EXTREME OPTICS AND THE SEARCH FOR EARTH-LIKE PLANETS

ROBERT J. VANDERBEI

Operations Research and Financial Engineering
Princeton University

Revised February 3, 2006

ABSTRACT. In this paper I describe a new and exciting application of optimization technology. The problem is to design a space telescope capable of imaging Earth-like planets around nearby stars. Because of limitations inherent in the wave nature of light, the design problem is one of diffraction control so as to provide the extremely high contrast needed to image a faint planet positioned very close to its much brighter star. I will describe the mathematics behind the diffraction control problem and explain how modern optimization tools were able to provide unexpected solutions that actually changed NASA's approach to this problem.

Date: February 3, 2006.

Key words and phrases. linear programming, convex optimization, nonlinear programming, Fourier transform, high-contrast imaging, Terrestrial Planet Finder.

The author was supported by a grant from the ONR (N00014-98-1-0036).

1. Introduction

Optimization often plays a central role in engineering design problems. In recent years, considerable attention has been directed toward optimizing the design of antenna arrays for radar [6, 8], finite impulse response filters [14, 2], and minimum weight structure design [4, 1], to name just a few areas. In this paper, I will discuss a new and exciting application of optimization technology. The general area is optical design. The specific problem is to design a telescope capable of achieving the extremely high contrast needed to image planets around nearby stars.

We have a close-up view of only one star, our Sun. As we all known this particular star has circling it a wealth of smaller objects such as planets, comets, and asteroids. And one of those planets is Earth, our home. All other stars, even the nearest ones, are so far away that it is impossible to see if there are planets around them by pointing a telescope at them. Yet, the stars themselves can be studied in some detail and we know from these studies that our Sun is a rather typical star. There are billions of others just like it throughout our galaxy. If our Sun isn't special, then probably planets aren't special either. With this logic astronomers have long expected that there are planets around other stars. But, detecting them is a major technological challenge.

However, in the early 1990's there was the mounting realization that large, Jupiter-sized planets with close-in orbits could be detected indirectly by looking for the wobble induced in the star by a heavy planet orbiting it. If our line of sight is perpendicular to the plane of the orbiting planet, then we could, in principle, detect a tiny circular motion of the planet against more distant background stars. This is possible but is considered very difficult. What proved to be more promising is to look for systems in which our perspective from Earth puts us more or less close to the planet's orbital plane. In that case, the planet moves toward us and then away from us as it orbits its star and in so doing it induces a similar to and fro wobble on the star itself. This wobble can be detected by very careful measurements of the Doppler shift of the star's spectral lines. Mayor and Queloz [7] were the first to show definitively that there is a massive planet orbiting a star. The star they studied is known as 51 Pegasi. Since then hundreds more stars have been shown to have planets and a few have even been shown to have multiple planets (see, e.g., [3]). Most of these discoveries have been made by the indirect *radial velocity method*. A few others have been discovered by another method called the *transit method*. These transiting planets have then been verified by the radial velocity method.

So, today, there is no longer any doubt that planets around other stars are common. However, no one has yet detected, by any means, an Earth-sized planet at a comfortable Earth-Sun-like distance from its star. The main reason that Earth-sized planets have not been discovered by the indirect methods described above is simply that such planets are too small and therefore don't induce much wobble. Jupiter's mass is about 300 times that of Earth's. Those two orders of magnitude make a big difference for indirect detection. Next, astronomers began to consider that maybe it is possible to image Earth-like planets directly. As mentioned at the beginning, this is an enormously difficult challenge but armed with proof that planets are out there it became natural to ask whether one could design a special-purpose telescope capable of detecting these planets. To answer this question, NASA/JPL has commissioned astronomers, engineers, and an optimization person (just one so far; more needed) to study this question. In this paper, I will describe a few of the approaches being considered and explain the optimization problems that arise in this design problem.

2. The Problem is Hard

The nearest star beyond our solar system is about 4 light years away. In order to have a large enough sample of stars to survey, we need to consider stars out to say 40 light years. That gives a sample of about 1000 stars.

Consider for a moment how our own Solar System would look if we could step away and look back at it from a distance of 10 parsecs (i.e., 33 light years). Assuming the most favorable configuration, Earth would be displaced 0.1 arcseconds from the Sun (this, by the way, follows immediately from the definition of parsec as the reciprocal of parallax). The resolving power of a telescope grows linearly with the diameter of its primary mirror. A 60 inch (1.5 m) primary mirror, which is not particularly big by today's standards, is in principle adequate to resolve a pair of objects 0.1 arcseconds apart. But, the story is complicated by two factors.

The first factor is that atmospheric turbulence blurs pictures taken from ground-based observatories. At sea level, this turbulence limits resolution to about 2 or 3 arcseconds. For this reason, all modern telescopes are built on mountain tops. Up there the air is thinner and the effects of turbulence are reduced. Even so, in the best locations atmospheric turbulence limits resolution to about 0.5 arcseconds. Despite modern advances in adaptive optics aimed at correcting for atmospheric turbulence, it is currently felt that only a space-based telescope will be capable of imaging planets around nearby stars.

The second challenge stems from the fact that the Sun shines on its own but a planet just reflects sunlight. The result is that, from a distance, the Sun appears 10^{10} times brighter than the Earth. The resolution discussion above applies only to pairs of objects of equal brightness. Such a huge disparity in brightness makes the problem dramatically more difficult. In fact, if one were to build a conventional telescope and just make it bigger as needed to achieve the required contrast, the primary mirror would have to about 1200 meters in diameter! Clearly, such a brute force solution is out of the question.

