OPTIMAL SHAPED PUPILS AND WAVEFRONT CONTROL FOR PLANET FINDING CORONAGRAPHY

N.J. Kasdin[1], M.G. Littman[1], A. Giveon[1], L. Pueyo[1], R.J. Vanderbei[2], D.N. Spergel[3], and M. Carr[3]

[1]Mechanical and Aerospace Engineering, Princeton University, Princeton, NJ, 08540, email: jkasdin@princeton.edu
[2]Operations Research, Princeton University, Princeton, NJ, 08540
[3]Astrophysics, Princeton University, Princeton, NJ, 08540

ABSTRACT

Motivated by the desire to image exosolar planets, recent work by us and others has shown that high-contrast imaging can be achieved using specially shaped pupil masks. The objective of a shaped pupil coronagraph is to design an aperture that results in a system point spread function (PSF) with the needed contrast allowing planet discovery at the smallest inner working distance (IWD) in the shortest integration time. To do this, we optimize throughput with contrast and IWD constraints. In this paper we summarize the various optimal shaped pupils we have created to date, comparing their performance. We also present preliminary results on stochastic wavefront estimation and control algorithms, an essential capability for any coronagraphic planet finding system.

1. INTRODUCTION

The key to direct detection and eventual spectroscopic characterization of extrasolar terrestrial planets is high-contrast imaging. Without employing novel approaches to suppress the bright diffraction rings of the main star it will be impossible to image a planet that is in the habitable zone of a nearby star. For instance, a planet at 10 pc that is 1 AU from its star appears separated from the star by 0.1 arcsecond when viewed from Earth. For a 5 m visible-light telescope, 0.1 arcseconds corresponds to about $4.8\lambda/D$—which is roughly the location of the fourth diffraction ring in a conventional telescope. The fourth diffraction ring is generally about 10^{-3} times as bright as the central Airy disk for the star—about seven orders of magnitude brighter than a terrestrial planet, which is expected to be only about 10^{-10} times as bright as the star. The imaging task is to reduce this "halo" to below the level of the planet.

For the past few years, we have been studying novel approaches to high-contrast imaging for NASA's Terrestrial Planet Finder (TPF). Only five years ago the consensus seemed to be that the only viable approach to imaging extrasolar planets was via a nulling interferometer in the infrared. Remarkably, the last few years have seen a tremendous outpouring of ideas for achieving high contrast in the visible, primarily via some form of coronagraphy. As a result, coronagraphy is now recognized as a viable alternative to nulling interferometry. Kasdin et al. (2003); Vanderbei et al. (2003a,b) provide a summary of our accomplishments in designing and analyzing apodized approaches to high-contrast imaging. In this paper, we present a brief summary of some of our pupil designs, showing how they achieve the needed contrast with reasonable throughput (and thus, integration time).

Pupil masking (or any apodization) provides the needed contrast in an ideal optical system. It has long been recognized, however, that performance of a visible light coronagraph will be limited by the halo of scattered light due to imperfections in the optics. Perhaps the most critical technology for TPF, therefore, is a wavefront sensing and control system to reduce this scatter to below the level of the planet. We also present a brief summary, therefore, of our progress in developing a stochastic phase retrieval and control system for use in a shaped pupil coronagraph.

2. PUPIL APODIZATION

It has long been known that pupil apodization provides a means to modify the Point Spread Function of a telescope. Jacquinot & Roizen-Dossier (1964) provides a particularly thorough summary of various apodization functions. Recently, Nisenson & Papaliolios (2001) suggested an apodized square aperture for TPF. The main advantage of a square aperture is it decouples the electric field into two one-dimensional Fourier transforms of the linear apodization function, though at the expense of either throughput or discovery space. Slepian (1965) showed that the *prolate spheroidal wavefunction* is the "optimal" apodization for a one-dimensional square aperture. Slepian also showed that the *generalized prolate spheroidal wavefunction* solves the circularly symmetric optimal apodization problem. That is, for a given inner working angle, ρ_{iwa}, it is the apodization that maximizes the following integral cost:

$$\text{maximize } \frac{\int_0^{\rho_{iwa}} E(\rho)^2 \rho d\rho}{\int_0^{\infty} E(\rho)^2 \rho d\rho} \tag{1}$$

