Adaptive Nulling: a new enabling technology for interferometric exo-planet detection

Oliver P. Lay*, Muthu Jeganathan, Robert Peters
Jet Propulsion Laboratory, California Institute of Technology
4800 Oak Grove Drive, Pasadena CA 91109

ABSTRACT

Deep, stable nulling of starlight requires careful control of the amplitudes and phases of the beams that are being combined. The detection of earth-like planets using the interferometer architectures currently being considered for the Terrestrial Planet Finder mission require that the E-field amplitudes are balanced at the level of ~ 0.1%, and the phases are controlled at the level of 1 mrad (corresponding to ~ 1.5 nm for a wavelength of 10 μm). These conditions must be met simultaneously at all wavelengths across the science band, and for both polarization states, imposing unrealistic tolerances on the symmetry between the optical beamtrains. We introduce the concept of a compensator that is inserted into the beamtrain, which can adaptively correct for the mismatches across the spectrum, enabling deep nulls with realistic, imperfect optics. The design presented uses a deformable mirror to adjust the amplitude and phase of each beam as an arbitrary function of wavelength and polarization. A proof-of-concept experiment will be conducted at visible / near-IR wavelengths, followed by a system operating in the Mid-IR band.

Keywords: Nulling interferometry, planet detection

1. MOTIVATION

1.1. Nulling requirements

The detection of Earth-like planets around nearby stars at Mid-Infrared wavelengths requires that the light from the star be suppressed by a factor of 10^5 or more over the bandwidth of interest, currently 7 - 17 μm. The technique of nulling interferometry[1] has been proposed for both the European Darwin mission[2] and NASA's Terrestrial Planet Finder (TPF)[3].

For the case where light incident on the science detector is first passed through a single-mode spatial filter (SMSF), the requirements for nulling the star become simple to express. The wavefront from the star is incident on the collecting apertures of the instrument and delivered by the respective beamtrains to a central beam combiner, that couples the light into the SMSF. The electric field within the SMSF is the vector sum of the electric field contributions from each collecting aperture. The starlight is nulled when the electric fields in the SMSF sum to zero, requiring specific combinations of the amplitude and phase.

When there are two collecting apertures, the electric fields must have equal amplitudes, and phases that differ by π radians. This requirement must be met simultaneously for both horizontal and vertical polarization states, and for all wavelengths across the science bandwidth. This is possible if the beamtrains are perfectly symmetric, except for an achromatic π-phase shift. Nulling by a factor of 10^5 requires amplitudes to be balanced at the 0.5% level (intensities equal to within 1%) and the phase to be matched to within 5 mrad (8 nm at a wavelength of 10 μm). Recent analysis shows that obtaining a starlight null that is sufficiently stable for detecting an earth at 10 pc leads to even tighter tolerances: approximately 0.1% for amplitude matching, and 1 mrad for the phase (1.5 nm at a wavelength of 10 μm). A number of effects can perturb the amplitude and the phase of the E-field from a collecting aperture that gets coupled into the SMSF.

* oliver.p.lay@jpl.nasa.gov; phone 1 818 354-2521; fax 1 818 393-4950

1.2. Amplitude perturbations

Amplitude imbalances can be grouped into two categories: throughput, which determines how much light gets to the entrance plane of the SMSF, and wavefront errors which determine how much of the light is coupled into the fundamental mode of the fiber (similar to Strehl ratio). Minor reflectivity variations in the myriad mirrors that make up the telescope, beam compressors and delay lines can lead to throughput errors on the order of a percent. Unprotected gold minimizes wavelength dependence of the throughput in the mid-IR. Contaminants (both pre- and post-launch), and manufacturing tolerances of beamsplitter coatings, however, are expected to have more complex spectral dependences.

Wavefront errors, on the other hand, affect how much of the light couples into the modal filter according to the Strehl ratio. Without any corrections, lightweight, meter-class telescopes operating at a temperature of 40K are expected to have several hundred nm of WFE. Large numbers of optics in the beam train and alignment errors will add to the WFE. Some form of correction, if not a full-blown adaptive optics system, will be needed to increase coupling efficiency and balance the paths. For example, actuated focus control can be used to correct for amplitude perturbations with λ^{-2} dependence, such as small errors in the optical figure.

In a fully symmetric architecture, Fresnel diffraction of the beams will be nearly identical for all beams. Though each beam may have significant spatial amplitude and phase fluctuations, the difference between beams will be small. Our analyses of Fresnel diffraction in the presence of realistic WFE and aberrations (modeled with phase screens) indicate differences in coupling efficiency could be in the order of 1% with a λ^{0} and λ^{-1} dependence.