The fact that such a large mirror would be required by the brute force solution is a result of the wave nature of light. A telescope does not focus light to a point. Instead, one gets a small blob of light with a system of concentric rings around it (see Figure 1). Each ring is fainter than the previous one but they don't stop and their intensity only decreases slowly with their distance from the central spot. The central spot is called the *Airy disk*, the rings are called *diffraction rings*, and the entire picture is called the *Airy pattern*, which is just one example

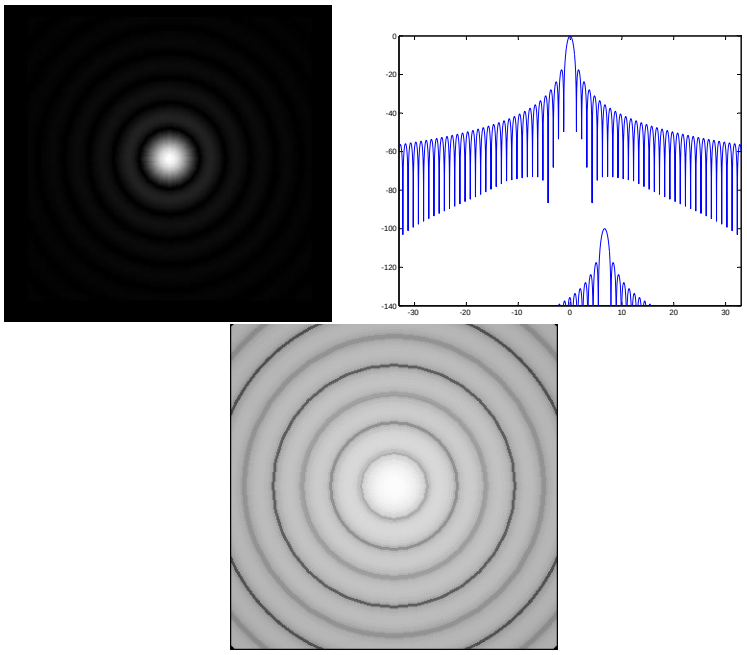

FIGURE 1. The standard Airy pattern. *Left.* This is how a star looks to someone peering through a telescope with a high-power eyepiece. *Middle.* Cross-sectional semi-log plot. The vertical axis is labeled in decibels. An intensity ratio of 10^{-10} corresponds to -100 dB. Also shown is a planet, shifted slightly to the right and 100 dB than the star. *Right.* In subsequent figures all 2-D plots are log-stretched so that black corresponds to 10^{-10} and white corresponds to 1. Shown here is how the Airy pattern looks under such a log stretch.

of what is generally called a *point spread function* (PSF). One needs to go to about the 750-th diffraction ring before it is 10^{-10} times as bright as the Airy disk. Fortunately, however, there are some proverbial knobs that one can tweak to sculpt a different diffraction pattern, one more amenable to planet detection.

3. A Brief Discussion of Diffractive Optics

The diffraction pattern depends on the shape of the opening through which light passes. This opening is called the entrance pupil. The reason that a simple telescope produces a circular Airy disk surrounded by circular diffraction rings is a direct consequence of the unstated assumption that the entrance pupil is an open circular disk. If it were something else, the diffraction pattern would be different. For example, many telescopes have a so-called secondary mirror positioned in front of the primary mirror and hung by a support structure called spiders. These supports partly obstruct the entrance pupil and therefore change the diffraction pattern. In this case, the result is bright diffraction "spikes". Diffraction spikes are not particularly useful for high contrast imagining, but it is possible to imagine other entrance pupil shapes that might have more advantageous diffraction patterns.

In order to explain how to find a pupil shape that delivers high contrast at locations very close to the central Airy disk, we need to describe briefly the relationship between the entrance pupil and the point spread function. The diffractive properties of light are a consequence of the fact that light is an electromagnetic wave. We can think of starlight as a plane wave arriving at the entrance pupil. The focusing element (i.e., primary mirror) of the telescope redirects the light so that, in principle, it all converges at a point in the focal plane. A more detailed analysis shows that all of the light that arrives at this focal point travels precisely the same distance no matter what part of the entrance pupil one considers the light coming from. Hence, if we think of the light as being re-emitted from every point across the entrance pupil, then all these so-called *wavelets* arrive at the focus point in the image plane in exactly the same phase. Hence, at this point there is no destructive interference. All of the waves contribute to an accumulation of signal at this point in the image plane. Now, other points in the image plane can be analyzed in a similar way. At other points, however, some of the waves will be in phase with others, while others will be out of phase and there will be some sort of mix between constructive and destructive interference. The result is that there will be some light seen at essentially all points in the image plane. The intensity depends on the amount of constructive interference.

It is not hard to turn the verbal description of the preceding paragraph into an explicit formula for the electric field in the image plane as an integral over the entrance pupil of the appropriate phase shifts that depend on the distance between the point in the entrance pupil and the point being considered in the image plane. The integrand is

a little bit messy but for points in the image plane close to the focus point, a linearization of the phase shifts reduces the integral to a simple Fourier transform for the electric field $E(\cdot)$ in the image plane:

$$E(\xi,\zeta) = \frac{1}{\lambda f}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{2\pi i(x\xi+y\zeta)/\lambda f} A(x,y)dydx.$$

Here, $A(\cdot)$ is the function that is one on the pupil opening and zero elsewhere, (ξ,ζ) is a point in the image plane, λ is the wavelength of visible light, and f is the *focal length*, i.e., the distance between the mirror and the image plane. For a circular pupil,

$$A(x,y) = 1_{x^2+y^2\le D/2}$$

where D is the diameter of the primary mirror. It is useful to change variables to unitless quantities right here at the start. Let

$$\tilde{x} = x/D, \quad \tilde{y} = y/D, \quad \tilde{A}(\tilde{x},\tilde{y}) = A(\tilde{x}D,\tilde{y}D),$$

$$\tilde{\xi} = \xi D/\lambda f, \quad \tilde{\zeta} = \zeta D/\lambda f, \quad \tilde{E}(\tilde{\xi},\tilde{\zeta}) = \frac{\lambda f}{D^2}E(\tilde{\xi}\lambda f/D, \tilde{\zeta}\lambda f/D).$$

If for notational convenience we then drop the tildes, the formula for the image-plane electric field reduces nicely to

$$E(\xi,\zeta) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{2\pi i(x\xi+y\zeta)} A(x,y)dydx$$

and the support of the function $A(\cdot)$ is contained in the unit disk. Rewriting using polar coordinates, we get

$$E(\rho,\phi) = \int_{0}^{\infty}\int_{0}^{2\pi} e^{-2\pi i r\rho\cos(\theta-\phi)} A(r,\theta)rd\theta dr.$$

Here, ρ is a radius and ϕ is a polar-coordinate angle in the image plane. If we further assume that $A(\cdot)$ is a function of r but not θ, then we get that the electric field is also a function only of radius:

$$E(\rho) = 2\pi\int_{0}^{\infty} J_0(2\pi r\rho)A(r)rdr.$$

Here J_0 denotes the 0-th order Bessel function of the first kind. It is important to note that in all these cases, the image-plane electric field $E(\cdot)$ depends linear on $A(\cdot)$.

