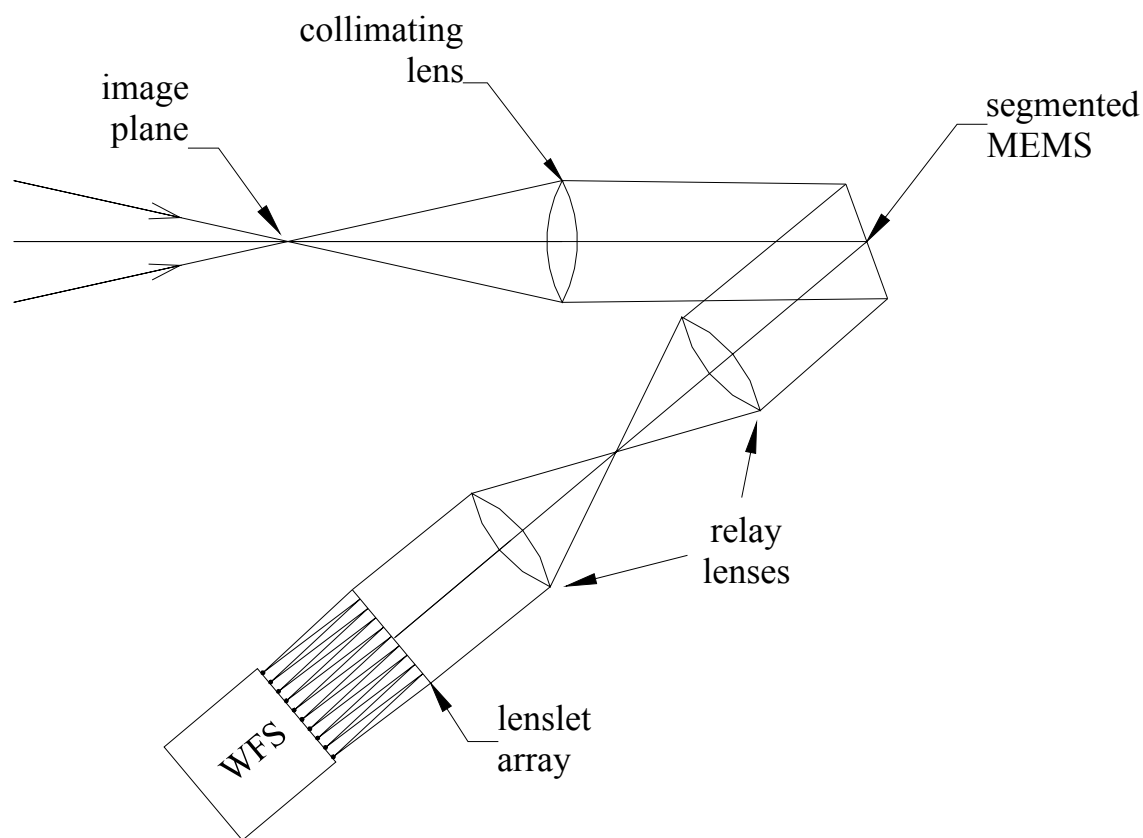
Figure 4- 9: An implementation of a focus-tracking, segmented MEMS into the WFS leg of an AO system; the scale is anamorphic. Converging light from the AO system is collimated by a lens, which forms a pupil. The segmented MEMS is placed at this pupil and is followed by a relay telescope, which creates a second pupil where the lenslet array of the SH WFS is located. Any relay optics between the lenslet array and WFS have been omitted for clarity.

Finally, one can expect that segmented MEMS will be attractive in their cost, relative to large programs like ELT AO. Current prices for MEMS are $100-300 per actuator (Macintosh), but it is likely that the price will drop significantly with time.

There is another approach that is possible: using a liquid crystal spatial light modulator rather than a MEMS. Spatial light modulators have the benefit of possessing very high spatial resolution, but there are two drawbacks. First, the stroke is typically quite limited—approximately one wavelength, thus requiring phase-wrapping. Second, current versions are slow (~video rates) since their development is driven by the display market.

4.5   Approaches to spot elongation with continuous-wave (CW) lasers

While dynamic refocusing mechanisms offer an excellent solution to the spot elongation problem, it may be that lasers to produce a sodium guide star with the appropriate pulse format will not exist or will not be practical within timescales of this project. In this case, it is helpful to have a strategy for using CW lasers. In fact, one might conclude that the CW strategy works well enough so that it is not worth the risk and cost in trying to develop sodium lasers with the right pulse format.

If our lasers are not pulsed, then we cannot use any of the previously mentioned pulse-tracking techniques, and so we should figure out how best to deal with elongated spots.

Looking at figure 4-1, we see that elongated spots will cover many pixels and that the elongated spot is, in general, not aligned to the pixel grid on the CCD. As mentioned earlier, this will necessitate arrays with more pixels per spot compared to the non-

elongated case; thus, the cost, complexity and perhaps noise of the measurement (if not photon-noise limited) will increase.

While the centroid of an elongated spot may be determined via a conventional center-of-mass calculation, there is another technique available. Poyneer has developed a cross-correlation technique to determine the position of an extended object (Poyneer 2003b). This method is sometimes called "scene-based wavefront sensing" because it can be used in sensing wavefronts from a surveillance scene. In applying this method, one cross-correlates the extended object with a template of the object. The method is quite robust, compared to conventional centroiding techniques, to variations in the object profile. Poyneer has modeled the performance of scene-based wavefront sensing in the LGS spot elongation case and has found that the technique works quite well, but performs best when the pixels are oriented along the elongated axis.

Thus, for cost/complexity, noise performance, and centroiding performance reasons, it is worthwhile to investigate methods that will align the CCD pixels to the local orientation of the spot elongation, on a subaperture-by-subaperture basis.