1.3. Phase perturbations

Since a white light fringe is localized to a few microns, stellar interferometers typically have delay lines to match static and dynamic path length errors. These delay lines can be used to correct for phase perturbations with a λ^{-1} dependence, such as a path length offset or a small optical figure error.

There are, however, other sources of phase errors that do not have a pure λ^{-1} dependence. For example, beamsplitter thickness variation of even a few microns between the different paths leads to tens of milliradians of phase error due to dispersion (see Fig. 1). Some of the nulling architectures require multiple beam combiners, each with multiple beamsplitters. Thickness variation of +/- 1 μm for each beamsplitter easily adds up to become a significant concern. For mid-IR materials such as ZnSe or KBr, we have found that correcting for a λ^{-2} variation can bring the error down to the required 1 mrad level. We continue to investigate the impact of index and thickness variations in multiplayer coatings (for beamsplitters, anti-reflection and dichroics) on phase errors.

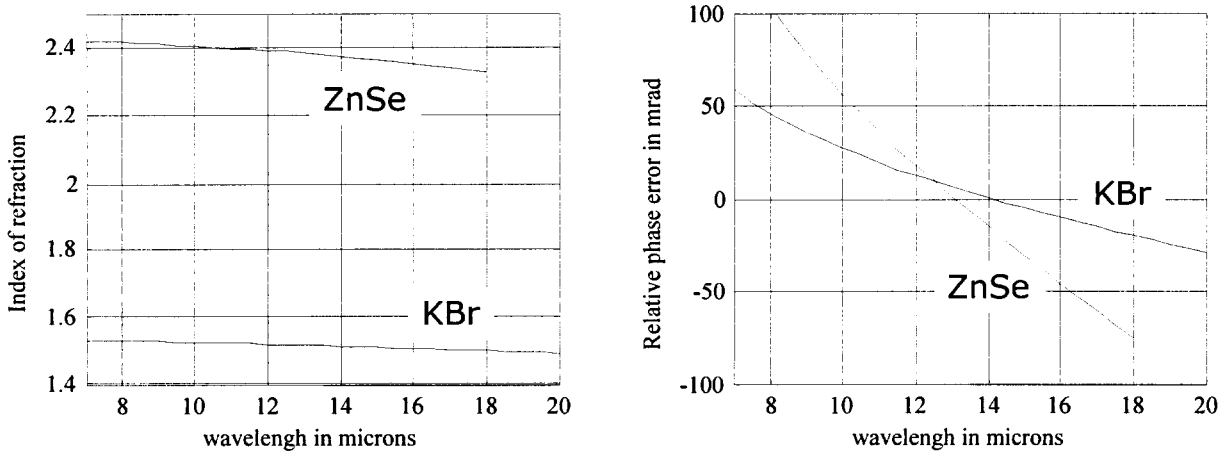


Fig.1. (a) Index of refraction of ZnSe and KBr vs. wavelength. (b) Relative phase error for 4 micron thickness differential.

As in the case of amplitude errors, diffraction can have the effect of amplifying WFE differences, and together can cause significant phase errors. In fact, a simple model of circularly symmetric beams and phase screens show phase errors in the order of 10 mrad with λ^0 and λ^{-1} dependence.

A summary of key contributors to amplitude and phase errors is listed in Table 1. Also included is the spectral dependence of the perturbation, many of which are not currently known. As can be seen, even with delay lines and focus control, additional correction is required for the λ^0 and λ^{-1} variations in amplitude, for the λ^0 and λ^{-2} variations in phase, and for those spectral dependences that are currently unknown. These requirements are challenging to meet for systems with identically sized collectors and beamtrains that are symmetric by design. They become even more difficult when there is intrinsic asymmetry, for example when the collectors have different sizes, or when the beam combination scheme requires different amplitudes from different collectors.

Table 1. Sources of asymmetry between input beams to nulling beam combiner.

Effect	Dependence of imbalance on wavelength	
	amplitude	phase
Mirror reflectivity	λ^0	?
Transmissive optics	λ^0	λ^{-2}
Beamsplitter coatings	λ^0, ?	?
Dichroic coatings	λ^0	?
Ground contamination	?	?
Outgassing	?	?
Thruster contamination	?	?
Path length offset	-	λ^{-1}
Wavefront figure error	λ^{-2}	λ^{-1}
Misalignment	λ^{-2}	λ^{-1}
Beam shear	λ^0	-
Diffraction & wavefront error	λ^0, λ^{-1}	λ^0, λ^{-1}

Two alternatives present themselves: (1) design the optical system to be symmetric with very tight tolerances on the transmissive optics, alignment, contamination, etc; (2) include a compensator that can correct for imbalances in the amplitude and phase, independently at each wavelength and polarization. The latter is what we call adaptive nulling (Fig. 2).

