

Eliminating Poisson's Spot with Linear Programming

Robert J. Vanderbei

Abstract A leading design concept for NASA's upcoming planet-finding space telescope involves placing an occulter 72,000 km in front of a 4-m telescope. The purpose of the occulter is to block the bright starlight thereby enabling the telescope to take pictures of planets orbiting the blocked star. Unfortunately, diffraction effects prevent a simple circular occulter from providing a sufficiently dark shadow—a specially shaped occulter is required. In this paper, I explain how to reduce this shape-optimization problem to a large-scale linear programming problem that can be solved with modern LP tools.

Key words: linear programming, large scale optimization, optical design, high-contrast imaging, extrasolar planets

1 Introduction

Over the last decade or so, astronomers have discovered hundreds of planets orbiting stars beyond our Sun. Such planets are called *exosolar planets*. The first such discovery Marcy and Butler (1998) and most of the subsequent discoveries involved an indirect detection method called the *radial velocity* (RV) method. This method exploits the fact that in a star-planet system, both the star and the planet orbit about their common center of mass. That is, not only does the planet move in an elliptical “orbit”, so does the star itself. Of course, planets are much less massive than their stars and therefore the center of mass of the system generally lies close to the center of the star itself. Nonetheless, a star with a planet orbiting it will show a sinusoidal wobble which, even if quite small, can be detected spectroscopically using clever modern techniques in spectroscopy assuming, of course, that we are not viewing the star-planet system “face-on”. Clearly, this technique is strongly biased toward very

Robert J. Vanderbei
Princeton University, Princeton NJ e-mail: rvdb@princeton.edu

massive (say Jupiter sized) planets orbiting very close to their star (say closer than Mercury is to our Sun).

It should be apparent that the radial velocity method of indirect planet detection is a rather subtle method. The results should be confirmed by other means and in some cases this has been done. The most successful alternative method for planet detection is the so-called *transit method*. This method only works for star-planet systems that we “see” almost perfectly edge-on. For such systems, from our perspective here on Earth the planet orbits the star in such a manner that at times it periodically passes directly in front of the star causing the star to dim slightly, but measurably, during the “transit”. The first such example discovered Charbonneau et al (2000) involves the star HD209458. Its planet was discovered by the radial velocity method but confirmed by detecting transits. This particular planet is 1.3 times larger than Jupiter in radius and it orbits HD209458 once every 3.5 days. During a transit event, the intensity of the starlight dips by 1.7%.

These discoveries have generated enormous interest among professional astronomers in actually detecting and even imaging Earth-like planets in Earth-like orbits around Sun-like stars. The planets detected by indirect methods tend to be much more massive than Earth and have orbits much closer to their star. Such planets would be rather uninhabitable by life-forms that have evolved here on Earth. The interest is in finding, and more specifically imaging, so-called habitable planets. This is called the *direct detection* problem. But it is a hard problem for three rather fundamental reasons:

1. **Contrast.** In visible light our Sun is 10^{10} times brighter than Earth.
2. **Angular Separation.** A planet at one Sun-Earth distance from its star viewed from a distance of 10 parsecs (32.6 light years) can appear at most 0.1 arcseconds away from the star itself.
3. **Paucity of Targets.** There are less than 100 Sun-like stars within 10 parsecs of Earth.

2 Large Space Telescope

The simplest design concept is just to build a telescope big enough to overcome the three obstacles articulated in the previous section. An Earth-like planet illuminated by a Sun-like star at a distance of about 10 parsecs will be very faint—about as faint as the dimmest objects imaged by the Hubble space telescope, which has a light-collecting mirror that is 2.4 meters in diameter. Hence, the telescope will need to be at least this big.