One approach is to design custom CCD's for this application. Beletic has proposed designing a CCD such that the pixels for each subaperture are aligned along the axis of the elongation; see figure 4-10. One possible mode in this approach can be used with

pulsed LGS's: the accumulated charge in the detector is moved along the CCD detector in time with the returned laser pulse so that the charge is integrated in a "time delay and integrate" manner.
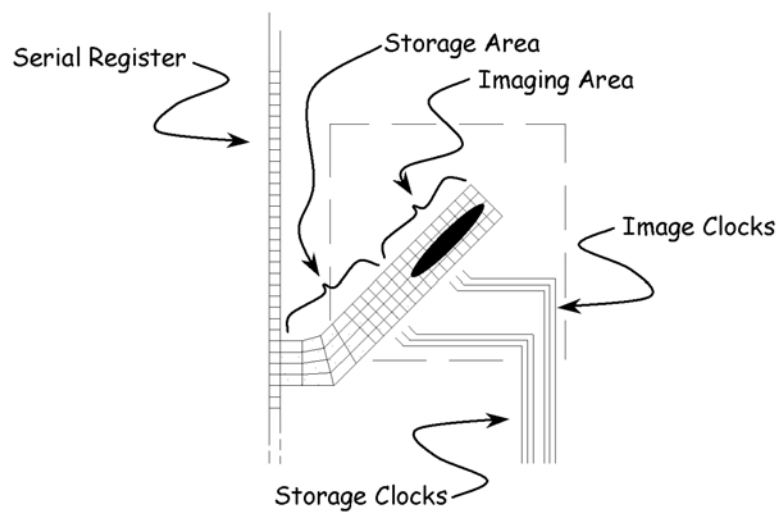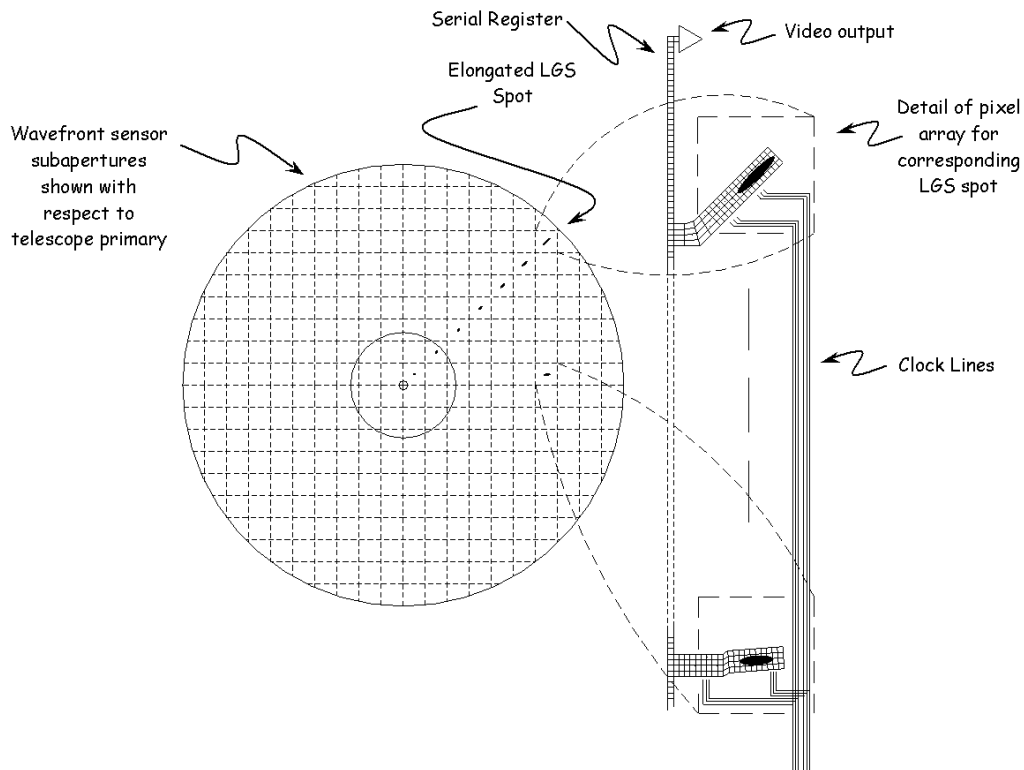
Figure 4- 10: Concept for WFS CCD customized for LGS spot elongation. Each subaperture presents a different size and orientation of LGS spot elongation. Each subaperture uses pixels that are aligned to the spot elongation for that subaperture. (from Beletic)

A novel solution is suggested here to use a custom lenslet array at the focal plane behind

the SH lenslets (figure 4-11). These image-plane lenslets are aligned along the local axis

of the elongation as shown in figure 4-12. The optical axis of each of the focal plane

lenslets can be chosen so that the pupil formed by each of the focal plane lenslets is

centered on a CCD pixel. The CCD pixels are arranged in a regular array. Thus, the

crucial component is the focal plane lenslet array, rather than the CCD. This may be

preferable since it is much easier to create a custom lenslet array (~$10,000) than a

custom CCD. Note that this approach bears some resemblance to lenslet-based PWFS's

in that the lenslets are used as "knife-edges" in the wavefront slope sensing. The

multicell diffraction considerations of section 3.9 apply here, but the task is much easier

since each lenslet need image only one subaperture (rather than a pupil full of

subapertures). The only requirement is that the image of the subaperture, including

diffraction effects, fits onto a CCD pixel; in other words, the subaperture images must not

have significant crosstalk between them. An example is shown in figure 4-13. The

example uses SH square lenslets, which are 500µ across and have a focal length of

10mm, and focal-plane lenslets that are 125µ by 250µ in size and have a focal length of

1mm. The CCD pixel is assumed to be 125µ across in the final pupil plane (or that relay

optics image the CCD pupil to 125µ across in the pupil plane). It is apparent in figure 4-

13 that the vast majority of the light is within the CCD boundary.