Proc. "Towards Other Earths: DARWIN/TPF and the Search for Extrasolar Terrestrial Planets", Heidelberg, Germany, 22-25 April 2003 (ESA SP-539, October 2003)

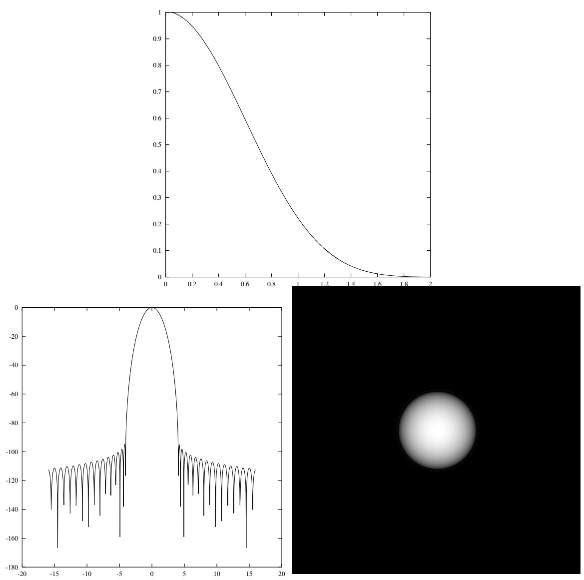

Figure 1. Top. *The generalized prolate spheroidal apodization.* Bottom. *The corresponding point-spread function. Note that high-contrast (*10^{-10} *extends from* $\rho_{iwa} = 4$ *to infinity.*

It turns out that $\rho_{iwa} = 4\lambda/D$ produces the needed contrast of 10^{-10}. Figure 1 shows a radial cut of the apodization and the resulting PSF.

3. OPTIMAL SHAPED PUPILS

While apodized pupils can, in principal, be used to ahieve the desired high contrast, they suffer from severe manufacturing difficulties. They also tend to have very low throughput. As an alternative, we have proposed using shaped pupils to taylor the PSF. In Kasdin et al. (2003) we compare the performance of a number of apodized approaches and contrast with that of a shaped pupil coronagraph.

Perhaps the simplest shaped pupil coronagraph, and the first we examined, uses Slepian's prolate spheroidal wavefunction. Figure 2 shows a single-opening shaped pupil whose outline is defined by the zero-order prolate spheroidal wavefunction. It can be shown that on axis this pupil produces the same optimal high contrast. That is, it reaches the needed contrast of 10^{-10} everywhere along the x-axis except within 4 λ/D of the center of the Airy disk.

It is clear from Figure 2 that the disadvantage of this mask is the extreme narrowness of the high-contrast region at close working distances from the central star. To open up the high-contrast discovery zone, we next considered using multiple-opening pupil masks. Several such masks are presented in Kasdin et al. (2003). These masks are computed as solutions to certain nonlinear optimization problems in which one maximizes some measure of throughput, such as open area of the mask, subject to constraints that ensure sufficiently high contrast throughout

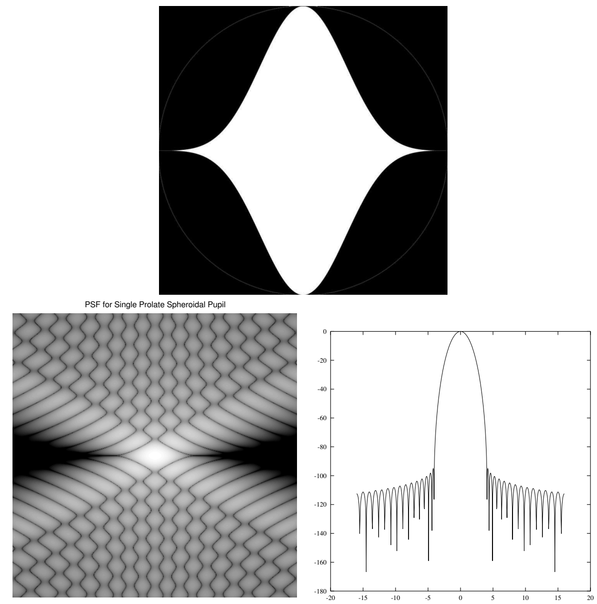

Figure 2. Top. *The Kasdin-Spergel prolate-spheroidal mask.* Bottom Left. *The associated* 2*-D point spread function and* Bottom Right. *its* x*-axis slice shown in decibels (*$10^{-10} = -100$*dB).*