But, the wave nature of light makes it impossible to focus all of the starlight exactly to a point. Instead, the starlight makes a diffraction pattern that consists of a small (in diameter) concentration of light, called the *Airy disk* surrounded by an infinite sequence of rings of light, called *diffraction rings*, each ring dimmer than the one before. The complete diffraction pattern associated with a point source is called the *Airy pattern*. About 83.8% of the light falls into the Airy disk. Another

7.2% lands on the first diffraction ring, 2.8% on the second diffraction ring, etc. Mathematically, the diffraction pattern is given by the square of the magnitude of the complex electric field at the image plane and the electric field is given by

$$\begin{aligned} E(\rho, \phi) &= \frac{1}{i\lambda f} \int_0^{D/2} \int_0^{2\pi} e^{\frac{i\pi}{\lambda f}(\rho^2 - 2r\rho \cos(\theta - \phi))} r d\theta dr \\ &= \frac{2\pi}{i\lambda f} e^{\frac{i\pi\rho^2}{\lambda f}} \int_0^{D/2} J_0(2\pi r\rho/\lambda f) r dr \\ &= \frac{D}{2i\rho} e^{\frac{i\pi\rho^2}{\lambda f}} J_1(\pi D\rho/\lambda f). \end{aligned}$$

Here, (ρ, ϕ) denotes polar coordinates of a point in the image plane, (r, θ) denotes polar coordinates of a point on the telescope's light-collecting mirror, f is the focal length of the telescope, D is the *aperture* (i.e., diameter) of the mirror, and J_0 and J_1 are the 0-th and 1-st Bessel functions of the first kind. Note that the electric field can in principle depend on both ρ and ϕ but in this case it depends only on ρ because the ϕ dependence disappeared after integrating the inner integral (over θ). When the electric field is manifestly a function of only ρ , we will denote it simply by $E(\rho)$.

From the last expression we see that, except for some scale factors, the Airy pattern for a point source is given by the square of the function J_1 (see Figure 1). Of course, the scale factors are critically important—we need them to figure out how large to make the telescope. In other words, we need to figure out values for D and f that will allow us to “see” a dim planet next to a bright star. With two point sources, a star and a planet, the combined image is obtained by adding the Airy patterns from each source together, after displacing one pattern from the other by an appropriate amount. Simple geometry tells us that the physical separation of the two images in the image plane is just their angular separation on the sky (typically 0.1 arcseconds) in radians times the focal length of the telescope. For example, 0.1 arcseconds is 4.85×10^{-7} radians and so a planet separated from its star by 0.1 arcseconds as viewed from Earth will form an image a distance $\rho = 4.85 \times 10^{-7} f$ from the star's image in the image plane of a telescope of focal length f . At this value of ρ , the intensity of the starlight $I(\rho) = |E(\rho)|^2$ must be reduced by 10 orders of magnitude otherwise the starlight will overwhelm the planet light. In other words, we need

$$|E(\rho)| \leq 10^{-5} |E(0)|.$$

Substituting

$$|E(\rho)| = \frac{D}{2\rho} |J_1(\pi D\rho/\lambda f)|$$

and

$$|E(0)| = \frac{2\pi}{\lambda f} \left(\frac{D}{2}\right)^2$$

into this inequality, we get

$$\frac{D}{2\rho} |J_1(\pi D\rho/\lambda f)| \leq 10^{-5} \times \frac{2\pi}{\lambda f} \left(\frac{D}{2}\right)^2$$

This inequality simplifies to

$$\frac{|J_1(u)|}{u} \leq 10^{-5},$$

where $u = \pi D\rho/\lambda f$. As most programming languages (e.g., C, C++, Matlab, and Fortran) have Bessel functions built in just like the trigonometric functions, it is easy to determine that the inequality is satisfied if and only if $u \geq 1853$. Now, as mentioned before, $\rho/f = 4.85 \times 10^{-7}$. Also for visible light $\lambda \approx 5 \times 10^{-7}$ meters. Hence, we can assume that $\rho/\lambda f$ is about 1 and therefore simplify our contrast criterion to