"SH" lenslets in
pupil plane

lenslets in
image plane

CCD pixels in
pupil plane
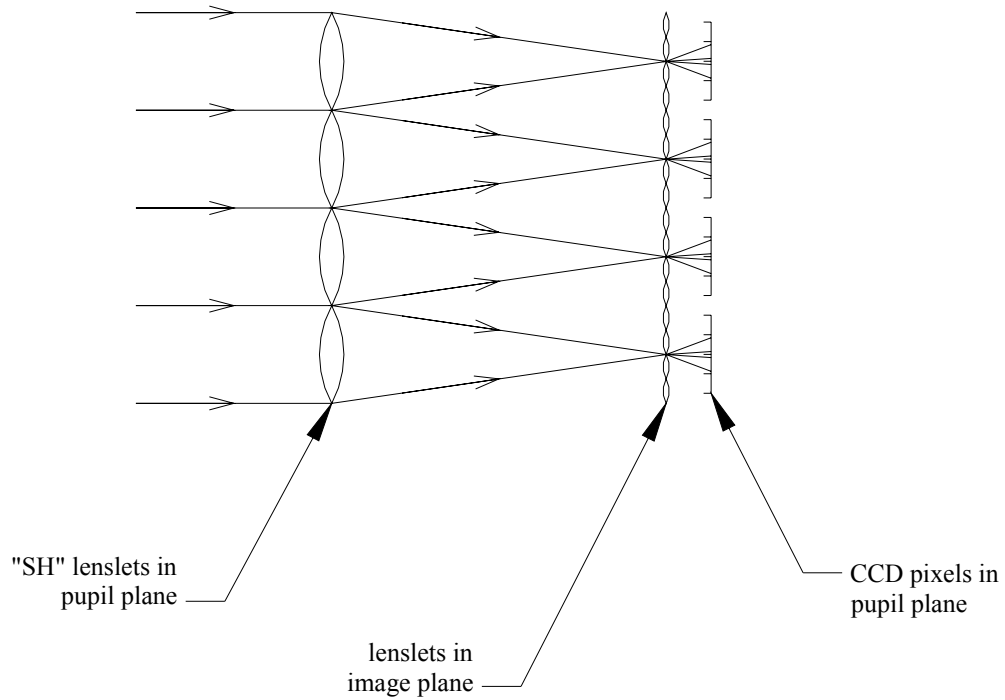
Figure 4- 11: SH WFS which uses a custom lenslet array to deal with LGS spot elongation. A nominally collimated beam is incident on a typical SH lenslet array. A second lenslet array is positioned at the lenslet focal plane; each lenslet images a pupil onto a single CCD pixel. The focal-plane lenslet array is customized to suit the LGS spot elongation for that subaperture, as shown in figure 4-12.
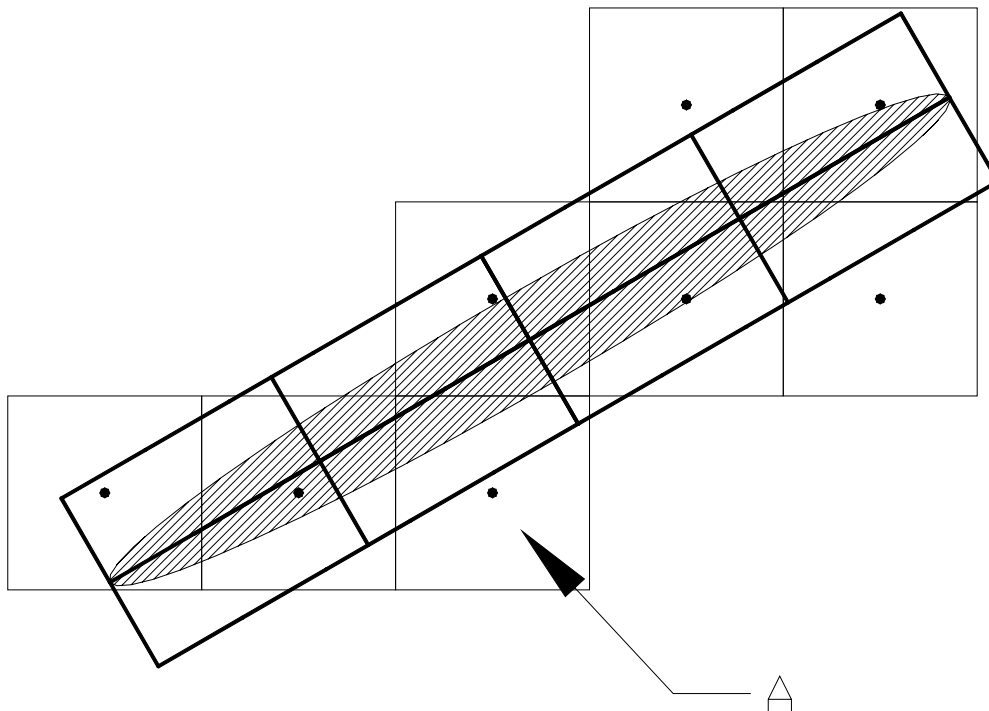
Figure 4- 12: Mapping of custom lenslets onto CCD pixels for an elongated spot from a typical subaperture in figure 4-6. The angled ellipse is the elongated spot for a typical subaperture; cf. elongated spots in figure 4-1. The heavy squares indicate the boundaries for each lenslet in the custom lenslet array. The optical axes of the lenslets are not in general in the center of the lenslet; rather, the optical axis (bullets) of each lenslet is placed in the center of an underlying CCD pixel (lightly-weighted squares). The CCD pixels are in a normal uniform grid array. In some cases, the optical axis of a lenslet may not be within the boundaries of the lenslet (e.g., the lenslet labeled "A", which has its optical axis just below the lenslet; the associated pixel is centered on this optical axis),

