Engineering Applications of Nonlinear Optimization

Robert Vanderbei

January 16, 2002

Operations Research and Financial Engineering, Princeton University

http://www.princeton.edu/∼rvdb

1 Outline

- LOQO: An Interior-Point Code for NLP
- Finite-Impulse-Response (FIR) Filter Design
- Antenna Array Design
- Telescope Pupil-Shape Optimization
- Stable Periodic Orbits for the Equal-Mass n-Body Problem
- Parking Orbits for Satellites

2 LOQO: An Interior-Point Code for NLP

LOQO solves problems in the following form:

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & b \leq h(x) \leq b + r, \\ & l \leq x \leq u \end{array}$$

The functions $f(x)$ and $h(x)$ must be twice differentiable (at least at points of evaluation).

The standard interior-point paradigm is used:

- Add slacks.
- Replace nonnegativities with barrier terms in objective.
- Write first-order optimality conditions.
- Rewrite optimality conditions in primal-dual symmetric form.
- Use Newton's method to get search directions...

3 Interior-Point Paradigm Continued

- Use Newton's method to get search directions:

$$\begin{bmatrix} -H(x,y) - D & A^T(x) \\ A(x) & E \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} = \begin{bmatrix} \nabla f(x) - A^T(x) y \\ -h(x) + \mu Y^{-1} e \end{bmatrix}.$$

  Here, D and E are diagonal matrices involving slack variables,

$$H(x,y) = \nabla^2 f(x) - \sum_{i=1}^{m} y_i \nabla^2 h_i(x) + \lambda I, \qquad \text{and} \qquad A(x) = \nabla h(x),$$

  where λ is chosen to ensure appropriate descent properties.

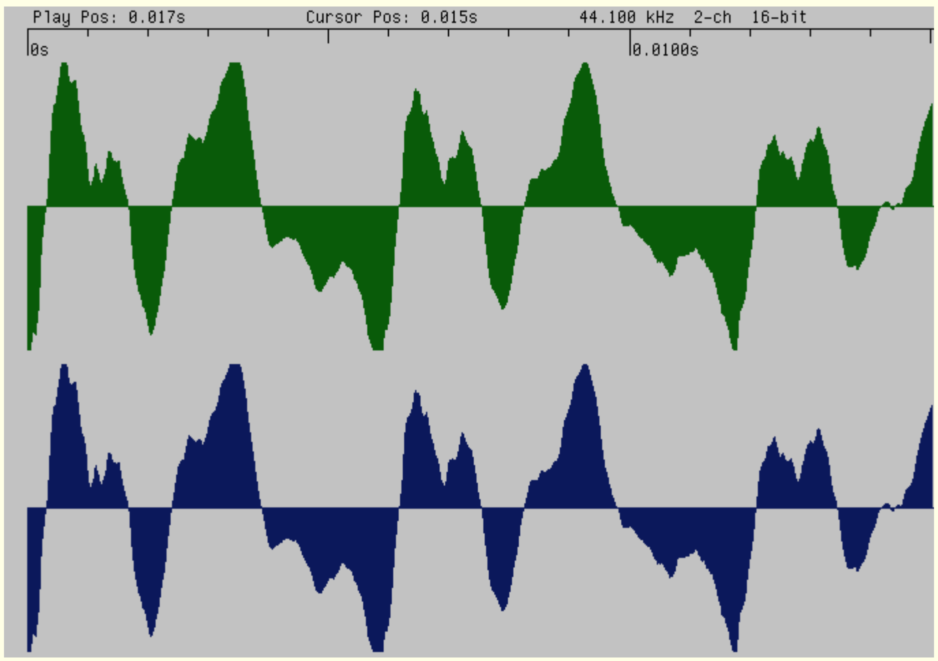
- Compute step lengths to ensure positivity of slack variables.
- Shorten steps further to ensure a reduction in either infeasibility or in the barrier function—a myopic, or Markov, filter. (N.B.: We no longer use a merit function.)
- Step to new point and repeat.

Outline

- LOQO: An Interior-Point Code for NLP
- Finite-Impulse-Response (FIR) Filter Design
- Antenna Array Design
- Telescope Pupil-Shape Optimization
- Stable Periodic Orbits for the Equal-Mass n-Body Problem
- Parking Orbits for Satellites

5 Finite Impulse Response (FIR) Filter Design

- Audio is stored digitally in a computer as a stream of short integers: u_k, $k \in \mathbb{Z}$.
- When the music is played, these integers are used to drive the displacement of the speaker from its resting position.
- For CD quality sound, 44100 short integers get played per second per channel.



0	-32768	8	-23681	16	12111
1	-32768	9	-18449	17	17311
2	-32768	10	-11025	18	21311
3	-30753	11	-6913	19	23055
4	-28865	12	-4337	20	23519
5	-29105	13	-1329	21	25247
6	-29201	14	1743	22	27535
7	-26513	15	6223	23	29471

6 FIR Filter Design—Continued

- A finite impulse response (FIR) filter takes as input a digital signal and convolves this signal with a finite set of fixed numbers $h_{-n}, \ldots, h_n$ to produce a filtered output signal:

$$y_k = \sum_{i=-n}^{n} h_i u_{k-i}.$$

- Sparing the details, the output power at frequency ν is given by

$$|H(\nu)|^2$$

where

$$H(\nu) = \sum_{k=-n}^{n} h(k) e^{2\pi i k \nu},$$

- Similarly, the mean squared deviation from a flat frequency response over a frequency range, say $\mathcal{L} \subset [0, 1]$, is given by

$$\frac{1}{|\mathcal{L}|} \int_{\mathcal{L}} \left| H(\nu) - 1 \right|^2 d\nu$$

7 Filter Design: Woofer, Midrange, Tweeter

minimize ρ

subject to $\int_0^1 \left(H_w(\nu) + H_m(\nu) + H_t(\nu) - 1\right)^2 d\nu \leq \epsilon$

$$\left(\frac{1}{|W|}\int_W H_w^2(\nu)d\nu\right)^{1/2} \leq \rho \qquad W = [.2, .8]$$

$$\left(\frac{1}{|M|}\int_M H_m^2(\nu)d\nu\right)^{1/2} \leq \rho \qquad M = [.4, .6] \cup [.9, .1]$$

$$\left(\frac{1}{|T|}\int_T H_t^2(\nu)d\nu\right)^{1/2} \leq \rho \qquad T = [.7, .3]$$

where

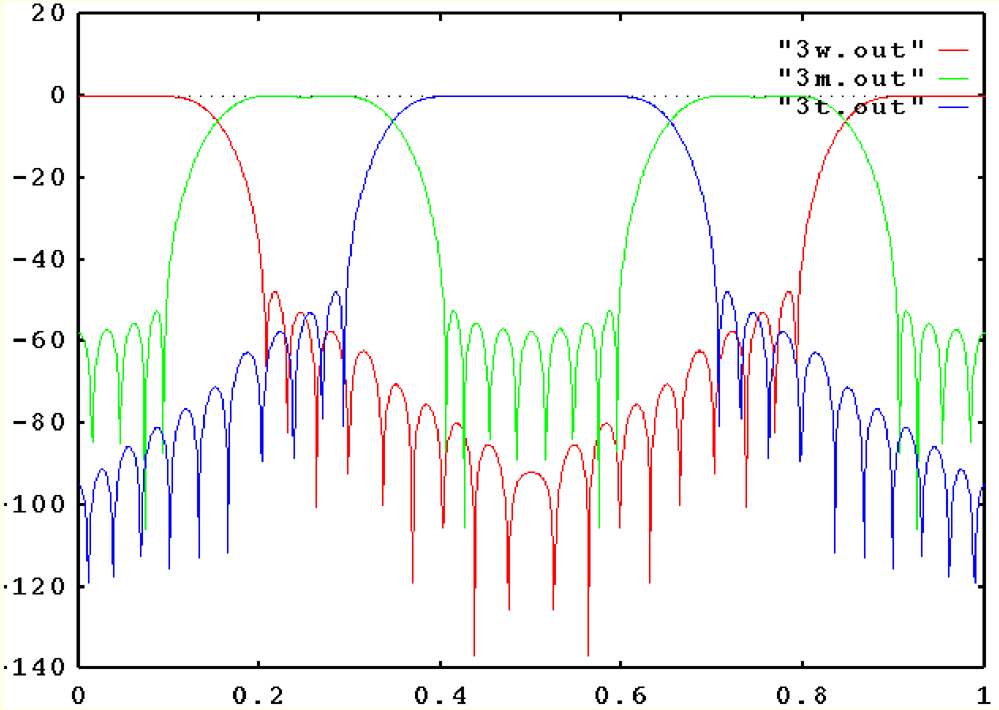
$$H_i(\nu) \quad = \quad h_i(0) + 2\sum_{k=1}^{n-1} h_i(k)\cos(2\pi k\nu), \quad i = W, M, T$$