a specified discovery zone. These optimization problems are solved using our own software, LOQO, which is described in Vanderbei (1999) (see also Vanderbei (2001)). One of the best masks found in this way is shown in Figure 3. With this mask it is possible to search for extrasolar

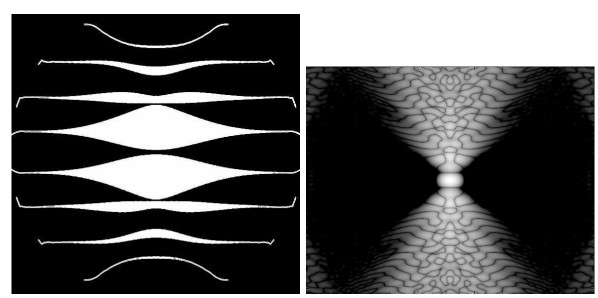

Figure 3. Left. *A mask with 8 openings.* Right. *The associated* 2*-D point spread function.*

planets as close as $4\lambda/D$ using only two integrations.

Of course, for discovery, it would be even better to have a pupil mask that provides high contrast simultaneously in all directions around the central lobe. To this end, we modified the optimization problem to look for circularly symmetric masks. Such masks consist simply of concentric rings. One such mask is shown in Figure 4. The natural question arises as to how to support these rings. As is well known, using a small number of support spiders creates strong diffraction spikes that will infiltrate the dark zone. We show in Vanderbei et al. (2003a), however, that a large number of such spiders has the desirable property that the spikes, though commensurately larger in number, also move away from the central lobe leaving the dark discovery zone essentially intact. Using 150 thin spiders,

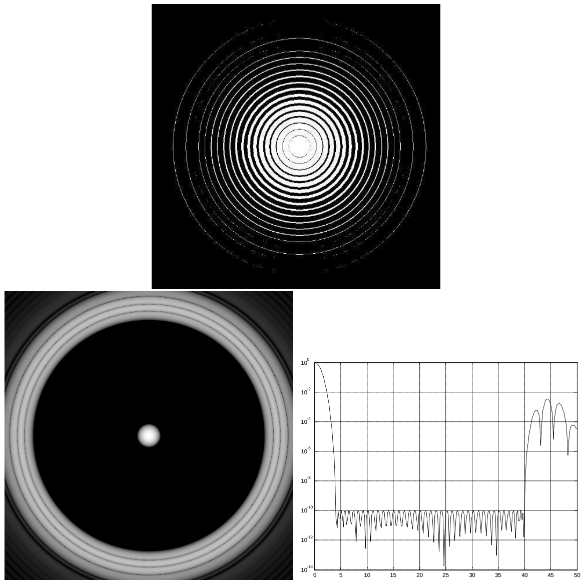

Figure 4. Top. *A concentric ring mask.* Bottom Left. *The associated* 2*-D point spread function and* Bottom Right. *a radial cross section.*