2. COMPENSATOR DESIGNS

Table 2 illustrates the requirements we have adopted on compensator performance. In addition to the science band of 6 – 17 μm, the compensator must pass the metrology wavelength, which is likely to be in the range of 0.5 – 2 μm. At least 6 spectral degrees of freedom are required (λ^0, λ^{-1}, λ^{-2} dependences for each of amplitude and phase), but more degrees of freedom are desirable given the number of unknown effects.

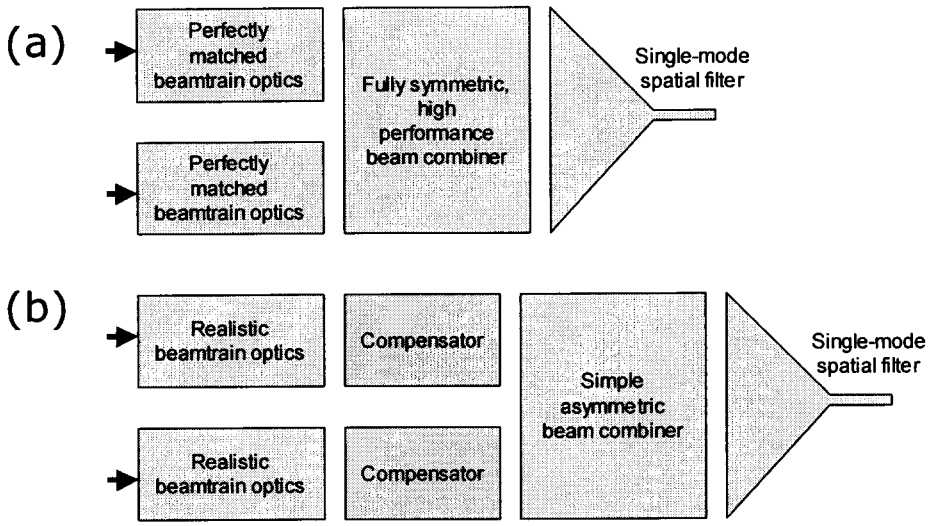


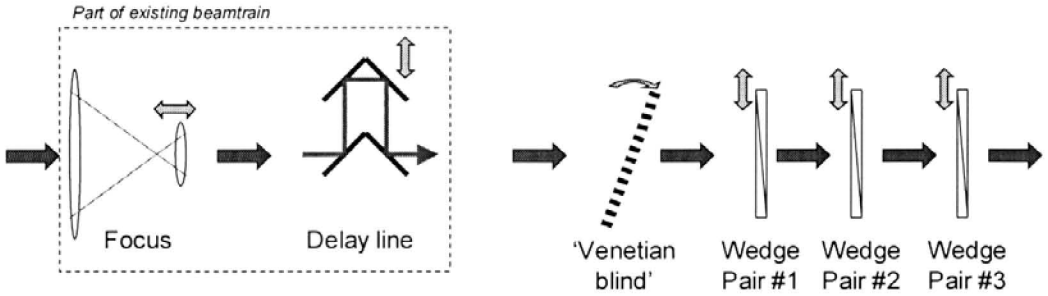
Fig.2. (a) A deep starlight null requires extremely well-matched beamtrain optics and a high performance symmetric beam combiner. (b) Inserting a compensator to correct for amplitude and phase perturbations relaxes the requirements on the beamtrains to more realistic levels and allows a simpler beam combiner design.

Table 2. Performance requirements for adaptive nulling compensator.

#	Requirement	
1	Wavelength range of operation	6-20 μm
2	Metrology wavelength	0.5-2 μm?
3	# independent spectral degrees of freedom	> 5 (20)
4	# independent polarization states	2
5	Null depth across the band	$< 10^{-5}$
6	Amplitude correction range	> 5%
7	Amplitude precision / stability (1 σ)	< 0.1%
8	Phase correction range	> 2 μm
9	Phase precision / stability (1 σ)	< 1 nm
10	Throughput reduction	< 20%
11	Polarization isolation	> 50 dB

The compensator should ideally act independently on the horizontal and vertical polarization states, and support a null depth of 10^{-5}. The null stability requirement leads to the amplitude and phase stability requirements of 0.1% and 1 nm, respectively. A maximum amplitude correction of 5% gives a dynamic range of 50 for amplitude, and the 2 μm phase correction range corresponds to a phase of 0.6 radians at a wavelength of 20 μm. The compensator should not be too lossy, and it should not mix the polarization states at more than the 10^{-5} level.