$$\pi D \geq 1853.$$

In other words,

$$D \geq 590 \text{ meters.}$$

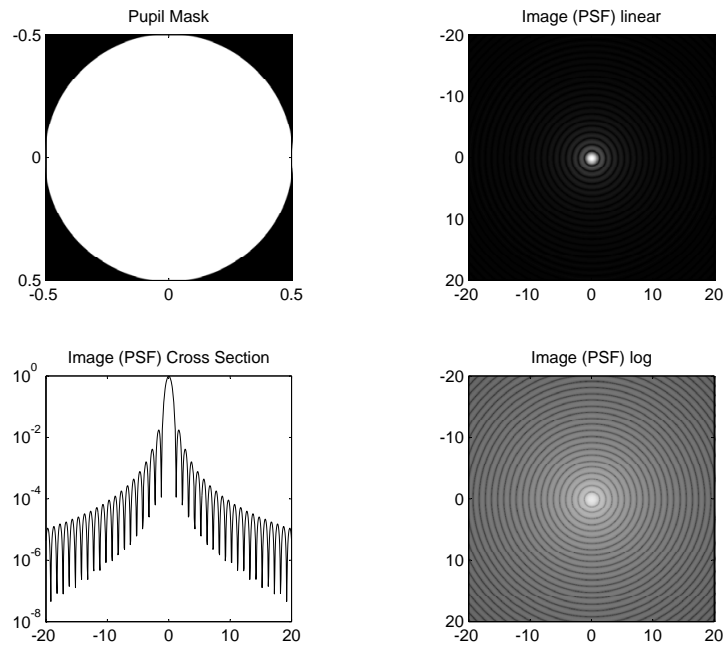


Fig. 1 The upper left figure shows that we are considering a telescope whose mirror has a circular profile (as is typical). The associated Airy pattern is shown in the two right-hand figures. The top one is shown with the usual linear scale. The bottom one is shown logarithmically stretched with everything below 10^{-10} set to black and everything above 1 set to white.

Of course, atmospheric turbulence is a major problem even for moderate sized telescopes. Hence, this telescope will need to be a space telescope. The Hubble space telescope (HST) has a mirror that is 2.4 meters in diameter. To image planets with a large but straight-forward telescope would require a space telescope 250 times bigger in diameter than HST. Of course, cost grows in proportion to the volume of the instrument, so one would imagine that such a telescope, if it were possible to build at all, would cost $250^3 = 16 \times 10^6$ times as much as HST. Such a large telescope will not be built in our lifetime.

3 Space Occulter

In the previous section, we saw that it is not feasible to make a simple telescope that would be capable of imaging Earth-like planets around nearby Sun-like stars. The problem is the starlight. One solution, first proposed by Spitzer (1965), is to prevent the starlight from entering the telescope in the first place. This could be done by flying a large occulting disk out in front of the telescope. If it is far enough away, one could imagine that it would block the starlight but the planet would not be blocked. Of course, as explained already the angular separation between the planet and the star is on the order of 0.1 arcseconds. Also, simple calculations about the expected brightness of the planet suggest that the telescope needs to have an aperture of at least 4 meters in order simply to get enough photons from the planet to take a picture in a reasonable time-frame. So, the occulting disk needs to be at least 4 meters in diameter. To get the planet at 0.1 arcsecond separation to be not blocked by the occulter, means that the occulter has to be positioned about 8250 kilometers in front of the telescope. But, this back-of-the-envelope calculation is based on the simple geometry of ray optics. As before, diffraction effects play a significant role. A 4 meter occulting disk at 8250 km will produce a shadow but the shadow will not be completely dark—some light will diffract into the shadow. In fact, at this small angle, lots of light will diffract into the shadow (see Figure 2). The occulter needs to be made bigger and positioned further away.

The first question to be addressed is: how much bigger and further? To answer that requires a careful diffraction analysis. The formula is similar to the one before. It differs because now there is no focusing element. Instead we have just an obstruction. Also, the disk of the occulter blocks light whereas the disk represented by the light-collecting mirror of a telescope transmits the light. Nonetheless, the formula for the downstream electric field can be explicitly written:

$$\begin{aligned} E(\rho, \phi) &= 1 - \frac{1}{i\lambda z} \int_0^{D/2} \int_0^{2\pi} e^{\frac{i\pi}{\lambda z}(\rho^2 - 2r\rho \cos(\theta - \phi) + r^2)} r d\theta dr \\ &= 1 - \frac{2\pi}{i\lambda z} e^{\frac{i\pi\rho^2}{\lambda z}} \int_0^{D/2} J_0(2\pi r\rho/\lambda z) e^{\frac{i\pi r^2}{\lambda z}} r dr, \end{aligned}$$

where z is the distance between the occulting disk and the pupil of the telescope.

