Exozodiacal Dust

Thesis by
Marc J. Kuchner

In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy

California Institute of Technology
Pasadena, California

2001
(Submitted 10 August 2000)

Acknowledgements

Mike Brown, for taking a chance on me and for teaching me to be relevant

Bill Reach, for teaching me that dust can be beautiful

Wal Sargent, for not forcing me to be a spectroscopist

Chris Koresko, for Bim-Bam-Bop, hamburger stars, and life on the other side of the Fourier transform

Sterl Phinney, for avuncular advice

Peter Goldreich, for teaching dynamics on demand

Gene Serabyn, Mark Colavita and Shri Kulkarni, for teaching me optical interferometry

Anand Sivaramakrishnan and Russ Makidon, for teaching me coronagraphy

Geoff Blake, for foolishly agreeing to be on my committee at the last minute

Anneila Sargent and Nick Scoville, for foolishly agreeing to be on my committee a long time ago

Robert Kirshner and Ramesh Narayan, for showing me that astronomers really do exist

Antonin Bouchez, Sarah Stewart, Eugene Chiang and Adam Burgasser, for caring about rocks

Marshall Cohen, for almost observing G 29-38 about 3 or 4 times

Ben Oppenheimer, for letting me ride in his convertible

Michelle Creech-Eakman, for being superwoman

John Cartwright, for all those girl scout cookies

Brad Behr, for proofreading lyrics while they were coming out of the printer

Brian Kern, for solving that problem on clusters and the Jacobi constant at 4 am with me

Roy Gal and Pensri Ho, for letting me move legos

David Vakil, for Mongolian food, with the sauce

Chris Fassnacht, for the good desk

Rudi Danner, for a place in Palomar Observatory history

Laura Grego, for her voice

Jon Sievers, Kurt Adelburger and Rob Simcoe, for being generally well behaved

Dave Schiminovich, Steve Kaye and Kristi Meade, for being Polly's Toybox

John Yamasaki, for putting the toys in the box

Pat Udomprasert, Bryan Jacoby, Edo Berger, Matthew Hunt, Mike Santos, Josh Bloom, Richard Ellis, Sharda Jogee, David Frayer, John Carpenter, Lori Lubin and Mark Metzger, for CPU cycles and for playing the find-the-brown dwarf video game

Alice Shapely, for being flexible about donuts

David Kaplan and Micol Christopher, for tennis

Tom Lloyd, Lou Madsen, Chris Kaestle, Gary Leskowitz and Mary Mason for various musical and non-musical distractions

Eric Kuchner, for making sure I didn't forget about birthdays

Olga Kuchner, for helping me go back to school

My Grandparents, how lucky I am to have you

Mom and Dad, nothing I could write would say enough. I love you both.

Abstract

Besides the sun, the most luminous feature of the solar system is a cloud of "zodiacal" dust released by asteroids and comets that pervades the region interior to the asteroid belt. Similar clouds of dust around other stars—exozodiacal clouds—may be the best tracers of the habitable zones of extra-solar planetary systems. This thesis discusses three searches for exozodiacal dust:

1) We observed six nearby main-sequence stars with the Keck telescope at 11.6 microns, correcting for atmosphere-induced wavefront aberrations and deconvolving the point spread function via classical speckle analysis. We compare our data to a simple model of the zodiacal dust in our own system based on COBE DIRBE observations and place upper limits on the density of exozodiacal dust in these systems.

2) We observed Sirius, Altair, and Procyon with the NICMOS Coronagraph on the Hubble Space Telescope to look for scattered light from exozodiacal dust and faint companions within 10 AU from these stars. We did not achieve enough dynamic range to surpass the upper limits set by IRAS on the amount of exozodiacal dust in these systems, but we did set strong upper limits on the presence of nearby late-type and sub-stellar companions. We explain the technique of coronagraphy with a discussion of the Fourier optics of a one-dimensional coronagraph.

3) The planned nulling capability of the Keck Interferometer should allow it to probe the region < 200 milliarcsecond from a bright star and to suppress on-axis starlight by factors of 10^{-3} to reveal faint circumstellar material. We model the response of the Keck Interferometer to hypothetical exozodiacal clouds to derive detection limits that account for the effects of stellar leakage, photon noise, noise from null depth fluctuations, and the fact that the cloud's shape is not known a priori. Our models show that the Keck Interferometer can detect an exozodiacal cloud with as little as 10

times the optical depth of the solar zodiacal cloud, even when the transmitted stellar signal is stronger than the signal from the dust cloud.

We also discuss the interaction of dust with planets. We used the COBE DIRBE Sky and Zodi Atlas and the IRAS Sky Survey Atlas to search for dynamical signatures of three different planets in the solar system dust complex:

1) We searched the COBE DIRBE Sky and Zodi Atlas for a wake of dust trailing Mars. We compare the DIRBE images to a model Mars wake based on the empirical model of the Earth's wake as seen by the DIRBE and place a 3-σ upper limit on the fractional overdensity of particles in the Mars wake of 18% of the fractional overdensity trailing the Earth.

2) We searched the COBE DIRBE Sky and Zodi Atlas for Trojan dust near Jupiter's L5 Lagrange point. We place a 3-σ upper limit on the effective emitting area of large (10-100 micron diameter) particles trapped at Jupiter's L5 Lagrange point of 6×10^{17} cm^2, assuming that these large dust grains are distributed in space like the Trojan asteroids. We would have detected the Mars wake if the surface area of dust in the wake scaled simply as the mass of the planet times the Poynting-Robertson time scale.

3) We compared the COBE DIRBE Sky and Zodi Atlas and the IRAS Sky Survey Atlas to search for dust created in the Kuiper Belt and trapped in mean-motion resonances with Neptune. We place a model-dependent upper limit of 0.6 MJy steradian^{-1} on the brightness of dust trapped in Neptune's 2:3 resonance.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 The Solar System, as it Would Appear to an Observer 10 Parsecs Away

Besides the sun, the most luminous component of the solar system is not Venus, not Jupiter, not any planet—but a cloud of "zodiacal dust" generated by asteroids and comets mostly within a radius of 4 AU from the sun. Figures 1.1 and 1.2 illustrate this point; they show the solar system as it would appear to an observer 10 parsecs distant at a wavelength of 20 microns, a crucial wavelength for searching for circumstellar material at terrestrial temperatures. The story is not much different at other optical or infrared wavelengths; our zodiacal cloud outshines all of the planets in total emitted or scattered light at any wavelength that is not much larger than the typical grain size (1–100 microns across).

Figure 1.1 serves as a reference point. It shows the solar system as it would appear to a telescope with a Gaussian beam with a full-width at half-maximum of 5 milliarcseconds (mas). At 20 microns, this corresponds to the diffraction limit of a telescope 1 kilometer in diameter. Earth appears as an unresolved point source 0.1 arcseconds to the right of center. In the face-on view, Mercury and Venus can also just barely be distinguished. The central blob of light is thermal emission from the zodiacal cloud, calculated from a model by Kelsall et al. (1998) based on observations by the Diffuse Infrared Background Experiment (DIRBE) aboard the Cosmic Background Explorer (COBE) satellite.

One kilometer is certainly larger than the diameter any single dish telescope, and it is also ten times longer than the baseline of any optical or infrared interferometer, existing or under construction, though this may be an appropriate baseline for the proposed Terrestrial Planet Finder interferometer (Beichman et al. 1999). Figure 1.2 shows the same scene, but convolved with a Gaussian with full width at half-maximum of 50 mas to illustrate how the solar system would appear to a telescope with a 100 meter baseline. Several interferometers now operating in the infrared have comparable baselines. At this practically attainable resolution, and at any coarser resolution, the zodiacal emission overwhelms the light from any single planet in any beam.

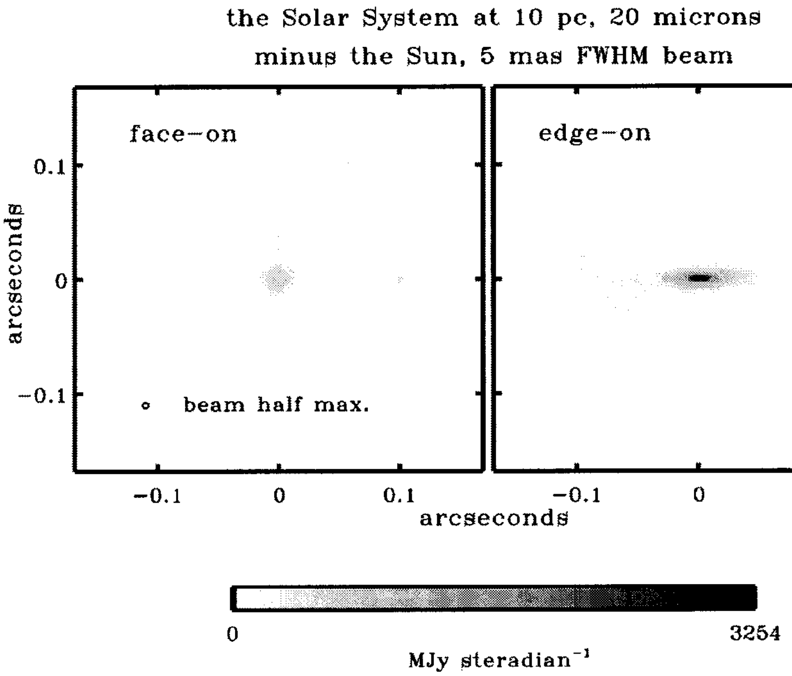

Figure 1.1: The Solar System viewed from 10 pc through a huge futuristic telescope with an aperture of 1000 meters. The Earth appears as a point source 0.1 arcseconds to the right of center.

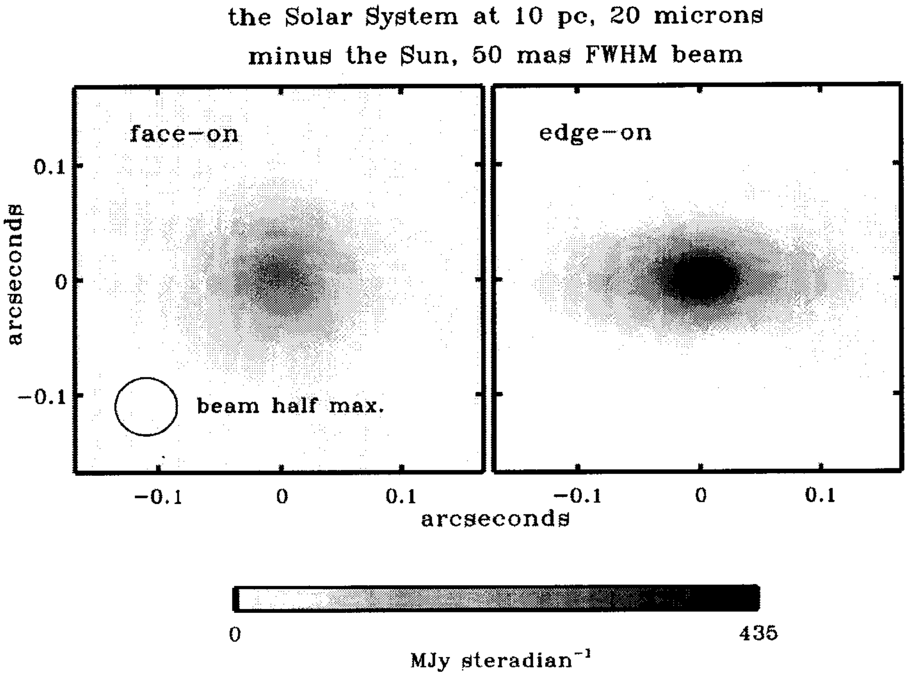

Figure 1.2: At the resolution of the biggest mid-infrared interferometers now under construction, the zodiacal cloud overwhelms the planets.

1.2 The Solar System Zodiacal Dust Cloud

The solar system dust cloud can be seen from inside the solar system too. On a clear night just after dark or before dawn, the zodiacal cloud often appears as a diffuse triangle of "zodiacal light" covering an impressively large fraction of the sky. The name "zodiacal" applies because the light is concentrated in the ecliptic plane within the constellations of the zodiac.

A simple example shows how dust in a planetary system can easily be much brighter than the planets themselves. Grinding a medium-sized asteroid (50 km diameter) into dust particles 10 microns in diameter would increase the surface area of the material by a factor of 5 billion. If you spread this much dust into an optically thin cloud, the total effective surface area would match the surface area of a solid ball with radius 2×10^6

km, 25 times the radius of Jupiter (7×10^4 km).

Planetary systems older than a few tens of millions of years tend to be nearly in equilibrium with the radiation field from the stars they orbit. Consequently, at a given circumstellar distance, the brightest component of a planetary system will be the one with the most effective surface area for emitting and scattering light. So it should not be surprising that our zodiacal cloud outshines all of the planets.

Unfortunately, the population of asteroids 50 km in diameter (and smaller) in the Solar System, which could potentially produce such large quantities of dust, is not well known, and larger asteroids collide rarely. However, simulations based on extrapolation of deep asteroid searches (Durda and Dermott 1997) suggest that asteroids 50 km in diameter suffer catastrophic collisions with other asteroids every $\sim$ 100 million years in the asteroid belt, and smaller asteroids collide much more often. These simulations show a size distribution punctuated by collisional cascades, where large asteroids are shattered into smaller chunks that grind each other into dust over the following 65 million years. It is clear from these calculations and from the appearance of extra-dense bands of dust mid-infrared images of the solar system that appear to correspond to Hirayama families of asteroids (Sykes and Greenberg 1986; Sykes and Greenberg 1986e; Reach et al. 1997) that asteroidal collisions are a major source of the zodiacal dust.

Comets clearly release quantities of dust which would suffice to maintain the current zodiacal cloud, if the dust survived long enough. It has even been suggested that most of the cloud comes from a single comet, Comet Encke (Whipple 1967). However, a large but uncertain fraction of cometary dust is released on hyperbolic orbits, and quickly escapes the solar system. Consequently, estimates of the cometary contribution to the zodiacal cloud are as uncertain as estimates of the asteroidal contribution. Probably the best indication that comets contribute significantly to the zodiacal cloud is the cloud's scale height, which is about ($r/4.14$ AU), where r is the distance from the sun. This corresponds to a "fan" shape with an opening scale angle of 14°, higher than what might be expected if all the dust were asteroidal (Dermott et al. 1992). Comets typically have higher orbital inclinations than asteroids; they could produce the high-lattitude dust particles (Liou et al. 1995).

Beyond the asteroid belt, zodiacal cloud peters out, according to particle counters aboard the Ulysses spacecraft (Mann and Grün 1995a). But farther from the sun—and beyond the reach of Ulysses—a third source of solar-system dust particles may become important. The high relative velocities of Kuiper Belt objects indicate that collisions among these bodies may create a second cloud of dust, a ring that orbits beyond the orbit of Neptune. This source of dust is the subject of Chapter 7.

Once the dust is created, it does not orbit the sun forever. Instead, the semimajor axis and eccentricity of the orbit slowly decay under the influence of a relativistic interaction with photons from the sun called Poynting-Robertson drag (P-R drag). Chapter 2 describes this effect in more detail, and shows how it influences the large-scale structure of the zodiacal cloud. P-R drag causes zodiacal dust particles released in the asteroid belt to spiral into the sun in timescales of 10^3–10^6 years.

1.3 Exozodiacal Dust: Zodiacal Dust Around Other Stars

Until we can resolve earth-like planets around nearby stars, zodiacal dust, by virtue of its luminosity, may be the best tracer of the central regions of evolved extrasolar planetary systems. Zodiacal dust around another main sequence star, the extrasolar analog of zodiacal dust, is called “exozodiacal” dust. This analogy implicitly assumes that other planetary systems are generally similar to the solar system—a dangerous assumtion. But perhaps there is a common class of planetary systems that are similar enough to the solar system to make the analogy useful.

Systems like our own have three regions sorted by circumstellar distance:

I. A region where giant planets orbit, clearing a zone that is relatively free of dust and small bodies.

II. A region exterior to the giant planet zone where presumably a primordial cloud of small bodies, like the solar system's Kuiper Belt, could remain.

III. A region interior to the giant planet zone where small bodies and possibly rocky planets might have orbits that are stable for the lifetime of the system.

I will call dust in zone II of a planetary system Kuiper Belt Dust, and dust in zone III zodiacal or exozodiacal dust.

This picture represents our solar system, but it clearly does not do justice to all planetary systems. For instance, systems with hot Jupiters orbiting at circumstellar distances of 0.1 AU may not have a zone III. Other planetary systems may not even have giant planets. However, there are a few good examples besides the solar system of nearby systems which seem to be well described by this model.

The A star HR 4796 appears to have a complex dust environment, including a narrow ring at ~ 70 AU that was imaged in scattered light with the NICMOS coronagraph on HST (Schneider et al. 1999). Subtracting a model of the emission from this ring from 12 micron images of HR 4796 made with the Keck telescope (Koerner et al. 1998) leaves a large amount of emission in excess of the stellar photosphere, some of which appears as marginally resolved light near the star. For an idea of the scale of the marginally resolved component, note that the diffraction limit of the Keck Telescope at 12 microns is 250 milliarcseconds, which corresponds to a transverse size scale for the cloud of ~ 17 AU given the distance of HR 4796 (67 ± 3.5 parsecs, determined by *Hipparcos*). No faint companions appear in the images interior to the ring, and no planets around this star have been detected by radial-velocity measurements because A stars generally have lines that are too sparse and too broad for precision doppler techniques. However, the ring of cold dust, the marginally resolved emission near the star, and the gap between the ring and the marginally resolved emission could be analogous to zones II, III and I respectively.

Maps of ϵ Eridani, a nearby K star, made with the Submillimeter Common-User Bolometer Array (SCUBA) on the James Clerk Maxwell Telescope (JCMT) at 850 microns revealed a large circumstellar disk (Greaves et al. 1998). Figure 2 of that paper shows an azimuthally averaged radial profile of the map, which indicates that the ring peaks at 18" from the star, then there is a flux minimum at a radius of 8", and a central rise in the flux density interior to 4.5". Observations with the Infrared Astronomical Satellite (IRAS) confirm the central flux peak at 60 microns interior to 8–11". The cavity extends too far from the star for P-R drag or ice sublimation to have excavated the large grains that must be producing this long-wavelength radiation. However, the three-zone model remains a plausible explanation for the observed structure; giant planets clear the observed gap, and the ring is zone II or Kuiper Belt dust and the central peak is exozodiacal dust.

Besides these two systems, which were resolved with ground-based 10-meter dishes, there are many other examples of main sequence stars with excess emission in the far infrared, the Vega-excess stars. Most of these stars do not have much excess emission in the mid-infrared. The common interpretation of this phenomenon is that the excess far infrared is emission from cold (< 150 K) dust analogous to Kuiper Belt dust in the solar system—or zone II dust. The decline of the excess in the mid-infrared is taken to indicate the presence of a gap that might be cleared by giant planets—zone I.

In general, zone III, the region nearest the star, is the most difficult zone to observe. Yet this the zone is intrinsically the most interesting to an earthling. In the solar system, HR 4796, and Epsilon Eridani, zone III contains the habitable zone, the region where liquid water might exist, where—should all the other conditions be right—a human could live, or perhaps, life resembling life on Earth might begin on its own.

1.4 Detecting Exozodiacal Dust

Even though exozodiacal dust clouds can be much easier than planets to detect directly, these faint diffuse clouds less than 1–2 arcseconds from bright nearby stars are still difficult to find. The clouds around HR 4796 and Epsilon Eridani and all known Vega-excess stars have 2–4 orders of magnitude more optical depth in dust than the solar system. It will be many years before we are capable of detecting a solar-level exozodiacal cloud.

Chapters 3, 4, and 5 represent three approaches to the problem of exozodiacal dust detection. The contrast between the light from the zodiacal cloud and the light from the solar photosphere is highest in the mid-infrared; it peaks around 25 microns. Chapter 3 describes a search for exozodiacal dust using the Keck Telescope in the mid-infrared. For this search, we used the technique of speckle interferometry to maximize the resolution of our observations.

Chapter 4 describes a search for exozodiacal dust in the near infrared using the Hubble Space Telescope (HST). At these wavelengths, zodiacal dust shines mostly in scattered light from the sun, and the contrast between dust and starlight should be less than at thermal wavelengths. To increase the dynamic range of our images, we observed through the coronagraph in the HST's Near Infrared Camera and Multi-Object Spectrometer (NICMOS).

Ultimately, the best technique for detecting exozodiacal dust may be nulling interferometry, a new technique which combines high resolution and high dynamic range. The Keck Interferometer, a facility which will combine the light from the two 10-meter Keck Telescopes, is designed with nulling capability for the purpose of investigating circumstellar matter. Chapter 5 is a discussion of how to use the Keck Interferometer to detect exozodiacal dust.

1.5 The Effect of Planets on Dust Clouds

Although dust easily dominates the surface area of a planetary system, the star and the planets dominate the dynamics. In a three-zone system where giant planets clear a swath that is nearly free of small bodies and dust around their orbits, interactions between giant planets and small particles are clearly crucial to the planetary system's design and appearance. Besides clearing gaps, giant planets can sweep dust clouds into complex shapes. The dust clouds around ϵ Eridani, β Pictoris and Vega all have dramatic asymmetries that are interpreted as the signatures of planetary perturbations (Dent et al. 2000; Ozernoy et al. 2000).

But even smaller, terrestrial-sized planets, though they do not clear holes in the disk, can also have a significant impact on the orbits of dust grains. For example, planets can temporarily trap particles in resonant orbits as they spiral past. Figure 1.3 shows the DIRBE model of the zodiacal cloud with the "smooth" component of the cloud subtracted. The asteroidal dust bands dominate the edge-on view. However, in the face-on view, we see a ring of dust trapped in resonant orbits by the Earth. An extra-overdense region trails the Earth, called the Earth's Wake. Chapter 2 provides some general background on the dynamics of dust in a planetary system and describes how planets can form rings and wakes.

Most interactions between planets and dust can be modeled with a combination of analytical theory and numerical orbit integrations. However, even for our solar system, which ought to be the easiest planetary system to study, observations of the planets' effect on the interplanetary dust are few. Consequently, we are far from a good dynamical description of the dust, and we are far from a good understanding of how our solar system would appear to an outside observer. I undertook to improve this situation by searching for some dynamical signatures of Mars, Jupiter and Neptune in the solar system dust clouds, using data from IRAS and DIRBE. Chapters 6 and 7 describe these searches in detail.

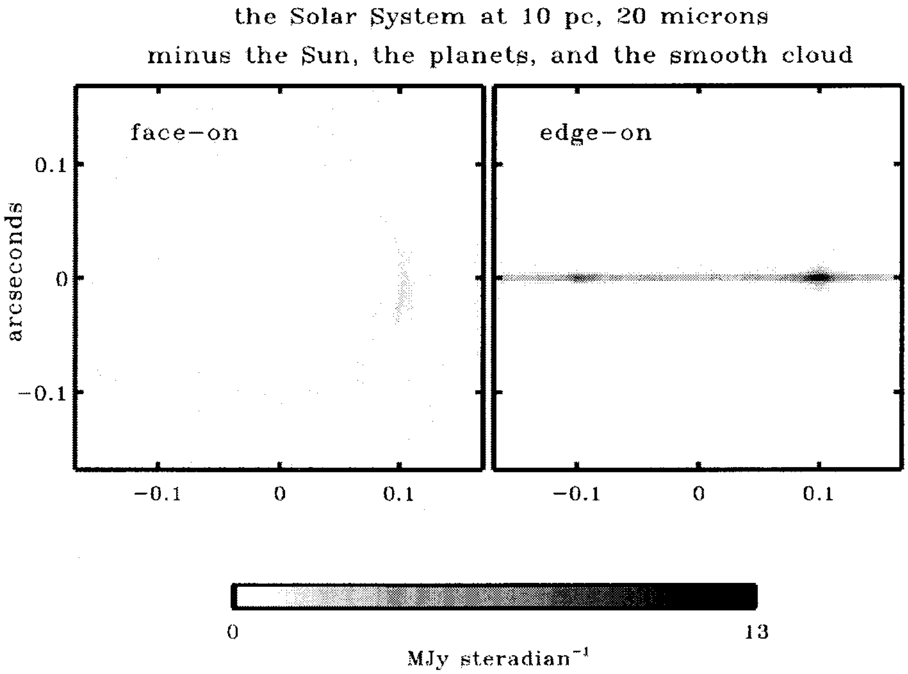

Figure 1.3: The DIRBE model of the Solar System zodiacal cloud shown with the "smooth" component of the cloud subtracted. What remains are the asteroidal dust bands, seen best in the edge-on view, and the Earth's ring and wake, seen best in the face-on view. The Earth's ring and wake are overdense regions of the cloud caused by the temporary trapping of dust in the Earth's mean motion resonances.

Chapter 2

Dust Dynamics

2.1 Solar Gravity and Radiation Pressure

A variety of different forces can affect charged, irregularly shaped, rotating dust particles in interplanetary space (Gustafson 1994). I will focus on the forces that are most important to particles that have long lifetimes in the solar zodiacal cloud: gravity, radiation pressure, and Poynting-Robertson drag. These forces apply even to neutral, spherical particles.

Most of the optical depth in the solar system dust cloud is due to particles which are gravitationally bound to the sun. However, for small particles, which have a small ratio of mass to cross section, radiation pressure from solar photons can balance or even exceed the gravitational force. Radiation pressure on a compact spherical grain is directly outward and inversely proportional to the distance from the star squared, so it counters the effect of the star's gravity. The ratio of the radiation pressure a particle feels to the gravitational force the particle feels is called β.

$$\beta = \frac{\sigma L_\star}{4\pi c G M_\star m}, \tag{2.1}$$

where $L_\star$ is the luminosity of the star, G is the gravitational constant, $M_\star$ is the mass of the star, m is the mass of the particle, and σ is the effective cross section of the particle for receiving stellar radiation. With this notation, the central force felt by a dust grain is

$$F_{central} = -\frac{GM_\star(1-\beta)m}{r^2}. \tag{2.2}$$

For spherical particles with density 2 g cm^{-3} that are large compared to the dominant wavelength of light, $\beta = 0.285/r_{dust}$, where r_{dust} is the radius of the particle in microns. This spherical approximation systematically underestimates β for real zodiacal dust particles, which can have complicated shapes. Gustafson (1994) reviews some measurements of β for more realistic dust particles. In general, however, small particles have large values of β. Radiation pressure forces particles about 1 micron in size onto highly eliptical or even parabolic orbits. These particles are called β-meteoroids. In the solar system, particles smaller than about 0.5 microns also feel strong electromagnetic forces from the solar wind.

When a particle with high β is released from a larger body, it may become unbound from the star. We can easily compute the criterion for this to occur. A particle has orbital energy

$$E_0 = \frac{1}{2}mv^2 - \frac{GM_\star m}{r}, \tag{2.3}$$

where v is the velocity of the particle. When we instantly turn on the radiation pressure and the particle continues to move with the same velocity, the energy is raised by $GM_\star \beta m/r$. The particle becomes unbound if $E > 0$, or

$$\frac{GM_\star \beta m}{r} > -E_0. \tag{2.4}$$

Now

$$E_0 = -\frac{GM_\star m}{2a_0}, \tag{2.5}$$

where a_0 is the particle's initial semi-major axis, and

$$r = \frac{a_0(1-e_0^2)}{1+e_0 \cos f_0} \tag{2.6}$$

for a particle on an eliptical orbit with eccentricity e_0 and true anomaly f_0. So the condition for unbinding a particle is

$$\beta > \frac{1}{2}\frac{1-e_0^2}{1+e_0\cos f_0}. \tag{2.7}$$

For a particle released from a circular orbit, this reduces to $\beta > 1/2$, or for a spherical particle with density 2 g cm^{-3}, $r_{dust} < 0.57$ microns.

2.2 Poynting-Robertson Drag

Imagine an isolated particle cooling and radiating energy isotropically into space. Since the radiation is emitted isotropically, the total reaction force on the particle due to the momentum lost to the radiation is zero. But now visualize the particle in a new Lorentz frame where it moves with some constant velocity. In this frame, the radiation is beamed in the direction of motion of the particle. The particle ought to be accelerated in the direction opposite its motion by the radiation pressure from the beamed radiation. Special relativity tells us, however, that if the particle's motion is unaccelerated in the frame where the particle is stationary, then the particle must be unaccelerated in a frame where it is moving. What is the resolution of this paradox?

The answer is that when the particle cools, it loses mass. In the frame where the particle is stationary, this makes no difference to our perception of the problem, because the total force is obviously zero anyway. But in the frame where the particle moves, we might think that since the particle's mass is decreasing, the particle ought to accelerate. Instead, the reaction force from the beamed radiation keeps the particle exactly on its unaccelerated path.

When a dust particle orbits a star, however, it soon attains equilibrium with the stellar radiation field. In other words, it can't cool. Since the particle can't cool, it can't lose mass. And since it can't lose mass, there is nothing to compensate for the reaction force from any beamed radiation it emits. So particles orbiting a source of photons feel a force in the opposite direction of their motion, a force called "Poynting-Robertson

Drag" (P-R drag).

The magnitude of this force in the case of complete absorption and isotropic re-emission is just the $h\nu/c^2$ "mass" of the light that the particle receives per time times the velocity of the particle, v. In other words,

$$F_{PR} = v \times \text{received power}/c^2. \tag{2.8}$$

Robertson (1937) first correctly derived this result from the relativistic formalism. It is also correct in the case of scattering from a perfectly reflective sphere. Interplanetary dust particles do not scatter much more than 20% of the visible light they receive, so a linear combination of a perfect absorber and a perfectly reflective sphere nearly describes a real particle. If the particle's absorptivity is constant over the band where most of the stellar radiation is emitted, then we can write

$$|F_{PR}| = \frac{L_\star \sigma v}{4\pi r^2 c^2}, \tag{2.9}$$

where $L_\star$ is the luminosity of the star, r is the distance from the particle to the star and σ is the cross section of the particle for receiving stellar radiation. Interplanetary dust particles are generally much larger than visible wavelengths, and made of amorphous materials, so this equation is a good approximation in the solar system.

Robertson (1937) also derived equations for the orbit of a particle experiencing P-R drag and calculated the secular variation of the orbital elements based on the orbit equations. In the case of a circular orbit, radius a, we can write down the perturbation in the semimajor axis directly. Using

$$\frac{d}{dt} mv = -F_{PR} \tag{2.10}$$

and

$$v = \left(\frac{GM_\star(1-\beta)m}{a}\right)^{1/2} \tag{2.11}$$

we find

$$\frac{da}{dt} = -\frac{L_\star \sigma}{2\pi a m c^2} = -\frac{2GM_\star \beta}{ac}. \tag{2.12}$$

This equation can be integrated using separation of variables to give

$$a(t) = \sqrt{a_0^2 - \frac{L_\star \sigma}{\pi m c^2} t} \tag{2.13}$$

where a_0 is the initial semimajor axis of the orbit. The time for the semimajor axis to decay to zero is the "Poynting-Robertson time":

$$T_{PR} = \frac{\pi a_0^2 m c^2}{\sigma L_\star} = 400 \text{ years } \frac{(a_0/1AU)^2}{\beta(L_\star/L_\odot)}. \tag{2.14}$$

To within a factor of a few, the P-R time is the time for a particle at distance a_0 from a star to intercept its own rest mass energy in photons. Low-β particles fall slowly into a star under the influence of P-R drag, while high-β particles are ejected from the vicinity of a star on a dynamical time scale. This is why our zodiacal cloud is much denser interior to the asteroid belt than exterior to the asteroid belt.