but this does not pose any difficulty. Note that the aspect ratio of the lenslets can be changed according to the amount of elongation for the particular subaperture.
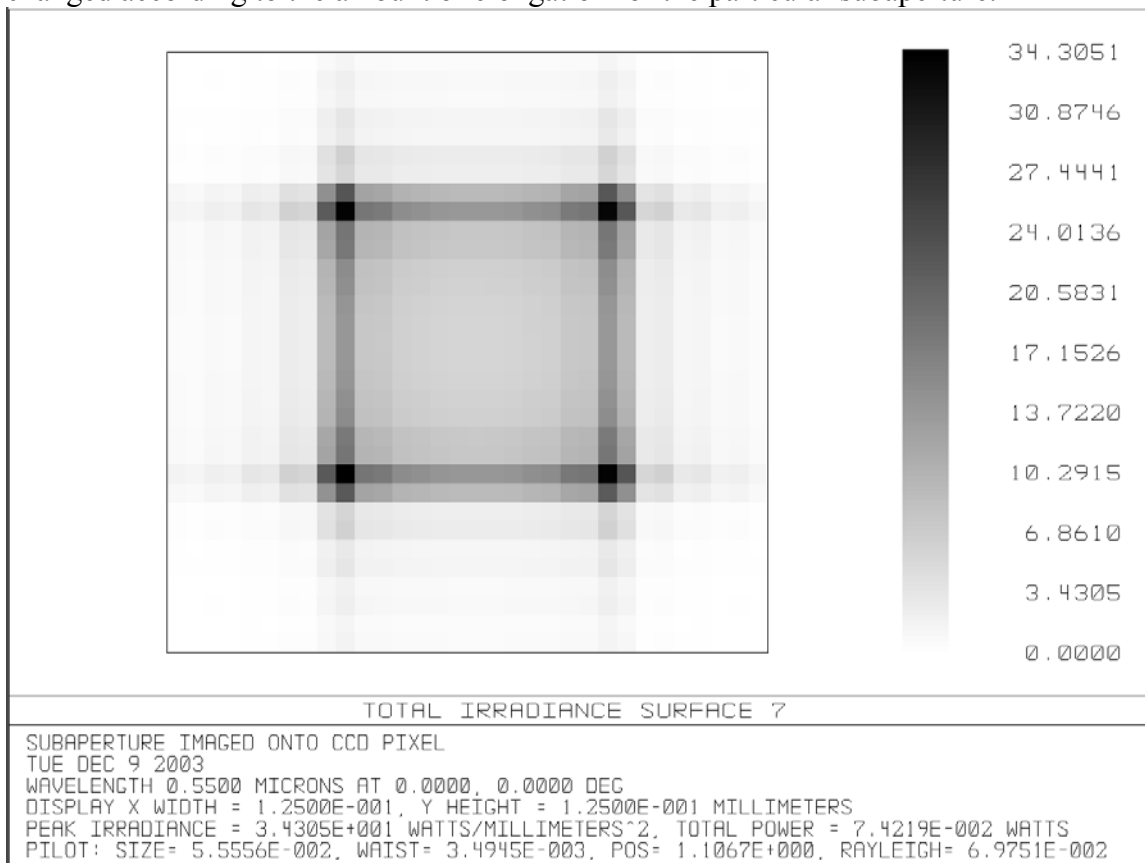
Figure 4- 13: Example of image of subaperture on a CCD pixel. The box around the diffracted spot pattern represents one CCD pixel.

4.6   Conclusions

Laser guide star elongation poses an important challenge to AO systems on ELT's. Two methods have previously been proposed: dynamic refocusing using a discrete mirror and custom CCD's. The two methods presented here, segmented MEMS (for use with pulsed

LGS's) and custom lenslet arrays (for use with CW LGS's) offer alternatives that are appealing in their simplicity, cost, and ease of future field-testing.

5. FURTHER WORK

The optical engineering required for an ELT AO system is a very large task, so there is no shortage of future work. The list below will be limited to those areas immediately related to this work, as opposed to the list of all optical engineering that needs to be done for ELT AO systems.

The y,y design method

The $y,\bar{y}$ method has been shown to be a powerful tool in designing MCAO systems and in considering the impact of various constraints on the design. In Chapter 2, the $y,\bar{y}$ method was applied to the case of relatively small DM's; large DM's such as those using "adaptive secondary" technology, were not considered. The $y,\bar{y}$ method can be applied fruitfully to cases using adaptive secondaries or other DM's. In any case, using the $y,\bar{y}$ design technique will be a useful tool in navigating the design space of MCAO systems and in feeding back to the project team what changes in requirements or in DM technology will have the greatest impact on the optical design..

Pyramid wavefront sensors

A number of extensions to existing PWFS techniques have been suggested: the lenslet-base PWFS, the multicell PWFS, and the multicell lenslet-based PWFS. Using lenslet

arrays is a promising approach to PWFS's. The chief question is whether the lenslet arrays are (or can be) fabricated accurately enough, particularly with respect to the edges where the "knife-edge" test is performed.

Such a test can be relatively straightforward. One can assemble a PWFS and test apparatus fairly easily; even from an existing SH WFS, as described in section 3.7. A series of tests is given below:

- Tip/tilt: Starting with a beam with very low aberrations, scan the beam around the center of the quad-cell of the PWFS. Do the data from the PWFS agree with the theoretical values?

- Low-order aberrations: Introduce known amounts of low-order aberrations such as focus or coma. Measure wavefront with PWFS for various tip/tilt amounts. Do the data from the PWFS agree with the theoretical values?

- Atmosphere-like aberrations: Introduce atmosphere-like, but known aberrations generated by phase plates or deformable mirror. Perform tests as described above.

These same tests can also be performed on multi-cell PWFS's. Of course, an on-sky test would be desirable, eventually. The method described in section 3.7 can be used to quickly create a PWFS on a telescope AO system.

Other areas for work include a comparison of the performance between multicell PWFS's and scanned conventional PWFS's, an examination of the relationship between multicell PWFS's and the wire test, and a complete analysis of the tolerances for a lenslet-based PWFS.