It is worth reviewing the change of variables made earlier. The unitless pupil-plane "length" r is given as a multiple of the aperture D and the unitless image-plane "length" ρ is given as a multiple of focal-length times wavelength over aperture $f\lambda/D$ or, equivalently, as an angular measure on the sky, in which case it is a multiple of just λ/D radians. For example, if $\lambda = 0.5\mu$m and $D = 10$m then $\lambda/D = 0.01$ arcseconds.

The *point spread function* is the image plane *intensity*, which is the square of the magnitude of the electric field.

In the derivation just given, the function $A(\cdot)$ was assumed to be zero-one valued so that it could simply represent an open aperture of some shape. Each such function $A(\cdot)$ represents a *shaped pupil*. However, it is also possible to relax this assumption by letting $A(\cdot)$ take on values between zero and one, inclusively. The physical interpretation is that a variable attenuating filter is placed in the incoming light beam. Places where the function is zero correspond to opaqueness, places where it is one correspond to openness, and places where it is between the extremes correspond to various shades of tinting. Such a nonuniform tinting is called an *apodization*.

4. Optimizing the Shape of the Point Spread Function

The problem is to find an apodization or, better yet, a pure shaped pupil for which the point spread function shows less than 10^{-10} intensity in some region very close to the central Airy disk and, given that constraint, the problem is to maximize the amount of light that gets through the apodizer. More specifically, we could specify inner and outer working angles ρ_{iwa} and ρ_{owa} and require that

$$\frac{E^2(\rho)}{E^2(0)} \leq 10^{-10} \quad \text{for } \rho_{\text{iwa}} \leq \rho \leq \rho_{\text{owa}}.$$

Expressed in this way, the constraint is a nonconvex nonlinear constraint. However, taking square roots and clearing the denominator, we can rewrite this constraint in a purely linear way:

$$-10^{-5}E(0) \leq E(\rho) \leq 10^{-5}E(0) \quad \text{for } \rho_{\text{iwa}} \leq \rho \leq \rho_{\text{owa}}.$$

Ideally, we would like to maximize the amount of light that lands on the central spot. That is, we would like to maximize

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

We call this the *Airy throughput*. The Airy throughput is a convex quadratic function of the apodization $A(\cdot)$. This is a bad choice for maximizing. If, however, we assume that ρ_{owa} is infinity (or, at least very large), then the contrast constraints essentially guarantee that no light will land outside the central spot. Hence, maximizing the amount of light on the central spot is almost the same as maximizing the total amount of light. In other words, the upper limit of integration in (1) can be set to infinity. But then Parseval's theorem tells us that

$$\int_0^{\infty} E^2(\rho)\rho d\rho = \int_0^{\infty} A^2(r) r dr.$$

(Physically, Parseval's theorem simply tells us that the total amount of light energy arriving at the image plane is equal to the total amount of light energy that passed through the entrance pupil.) In the case of a shaped pupil, $A(\cdot)$ takes values only zero and one and hence

$$\int_0^{\infty} A^2(r) r dr = \int_0^{\infty} A(r) r dr.$$

This last functional is linear. Whether we are working with apodizations or shaped pupils, this last functional serves as a useful surrogate for the Airy throughput. Finally, since the support of the apodization function is contained in the unit disk, the upper integration limit can be changed from ∞ to $1/2$.

We are now ready to formulate an optimization problem for the best high-contrast apodization:

$$\begin{aligned} \text{maximize} \quad & \int_0^{1/2} A(r) r dr \\ \text{subject to} \quad & -10^{-5}E(0) \le E(\rho) \le 10^{-5}E(0), \ \rho_{\text{iwa}} \le \rho \le \rho_{\text{owa}}, \\ & 0 \le A(r) \le 1, \qquad\qquad 0 \le r \le 1/2. \end{aligned} \tag{2}$$

This is an infinite dimensional linear programming problem. But by discretizing the set of r's and the set of ρ's, one can formulate a finite dimensional linear program. In practice, discretizations consisting of hundreds of points are generally adequate to get solutions that closely approximate the infinite dimensional problem.

Part of the art of optimization is to pick certain auxilary parameters appropriately. The choice of ρ_{owa} is an example. While setting it to infinity has a certain mathematical appeal, for the real problem of planet finding it suffices to set it to a value say about 10 times ρ_{iwa}. And, of course, as a practical matter one needs to set it to a finite value. Anyway, making it finite introduces an unexpected consequence. Namely, the optimal solution to (2) turns out to be a so-called *bang-bang solution*. That is, even though we allow $A(\cdot)$ to take values between zero and one, essentially all of the values end up at their limits. This is good if we wish to design a pupil mask but it is bad if our aim is to find the best smooth apodization. We will discuss both of these cases.

5. Optimal Smooth Apodization

One way to find a smooth solution is to impose smoothness constraints. For example, we can impose upper and lower bounds on the first and/or second derivatives of $A(\cdot)$. It turns out that the following problem gives good apodizations:

$$\begin{aligned} \text{maximize} \quad & \int_0^{1/2} A(r) r dr \\ \text{subject to} \quad & -10^{-5}E(0) \le E(\rho) \le 10^{-5}E(0), \qquad \rho_{\text{iwa}} \le \rho \le \rho_{\text{owa}}, \\ & 0 \le A(r) \le 1, \qquad\qquad 0 \le r \le 1/2, \\ & -50 \le A''(r) \le 50, \qquad\qquad 0 \le r \le 1/2. \end{aligned}$$

An AMPL model for this problem is shown in Figure 2. As shown in the AMPL model, the pupil-plane integration interval $[0, 1/2]$ has been discretized using 400 points uniformly spaced over the interval. Similarly, the image-plane region of interest, $0 \leq \rho \leq 60$, has also been discretized into 400 uniformly distributed points. Modern optimization tools can solve this linear programming problem in well less than a minute on current generation computers. Of course, one could, and perhaps should, use a much larger value for the discretization parameter `N`. In fact it would also be interesting to consider discretizing just one of these two domains while leaving the other one infinite. The result is then a semi-infinite linear programming problem which might prove to be tractible. However, we leave as a topic for future research. Instead, we compute the solution just for `N=400` and then we spline the resulting apodization function using many more points and compute the resulting PSF at a much finer resolution in the image plane to verify that this finer approximation is still a valid solution; i.e., that it satisfies the contrast constraints.