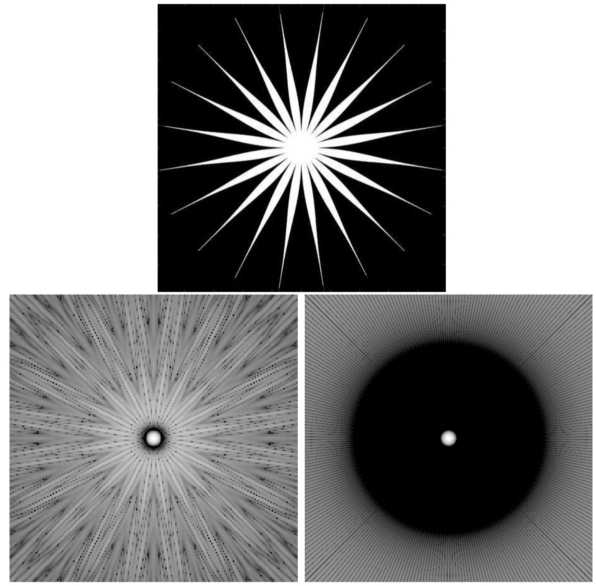

Figure 5. Top. *A* 20*-point starshape mask.* Bottom Left. *The associated* 2*-D point spread function for the* 20 *point mask.* Bottom Right. *The associated* 2*-D point spread function for the analogous* 150 *point mask.*

for example, it is possible to preserve the 10^{-10} level of contrast throughout the design dark zone. We call such masks *spiderweb masks*.

We have also looked at another mask design, which we call *starshape masks*, that also provides single-image discovery capability. With these starshape masks, we dispense with the concentric rings and let the shape of the spider vanes themselves provide the desired diffraction control. The computational optimizations that lead to these masks are detailed in Vanderbei et al. (2003b). One such mask is shown in Figure 5. As with spiderweb masks, a large number of star-points are needed to push out the diffraction spikes. The 20-point mask is shown just for clarity; a 150 point mask is needed for terrestrial planet discovery. The figure shows the point spread function for such a many pointed mask.

Since our techniques for optimizing masks are very mature, we have begun exploring methods for manufacturing and testing pupil masks. Figure 6 shows a photograph of a sample pupil mask we recently had manufactured by Max Levy Autograph in Philadelphia, PA of an earlier multipupil deisgn. The pupil is made using a nickel electroforming process where photolithography is used to create a mandrel out of photoresist and then nickel is grown on the glass substrate in a two layer process. The nickel mask is then removed, resulting in a free-standing metal mask with no substrate. We have made three masks, of different thicknesses and hope to begin testing them in the laboratory soon.

Also in Figure 6 is a photograph of one edge of the mask taken from a scanning electron microscope. Our expectations were that the edges could be manufactured with better than 1 micron tolerance, more than adequate for a 10^{-10} contrast. We note from the image, however, that the mask edge has a cupped and scalloped residual.

Figure 6. Left *A four opening pupil mask. This mask was manufactured using electroforming techniques. It is a 2 by 4 binary mask made out of nickel. Edge accuracy is approximately 1-2 microns.* Right *A scanning electron microscope image of the mask edge, showing details of the manufacturing process and errors.*

This is due to imperfections where the nickel meets the photoresist. We are currently investigating modifications to the process to generate improved edges or additional steps, such as electropolishing, to clean the edges.

4. WAVEFRONT CONTROL

The most critical technology for a coronagraph implementation of TPF is wavefront sensing and control. The scatter and ghosts due to imperfections in the mirrors or distortions on orbit must be corrected to below the level of the planet. Scattered light from ripples in the mirrors or amplitude changes due to reflectivity variations can be orders of magnitude larger than the planet in the mid-spatial frequency range where discovery takes place. Wavefront accuracy on the order of $\lambda/10,000$ is necessary to reduce the "halo" to acceptable levels.

A visible light TPF must be equipped with an active wavefront control system. Most advocates of a visible light TPF assume the utilization of a deformable mirror at a reimaged pupil (perhaps just following the shaped aperture mask) to correct the light. A variety of wavefront sensing algorithms are being examined for determining the needed correction (e.g., Redding et al. (2002); Shaklan et al. (2002); Trauger et al. (2002)). Piezo-actuator based deformable mirror technology has progressed far at JPL and other locations such that the needed stability and precision has been demonstrated in the laboratory (Trauger et al. (2002)).

While there is a long history of using DMs for wavefront control in the adaptive optics (AO) community, the problem of correcting phase and amplitude for TPF is unique. In particular, rather than trying to correct focus, strehl, and other low order aberrations, our primary concern is the mid-spatial frequencies. Bandwidths are also orders of magnitude lower because of the absence of atmospheric distortion, providing time for more complex control algorithms. Finally, it is expected that sensing will take place in the image plane rather than a diverted pupil plane to eliminate non-common path errors, introducing more complexity to the estimation algorithm.