One approach to implementing a compensator is to use a "serial cascade" of correcting elements. Each element has a different spectral response to amplitude and/or phase, with an adjustable gain. An example is depicted schematically in Fig. 3.



Amp	λ^{-2}		λ^{0}	λ^{-1}	
Phase		λ^{-1}		λ^{-2}	λ^{0}

Fig. 3. Example of a serial cascade design for the compensator. The table below indicates the spectral dependence controlled by each element. The wedge pairs provide an adjustable thickness of material through which the beam propagates, with properties chosen to provide the desired spectral dependence.

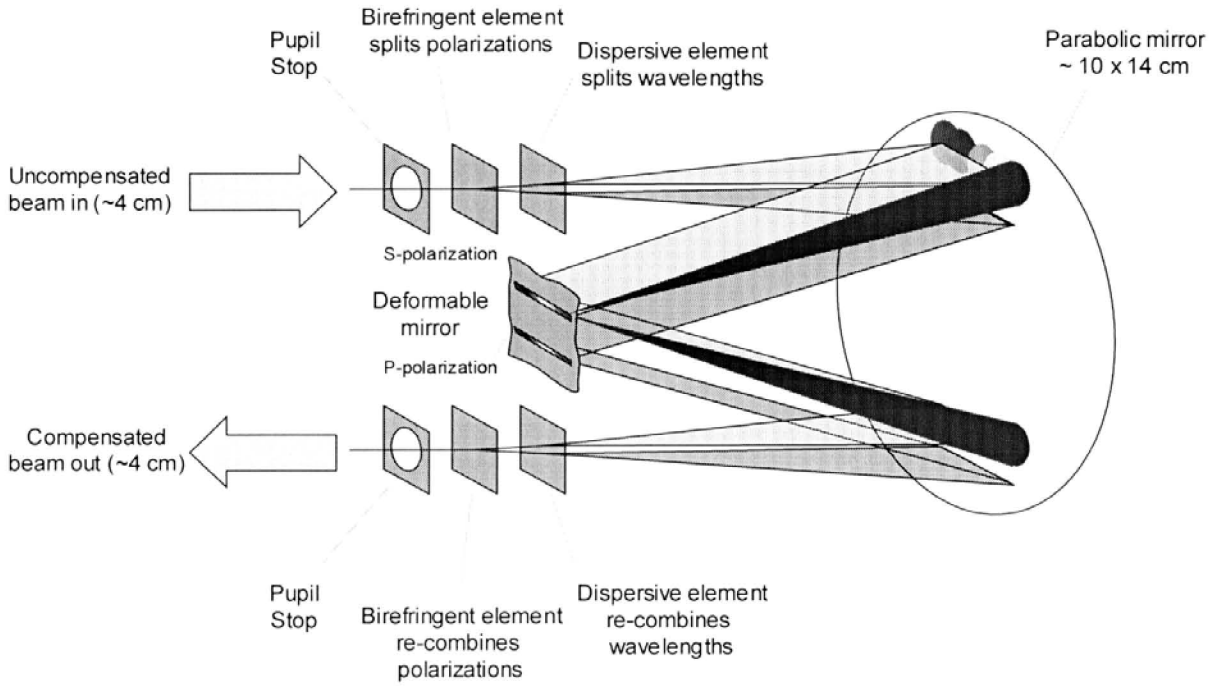


Fig. 4. Parallel, high-order compensator design, using a deformable mirror to control amplitude and phase.

The serial cascade approach is best suited to a low-order compensator, i.e. one with relatively few degrees of freedom. As shown in Fig. 2, it does not provide independent control of the different polarization states, although this capability could be added by introducing more elements. The throughput is clearly impacted as more elements are added, since all the photons, irrespective of their wavelength, must pass through all the elements.

An alternative approach is to split the light into the different wavelength and polarization states and operate on them in parallel before recombining them. In principle, it is then possible to implement a high order compensator (many spectral degrees of freedom) without a large impact on throughput. Such a compensator is illustrated in Fig. 4.