For reference, in the general case with non-zero eccentricity, the secular perturbations are (Wyatt and Whipple 1950):

$$\left[\frac{da}{dt}\right]_{P-R} = -\frac{GM_\star \beta}{ac} \frac{2+3e^2}{(1-e^2)^{3/2}} \tag{2.15}$$

$$\left[\frac{de}{dt}\right]_{P-R} = -\frac{5GM_\star \beta}{2a^2 c} \frac{e}{(1-e^2)^{1/2}}. \tag{2.16}$$

Using Kepler's 3rd law, we can rewrite equation 2.16 as

$$\left[\frac{dn}{dt}\right]_{P-R} = \frac{3GM_\star\beta n}{2a^2c}\frac{2+3e^2}{(1-e^2)^{3/2}}. \tag{2.17}$$

2.3 The Large-Scale Structure of the Zodiacal Cloud

Imagine that all the solar system dust particles were released in a narrow ring at exactly 3 AU. If particles are neither created or destroyed, the number density of particles, N, satisfies the continuity equation. In a system with cylindrical symmetry, this equation is

$$\frac{1}{r}\frac{\partial}{\partial r}(rNv_r) + \frac{\partial}{\partial t}(Nv_r) = 0, \tag{2.18}$$

where r is the distance from the sun and v_r is the rate that a particle's distance from the sun changes due to P-R drag. In steady state, we set the term with the time-derivative equal to zero, and integrate to get

$$rNv_r = \text{constant} \tag{2.19}$$

for a constant source of dust. For particles on circular orbits, $v_r = da/dt \propto 1/r$, so we have N = constant. In other words, a steady-state zodiacal cloud with dust on circular orbits has a constant face-on optical depth interior to the region where the dust is released.

When the P-R force comes from thermal emission or scattering from a uniform sphere, the force vector lies in the plane of the orbit, so it does not affect the orbit's inclination. So particles released with a variety of orbital inclinations and phases spiraling into the sun form a fan-shaped structure, with a thickness proportional to r. Since the face-on optical depth of the fan is roughly constant, the number density of the fan falls roughly off as r^{-1}.

The most detailed empirical model of the solar system zodiacal cloud is the 88-

parameter model that Kelsall et al. (1998) fit to the data from the Diffuse Infrared Background Experiment (DIRBE) aboard the Cosmic Background Explorer (COBE) satellite. We will use this model over and over again throughout this thesis. In the Kelsall et al. (1998) model, the dust density falls off as $r^{-1.34}$.

The DIRBE observations do not probe the solar system interior to a heliocentric radius of 0.87 AU. However, observations by the Heilos satellites of the scattered zodiacal light suggest that the dust number density is proportional to $r^{-1.3}$ all the way into heliocentric distances of 0.3 AU (Lienert et al. 1981). The radial power laws from the DIRBE and Helios models are remarkably close to r^{-1}; the differences may be due to the cometary dust component, which is supplied throughout the inner solar-system, the gradual collisional evolution as the dust, and in the case of Helios data, the scattering properties of the dust.

2.4 Interactions With Planets

A planet in a dust cloud can dramatically alter the cloud's structure, or subtly perturb the dust orbits. A dust particle interacting with a planet can be considered a special case of the restricted three-body problem, which is discussed in detail in many textbooks on dynamics (e.g. Brouwer and Clemence 1961, Murray and Dermott 1999). Here is a rough sketch of the big picture.

A dust grain orbiting a star together with a planet feels the potential

$$V_\star + V_{planet}, \tag{2.20}$$

where

$$V_\star = \frac{GM_\star(1-\beta)m}{r} \qquad \text{and} \qquad V_{planet} = \frac{GM_\star m}{r_{planet}}. \tag{2.21}$$

It is traditional to expand V_{planet} in a Fourier series. This makes sense because when the perturbations from the planet are small, the coordinates of objects on eliptical orbits about the star are periodic in time. The important quantities are not the absolute positions of the objects, but their relative positions—and not the cartesian coordinates, but their orbital elements:

a	semimajor axis	λ	mean longitude
e	eccentricity	ϖ	longitude of pericenter
i	inclination	Ω	ascending node

The expansion is made in cosines of angles which are linear combinations of the orbital elements of the particle and the orbital elements of the planet.

$$V_{planet} = GM_{\star}m_{planet}\sum S(a, a', e, e', I, I')\cos\phi \tag{2.22}$$

$$\phi = j_{\lambda'}\lambda' + j_{\lambda}\lambda + j_{\varpi'}\varpi' + j_{\varpi}\varpi + j_{\Omega'}\Omega' + j_{\Omega}\Omega \tag{2.23}$$

The unprimed quantities refer to the orbit of the planet; the primed quantities refer to the orbit of the dust particle. The j's in this linear combination are strictly integers. When V_{planet} is expanded this way, it is called the "disturbing function." We can learn about the interaction of a planet and a small object, like a dust grain, by classifying the terms in this Fourier Series.

Secular Terms

$$j_{\lambda'} = j_{\lambda} = 0 \tag{2.24}$$

Secular terms are those terms which do not depend on λ or λ', the mean longitudes of the planet and the particle, the angles that indicate where in their orbits the particle and planet are *right now.* These terms can alter the eccentricity, inclination, longitude of pericenter, and the longitude of ascending node of zodiacal dust orbits, but do not affect the semimajor axis or epoch. In the limit where particles spend a long time very close to a single perturber, the perturber will add a "forced eccentricity" and "forced inclination" to the orbits of the particles which match the eccentricity and inclination of the perturber. In the solar system, secular perturbations offset the center of the zodiacal cloud from the sun by ~ 0.012 AU (Kelsall et al. 1998). The offset is roughly equal to the product of the semimajor axis and the eccentricity of the orbit of Jupiter, which presumably causes the perturbation.

Secular perturbations may also create the warp in the β Pictoris disk (Mouillet et al. 1997) and shape the ring around HR 4796 (Wyatt et al. 1999). Jupiter-mass planets appear to be common around G stars, and their orbits are often eccentric (see, e.g., Marcy and Butler 1998). It is easy to imagine an exozodiacal dust disk dramatically reshaped by secular perturbations from these massive planets, even if the planet is too far from the star to make a detectable radial-velocity signature.

Resonant Terms

$$j_{\lambda'} n' + j_{\lambda} n \approx 0 \tag{2.25}$$

where n is planet's orbital frequency, or mean motion, and n' is the particle's mean motion. These are the terms whose effects accumulate over time when a particle's periodic close encounters with the planet, or conjunctions, occur repeatedly at the same places in the particle's orbit. This requires that particle's mean motion divided the planet's mean motion is a ratio of whole numbers.

If a particle is spiralling inwards, it is easier to trap the particle into a resonance where the particle is exterior to the planet. In this case it is common to write

$$\frac{n'}{n} = \frac{p}{p+q}, \tag{2.26}$$

where p and q are whole numbers, instead of using the j's. Using Kepler's law, the above condition can be rewritten as

$$\frac{a'}{a} = (1-\beta)^{1/3}\left[\frac{p+q}{p}\right]^{2/3}. \tag{2.27}$$

Resonances are associated with special values of a particle's semimajor axis.

A particle in a $p{:}p+q$ resonance has a close encounter with the planet once every p orbits of the particle, or $p+q$ orbits of the planet. In general, the smaller p and q, the more frequent the close encounters, and the stronger the time-averaged resonant perturbations. When $q = 1$, the dust particle completes a whole number of orbits between successive conjunctions with the planet, and every conjunction occurs at the same longitude in inertial space. These sorts of resonances are called first order resonances.

Short-Period Terms

All the other terms are collectively called "short period terms." They are difficult and often impossible to include in analytic calculations because they are infinite in number. In many situations, their perturbative effects are uncorrelated, and they average away. However, sometimes their effects are crucial. Short period terms can dominate the dynamics whenever there is a close encounter between a particle and a planet. Their cumulative effects over many orbital periods contribute to the chaotic nature of the three body problem. Short period terms provide the ultimate mechanism for removing particles trapped in resonances.

2.5 Resonant Rings and Wakes

As a particle's semimajor axis passes through a resonance, the resonant perturbations can temporarily balance P-R drag, halting the particle's inward spiral. The particle is then said to be "trapped" in the resonance, though since the situation is temporary, "detained" might be a better word. Whatever the terminology, the result is that a particle spends more time in the vicinity of a resonance than elsewhere on its inward spiral; resonances create overdense regions in dust clouds.

The strength of a resonance depends on frequency and strength of the closest encounters between the particle and the planet. Resonances that are close to the planet have $a^{'}/a \approx 1$. According to equation 2.27, this means q is small compared to p. Indeed, the most important resonances for trapping dust are the first order resonances, where $q = 1$, i.e., resonances of the form p:p+1.

As a dust particle spirals past a planet, first it encounters the 1:2 resonance, then the 2:3, then the 3:4 and so on. At first, the resonances get stronger as the particle moves closer to the planet. But then as p continues to increase, the particles become more susceptible to close encounters with the planet, and the trapping times decrease. Eventually at high values of p, the resonances start to overlap, and chaos sets in (Wisdom 1980). Generally, when many particles spiral past a planet, several resonances near a special value of p become well populated.

Figure 2.1 shows how a 3:4 resonant orbit appears in a reference frame that is rotating with the planet (indicated by the $\oplus$) around a star (indicated by the $\odot$). In Figure 2.1, the dust particle is exactly on resonance. As the particle librates about this equilibrium orbit, it traces out a path something like the one in Figure 2.2. Figure 2.3 illustrates how dust trapped in mean-motion resonances of the form p:p+1 can create a ring-like density enhancement in a zodiacal cloud. Overlaying plots of librating resonant orbits for a few different resonances (p=3,4, and 5) produces the characteristic ring shape, with a gap at the location of the planet.

The effect of P-R drag on resonant orbits is to displace their equilibrium pericenters by a small angle. Smaller particles with larger βs require a bigger kick to counter the effect of P-R drag, so they prefer to have their resonant encounters closer to the planet on its trailing side, shown in Figure 2.4. The result is that for smaller particles, the ring shifts so that the hole leads the planet, and a "wake" forms, trailing the planet. This wake is just the region where the pericenters of the orbits of the small particles align so these particles can receive their resonant perturbations directly behind the planet.

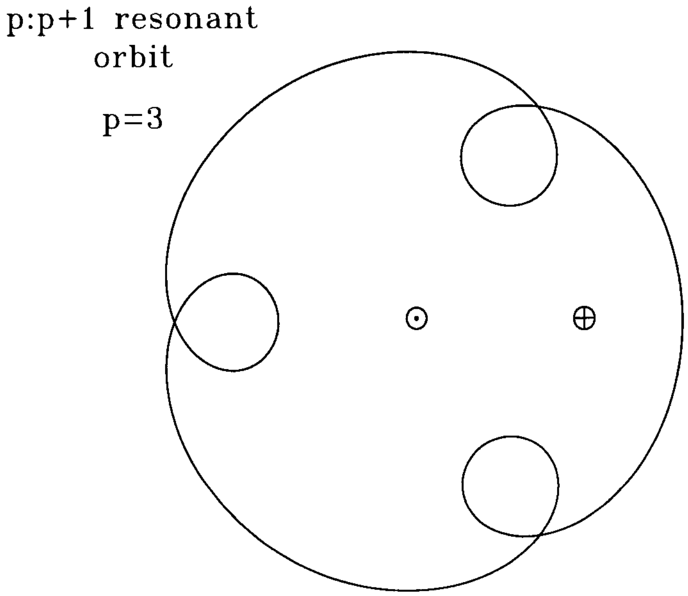

Figure 2.1: A resonant orbit shown in a reference frame corotating with the planet.

When a particle is trapped in a resonance, the resonant perturbations maintain the particle's semimajor axis, and thereby the energy of its orbit (equation 2.5). But P-R drag continues to decrease the particle's angular momentum ($h = na^2\sqrt{1-e^2}$) and increase the particle's eccentricity. In the absence of other perturbations the particle's eccentricity would approach a preferred value, where the orbit has just become planet crossing (Weidenschilling and Jackson 1993) in a time period proportional to the P-R time. What actually happens is that as the particle's orbit becomes more nearly planet crossing, it becomes more susceptible to the effects of other terms in the disturbing function. One day it has a close encounter with the planet that knocks it out of resonance, sometimes into a safe orbit that lets it spiral past the planet, sometimes into an unbound orbit, and sometimes into another resonance.

The more massive the planet, the more the low p resonances dominate the trapping.

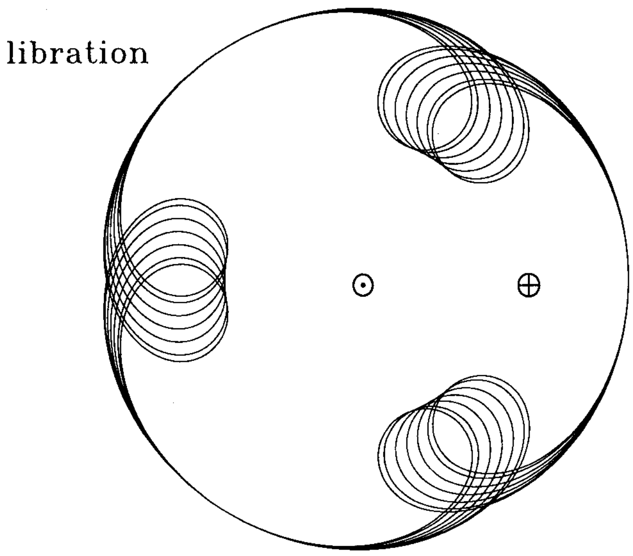

Figure 2.2: A resonant orbit with libration.

For massive planets the more distant resonances become stronger and the interior resonances overlap sooner. In numerical models of Earth's ring (Dermott et al. 1994), the most populated resonance is $p = 5$. Numerical models of dust trapped in resonances by Neptune (Liou and Zook 1999) suggest the most populated resonance is $p = 2$ or 3.

The situation becomes more complicated when the planet is on an eccentric orbit. The small eccentricity of the Earth's orbit causes the Earth to move with respect to its ring and wake, causing seasonal variations in the DIRBE and IRAS maps. When planets have very eccentric orbits, they can create structures that are better described as "arcs" than rings. For some examples of the interesting ways particles can behave when perturbed by a planet on an eccentric orbit, see Roques et al. (1994).

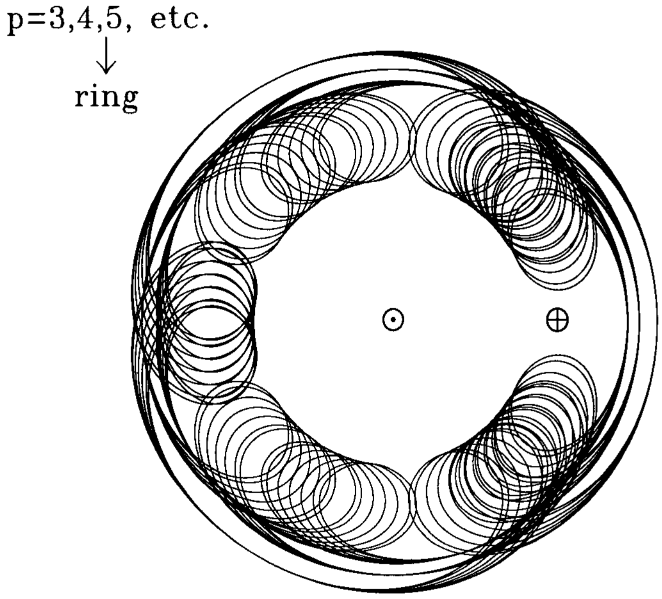

Figure 2.3: If several resonant orbits are occupied, the result is a ring with a gap at the location of the planet.

2.6 The Pericenter Shift

Here is a brief discussion of the pericenter shift that assumes that the planet is on a circular orbit.

In general, a resonant term has the following form:

$$R = \frac{Gm}{a} F_{j_{\lambda'} j_{\lambda} j_{\varpi'} j_{\varpi} j_{\Omega'} j_{\Omega}}(a, a', e, e', i, i') \cos\phi, \tag{2.28}$$

where ϕ is called the resonant argument. The function F can be read from tables in Murray and Dermott (1999) for instance. We can compute the time variation of any orbital element, and hence of ϕ using Lagrange's Planetary Equations (Brouwer and

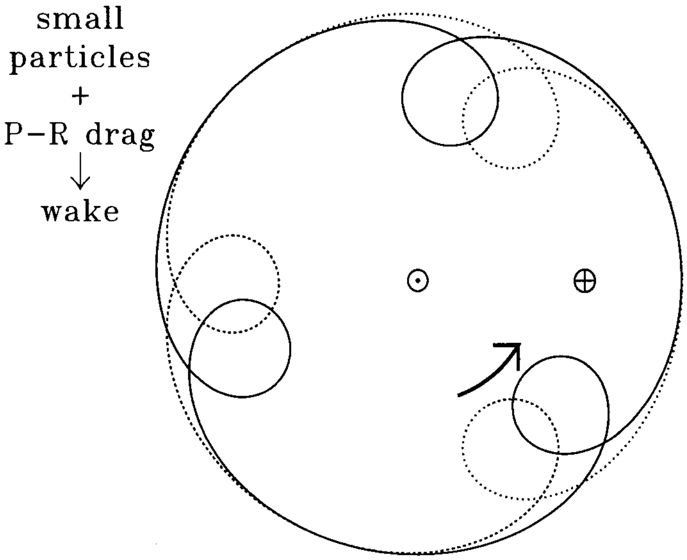

Figure 2.4: P-R drag shifts the pericenters of the equilibrium resonant orbits for particles with small β.

Clemence 1961). For instance, Lagrange's equations tell us

$$\frac{dn'}{dt} = -\frac{3}{a'^2}\frac{\partial R}{\partial \lambda}. \tag{2.29}$$

In other words,

$$\left[\frac{dn'}{dt}\right]_{resonant} = \frac{3Gm}{a'^3} F \sin\phi. \tag{2.30}$$

When the planet's orbit is circular

$$\phi = j_{\lambda'}\lambda' + j_{\lambda}\lambda + j_{\varpi'}\varpi. \tag{2.31}$$

This angle can be interpreted as the angle between the longitude of pericenter of the particle's orbit and the longitude of the particle's conjunction with the planet. Differentiating twice, and assuming that the planet's orbit is unchanging, we find

$$\ddot{\phi} = j_{\lambda'}(\dot{n}' + \dot{\epsilon}') + j_{\varpi'}\ddot{\varpi}', \tag{2.32}$$

where ϵ is the particle's mean longitude at epoch. If we neglect $\ddot{\epsilon}'$ and $\ddot{\varpi}'$ (they are are generally small), we find

$$\ddot{\phi} = j_{\lambda'}\dot{n}' = \frac{3Gm}{a'^3} F \sin\phi. \tag{2.33}$$

This differential equation is analogous to the equation of motion for a simple pendulum. There are two families of solutions. Some solutions correspond to the pendulum oscillating about an equilibrium position; ϕ oscillates about a stable point, $\phi_0 = \pi$, and the particle is trapped in a resonant orbit with the planet. The other solutions correspond to the pendulum swinging around in a circle; in this case, ϕ circulates, and the particle is not trapped. When the particle is trapped, the oscillations of ϕ around its equilibrium value are called "librations."

P-R drag adds another term to this equation of motion; it tries to slowly increase ϕ. Using equation 2.17,

$$\ddot{\phi} = j_{\lambda'}\left(\left[\frac{dn'}{dt}\right]_{resonant} + \left[\frac{dn'}{dt}\right]_{PR}\right) = \frac{3Gm}{a'^3} F \sin\phi + \frac{3GM_\star \beta n'}{2a'^2 c}\frac{2 + 3e'^2}{(1 - e'^2)^{3/2}}. \tag{2.34}$$

In the vicinity of the equilibrium point, this system is analogous to a mass hanging on a spring in a gravitational field; the gravitational field doesn't change the character of the allowed oscillations, but it shifts the position of the equilibrium point so that at rest, the mass hangs lower, where the force from the stretched spring balances the gravitational force. Likewise, when the P-R term is added to the equations, it doesn't destroy the oscillations, but instead, it shifts the location of the equilibrium position,

ϕ_0. Setting $\ddot{\phi} = 0$ at ϕ_0, we find

$$\sin\phi_0 = -\frac{M_\star \beta n' a'}{2Fmc} \frac{2 + 3e'^2}{e'(1 - e'^2)^{3/2}}. \tag{2.35}$$

In the reference frame rotating with the planet, this translates into a shift in the pericenters of the equilibrium orbit by $\phi_0/(p+1)$.

Chapter 3

A Single-Dish Search for Thermal Emission from Exozodiacal Dust

3.1 Speckle Interferometry

Passive ground-based telescopes are designed to form compact images from plane electromagnetic waves. However, a plane wave accumulates phase errors as it passes through the atmosphere, and these errors blur the image of a point source into a seeing disk. The size of the seeing disk indicates the effective aperture limit that the atmosphere imposes. That aperture limit is roughly what those who study the atmosphere call R_0.

A typical seeing disk at optical wavelengths has a full width at half-maximum of 0.5-3 arcseconds, much larger than the $\sim$ 12 milliarcsecond diffraction limit of a 10-meter telescope at V band. For Kolmogorov turbulence, the size of the seeing disk scales as $\lambda/R_0 \propto \lambda^{-1/5}$ (e.g., Hardy 1998), so at 12 microns, the seeing disk is about half the size it would be in V band, and the diffraction limit of the telescope is about 240 milliarcseconds. In other words, at 12 microns, on a night of good seeing, R_0 can be about 10 meters and a 10-meter telescope can be nearly diffraction limited.

However, for our mid-infrared search for exozodiacal dust, we wanted the ability to investigate the size of sources not much larger than the diffraction limit of the telescope. So even the small atmosphere-induced phase errors that a 10-meter telescope is sensitive to at 12 microns in good seeing could be a significant obstacle—and we did not have good seeing. A conceivable solution to the problem would be to intersperse observations of our target stars with observations of calibrator stars which we expect to be unresolved. Then one might imagine using the calibrator images to deconvolve the seeing disk from the images of the target stars. However, that approach has limited potential, because the time-averaged seeing disk has very little power at spatial frequencies beyond λ/R_0. So dividing one Fourier transform of a seeing-blurred image by another only serves to amplify the background noise at high spatial frequencies.

More sophisticated approaches to the deconvolution problem do exist. For our search for exozodiacal dust in the mid-infrared we decided to try a technique called speckle interferometry, which Labeyrie (1970) invented specifically to overcome the effects of phase errors introduced by the atmosphere. Here is a description of the technique which expands upon a paper by Roddier (1986).

Consider a monochromatic plane wave with wavenumber k incident on a telescope primary. We can write the electric field in the image plane, E as a sum of electric fields arriving from many infinitesimal sub-apertures of the primary. The field from each sub-aperture i changes phase as it propagates the distance x_i to the image plane. If we ignore the time variations, we have

$$E = \sum_i E_i e^{ikx_i}. \tag{3.1}$$

The optics conspire to make the paths x_i proportional to some constant plus the dot-product of the position coordinates in the pupil plane, $\vec{p}$, and the position coordinates in the image plane, $\vec{\xi}$, at least for small deviations from on-axis propagation (e.g., Schreoder 1987);

$$E \propto \sum_i E_i e^{ik\vec{p}\cdot\vec{\xi}}. \tag{3.2}$$

The result is that the E for a point source $\propto$ the Fourier transform of the aperture function. This means that the intensity in the image plane, $I = |E|^2$, is the autocorrelation of the aperture function;

$$I = \sum_{i} \sum_{j} E_i E_j e^{ik(x_i - x_j)} = \sum_{\Delta P} E_i E_j e^{ik\Delta P \cdot \vec{\xi}}, \tag{3.3}$$

where $\Delta\vec{P}$ is the separation vector between sub-aperture i and sub-aperture j. The auto-correlation of the aperture function is the number of pairs, N, of infinitesimal subapertures that have the same vector spacing on the primary, also called the Pupil Redundancy.

Here is a summary of these relationships for a monochromatic plane wave using the hat accent to indicate Fourier conjugates.

$$\hat{E} = \text{aperture function} \tag{3.4}$$

$$\hat{I} = N(\Delta\vec{p}) \tag{3.5}$$

The effect of the atmosphere is well modeled by grouping the terms in equations 3.3 and 3.5 into macroscopic sub-apertures, each the size of R_0, and multiplying each group by a random phasor, $e^{i\phi}$. The result is that at any instant, the image of a point source appears like a cluster of diffraction-limited images, called speckles. This pattern changes on a time scale called the "coherence time" as winds blow the atmospheric turbulence across the sky.

$$\hat{I}_{speckle} = N(\Delta\vec{p})e^{i\phi_{\Delta\vec{p}}} \tag{3.6}$$

Speckle interferometry offers a way to use the telescope-atmosphere transfer function, $\hat{I}_{speckle}$, to estimate the $N(\Delta\vec{p})$, the telescope transfer function in the absence of the atmosphere. The trick is that when you sum a series of N identical phasors, the sum grows as N, but when you sum an uncorrelated series of N random phasors, the magnitude of the result grows as $\sqrt{N}$. At any instant, $\hat{I}_{speckle}$ itself is a sum of random phasors, at least at $\Delta\vec{P} > R_0$, so it approaches a constant times $\sqrt{N(\Delta\vec{P})}$. But if we capture $\hat{I}_{speckle}$ at several different times, and let $e^{i\phi_{\Delta\vec{P}}}$ become un-correlated between them, we can form a sum that grows as $N(\Delta\vec{P})$ by squaring each $\hat{I}_{speckle}$. This function has power at spatial frequencies up to the diffraction limit of the telescope, which means that deconvolving it from an image will not simply serve to amplify noise at high spatial frequencies.

The practical recipe for speckle interferometry is not much more complicated than this. If a source has brightness T, then in the image plane, we have the convolution of T and $I_{speckle}$, and in the pupil plane, we have $\hat{I}_{speckle}\hat{T}$. We followed this procedure:

1) Take many short integrations, shorter than the atmospheric coherence time.

2) Fourier transform all the images, and square the Fourier transforms to get power spectra, which are $|\hat{I}_{speckle}|^2$ times the power spectrum of the target source, $|\hat{T}|^2$.

3) Sum all the power spectra to get a function which is proportional to $N(\Delta P)|\hat{T}|^2$.

4) Do the same for a point-like calibrator star to get a function proportional to $N(\Delta P)|\hat{C}|^2$ where $\hat{C}$ is the power spectrum of the calibrator. If the calibrator is much smaller than the diffraction limit, then $\hat{C}$ is a constant at $\Delta\vec{P} <$ the diffraction limit.

5) Divide the sum of the target power spectra by the sum of the calibrator power spectra. If the seeing has not changed substantially over the course of the observations, this step yields the spatial power spectrum of the target, $|\hat{T}|^2$, at $\Delta\vec{P} <$ the diffraction limit.

In the infrared, the high level of thermal background necessitates subtracting images of the background light from images of astronomical objects to make useful final images. To correct for the contributions of the non-uniform background images to the power spectra of the astronomical objects in our mid-infrared search for exozodiacal dust, we also tallied power spectra for our off-source exposures, and subtracted these from the on-source power spectra before we compared targets and calibrators.

3.2 An 11.6 Micron Keck Search For Exozodiacal Dust

Marc J. Kuchner, Michael E. Brown and Chris D. Koresko
California Institute of Technology, Pasadena, CA 91125

We have begun an observational program to search nearby stars for dust disks that are analogous to the disk of zodiacal dust that fills the interior of our solar system. We imaged six nearby main-sequence stars with the Keck telescope at 11.6 microns, correcting for atmosphere-induced wavefront aberrations and deconvolving the point spread function via classical speckle analysis. We compare our data to a simple model of the zodiacal dust in our own system based on COBE DIRBE observations (Kelsall et al. 1998) and place upper limits on the density of exozodiacal dust in these systems.[1]

3.2.1 Introduction

Our sun is surrounded by a disk of warm (>150 K) "zodiacal" dust that radiates most of its thermal energy at 10–30 microns. This zodiacal dust is produced largely in the inner part of the solar system by collisions in the asteroid belt (Dermott et al. 1992) and cometary outgassing (Liou and Zook 1996). Zodiacal dust is interesting as a general feature of planetary systems, and as an indicator of the presence of larger bodies which supply it; dust orbiting a few AU from a star is quickly removed as it loses angular momentum to Poynting-Robertson drag (Robertson 1937). Understanding the extra-solar analogs of zodiacal dust may also be crucial in the search for extra-solar planets (Beichman et al. 1996) since exozodiacal dust in a planetary system could easily outshine the planets and make them much harder to detect.

The best current upper limits for the existence of exozodiacal dust disks come from IRAS measurements of 12 and 25 micron excesses above photospheric emission. Seen from a nearby star, solar system zodiacal dust would create only a 10^{-4} excess over the sun's photospheric emission at 20 microns. IRAS measurements, however, have typical

---

[1] Adapted from Kuchner et al. (1998)

measurement errors of 5 percent (Moshir et al. 1992) and display systematic offsets of a similar magnitude when they are compared to other photometry (Cohen et al. 1996). If there were a solar-type zodiacal disk with 1000 times the density of the disk around the sun around Tau Ceti, the nearest G star, the excess infrared emission would barely exceed the formal 68% confidence intervals of the IRAS photometry. Moreover, all photometric detection schemes of this sort are limited by how accurately the star's mid-infrared photospheric emission is known. For farther, fainter stars than Tau Ceti, inferring the presence of dust from the IRAS data becomes still harder.

The detection of faint exozodiacal-dust emission is more feasible if one can resolve the dust emitting region. The high resolution and dynamic range needed for these observations will generally require large interferometers like the Keck Interferometer, the Large Binocular Telescope, and the Very Large Telescope Interferometer. But it is already possible to resolve the zodiacal dust mid-infrared emitting regions of the nearest stars. A 10-meter telescope operating at 12 microns has a diffraction-limited resolution of 0.25 arc seconds, corresponding, for example, to a transverse distance of 2 AU at 8 parsecs.

We have begun a search for zodiacal dust around the nearest stars using the mid-infrared imaging capabilities of the Long Wavelength Spectrometer (LWS) (Jones & Peutter 1993) on the W. M. Keck telescope. The large aperture of the telescope allows us to make spatially resolved images of the zodiacal dust 11.6 micron emitting region around the stars so that we can look for dust emission above the wings of the point-spread function (PSF) rather than as a tiny photometric excess against the photosphere. We present here the results of two nights of observations, and compare them with a simple model of exozodiacal thermal emission to place upper limits on the amount of dust present in the systems we observed.

3.2.2 Observations

We observed six nearby stars with LWS on the W. M. Keck telescope on August 3 and 4, 1996, using standard mid-infrared imaging techniques. The target stars were the nearest A–K main-sequence stars observable from Mauna Kea on those dates. With the object on-axis, we took a series of frames lasting 0.8 ms each, chopping the secondary mirror between the object and blank sky 8 arcseconds to the north at a frequency of 10 Hz. Then we nodded the primary mirror for the next series of frames so that the sky was on-axis and the object off-axis. We repeated this process for 3 nods over a period of 5 minutes, for an on-source integration time of 1.1 minutes, and a typical noise of 2 mJy in one 0.11 by 0.11 arcsecond pixel due to the thermal background. The seeing was poor both nights, up to 2 arc seconds in the visible. To measure the atmosphere-telescope transfer function, we made similar observations of seven distant, luminous calibrator stars near our targets on the sky, alternating between target and calibrator every 5–10 minutes.

We increased our frame-rate for the second night of observations so that we could compensate for the seeing using speckle analysis. Figure 3.3 shows a cut through a single 84 ms exposure of Altair on August 4, compared to an Airy function representing the diffraction-limited PSF of a filled 10-meter aperture at 11.6 microns. The cores of the images are diffraction-limited, but the wings are sensitive to the instantaneous seeing, making speckle analysis necessary. Table 3.1 provides a summary of our observations.