LGS spot elongation

Two approaches for dealing with LGS spot elongation have been proposed: segmented MEMS and custom lenslet arrays. Both approaches can be prototyped fairly easily and tested in the lab and on the sky. Such a test would likely use fewer subapertures than in specification table in Chapter 4 since large actuator count MEMS are not yet available. The testing program used by Lloyd-Hart for the discrete resonating mirror can be applied to these proposed techniques.

In addition, the proposed methods could be applied to Rayleigh guide stars. The elongation problem is much harder for Rayleigh guide stars, so it is not clear that either of these approaches will work here.

APPENDIX: DEFORMABLE MIRROR PACKAGING

CONSIDERATIONS

For an MCAO system where the correction is done in the wrong order (from high altitudes to low altitudes, as explained in Chapter 2), the DM's will likely be placed in the same optical space. There is a question of how closely we will be able to package the DM's, and this can have implications for the size of DM's that will be chosen.

In general, the longitudinal magnification equals the square of the lateral magnification, which is in turn proportional to the pupil size in the AO relay. As the pupil size decreases, the DM's become closer together. Since the angle of incidence on the DM's is approximately (diameter of DM)/(distance between DM's), the angle of incidence will increase as the pupil size decreases. Of course, as the angle of incidence increases, the DM elongates in one axis, and so eventually, the DM will start increasing in size even as the pupil size decreases. Eventually, as the pupil decreases, the DM's become so close together that it is not possible to package the DM's together.

The following shows triads of 3 DM's (0km, 4km, and 8km) using different sizes of the pupil: 225mm, 300mm, and 350mm. The minimum size of the DM long axis is found at a pupil size of ~300mm.
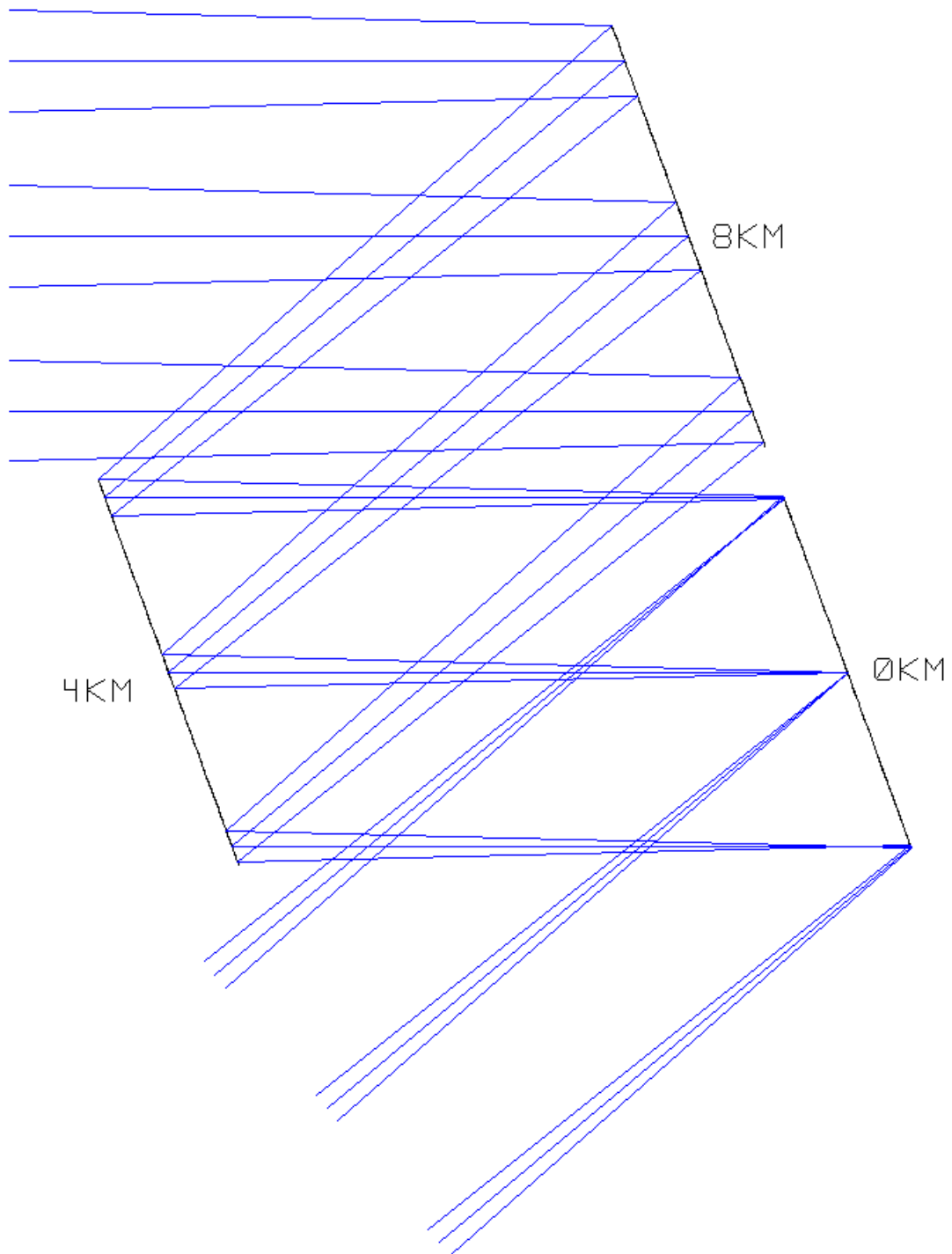
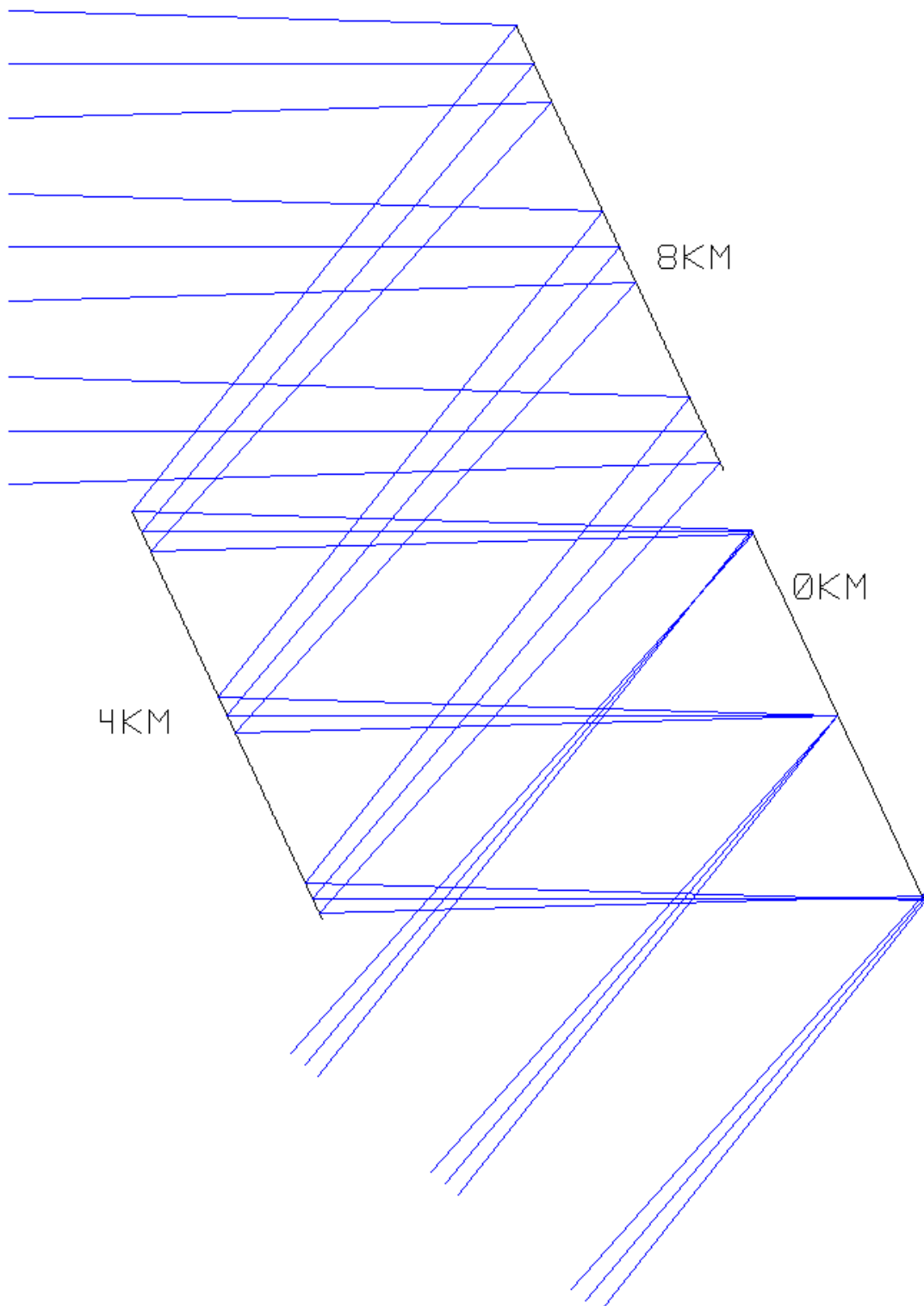Figure A- 1: DM's at 8km, 4km, and 0km with a pupil size of 350mm.