$h_i(k)$ = filter coefficients, i.e., **decision variables**

Specific Example: Pink Floyd's "Money"

filter length: $n = 14$

integral discretization: $N = 1000$



Demo: orig-clip woofer midrange tweeter reassembled

Ref: J.O. Coleman, U.S. Naval Research Laboratory,

CISS98 paper available: engr.umbc.edu/∼jeffc/pubs/abstracts/ciss98.html

Outline

- LOQO: An Interior-Point Code for NLP

- Finite-Impulse-Response (FIR) Filter Design

- Antenna Array Design

- Telescope Pupil-Shape Optimization

- Stable Periodic Orbits for the Equal-Mass n-Body Problem

- Parking Orbits for Satellites

10 Antenna Array Design

- Given: an array (linear or 2-D) of radar antennae.
- An incoming signal generates an output signal at each antenna.
- A linear combination of the signals is made to produce one total output signal.
- Coefficients of the linear combination can be chosen to accentuate and/or attenuate the output signal’s strength as a function of the input signal’s source direction.

11 2-D Antenna-Array Design Problem

$$\text{minimize} \quad \int_S |A(p)|^2 ds$$

$$\text{subject to} \quad A(p_0) = 1,$$

where

$$A(p) = \sum_{l \in \{\text{array elements}\}} w_l e^{-2\pi i p \cdot x_l}, \quad p \in S$$

w_l = complex-valued **design weight** for array element l

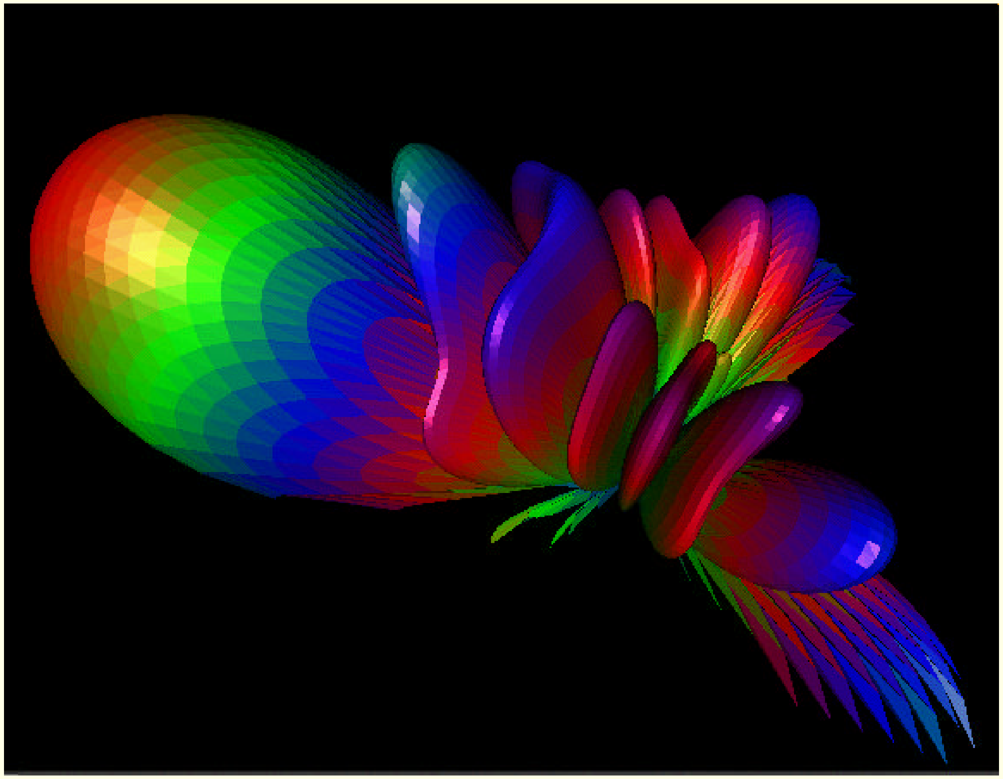
S = subset of unit hemisphere: sidelobe directions

x_l = spatial coord vector for array element l

p_0 = “look” direction

Specific Example: Hexagonal Lattice of 61 Elements

$$\begin{aligned} \rho &= -20 \text{ dB} = 0.01 \\ S &= \text{889 points outside } 20^\circ \text{ from look direction} \\ p_0 &= 40^\circ \text{ from zenith} \end{aligned}$$



Outline

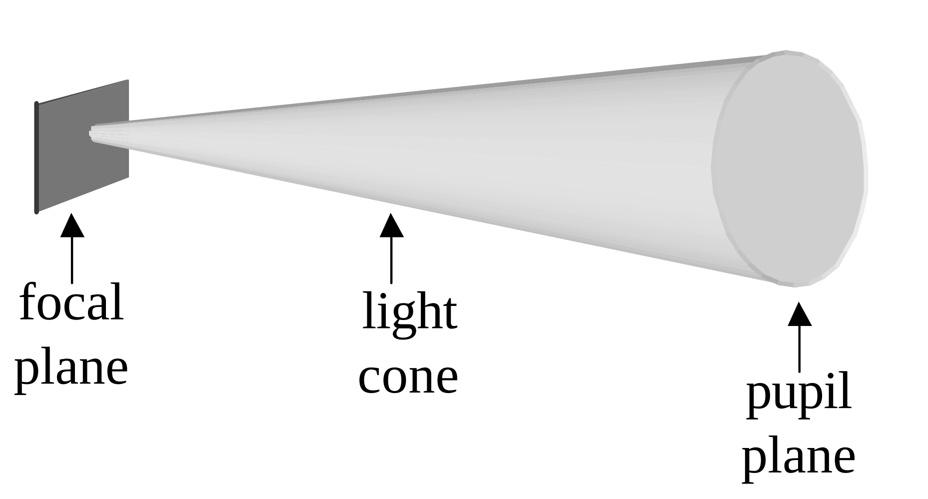
- LOQO: An Interior-Point Code for NLP
- Finite-Impulse-Response (FIR) Filter Design
- Antenna Array Design
- Telescope Pupil-Shape Optimization
- Stable Periodic Orbits for the Equal-Mass n-Body Problem
- Parking Orbits for Satellites

A Terrestrial Planet Finder Space Telescope

There is an effort currently under way to design and build a space telescope that will be able to “see” planets around nearby stars (other than the Sun).

Consider a telescope. Light enters the front of the telescope. This is called the pupil plane.

The telescope focuses all the light passing through the pupil plane from a given direction at a certain point on the focal plane, say $(0, 0)$.



However, the wave nature of light makes it impossible to concentrate all of the light at a point. Instead, a small disk, called the Airy disk, with diffraction rings around it appears.

These diffraction rings are bright relative to any planet that might be orbiting a nearby star and so would completely hide the planet. The Sun, for example, would appear 10^{10} times brighter than the Earth to a distant observer.