We flat-fielded the images by comparing the response of each pixel to the response of a reference pixel near the center of the detector. First we plotted the data number (DN) recorded by a given pixel against the DN in the reference pixel for all the frames in each run. Since the response of each pixel is approximately linear over the dynamic range of our observations and most of the signal is sky background, which varies with time but is uniform across the chip, the plotted points for each pixel describe a straight line; if all the pixels had the same response, the slope of each line would equal 1. We divided each pixel's DN by the actual slope of its response curve relative to the reference

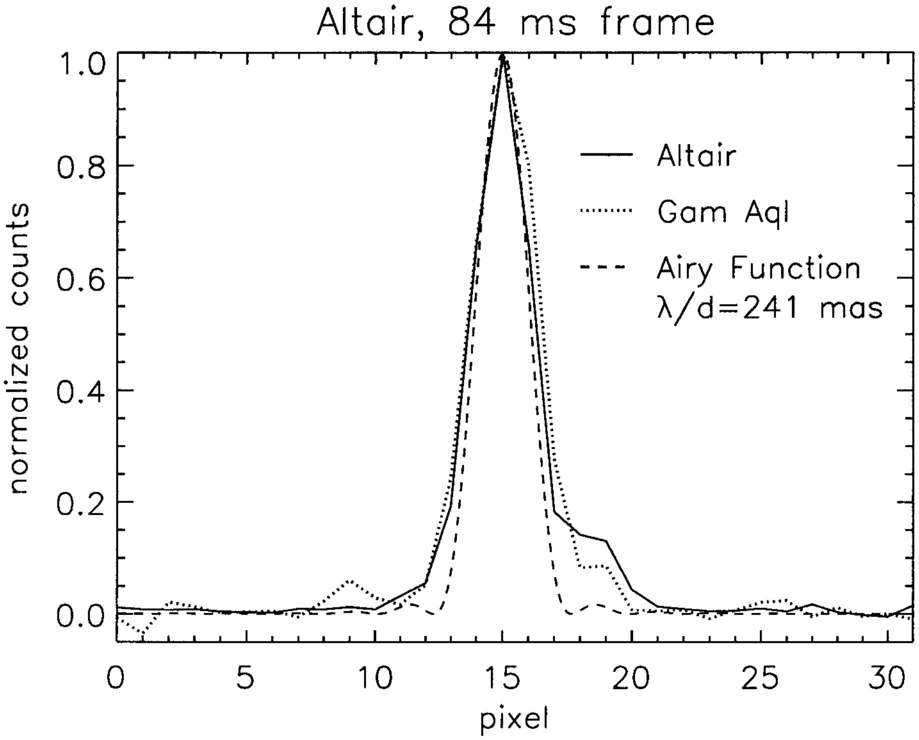

Figure 3.1: A cut through a single 4 ms image of Altair, compared to a similar cut through an image of a calibrator star, Gamma Aquila, and an Airy function representing the PSF of an ideal, filled, 10-meter aperture at 11.6 microns. The cores of the images are diffraction-limited, but the wings are sensitive to the instantaneous seeing, making speckle analysis necessary.

pixel, effectively matching all pixels to the reference pixel. We then interpolated over bad pixels, frame by frame.

To compensate for the differences in the thermal background between the two nod positions, we averaged together all the on-axis sky frames to measure the on-axis thermal background and subtracted this average from all of the on-axis frames—both object and sky. We used the same procedure to correct the off-axis frames.

Next, we chose subframes of 32 by 32 pixels on each image, centered on the star (or for sky frames, the location of the star in an adjacent object frame), and processed these according to classical speckle analysis (Labeyrie 1970). We Fourier transformed

Table 3.1: Observations with LWS

Date	Target	Calibrator	Time Per Frame (ms)	Object Frames	Pairs	Log Disk Density Solar Disk = 0
Aug. 3	Vega	R Lyr	800	90	2	≤ 4.0
		κ Lyr	800	90	1	
	61 Cyg A	ζ Cyg	800	90	4	...
	61 Cyg B	ζ Cyg	800	90	5	...
	τ Cet	ν Cet	800	90	2	...
			210	348	1	
Aug. 4	70 Oph B	β Oph	84	864	5	...
		74 Oph	84	864	4	
	Altair	γ Aql	84	864	6	≤ 3.2
		β Aql	84	864	5	

them, and summed the power spectra, yielding a sky power spectrum and an object power spectrum for each series. Then we azimuthally averaged the power spectra in the u-v plane—that is, we averaged over all the frequency vectors of a given magnitude, $\sqrt{u^2 + v^2}$. This azimuthal averaging corrects for the rotation of the focal plane of the alt-az-mounted Keck telescope with respect to the sky. We then subtracted from every object power spectrum the corresponding sky power spectrum and divided each corrected target power spectrum by the corrected power spectrum of a calibrator star observed in the same manner as the target star immediately before or after the target star. Figure 3.4 shows an azimuthally-averaged power spectrum of Altair and the corresponding sky power spectrum, compared with a power spectrum of calibrator Gamma Aquila and its corresponding sky power.

We then averaged all the calibrated power spectra for a given target. If the object and calibrator are both unresolved, the average calibrated power spectrum should be the power spectrum of the delta function: a constant. We found that the pixels along the u and v axes of the power spectra were often contaminated by noise artifacts from the detector amplifiers, so we masked them out.

Figures 3.3–3.8 show the calibrated azimuthally-averaged power spectra for our

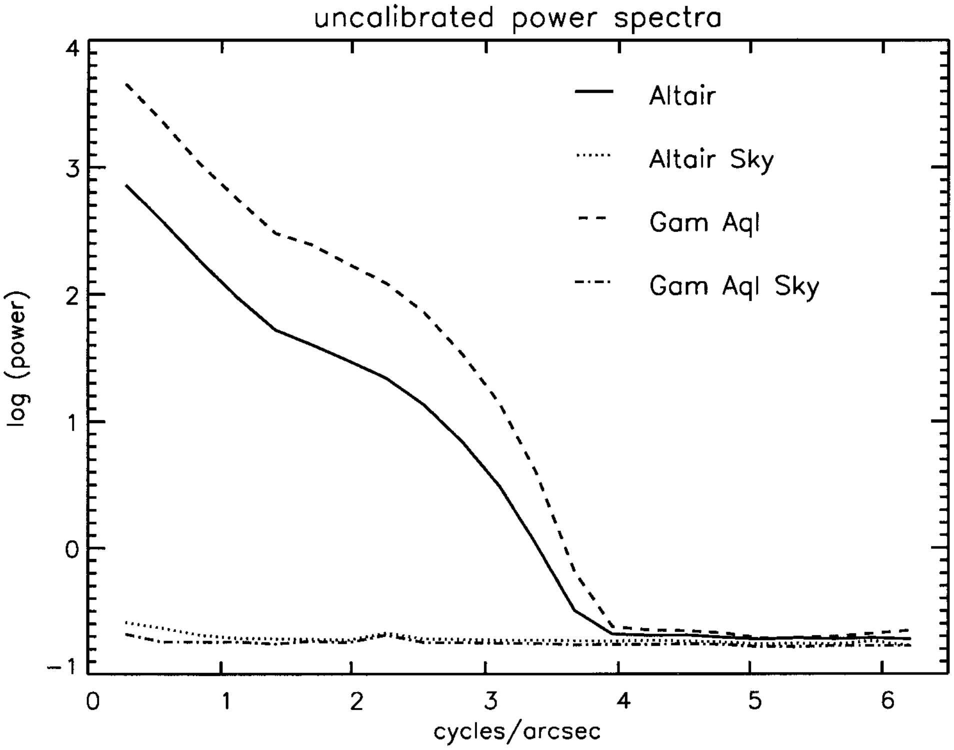

Figure 3.2: An azimuthally-averaged power spectrum of Altair and the corresponding sky power spectrum, compared with a power spectrum of calibrator Gamma Aquila and its corresponding sky power. The power in the star images approaches the sky power near the diffraction limit at 4 cycles per arcsecond.

target stars. To compare different power spectra from the same target, we normalized each azimuthally-averaged power spectrum so that the geometric mean of the first 10 data points in each spectrum equals 1. For Altair and 61 Cygni A and B we had more than three pairs of target and calibrator observations, i.e., calibrated power spectra, so we show the average of all the spectra and error bars representing the 68% confidence interval for each datum, estimated from the variation among the individual power spectra. The error is primarily due to differences in the atmosphere-telescope transfer function between object and calibrator. None of the calibrated power spectra deviate from a straight line by more than a typical error; all the targets are unresolved to the accuracy of our measurements.

3.2.3 Discussion

To interpret our observations we compared them to models of the IR emission from the solar zodiacal cloud. We constructed a model for exozodiacal emission based on the smooth component of the Kelsall et al. (1998) model of the solar system zodiacal cloud as seen by COBE DIRBE, with emissivity $\epsilon \propto r^{-0.34}$ and a temperature $T = 286\ \mathrm{K}\ r^{-0.467} L^{0.234}$, where r is the distance from the star in AU, and L is the luminosity of the star in terms of $L_{\odot}$. For a dust cloud consisting entirely of a single kind of dust particle of a given size and albedo, the L exponent in the expression for the temperature is simply $-1/2$ times the r exponent (Backman & Paresce 1993).

The physics of the innermost part of the solar zodiacal dust is complicated (see Mann & MacQueen 1993), but our results are not sensitive to the details, because the hottest dust is too close to the star for us to resolve. We assume that the dust sublimates at a temperature of 1500 K, and allow this assumption to define the inner radius of the disk. We set the outer radius of the model to 3 AU, the heliocentric distance of the inner edge of our own main asteroid belt. Our conclusions are not sensitive to this assumption; decreasing the outer radius to 2 AU or increasing it to infinity makes a negligible difference in the visibility of the model, even for A stars.

The assumed surface density profile, however, does make a difference. A collisionless cloud of dust in approximately circular orbits spiraling into a star due to Poynting-Robertson drag that is steadily replenished at its outer edge attains an equilibrium surface density that is independent of radius (Wyatt and Whipple 1950, Briggs 1962). Models that fit data from the Helios space probes (Lienert et al. 1981), the fit by Kelsall et al. (1998) to the COBE DIRBE measurements and Good's (1997) revised fit to the IRAS data all have surface densities that go roughly as $r^{-0.4}$. This distribution appears to continue all the way in to the solar corona (MacQueen & Greely 1995). We find that, in general, if we assume an $r^{-\alpha}$ surface density profile, our upper limit for the 1 AU density of a given disk scales roughly as $10^{\alpha/2}$; disks with more dust towards the outer edge of the 11.6 micron emitting region are easier to resolve.

Likewise, the assumed temperature profile strongly affects our upper limits. Unfortunately, we know little about the temperature profile of the solar zodiacal cloud. COBE DIRBE and IRAS only probed the dust thermal emission near 1 AU, and Helios measured the solar system cloud in scattered light, which does not indicate the dust temperature. We found that a dust cloud model with the IRAS temperature profile ($T = 266$ K $r^{-0.359}L^{0.180}$) was much easier to resolve than the model based on DIRBE measurements that we present here, especially for G and K stars.

To compare the models with the observations, we synthesized high resolution images of the model disks at an inclination of 30 degrees. We calculated the IR flux of the stars from the blackbody function, and obtained the parallaxes of the stars from the Hipparcos Catalog (ESA 1997). We inferred stellar radii and effective temperatures for each star from the literature and checked them by comparing the blackbody fluxes to spectral energy distributions based on photometry from the SIMBAD database (Egret et al. 1991). For Altair and Vega, we use the interferometrically measured angular diameters (1974) (they are 2.98 +/-0.14 mas and 3.24 mas). Stellar fluxes typically disagree with fitted blackbody curves by $\sim 10\%$ in the mid-infrared (Engelke 1990), but our method does not require precise photometry, and the blackbody numbers suffice for determining conservative upper limits. We computed the power spectra of the images, and normalized them just like the observed power spectra. In Figures 3.3–3.8, the azimuthally-averaged power spectra for our target stars are compared to the extrapolated COBE DIRBE model at a range of model surface densities. Disks with masses as high as 10^3 times the mass of the solar disk will suffer collisional depletion in their inner regions, so they are unlikely to have the same structure as the solar disk. By neglecting this effect we are being conservative in our mass limits. The density of the densest model disk consistent with the data in each case is listed in Table 3.1.

Altair

Our best upper limit is for Altair (spectral type A7, distance 5.1 pc); with 11 pairs of object and calibrator observations we were able to rule out a solar-type disk a few

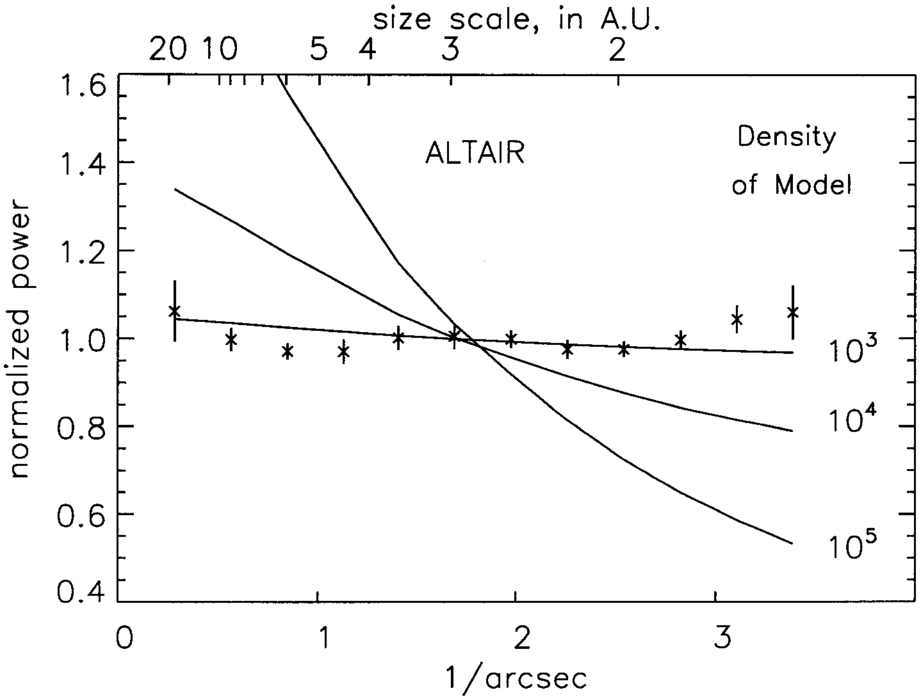

Figure 3.3: Azimuthally-integrated power spectrum of Altair compared to simulated power spectra of model disks with various densities (1 = the solar disk). An unresolved point-source would appear as a straight line at a normalized power of 1.0. The densest model disk consistent with the observations has a density of roughly 10^3 times that of the solar disk.

times 10^3 as dense as our zodiacal cloud. Such a disk would have been marginally detectable by IRAS as a photometric excess.

Vega

IRAS detected no infrared excess in Vega's spectral energy distribution at 12 microns, with an uncertainty of 0.8 Jy. This may be due to a central void in the disk interior to about 26 AU (Backman & Paresce 1993). Aumann et al. (1984) suggested that Vega (A0, 7.8 pc) could have a hot grain component (500 K) with up to 10^{-3} of the grain area of the observed component and not violate this limit. The apparent upward trend in the visibility data may be a symptom of resolved flux in the calibrator stars.

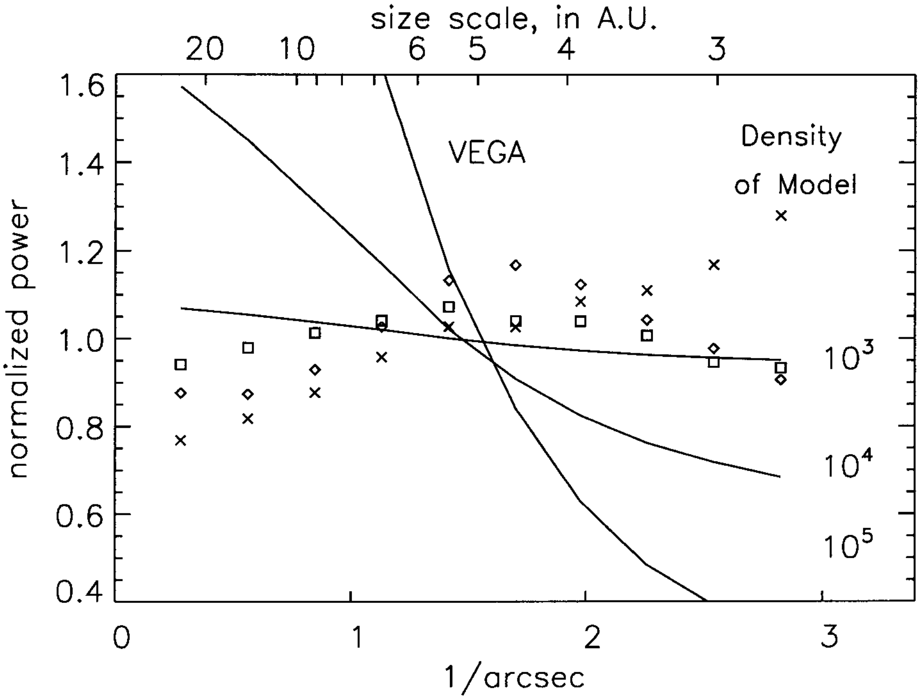

Figure 3.4: Azimuthally averaged power spectrum of Vega compared to simulated power spectra of model disks with various densities.

We have only 3 object/calibrator pairs for Vega, not enough to test this hypothesis. Our upper limit is a solar-type disk with approximately 3×10^3 times the density of the solar disk. This disk would have a ≥500 K emitting area of 10^{24} cm^2, about 10^{-3} of the grain area of the observed component.

61 Cygni A and B

Though 61 Cygni is close to the galactic plane and surrounded by cool cirrus emission, Backman, Gillett and Low (1986) identified an IRAS point source with this binary system and deduced a far-infrared excess not unlike Vega's. The color temperature of the excess suggests the presence of dust at distances > 15 AU from either star. However, these stars are dim (spectral types K5 and K7) and the region of the disk hot enough to emit strongly at 11.6 microns is close to the star and difficult to resolve; we

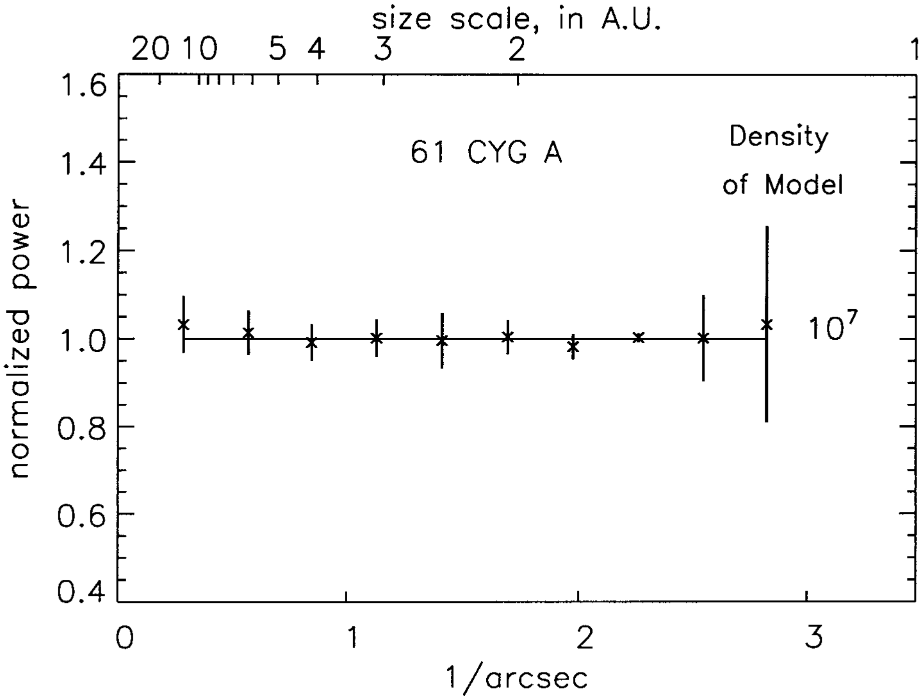

Figure 3.5: Azimuthally averaged power spectrum of 61 Cygnus A compared to simulated power spectra of model disks with various densities.

could not detect a solar-type dust disk around either of these objects at any density, assuming the COBE DIRBE model, or unless it had 10^5 times the density of the solar disk, assuming the IRAS model.

70 Oph B

70 Oph is a binary (types K0 and K4) with a separation of 24 pixels (2.6 arcsec). We were able to assemble a power spectrum for B from 9 object/calibrator pairs, but the image of A fell on a part of the LWS chip that suffered from many bad pixels and was unusable. The image of A may also have been distorted by off-axis effects. 70 Oph B, like 61 Cygni A and B, is dim, making any dust around it cool and hard to detect at 11.6 microns.

τ Ceti

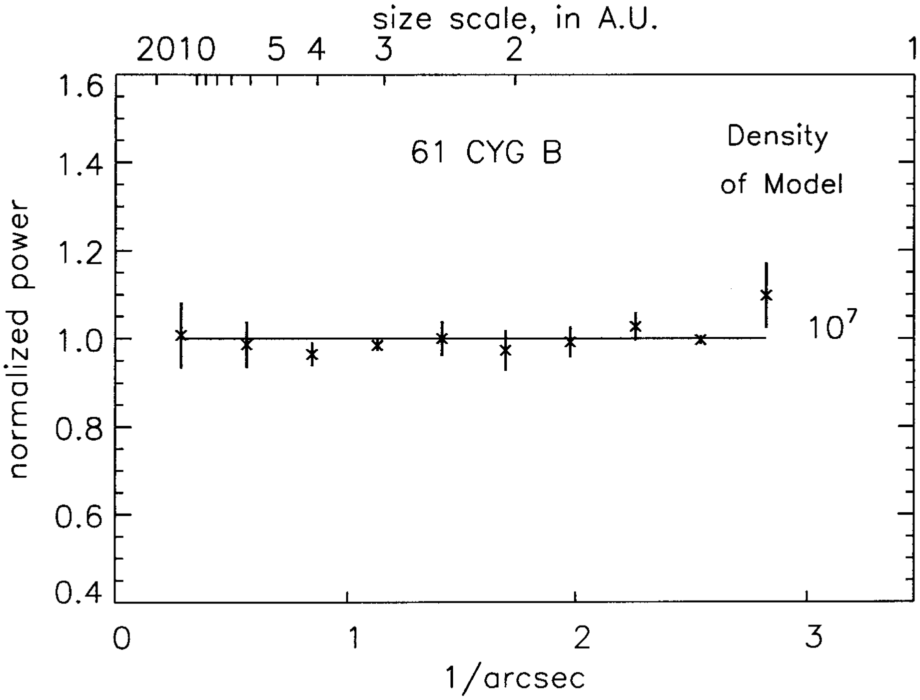

Figure 3.6: Azimuthally averaged power spectrum of 61 Cygnus B compared to simulated power spectra of model disks with various densities.

IRAS could have barely detected a disk with ~ 1000 times the emitting area of the solar disk around Tau Ceti (G8, 3.6 pc), the nearest G star. We have only three object/calibrator pairs for this object, not enough data to improve on this limit.

We are grateful to Dana Backman, Alycia Weinberger, Keith Matthews and Eric Gaidos for helpful discussions, and to Keith Matthews and Shri Kulkarni for assistance with the observations. This research has made use of the Simbad database, operated at CDS, Strasbourg, France. The observations reported here were obtained at the W. M. Keck Observatory, which is operated by the California Association for Research in Astronomy, a scientific partnership among California Institute of Technology, the University of California, and the National Aeronautics and Space Administration. It was made possible by the generous financial support of the W. M. Keck Foundation.

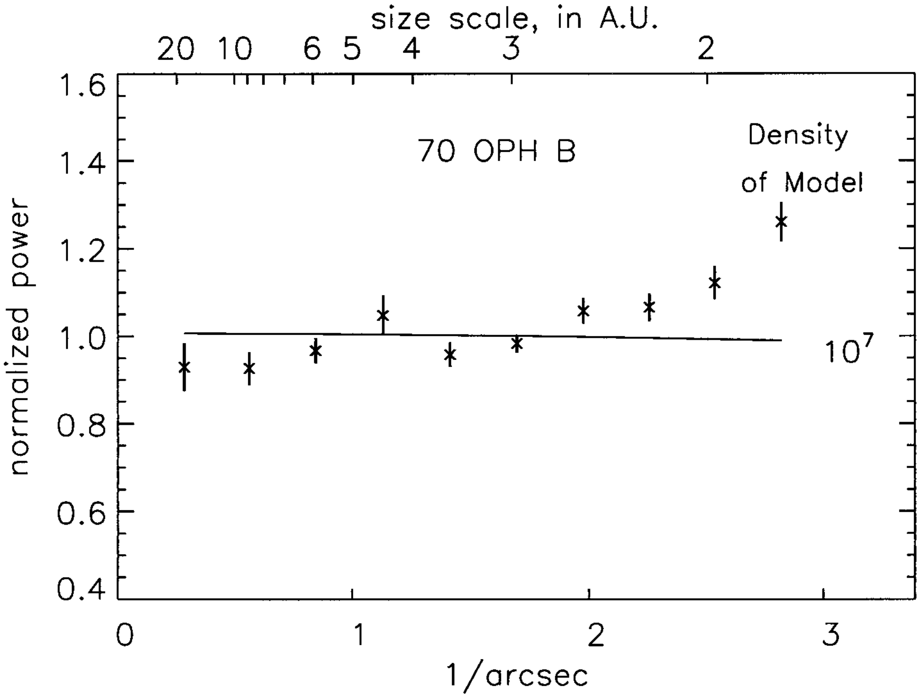

Figure 3.7: Azimuthally averaged power spectrum of 70 Oph B compared to simulated power spectra of model disks with various densities.

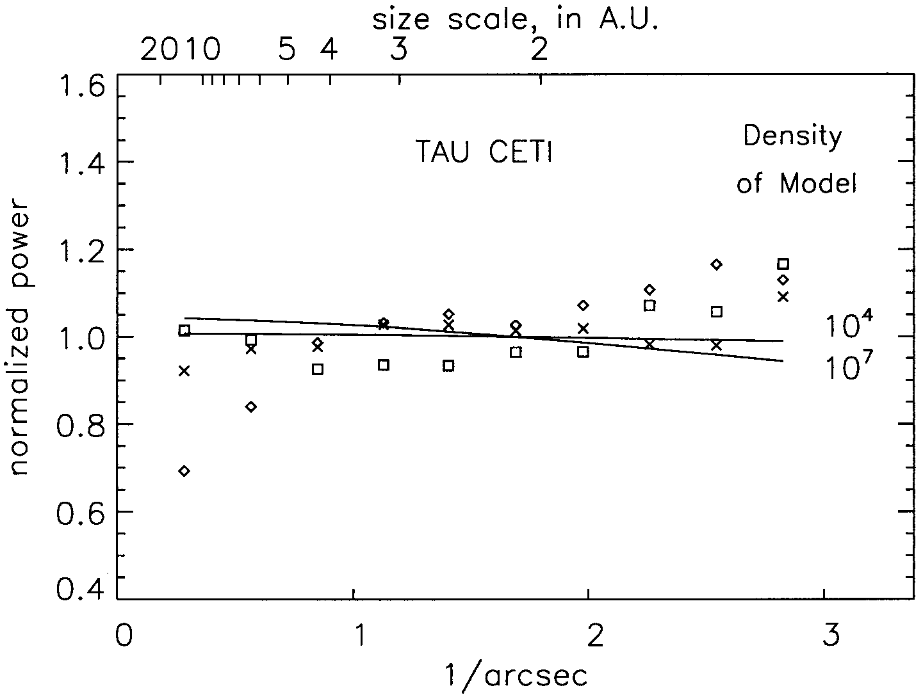

Figure 3.8: Azimuthally averaged power spectrum of Tau Ceti compared to simulated power spectra of model disks with various densities.

Chapter 4

A Search For Scattered Light from Exozodiacal Dust

4.1 The One-dimensional Coronagraph Without an Atmosphere

[1] To set the stage for discussion of a more realistic instrument, we illustrate the Fourier optics of a one-dimensional coronagraph. Other more formal expositions of coronagraphic imaging can be found in the literature. Our analysis assumes that the Fraunhofer approximation applies, that is the transverse electric field in the image plane is the Fourier transform of the phasor of the wavefront phase in the pupil plane (if $\phi(x, y)$ is the phase, then $e^{i\phi(x,y)}$ is the corresponding phasor). We make use of the standard Fourier analysis results which can be found in Bracewell (1986).

For this analysis we consider only monochromatic imaging, but note that in typical broadband imaging, the final image can be described by the sum or integral of several monochromatic images, weighted by the instrumental transmission function. Wavelength variation across the band will act in such a way as to smear image features radially by the same factor as the fractional bandwidth, since the wavelength enters into diffraction-limited image formation only in the combination (λ/D). As a result, bright Airy rings will get wider, but coronagraphic suppression of such rings will persist. By treating the monochromatic case we can distinguish clearly between halo suppression and Airy ring suppression. Secondary support spiders and scintillation (field strength variation across the wavefront) are not modeled here.

In the absence of atmospheric degradation, a monochromatic on-axis source at

---

[1] Section 4.1 is adapted from Sivaramakrishnan et al. (2000)

infinity produces a transverse electric field at the telescope pupil

$$E = E_0\, Re(e^{i(kz-\omega t)}). \tag{4.1}$$

We follow the passage of the incident wave's field through a one-dimensional coronagraph. Figure 4.1 shows a diagram of the optical path.

We label eight key locations along this path with letters a–h. Eight plots in the figure show the electric field due to an on-axis source at these key locations, and the transmission functions of the optical stops that affect the incident wave as it passes through the coronagraph.

First, the incoming wave passes through the telescope aperture (Figure 4.1a). We represent this interaction by multiplying the field by the aperture stop function, so that in the pupil plane

$$E_a = E\,\Pi(x/D_\lambda), \quad \text{where} \quad \begin{aligned} \Pi(x) &= 1 \quad \text{for } |x| < 1/2, \\ \Pi(x) &= 0 \quad \text{elsewhere.} \end{aligned} \tag{4.2}$$

Here $D_\lambda = D/\lambda$ is the number of wavelengths across the telescope aperture. We denote a pupil plane coordinate by x, and an image plane coordinate by θ. The telescope optics then form the wave into an image (Figure 4.1b). The electric field in the image plane is the Fourier transform of the aperture field E_a:

$$E_b \propto \mathrm{sinc}(D_\lambda \theta/2), \tag{4.3}$$

where θ is the field angle in radians in the first image plane. We omit the constants of proportionality for simplicity.

In a conventional imaging camera this image field would fall on a detector. However, in a coronagraph, the star is occulted by a field stop in this image plane. We describe the stop in terms of a shape function $w(D_\lambda\theta/s)$, which is unity where the stop is opaque and zero where the stop is absent. If $w(\theta)$ has a width of order unity, the stop size will be of the order of s resolution elements. The transmission function in the image plane is therefore $1 - w(D_\lambda \theta/s)$ (Figure 4.1c). To illustrate the present discussion, we take $w(\theta) = \exp(-\theta^2/2)$. The field in the first imaging plane after the occulting stop (Figure 4.1d) can be written as

$$E_d \propto \operatorname{sinc}(D_\lambda \theta/2)(1 - w(D_\lambda \theta/s)). \tag{4.4}$$

This image is relayed to a detector through a second pupil plane. The electric field at this second pupil is the Fourier transform of the occulted image field (see Figure 4.1e):

$$E_e \propto \Pi(x/D_\lambda) * (\delta(x) - \frac{s}{D_\lambda} W(s\,x/D_\lambda)) \tag{4.5}$$

Here W is the Fourier transform of the image stop function w, $\delta(x)$ is the Dirac delta function, and $*$ denotes convolution. W has width of order unity, although it will not have bounded support for occulting stop shape functions of finite extent, such as the hard-edged stop $w(\theta) = \Pi(\theta)$ (the support of a function is the set of points at which the function is non-zero). The geometrical significance of equation 4.5 becomes clear if we rewrite it as

$$E_e \propto \Pi(x/D_\lambda) - \frac{s}{D_\lambda} \Pi(x/D_\lambda) * W(s\,x). \tag{4.6}$$

If the image stop is completely opaque at its center, $w(0) = 1$. This means that its transform, $(s/D_\lambda)\ W(sx/D_\lambda)$, has unit area, regardless of any scaling we have performed on the argument of w. This makes for cancellation of the field across most of the pupil when $s >> 1$. In Figure 4.2 we show how the equation 4.6 is constructed graphically, using a Gaussian image stop whose width is $5\lambda/D$ (i.e., $s = 5$). This is why the Lyot stop must mask out a border of order D/s wide around the pupil boundary to produce significant reduction in the throughput of unocculted light from the on-axis source. It is only at this stage that the coronagraph increases the dynamic range of the final image.