Conceptually, there are two parts of the wavefront control problem—estimating the phase error, and thus the corresponding mirror errors (*phase retrieval*), and determining the best deformable mirror shape to correct the resulting aberrations in the image plane. In the remaining sections we describe our efforts to develop optimal algorithms for determining phase and amplitude and to implement them in a correction system.

4.1. Phase Estimation

The goal of the phase retrieval algorithm is to estimate the phase aberration, ϕ, given image plane measurements of only the intensity. Because we are searching for minimum integration time algorithms, we assume very small photon counts and thus a Poisson distribution for the arriving light. This leads us to stochastic algorithms for determining phase. In addition, because the detector measures only the square of the electric field, multiple, diverse measurements are required to infer the phase and eliminate ambiguities. Some commonly used diversity methods are *focus*, where the position of the image plane is moved, and *DM*, where the actuators on the DM are moved in a predictable way. We are investigating *pupil diversity* algorithms. By pupil diversity we mean a change in the shape of the pupil among multiple measurements to provide the added degrees of freedom to determine phase. The simplest form of pupil divesity is a scaling of the pupil function.

Given and $N \times N$ CCD array, the relative arrival rate, $\lambda_m(\xi, \eta)$, of photons as a function of position in the image plane (ξ, η) for the m-th pupil function is given by the square of the Fourier Transform of the pupil apodization, $A_m(x, y)$:

$$\lambda_m(\xi, \eta) \propto \left| \mathcal{F} \left\{ A_m(x, y) e^{i\phi(x,y)} \right\} \right|^2 \qquad (2)$$

We then take a maximum likelihood approach to estimating the phase based on a Poisson model of this photon density. If z_{ij} is the photon count on the (i, j) pixel, then the best estimate of the phase function is given by the optimization problem:

$$\max_{\phi(x,y)} \left\{ \sum_m \sum_{i=1}^{N} \sum_{j=1}^{N} \left[-(\lambda_{ij})_m + (z_{ij})_m \log((\lambda_{ij})_m) \right] \right\} \qquad (3)$$

The optimization in Eq. 3 is applied in two steps. For the first step, we use a genetic algorithm to perform a global search because of the large number of parameters necessary to approximate the signal and due to the unconstrained nature of the problem. In the second step we locally optimize the cost function using a BFGS Quasi-Newton method with a mixed quadratic and cubic line search procedure. The resulting phase estimate is then used in a wavefront correction algorithm.

4.2. Correction Algorithm

Given an estimate of the phase aberration, $\phi_{est}(x, y)$, the goal of the correction algorithm is to determine the feasible shape of the deformable mirror, $\phi_{DM}(x, y)$ such that the point spread function of the corrected system,

$$I_{cor}(\xi, \eta) = \left| \mathcal{F} \left\{ A(x, y) e^{i\phi(x,y)} e^{-i\phi_{DM}(x,y)} \right\} \right|^2 \qquad (4)$$

will be as close as possible to the point spread function of the system without aberrations (or, alternatively, as dark as possible in a specified region). We do this by incorporating a model of the deformable mirror shape given actuator commands and a model for the PSF of the optical system including the DM. For the work to date, we have assumed we can model the DM via superposition of Gaussian influence functions.

The correction algorithm consists of three steps. The first step is to use the phase estimate from the phase retrieval algorithm above as an initial guess for the actuator settings (this is not the "optimal" setting as the uncorrected higher frequency phase error can intermodulate and produce beat frequencies in the region of the image we are interested in). The second step is to find the feasible deformable mirror shape that minimizes the square of the difference between the deformable mirror shape and the estimated phase aberration shape with the above initial guess. In other words, we minimize the cost:

$$J_2 = \min_{h_{mn}} \left\{ \sum_x \sum_y \left[\phi_{DM}(x, y) - \phi_{est}(x, y) \right]^2 \right\} \qquad (5)$$

where h_{mn} are the DM settings. This will produce a deformable mirror shape that is closest to the shape of the phase aberration model. However, the ultimate goal of the correction algorithm is the nulling of the point spread function. Therefore, the last step is to minimize the square of the difference between the corrected point spread function and the ideal point spread function within a defined region of the image plane. Both steps 2 and