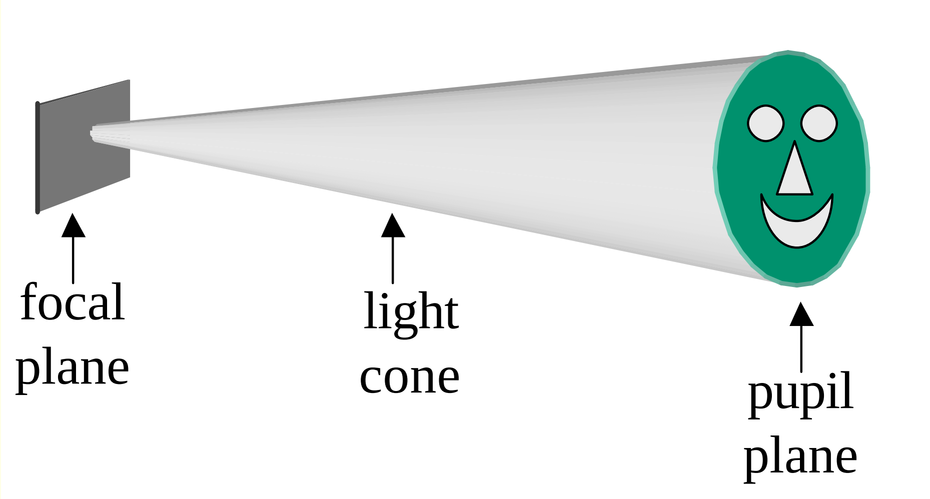
By placing a mask over the pupil, one can design the shape and strength of the diffraction rings. The problem is to find an optimal shape so as to put a very deep null very close to the Airy disk.

14 A Terrestrial Planet Finder Space Telescope

There is an effort currently under way to design and build a space telescope that will be able to "see" planets around nearby stars (other than the Sun).

Consider a telescope. Light enters the front of the telescope. This is called the pupil plane.

The telescope focuses all the light passing through the pupil plane from a given direction at a certain point on the focal plane, say $(0, 0)$.

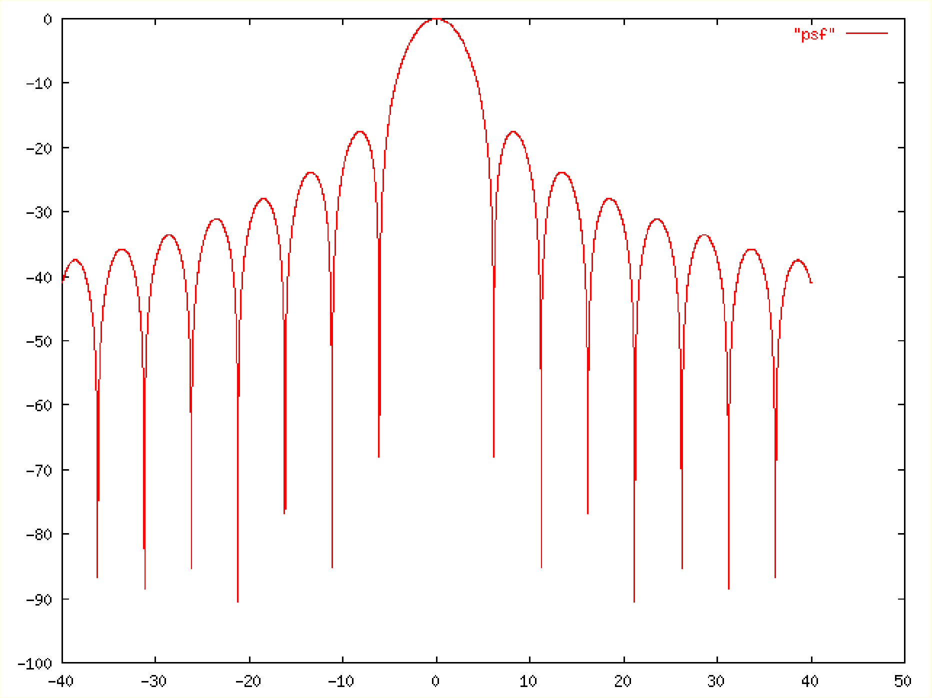


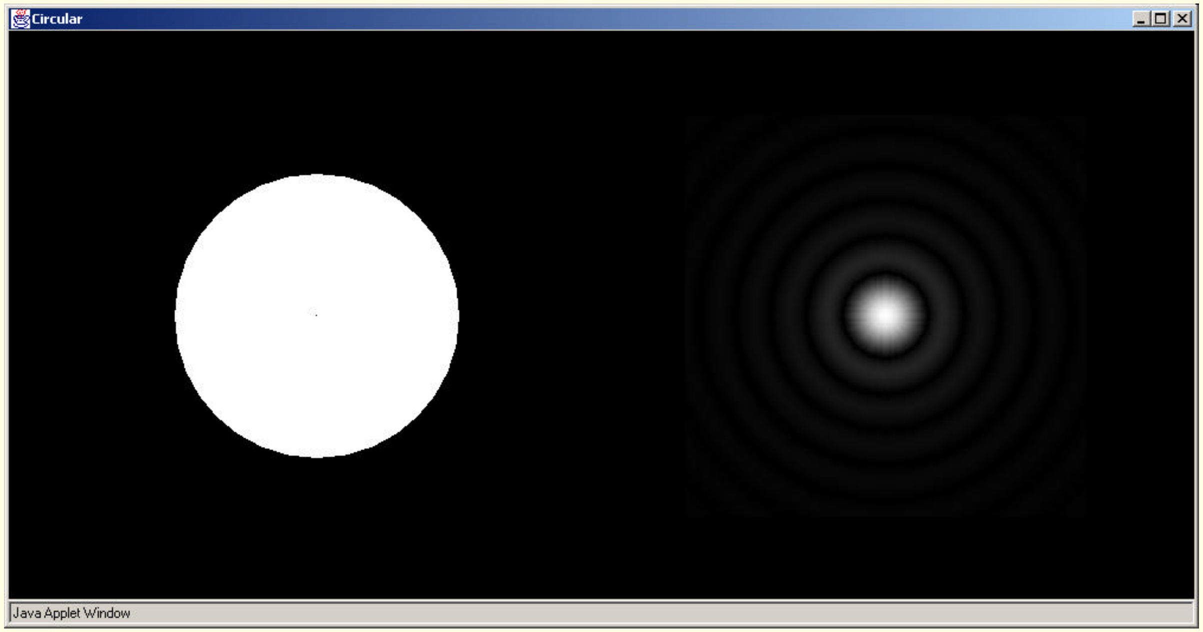
However, the wave nature of light makes it impossible to concentrate all of the light at a point. Instead, a small disk, called the Airy disk, with diffraction rings around it appears.

These diffraction rings are bright relative to any planet that might be orbiting a nearby star and so would completely hide the planet. The Sun, for example, would appear 10^{10} times brighter than the Earth to a distant observer.

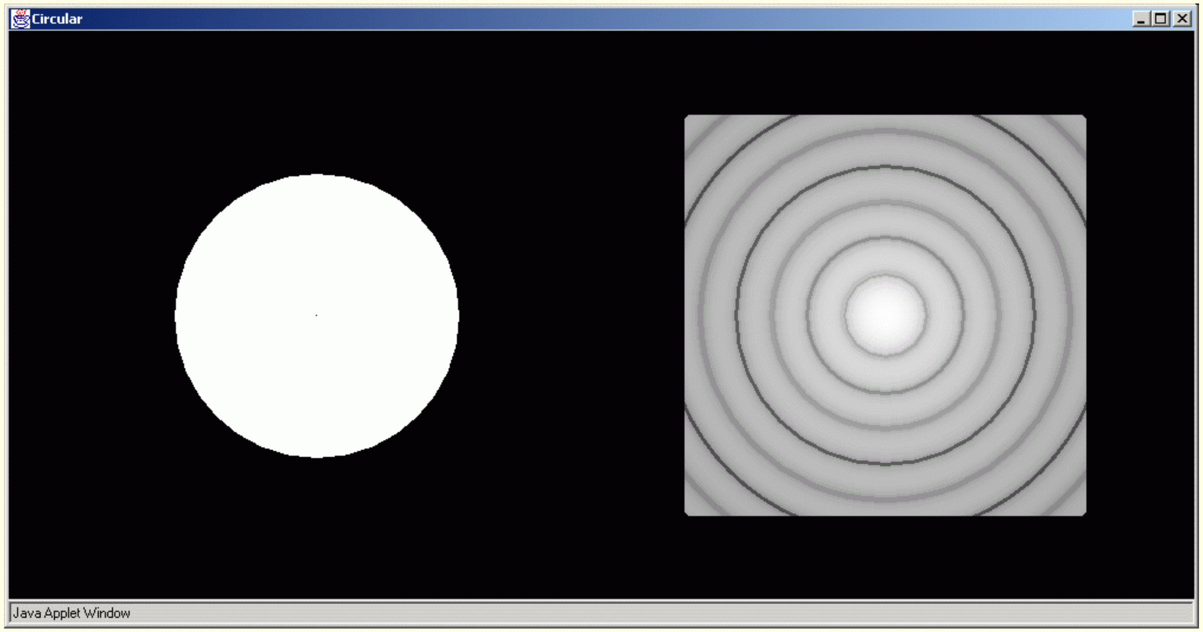
By placing a mask over the pupil, one can design the shape and strength of the diffraction rings. The problem is to find an optimal shape so as to put a very deep null very close to the Airy disk.