The unocculted light (Figure 4.1d) has a highly periodic distribution, with periodicity $\sim \lambda/D$. In the following pupil plane (which is the transform space of the image plane), this energy is concentrated near $\pm D/2$. The larger the occulting stop diameter, the more E_d looks like a pure sinusoid, and the more the unocculted energy is localized in the neighborhood of the boundary of the following pupil.

In seeing-limited coronagraphs, the occulting stop is typically many diffraction widths in size ($s \geq 10$). Consequently, the Lyot stop need only be undersized by a small fraction of the pupil diameter (e.g., 10% or less), resulting in minimal loss of throughput for unocculted, off-axis sources. In contrast, off-axis throughput in an optimized, diffraction-limited coronagraph with significant rejection of on-axis light must fall dramatically as the image plane stop shrinks to a few diffraction widths. This is because the spillover of unocculted on-axis light occurs in a wide border around the pupil boundary in the plane of the Lyot stop. Thus the Lyot stop must obscure a significant fraction of the re-imaged primary mirror to remove the on-axis spillover, and as a consequence reduce the off-axis throughput.

Since the scale of the Lyot stop oversizing is D/s, we fine-tune the Lyot stop diameter so that it obscures a border $\mathcal{F}D/s$ around the perimeter of the primary. Therefore, the Lyot stop diameter is

$$D_{Lyot} = D - 2\mathcal{F}D/s. \tag{4.7}$$

If a secondary obstruction is present, then the Lyot stop must block out a similar border around the inside edge of the annular pupil. This is why small secondary mirrors benefit diffraction-limited coronagraphy.

For the case of an unobstructed primary aperture with an image plane stop of $5\lambda/D$ (i.e., $s = 5$), then, using the theory outlined above, approximately $16/25$ of the aperture should be obscured by a matched Lyot stop: when projected back onto the primary pupil, the Lyot stop is opaque outside a circle of diameter $\sim 3D/5$. Rejection of unwanted on-axis light must be balanced by signal-to-noise considerations pertaining to the off-axis source brightness. This places a practical lower limit on the angular size of the occulting spot in the first image plane.

The above arguments hold for two-dimensional apertures as well. The derivation is analogous to the one-dimensional case, although the functions and transforms become two-dimensional (e.g., for a circular telescope aperture, the sinc function is replaced by the Airy function). Wang & Vaughan (1988) describe the two-dimensional case, and Malbet (1996) treats the PSF's of off-axis sources in such coronagraphs.

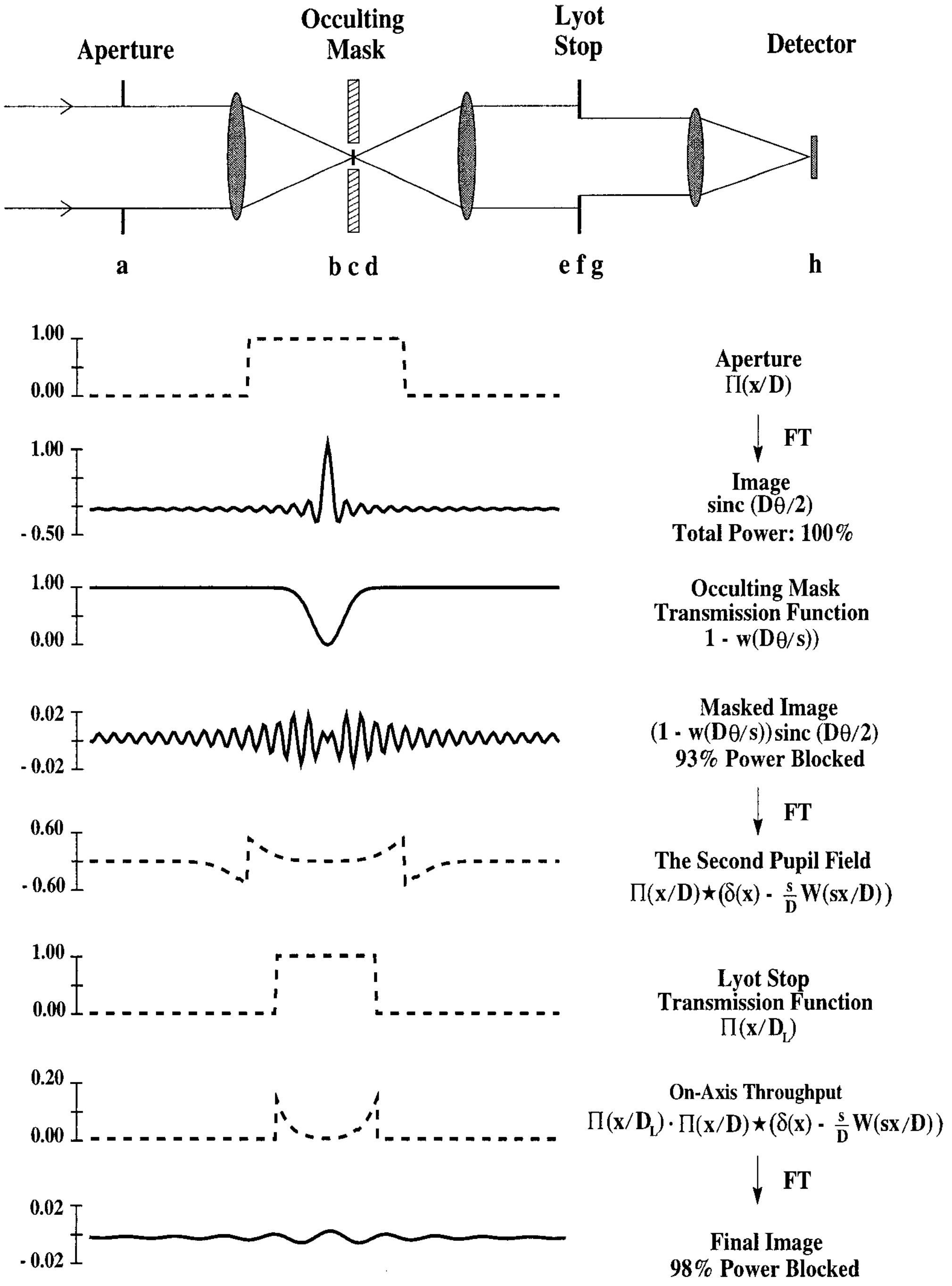

Figure 4.1: One-dimensional coronagraph summary, with locations and field or stop profiles of: (a) primary pupil for on-axis source; (b) image before field stop; (c) field stop; (d) image after field stop; (e) pupil before Lyot stop; (f) Lyot stop; (g) pupil after Lyot stop; and (h) final on-axis image. We note that, in this example, 98% of the incident power is blocked by the coronagraph.

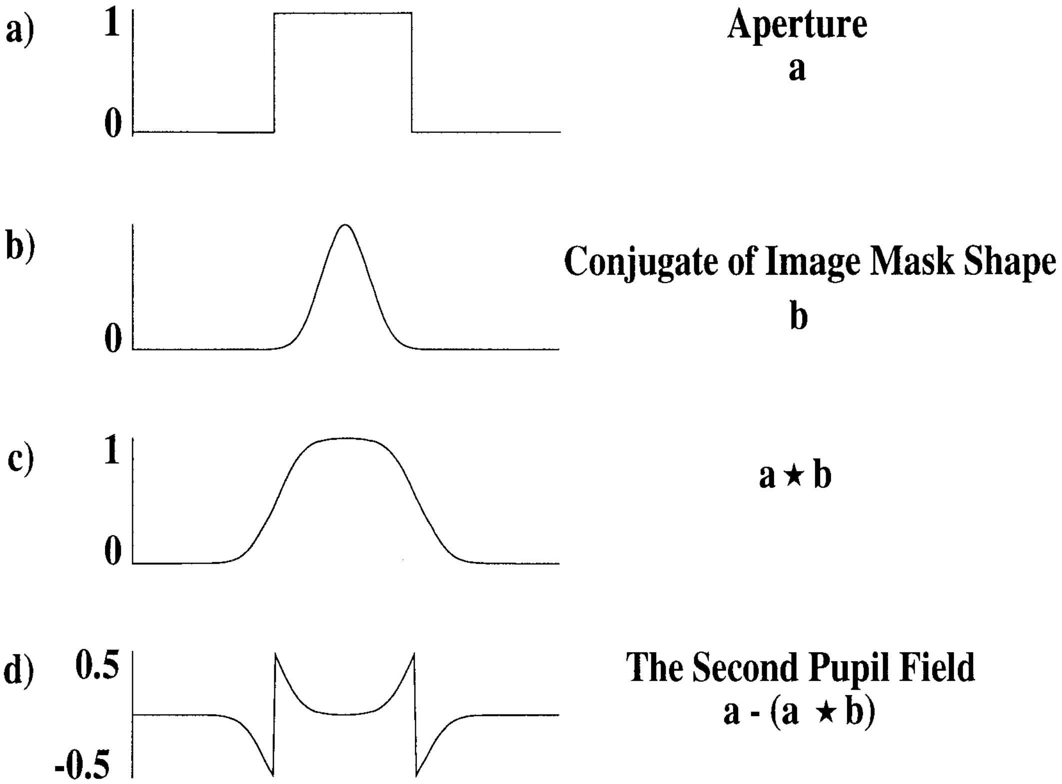

Figure 4.2: Graphical representation of the Lyot plane field calculation. (a) Pupil function of width D; (b) Gaussian profile field stop with $4\lambda/D$ standard deviation produces a Gaussian with standard deviation $D/4$ in the Lyot plane; (c) the convolution of the pupil function with the transform of the stop profile; (d) the final Lyot stop field showing bright edges and no energy in the center.

4.2 A Search for Exozodiacal Dust and Faint Companions Near Sirius, Procyon, and Altair with the NICMOS Coronagraph

Marc J. Kuchner
Palomar Observatory, California Institute of Technology, Pasadena, CA 91125
Michael E. Brown
Division of Geological and Planetary Sciences, California Institute of Technology,
Pasadena, CA 91125

We observed Sirius, Altair, and Procyon with the NICMOS Coronagraph on the Hubble Space Telescope to look for scattered light from exozodiacal dust and faint companions within 10 AU from these stars. We did not achieve enough dynamic range to surpass the upper limits set by IRAS on the amount of exozodiacal dust in these systems, but we did set strong upper limits on the presence of nearby late-type and sub-stellar companions. [2]

4.2.1 Introduction

Several main sequence stars are close enough that a large telescope operating at the diffraction limit can resolve the terrestrial temperature zone within 10 AU from the star (Kuchner, Brown & Koresko 1998). We used the NICMOS coronagraph to image three of the nearest main-sequence stars, Sirius (α CMa; HR 2491), Procyon (α CMi; HR 2943), and Altair (α Aql; HR 7557) in the near-infrared to look for circumstellar material in this relatively uncharted circumstellar region.

Besides the Sun, the most luminous component of the central region of our solar system is a cloud of dust which forms when asteroids collide and when comets outgass. Similar dust around other stars is called "exozodiacal" dust. Any dust orbiting close to one of our targets must have been generated recently by some population of larger

---

[2] Adapted from Kuchner and Brown (2000)

bodies, since small grains near a star quickly spiral into the star due to Poynting-Robertson drag (Robertson 1937). Ten micron diameter dust 3 AU from a G star spirals into the star on time scales of $\sim 10^5$ years; this time scale is inversely proportional to the stellar luminosity. A search for exozodiacal dust is therefore implicitly a search for extra-solar asteroid or comet-like bodies that make dust.

Several disks around nearby main-sequence stars appear to have exozodiacal components. For some systems, like β Pictoris, the presence of warm dust is inferred from analysis of a silicate emission feature at 10 microns (Telesco & Knacke 1991; Knacke et al. 1993). Others, like the disk around HR 4796, show marginally resolved emission at 10 microns that is interpreted as exozodiacal (Koerner et al. 1998). Dust clouds like these, which have ~ 1000 times as much warm dust as our sun, emit thermal radiation substantially in excess of the stellar photospheric emission, and can often be detected photometrically by studying the spectral energy distribution of the star in the mid-infrared. However, many less massive exozodiacal clouds may never be detectable photometrically because no stellar spectrum is known to better than $\sim 3\%$ in the mid-infrared (Cohen et al. 1996). We have begun to search for disks that are too faint to be detected photometrically by spatially resolving the critical regions less than 10 AU from nearby stars.

Coronagraphic images can also reveal faint companions to nearby stars. Such companions can go undetected by radial velocity surveys because of their small masses or long orbital periods. Siruis and Procyon both have white dwarf companions whose orbits are well studied, but analyses of the orbital motion in these systems leave room for additional low mass companions.

The Sirius system in particular, so prominent in the night sky, has spurred much debate in the last century over its properties. Three analyses of the proper motion of Sirius have suggested that there may be a perturbation in the orbit of Sirius B with a ~ 6 year period (Volet 1932; Walbaum & Duvent 1983; Benest & Duvent 1995). These analyses do not indicate whether the perturbing body orbits Sirius A or B, and dynamical simulations indicate that stable orbits exist around both Sirius A and B at circumstellar distances up to more than half the binary's periastron separation (Benest 1989). If such a companion were in a simple face-on circular orbit, it would appear at a separation of 4.2 AU (1.6 arcsec) from Sirius A or a separation of 3.3 AU (1.3 arcsec) from Sirius B assuming that the masses for Sirius A and B are 2.1 and 1.04 $M_\odot$ respectively (Gatewood & Gatewood 1978). Benest & Duvent (1995) do not derive a mass for the hypothetical companion from observations of the system, but they estimate that a perturber much more massive than 0.05 $M_\odot$ would rapidly destroy the binary.

Perhaps the most interesting debate about Sirius is whether or not the system appeared red to ancient observers $\sim$ 2000 years ago. Babylonian, Graeco-Roman and Chinese texts from this time period have separately been interpreted to say that Sirius was a red star (Brecher 1979; Schlosser & Bergman 1985, Bonnet-Bidaud & Gry 1991). Tang (1986), van Gent (1984) and McCluskey (1987) have attacked some of these reports, claiming that they represent mistranslations or misidentifications of the star, and Whittet (1999) suggests that the coloration was reddening due to the Earth's atmosphere. However, Bonnet-Bidaud & Gry (1991) claim that if the Sirius system did indeed appear red, the existence of a third star in the group interacting periodically with Sirius A could explain the effect.

The low mass companions (< 0.1 $M_\odot$) that we could hope to detect with NICMOS are late-type stars or warm brown dwarfs, shining with their own thermal power in the near-infrared. Schroeder et al. (2000) recently imaged Sirius, Procyon and Altair at 1.02 microns with the Planetary Camera on HST in a search for faint companions to nearby stars. Our observations are $\sim$ 3 times more sensitive to faint companions because of the coronagraph, but their images extend to angular separations of 17 arcseconds, while we only have nearly complete coverage of the central 3.5 arcseconds.

4.2.2 Observations

We observed our target stars with the NICMOS Camera 2 coronagraph on five dates during 1999 October. We used the F110 filter, the bluest available near-infrared filter, with an effective wavelength of 1.104 microns, to take advantage of the higher dynamic range the coronagraph has at shorter wavelengths. We took images of Sirius and Procyon at two different position angles, effectively rolling the telescope about the axis to the star by 15° between them. When we searched for faint companions in the images, we subtracted the images taken at one roll angle from the images taken at the other angle to cancel the light in the wings from the image of the occulted star. We planned to image Altair at a second roll angle, but on our second visit to the star the telescope's Fine Guidance Sensors failed to achieve fine lock on the guide star due to "walkdown" failure.

At each roll angle we took 50 short exposures in ACCUM mode, lasting 0.6 seconds each, and we co-added them, for total integration times of 30 seconds. Even though we used the shortest available exposure times, our images saturated interior to about 1.9 arcseconds for Sirius, 1.4 arcseconds for Procyon, and 0.7 arcseconds for Altair. The actual coronagraphic hole is only 0.3 arcseconds in radius. Table 4.1 summarizes the timing of our observations and the position angles (East of North) of the Y-axes of the images. This table also lists the distances to the targets, taken from the Hipparcos Catalogue (Perryman et al. 1997), and their J magnitudes, from the SIMBAD online database.

Table 4.1: Observations with NICMOS

Target	Spectral Type	Distance (pc)	J	UT Date	Position Angle
Sirius	A1V	2.64	-1.34	October 20	64.51°
				October 22	81.51°
Procyon	F5IV	3.50	-0.40	October 9	40.51°
				October 21	55.51°
Altair	A7V	5.14	0.39	October 14	-126.24°

Figure 4.1 shows an image of Sirius taken at one roll angle. The white dwarf Sirius B appears to the left of Sirius A, at a separation of 3.79 arcseconds. We derived photometry of Sirius B at 1.1 microns from the roll subtracted image of Sirius using a prescription from Rieke (1999). We measured the flux in circular apertures with radii 7.5 pixels around the positive and negative images and multiplied the flux in those regions by an aperture correction of 1.110 to extrapolate to the total flux. Then we used a factor of 1.996×10^{-6} Jy/ADU/Sec to convert from ADU to Janskys. In this manner, we measured the flux in Sirius B to be 0.503 ± 0.15 Jy. Procyon also has a white dwarf companion, Procyon B, that has been previously detected by HST (Provencal et al. 1997). It is not visible in our images, because it is currently at a separation of ~ 5 arcseconds from Procyon A.

4.2.3 Exozodiacal Dust

We compared our observations of Sirius, Procyon and Altair to a simple model for what our zodiacal cloud would look like if it were placed around these stars. Kelsall et al. (1998) fit an 88-parameter model of the zodiacal cloud to the maps of the infrared sky made by the Diffuse Infrared Background Experiment (DIRBE) aboard the Cosmic Background Explorer (COBE) satellite. We used the smooth component of this model, which has a face-on optical depth of $7.11 \times 10^{-8} (r/1\mathrm{AU})^{-0.34}$, and extrapolated it to an outer radius of 10 AU.

Kelsall et al. (1998) fit a scattering phase function and an albedo to the DIRBE zodiacal cloud observations at 1.25 microns, but because DIRBE surveyed a narrow range of solar elongation, these functions are wrong at scattering angles less than 60 degrees. Instead, we used a scattering phase function consisting of a linear combination of three Henyey-Greenstein functions that Hong (1985) fit to visible light observations of the zodiacal cloud with the Helios Satellite, and we assumed an albedo of 0.2, from the Kelsall et al. (1998) fit to the 1.25 micron DIRBE maps. This phase function describes observations at scattering angles as low as 15°. We trust this extrapolation

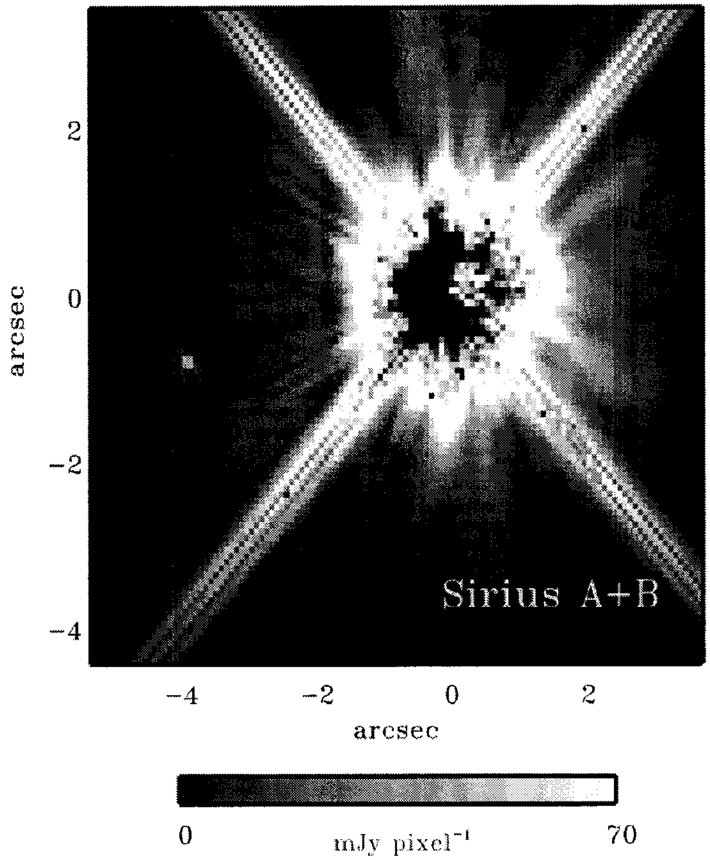

Figure 4.3: A coronagraphic image of the Sirius system. The white dwarf Sirius B appears to the left of the residual light from Sirius A. Even though we used the shortest available exposure time, the region < 1.9 arcsec from Sirius A is saturated.

because zodiacal dust is nearly grey between 0.5 and 1.25 microns.

For this part of our search, we could not use roll-subtraction to cancel the light in the images of our target stars, because this approach would also cancel most of the light from an exozodiacal disk, even if the disk were edge-on. Instead we subtracted images of Altair from the images of Procyon and Sirius, with the assumption that all three of our stars would not have identical circumstellar structures. We used the IDP3 data analysis software (Lytle et al. 1999) to perform sub-pixel shifts on the images of Altair before we subtracted them from our images of Sirius and Procyon to compensate for the slightly different relative alignments of the three stars and the coronagraphic hole.

Figures 4.2a and 4.3a show our images of Sirius and Procyon minus our image of Altair. Software masks hide the regions where the images are saturated and the four main diffraction spikes. The bright horn just above the masked area in the Procyon image is a well known NICMOS artifact.

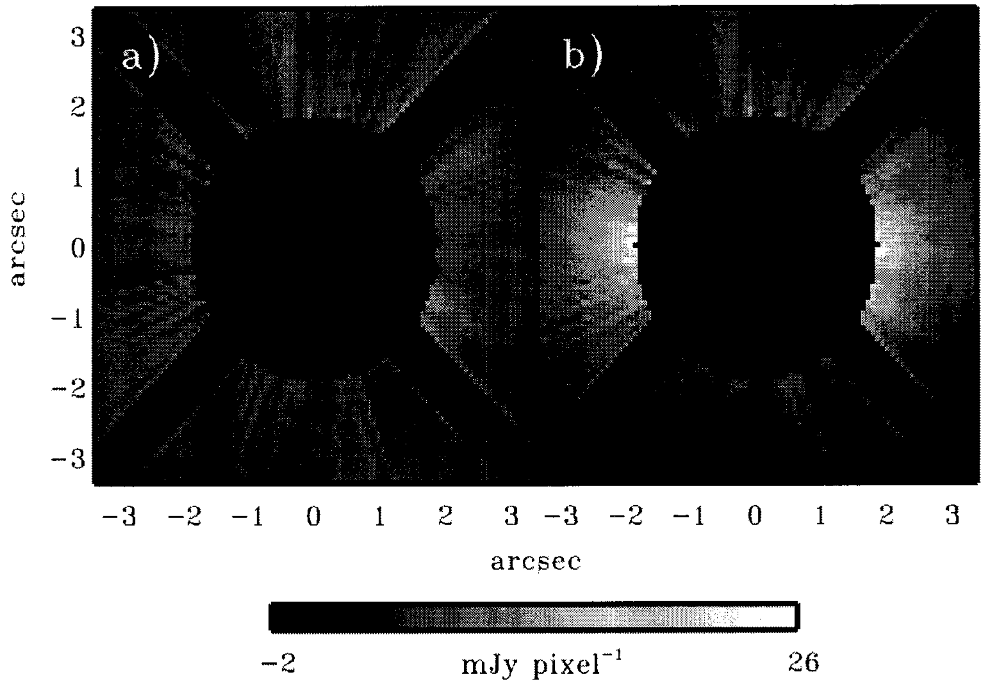

Figure 4.4: a) An image of Sirius made using our coronagraphic image of Altair to cancel the wings of the occulted PSF. The saturated regions of the image are hidden with a software mask. b) The same image plus a model of the scattered light from an exozodiacal cloud similar to the solar zodiacal cloud but 2.5×10^5 times as bright.

Figures 4.2b and 4.3b show the same images plus synthesized images of exozodiacal clouds seen in scattered light. The models are brightest immediately to the left and

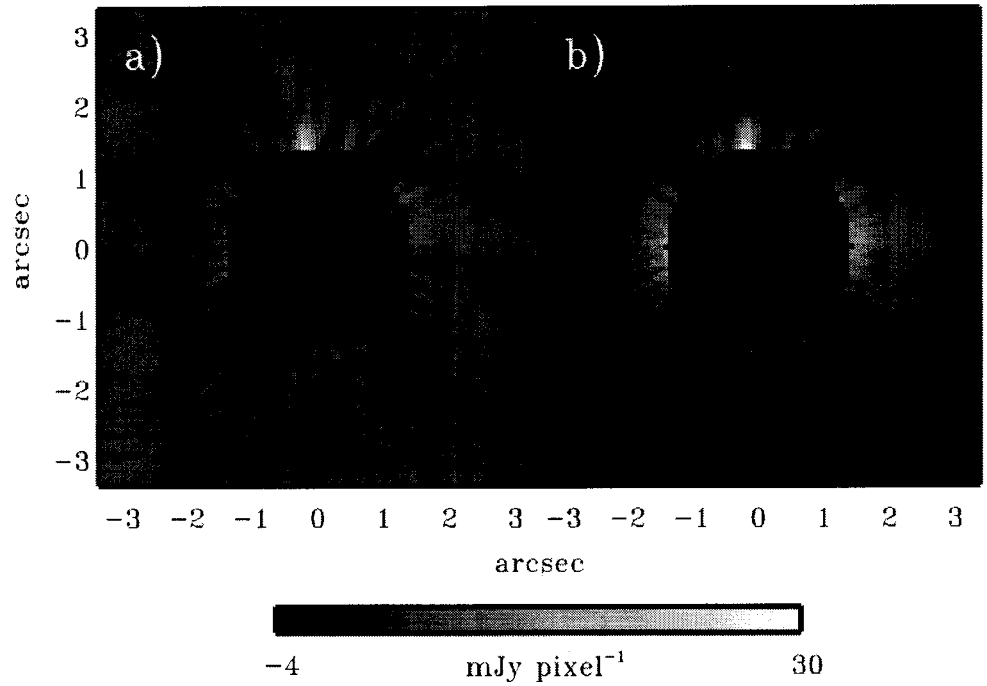

Figure 4.5: a) An image of Procyon using Altair as a PSF calibrator. b) The same image plus a model of the scattered light for an exozodiacal cloud 7×10^5 times as bright as our own zodiacal cloud.

the right of the circular masked regions. The symmetry planes of the model disks are inclined 30° from edge-on. The dust densities in these have been enhanced to $> 10^5\times$ solar levels so they are marginally discernible from the residuals from the PSF subtraction. We used these models for the sake of comparison with the solar zodiacal cloud; real disks with this much dust would be severely collisionally depleted, unlike the solar cloud, and would be unlikely to have the same radial structure as the solar cloud. Despite the high dynamic range of the NICMOS coronagraph and our efforts at PSF calibration, we were not able to improve upon photometric detection limits for exozodiacal dust around these stars; if the stars actually had this much circumstellar dust, the thermal emission from the dust would have been seen as a photometric excess by IRAS.

Our study demonstrates the difficulty of detecting exozodiacal dust in the presence of scattered light from a bright star in a single-dish telescope. Faint companions can be differentiated from the wings of the telescope PSF by techniques like roll subtraction, but if exozodiacal clouds resemble the solar zodiacal cloud, light from these clouds will resemble the PSF wings. Even though coronagraphs can suppress the PSF wings from an on-axis source by as much as an order of magnitude, the dynamic range obtainable with a coronagraph on a large, diffraction-limited telescope in the near-infrared is far from that required to probe dust levels comparable to the solar cloud.

4.2.4 Faint Companions

For our faint companion search we created roll-subtracted images of Sirius and Procyon using the IDP3 software. To find the detection limits for faint companions among the non-Gaussian PSF residuals, we tested our abilities to see artificial stars added to our images. We examined roughly 350 copies of the PSF-subtracted images of each of Sirius, Procyon and Altair with help from a few of our patient colleagues. To five-sixths of the images, we added images of artificial stars, copied from our image of Sirius B, at random positions and magnitudes that were unknown to the examiner. The other images were left unaltered and mixed with the images that contained artificial stars. The examiners were shown each image one at at time, and asked whether they could say confidently that the image they were shown had an artificial star. Only 2% of the time did an examiner claim to see an artificial star when none had been added to the image. We quote as our detection limit the threshold for finding 90% of the artificial companions; that is, the examiners reported 90% of the artificial companions brighter than our detection limit at a given separation. Figure 4.4 shows these detection limits. For comparison, we plot the expected magnitudes of two kinds of possible companions to these objects: an L0 dwarf like 2MASP J0345432+254023 (Kirkpatrick et al. 1999) and a cool brown dwarf, Gl229B (Matthews et al. 1996).

Our apparent detection limits for Procyon are somewhat better than our detection limits for Sirius because Procyon is almost a magnitude fainter in the near infrared; the two sets of observations yielded about the same dynamic range. Although Altair is fainter than Procyon, our absolute detection limits for faint companions to Altair are not much better than our detection limits for companions around Procyon because we have only exposures at only one roll angle for Altair. If we compare our upper limits to the J magnitude of T Dwarf Gl229B (Matthews et al. 1996), we find that we can rule out dwarfs hotter than this object—including all L dwarfs—farther than 2.3 arcseconds from Procyon and $\sim$ 3.0 arcseconds from Sirius and Altair. For comparison, note that Gl 229B was discovered 7.7 arcseconds from a M1V star with an intrinsic luminosity 5 magnitudes fainter than Sirius in the J band (Nakajima et al. 1995). Our limits are weaker closer to the stars. We can rule out the existence of M dwarf companions farther than 1.4 arcseconds from Altair, 1.6 arcseconds from Procyon and 1.8 arcseconds from Sirius at greater than the 90% confidence level. In general, it should be noted that the coronagraph hole is only 3.5 arcseconds from the edge of the chip, and that artificial faint companions that were behind one of the four main diffraction spikes at one roll angle were harder to detect than artificial companions at other position angles; Figure 4.4 is averaged over position angle.

We do not see any evidence for previously undetected faint companions in our images. If there were a low-mass companion orbiting Sirius at 4.2 AU we could not detect it because it would lie in the saturated parts of our images. However, we did survey a large fraction of the space where a companion orbiting Sirius B might be found, and we could have detected a brown dwarf like Gl 229B throughout most of this zone. If there is a third object in the Sirius system, and it orbits Sirius B with a 6 year orbit, it is probably fainter than a brown dwarf.

We thank Glen Schneider and Aaron Evans for help with the data reduction, and Edo Berger, John Carpenter, Micol Christopher, Ulyana Dyudina, David Frayer, Roy

Gal, Pensri Ho, Matthew Hunt, Shardha Jogee, Olga Kuchner, Charlie Qi, Michael Santos, Alice Shapley and David Vakil for searching for artificial stars in our data.

This research made use of the Simbad database, operated at the Centre de Donnees de Strasbourg (CDS), Strasbourg, France.

Support for this work was provided by NASA through grant number GO-07441.01-96A from the Space Telescope Science Institute, which is operated by AURA, Inc., under NASA contract NAS5-26555.