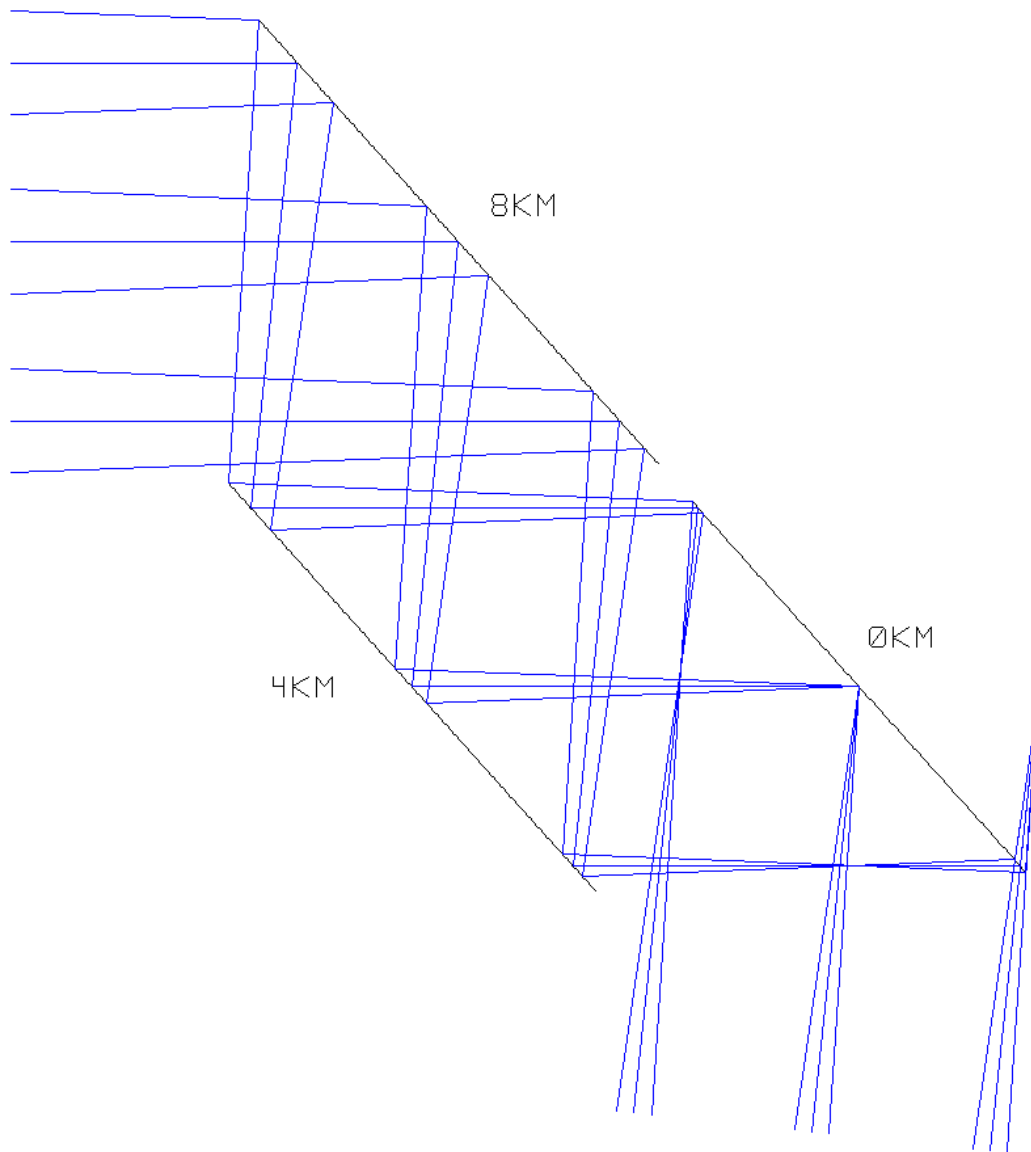Figure A- 2: DM's at 8km, 4km, and 0km with a pupil size of 300mm.