15 Airy Disk and Diffraction Rings





Airy Disk and Diffraction Rings—Log Scaling



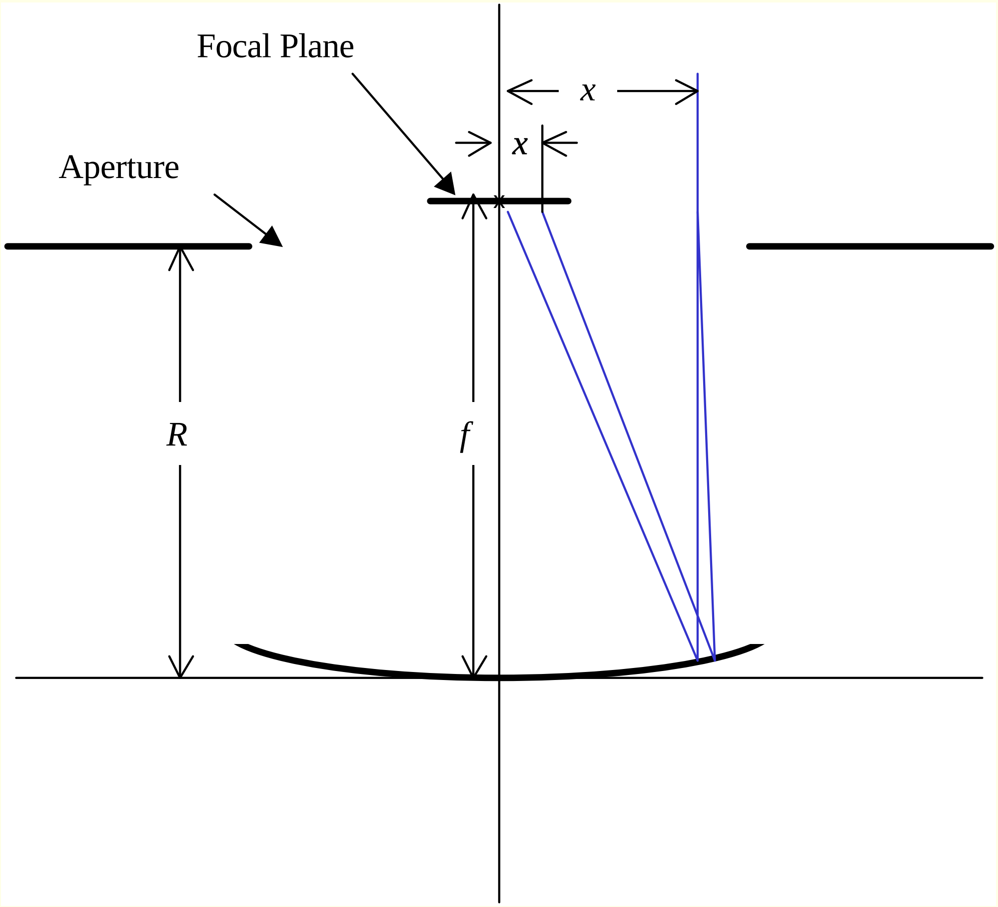
The problem is to find an aperture mask, i.e. a pupil plane mask, that yields a -100 dB null somewhere near the first diffraction ring. A hard problem! Such a null would appear black in this log-scaled image.

17 Diffraction Limited Optics

A coherent plane wave passing the pupil plane produces an interference pattern at the focal plane given by

$$E(\xi,\zeta) = \iint e^{-2\pi i k \Delta_{x,y}(\xi,\zeta)} dy dx$$

where the double integral extends over the open portion of the mask, $k = 1/\lambda$ is the wave number, and $\Delta_{x,y}(\xi,\zeta)$ denotes the difference in the distance from the point (x,y) on the pupil plane to $(0,0)$ and from (x,y) to (ξ,ζ). The formula for $\Delta_{x,y}(\xi,\zeta)$ is complicated:



$$\Delta_{x,y}(\xi,\zeta) \;=\; \sqrt{R^2 - \frac{8Rf}{x^2+y^2+4f^2}(x\xi + y\zeta) + \xi^2 + \zeta^2} - R.$$

18 Exact Optimization Problem

We assume the mask is symmetric about the x-axis and can be represented by a function $A(x)$:

$$E(\xi, \zeta) = \int_{-a/2}^{a/2} \int_{-A(x)}^{A(x)} e^{-2\pi i k \Delta_{x,y}(\xi,\zeta)} dy dx.$$

Here, a is the “width” of the aperture.

The goal is to find the mask function $A(x)$ that minimizes the squared modulus of E over a range of points on the x-axis of the focal plane:

$$\begin{aligned} \text{minimize} \quad & \int_{\xi_0}^{\xi_1} |E(\xi, 0)|^2 d\xi \\ \text{subject to:} \quad & \int_{-a/2}^{a/2} A(x) dx \geq a^2/6 \\ & 0 \leq A(x) \leq a/2 \end{aligned}$$

Note: Kasdin, Littman, & Spergl (2001) showed that the prolate spheroidal wave functions are optimal when $\xi_1 = \infty$. These functions look like Gaussians with compact support.

19 # A Simple Approximation

It is easy to check that the following approximation holds to first order:

$$\mathbf{\Delta}_{x,y}(\xi,\zeta) \approx (x\xi + y\zeta)/f$$

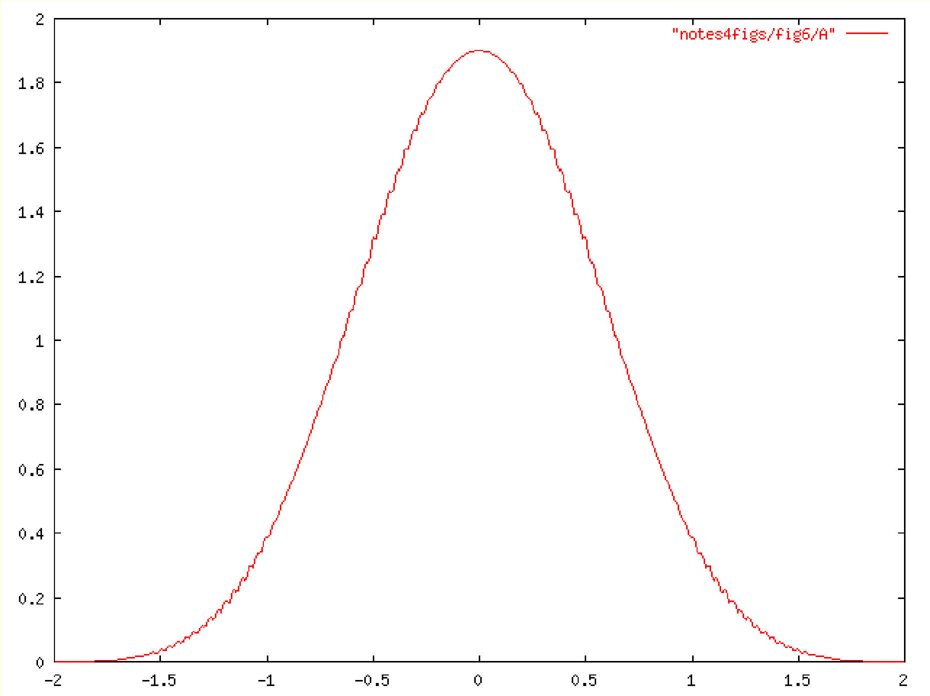
Hence, the amplitude function $E(\xi,\zeta)$ simplifies nicely:

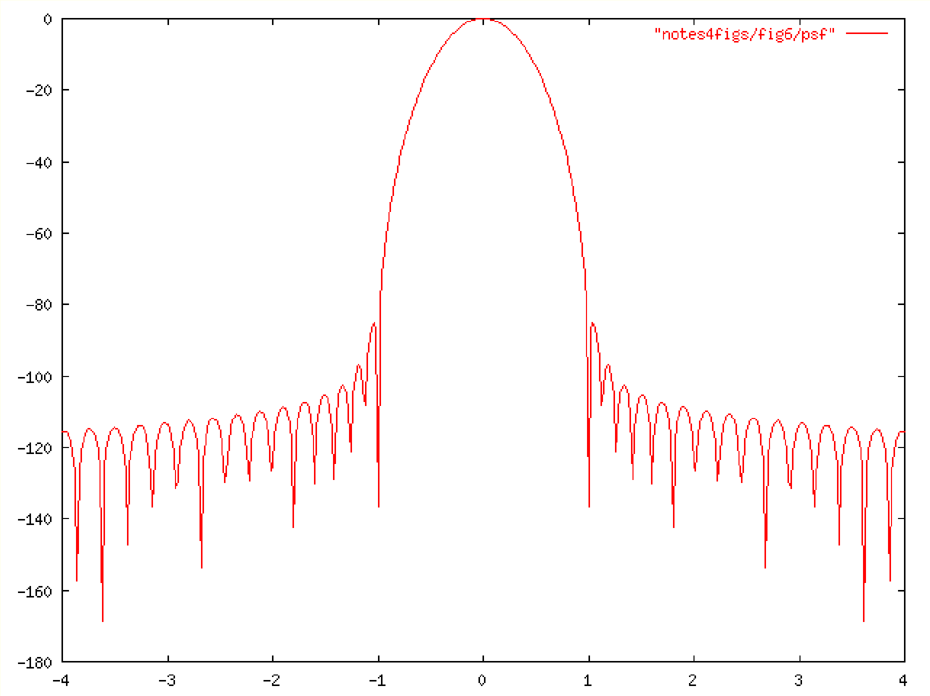
$$\begin{aligned} E(\xi,\zeta) &= \int_{-a/2}^{a/2} \int_{-A(x)}^{A(x)} e^{-2\pi i k (x\xi + y\zeta)/f} dy dx \\ &= \int_{-a/2}^{a/2} e^{-2\pi i k x \xi/f} 2A(x) \mathrm{sinc}(kA(x)\zeta/f) dx. \end{aligned}$$

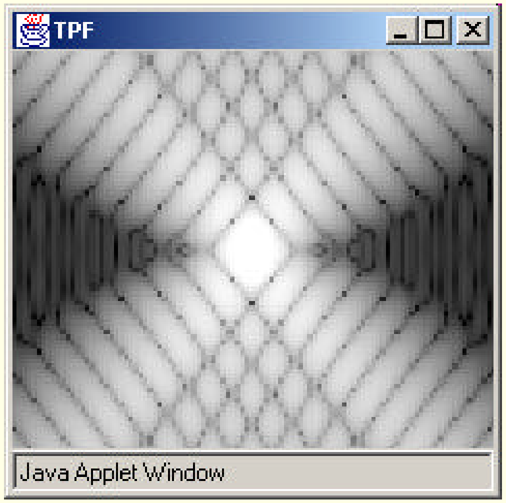
Along the horizontal axis we get further simplification:

$$\begin{aligned} E(\xi,0) &= \int_{-a/2}^{a/2} e^{-2\pi i k x \xi/f} 2A(x) dx \\ &= \int_{-a/2}^{a/2} 2\cos(2\pi k x \xi/f) A(x) dx. \end{aligned}$$

$\xi_0 = 4\lambda f/a$ and $\xi_1 = 100\lambda f/a$

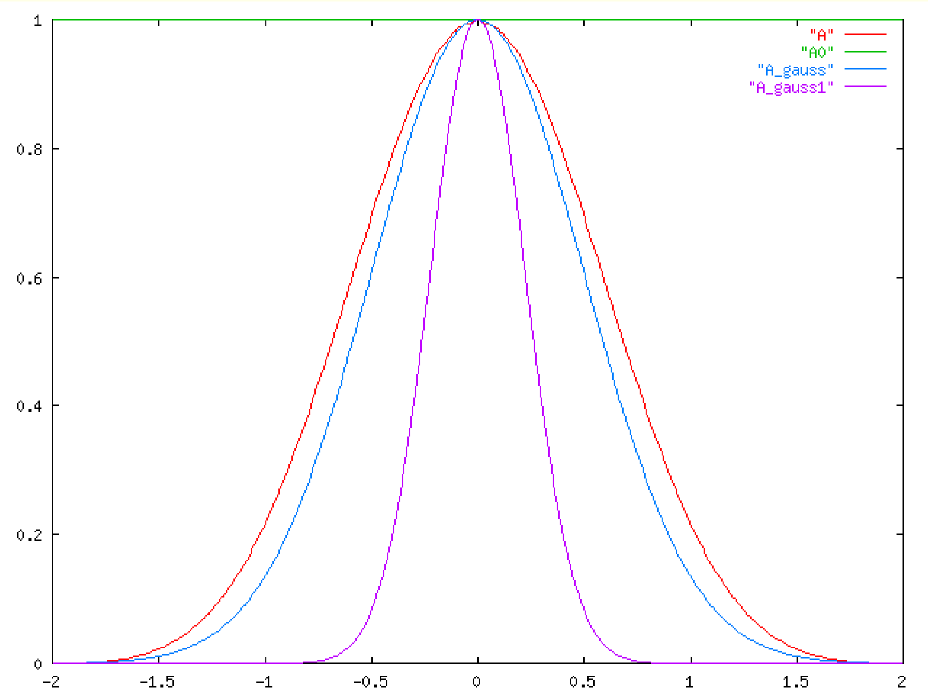


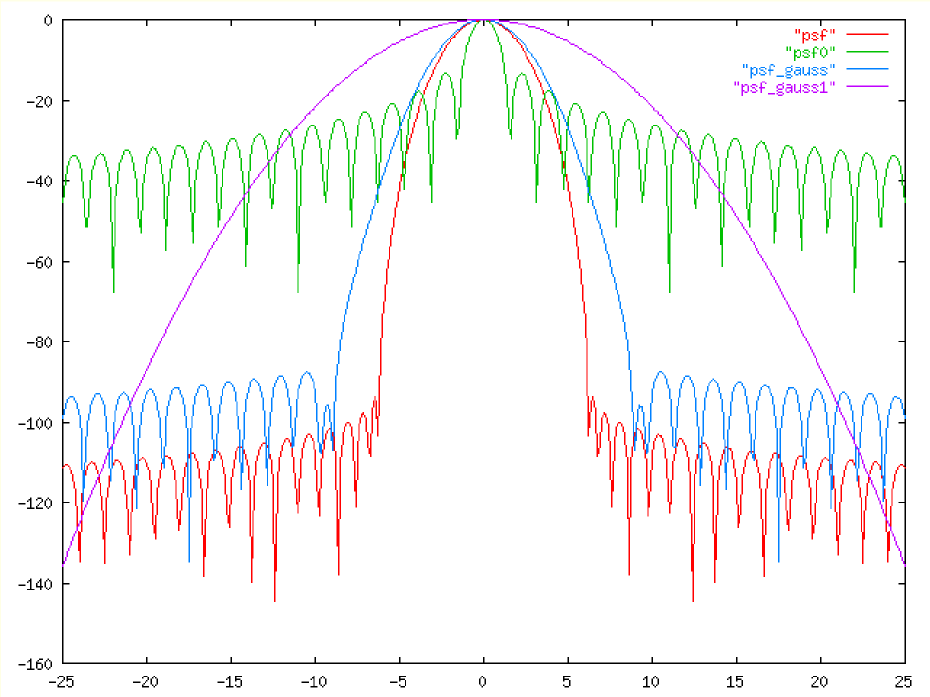




21 Robustness

How good are other similar designs? We compared with a full clear aperture (i.e., a square-shaped aperture) and two Gaussian masks:





Although the Gaussian masks are inferior, they come close.