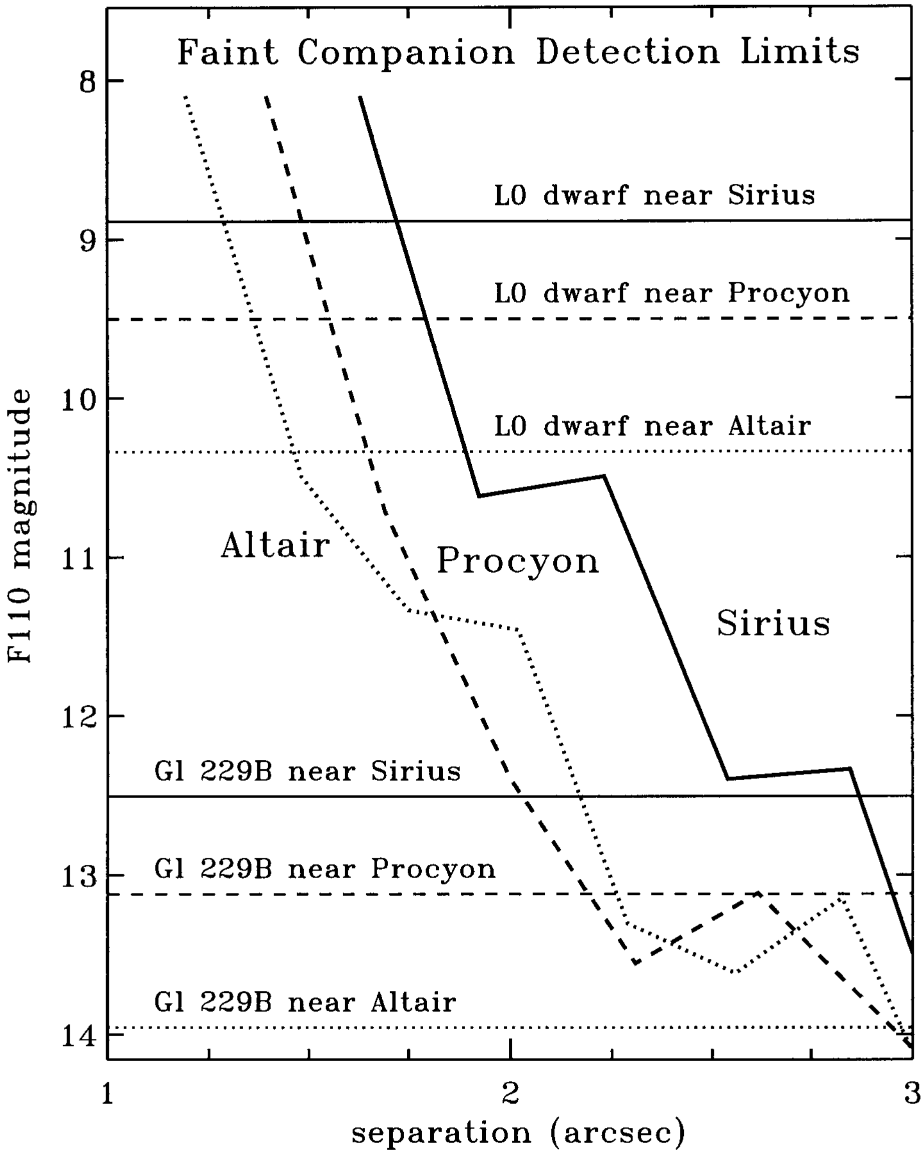

Figure 4.6: Detection limits for faint companions around our three target stars as a function of separation from the stars. The magnitudes of some representative cool objects, GL 229B and an L0 dwarf, are shown for comparison.

Chapter 5

Modeling Exozodiacal Dust Detection with the Keck Interferometer

Marc J. Kuchner

Palomar Observatory, California Institute of Technology, Pasadena, CA 91125

Eugene Serabyn

Jet Propulsion Laboratory, Pasadena, CA 91109

The planned nulling capability of the Keck Interferometer should allow it to probe the region within 200 milliarcseconds of a bright star while suppressing on-axis starlight by a factor of $\sim 10^{-3}$ for a G star at 10 parsecs. This high-resolution, high-dynamic range mid-infrared experiment will search nearby stars for faint emission from exozodiacal dust, a tracer of extrasolar comets and asteroids, taking a step toward direct observations of extrasolar terrestrial planets. We derive general relations between the visibility of an astronomical source and quantities the Keck Nuller is designed to measure. Then we use these relations to simulate Keck Nuller data sets and compute detection limits for exozodiacal clouds analogous to the solar system cloud. We explore how the distance, spectral type, and declination of the target star affect the sensitivity of the experiment, and consider a variety of cloud morphologies. We show that stellar leak can easily be distinguished from the signal from a dust cloud by its spectral signature, so that a $10\times$ solar cloud around a G star at 10 parsecs is detectable even when the stellar leak is three times the signal from the cloud. We also find that the interferometer can constrain the cloud morphology given a long track at ~ 10 times the detection limit.

5.1 Introduction

Many important astrophysical phenomena may remain undetected not because they are too faint for large modern telescopes but because they are close to a bright source, like a nearby star. In a single-aperture telescope—even one with adaptive optics—diffracted light and a halo of scattered light from small imperfections in the primary mirror surround the image of a bright star and easily overpower circumstellar sources. An interferometer offers much higher resolution than a single telescope, but the stellar fringes in a conventional interferometer likewise dwarf the fringes from faint off-axis sources.

The technique of nulling interferometry can potentially avoid these problems by offering a way to selectively reject the on-axis stellar signal. A nulling interferometer cancels on-axis starlight by introducing a relative phase shift of π radians between the light signals from two separated telescopes, so when the signals are combined at zero optical path difference, the electric fields subtract. Off-axis light acquires a different phase shift, so it "leaks" through. In a conventional interferometer, the finite bandpass results in a finite transmitted signal at the fringe minimum, even for an unresolved source. However, a nulling interferometer can be designed to be achromatic over a broad band, so that the off-axis sensitivity is high, while the total on-axis transmission is identically zero over the entire band.

Baudoz et al. (1998) and Hinz et al. (1998) have both achieved stellar rejection ratios of $\sim$ 20:1 using small-aperture interferometers. However, the vast potential of nulling interferometry remains unrealized. In the near term, nulling beam combiners are planned for a number of facilities, including the Keck Interferometer (Colavita 1998; Colavita & Wizinowich 2000) and the Large Binocular Telescope (Angel & Woolf 1997). These interferometers are designed to produce rejection rations of 1000:1 or better, and to harness the collecting area of pairs of 8–10-meter class telescopes. In the long term, nulling interferometry may be practiced in space; nullers may serve the Space Interferometry Mission (Danner & Unwin 1999) and the Terrestrial Planet Finder, a proposed space mission designed to investigate inhabitable planets around other stars (Beichman et al. 1999). The Keck Interferometer, currently under development at the Jet Propulsion Laboratory, is scheduled to be the first nulling interferometer to link 10-meter class telescopes. It will combine the beams from the two Keck telescopes, and offer nulling capability in the 10-micron atmospheric window in late 2001.

An important motivation for nulling interferometry is to search nearby stars for circumstellar exozodiacal dust—the extrasolar analog of the asteroidal and cometary dust that orbits the sun. The solar zodiacal cloud, a sparse disk of 10–100 micron diameter silicate grains, is the most luminous component of the solar system after the sun. Its optical depth is only $\sim 10^{-7}$, but a patch of the solar zodiacal cloud only 0.3 AU across has roughly the same emitting area as an Earth-sized planet. Our zodiacal cloud is comparable in surface area to a single 50 km diameter asteroid ground to dust, so it is easy to imagine that similar and even brighter clouds may be common in other planetary systems and that exozodiacal dust may often be the most luminous feature of the habitable zones of nearby stars. In this case, exozodiacal dust could present a severe obstacle for the direct detection of extra-solar terrestrial planets (Beichman 1996; Beichman et al. 1999). On the other hand, exozodiacal dust traces small bodies—extrasolar asteroids and comets—which may provide vital clues to the configuration, formation and even the migration history of a planetary system.

Exozodiacal clouds may be easier to detect than extra-solar terrestrial planets, but finding an exozodiacal cloud is still not easy. The total emission from our zodiacal cloud is no more than $\sim 10^{-4}$ of the Sun's at any wavelength (see Section 5.2), so photometric surveys like the IRAS (Infrared Astronomical Satellite) survey can only detect exozodiacal clouds that are > 500 times as bright as the solar cloud (see Backman & Paresce 1993). The planned Space Infrared Telescope Facility (SIRTF) should have much more sensitivity than IRAS and it will perform photometry of nearby G stars in the mid-infrared with accuracies of a few percent. However, this experiment will also be limited to detecting exozodiacal clouds that are hundreds or thousands of times as bright as the solar cloud, because of the uncertainty in the flux from the stellar photosphere. Attempts to spatially resolve faint exozodiacal clouds with single-dish telescopes in the mid-infrared (Kuchner, Koresko & Brown 1998) and near-infrared (Kuchner & Brown 2000) have not yielded significantly better detection limits. A few Vega-like stars appear to host distinct populations of dust orbiting at circumstellar distances of < 10 AU, but these massive clouds have > 1000 times the optical depth of the solar zodiacal cloud (e.g., Koerner et al. 1998; Greaves et al. 1998, Fajardo-Acosta, Beichman & Cutri 2000).

Serabyn et al. (2000) estimate that with a stellar rejection ratio of $\sim 1000 : 1$, and a broadband sensitivity of ~ 0.2 mJy at $\lambda = 10$–13 microns in an hour of integration, the Keck Nuller (KN) should be capable of detecting exozodiacal clouds as faint as the solar zodiacal cloud. However, Serabyn et al. (2000) did not include all the relevant noise sources and did not address the interpretation of the interferometric data—separating the circumstellar signal from the residual starlight, and inferring the morphology of the circumstellar cloud. We approach this aspect of the problem of exozodiacal dust detection with a nulling interferometer by simulating KN data sets—both signal and noise—as realistically as presently possible, and fitting models of the interferometer response to the simulated data.

We begin this paper with a description of our exozodiacal cloud modelling package, "ZODIPIC," in Section 5.2. We review standard and nulling interferometry in Section 5.3 and examine some relevant brightness distributions in the (u, v) plane in Section 5.4. The Keck Nuller differs from an ordinary nulling interferometer because it uses an interferometric chopping scheme. We describe this feature of the Keck Nuller and derive relations between the visibility of an astronomical source and quantities the Keck Nuller is designed to meaure in Section 5.5. In Sections 5.6 and 5.7 we discuss two important noise sources for the Keck Nuller, photon noise and null depth fluctuations. Then, in Section 5.8, we calculate the response of the Keck Nuller to hypothetical exozodiacal clouds and simulate the recovery of information from Keck Nuller data to compute detection limits for exozodiacal clouds. With this simulation, we explore how the distance, spectral type, and declination of the target star affect the sensitivity of the experiment. We also consider a range of cloud morphologies. Section 5.9 concludes.

5.2 Exozodiacal Dust: ZODIPIC

We synthesized a variety of images of model exozodiacal clouds using "ZODIPIC," an IDL package written by one of the authors (Kuchner), available online at http://cfa-www.harvard.edu/~mkuchner/. ZODIPIC evaluates the empirical model of the solar zodiacal cloud that Kelsall et al. (1998) fit to detailed maps of the infrared sky at 10 wavelengths from 1 micron to 200 microns from the Diffuse Infrared Background Experiment (DIRBE) aboard the Cosmic Background Explorer (COBE) satellite. This model has a face-on optical depth of $\tau = 7.11 \times 10^{-8}(r/1\mathrm{AU})^{-0.34}$. Although DIRBE never mapped solar elongations $< 60^\circ$, observations made by the Helios Satellite as far in as 0.3 AU (Lienert et al. 1981) and maps assembled from data from the Clementine mission (Hahn et al. 2001) extending to solar elongations as small as 3° yield consistent power-laws. ZODIPIC is designed to be a general purpose dust cloud modelling tool; it can compute images at scattering as well as thermal wavelengths, and offers a variety of tweakable parameters accessible from the IDL comand line.

A challenge for those wishing to compute accurate images of dust disks is the large range in radius often involved; most of the flux from a cloud tends to come from the innermost regions, which may have a size scale orders of magnitude smaller than the cloud outer radius. For example, we truncate our model at an outer radius of 3.28 AU and at an inner radius set by a temperature of 1500 K, where we assume the dust sublimates. Since the dust temperature, T, in the DIRBE model follows the radial power law $T = 286K(r/1\mathrm{AU})^{-0.467}$, this inner radius corresponds to 0.03 AU. ZODIPIC handles this factor of ~ 100 in radius by iterating to distribute computing power evenly in logarithmic radial intervals. In one iteration, ZODIPIC calculates the volumetric emissivity everywhere in a region defined by a cube, with a central cubic region excluded, and integrates the emissivity along lines of sight to produce an image. In the next iteration, ZODIPIC confines itself to the cubic region that was excluded from the first iteration, and repeats the process using the same number of grid points. After two or three iterations, the spatial resolution of the grid in the innermost cube generally suffices to accurately calculate the flux from the inner part of the cloud, and the program re-bins and totals the images from all the iterations to yield a final output image.

Figure 5.1 shows three cuts through model images computed by ZODIPIC. The models represent the solar zodiacal cloud seen at 10 microns, face-on and edge-on, not including the stellar flux. The units (mega-Janskys per steradian) are distance-independent, but the angular scale on the top axis is appropriate for a disk 10 parsecs distant. Vertical grey stripes in the background indicate the size scale of the nulling fringes of the Keck Interferometer at 10 microns, assuming this distance. The Earth temporarily traps inspiralling particles in a series of first order mean motion resonances, forming an extra concentration of particles in a circumsolar ring (Jackson and Zook 1989; Dermott et al. 1994; Reach et al. 1995), which appears as a small bump at 1 AU in the figure. Dynamical perturbations from a Neptune mass planet can dramatically reshape a circumstellar cloud (e.g., Roques et al. 1994, Liou and Zook 1999). However, as this figure shows, the dynamical perturbations from an Earth mass planet are not likely to be consequential to the Keck Nuller experiment. What is more important is the gross structure of the exozodiacal cloud near or even interior to the first fringe maximum, where most of the detectable signal arises—in this case the critical zone is at roughly 0.1 AU.

Figure 5.2 compares the DIRBE model with an earlier and less sophisticated model of the zodiacal cloud fit to the all-sky maps in the IRAS Sky Survey Atlas (ISSA). It is described in appendix G of the ISSA Explanatory Supplement (Wheelock et al. 1994) by J. Good. The figure shows the encircled flux as a function of radius for face-on images of the two models viewed at a distance of 10 parsecs at 5, 10 and 20 microns.

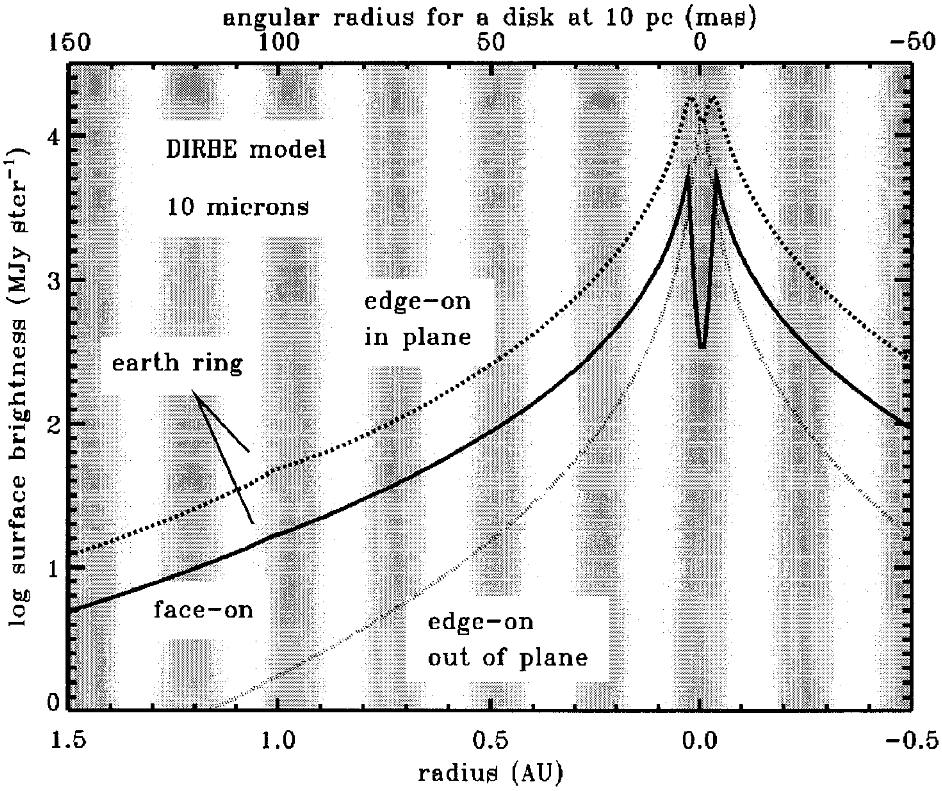

Figure 5.1: The solar zodiacal cloud seen at 10 microns at a distance of 10 parsecs, based on the DIRBE model as realized by ZODIPIC. The vertical stripes in the background represent nuling fringes for the 85 meter Keck Interferometer baseline at 10 microns. What is most important to the Keck Nuller experiment is the structure of the exozodiacal cloud at 0.1 AU, near the first fringe maximum.

At 10 microns, the total flux of the DIRBE model is 0.11 Jy, while the total flux in the IRAS model is 0.20 Jy. In our simulation of exozodiacal cloud detection, the fainter DIRBE model yields the more conservative signal-to-noise estimates.

Of course, a real exozodiacal cloud may not resemble the solar zodiacal cloud. For instance, an exozodiacal cloud may flow from a cloud of small bodies with a very different distribution than the small bodies in our solar system. Another possible complication is that exozodiacal clouds that are > 100 times denser than our zodiacal cloud should be collisionally depleted in their centers.

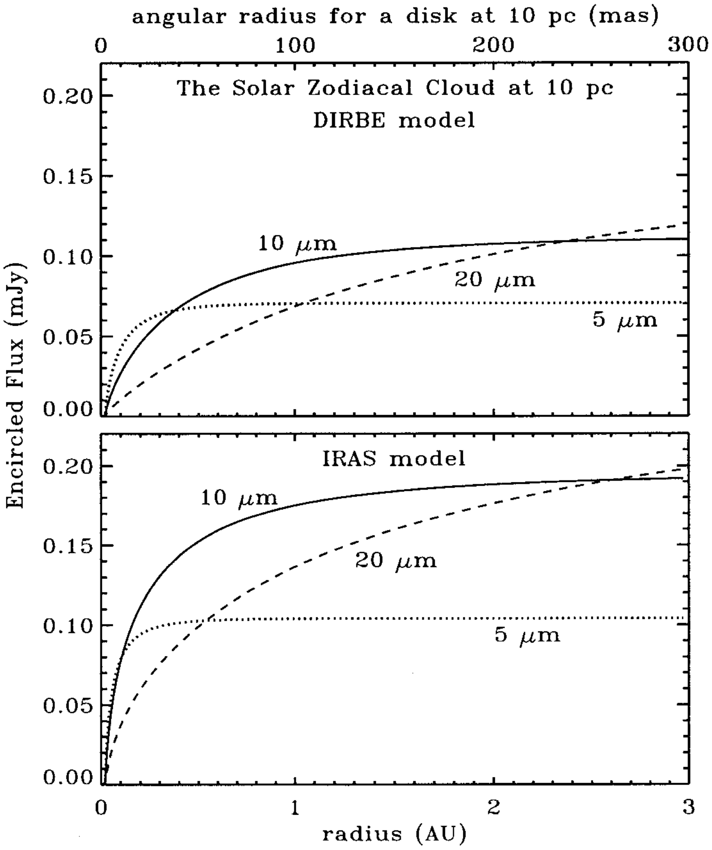

Figure 5.2: The encircled flux as function of radius for two popular models for the Solar System Zodiacal Cloud viewed face-on from a distance of 10 parsecs. The DIRBE model is more accurate, but some discussions refer to the older IRAS model, which is brighter by a factor of ~ 2 in the 10-micron window.

A saving grace is that any cloud that is not collisional approaches a face-on optical depth $\tau =$ constant law far interior to the source of the particles. This law is derived by combining the continuity equation for the dust number density with the equations of motion for a particle under P-R drag (e.g., Wyatt and Whipple 1950). The Keck Nuller will be most sensitive to the interior 0.3 AU of a dust cloud at 10 pc. So the DIRBE model of the solar cloud, which is not very different from a $\tau =$ constant disk, makes a good reference point for our simulations.

Another potentially important difference between exozodiacal clouds and the solar system cloud is the spectrum. Silicate dust has an Si-O feature at 8–12 microns which is strong in comet tails, which are dominated by small rystalline particles. This feature is no more than a 10% effect in the zodiacal cloud (Briotta 1976; Reach et al. 1996) however, in the spectrum of the β Pictoris dust cloud, fully half of the 10 micron emission is in the silicate feature (Telesco and Knacke 1991; Knacke et al. 1993). We neglect this feature entirely.

5.3 Single Baseline Nulling Interferometry

We begin our discussion of interferometry with a brief description of "standard" interferometry at radio wavelengths. Using the notation of Thompson, Moran and Swenson (1998), the fringe pattern of a standard radio interferometer for small $\boldsymbol{\sigma}$ is

$$F_{radio}(\boldsymbol{\sigma}, \mathbf{s_0}) = \cos\left(2\pi \mathbf{D}_\lambda \cdot (\mathbf{s_0} + \boldsymbol{\sigma})\right), \tag{5.1}$$

where $\mathbf{D}_\lambda = (u, v)$ is the projected baseline vector measured in wavelengths of light, with components u and v defined in the standard way, and $\boldsymbol{\sigma}$ is the vector angular distance from the phase center on the celestial sphere. The vector $\mathbf{s_0}$ represents an instrumental phase delay that allows the fringe to be scanned across the sky. The monochromatic interferometer response to a source brightness distribution $B(\boldsymbol{\sigma})$ is

$$r_{radio}(u, v) = \int d\Omega\ B(\boldsymbol{\sigma})A(\boldsymbol{\sigma})F_{radio}(\boldsymbol{\sigma}, \mathbf{s_0}) = \int d\Omega\ B(\boldsymbol{\sigma})A(\boldsymbol{\sigma}) \cos\left(2\pi \mathbf{D}_\lambda \cdot (\mathbf{s_0} + \boldsymbol{\sigma})\right), \tag{5.2}$$

where $A(\boldsymbol{\sigma})$ is the telescope beam pattern, $\int d\Omega$ represents integration over the celestial sphere. If we define the complex visibility in the traditional way as the Fourier transform of the product of the source brightness distribution and the telescope beam pattern,

$$\mathcal{V}(\mathbf{D}_\lambda) = \int d\Omega\ B(\boldsymbol{\sigma})A(\boldsymbol{\sigma})e^{-2\pi i\mathbf{D}_\lambda\cdot\boldsymbol{\sigma}}, \tag{5.3}$$

then for a standard radio interferometer, we can write the response in terms of two spatial frequency components:

$$r_{radio}(u,v) = \frac{1}{2}\mathcal{V}(\mathbf{D}_\lambda)e^{2\pi i\mathbf{D}_\lambda\cdot\mathbf{s_0}} + \frac{1}{2}\mathcal{V}(-\mathbf{D}_\lambda)e^{-2\pi i\mathbf{D}_\lambda\cdot\mathbf{s_0}}. \tag{5.4}$$

In general, the visibility, $\mathcal{V}(\mathbf{D}_\lambda) = \mathcal{V}(u,v)$, of a real source is complex and Hermitian. If the source is symmetrical about the origin, or "even," the visibility is real and symmetrical about the origin. In this case, equation 5.4 reduces to

$$r_{radio,even}(u,v) = \mathcal{V}(\mathbf{D}_\lambda)\cos(2\pi\mathbf{D}_\lambda\cdot\mathbf{s_0}). \tag{5.5}$$

If source brightness distribution is anti-symmetric about the origin, or "odd", the visibility is imaginary and odd, and equation 5.4 reduces to

$$r_{radio,odd}(u,v) = i\mathcal{V}(\mathbf{D}_\lambda)\sin(2\pi\mathbf{D}_\lambda\cdot\mathbf{s_0}). \tag{5.6}$$

Naturally, a purely odd source brightness distribution is unphysical, but source brightness distributions generally have even and odd components. These observations allow us to combine equation 5.6 and equation 5.6 into

$$r_{radio}(u,v) = \Re\Big\{\mathcal{V}(\mathbf{D}_\lambda)e^{2\pi\mathbf{D}_\lambda\cdot\mathbf{s_0}}\Big\}. \tag{5.7}$$

Ordinarily, an interferometer operating in the optical/infrared regime uses square-law detectors, so it produces a positive-definite fringe pattern

$$F_{optical/IR}(\boldsymbol{\sigma},\mathbf{s_0}) = \cos^2(\pi\mathbf{D}'_\lambda\cdot(\mathbf{s_0}+\boldsymbol{\sigma})) = \frac{1}{2}\Big(1+\cos\left(2\pi\mathbf{D}_\lambda\cdot(\mathbf{s_0}+\boldsymbol{\sigma})\right)\Big). \tag{5.8}$$

A single-baseline optical/infrared nulling interferometer operating in the same regime blocks most of the light from an on-axis source by centering it in a fixed central destructive fringe. The chief advantages of nulling interferometry are realized by constructing a nulling interferometer so that it operates achromatically over a broad band. However, viewed monochromatically, the optical/infrared nulling interferometer is equivalent to a conventional optical/infrared interferometer with a fixed phase. The positive-definite fringe pattern from a monochromatic single-baseline optical/infrared nulling interferometer is

$$F_{nulling}(\boldsymbol{\sigma}) = \sin^2(\pi \mathbf{D}_\lambda \cdot \boldsymbol{\sigma}) = \frac{1}{2}\Big(1 - \cos\left(2\pi \mathbf{D}_\lambda \cdot \boldsymbol{\sigma}\right)\Big). \tag{5.9}$$

If the combined beams are detected in a way that integrates over the fringe, as in a conventional radio interferometer, the monochromatic nulling interferometer response is

$$r_{nulling}(u, v) = \int d\Omega \; B(\boldsymbol{\sigma}) A(\boldsymbol{\sigma}) F_{nulling}(\boldsymbol{\sigma}). \tag{5.10}$$

If we write the fringe pattern in terms of complex exponentials,

$$F_{nulling}(\boldsymbol{\sigma}) = \frac{1}{2} - \frac{1}{4}\left(e^{2\pi i \mathbf{D}_\lambda \cdot \boldsymbol{\sigma}} + e^{-2\pi i \mathbf{D}_\lambda \cdot \boldsymbol{\sigma}}\right), \tag{5.11}$$

we see that the response of the nulling interferometer can be expressed in terms of three spatial frequency components:

$$r_{nulling}(u, v) = -\frac{1}{4}\mathcal{V}(\mathbf{D}_\lambda) + \frac{1}{2}\mathcal{V}(\mathbf{0}) - \frac{1}{4}\mathcal{V}(-\mathbf{D}_\lambda). \tag{5.12}$$

The term at $\mathbf{D}_\lambda = 0$ is a common feature of optical interferometers, which use quadratic detectors. It guarantees that the noise-free fringe power is never negative. The $\mathbf{s_0}$ phase terms are missing here because it is assumed that the phase is fixed; the null fringe is always centered on the central source.

If we divide the source brightness distribution into odd and even parts, we find:

$$r_{nulling,even}(u,v) = \frac{1}{2}\big(\mathcal{V}(\mathbf{0}) - \mathcal{V}(\mathbf{D}_\lambda)\big), \tag{5.13}$$

$$r_{nulling,odd}(u,v) = 0. \tag{5.14}$$

Nulling interferometers can not detect the odd component of a brightness distribution. With this observation, we can rewrite equation 5.12 as

$$r_{nulling}(u,v) = \frac{1}{2}\Big(\mathcal{V}(\mathbf{0}) - \Re\{\mathcal{V}(\mathbf{D}_\lambda)\}\Big). \tag{5.15}$$

Another important quantity is the fraction of the source's total flux that is transmitted through the nulled fringe. The fringe *transmission* is

$$T(u,v) = \frac{\int d\Omega\ B(\boldsymbol{\sigma})A(\boldsymbol{\sigma})F_{nulling}(\boldsymbol{\sigma})}{\int d\Omega\ B(\boldsymbol{\sigma})} = \frac{r_{nulling}(u,v)}{\text{source flux}} \tag{5.16}$$

The visibility can be used to calculate the transmission, using equation 5.12. Since the overall sensitivity of a real nulling interferometer will probably not be constant and need to be calibrated by regularly measuring the total source flux in the absence of the nulling fringe, the transmission is the astrophysically relevant parameter that a single-baseline monochromatic nulling interferometer can observe.

5.4 Starlight and Exozodiacal Clouds in the (u,v) Plane

We can visualize the operation of a single baseline nulling interferometer by examining some representative sources in the (u,v) plane. We could plot the complex visibility of the sources and leave it to the reader to sum some appropriate spatial frequency components according to the prescriptions in Section 5.3. Instead we prefer to plot the corresponding transmissions, since a monochromatic nulling interferometer can measure that quantity directly. To first approximation, the Keck Nuller behaves like a single-baseline nulling interferometer, so this approach can also provide an intuitive feeling for what the Keck Nuller can measure.

Consider a star that is a small uniform disk with angular diameter $\theta_\star$, and brightness $B_\star(\boldsymbol{\sigma}) = B_\star$ for $\boldsymbol{\sigma}_\mathbf{r} < \theta_\star/2$ and $B_\star(\boldsymbol{\sigma}) = 0$ elswhere. This star has a total flux $F_\star = B_\star \pi (\theta_\star/2)^2$. Performing in the integral in equation 5.3 and using the fact that the single-aperture power pattern, $A(\boldsymbol{\sigma})$, is nearly unity at the phase center, we find the standard expression for the visibility of a uniform disk,

$$\mathcal{V}_\star(\mathbf{D}_\lambda) = B_\star \frac{\theta_\star J_1(\pi\theta_\star|\mathbf{D}_\lambda|)}{2|\mathbf{D}_\lambda|} = F_\star \frac{2J_1(\pi\theta_\star|\mathbf{D}_\lambda|)}{\pi\theta_\star|\mathbf{D}_\lambda|}, \tag{5.17}$$

where J_1 is the Bessel function of first order. Expanding the visibility to third order in the small angle, we get

$$\mathcal{V}_\star(\mathbf{D}_\lambda) \approx F_\star\Big(1 - \frac{1}{8}\big(\pi\theta_\star\mathbf{D}_\lambda\big)^2\Big). \tag{5.18}$$

Using equation 5.13, we find that in this limit, the response of a single-baseline optical/infrared nulling interferometer to the star is

$$r_{nulling,\star}(u,v) \approx \frac{1}{16} F_\star \Big(\pi\theta_\star\mathbf{D}_\lambda\Big)^2 \tag{5.19}$$

and from equation 5.16, we see that the transmission—the fraction of the starlight transmitted by the fringe—is

$$T_\star(u,v) \approx \frac{1}{16}\Big(\pi\theta_\star\mathbf{D}_\lambda\Big)^2 \tag{5.20}$$

Equation 5.20 shows that when a star is much smaller than the fringe spacing, a single-baseline monochromatic nulling interferometer can only measure one stellar parameter, $\theta_\star^2$, and that for a star of a given type, the amount of starlight a nulling interferometer detects varies inversely as the distance to the star to the fourth power. If the flux from the star follows the Rayleigh-Jeans law at the observed wavelengths, then $r_{\star,nulling} \propto \lambda^{-4}$ in the mid-infrared (see also Serabyn et al. 2000). This steep function of λ can be very different from the wavelength dependence of the signal from a circumstellar disk. This contrast makes it easier to distinguish stellar leakage from other signal components given even a small amount of spectral coverage. For example, $r_{\star,nulling}$ varies by almost a factor of 3 over the 10–13 micron band.