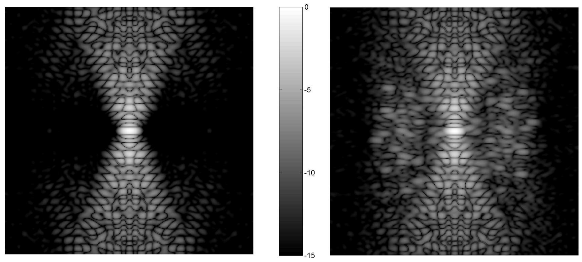

Figure 7. Left *The ideal point spread function.* Right *The aberrated point spread function.*

3 are accomplished using a Quasi-Newton method optimizer.

Figures 7 and 8 show a simulation of the correction algorithm. For this simulation, a 64×64 matrix was used to represent the pupil embedded in a larger 256×256 array. Figure 7 shows the ideal and aberrated PSF of the simulation. Figure 8 shows the corrected PSF after the first, second, and third steps of the algorithm. The DM model used for the simulation consisted of a 30×30 array of actuators.

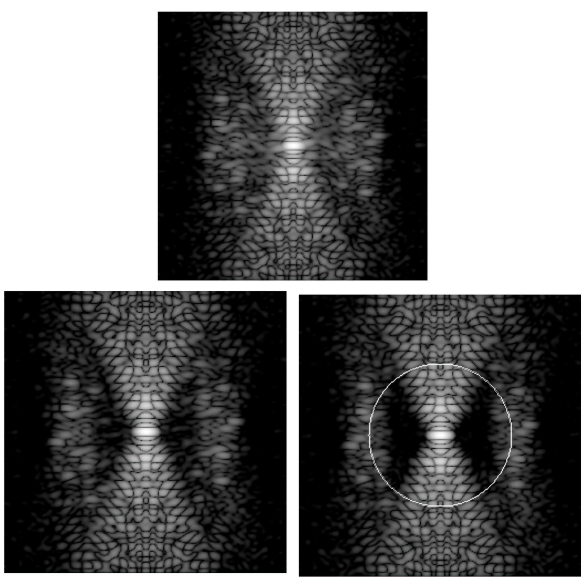

Figure 8. Top *The PSF of an aberrated system.* Bottom Left *The corrected PSF using an optimal phase estimate.* Bottom Right *The final corrected PSF after an optimal dark hole algorithm.*

4.3. Amplitude Control

The correction of phase error is routine in the adaptive optics community, but little attention has been given to amplitude control. For TPF, amplitude control is of equal importance. Amplitude (reflectivity) errors on the mirrors (or phase errors at conjugate planes) can introduce scatter of intensity equal to the phase errors. Amplitude

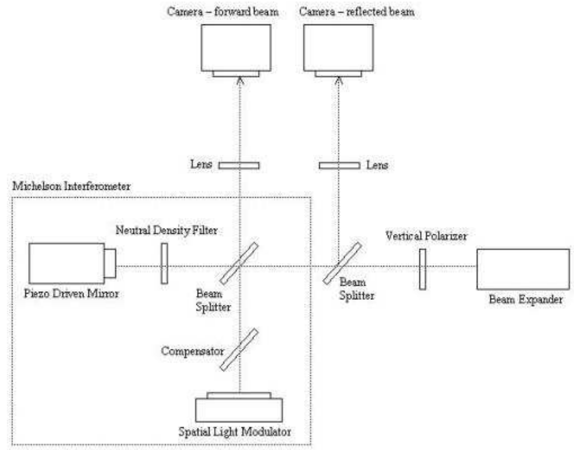

Figure 9. Experimental setup of a Zero-Path Difference Michelson Interferometer for phase and amplitude control.

correction is also particularly difficult to achieve polychromatically. In the original CODEX proposal, Brown and Burrows suggested a phase based amplitude correction utilizing only half the available image plane. In Littman et al. (2002), we presented a new concept for simultaneously correcting phase and amplitude using two deformable mirrors arranged in a Zero Path Difference Michelson interferometer (ZPDM). By adjusting the mirrors in common mode, phase distortions can be corrected while by adjusting the mirrors differentially, amplitude corrections can be made (with the excess light returning along the other leg of the interferometer).