Manufacturing Precision

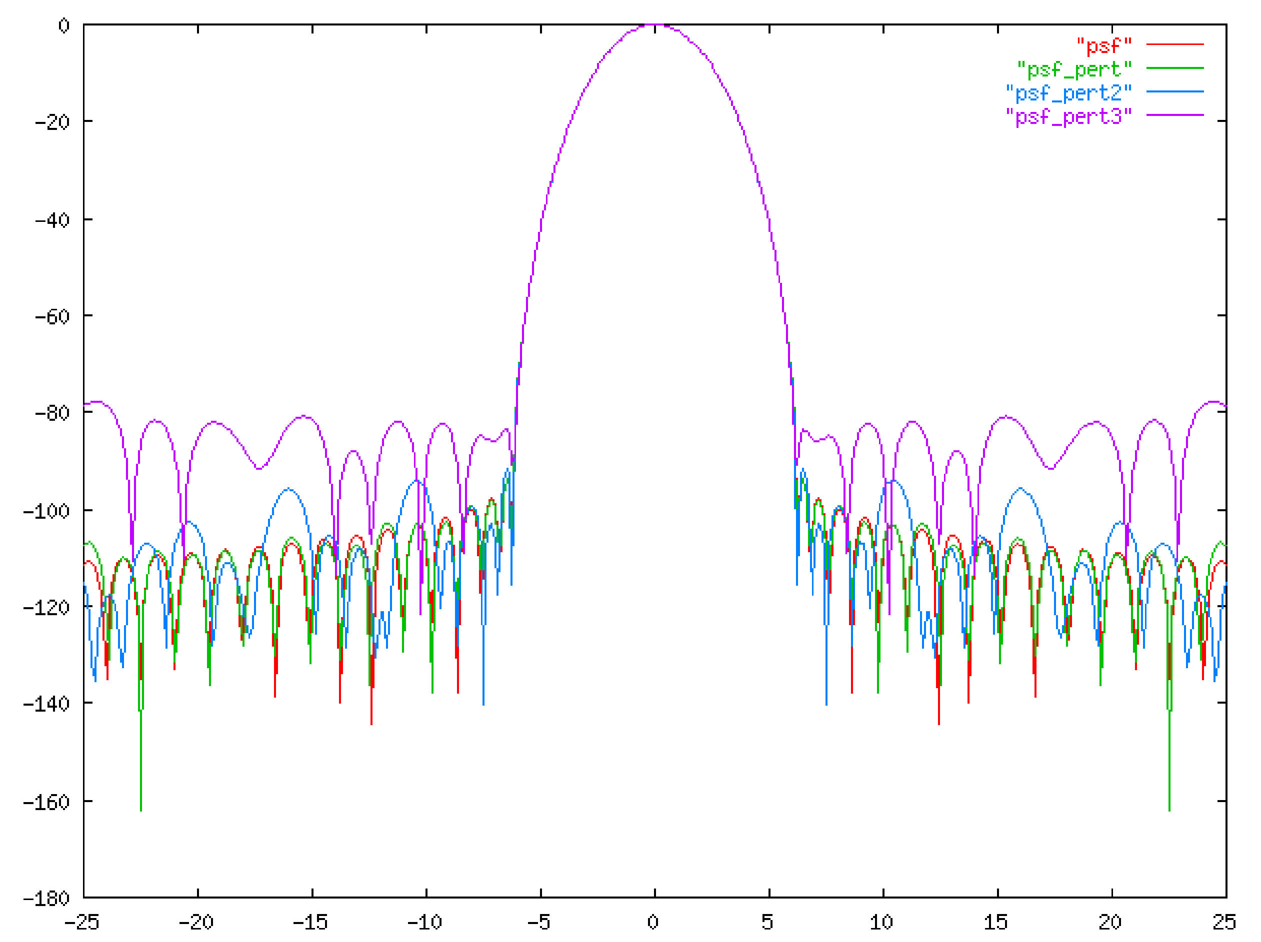
How accurately does one need to cut the mask? I tried three perturbations:

psf_pert is the optimal mask with perturbation:

$$\max((A[r] + (2\text{Uniform01}() - 1)/10^5), 0).$$

psf_pert2 is 10 times bigger.

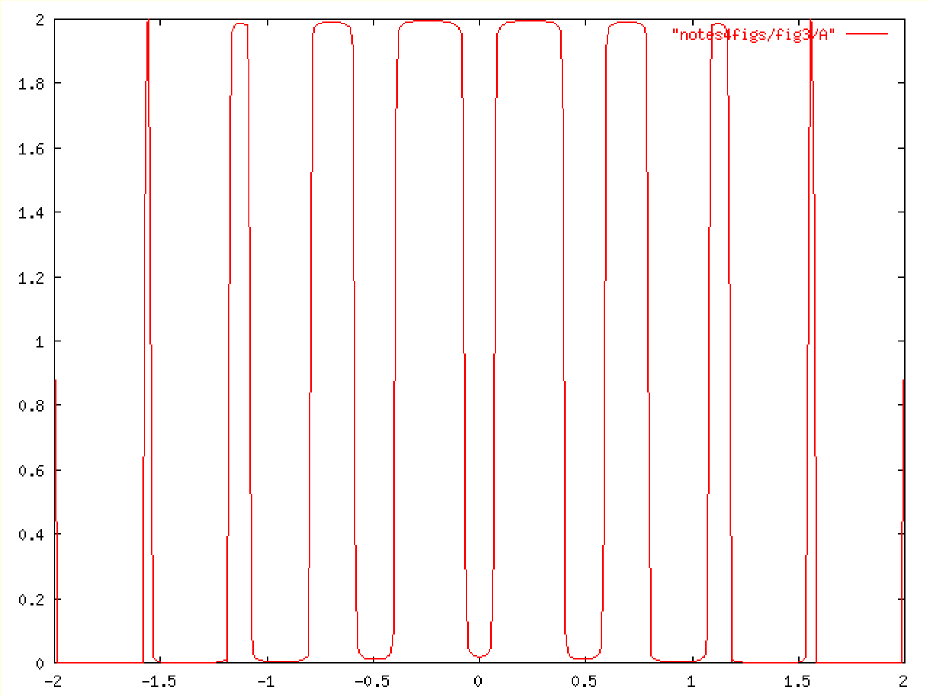
psf_pert3 is 10 times bigger again.