The tightest inner working angle for which a feasible solution exists turns out to be $\rho_{\text{iwa}} = 4$. The result for this inner working angle is shown in Figure 3. For this value of ρ_{iwa}, the optimal Airy throughput is about 9% (Airy throughput are given as percentages of the total possible throughput). For comparison, let's consider again a clear open circular aperture; i.e., the standard Airy pattern. The first null in the Airy pattern occurs at $\rho_{\text{iwa}} = 1.24$ and 84.2% of the light lands on the central Airy spot. But, the first diffraction ring is only about two orders of magnitude less intense than the peak value. To get ten orders of magnitude reduction, one needs to look beyond $\rho_{\text{iwa}} = 748$.

We end this section by pointing out that apodization is a rather old concept. In fact, Bell Labs mathematician David Slepian studied in [9] the problem of finding an apodization function that minimizes the total energy falling outside some radius. That problem is very similar to the one presented here. Because he only introduced a single quadratic constraint, he was able to solve his problem more or less explicitly in terms of what he called *prolate spheroidal wave functions*.

```
function J0;

param pi := 4*atan(1);
param N := 400; # discretization parameter
param rho0 := 4;
param rho1 := 60;

param dr := (1/2)/N;
set Rs ordered := setof {j in 0.5..N-0.5 by 1} (1/2)*j/N;

var A {Rs} >= 0, <= 1, := 1/2;

set Rhos ordered := setof {j in 0..N} j*rho1/N;
set PlanetBand := setof {rho in Rhos: rho>=rho0 && rho<=rho1} rho;

var E0 {rho in Rhos} = 2*pi*sum {r in Rs} A[r]*J0(2*pi*r*rho)*r*dr;

maximize area: sum {r in Rs} 2*pi*A[r]*r*dr;
subject to sidelobe_pos {rho in PlanetBand}: E0[rho] <= 10^(-5)*E0[0];
subject to sidelobe_neg {rho in PlanetBand}: -10^(-5)*E0[0] <= E0[rho];

subject to smooth {r in Rs: r != first(Rs) && r != last(Rs)}:
    -50*dr^2 <= A[next(r)] - 2*A[r] + A[prev(r)] <= 50*dr^2;

solve;
```

FIGURE 2. An AMPL model for finding the optimal apodiation function.

6. Concentric Ring Masks

The optimal apodization given in the previous section is in many respects ideal. However, it has a major drawback. With current technology it is not possible to make an apodization with the level of precision needed to achieve the requisite high contrast. From a manufacturing perspective, it is much easier to make a shaped pupil instead. That is, we seek a bang-bang solution. As already mentioned, without explicit smoothness constraints, the solution to the optimal apodization problem produces just such a binary solution. However, the discretization limits the places where the binary mask can transition between zero and one. It turns out that these transition points need to be computed with a high degree of precision, requiring at least some tens of thousands of discrete sample points, if not more. Such an approach might be tractible but there is in fact an easier way which we now describe.

Recall that for circularly symmetric apodizations

$$E(\rho) = 2\pi \int_0^{1/2} J_0(2\pi r\rho) A(r) r dr,$$

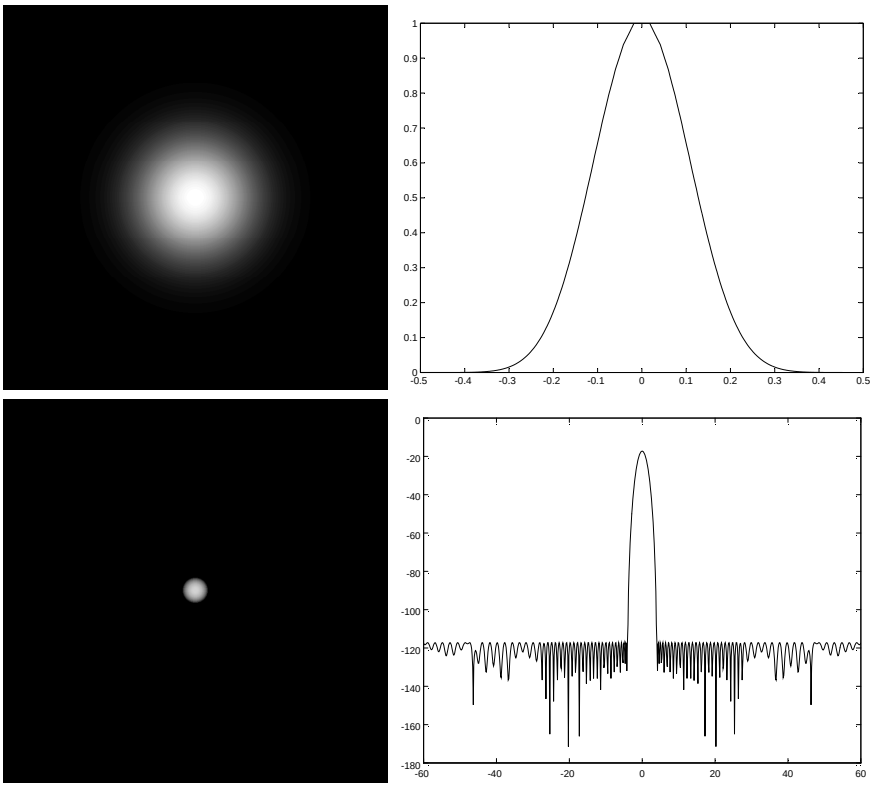

FIGURE 3. *Top.* The optimal apodization. *Bottom.* The PSF. The inner working angle is $\rho_{\text{iwa}} = 4$. The Airy throughput is 9%.