Figure A- 3: DM's at 8km, 4km, and 0km with a pupil size of 250mm.

REFERENCES

Angel, J. R. P., "Ground-based imaging of extrasolar planets using adaptive optics", Nature, vol. 368, p.203-207 (1994).

Angel, J.R.P., "Dynamic refocus for laser beacons", in *Science with LBT*, p.21-25 (Max-Planck-Institut für Astronomie, 2000).

Babcock, H., "The Possibility of Compensating Atmospheric Seeing", Publications of the Astronomical Society of the Pacific, vol. 65, p.229-236 (1953).

Barakat, R., "General Diffraction Theory of Optical Aberration Tests, from the Point of View of Spatial Filtering", Journal of the Optical Society of America, vol. 59, p.1432-1439 (1969).

Bauman, B., et al., "New optical design of adaptive optics system at Lick Observatory", in *Proceedings of SPIE*, vol. 3762, p.194-200 (1999).

Bauman, B., "First-Order Optical Design with the $y,\bar{y}$ Technique", in *Center for Adaptive Optics Proceedings: Summer School on Adaptive Optics 2000* (in preparation).

Beckers, J.M., "Increasing the size of the isoplanatic patch with Multi-Conjugate adaptive optics", in *Proceedings of the ESO Conference on Very Large Telescopes and their instrumentation*, p.693-703 (1988).

Beletic, J., et al., "Development of the Next Generation Optical Detectors for Wavefront Sensing", proposal to The Adaptive Optics Development Program at National Optical Astronomical Observatory, October 3, 2003 (unpublished).

California Association for Research in Astronomy, "Keck Observatory Report No. 208: Adaptive Optics for Keck Observatory" (Keck Observatory, Hawaii, January 1996).

Cornelissen, S of Boston Micromachines, personal communication, November 11, 2003.

Delano, E., "First-Order Design and the $y, \bar{y}$ Diagram", Applied Optics, vol. 12, p.1251-1256 (1963).

Diolaiti, E., et al, "Some novel concepts in multipyramid wavefront sensing", in *Proceedings of SPIE,* vol. 4839, p.299-306.

Ditteon, R., *Modern Geometrical Optics* (Wiley, New York, 1998).

Doble, N. of Iris AO, personal communication, November 11, 2003.

Ellerbroek B., et al., "Multiconjugate adaptive optics for Gemini-South", in *Proceedings of SPIE*, vol. 4839, p.55-66 (2003).

Esposito, S. and Riccardi, A., "Pyramid Wavefront Sensor behavior in partial correction Adaptive Optics systems", Astronomy & Astrophysics, 369, L9-L12 (2001).

Feeney, O., "Theory and Laboratory Characterization of Novel Wavefront Sensor for Adaptive Optics Systems", Ph.D. dissertation, National University of Ireland, Galway, Ireland (2001).

Flicker, R., "Sequence of phase correction in multiconjugate adaptive optics", Optics Letters, vol. 26, p.1743-1745 (2001).

Foucault, M.L., "Description of Various Processes mad use of for Finding out the Configuration of Optical Surfaces", Monthly Notices of the Royal Astronomical Society, vol. 19, p.284-287 (1859).

Foy, R., and Labeyrie, A., "Feasibility of adaptive telescope with laser probe", Astronomy & Astrophysics, vol. 152, p.L29-L31 (1985).

Fried, D., "Optical Resolution through an Inhomogeneous Medium for Very Long and Very Short Exposures", Journal of the Optical Society of America, vol. 56, p.1372-1379 (1966)

Fried, D., "Anisoplanaticism in Adaptive Optics", Journal of the Optical Society of America, vol. 72, p.52-61 (1982).

Fried, D., and Besher, J., "Analysis of fundamental limits to artificial guide star adaptive optics system performances for astronomical imaging", Journal of the Optical Society of America A, vol. 11, p.277-287.

Gaskill, J., *Linear Systems, Fourier Transforms, and Optics* (Wiley, New York, 1978).

Gavel, D., personal communication.

Gaviola, E., "On the quantitative use of the Foucault knife-edge test", Journal of the Optical Society of America, vol. 26, p.163-169 (1936).

Georges III, J., Mallik, P., Stalcup, T., Angel, R, Sarlot, R., "Design and Testing of a Dynamic Refocus System for Rayleigh Laser Beacons", in *Proceedings of SPIE*, vol. 4839, p.473-483 (2003).

Ghigo, M., Crimi, G., Perennes, F., "Construction of a pyramid wavefront sensor for adaptive optics compensation", in *Venice 2001 Beyond Conventional Adaptive Optics* (Venice, Italy, 2001).