Figure 5.3 shows some sample (u, v) tracks for the long baseline of the Keck Interferometer, plotted for declinations $-10°$, $20°$ and $50°$, assuming a zenith angle limit of $45°$. Figure 5.4 shows a contour plot of the transmission $T_{\star}(u, v)$ for a star with $\theta_{\star} = 0.930$ milliarcseconds, representing the Sun seen at 10 parsecs. Here we have assumed a single-aperture power pattern corresponding to a 4-meter diameter aperture, our approximation for the Keck Nuller power pattern, (see Section 5.5). We have overlaid on this contour plot the (u, v) tracks for the Keck Interferometer's long baseline shown in Figure 5.3. The fraction of the stellar light that leaks through the nulling fringes of the Keck Nuller can be read directly from the contour map at the location of the (u, v) tracks.

Figures 5.5, 5.6, and 5.7 show contour plots of the transmission, $T(u, v)$, as defined in Section 5.2 for the solar zodiacal cloud orbiting a star just like the sun, viewed at a distance of 10 parsecs, the same model shown in Figure 5.1. These figures show how the limited (u, v) coverage of a single baseline interferometer like the Keck Interferometer affects its ability to measure the geometry of a zodiacal cloud. Figure 5.5 shows a face-on cloud and Figures 5.6 and 5.7 show an edge-on cloud at two different position angles. At some orientations like the one shown in Figure 5.6, an edge-on cloud will be difficult to distinguish from a face-on cloud, because in each case, the (u, v) tracks do not cross any contour lines so the transmission depends little on hour angle. However, when the position angle of the cloud is roughly northwest, perpendicular to the baseline, as in Figure 5.7, the transmission will be a stronger function of hour angle, and the Keck Nuller can constrain the dust cloud inclination. At most, the variation in transmission

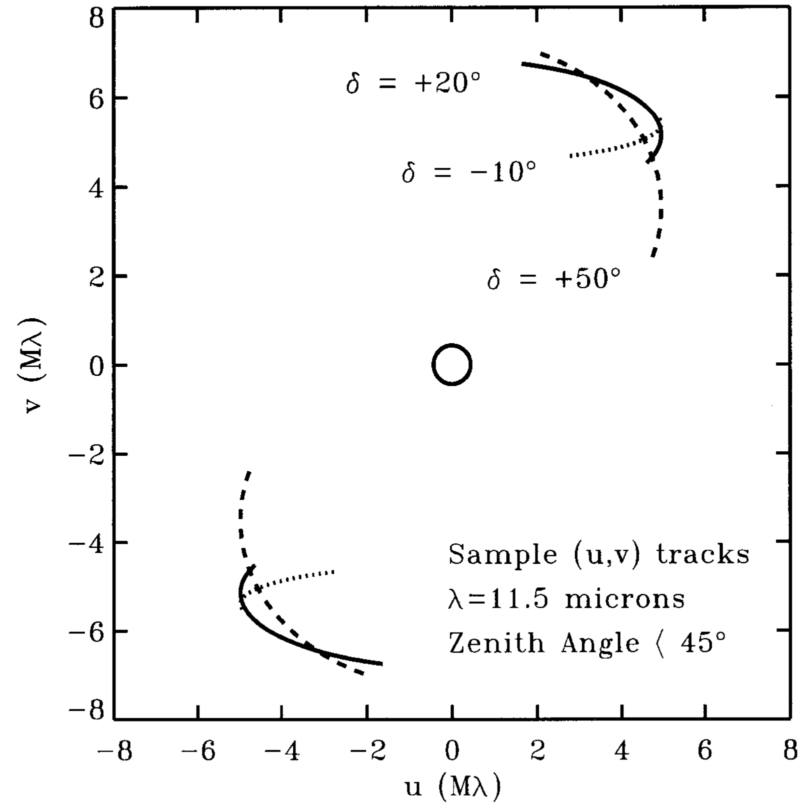

Figure 5.3: Some sample (u, v) tracks for the Keck Interferometer long baseline at 11.5 microns for targets at declinations −10°, 20° and 50°.

as a function of hour angle is $\sim$ 10%.

5.5 The Keck Interferometer in Nulling Mode

The experimental arrangement for the Keck Nuller (KN) will be described thoroughly elsewhere; here we give a brief overview. The Keck Interferometer site, the summit of Mauna Kea, HI, is at 19.83° North latitude, and the baseline between the two 10-meter telescopes is 85 meters in length, positioned 37.7° East of North. Because of the incompatibility of a spatial chopper with the Keck adaptive optics system, the KN must use an interferometric chopping scheme to remove the thermal background. To

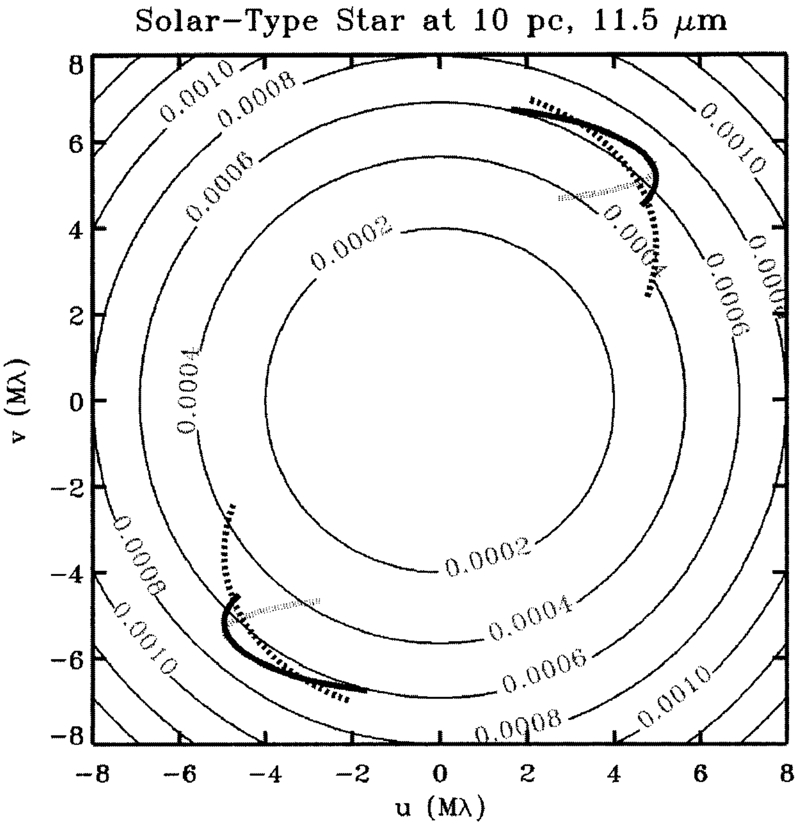

Figure 5.4: A contour map of the transmission for a solar-type star at 10 parsecs in the (u, v) plane. Overlaid are sample (u, v) tracks for the long baseline of the Keck interferometer, shown in Figure 5.3. At $u = v = 0$, the transmitted signal is zero, because there the null covers the entire sky.

accomplish this, the KN will divide each of the Keck telescopes into two subapertures to produce four beams on the sky. The post-telescope beam train will propagate the four subaperture beams independently to the beam combiners in the observatory basement. The result is a set of baselines of two different types: long (∼ 85 meter) baselines between the corresponding subapertures on the two dishes, and short (∼ 5 meter) baselines between the two halves of each dish. The circle at the center of Figure 5.3 has a radius of 5 meters/λ; it indicates the locus of possible short baselines.

The stellar light from the two parallel long baselines will be nulled by a pair of

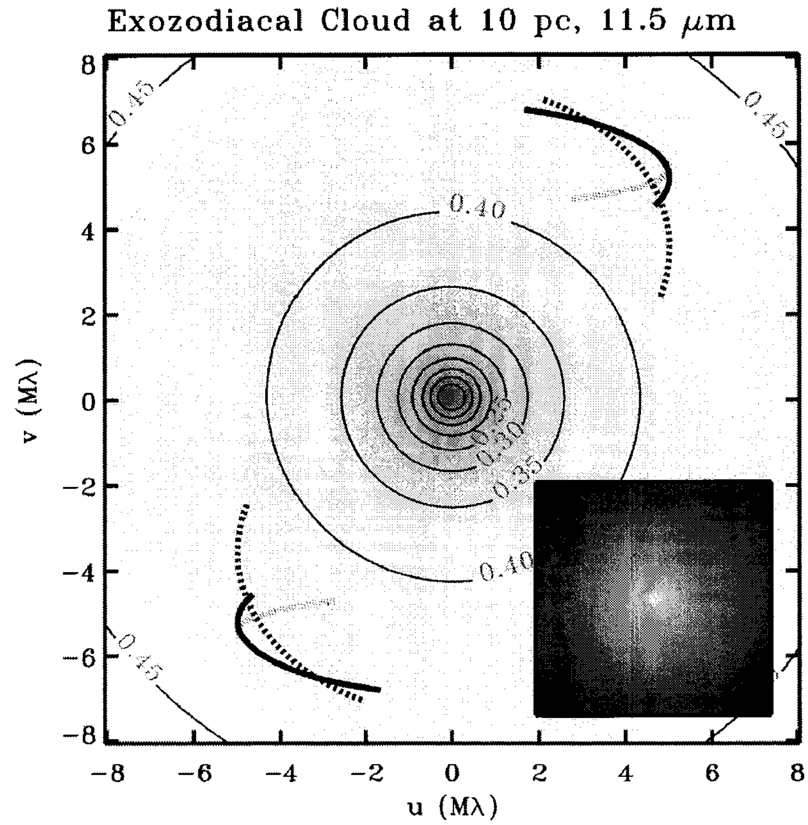

Figure 5.5: A contour map of the transmitted signal from a face-on zodiacal cloud in the (u, v) plane and the sample tracks for the Keck Interferometer from Figure 5.3. The inset shows the logarithm of an image of the cloud. Notice that the contours are elongated in the direction perpendicular to the plane of the disk.

nulling beam combiners, generating the fringe pattern given by equation 5.9 for both telescope pairs. Then the outputs of the two nullers will be fed into a pair of conventional scanning interferometric beam combiners, which will combine the nulled beams, creating a second fringe pattern. This second fringe pattern is standard fringe pattern for a single-baseline optical/infrared interferometer:

$$F_{cross}(\boldsymbol{\sigma}, \mathbf{s_0}) = \cos^2(\pi \mathbf{D}'_\lambda \cdot (\mathbf{s_0} + \boldsymbol{\sigma})) = \frac{1}{2}\Big(1 + \cos\left(2\pi \mathbf{D}'_\lambda \cdot (\mathbf{s_0} + \boldsymbol{\sigma})\right)\Big). \tag{5.21}$$

Here $\mathbf{D}'_\lambda$ represents the short baseline across the aperture, measured in wavelengths

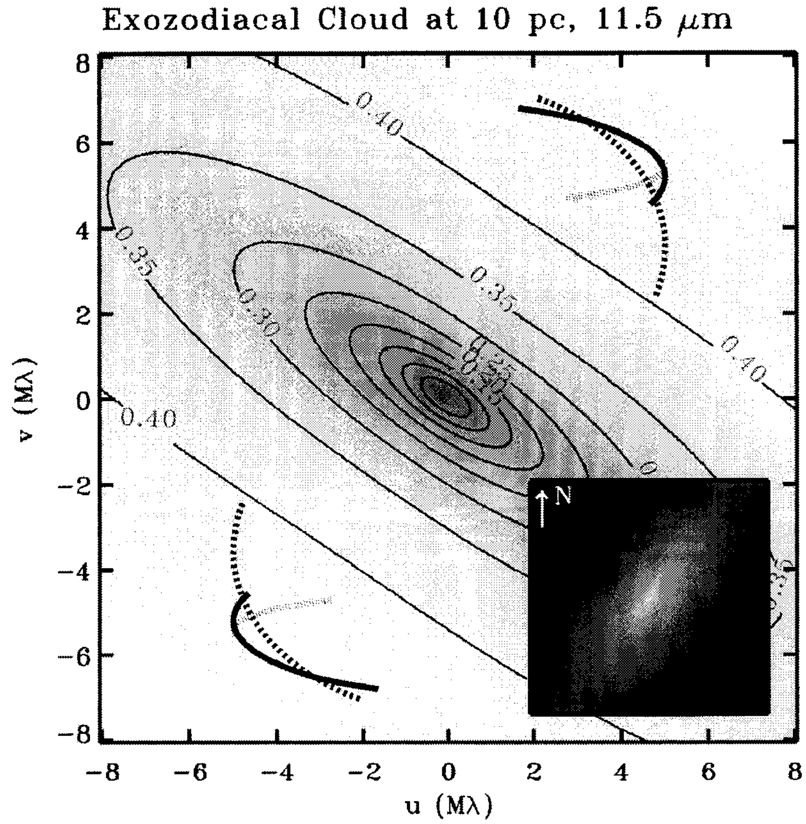

Figure 5.6: The transmitted signal for the same cloud as in Figure 5.5, but edge-on. An edge-on cloud at this position angle will be difficult for the KN to distinguish from a face-on cloud.

of light, and again, the addition of the vector $\mathbf{s}_0$ to $\boldsymbol{\sigma}$ indicates that this secondary fringe will not necessarily stay locked on the star, but can scan across the sky. The scanning secondary fringe system will detect the off-axis signal which has not been nulled by integrating over the narrow fringes. Figure 5.8 illustrates the experimental arrangement schematically.

After the beams are combined, they are spatially filtered at a pinhole, the output of which is dispersed in a mid-infrared camera. The result is that the monochromatic

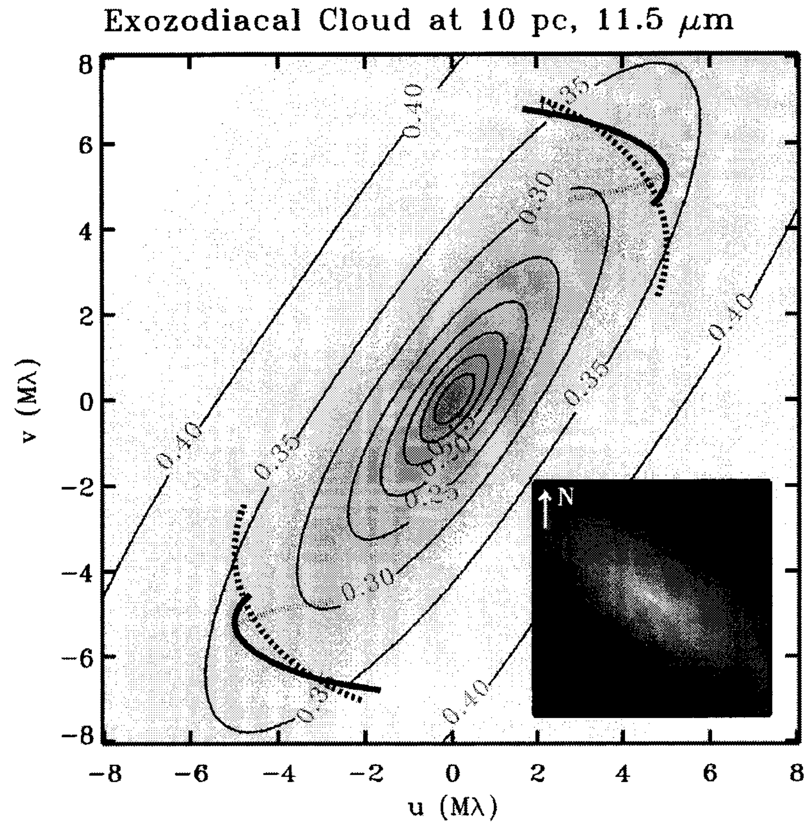

Figure 5.7: The transmitted signal for the same cloud as in Figure 5.6, but at a different position angle. At this position angle the KN will transmit less signal from the cloud. However, at this position angle, the transmitted signal will have a strong dependence on hour angle, so only for clouds with roughly this orientation can the KN constrain the inclination of the disk.

interferometer response to a source brightness distribution $B(\boldsymbol{\sigma})$ is

$$r_{KN}(u,v) = \int d\Omega \; B(\boldsymbol{\sigma})A(\boldsymbol{\sigma})F_{nulling}(\boldsymbol{\sigma})F_{cross}(\mathbf{s}_0,\boldsymbol{\sigma}), \tag{5.22}$$

where $A(\boldsymbol{\sigma})$ is now the subaperture single-mode beam pattern. Explicitly, the response is

$$r_{KN}(u,v) = \int d\Omega\ B(\boldsymbol{\sigma})A(\boldsymbol{\sigma})\frac{1}{2}\Big(1-\cos\left(2\pi\mathbf{D}_\lambda\cdot\boldsymbol{\sigma}\right)\Big)\frac{1}{2}\Big(1+\cos\left(2\pi\mathbf{D}'_\lambda\cdot(\mathbf{s}_0+\boldsymbol{\sigma})\right)\Big). \tag{5.23}$$

The single-mode subaperture diffraction pattern is somewhat complicated, but not important for our simulations, since the angular scale is large compared to the size of our targets. For simplicity we will assume that $A(\boldsymbol{\sigma})$ corresponds to a circular aperture 4 meters in diameter.

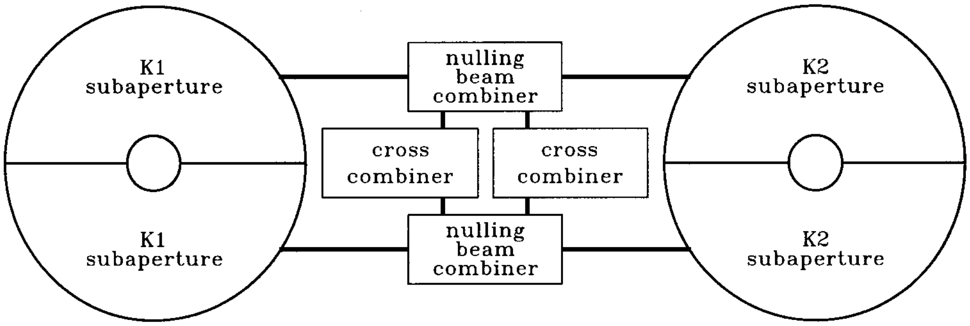

Figure 5.8: A diagram of the Keck Interferometer showing how the beams from corresponding subapertures are first combined with nulling beam combiners, then the residual outputs of the nullers are detected with conventional scanning beam combiners. The relative separation of the two Keck primaries (K1 and K2) is much larger than suggested by this figure.

Figure 5.9 illustrates the integral in equation 5.23. It shows a logarithmic contour plot of the brightness of a representative exozodiacal cloud and the product of the fringe patterns $F_{nulling}(\sigma)$ and $F_{cross}(s_0, \sigma)$ and the single-aperture power pattern, $A(\sigma)$ at four different values of s_0, at 11.5 microns. The thin stripes are due to $F_{nulling}(\sigma)$ and the broad bands are due to $F_{cross}(\sigma, s_0)$.

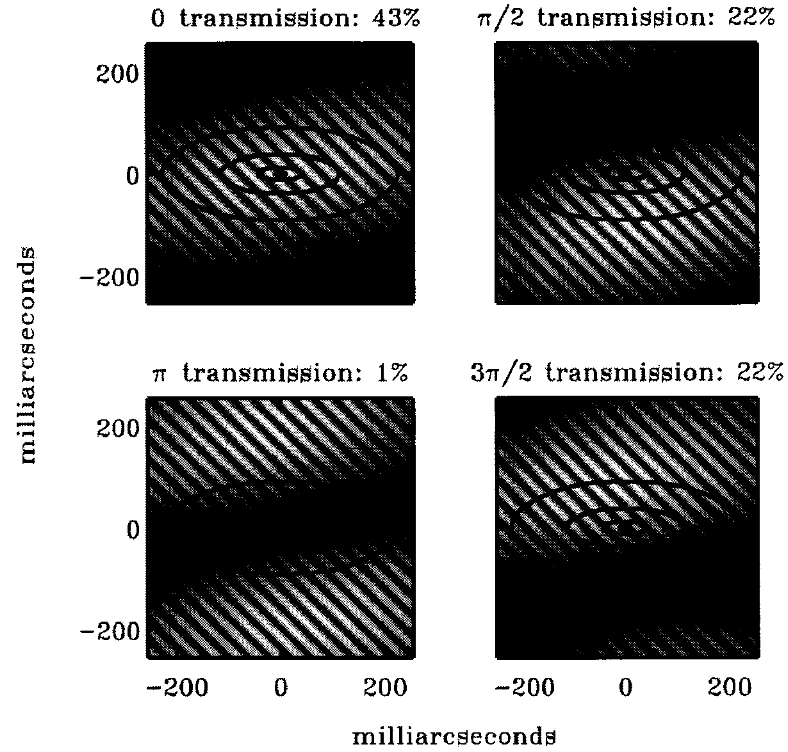

Figure 5.9: An illustration of the fringe patterns for the KN. The contours show the brightness of a hypothetical exozodiacal cloud at 10 pc and 11.5 microns and the shading shows the product of the single-dish beam, the long-baseline fringe, and the scanning short-baseline fringe. The phase of the long-baseline fringe is fixed in nulling mode, but the short baseline fringe scans; the four panels show phases 0,$\pi/2$, π, and $3\pi/2$ of the short baseline fringe.

The Keck Nuller is more complicated than the single-baseline nulling interferometer due to the interferometric chopping, but that does not prevent us from describing this experiment using the notion of complex visibility. This approach allows us to model the response of the interferometer to a given source using a single Fourier transform. Since

$$F_{cross}(\boldsymbol{\sigma}, \mathbf{s_0}) = \cos^2(\pi \mathbf{D}_\lambda \cdot (\mathbf{s_0} + \boldsymbol{\sigma})) = \frac{1}{4} e^{2\pi \mathbf{D}'_\lambda \cdot (\mathbf{s_0} + \boldsymbol{\sigma})} + \frac{1}{2} + \frac{1}{4} e^{-2\pi i \mathbf{D}'_\lambda \cdot (\mathbf{s_0} + \boldsymbol{\sigma})}, \tag{5.24}$$

we have

$$\begin{aligned} &F_{nulling}(\boldsymbol{\sigma}) F_{cross}(\boldsymbol{\sigma}, \mathbf{s_0}) = \\ &\quad \begin{array}{lll} -\frac{1}{16} e^{2\pi i((\mathbf{D}_\lambda + \mathbf{D}'_\lambda)\cdot\boldsymbol{\sigma})} e^{2\pi i \mathbf{D}'_\lambda \cdot \mathbf{s_0}} & -\frac{1}{8} e^{2\pi i \mathbf{D}_\lambda \cdot \boldsymbol{\sigma}} & -\frac{1}{16} e^{2\pi i((\mathbf{D}_\lambda - \mathbf{D}'_\lambda)\cdot\boldsymbol{\sigma})} e^{-2\pi i \mathbf{D}'_\lambda \cdot \mathbf{s_0}} \\ +\frac{1}{8} e^{2\pi i \mathbf{D}'_\lambda \cdot \boldsymbol{\sigma}} e^{2\pi i \mathbf{D}'_\lambda \cdot \mathbf{s_0}} & +\frac{1}{4} & +\frac{1}{8} e^{-2\pi i \mathbf{D}'_\lambda \cdot \boldsymbol{\sigma}} e^{-2\pi i \mathbf{D}'_\lambda \cdot \mathbf{s_0}} \\ -\frac{1}{16} e^{-2\pi i((\mathbf{D}_\lambda - \mathbf{D}'_\lambda)\cdot\boldsymbol{\sigma})} e^{2\pi i \mathbf{D}'_\lambda \cdot \mathbf{s_0}} & -\frac{1}{8} e^{-2\pi i \mathbf{D}_\lambda \cdot \boldsymbol{\sigma}} & -\frac{1}{16} e^{-2\pi i((\mathbf{D}_\lambda + \mathbf{D}'_\lambda)\cdot\boldsymbol{\sigma})} e^{-2\pi i \mathbf{D}'_\lambda \cdot \mathbf{s_0}}. \end{array} \end{aligned} \tag{5.25}$$

Using the definition of visibility given above, we find the net response of the Keck Nuller:

$$\begin{aligned} &r_{KN}(\mathbf{D}_\lambda, \mathbf{D}'_\lambda, \mathbf{s_0}) = \\ &\quad \begin{array}{lll} -\frac{1}{16}\mathcal{V}(\mathbf{D}_\lambda + \mathbf{D}'_\lambda) e^{2\pi i \mathbf{D}'_\lambda \cdot \mathbf{s_0}} & -\frac{1}{8}\mathcal{V}(\mathbf{D}_\lambda) & -\frac{1}{16}\mathcal{V}(\mathbf{D}_\lambda - \mathbf{D}'_\lambda) e^{-2\pi i \mathbf{D}'_\lambda \cdot \mathbf{s_0}} \\ +\frac{1}{8}\mathcal{V}(\mathbf{D}'_\lambda) e^{2\pi i \mathbf{D}'_\lambda \cdot \mathbf{s_0}} & +\frac{1}{4}\mathcal{V}(\mathbf{0}) & +\frac{1}{8}\mathcal{V}(-\mathbf{D}'_\lambda) e^{-2\pi i \mathbf{D}'_\lambda \cdot \mathbf{s_0}} \\ -\frac{1}{16}\mathcal{V}(-\mathbf{D}_\lambda + \mathbf{D}'_\lambda) e^{2\pi i \mathbf{D}'_\lambda \cdot \mathbf{s_0}} & -\frac{1}{8}\mathcal{V}(-\mathbf{D}_\lambda) & -\frac{1}{16}\mathcal{V}(-\mathbf{D}_\lambda - \mathbf{D}'_\lambda) e^{-2\pi i \mathbf{D}'_\lambda \cdot \mathbf{s_0}}. \end{array} \end{aligned} \tag{5.26}$$

To make sense of these equations, we define the *Keck Nuller response phasor*

$$Z(\mathbf{D}_\lambda, \mathbf{D}'_\lambda) = -\frac{1}{4}\mathcal{V}(\mathbf{D}'_\lambda + \mathbf{D}_\lambda) + \frac{1}{2}\mathcal{V}(\mathbf{D}'_\lambda) - \frac{1}{4}\mathcal{V}(\mathbf{D}'_\lambda - \mathbf{D}_\lambda), \tag{5.27}$$

and the *corresponding single-baseline nuller response*

$$r_{nulling}(\mathbf{D}_\lambda) = \frac{1}{2}\Big(\mathcal{V}(\mathbf{0}) - \Re\big\{\mathcal{V}(\mathbf{D}_\lambda)\big\}\Big). \tag{5.28}$$

If the source is symmetrical about the origin, like the simple models we use for stars and exozodiacal clouds, the visibility is real and symmetrical about the origin. In this case, $Z(\mathbf{D}_\lambda, \mathbf{D}'_\lambda)$ is real, and equation 5.26 reduces to

$$r_{KN,even}(\mathbf{D}_\lambda, \mathbf{D}'_\lambda, \mathbf{s_0}) = \frac{1}{2}\Big(r_{nulling}(\mathbf{D}_\lambda) + \cos(2\pi\mathbf{D}'_\lambda \cdot \mathbf{s_0}) Z(\mathbf{D}_\lambda, \mathbf{D}'_\lambda)\Big). \tag{5.29}$$

If source brightness distribution is anti-symmetric about the origin, then $Z(\mathbf{D}_\lambda, \mathbf{D'}_\lambda)$ is imaginary and equation 5.26 reduces to

$$r_{KN,odd}(\mathbf{D}_\lambda, \mathbf{D}'_\lambda, \mathbf{s_0}) = \frac{i}{2}\sin(2\pi\mathbf{D}'_\lambda \cdot \mathbf{s_0}) Z(\mathbf{D}_\lambda, \mathbf{D'}_\lambda). \tag{5.30}$$

With these observations, we can rewrite equation 5.26 as

$$r_{KN}(\mathbf{D}_\lambda, \mathbf{D}'_\lambda, \mathbf{s_0}) = \frac{1}{2}\Big(r_{nulling} + \Re\Big\{Z(\mathbf{D}_\lambda, \mathbf{D}'_\lambda) e^{-2\pi i \mathbf{D}'_\lambda \cdot \mathbf{s_0}}\Big\}\Big), \tag{5.31}$$

or equivalently

$$r_{KN}(\mathbf{D}_\lambda, \mathbf{D}'_\lambda, \mathbf{s_0}) = \frac{1}{2} r_{nulling} + A(\mathbf{D}_\lambda, \mathbf{D}'_\lambda)\cos\Big(2\pi(\mathbf{D}'_\lambda \cdot \mathbf{s_0}) + \phi(\mathbf{D}_\lambda, \mathbf{D}'_\lambda)\Big). \tag{5.32}$$

At a given source declination and hour angle, the Keck Nuller response to a given source is a constant plus a cosine function of the phase of the cross fringe, with an amplitude equal to half the modulus of $Z(\mathbf{D}_\lambda, \mathbf{D}'_\lambda)$

$$A(\mathbf{D}_\lambda, \mathbf{D}'_\lambda) = \frac{1}{2}|Z(\mathbf{D}_\lambda, \mathbf{D}'_\lambda)| \tag{5.33}$$

and a phase equal to the phase of $Z(\mathbf{D}_\lambda, \mathbf{D}'_\lambda)$,

$$\phi(\mathbf{D}_\lambda, \mathbf{D}'_\lambda) = \arctan\left(\frac{\Im\{Z(\mathbf{D}_\lambda, \mathbf{D}'_\lambda)\}}{\Re\{Z(\mathbf{D}_\lambda, \mathbf{D}'_\lambda)\}}\right). \tag{5.34}$$

In reality, the response will also contain a large term due to the thermal background. The known scanning rate of the cross fringe should allow the detection of the cosine response term despite the presence of this background. However, the magnitude and variability of the background will make measuring the absolute value of the response impractical. The amplitude and phase of the cosine term, given by equations 5.33 and 5.34, are the primary observable quantities.

Moreover, it will probably be necessary to calibrate the overall sensitivity of the interferometer by regularly measuring the source flux detected by the interferometer in the absence of the nulling fringe. The result is that the astrophysically relevant measurable quantity will be the fraction of the flux from the source that translates into measurable signal, which we will call the *normalized Keck Nuller response phasor*

$$Z_{norm}(\mathbf{D}_\lambda, \mathbf{D}'_\lambda) = \frac{Z(\mathbf{D}_\lambda, \mathbf{D}'_\lambda)}{\text{source flux}}. \tag{5.35}$$

This quantity is analogous to the transmission of a single-baseline nulling interferometer, equation 5.16. The magnitude of Z_{norm} is the maximum fraction of the source flux transmitted by the KN fringe minus the minimum fraction of the source flux transmitted by the fringe. The phase of Z_{norm} is the same as the phase of Z. In Figure 5.9, since the source has even symmetry, the phase of Z_{norm} is zero, and $|Z_{norm}|$ is the transmission at cross fringe phase 0 minus the transmission at cross fringe phase π, or 43% - 1% = 42%.

Note that since the brightness distribution is real, the amplitude and phase of the response are determined completely by $Z(\mathbf{D}_\lambda, \mathbf{D}'{}_\lambda)$, so only three spatial frequency components are actually relevant to modelling Keck Nuller data. Also, notice that if we set $\mathbf{D}'_\lambda = 0$, $Z(\mathbf{D}_\lambda, \mathbf{D}'_\lambda)$ becomes the response that the Keck Nuller would have, if it were a simple single-baseline nulling interferometer. In a sense, the Keck Nuller can be understood as a single baseline nulling interferometer that is offset by the vector $\mathbf{D}'_\lambda$ in the (u, v) plane. For sources smaller than the cross fringe, $|Z_{norm}|$ approaches the transmission of the corresponding single-baseline nuller.

A single-baseline nulling interferometer can not detect the odd component of a brightness distribution. For example, a single-baseline nulling interferometer maybe be adept at detecting faint companions to bright stars, but it can not tell which side of a star a faint companion is on. However, the asymmetry of a source—the odd component of the brightness distribution—appears in the phase of the Keck Nuller response.