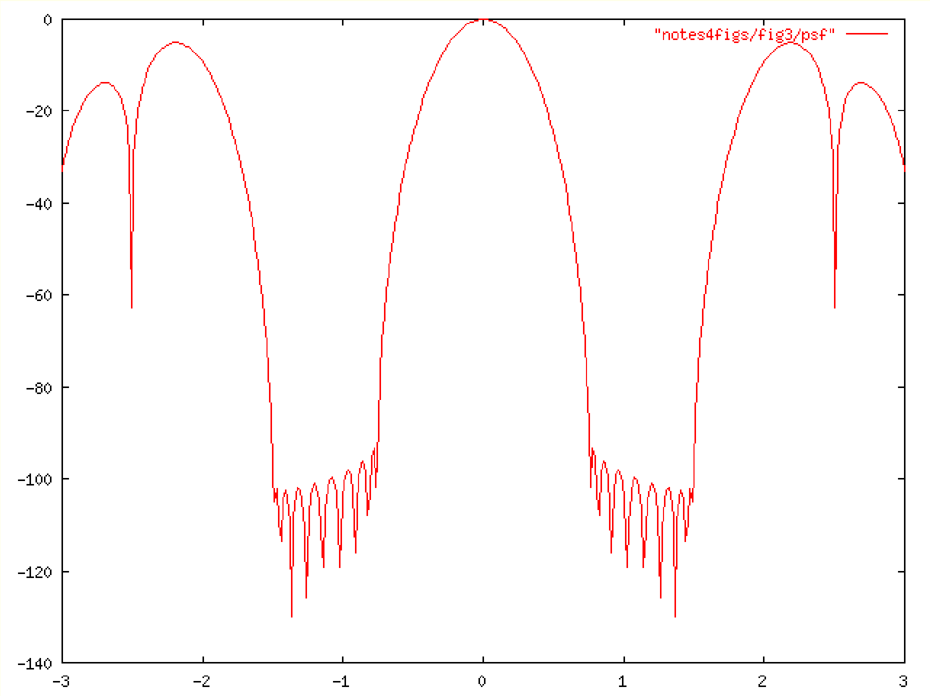


Decreasing the Inner Working Distance

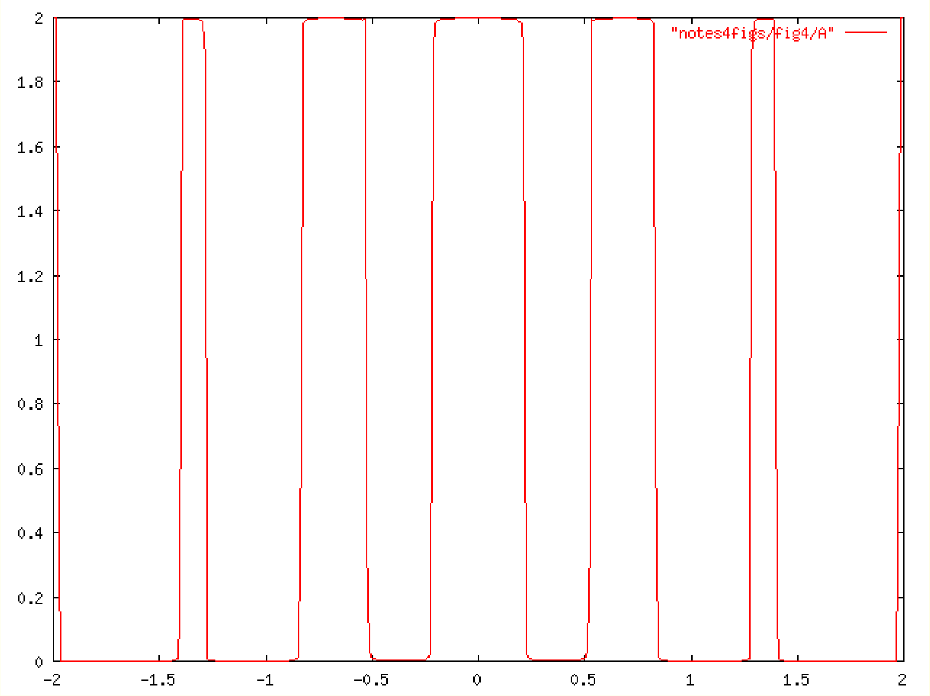
And then decreasing the outer working distance as needed to get -100 dB's in the working band:

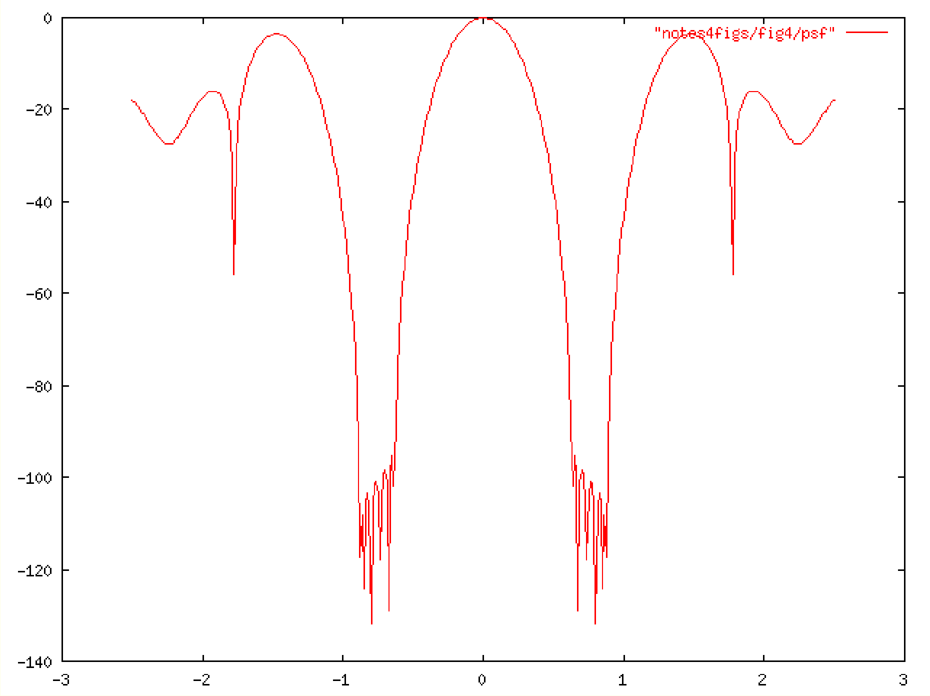
$$\xi_0 = 3\lambda f/a \qquad \text{and} \qquad \xi_1 = 6\lambda f/a$$





 $\xi_0 = 2.5\lambda f/a$ and $\xi_1 = 3.5\lambda f/a$





25 Circularly Symmetric Masks

- My original question was "Why not work with circularly symmetric optics?" In this case, one could think of making a variable filter. That is, at point (x, y) have the filter transmit a fraction $A(x, y)$ of the light.
- Such a filter is called an apodization.
- The answer is that apodizations are hard to make accurately.
- For small working bands, the square-aperture masks are essentially bang-bang all-or-nothing masks.
- It suggests looking for similar circularly symmetric masks.
- They can be thought of as apodizations in which the apodizing function $A(r)$ is zero-one valued.
- On the next few slides we derive the formulas for circularly symmetric apodization and then restrict attention to the zero-one valued case.

Circularly Symmetric Apodization

Instead of a square mask, we consider now a circularly symmetric apodized aperture:

$$E(\xi, \zeta) = \int_0^{a/2} \int_{-\pi}^{\pi} A(r) e^{-2\pi i k(x\xi + y\zeta)/f} r d\theta dr$$

where, of course, $x = r\cos\theta$ and $y = r\sin\theta$.

WLOGWMAT, $\zeta = 0$ and hence we look at

$$\begin{aligned} E(\xi, 0) &= \int_0^{a/2} r A(r) \left(\int_{-\pi}^{\pi} e^{-i \frac{2\pi k \xi}{f} r \cos\theta} \right) d\theta dr \\ &= \int_0^{a/2} 2\pi r A(r) J_0\left(\frac{2\pi k r \xi}{f}\right) dr \end{aligned}$$

27 Circularly Symmetric Masks

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

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

The integral on the previous slide can now be written as a sum of integrals and each of these integrals can be explicitly integrated to get:

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

Circularly Symmetric Masks Optimization Problem

$$\text{minimize} \quad \int_{\xi_0}^{\xi_1} |E(\xi)|^2 \xi d\xi$$

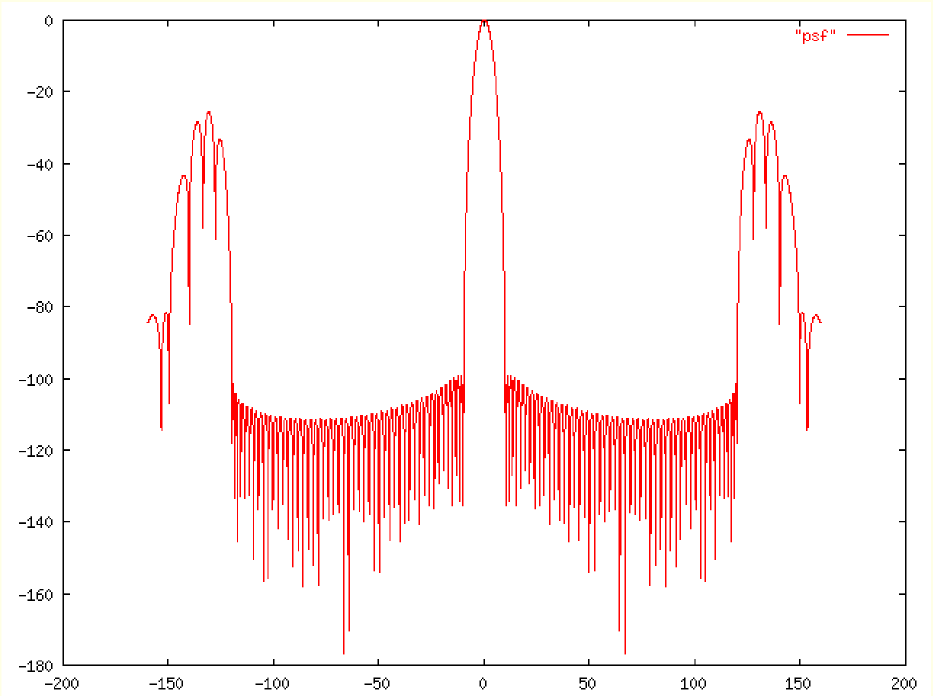
$$\text{subject to:} \quad \sum_{j=0}^{m-1} \pi (r_{2j+1}^2 - r_{2j}^2) \geq \pi a^2/16$$