The uncorrected beam, with diameter ~ 4 cm, enters at the upper left, passes through a pupil stop, and then through a birefringent element that splits the polarization states by a small angle. The light is then dispersed by a prism and is incident on a parabolic mirror that focuses the collimated beams onto a deformable mirror (DM). At this point the input light is spread into two focused lines, one for each polarization state, dispersed by wavelength. After reflection from the DM, the light is re-collimated by the parabolic mirror, de-dispersed and the two polarization states are re-combined before passing through the exit pupil stop.

The DM allows independent control of the amplitude and phase for each polarization and wavelength, as illustrated in Fig. 5. Piston of the deformable mirror adjusts the phase of the output beam (Fig. 5a); changing the local slope of the deformable mirror at the focal point introduces a shear of the outgoing collimated beam, which is then converted into a reduction of amplitude by the exit pupil stop (Fig. 5b). The piston and local slope are adjusted independently for the different wavelengths and polarization.

This compensator is part of a control system for balancing the amplitudes and phases of the incoming beams. Also needed is a sensor for detecting the imbalances and an algorithm to make the appropriate adjustment at the DM. Since we are correcting for imbalances across the science band, the sensor must operate over the same range of wavelengths. There are at least 3 options:

1. Monitor the null depth directly at the science detector, in each of the spectral channels. The advantages are that no additional sensors are needed, there are no uncommon path effects, and there is no interruption of the science data. One disadvantage is that an iterative adjustment of amplitude and phase is needed to minimize the null depth. Another is that the null depth for the star is masked by the photons from the exo-zodiacal and local zodiacal dust, and by the thermal emission of the instrument, so that the measurement is not particularly sensitive.
2. Measure the amplitudes and phases of the different beams at regular intervals of time. The science observing must be interrupted to do this. Amplitude is obtained by measuring the photon rate at the science detector for each beam in turn (i.e. block all but one of the beams). Phase is determined by measuring the photon rates obtained for pairs of beams. No additional sensors are needed, there are no uncommon path effects, and the amplitude and phase are being sensed separately. The main disadvantage is that the time available for science observations is reduced.
3. Monitor the amplitudes and phases of the beams by splitting off some of the science light before the beams are combined and using separate detectors. As for the previous option, the amplitudes and phases are measured directly. Disadvantages are that science photons are being diverted to this purpose, there can be substantial uncommon path effects, and additional sensors are required.

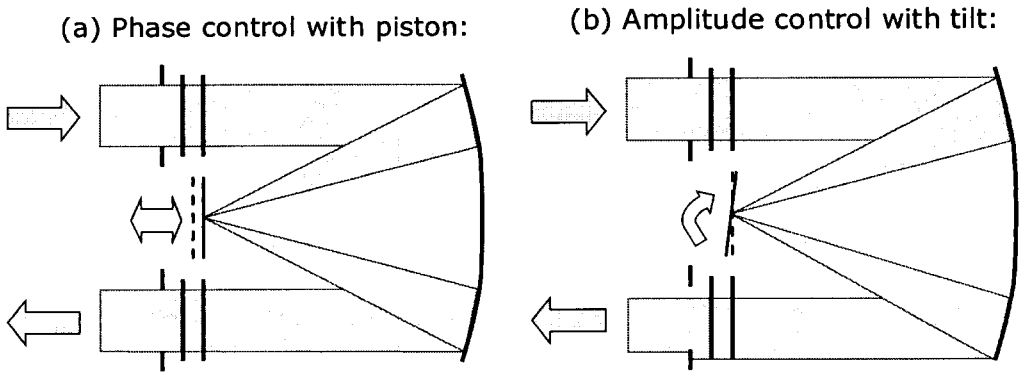


Fig. 5. Phase and amplitude control with a deformable mirror. Schematic represents a side view of Fig. 3 with beams shown for a single polarization and wavelength.

Achieving the 0.1% and 1 mrad levels of performance will only be possible at low bandwidth, $f << 0.1$ Hz, much slower than the bandwidths for path and pointing control ($f \sim 100$ Hz). Adaptive nulling is a quasi-static correction, which is not a problem if the sources of asymmetry (Table 1) are not changing on short timescales. The correction algorithm will depend on the nature of the sensor, and will need to account for any cross-coupling effects between the amplitude and phase control.

3. DEVELOPMENT PLANS

A demonstration of adaptive nulling has been funded as part of the TPF technology development effort over the next 3 years. The current plan starts with a proof-of-concept demonstration at visible/near-IR wavelengths. This would include a DM and demonstrate full functionality with commercial components, but at a reduced level of performance. If successful, it would be followed by a demonstration at mid-IR wavelengths, at levels close to the requirements in Table 2.