Figure 3 shows that even an occulter 50 meters in diameter 72000 km away does not give a dark enough shadow. Even an occulter 100 times larger and further away than that achieves only 4 orders of magnitude suppression.

4 Poisson's Spot

Note that both Figures 2 and 3 show a bright spot at the center of the shadow. One would expect that a shadow would be darkest at its center and this spot might make one think there is something wrong with the formulae used to compute the shadow. In fact, when Poisson realized in 1818 that the wave description of light would imply the existence of just such a bright spot, he used this as an argument against the theory, not well-accepted at the time, that light is a wave. But Dominique Arago verified the existence of the spot experimentally. This was one of the first, and certainly one of the most convincing, proofs that light is indeed a wave. Today, the bright spot is called either *Poisson's spot* or the *spot of Arago*.

Mathematically, it is easy to see where the spot comes from. Putting $\rho = 0$ in the formula for the electric field, the field can be computed explicitly:

$$\begin{aligned} E(0, \phi) &= 1 - \frac{2\pi}{i\lambda z} \int_0^{D/2} e^{\frac{i\pi r^2}{\lambda z}} r dr \\ &= 1 + \int_0^{\frac{i\pi D^2}{4\lambda z}} e^u du \\ &= e^{\frac{i\pi D^2}{4\lambda z}}. \end{aligned}$$

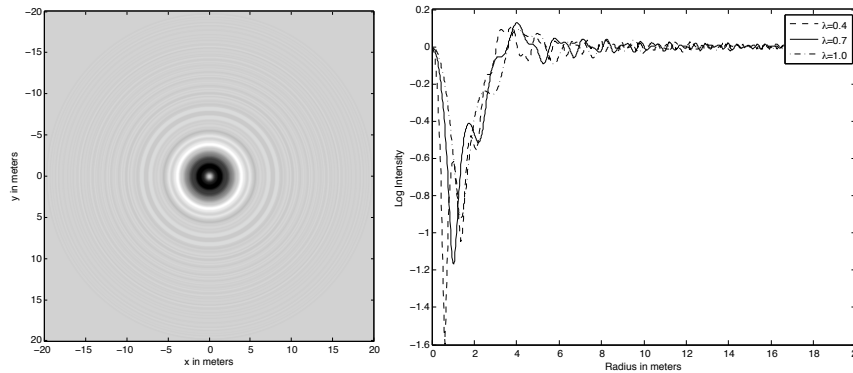


Fig. 2 A circular occulting disk 4 meters in diameter and 8250 km away gives a shadow that provides less than an order of magnitude of starlight suppression in its shadow. The left image is plotted on a linear scale. The right image shows a semilog plot of intensity as a function of radius. Note that the shadow is darkened by less than an order of magnitude. Also, note the small but bright spot at the center—this is *Poisson's spot*.

Since the intensity of the light is the magnitude of the electric field squared, we see that the intensity at the center is unity.

5 Apodized Space Occulter

Conventional wisdom, developed over centuries, is that unpleasant diffraction effects such as Poisson's spot can be mitigated by "softening" hard edges. This suggests replacing the disk occulter, which is assumed to be completely opaque to its outer edge, with a translucent occulter, opaque in the center but becoming progressively more transparent as one approaches the rim (Copi and Starkman (2000) were the first to propose this). Such variable transmissivity is called *apodization*. The question is: can we define an apodization that eliminates Poisson's spot and in addition guarantees a very dark shadow (10 orders of magnitude of light suppression) of at least 4 meters diameter using an apodized occulting disk that is not too large and reasonably close by (if anything flying tens of thousands of kilometers away can be considered close by)?