where J_0 denotes the 0-th order Bessel function of the first kind. Let

$$A(r) = \begin{cases} 1 & r_{2j} \le r \le r_{2j+1}, \qquad j = 0, 1, \ldots, m-1 \\ 0 & \text{otherwise,} \end{cases}$$

where

$$0 \le r_0 \le r_1 \le \cdots \le r_{2m-1} \le 1/2.$$

The integral defining the image-plane electric field can now be written as a sum of integrals and each of these integrals can be explicitly integrated to get:

$$E(\rho) = \sum_{j=0}^{m-1} \frac{1}{\rho} \left(r_{2j+1} J_1 \left(2\pi\rho r_{2j+1} \right) - r_{2j} J_1 \left(2\pi\rho r_{2j} \right) \right). \tag{3}$$

```
function intrJ0;

param pi := 4*atan(1);
param N := 400; # discretization parameter
param rho0 := 4;
param rho1 := 60;

var r {j in 0..M} >= 0, <= 1/2, := r0[j];

set Rhos2 ordered := setof {j in 0..N} (j+0.5)*rho1/N;
set PlanetBand2 := setof {rho in Rhos2: rho>=rho0 && rho<=rho1} rho;

var E {rho in Rhos2} =
    (1/(2*pi*rho)^2)*
    sum {j in 0..M by 2} (intrJ0(2*pi*rho*r[j+1])-intrJ0(2*pi*rho*r[j]));

maximize area2: sum {j in 0..M by 2} (pi*r[j+1]^2 - pi*r[j]^2);
subject to sidelobe_neg2 {rho in PlanetBand2}:
    -10^(-5)*E[first(rhos2)] <= E[rho];
subject to sidelobe_pos2 {rho in PlanetBand2}:
                                E[rho] <= 10^(-5)*E[first(rhos2)];

subject to order {j in 0..M-1}: r[j+1] >= r[j];

solve mask;
```

FIGURE 4. An AMPL model for finding an optimal concentric ring mask.

Letting the r_j's be design variables, the resulting optimization problem for a shaped pupil is

$$\text{maximize} \quad \sum_{j=0}^{m-1} \pi(r_{2j+1}^2 - r_{2j}^2)$$

$$\text{subject to} \quad -10^{-5}E(0) \le E(\rho) \le 10^{-5}E(0), \quad \text{for } \rho_0 \le \rho \le \rho_1$$

where $E(\rho)$ is given by (3). This problem has an infinite number of constraints but only a finite number of variables. It is, however, a nonconvex optimization problem. Nonconvex optimization problems are hard to solve to a provably optimal solution. But, locally optimal solutions are often easy to find, especially if a good starting point is given. In this case, the bang-bang solution to the linear problem serves as a good starting point for the nonlinear problem. Figure 4 shows the AMPL model and Figure 5 shows the result.

The difficulty with the concentric-ring mask is the problem of how to support the inner rings. If each ring "hung" from the next one further

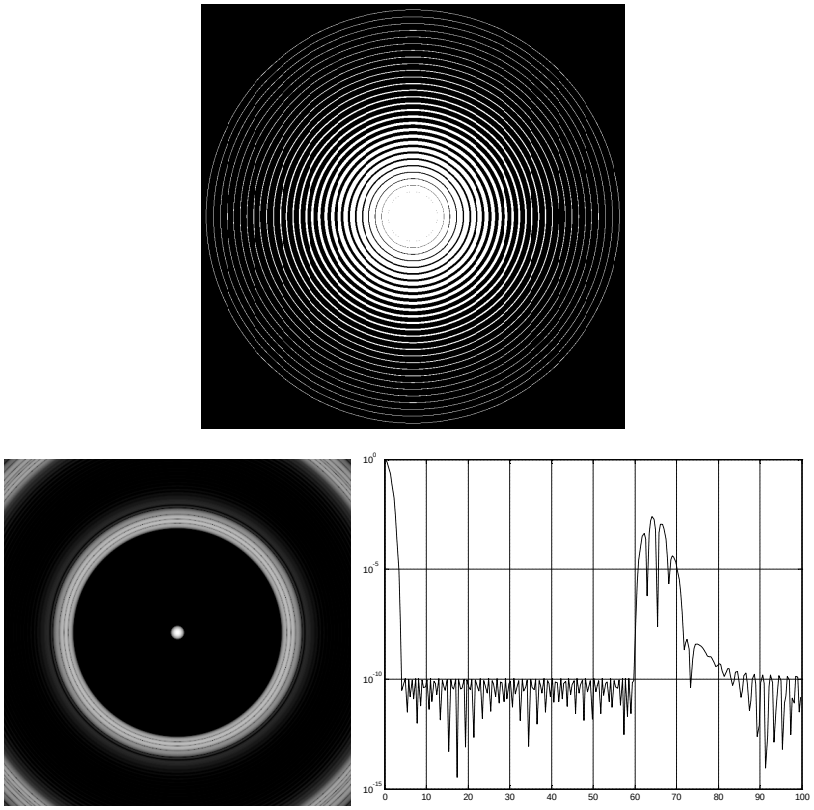

FIGURE 5. The optimal concentric ring pupil mask. The inner working angle is $\rho_{\text{iwa}} = 4$. The Airy throughput is 9%.

out, then this support structure changes the shape of the shaped pupil and therefore affects the resulting PSF. We showed in [13] that such masks can be okay but they require a large number of so-called support spiders and again manufacturability becomes an issue. Of course, the other possibility is simply to lay these rings on high quality glass. But even the best glass might not be good enough. Glass always introduces some scatter due to imperfections and even a tiny amount of scatter will destroy the contrast. In the next section, we look for masks that are more manufacturable. As will be seen, there is a trade-off as some of the best properties of the masks persented so far become compromised.

7. Other Masks

Consider a binary apodization (i.e., a mask) consisting of an opening given by

$$A(x,y) = \begin{cases} 1 & |y| \le a(x) \\ 0 & \text{else} \end{cases}$$

We only consider masks that are symmetric with respect to both the x and y axes. Hence, the function $a(\cdot)$ is a nonnegative even function. In such a situation, the electric field $E(\xi,\zeta)$ is given by

$$\begin{aligned} E(\xi,\zeta) &= \int_{-\frac12}^{\frac12}\int_{-a(x)}^{a(x)} e^{2\pi i(x\xi+y\zeta)}dydx \\ &= \begin{cases} \dfrac{2}{\pi}\displaystyle\int_0^{\frac12}\cos(2\pi x\xi)\frac{\sin(2\pi a(x)\zeta)}{\zeta}dx & \text{if } \zeta\neq 0 \\ 4\displaystyle\int_0^{\frac12}\cos(2\pi x\xi)a(x)dx & \text{otherwise.} \end{cases} \end{aligned} \tag{4}$$

Because of the symmetry, it suffices to optimize in the first quadrant only:

$$\begin{aligned} &\text{maximize} && \int_0^{\frac12} 4a(x)dx \\ &\text{subject to} && -10^{-5}E(0,0) \le E(\xi,\zeta) \le 10^{-5}E(0,0), && \text{for } (\xi,\zeta)\in\mathcal{O} \\ & && 0 \le a(x) \le 1/2, && \text{for } 0\le x\le 1/2 \end{aligned}$$

The objective function is the total open area of the mask. The first constraint guarantees 10^{-10} light intensity throughout a specified region of the focal plane, and the remaining constraint ensures that the mask is really a mask.