3.1. Visible proof-of-concept experiment

The first adaptive nuller will be built to operate at visible/near-IR wavelengths. Initially, a single wavelength HeNe laser will be used as the light source, and a PZT actuated flat mirror will be used in place of a DM. A simple photodetector will be used as the sensor and will be sampled by a PC to generate control signals to the PZT amplifier. Two shutters allow us to block either the nuller or reference arm so the amplitude of each may be measured independently. We can dither the nuller arm mirror to find the best nulled output. A perturbation to the beam's amplitude and/or phase may be added to either arm to demonstrate control. We expect to be able to control the amplitude and phase of the adaptive arm compared to a static reference arm and achieve a null better than 2×10^{-2}.

The system will then be augmented to full functionality, as depicted in Fig. 6. The laser source will be replaced by a wide bandwidth near-IR source consisting of two fiber coupled LEDs. The simple photodetector will be upgraded to an optical spectrometer and the PZT-actuated flat mirror will be replaced with a square MEMS type DM with 140 actuators (see Table 3 for list of specifications).

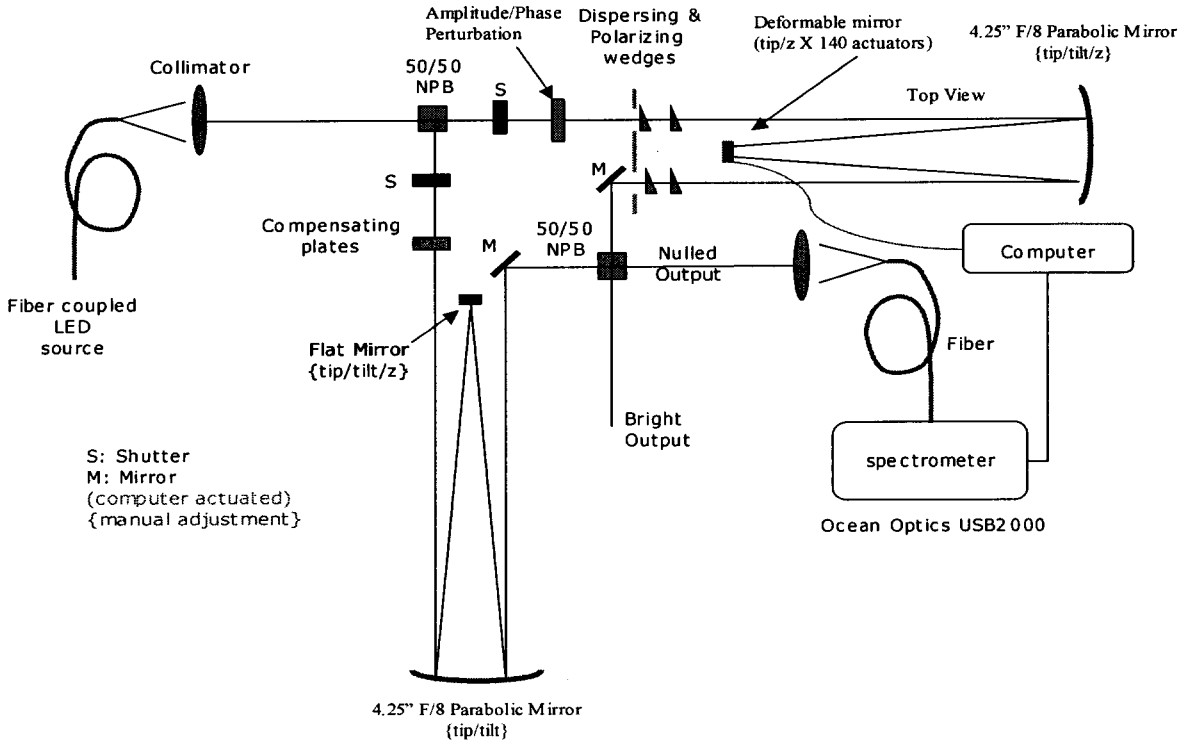


Fig 6. Full visible proof-of-concept experiment schematic.

Figure 7 illustrates how we intend to demonstrate the performance of the proof-of-concept experiment. In Fig. 7a, the photon rates are measured independently for each arm of the interferometer by closing the shutter on the other arm. The DM is then commanded to match the photon rates in each spectral channel. Figure 7b shows the phase correction process. A small delay offset is introduced between the arms, producing a set of fringes in the spectrometer output. Applying piston to a single element of the DM will manifest itself as a slight shift in the fringe for that channel in the spectrometer. Figure 7c shows the uncorrected and corrected spectrometer output when the two arms of the interferometer interfere with zero added path offset. The goal is to demonstrate a null level of 2% over all spectral channels.