Figure 9 displays the experimental configuration of our interferometer using a spatial light modulator in place of a deformable mirror. Figure 10 presents a demonstration of amplitude adjustment in the laboratory using white light from a black body source in a ZPDM. The main limitation of the Michelson approach is its sensitivity to bandwidth. Exact correction is achievable only at a single wavelength; however, the zero path difference makes it minimally sensitive to small variations in wavelength. We are currently studying methods for improving the frequency bandwidth. For instance, it may be possible to coat the mirrors with a dielectric having the proper phase delay vs. frequency to compensate for the frequency sensitivity of the ZPDM.

5. CONCLUSION

We, and others, have made great progress in the last three years studying and designing coronagraphs for terrestrial planet finding. We believe that shaped pupils hold the most promise for achieving the highest contrast at the shortest integration time and with the least sensitivity to errors. We believe that within a year we will have a final optimal pupil design for TPF.

That leaves wavefront control as the most critical technology for visible light planet finding. Progress is being

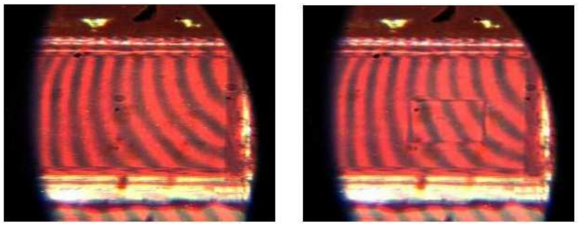

Figure 10. Demonstration of amplitude control using a white light source. By using two deformable mirrors arranged in a Michelson interferometer, both amplitude and phase errors can be corrected. Common mode adjustments of the two mirrors on each leg corrects phase, and differential adjustment corrects amplitude. In the figure, the left image is before adjustment and right image is after adjustment. (Littman et al. (2002))

made there as well, and shaped pupils provide a novel and unique method for providing the diversity needed in the phase retrieval algorithm. We have shown here preliminary results of a stochastic phase retrieval and wavefront control algorithm that demonstrates the feasibility of the needed control. We have also shown that amplitude control will be critical to achieving the needed high contrast and have presented a novel approach to simultaneous phase and amplitude correction with two DMs.

REFERENCES

Jacquinot P., Roizen-Dossier B., 1964, Progess in Optics, 3, 29

Kasdin N., Vanderbei R., Spergel D., Littman M., 2003, Astrophysical Journal, 582, 1147

Littman M., Carr M., Kasdin N., Spergel D., 2002, In: Proceedings of SPIE Conference on Astronomical Telescopes and Instrumentation, no. 12 in 4854

Nisenson P., Papaliolios C., 2001, The Astrophysical Journal, 548, L201

Redding D., Green J., Basinger S., et al., 2002, In: Proceedings of SPIE Conference on Astronomical Telescopes and Instrumentation, no. 40 in 4854

Shaklan S., Green J., Moody D., 2002, In: Proceedings of SPIE Conference on Astronomical Telescopes and Instrumentation, no. 42 in 4860

Slepian D., 1965, Journal of the Optical Society of America, 55, 1110

Trauger J., Moody D., Gordon B., Gursel Y., 2002, In: Proceedings of SPIE Conference on Astronomical Telescopes and Instrumentation, no. 01 in 4854

Vanderbei R., 1999, Optimization Methods and Software, 12, 485

Vanderbei R., 2001, Linear Programming: Foundations and Extensions, Kluwer Academic Publishers, 2nd edn.

Vanderbei R., Spergel D., Kasdin N., 2003a, Astrophysical Journal, to appear

Vanderbei R., Spergel D., Kasdin N., 2003b, Astrophysical Journal, submitted