Ghigo, M., Perennes, F., Ragazzoni, R., "Manufacturing by deep x-ray lithography of pyramid wavefront sensors for astronomical adaptive optics", in *Proceedings of SPIE*, vol. 4839, p.259-265 (2003).

Goncharov, A., Ph.D. dissertation, Lund University, Sweden (2003).

Hardy, *Adaptive Optics for Astronomical Telescopes* (Oxford University Press, 1998).

Hofer, H., Artal, P., Singer, B., Aragon, J.L. & Williams, D.R., "Dynamics of the eye's wave aberration", Journal of the Optical Society of America A, vol. 18 , p.497-506. (2001).

Horwitz, B., "A spatial filter for phase object differentiation", Optics Communications, vol. 17, p.231-233 (1976).

Horwitz, B., "A new, pupil plane wavefront gradient sensor", in *Proceedings of SPIE*, vol. 2201, p.495-501 (1994).

Katzoff, S., *Quantitative Determination of Optical Imperfections by Mathematical Analysis of the Foucault Knife-Edge Test Pattern*, NASA Technical Note D-6119 (August 1971).

Linfoot, E.H., *Recent Advances in Optics* (Oxford University Press, New York, 1958).

Lloyd, J.M., *Thermal Imaging Systems* (Plenum Publishers, New York, 1975).

Lloyd-Hart M., personal communication, August 18, 2003.

Lloyd-Hart, M., Georges III, J., Angel, J.R.P., Brusa, G., Young, P., "Dynamically refocused Rayleigh laser beacons for atmospheric tomography", in *Proceedings of SPIE*, vol. 4494, p.259-270 (2002).

Macintosh, B., et al., "Extreme Adaptive Optics Planet Imager: XAOPI", in *Proceedings of SPIE*, vol. 5170 (in press).

Malacara, D., *Optical Shop Testing*, 2nd edition (Wiley, New York, 1992), p.265-320.

Max, C.E., Avicola, K., Bissinger, H., Brase, J., Gavel, D., Friedman, H., Morris, J., Olivier, Salmon, J., and Waltjen, K., "Feasibility experiment for sodium-layer guide stars

at Lawrenece Livermore National Laboratory", in *Laser Guide Star Adaptive Optics Workshop: Proceedings*, p.535-541 (1992).

Nelson, J., personal communication.

Pennington, D., et al., "Compact fiber laser approach to 589 nm laser guide stars" in *Technical Digest, European Conference on Lasers and Electro-Optics 2003*

Perreault, J., Bifano, T., Levine, B.M., Horenstein, M., "Adaptive optic correction using microelectromechanical deformable mirrors", Optical Engineering, vol. 41, p.561-566 (2002).

Poyneer, L. (2003a), "Scene-based Shack-Hartmann wave-front sensing: analysis and simulation", Applied Optics, vol. 42, p.5807-5815 (Oct 10, 2003).

Poyneer, L. (2003b), LaFortune, K., Awwal, A., "Correlation wave-front sensing for Shack-Hartmann-based Adaptive Optics using a point source", Lawrence Livermore National Laboratory Report UCRLJC 152975 (Lawrence Livermore National Laboratory, 2003).

Ragazzoni, R. (2000a), Marchetti, E., Valento, G., "Adaptive-optics corrections available for the whole sky", Nature, vol. 403, no.6765, p.54-6 (Jan. 6, 2000).

Ragazzoni, R., "Pupil plane wavefront sensing with an oscillating prism", Journal of Modern Optics, vol. 43, p.289-293 (1996).

Ragazzoni, R., Diolaiti, E., Vernet, E., "A pyramid wavefront sensor with no dynamic modulation", Optics Communications, vol. 208, p.51-60 (2002).

Ragazzoni R. (2000b), Farinato, J., and Marchetti, E., "Adaptive optics for 100-m-class telescopes: new challenges require new solutions", in *Proceedings of SPIE*, vol. 4007, 1076 (2000).

Ragazzoni, R., et al., "The pyramid wavefront sensor aboard AdOpt @ TNG and beyond: a status report", in *Proceedings of SPIE*, vol. 4494, p.181-187 (2002).

Shack R.V., Platt B.C., "Production and use of a lenticular Hartmann screen", in *Programme of the 1971 spring meeting of the Optical Society of America,* abstracts only received (Washington, DC, 1971), p.15.

Shack, R.V., "Analytic system design with pencil and ruler: the advantages of the y-y diagram", in *Proceedings of SPIE*, vol. 39, p.127-140 (1973).

Shack, R.V., class notes for "Aberration Theory" course at University of Arizona Optical Sciences Center, 1991 (unpublished).

Sprague, R., Thompson, B., "Quantitative Visualization of Large Variation Phase Objects", Applied Optics, vol. 11, p.1469-1479 (1972).

University of California and California Institute of Technology, *California Extremely Large Telescope: Conceptual Design for a Thirty-Meter Telescope* (Center for Adaptive Optics, University of California Santa Cruz, June 2002).

Verinaud, C., Arcidiacono, C., Carbillet, M., Diolaiti, E., Ragazzoni, R., Vernet-Viard, E., Espostio, E., "Layer Oriented Multi-Conjugate Adaptive Optics systems: Performance analysis by numerical simulations", in *Proceedings of SPIE*, vol. 4839, p.524-535 (2002).

van Dam, Marcos and Le Mignant, David, "Investigation into the size of the Shack-Hartmann spots", Keck Adaptive Optics Note No. 226 (Keck Observatory, Hawaii, 2003).

Williams; D.R., Vaughn, W.J., Singer; B.D., Hofer, H., Yoon, G.-Y., U.S. Patent # 6,299,311 (October 9, 2001).

Wilson, R. Gale, "Wavefront-error evaluation by mathematical analysis of experimental Foucault-test data", Applied Optics, vol. 14, p.2286-2297 (1975).