Table 3: Specifications for Boston Micromachines μDM 140 deformable mirror

Number of elements	140
Format	12 x 12 minus 4 corners
Continuous facesheet	Yes
Stroke	2 μm
Surface rms	30 nm
Repeatability	4 nm

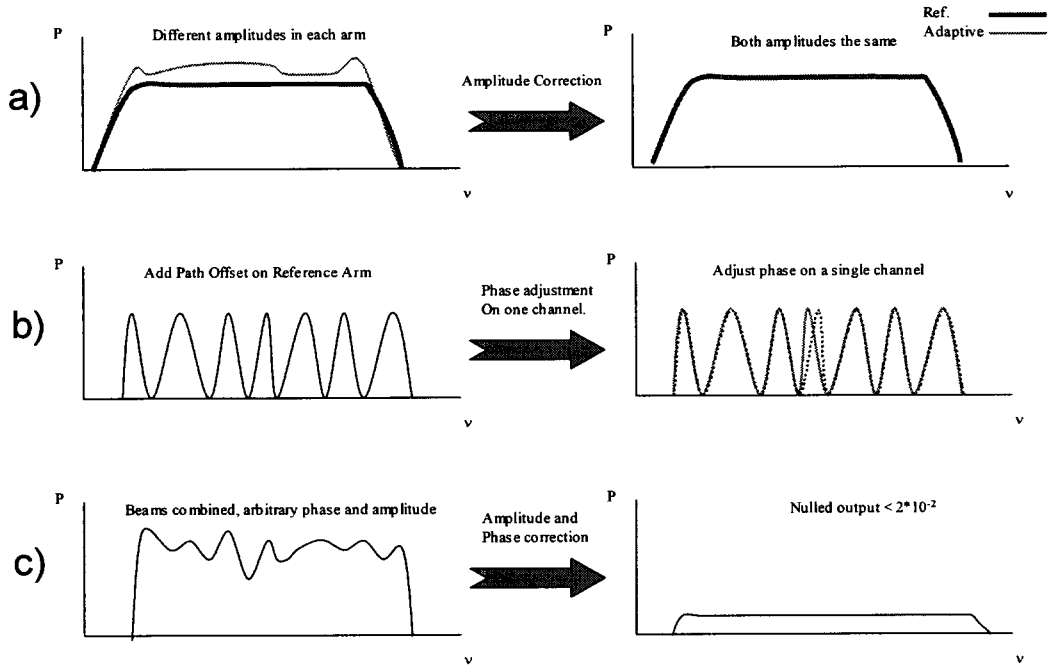


Fig 7. Amplitude, phase and nulling control with the proof-of-concept experiment. Each set of axes represents photon rate vs. optical frequency.

3.2. Mid-IR design

The feasibility of a mid-IR compensator is largely determined by the availability of appropriate material for spectral and polarization separation. The wavelength dispersing prism must not only be transparent in the entire science band (6-17 μm), fringe tracking band (2-4 μm) and metrology wavelength (1-2 μm), but also must not disperse the spectrum so much that the metrology beam traverses a very different path. KBr (Potassium Bromide) appears to be a reasonable choice based on preliminary modeling in Zemax (see Fig 8). Although KBr is hygroscopic, its non-toxic nature and the fact that it has flight history make it an attractive choice.

Polarization separation appears to be more of a challenge. Cadmium Selenide (CdSe) is the only birefringent crystal listed in the Handbook of Optics that is transparent from 1 - 20 microns. Initial modeling in Zemax indicates that CdSe will in fact sufficiently separate the polarization for the compensator to work in the mid-IR (Fig 9). Polarization splitting with a wire grid on a wedge is being considered as an alternate option. We plan to investigate these mid-IR materials in parallel with the visible experiment to prepare for a mid-IR compensator demonstration in 2005.