Let $A(r)$ denote the level of attenuation at radius r along the occulter. Then the formula for the electric field has to be modified to this formula:

$$E(\rho) = 1 - \frac{2\pi}{i\lambda z} e^{\frac{i\pi\rho^2}{\lambda z}} \int_0^{D/2} J_0(2\pi r\rho/\lambda z) e^{\frac{i\pi r^2}{\lambda z}} A(r) r dr,$$

We can formulate the following optimization problem:

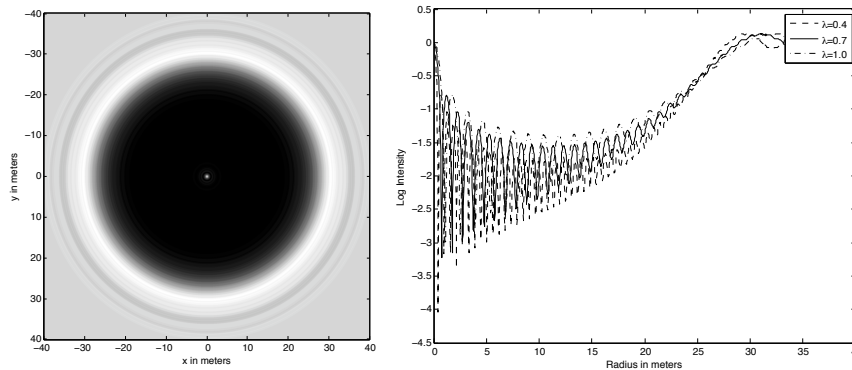


Fig. 3 A circular occulting disk 50 meters in diameter and 72,000 km away gives a shadow that provides only a few orders of magnitude of starlight suppression in its shadow. The left image is plotted on a linear scale. The right image shows a semilog plot of intensity as a function of radius. Note that the shadow is only about 2 orders of magnitude suppressed, a far cry from the 10 orders needed for planet finding. Also, note that Poisson's spot is again visible.

$$\begin{aligned} & \text{minimize} && \gamma \\ & \text{subject to} && |E(\rho)| \leq \gamma, \quad \text{for } \rho \in [0, 3], \lambda \in [0.4, 1.0] \times 10^{-6} \end{aligned}$$

All lengths are in meters. The wavelength λ varies over visible and near infra-red wavelengths. The radius of the deep shadow is 3 meters to leave some wiggle-room for a 4 meter telescope to position itself. We have chosen the apodization function A to depend only on radius r . Hence, the electric field depends only on radius too. We have therefore written $E(\rho)$ for the electric field at radius ρ . The variables in the optimization are the suppression level γ and the apodization function A . Since the electric field E depends linearly on the apodization function A , the problem is an infinite-dimensional second-order cone-programming problem.

We can make it an infinite dimensional linear programming problem if we replace the upper bound on the magnitude of the electric field E with upper and lower bounds on the real and imaginary parts of E :

$$\begin{aligned} & \text{minimize} && \gamma \\ & \text{subject to} && -\gamma \leq \Re(E(\rho)) \leq \gamma \quad \text{for } \rho \in [0, 3], \lambda \in [0.4, 1.0] \times 10^{-6} \\ & && -\gamma \leq \Im(E(\rho)) \leq \gamma \quad \text{for } \rho \in [0, 3], \lambda \in [0.4, 1.0] \times 10^{-6}. \end{aligned}$$

Finally, we can discretize the set of shadow radii $[0, 3]$ into 150 evenly spaced points, the set of occulter radii $[0, D]$ into 4000 evenly spaced points, and the wavelength band $[0.3, 1.1] \times 10^{-6}$ in increments of 10^{-7} and replace the integral defining E in terms of A with the appropriate Riemann sum to get a large-scale (but finite) linear programming problem:

$$\begin{aligned} & \text{minimize} && \gamma \\ & \text{subject to} && -\gamma \leq \Re(E(\rho)) \leq \gamma \quad \text{for } \rho \in [0, 3], \lambda \in [0.4, 1.0] \times 10^{-6} \\ & && -\gamma \leq \Im(E(\rho)) \leq \gamma \quad \text{for } \rho \in [0, 3], \lambda \in [0.4, 1.0] \times 10^{-6} \\ & && A'(r) \leq 0, \quad \text{for } r \in [0, D/2] \\ & && -d \leq A''(r) \leq d, \quad \text{for } r \in [0, D/2]. \end{aligned}$$

For the linear programming problem, we add additional constraints to stipulate that the function A be monotonically decreasing and have a second derivative (difference) that remains between given upper and lower bounds. These additional constraints help to ensure that the solution to the discrete problem is a good approximation to the solution to the underlying infinite dimensional problem.