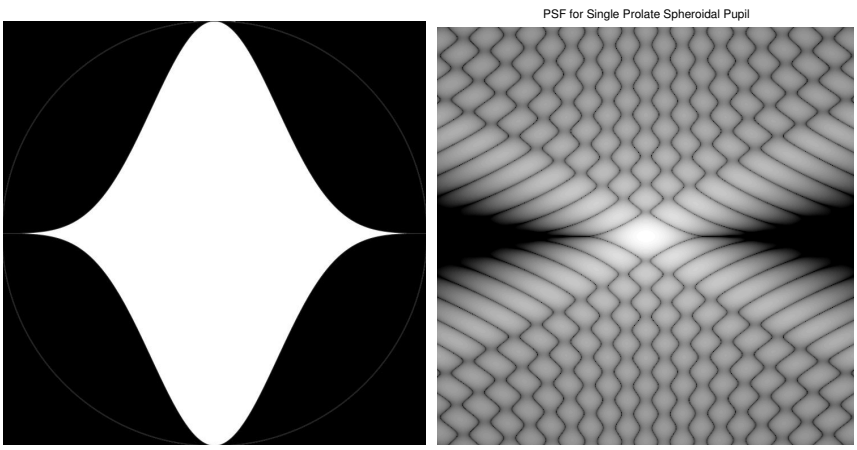

FIGURE 6. *Left.* Single opening shaped pupil. *Right.* Corresponding PSF exhibiting high-contrast along the horizontal axis for $\rho \geq 4$. The Airy throughput is 43%.

7.1. **One Pupil with On-Axis Constraints.** From (4) it follows that the problem is an infinite dimensional linear programming if the set $\mathcal{O}$ is a subset of the x-axis. The resulting shaped pupil and its PSF are shown in Figure 6.

Note that the pupil looks very much like an opening enclosed by a pair of gaussian functions. In fact, Princeton astrophysicist David Spergel realized early on that, given the Fourier transform connection between the pupil plane and the image plane, a gaussian profile in the pupil plane ought to produce a gaussian profile for the electric field in the image plane. As the gaussian function only decreases, there would be no side-lobes and hence with a proper choice of parameters a gaussian solution could be quite a good one. See [10]. Shortly after having this insight, Spergel won a McArthur "genius" prize. The citation mentions this work together with his work in cosmology.

7.2. **Multiple Pupil Mask.** The obvious downside of the single pupil mask is its very narrow region of high-contrast. In fact, at $\rho = \rho_{\text{iwa}}$, the high-contrast region is just a cusp. One would normally imagine taking multiple exposures at different rotation angles in order to survey all around a given star with a mask like this. But, because the high-contrast region is cusp-like, the number of exposures required is really unbounded. Hence, it is necessary to find a solution with a larger dark zone. Asking for a larger dark zone using just one mask opening almost immediately results in an infeasible problem. But, if one considers

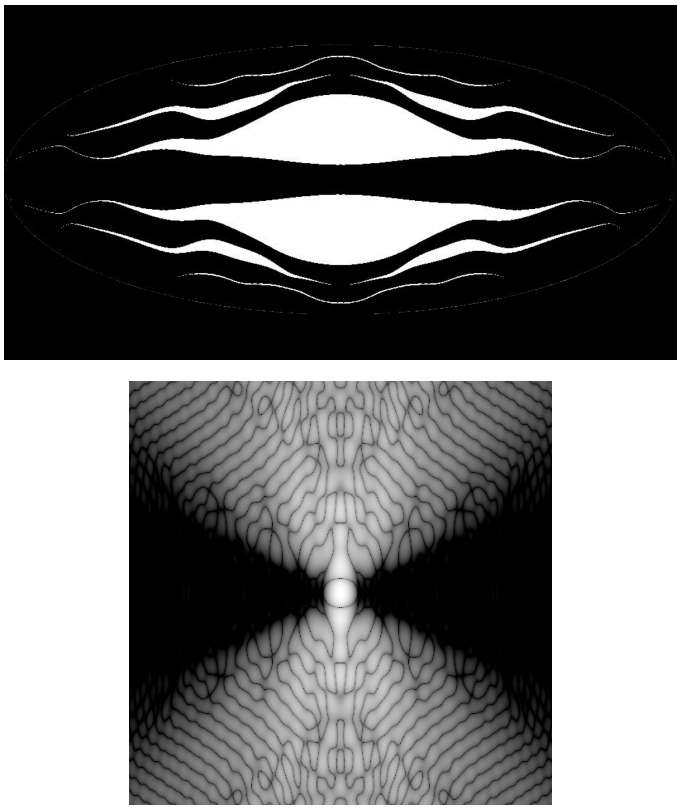

FIGURE 7. *Top.* Six-opening shaped pupil mask. *Bottom.* Corresponding PSF. Note that $\rho_{\text{iwa}} = 4$ and the Airy throughput is 12% (relative to ellipse). Easy to make. Only requires a few rotations.

multiple openings, the added degrees of freedom indeed make it possible to find solutions with larger dark zones in the PSF. Figure 7 shows the result for a six-opening design made to fit into an elliptical aperture.

8. Lab Results

NIST has made us a mask according to the six-opening elliptical design shown in Figure 7. The mask was cut into a silicon wafer using deep ion etching technology. Figure 8 shows microscope images of the mask. Figure 9 shows the mask installed in the Princeton TPF laboratory. Figure 10 shows an early result comparing the actual PSF to the simulated PSF.