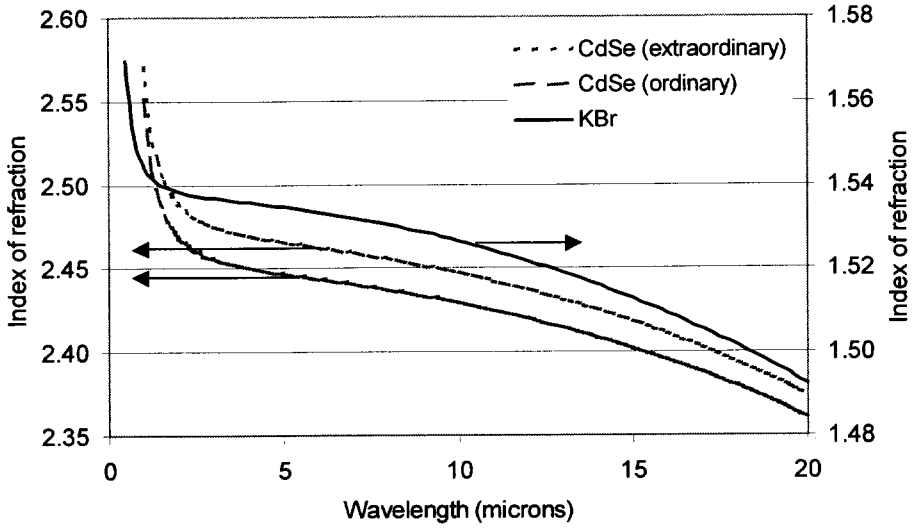


Fig. 8: Refractive index curves for Potassim Bromide (KBr) and the birefringent material Cadmium Selenide (CdSe). The low dispersion of KBr at short wavelengths is useful for minimizing the angular deviation of the metrology beam.

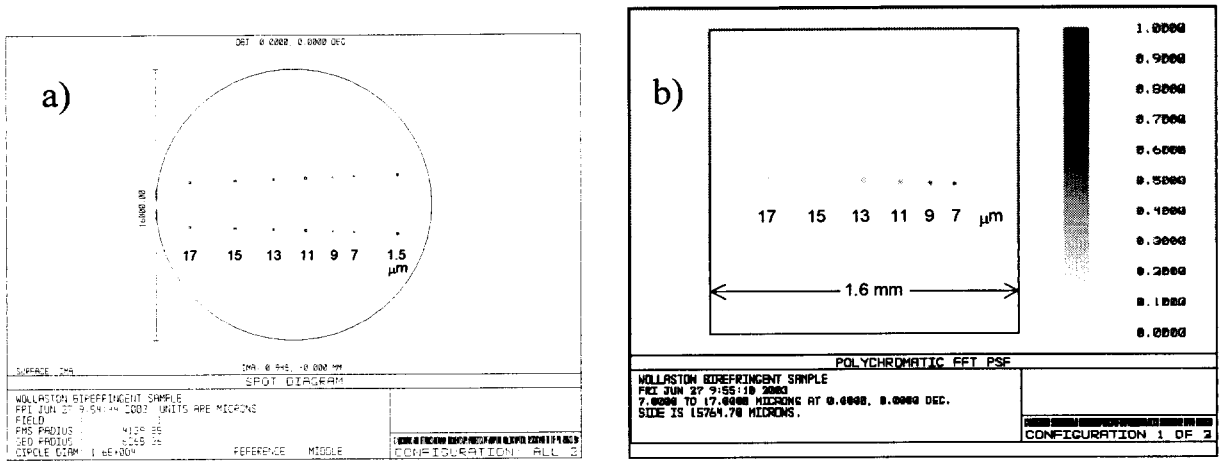


Fig. 9: Results from Zemax with a 15.3-degree KBr dispersing prsim and Wollaston type CdSe prism. (a) Spot diagram at the focal plane of a 1000 mm FL paraxial lens at discrete wavelengths every 2 microns. The m,etrology wavelength is not dispersed too far from the mid-IR band. The polarization states are separated in the vertical direction. (b) Point spread functions for a single polarization at different mid-IR wavelengths. The reduced peak intensity is due to the larger diffraction limited spot size.

4. SUMMARY

This paper describes the need for a compensator that can adaptively correct for mismatches between the optical beamtrains of a Mid-IR nulling interferometer, across the science bandwidth with a range of spectral dependences. By easing the tight requirements on symmetry, such a compensator also enables more flexibility in the optical design and the use of much simpler, asymmetric nulling beam combiners. One possible compensator design uses a deformable mirror to

provide independent control of amplitude and phase for each wavelength and polarization. This will be demonstrated in a proof-of-concept experiment at JPL.

The work described in this paper was performed at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

REFERENCES

1. E. Serabyn, "Nulling interferometry progress", *Proc. SPIE*, Vol. 4838, 594-608, 2003
2. C. V. M. Fridlund & P. Gondoin, The Darwin mission, *Proc. SPIE*, Vol. 4852, 394-404, 2003
3. *Terrestrial Planet Finder*, eds. Beichman, C. A., Woolf, N. J., and Lindensmith, C. A., JPL Publication 99-3, 1999