The linear programming problem was formulated in AMPL (Fourer et al (1993)) and solved using LOQO (Vanderbei (1999)). The AMPL model is shown in Figure 4. The linear programming problem has 4001 variables and 8645 constraints. Using its default stopping criteria, LOQO solves the problem in 51 iterations which takes 224 seconds on a current generation Macbook Pro. The optimal apodized occulter is shown in Figure 5.

By modern computational standards, this is not a very large linear programming problem. But, there are several issues at play here not the least of which is that this particular instance is only a discrete approximation to a much larger, in fact infinite dimensional, problem. It is of paramount importance to verify that the discretization provides a meaningful solution to the real problem. To do this, we have code that uses a spline to resample the apodization function at a finer level and then compute the electric field at the telescope pupil also at a finer level of discretization. We also need to check at wavelengths falling between the discrete set at which we have stipulated high contrast. All of these checks have been performed. It turns out that the discrete sampling of wavelengths creates the biggest errors. Hence, at times we have run the model with twice as many wavelength sample points. But, we should also bear in mind that there are many other issues. For one, we are still in the design phase and are therefore considering many different designs. This occulter becomes transparent beyond 0.075 arcseconds. Is that adequate? Do we need to be able to detect planets closer to their star than that? Is 10^{-10} a strong enough level of contrast? Is it stronger than it needs to be? These issues are still being debated by the scientists who have a feel for what is needed. Also, what about sensitivity? There will undoubtedly be manufacturing error. How precisely must the occulter be made to ensure the design level of performance? This is one of the most critical questions and will be addressed in more detail in the next section. Finally, how accurately can two spacecraft separated by tens of thousands of kilometers be positioned with respect to each other? Can we get the telescope centered in the shadow and can we keep it there long enough to take a multi-hour exposure? These are difficult engineering questions. Optimization plays a critical role in assessing various trade-offs. We have produced a large number of scenarios for consideration. The one presented here is currently the one that seems to provide the best balance of the various trade-offs that must be considered. But, surely next month a slightly different version will be considered better.

6 Petalized Space Occulter

The apodized occulter design given in the previous section provides the starlight suppression necessary to image an Earth-like planet. But, sensitivity analyses show that it is not possible to build such an occulter with adequate precision. It is shown in Vanderbei et al (2007) that a binary, i.e. unapodized, non-circular occulter that consists of multiple petals designed so that the average level of opacity at radius r matches $A(r)$ can provide a dark shadow as long as the number of petals is large enough. Simulations show that 16 petals is adequate—see Figure 6. This is the current baseline design for the occulter concept for planet finding. While it may seem like a daunting task to build such a large occulter (50 meters tip-to-tip) with the required precision (millimeter) and fly it 72000 kilometers in front of a 4 meter telescope with a positioning accuracy of 1 meter, this design is currently considered

the most promising concept for NASA's eventual Terrestrial Planet Finder space telescope—see Simmons et al (2003) and Cash (2006).