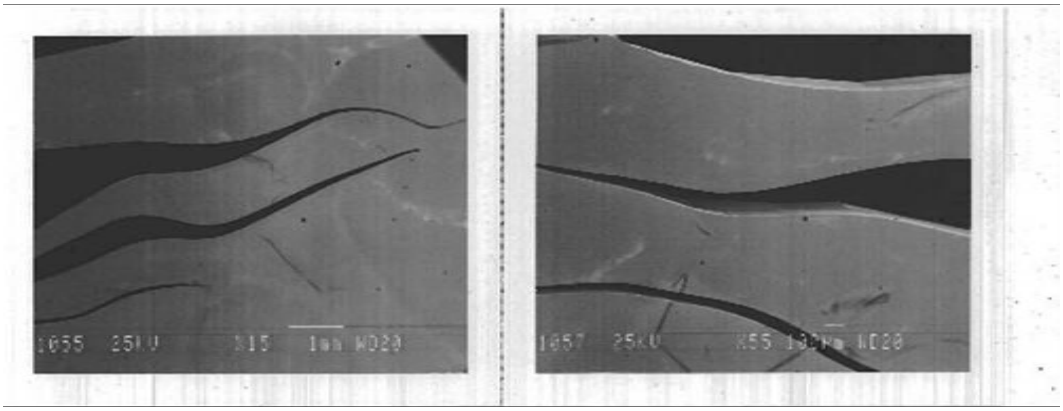

FIGURE 8. Images of an actual mask cut from Silicon.

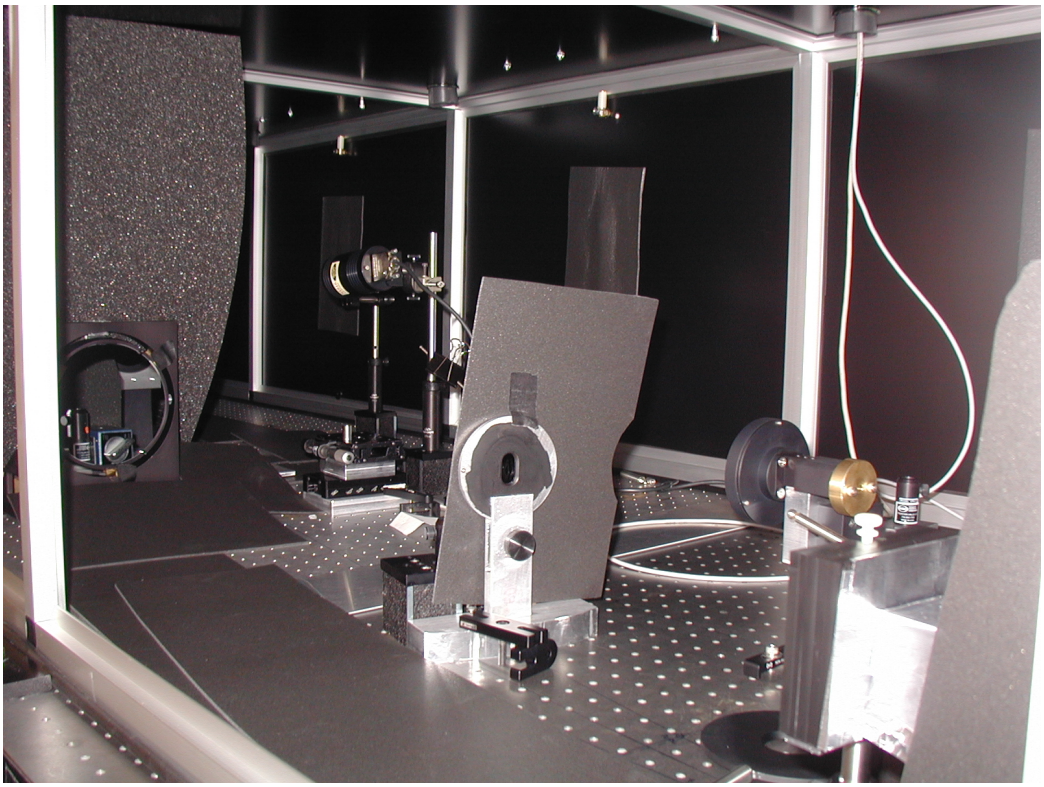

FIGURE 9. The six-opening elliptical mask installed in our laboratory set up designed to generate and then image an artificial (laser generated) star.

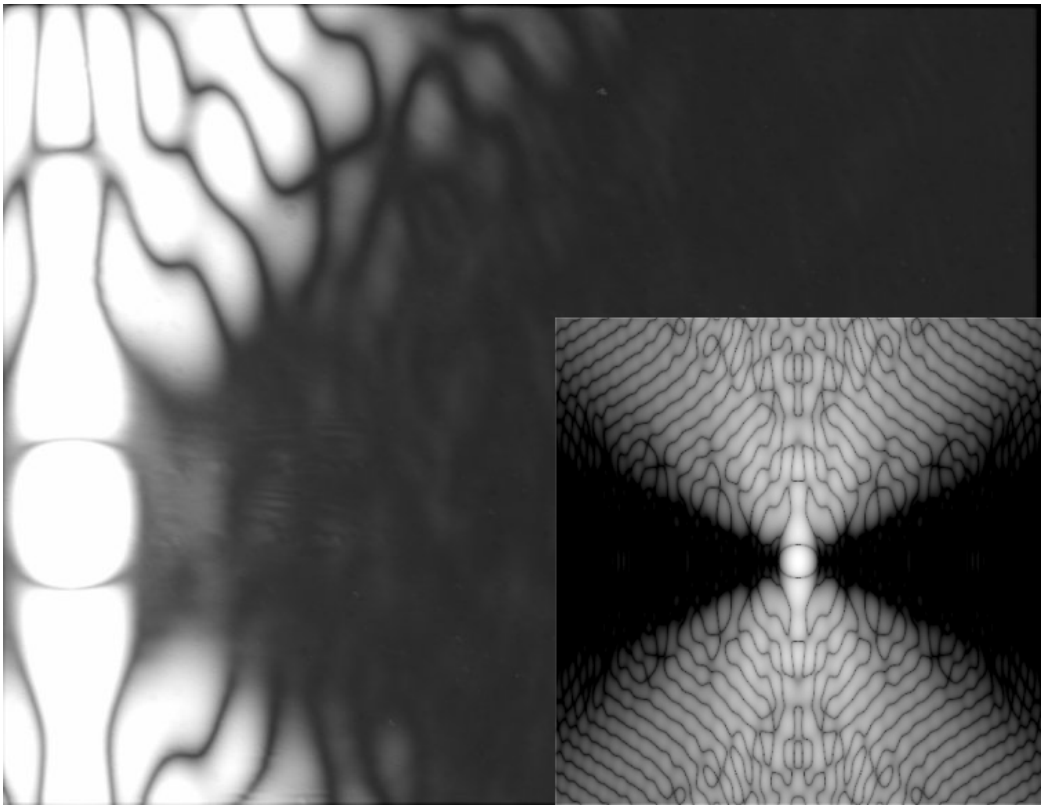

FIGURE 10. Simulation vs. reality. The background image is an actual CCD image. In that image, the brightest pixel represents approximately $1,642,000,000$ recorded photons. The image is a sum of 21 one-hour exposures.