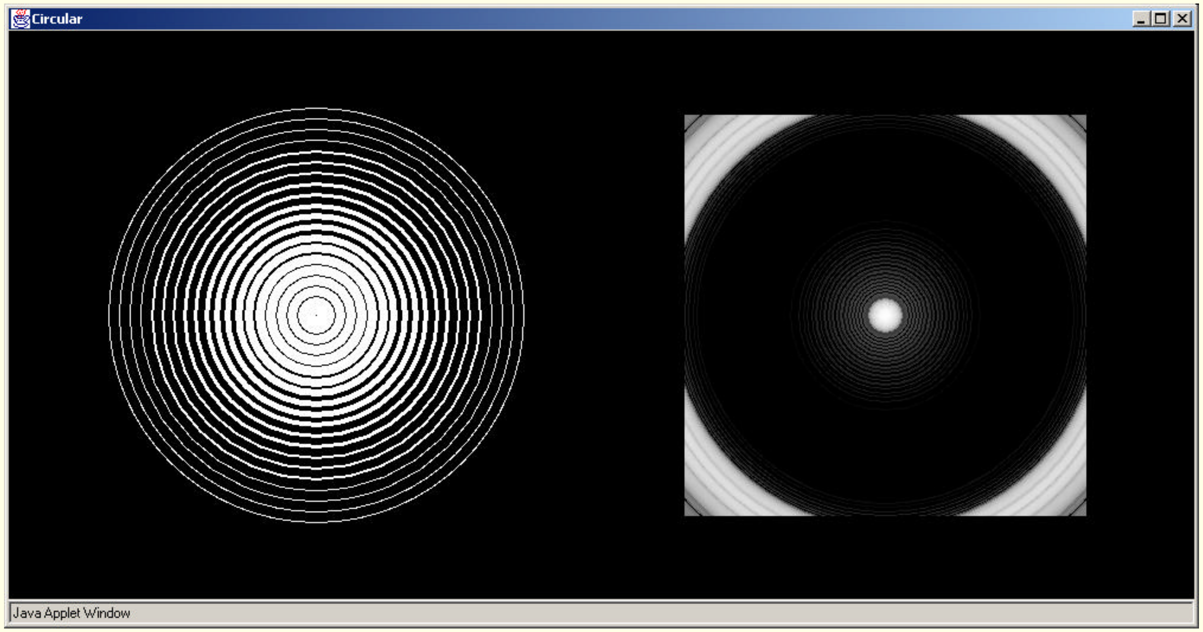
$$0 \leq r_0 \leq r_1 \leq \cdots \leq r_{2m-1} \leq a/2.$$

where $E(\xi)$ is the function of the r_j's given on the previous slide.

Note: Here there be dragons. Until now all models have been convex quadratically constrained quadratic programs. This model is nonconvex and nonquadratic — the objective function involves Bessel functions and the constraints involve differences of quadratic terms.

 $\xi_0 = 4\lambda f/a$ and $\xi_1 = 48\lambda f/a$ and $m = 26$





Outline

- LOQO: An Interior-Point Code for NLP
- Finite-Impulse-Response (FIR) Filter Design
- Antenna Array Design
- Telescope Pupil-Shape Optimization
- Stable Periodic Orbits for the Equal-Mass n-Body Problem
- Parking Orbits for Satellites

Celestial Mechanics—Periodic Orbits

- Find periodic orbits for the planar gravitational n-body problem.
- Minimize action:

$$\int_0^{2\pi} (K(t) - P(t))dt,$$

- where $K(t)$ is kinetic energy,

$$K(t) = \frac{1}{2} \sum_i m_i \left(\dot{x}_i^2(t) + \dot{y}_i^2(t) \right),$$

- and $P(t)$ is potential energy,

$$P(t) = -\sum_{i<j} \frac{m_i m_j}{\sqrt{(x_i(t) - x_j(t))^2 + (y_i(t) - y_j(t))^2}}.$$

- Subject to periodicity constraints:

$$x_i(2\pi) = x_i(0), \qquad y_i(2\pi) = y_i(0).$$

Periodic Solutions

We assume solutions can be expressed in the form

$$
\begin{aligned}
x(t) &= a_0 + \sum_{k=1}^{\infty} \left(a_k^c \cos(kt) + a_k^s \sin(kt) \right) \\
y(t) &= b_0 + \sum_{k=1}^{\infty} \left(b_k^c \cos(kt) + b_k^s \sin(kt) \right)
\end{aligned}
$$

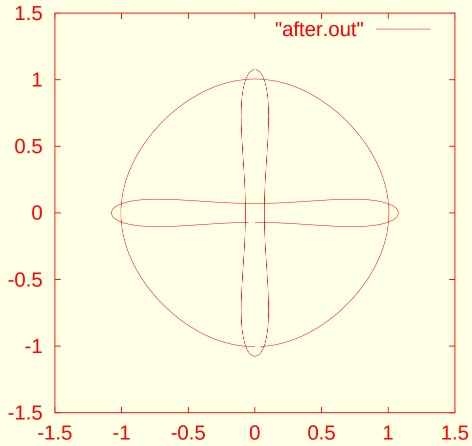
The variables a_0, a_k^c, a_k^s, b_0, b_k^c, and b_k^s are the decision variables in the optimization model.

Choreographies and the Ducati Orbit

Recently, Montgomery and Chencinier (2001) and then others discovered a host of new solutions to the equimass n-body problem. They call these solutions choreographies because all of the bodies follow the same path — they are simply spread out uniformly along this path. They found these orbits by minimizing the action functional.

I reproduced these choreographies using AMPL and LOQO and noticed that all except one are unstable.

This inspired me to look for stable solutions. I found a few, including the one that M. Todd called the Ducati solution.



34 Limitations of the Model

- The infinite sum gets truncated to a finite sum. This amounts to adding constraints. Hence, the solution might be suboptimal. That is, the trajectory obtained might not satisfy the equations of motion.
- Masses must be positive.
- Model can't solve 2-body problem w/ eccentricity (see next section).

Outline

- LOQO: An Interior-Point Code for NLP
- Finite-Impulse-Response (FIR) Filter Design
- Antenna Array Design
- Telescope Pupil-Shape Optimization
- Stable Periodic Orbits for the Equal-Mass n-Body Problem
- Parking Orbits for Satellites

Elliptic Solutions to the 2-Body Problem

An ellipse with semimajor axis a, semiminor axis b, and having its left focus at the origin of the coordinate system is given parametrically by:

$$x(t) = f + a\cos t, \qquad y(t) = b\sin t,$$

where $f = \sqrt{a^2 - b^2}$ is the distance from the focus to the center of the ellipse.

However, this is not the trajectory of a mass in the 2-body problem. Such a mass will travel faster around one focus than around the other. We need to introduce a time-change function $\theta(t)$:

$$x(t) = f + a\cos\theta(t), \qquad y(t) = b\sin\theta(t).$$

This function θ must be increasing and must satisfy $\theta(0) = 0$ and $\theta(2\pi) = 2\pi$.

The optimization model can be used to find (a discretization of) $\theta(t)$ automatically by letting it be a vector of variables and adding appropriate monotonicity and boundary constraints.

A Hill-Type Solution to the Eccentric Sun-Earth System

Using an eccentricity $e = f/a = \mathbf{0.0167}$ and appropriate Sun and Earth masses, we can find a periodic Hill-Type satellite trajectory in which the satellite orbits the Earth once per year.