Applying the preceeding equations to the case of the small uniform disk described in Section 5.4, we find:

$$\begin{aligned} r_{\star,KN}(u,v) &\approx \frac{1}{16}F_\star\left(\pi\theta_\star\mathbf{D}_\lambda\right)^2\frac{1}{2}\left(1+\cos(2\pi\mathbf{D}_\lambda^{'}\cdot\mathbf{s_0})\right) && (5.36)\\ Z_\star(\mathbf{D}_\lambda,\mathbf{D}_\lambda^{'}) &\approx \frac{1}{16}F_\star\left(\pi\theta_\star\mathbf{D}_\lambda\right)^2 && (5.37)\\ \phi_\star(\mathbf{D}_\lambda,\mathbf{D}_\lambda^{'}) &\approx 0 && (5.38)\\ Z_{\star,norm}(\mathbf{D}_\lambda,\mathbf{D}_\lambda^{'}) &\approx \frac{1}{16}\left(\pi\theta_\star\mathbf{D}_\lambda\right)^2. && (5.39) \end{aligned}$$

In other words the Keck Nuller response to a small star at the phase center is just the corresponding single-baseline nulling interferometer response times the cross fringe at the phase center.

5.6 Photon Noise

Because it operates in the mid-infrared, photon noise in the thermal background will be the limiting noise source for the KN. Our discussion of the KN photon noise budget is based on Serabyn et al. (2000), and follows an outline which is standard for cryogenic cameras. We will call emissivities by the letter ϵ and transmissivities by the letter η.

Consider a beam of light propagating through the interferometer from a primary mirror subaperture to a detector. First the beam encounters some warm optics: the Keck telescopes themselves and the mirrors that relay the beams into the mid-infrared camera. These optics add to the beam a thermal background $\epsilon_n B_\nu(T)$. Given the large number of warm mirrors in the Keck beamtrain, most of the thermal background is generated here in the warm optics, rather than in the atmosphere. Next the beam enters the cryogenic camera dewar, which houses some cold optics and the detector. The cold optics transmit a fraction of the photons η_c and the detector operates with efficiency

η_d, so the background that the detector finally detects is $\eta_c \eta_d \epsilon_n B_\nu(T)$. Serabyn et al. (2000) estimate that $\eta_c \eta_d \epsilon_n$ (the Net Detected Emissivity) will be in the range 0.09–0.17 for various assumed efficiencies.

The KN couples to only a single mode of the radiation field because of a spatial filter in the camera. So the number of detected background photons per second per Hz contributed by each subaperture is

$$\dot{N}_\nu = \eta_c \eta_d \epsilon_n \frac{2}{e^{h\nu/k_b T} - 1}. \tag{5.40}$$

The corresponding incident flux density is

$$F_\nu = \frac{h\nu}{A_{SA} \eta_w \eta_c \eta_d} \dot{N}_\nu, \tag{5.41}$$

where A_{SA} is the collecting area of the subaperture. After an integration of time Δt in a small bandwidth $\Delta\nu$, the photon noise is

$$\Delta \dot{N} = \sqrt{\frac{\dot{N}_\nu}{\Delta\nu \Delta t}} \tag{5.42}$$

photons, so the Noise Equivalent Flux Density (NEFD) for a single subaperture would be

$$\mathrm{NEFD}_{SA} = \frac{h\nu}{A_{SA}\ \eta_w \eta_c \eta_d} \sqrt{\frac{\dot{N}_\nu}{\Delta\nu \Delta t}}. \tag{5.43}$$

However, there are four subapertures to consider, each with signal and background photons. The background is not coherent, so half of it is rejected. We must also account for the fraction of the coherent source flux that translates into detectable signal, $|Z_{norm}|$. Including these factors, we find that the signal to noise ratio for the

KN is $\frac{4}{\sqrt{2}}|Z_{norm}|$ times the signal-to-noise ratio for a single subaperture, or

$$\mathrm{NEFD}_{KN} = \frac{h\nu}{2A_{SA}|Z_{norm}|\eta_w\eta_c\eta_d}\sqrt{\frac{\dot{N}_\nu}{2\Delta\nu\Delta t}}. \tag{5.44}$$

For the KN, a few different possible subaperture geometries would make the effective collecting area of each subaperture about 35% of the area of the primary mirror, or $A_{SA} = 2.7 \times 10^5$ cm^2. Using the estimate $\eta_w\eta_c\eta_d \sim 0.1$ from Serabyn et al. (2000) and taking $\eta_c\eta_d\epsilon_n = 0.13$, we find that the NEFD over a 1 micron bandwidth is $3.2/|\langle Z_{norm}\rangle|$, $6.1/|\langle Z_{norm}\rangle|$, and $11.1/|\langle Z_{norm}\rangle|$ mJy in one second, at 8, 10, and 13 microns, respectively, or $0.084/|\langle Z_{norm}\rangle|$ mJy in one hour in the 10–13 micron band, where the brackets indicate bandwidth averaging.

5.7 Null Depth Fluctuations

So far, our discussion has assumed that the interferometer is stable and symmetrical enough to match the electric fields from the two telecopes so that they exactly cancel at the phase center. In reality, fluctuations in the optical path difference (OPD) between the beams, and asymmetries in the beam paths will mar this cancellation. Serabyn (2000) describes in detail the way these errors translate into fluctuations in the null depth. The net result will be that more on-axis light—starlight—will make its way onto the detector at any instant than in the ideal case. Active control loops can stabilize these fluctuations so that they have a zero mean. However, because the null depth goes as the square of the phase error, some time-averaged leak will remain even with these corrections.

To good approximation, the extra leak does not depend on the size of the star, only on its flux, $F_\star$. The dominant effect will probably be the OPD fluctuations, since they have the short, atmospheric timscales. In this case, the time-averaged extra signal is

$\bar{N}F_\star$ and

$$\bar{N} = \frac{1}{4}\sigma_\phi^2 \tag{5.45}$$

where σ_ϕ^2 is the variance of the phase. The Keck Interferometer design goal is to constrain OPD fluctuations to $\sigma_{\Delta OPD} < 0.032$ microns, which translates into a typical time-averaged leak of $\bar{N} = (0.10\ \mu m/\lambda)^2$, or about 10^{-4} of the stellar flux in the 10-micron band.

Figure 5.10 shows our estimates for the photon noise and the typical time averaged leak, assuming 30 minute integrations. Also plotted in Figure 5.10 are the ranges of signal levels, $|Z|$, that could be expected during a 6-hour track on a single target system at 10 pc, a solar type star and a surrounding exozodiacal cloud with 30 times the density of the solar-system cloud. With the assumptions just described, null depth fluctuations become an important noise source only at wavelengths shorter than 10 microns. It may also be possible to measure and correct for these effects by monitoring a calibrator star. For these reasons, we do not try to simulate null depth fluctuations in any detail.

5.8 Simulation of Exozodiacal Dust Detection

Before we began our simulation of exozodiacal dust detection with the Keck Nuller, we synthesized a library of images of exozodiacal clouds using ZODIPIC on 512 × 512 grids of pixels, where each pixel is one milliarcsecond across. We then multiplied these images by the single-aperture power pattern, padded the grids with zeros to 2048 × 2048 pixels and Fourier transformed the images. Only the lowest frequencies of the resulting visibility matrices are of interest, so we discard all but the lowest frequency 256 × 256 pixels of the transforms, and perform cubic spline interpolation on these matrices to produce our final 1024 × 1024 model visibility files. Our library covers three wavelengths in the 10–13 micron region (10.5, 11.5, and 12.5 microns), and 11 disk inclinations that

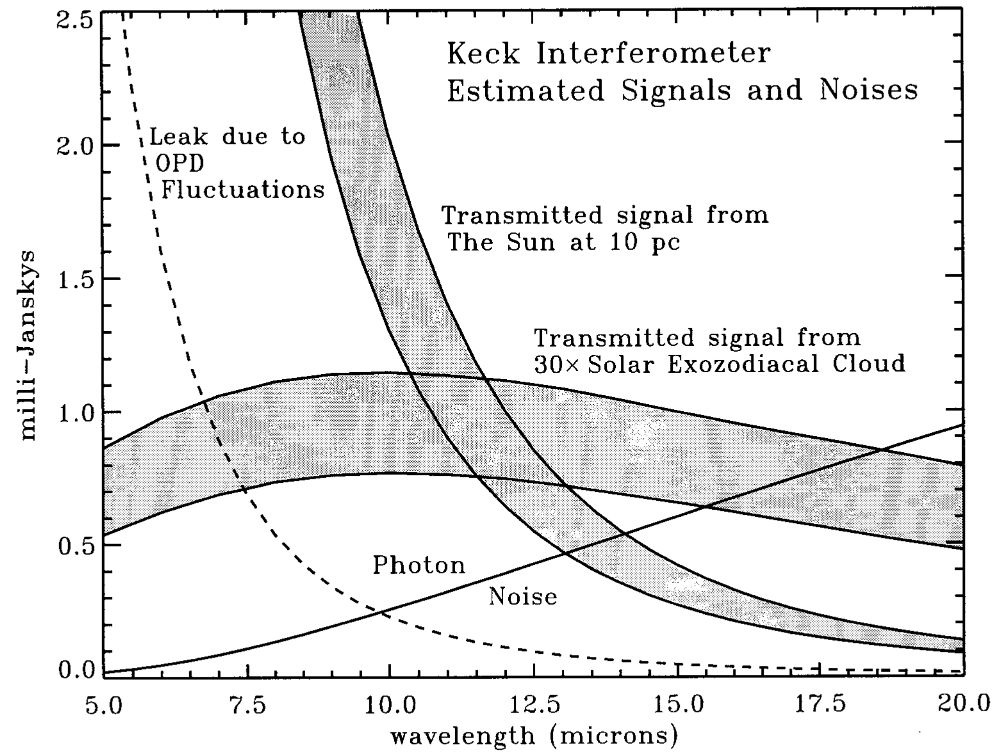

Figure 5.10: A plot of the noise and signal levels relevant to the problem of exozodiacal dust detection with the Keck Interferometer. The shaded region shows the range of possible signals that the KN will transmit from a 30 × solar exozodiacal cloud.

are evenly distributed according to the probability of finding a disk at that inclination, assuming a random three-dimensional distribution of disk axes (most disks are close to edge-on).

The simulation itself has four steps.

1) Choose the parameters of an "observed" model disk: optical depth, inclination, position angle, declination, and so on. The optical depth of the disk is assumed to be some multiple of the optical depth of the solar system cloud.

2) Compute the magnitude of the KN response phasor, $|Z|$, for this disk at 10.5, 11.5, and 12.5 microns as a function of hour angle using equation 5.27 and our library of

visibilities. We divide the night into thirty-minute integrations and assume a zenith angle limit of 45°. We also assume that the short baselines are fixed on the primary mirrors so that they are vertical at altitude 0°. Since the Keck telescopes are alt-az, the primary mirrors rotate as they track, and consequently so do our assumed short baselines. We include the short baseline fringes for completeness, but they are generally unimportant to our detection limits.

3) Add random numbers chosen from a Gaussian distribution to represent photon noise as described in Section 5.6.

4) Try to infer the parameters of the dust disk from the noisy simulated data set. To do this, we try, one after another, all the disks in our library at each of 20 position angles $(0^\circ, 9^\circ, 18^\circ, \ldots)$ to see how well they match the simulated data set. For each trial disk, we compute a response data set without any noise. We then perform a least-squares fit to the simulated data to solve for the optical depth of the dust cloud and note the disk parameters of the best fitting model out of the 220 trial disks.

To model the detectability of a given cloud, we run the simulation over and over and plot the distribution of the optical depths of the best fitting models divided by the optical depth of the input cloud. Figure 5.11 compares three such distributions, each representing 220 runs of the simulation. The distributions are normalized so the area under the curve is unity. The input clouds for each of these distributions are the clouds shown in Figures 5.5, 5.6, and 5.7.

Each of the three input clouds in this figure have optical depth 10 times that of the solar cloud; the only difference among them is the disk orientation. However, the simulation uses trial disks at all possible disk orientations and the mean transmitted flux levels are different for different disk orientations. At this dust level, there is not enough signal to constrain the orientation, so the differences in mean transmitted flux levels at different orientations translate into biases in the distributions of measured optical depths.

Figure 5.12 shows the distributions of simulated measurements of cloud optical

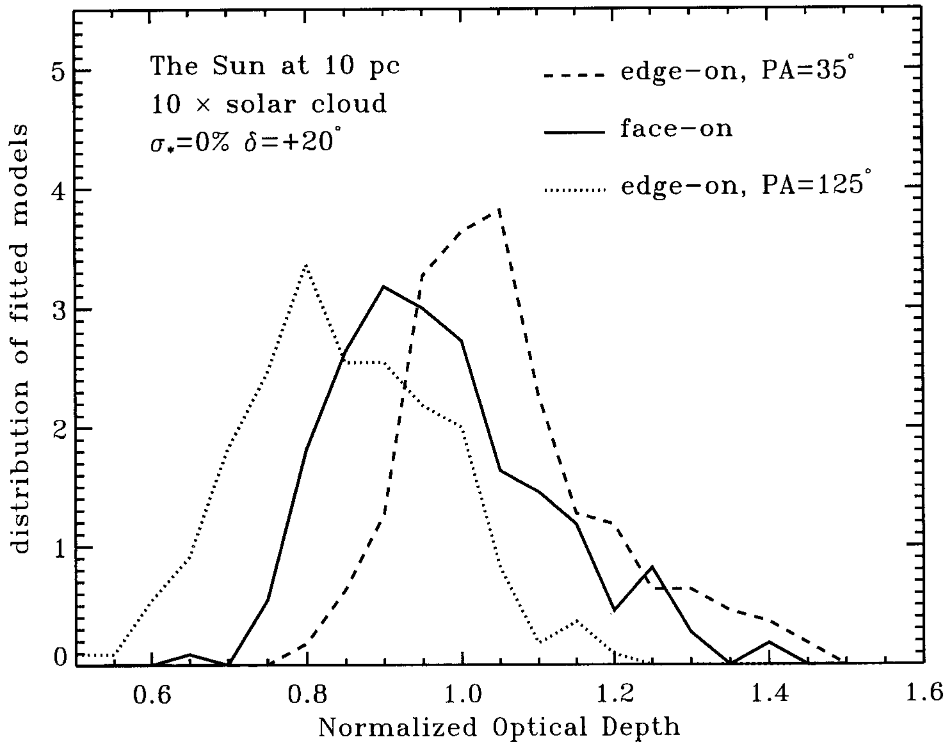

Figure 5.11: The distributions of simulated measurements of dust cloud optical depths for a 10 × solar cloud at 10 parsecs tallied over 220 runs of the simulation. The three curves represent the three different cloud viewing angles shown in Figures 5.5, 5.6, and 5.7. The simulation assumes the disk orientation to be unknown, and at this dust level, there is not enough signal to constrain the orientation. So these three input models yield biased distributions.

depths for the same three disk orientations as in Figure 5.11, but for a 30 × solar cloud. Now, there is enough signal to constrain the orientation. The distributions for the face-on cloud (Figure 5.5) and the edge-on cloud with position angle 35° (Figure 5.6) show only small biases. The distribution for the edge-on cloud with position angle 125° (Figure 5.7) is broad and slightly bimodal because this cloud yields relatively little transmitted signal and resembles the face-on cloud to the interferometer. This signal level is 10 times the detection limit assuming perfect subtraction of the starlight. This level is the minimum level where a night of data can constrain the disk orientation, since as we showed in Section 5.4, the maximum variation in transmitted signal for an exozodiacal cloud during one track is about 10% of the mean transmitted signal.

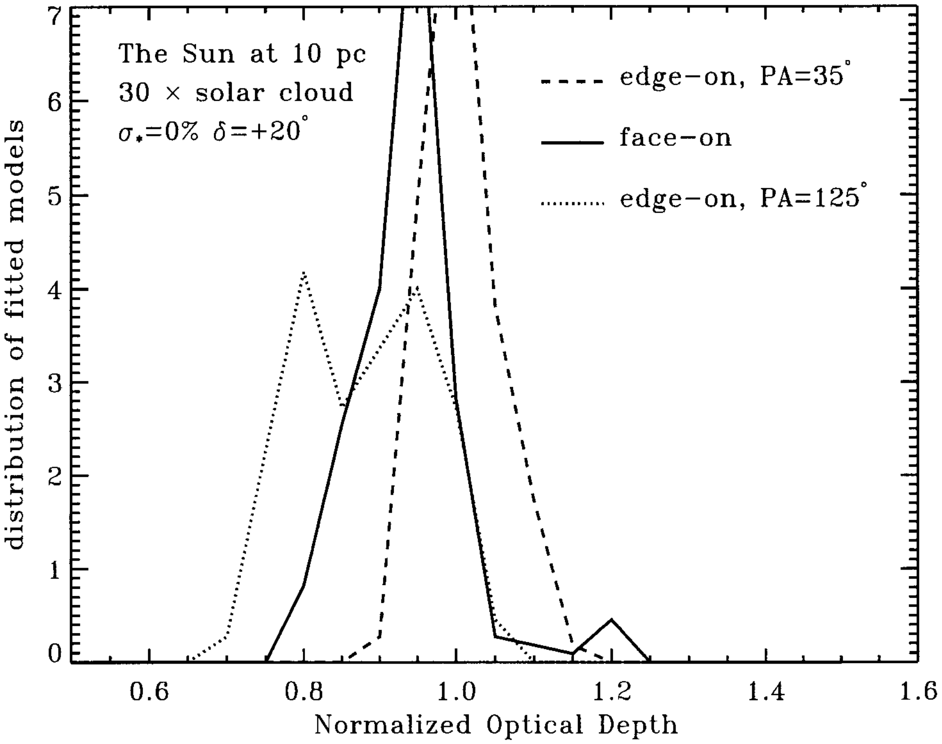

Figure 5.12: The same as Figure 5.6, but for a 30 × solar cloud (10 × the detection limit assuming perfect star subtraction). Now, there is enough signal to constrain the orientation. The distributions for the face-on cloud (Figure 5.5) and the edge-on cloud with position angle 35° (Figure 5.6) show only small biases. The distribution for the edge-on cloud with position angle 125° (Figure 5.7) is broad and slightly bimodal because this cloud yields relatively little transmitted signal and resembles the face-on cloud to the interferometer.

Since we don't know the inclination or position angle of the disk a priori, it makes sense to use a wide variety of input inclinations and position angles and plot the distribution of all the fitted optical depths together. The biases shown in Figure 5.11 then become subsumed by overall width of the distribution. Now the standard deviation of the distribution indicates the accuracy of the simulated measurement. For the remainder of the figures and tables in this paper, we followed this procedure:

1) Choose an input model (optical depth, distance, etc., but not inclination or position angle)

2) Run the simulation described above 220 times for each model. Those 220 simulations cover the 11 inclinations in our library, and 20 different position angles $(0°, 9°, 18°, \ldots)$. This entails fitting 48400 models!

3) Plot the distribution of all the inferred optical depths combined, normalized to the input depth. If the standard deviation of the distribution (σ) is less than 1/3, then we consider the disk to be detectable.

Figure 5.13 shows the distributions of inferred optical depths for 220 runs of the simulation for 3, 10, 30, and 100 × solar clouds covering the full range of input inclinations and position angles, as described above. The distribution for the 3× solar cloud has a standard deviation of 30%; this cloud is marginally detectable by our criterion.

The detection limits described by Figure 5.13 are not much more severe than expectations based on photon noise considerations alone. We find that a 10× solar cloud is clearly detectable, while a 3× solar cloud is clearly not detectable given a single night of data. However, notice that the standard deviations of the distributions do not scale simply as the inverse optical depth of the dust, as you would expect if the cause of the spread in the simulated measurements were purely Gaussian error.

5.8.1 A Priori Knowledge of the Stellar Leak

In reality, as Figure 5.10 shows, the data will contain stellar leak proportional to $F_\star \theta_\star^2/\lambda^2$. For some stars, the angular diameter squared $(\theta_\star^2)$ is known quite precisely. However, for many stars, the best available estimate of $\theta_\star^2$ is found by combining an effective temperature known to a few hundred degrees ($\sim$ 4%) with an infrared flux measurement with similar or better accuracy. To be safe, we assume that existing measurements constrain the leak from most stars to no better than 6%.

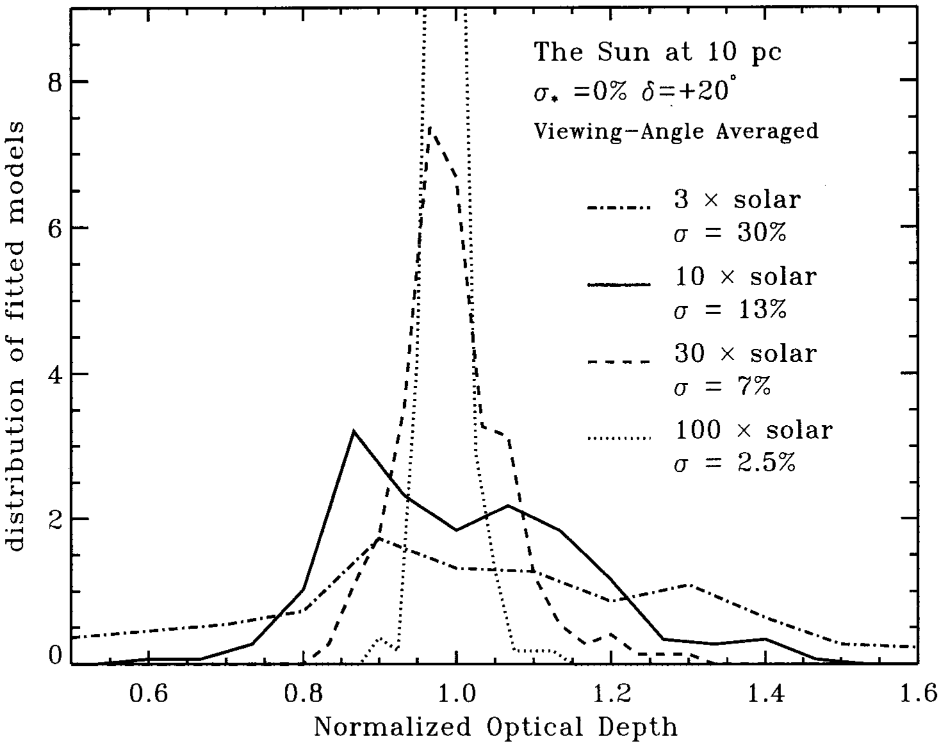

Figure 5.13: The distributions of simulated measurements of dust cloud optical depths for a tallied over 220 runs of the simulation, divided by the optical depth of the input cloud, which was 3, 10, 30, or 100 × solar. For these runs, the input models explored all possible input inclinations and position angles.

Moreover, as we discussed in Section 5.7, null depth fluctuations permit extra starlight to leak through the null. This extra leak should be 10–15% of the stellar signal at 10 microns for a G star at 10 pc and it will have a $\lambda^{-2}F_\star$ wavelength dependence, making it spectroscopically identical to the star. It is uncertain how much of this extra leak can be removed by calibration.

For these reasons, we tested the possibility of deducing the stellar leak from the data itself assuming limited a priori knowledge of the $\lambda^{-2}F_\star$ component of the signal. We ran a set of simulations where we added a term representing the Keck Nuller response phasor due to the star calculated from equation 5.39 to the response phasor calculated for the dust cloud in step two of the simulation, and then performed a two-parameter fit in step four to solve simultaneously for the cloud optical depth and for the stellar leakage. For these simulations we assume that the stellar leak is known a priori to $\sigma_* = 10\%$. Deriving both the dust signal and the star signal from the data works mostly because of the strong wavelength dependence of $Z_\star$ compared to the exozodiacal dust signal as shown in Figure 5.10. We also tried some simulations where we assumed no a priori knowledge of the stellar leak at all, that is, $\sigma_\star = \infty$, and found no significant difference between that and the $\sigma_* = 10\%$ case.

Figure 5.14 shows how the requirement of deriving information about the stellar leak from the data affects the measurement of the zodiacal cloud. Again, the input cloud optical depths are 3,10,30 and 100 × solar. The distributions of best fit optical depths are broader, but the simulations suggest that it is possible to derive the stellar leak from the data, and detect a 10× solar exozodiacal cloud around a solar-type star at 10 pc.

5.8.2 Source Declination

Figure 5.15 illustrates the duration and length of the (u, v) tracks as a function of declination. The source declination affects the (u, v) coverage of the interferometer; at different declinations, the star leak and the signal from the cloud are different functions of hour angle. Since detecting an exozodiacal cloud requires deciding how much of the signal is due to the cloud and how much is due to the star, it is conceivable that changing the functional form of the star leak and the cloud signal could affect the detection limits.

We explored how the declination of a star influences the Keck Nuller's ability to detect exozodiacal dust around that star. We simulated the detection of 10, 30, and 100 × solar clouds at declinations of $-10°$, $20°$, and $50°$ and found little or no difference in the sensitivity of the experiment among the choices. Figure 5.16 compares the distributions of fitted model optical depths for these three choices of source declination

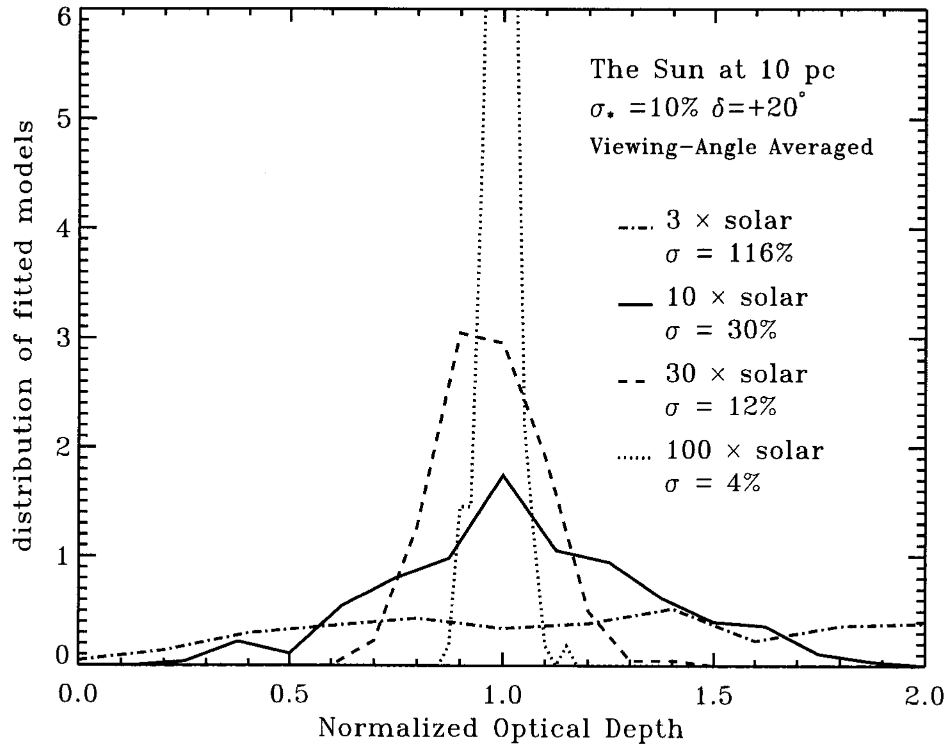

Figure 5.14: A recalculation of the distributions in Figure 5.13 with the stellar leak added to the signal. We assumed an a priori knowledge of the stellar leak of ten percent ($\sigma_\star$ =10%), and performed a two parameter fit to the simulated data sets to recover the dust signal and the stellar signal simultaneously from the data. These simulations suggest that it is still possible to confidently detect a 10× solar exozodiacal cloud in a night of observation given an uncertain star leak.

for a 30× solar input cloud around a solar-type star at 10 pc, assuming a 10% a priori knowledge of the stellar leak. There is no significant difference among these distributions. Of course, sources significantly north of 50° or south of −10° will only be observable for short periods of time, as Figure 5.15 shows.

5.8.3 Stellar Distance and Spectral Type

Finally, we explored the effects of the distance and spectral type of the central star. Naturally, more luminous stars have more luminous disks. To scale the dust models

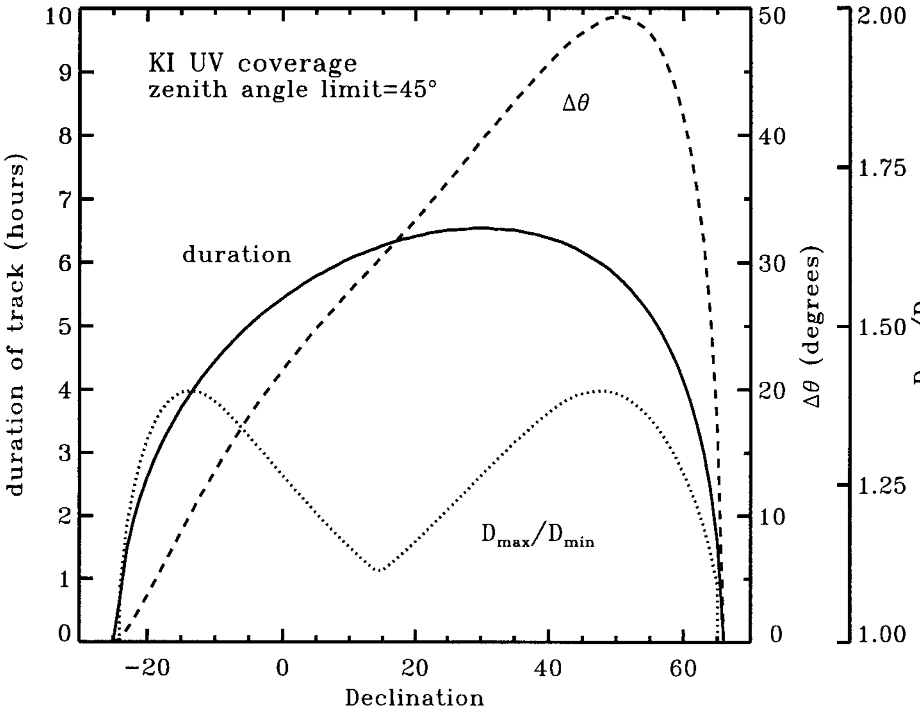

Figure 5.15: The (u, v) coverage of the Keck Interferometer as a function of declination assuming that the interferometer operates at zenith angles $< 45°$. The solid line shows the maximum duration of the track in hours. The dashed line shows the extent of the tracks in polar angle, θ in the (u, v) plane. The dotted line shows the ratio of the maximum projected baseline length to the minimum projected baseline length, D_{max}/D_{min}.

appropriately, we assumed that the temperature of the dust is proportional to the luminosity of the star to the power $\delta/2$, where δ is the negative exponent of the radial power law for the temperature. For the DIRBE model, $T = 286.0K(r/1AU)^{-0.467}$, so ZODIPIC uses $T = 286.0K(r/1AU)^{-0.467}(L/L_{\odot})^{0.234}$. Rather than show more distributions of fitted models, we summarize our results in a table of minimum detectable optical depths (Table 5.1) as defined by the above criterion; a disk is considered detectable when the standard deviation of the distribution of simulated measurements is less than 1/3 of the input dust level.