9. Optimization Success Story

The original design concept for TPF involved a cluster of space telescopes flying in formation, collecting infrared light, and beaming that light to a collector spacecraft to be interfered so as to null out the starlight without nulling the planet light. This design is called an *infrared nulling interferometer*. Building such a thing would be an enormously difficult enterprise. The Princeton TPF team has advocated from the beginning that a monolithic visible light telescope with creative diffraction control would be a much more tractible and cost-effective way to search for Earth-like planets around other stars.

At first there was great skepticism about the feasibility of a monolithic visible light telescope. The thinking was that the diffraction control problem was simply too hard to solve. But, using some good

modeling ideas and state-of-the-art optimization tools, we were able to develop the designs presented in this paper (as well as others, see [5, 13, 12, 11]). These designs prompted NASA to rethink the situation and on April 12, 2004, Charles Beichman sent the following letter to the TPF Science Working Group:

> Dear TPF-SWG,
>
> I am writing to inform you of exciting new developments for TPF. As part of the Presidents new vision for NASA, the agency has been directed by the President to *conduct advanced telescope searches for Earth-like planets and habitable environments around other stars.* Dan Coulter, Mike Devirian, and I have been working with NASA Headquarters (Lia LaPiana, our program executive; Zlatan Tsvetanov, our program scientist; and Anne Kinney) to incorporate TPF into the new NASA vision. The result of these deliberations has resulted in the following plan for TPF:
>
> 1. Reduce the number of architectures under study from four to two: **(a) the moderate sized coronagraph, nominally the 4x6 m version now under study**; and (b) the formation flying interferometer presently being investigated with ESA. Studies of the other two options, the large, 10-12 m, coronagraph and the structurally connected interferometer, would be documented and brought to a rapid close.
>
> 2. Pursue an approach that would result in the launch of BOTH systems within the next 10-15 years. The primary reason for carrying out two missions is the power of observations at IR and visible wavelength regions to determine the properties of detected planets and to make a reliable and robust determination of habitability and the presence of life.
>
> 3. **Carry out a modest-sized coronagraphic mission, TPF-C, to be launched around 2014**, to be followed by a formation-flying interferometer, TPF-I, to be conducted jointly with ESA and launched by the end of decade (2020). This ordering of missions is, of course, subject to the readiness of critical technologies and availability of funding. But in the estimation of NASA HQ and the project, the science, the technology, the political will, and the budgetary resources are in place to support this plan.

10. Epilogue

The expensive war, the tax cuts, hurricane Katrina, the re-entry loss of the space shuttle Columbia, the huge cost overruns of the Space Station, and other current events have put an enormous strain on the NASA's ability to pursue its scientific missions. As a consequence, there is currently no new funding for TPF and all the scientists and engineers at JPL who were working on the project have been reassigned to other things. This is an unfortunate turn of events but I'm told these things are fairly common in the real world in which NASA operates. Let's hope the situation reverses sooner rather than later.

Acknowledgement. The author would like to thank his many colleagues at Princeton who have in various ways contributed to the work described here. In particular, he would like to acknowledge Jeremy Kasdin, David Spergel, Mike Littman, Ed Turner and all the students and postdocs for the many many engaging Friday morning conversations.

References

[1] M.P. Bendsøe, A. Ben-Tal, and J. Zowe. Optimization methods for truss geometry and topology design. *Structural Optimization*, 7:141–159, 1994.

[2] J.O. Coleman and D.P. Scholnik. Design of Nonlinear-Phase FIR Filters with Second-Order Cone Programming. In *Proceedings of 1999 Midwest Symposium on Circuits and Systems*, 1999.

[3] A. Cumming, G.W. Marcy, R.P. Butler, and S.S. Vogt. The statistics of extrasolar planets: Results from the keck survey. In D. Deming and S. Seager, editors, *Scientific Frontiers in Research on Extrasolar Planets, ASP Conference Series*, 294, pages 27–30, 2002.

[4] J.K. Ho. Optimal design of multi-stage structures: a nested decomposition approach. *Computers and Structures*, 5:249–255, 1975.

[5] N.J. Kasdin, R.J. Vanderbei, D.N. Spergel, and M.G. Littman. Extrasolar Planet Finding via Optimal Apodized and Shaped Pupil Coronagraphs. *Astrophysical Journal*, 582:1147–1161, 2003.

[6] H. Lebret and S. Boyd. Antenna array pattern synthesis via convex optimization. *IEEE Transactions on Signal Processing*, 45:526–532, 1997.

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

[8] D. Scholnik and J. Coleman. Optimal design of wideband array patterns, 2000.

[9] D. Slepian. Analytic solution of two apodization problems. *Journal of the Optical Society of America*, 55(9):1110–1115, 1965.

[10] D. N. Spergel. A new pupil for detecting extrasolar planets. *astro-ph/0101142*, 2000.

[11] R. J. Vanderbei, N. J. Kasdin, and D. N. Spergel. Checkerboard-mask coronagraphs for high-contrast imaging. *Astrophysical Journal*, 615(1):555, 2004.

[12] R.J. Vanderbei, D.N. Spergel, and N.J. Kasdin. Circularly Symmetric Apodization via Starshaped Masks. *Astrophysical Journal*, 599:686–694, 2003.

[13] R.J. Vanderbei, D.N. Spergel, and N.J. Kasdin. Spiderweb Masks for High Contrast Imaging. *Astrophysical Journal*, 590:593–603, 2003.

[14] S. Wu, S. Boyd, and L. Vandenberghe. Fir filter design via semidefinite programming and spectral factorization, 1996.

Princeton University, Department of Operations Research and Financial Engineering, Princeton, NJ 08544, USA (rvdb@princeton.edu).