References

- Cash W (2006) Detection of earth-like planets around nearby stars using a petal-shaped occulter. *Nature* 442:51–53
- Charbonneau D, Brown T, Latham D, Mayor M (2000) Detection of planetary transits across a sun-like star. *The Astrophysical Journal* 529:45–48
- Copi C, Starkman G (2000) The big occulting steerable satellite (boss). *The Astrophysical Journal* 532:581–592
- Fourer R, Gay D, Kernighan B (1993) *AMPL: A Modeling Language for Mathematical Programming*. Scientific Press
- Marcy G, Butler P (1998) Detection of extrasolar giant planets. *Annual Review of Astronomy and Astrophysics* 36:57–97
- Simmons W, Kasdin N, Vanderbei R, Cash W (2003) System concept design for the new worlds observer. *Bulletin of the American Astronomical Society* 35:1205
- Spitzer L (1965) The beginnings and future of space astronomy. *Am Scientist* 50(3):473
- Vanderbei R (1999) LOQO user's manual—version 3.10. *Optimization Methods and Software* 12:485–514
- Vanderbei R, Cady E, Kasdin N (2007) Optimal occulter design for finding extrasolar planets. *Astrophysical Journal* 665(1):794–798

```

function J0;

param pi := 4*atan(1);
param pi2 := pi/2;
param N := 4000;      # discretization parameter at occulter plane
param M := 150;      # discretization parameter at telescope plane
param c := 25.0;     # overall radius of occulter
param z := 72000e+3; # distance from telescope to apodized occulter
param lambda {3..11}; # set of wavelengths
param rho1 := 25;    # max radius investigated at telescope's pupil plane
param rho_end := 3; # radius below which high contrast is required

# a few convenient shorthands
param lz {j in 3..11 by 1.0} := lambda[j]*z;
param pi2lz {j in 3..11 by 1.0} := 2*pi/lz[j];
param pilz {j in 3..11 by 1.0} := pi/lz[j];

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

var A {r in Rs} >= 0, <= 1;

set Rhos ordered;
let Rhos := setof {j in 0..M} (j/M)*rho1;

var contrast >= 0;

var Ereal {j in 3..11 by 1.0, rho in Rhos} =
    1-pi2lz[j]*
    sum {r in Rs} sin(pilz[j]*(r^2+rho^2))*A[r]*J0(-pi2lz[j]*r*rho)*r*dr;
var Eimag {j in 3..11 by 1.0, rho in Rhos} =
    pi2lz[j]*
    sum {r in Rs} cos(pilz[j]*(r^2+rho^2))*A[r]*J0(-pi2lz[j]*r*rho)*r*dr;

minimize cont: contrast;

subject to main_real_neg {j in 3..11 by 1.0, rho in Rhos: rho < rho_end}:
    -contrast <= Ereal[j,rho];
subject to main_real_pos {j in 3..11 by 1.0, rho in Rhos: rho < rho_end}:
    Ereal[j,rho] <= contrast;
subject to main_imag_neg {j in 3..11 by 1.0, rho in Rhos: rho < rho_end}:
    -contrast <= Eimag[j,rho];
subject to main_imag_pos {j in 3..11 by 1.0, rho in Rhos: rho < rho_end}:
    Eimag[j,rho] <= contrast;

subject to monotone {r in Rs: r>first(Rs)}: A[prev(r)] >= A[r];
subject to smooth {r in Rs: r>first(Rs) && r<last(Rs)}:
    -0.044 <= (A[next(r)]-2*A[r]+A[prev(r)])/dr^2 <= 0.044;

let {j in 3..11 by 1.0} lambda[j] := (j-0.5)*1e-7;

solve;

printf: "%12e, %f, %f \n", contrast^2, 2*c, z/1000;

printf {r in Rs}: "%10.7f %10.7f \n", r, A[r] > "A";

```

Fig. 4 AMPL model for optimal apodization.

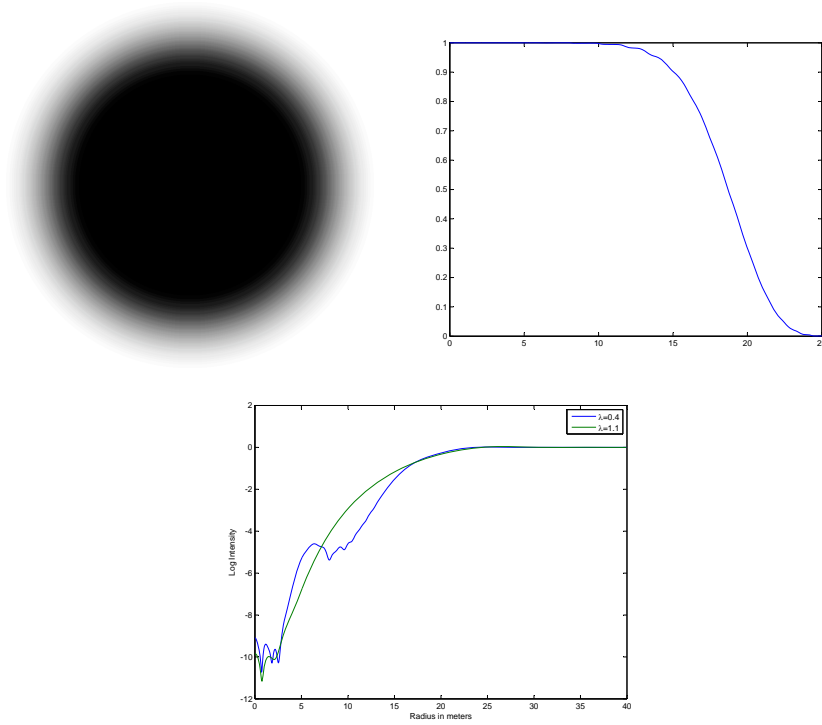


Fig. 5 An apodized circular occulting disk 50 meters in diameter and 72,000 km away gives that achieves roughly 10 orders of magnitude suppression over the entire desired band of wavelengths. The top-left image shows the apodized occulter. The top-right graph shows its attenuation as a function of radius. The bottom graph shows the intensity at the pupil plane at the two extreme wavelengths 0.4 and 1.0 microns.

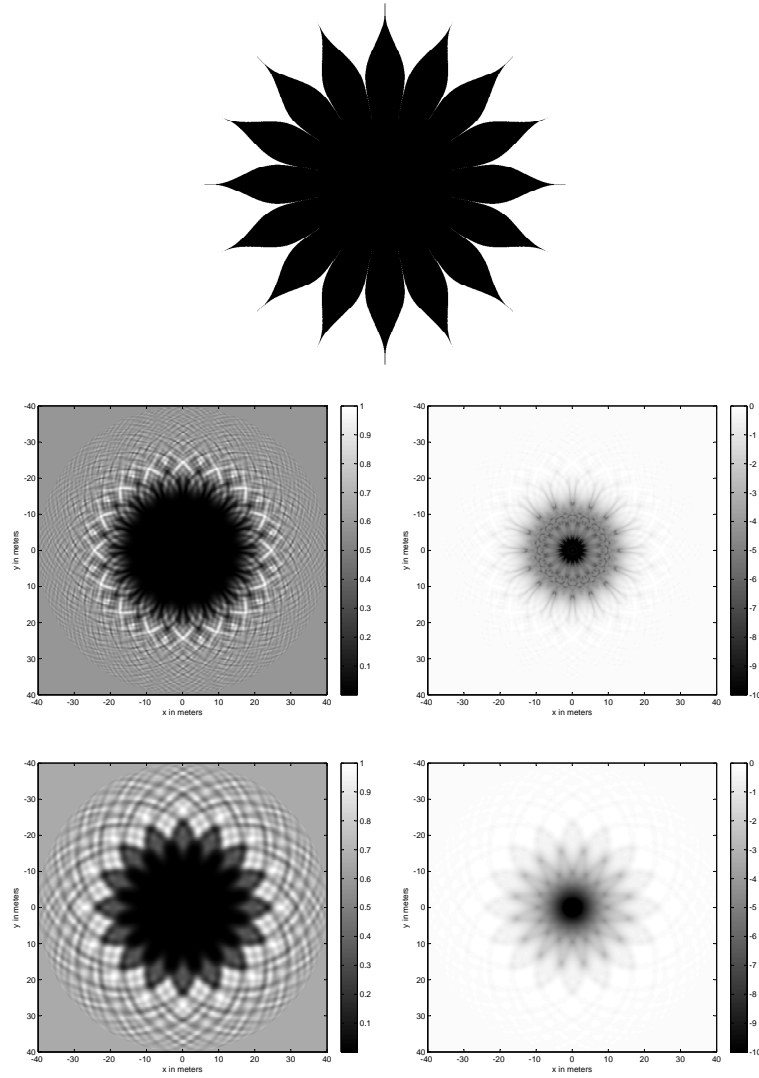


Fig. 6 The top image shows a 16-petal approximation to the optimal apodization. The second row shows the downstream shadow for $\lambda = 0.4$ microns using both a linear (on the left) and a log (on the right) stretch. The third row shows the downstream shadow for $\lambda = 1.0$ microns.