For each disk, we give three detection limits: the limit for when a single star is

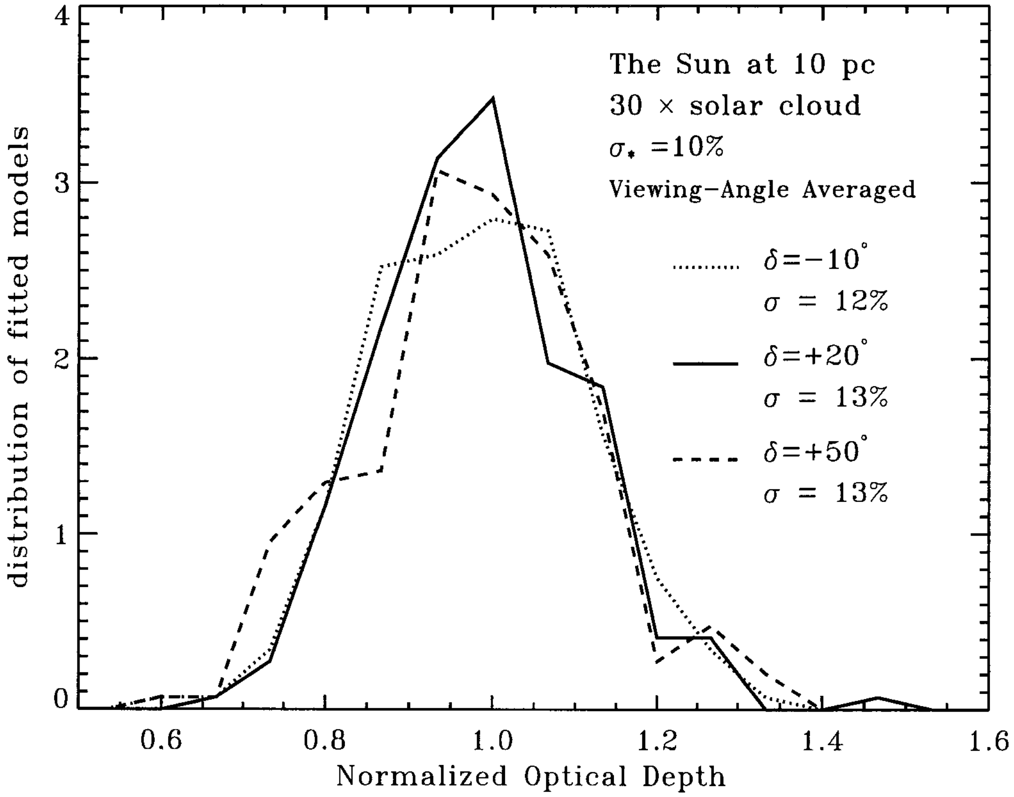

Figure 5.16: Over a wide range, the declination of the star has little effect on dust cloud detection limits. For example, this figure shows that the distributions of best fit models for 30× solar input clouds at −10°, 20°, and 50° have no significant differences.

observed for an entire track and there is no uncertainty in the stellar leak, the limit when a single star is observed for an entire track and the stellar leak is known a priori to 10%, and the limit when the star is observed for only one hour, and the stellar leak is known a priori to 10

Note that the detection limits for stars at 5 parsecs are all better than the detection limits for more distant stars, even though moving from 10 parsecs to 5 parsecs increases the stellar transmission by a factor of 4 and the stellar leak (in photons) by a factor of 16. For the nearest stars, the stellar signal dominates the dust signal even more than for stars at 10 parsecs, but the stellar signal can easily be modeled and subtracted from

Table 5.1: Detection limits, $\delta = 20°$

Spectral Type	Distance (pc)	Detection Limits (solar zodiacal cloud=1) one night, no star	one night, σ_*=10%	one hr., σ_*=10%
F2 V	5	0.5	1.8	6
	10	1.5	5	17
	15	3	10	33
G2 V (The Sun)	5	0.8	4	12
	10	3	9	35
	15	8	22	75
K7 V	5	5	16	55
	10	30	80	430
	15	140	230	800

a data set.

5.9 Conclusion

Our simulations show that the Keck Interferometer can measure faint zodiacal dust signals even when they are mixed with a more powerful transmitted signal from a central star. We succeeded in deriving the strength of the stellar leakage from the simulated interferometric data itself, assuming only imperfect knowledge of the stellar parameters. However, since we derived much lower detection limits when we considered the stellar leak to be perfectly subtractable from the data, we suggest that careful calibration of the leak due to null depth fluctuations and independent measurements of the stellar parameters might improve the experiment.

We found that the KN can begin to constrain the inclination and position angle of a dust cloud when the signal is roughly 10 times the detection limit. As Figure 5.10 shows, varying the disk orientation results in a $\sim$ 50% range in the transmitted signal. At low dust levels, all of this range propagates into the measurement error, but not at high dust levels.

The detection limits for the Keck Nuller are 2–3 orders of magnitude better than our current ability to resolve clouds or detect them photometrically. This large newly accessible region of parameter space may hold many surprises. Our simulation assumed that the exozodiacal cloud is identical to the solar zodiacal cloud. This may be a common appearance for zodiacal clouds, as we have suggested in Section 5.2, but it is far from being general. For clouds near the detection limit, we will be forced to assume many of the cloud properties, and declaring the cloud to be analogous to the solar cloud is as good an assumption as any. However, when there is enough signal and (u, v) coverage for the Keck Nuller to constrain the cloud morphology given a long track, a dust cloud that is a simple solar-system analog is inadequate. The dynamics of sublimating dust grains, the interaction of dust with planets, the collisional processing of dust grains, and the emission of a 10-12 micron silicate feature are all phenomena we have not explored that may influence the appearance of these higher mass clouds.

Chapter 6

A Search for Resonant Structures in the Zodiacal Cloud with COBE DIRBE: The Mars Wake and Jupiter's Trojan Clouds

Marc J. Kuchner
Palomar Observatory
California Institute of Technology, Pasadena, CA 91125
William T. Reach
Infrared Processing and Analysis Center
Caltech, Pasadena, CA 91125
Michael E. Brown
Division of Geological and Planetary Sciences
California Institute of Technology, Pasadena, CA 91125

We searched the COBE DIRBE Sky and Zodi Atlas for a wake of dust trailing Mars and for Trojan dust near Jupiter's L5 Lagrange point. We compare the DIRBE images to a model Mars wake based on the empirical model of the Earth's wake as seen by the DIRBE and place a 3-σ upper limit on the fractional overdensity of particles in the Mars wake of 18% of the fractional overdensity trailing the Earth. We place a 3-σ upper limit on the effective emitting area of large (10-100 micron diameter) particles trapped at Jupiter's L5 Lagrange point of 6×10^{17} cm^2, assuming that these large dust grains are distributed in space like the Trojan asteroids. We would have detected the Mars wake if the surface area of dust in the wake scaled simply as the mass of the planet times the Poynting-Robertson time scale.[1]

---

[1] Adapted from Kuchner et al. (2000)

6.1 Introduction

A planet interacting with a circumstellar dust cloud can produce a variety of dynamical structures in the dust. Planets can clear central holes and create large-scale asymmetries, such as arcs and warps (Roques et al. 1994). Planets can also detain dust in mean motion resonances, forming structures such as the circumsolar ring and wake of dust trailing the Earth in its orbit (Jackson and Zook 1989; Dermott et al. 1994) and clouds of dust at the planet's Lagrange points (Liou and Zook 1995).

Understanding the structures of circumstellar debris disks is vital to the search for extra-solar analogs of our solar system. Concentrations in circumstellar dust clouds may confuse planet-finding interferometers like the Keck Interferometer or the proposed Terrestrial Planet Finder (Beichman et al. 1999). Smooth exozodiacal clouds can be identified by their symmetry and subtracted from the signal of a Bracewell interferometer (MacPhie and Bracewell 1979), but cloud asymmetries can be difficult to distinguish from planets (Beichman 1998). On the other hand, planet-induced asymmetries can serve to reveal the presence of a planet that is otherwise undetectable (Wyatt et al. 1999).

If we understand the inhomogeneities in our own zodiacal dust cloud, we will be better prepared to interpret observations of other planetary systems. The Diffuse Infrared Background Experiment (DIRBE) aboard the Cosmic Background Explorer (COBE) satellite has provided detailed, revealing images of the zodiacal cloud (e.g., Spiesman et al. 1995, Reach et al. 1997). It surveyed the entire sky from near-Earth orbit in 10 broad infrared bands simultaneously with a 0.7° by 0.7° field of view over a period of 41 weeks (Boggess et al. 1992), and imaged the Earth's ring and wake (Reach et al. 1995). We investigated the COBE DIRBE data set as a source of information about structure in the solar zodiacal cloud associated with planets other than Earth.

6.2 The Data Set

We worked with a version of the DIRBE data set which contains the sky brightness with a model for the background zodiacal emission subtracted: the zodi-subtracted weekly data set from the DIRBE Sky and Zodi Atlas (DSZA). The DIRBE team created the DZSA by fitting an 88-parameter model of the zodiacal dust emission to the observed sky brightness (Kelsall et al. 1998). The model includes a smooth widened-fan component, three pairs of dust bands near the ecliptic plane, and the Earth's ring and trailing wake, but no structures associated with Mars, Jupiter, or any other planets. The zodi-subtracted weekly data set contains 41 files for each DIRBE band, spanning a period from 10 December 1989 to 21 September 1990, covering 10 bands, centered at 1.25, 2.2, 3.5, 4.9, 12, 25, 60, 100, 140, and 240 microns.

Galactic emission dominates the zodi-subtracted maps in the mid and far-infrared near the galactic plane. Near the ecliptic plane, the zodi-subtracted maps are dominated by residuals from the subtraction of the dust bands that are associated with prominent asteroid families (Reach et al. 1997; Kelsall et al. 1998). The presence of these bands makes searching for smooth, faint heliocentric rings of dust near the ecliptic plane impossible. However, we could hope to distinguish a blob of dust following a planet across the sky from other cloud components and from the galactic background by the apparent motion of the blob during the COBE mission.

The COBE satellite orbited the Earth near the day/night terminator and repeatedly mapped a swath of the sky extending about 30 degrees before and behind the terminator (see the COBE DIRBE Explanatory Supplement (1997) for details). Each weekly map contains a robust average of all the week's data and covers a region a little larger than the daily viewing swath. This weekly averaging tends to exclude transient events that would contaminate our final maps, but should not otherwise significantly affect a search for large features that move only a few degrees per week. Figure 6.1 is a schematic view of the solar system during week 34 of the mission (9–16 July 1990) showing the positions of Earth, Mars and Jupiter, and the DIRBE viewing swath for that week.

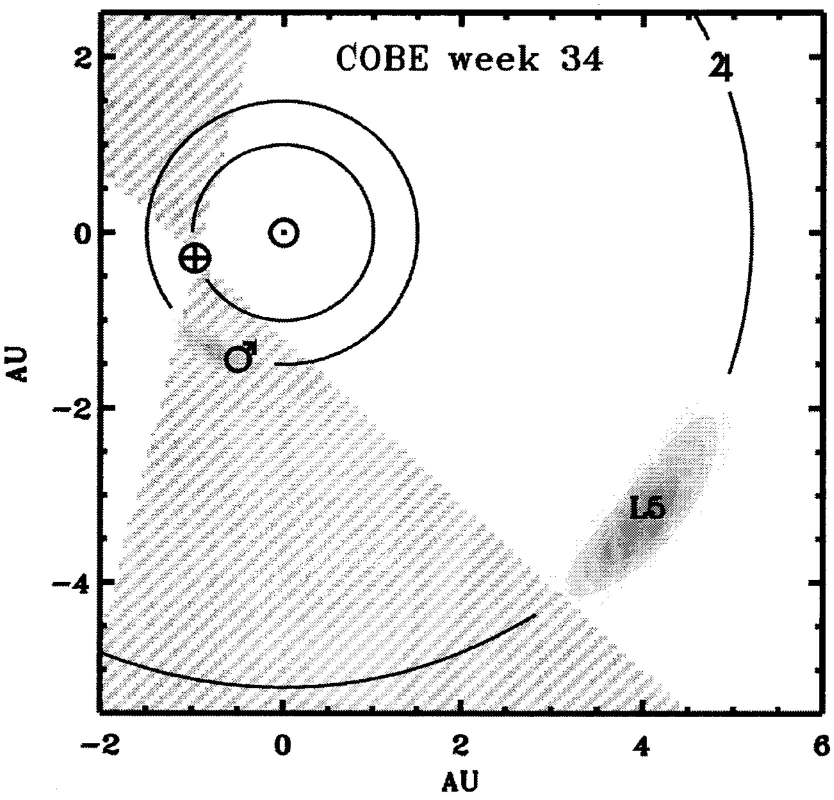

Figure 6.1: The solar system during week 34. The shaded regions following Mars represents our model for the Mars wake; the shaded region centered on L5 represents our model for the Trojan cloud. The hatched area represents the DIRBE viewing swath for that week.

Because DIRBE never imaged the sky within 60° of the sun, the orbits of Mercury and Venus, for instance, do not appear in the data. Mars appeared in the DIRBE viewing swath for 25 weeks of the mission, and moved 111° in ecliptic longitude during those weeks. Jupiter moved only 40° in ecliptic longitude during the entire mission, but this is sufficient to allow some crude background subtraction. More distant planets moved less. Based on these constraints, we decided to search the weekly maps for dust features following the orbital paths of Mars and Jupiter. Figure 6.2 shows the intersection of the DIRBE viewing swath with the ecliptic plane throughout the 41 weeks of the mission, and the ecliptic longitudes of Mars, Jupiter and the Sun during those weeks.

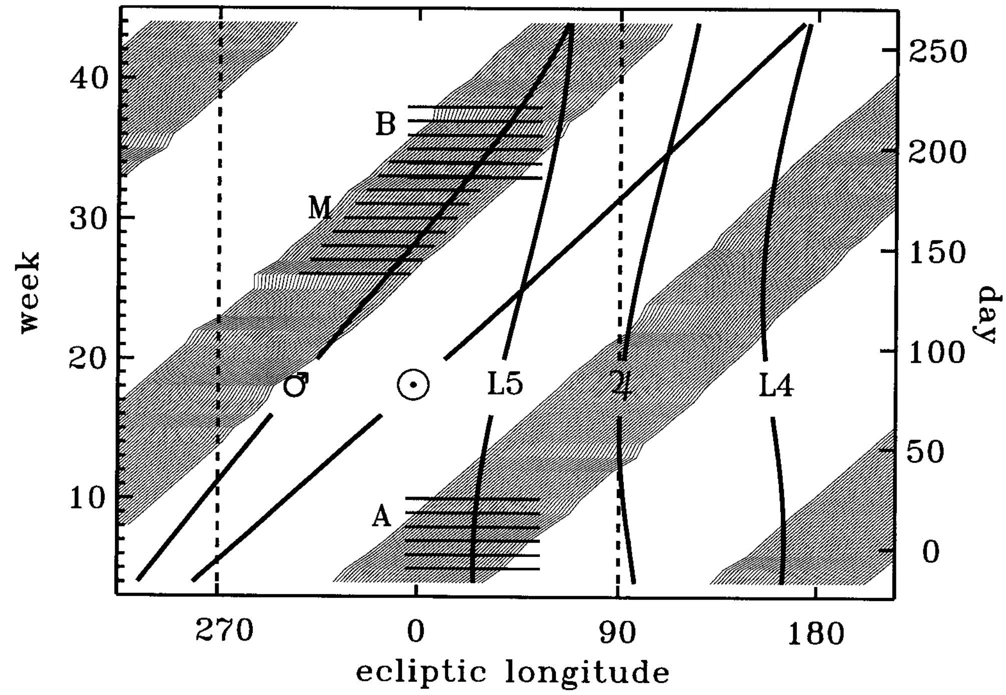

Figure 6.2: The ecliptic longitudes of the Sun, Mars, Jupiter, and Jupiter's L4 and L5 Lagrange points during 40 weeks of the COBE mission when DIRBE was recording. The shaded diagonal stripes show the intersection of the DIRBE viewing swath with the ecliptic plane. The vertical dashed lines show where the galactic plane crosses the ecliptic. The horizontal bars show the data sets used to construct Figures 6.4, 6.5, and 6.6.

6.3 The Mars Wake

A ring of zodiacal dust particles detained in near-Earth resonances follows the Earth around the sun (Jackson and Zook 1989; Dermott et al. 1994). This ring consists mainly of dust in mean-motion resonances where the particles orbit the sun j times every $j + 1$ Earth years (j is a whole number). Smaller trapped particles experience greater Poynting-Robertson acceleration, so the equilibrium locations of their orbital pericenters shift closer to the Earth on the trailing side, where the component of Earth's gravity that opposes Poynting-Robertson drag is stronger. The result, averaged over many particles, appears as a density enhancement in the ring behind the Earth—a trailing dust wake. The Earth's wake was detected by by IRAS (Dermott et al. 1988; Reach 1991), and later, by DIRBE as an asymmetry in the near-Earth dust brightness of $\sim$ 1.1 MJy ster^{-1} at 12 microns and $\sim$ 1.7 MJy ster^{-1} at 25 microns (Reach et al. 1995). We searched the DIRBE data set for a similar wake of dust trailing Mars.

Blackbody dust at the heliocentric distance of Mars has a typical temperature of $\sim$ 220 K; it emits most strongly in the 12 and 25 micron DIRBE bands. We restricted our exploration to data from these two bands. We began by assembling a composite map of the emission from beyond the solar system, mostly due to stars and dust in the Galactic plane, by averaging together all the zodi-subtracted weekly maps in their native COBE quadrilateralized spherical cube coordinates, a coordinate system that is stationary on the celestial sphere. We subtracted this composite map from each of the zodi-subtracted weekly maps, effectively removing most of the galactic emission and any other stationary emission except within a few degrees of the galactic plane, where the emission is so high that detector and pointing instabilities make our linear subtraction method ineffective.

The remaining maps, with outlying data removed, had surface brightness residuals in the range of -1.7 to +1.0 MJy ster^{-1} at 12 microns, and -1.6 to 2.1 MJy ster^{-1} at 25 microns. For comparison, the typical total zodiacal background near Mars during the mission is $\sim$ 35 MJy ster^{-1} at 12 microns and $\sim$ 66 MJy ster^{-1} at 25 microns. The most prominent remaining features were the stripes parallel to the ecliptic plane within a few degrees of the ecliptic plane that are associated with the asteroidal dust bands. The next most prominent remaining features were wide bands extending ±30 degrees from the ecliptic that appeared to follow the sun. The $12 - 25$ micron color temperature of the wide bands was $\sim$ 280 K; they are probably residuals resulting from imperfect subtraction of the Earth's ring and wake. We assembled a crude map of the residual near-Earth flux by averaging together the galaxy-subtracted maps in geocentric ecliptic coordinates referenced to the position of the Sun. Subtracting this from the weekly maps cancelled most of the signal in the wide bands. Mars moved 87° with respect to the Sun during the mission, allowing us to subtract this composite map without subtracting a significant flux from a wake moving with Mars. Figure 6.3 shows our map of the galactic background; Figure 6.4 shows the near-Earth residuals.

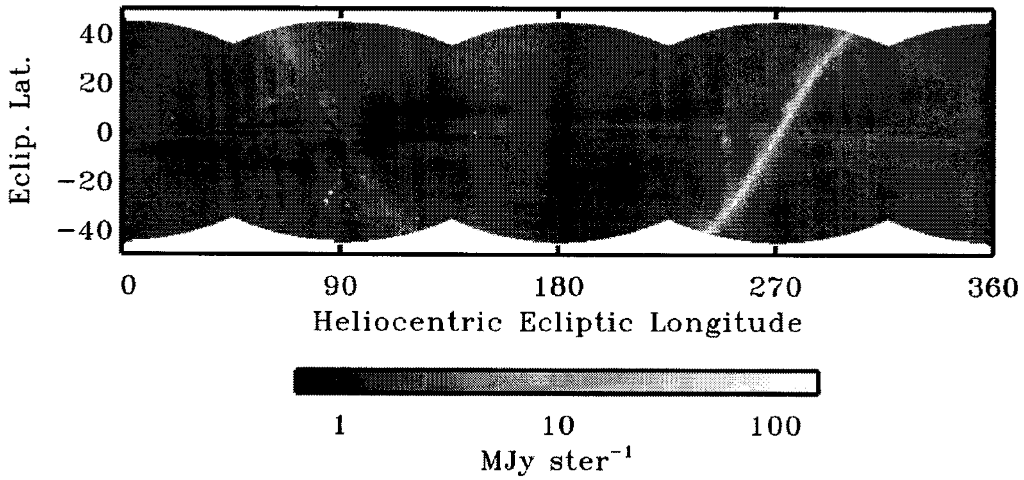

Figure 6.3: The galactic background at 25 microns, constructed by averaging all the weekly maps in their native quadrilateralized spherical cube coordinates.

Next we chose subframes of each weekly map centered on the ecliptic coordinates of Mars in the middle of the week, and inspected them visually. No structure in the data appeared to move with Mars from week to week.

In order to understand the data better, we constructed a simple model of the Mars

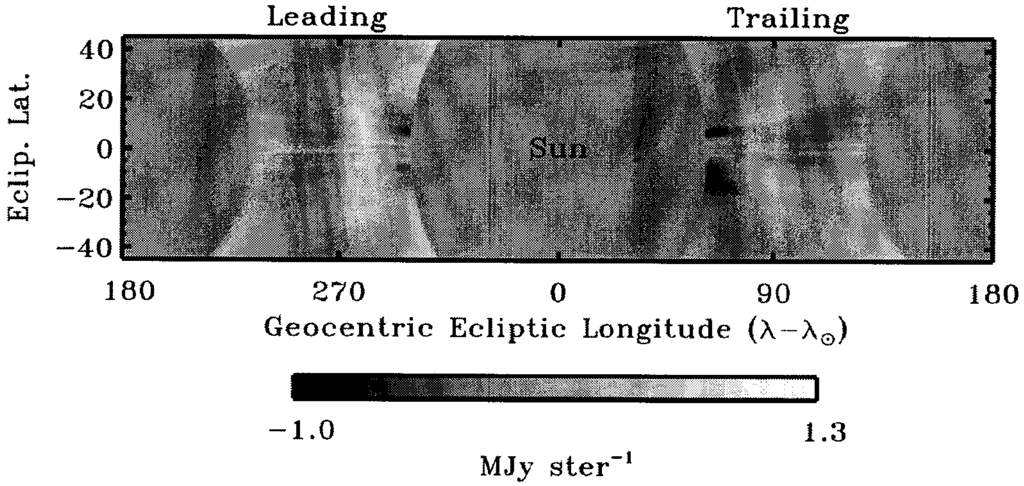

Figure 6.4: Residuals from the Earth's wake at 25 microns, constructed by averaging all the weekly galaxy-subtracted maps in a geocentric ecliptic coordinate system with the sun at the origin.

wake from the empirical model of the Earth's trailing wake fit to the DIRBE data by Kelsall et al. (1998). The model has the following form:

$$n = n_0 \exp\left[-\frac{(r-r_0)^2}{2\sigma_r^2} - \frac{|z|}{\sigma_z} - \frac{(\theta-\theta_0)^2}{2\sigma_\theta^2}\right] \tag{6.1}$$

where n is the local average of particle number density times particle cross section, and r,z, and θ are cylindrical coordinates in the plane of the orbit of Mars centered on the sun. Mars is located at $r = r_0$, $z = 0$, $\theta = 0$. The parameters of the model, θ_0, σ_r, σ_z, σ_θ, and n_0, are the same as the corresponding parameters for the Earth's wake: $\theta_0 = -10°$, $\sigma_r = 0.10$ AU, $\sigma_z = 0.091$ AU, $\sigma_\theta = 12.1°$. The shaded area trailing Mars in Figure 6.1 shows how this model would appear viewed from above the ecliptic plane. The Kelsall et al. (1998) Earth wake has $n_0 = 1.9 \times 10^{-8}$ AU^{-1}, but we chose $n_0 = 1.08 \times 10^{-8}$ AU^{-1} so that the density of the model would be proportional to the local background dust density at the orbit of Mars. The model represents what the Earth wake would look like if it were trailing Mars instead of Earth.

We evaluated the model's surface brightness by computing the line-of-sight integral

$$I_\lambda = E_\lambda \int n(r, z, \theta) B_\lambda(T) ds \tag{6.2}$$

where $B_\lambda(T)$ is the Planck function and E_λ is an emissivity modification factor prescribed by the COBE model to account for the deviation of the Earth wake's spectrum from a blackbody; $E_{12\mu\mathrm{m}} = 1.06$, $E_{25\mu\mathrm{m}} = 1.00$. The temperature of the dust varies with heliocentric distance, R, as $T = 286$ K $R^{-0.467}$, following the DIRBE model. This expression is similar to what you would expect for grey-body dust ($T = 278$ K $R^{-0.5}$).

In Figure 6.5, we compare a synthesized image of the model wake with a background-subtracted image of the infrared sky around Mars. The image shows the flux in the 25 micron band averaged over weeks 26–34 (14 May 1990 to 15 July 1990) in ecliptic coordinates referenced to the position of Mars. The subset of the weekly 25-micron data used in this image is indicated by the horizontal stripes labeled "M" in Figure 6.2. Mars moved 40 degrees in ecliptic longitude over this period. The DIRBE team blanked the data within a square about 2.5° on a side centered on Mars, and within a 1.5° radius circle centered on Jupiter. A software mask in Figure 6.5 covers the region around Mars affected by this processing. This whole region, up to 40° behind Mars, shows no sign of a brightness enhancement that we would associate with a wake of dust trailing Mars.

Maps from the later weeks suffer from an oversubtraction due to imperfections in the Kelsall et al. (1998) zodi model, visible as the dark region to the lower right. Weeks earlier in the mission suffer from a similar undersubtraction. These artifacts, our primary source of noise, appear to arise from dust bands at latitudes of $\sim \pm 10^\circ$, where bands associated with the Eos asteroids are prominent in the raw DIRBE data (Reach

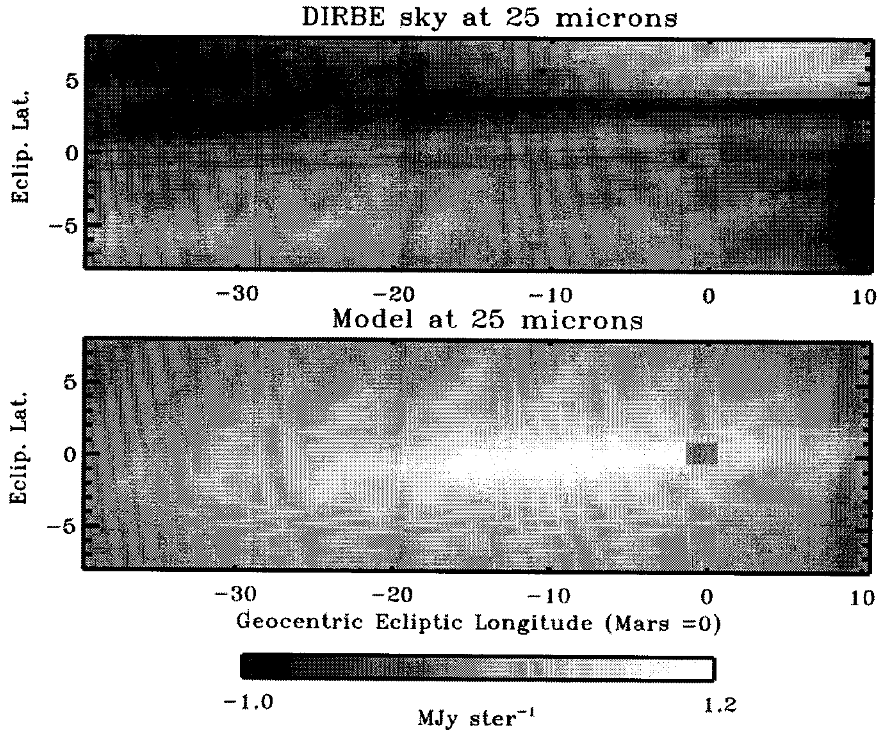

Figure 6.5: An image of the sky near Mars at 25 microns, compared to a model based on the COBE DIRBE empirical model for the wake trailing Earth. The image is averaged over weeks 26–34 (data set M). The region within 1.5° of Mars has been covered by a software mask.

et al. 1997). As the dust from the asteroid belt spirals towards the sun, perturbations from planets deform the bands. The Kelsall et al. model includes a simple model of this dust band which could not take these perturbations into account. We chose the span of weeks used to create Figure 6.4 to minimize these artifacts, which are easily discernible by their extent in latitude and longitude.

To better compare the model with the data, we focused on a narrow strip with a height of 3° in ecliptic latitude, extending from 8° ahead of Mars to 39° behind Mars in ecliptic longitude. This strip contains most of the flux in the model wake. We averaged together maps from weeks 26–34 prepared as described above to produce an image of this strip. In Figure 6.6, we plot a cut through this strip, and we compare it with the model, processed in the same manner as the data. The data are dominated by residuals from the ecliptic bands and the Earth's ring, smeared out in the ecliptic plane by the orbital motion of Mars. The standard deviation of the data is 0.54 MJy ster^{-1}; although the distribution of the residuals is not Gaussian, based on this comparison we can place a rough 3-σ upper limit on the central peak of the Mars wake of 18% of the flux expected from our simple model.

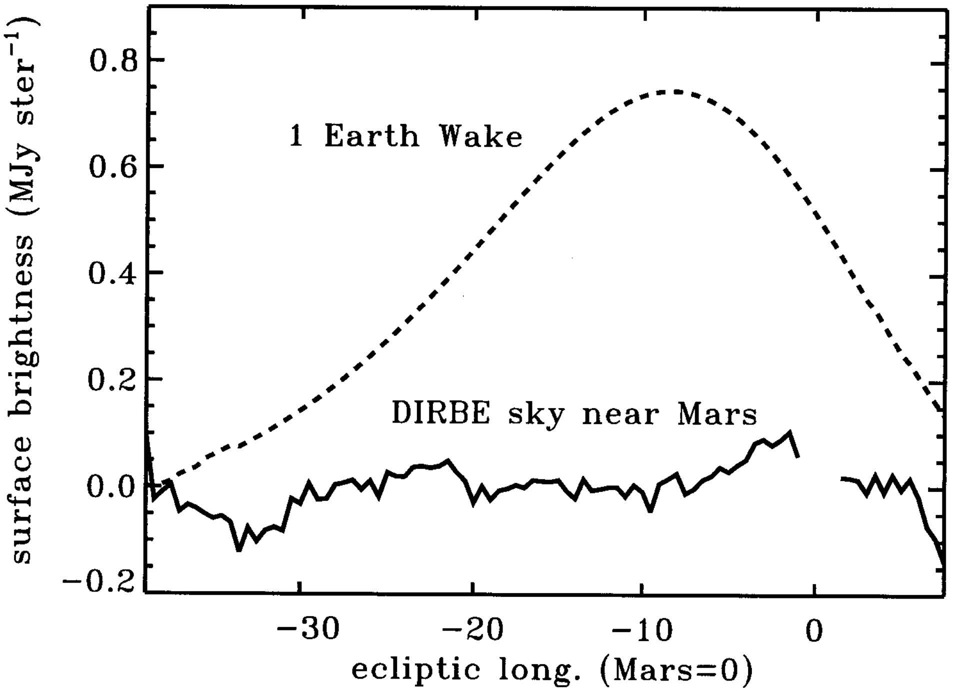

Figure 6.6: A cut through the image of the 25-micron sky near Mars shown in Figure 6.5, compared to the same model.

The empirical model of the Earth's wake we have used for comparison to the Mars dust environment is not an ideal model for the Mars wake. It may not even be a good representation of the Earth's wake. Since COBE viewed the Earth wake from near the Earth only, the observations constrain the product $n_0\sigma_\theta$ for the Earth wake, but do not provide good constraints on either of these parameters alone. Kelsall et al. (1998)

quote a formal error of 28% on the determination of σ_θ. Calculations for 12 micron particles suggest that σ_θ for the Earth wake might be 40% lower than the Kelsall et al. (1988) number; this figure is based on Figure 5 in Dermott et al. (1994). Since we are sensitive to the wake's surface brightness peak as seen from the Earth, not Mars, using a more compact wake model affects our upper limits. Holding $n_0\sigma_\theta$ constant and decreasing σ_θ by 40% translates into a decrease of our upper limit to 11% of one Earth wake.

Mars has 11% of the mass of the Earth, so we expect it to trap less dust than the Earth, but not simply 11% as much dust. In fact, there is no simple scaling law that describes how the density of a dust ring relates to the size of the planet that traps it. The density of the Mars ring is proportional to the capture probability times the trapping time for each resonance summed over all relevant resonances and the distribution of particle sizes. In the adiabatic theory for resonant capture due to Poynting-Robertson drag, the capture probabilities depend on the mass of the planet compared to the mass of the star and on the eccentricity of the particle near resonance and (Beaugé and Ferraz-Mello 1994). So one complicating factor is that the orbits of the dust particles are slightly more eccentric when they pass Mars than when they pass the Earth; a particle released on the orbit of a typical asteroid, at 2.7 AU with an eccentricity of 0.14, will have an eccentricity of 0.07 as it passes Mars, and an eccentricity of 0.04 when it passes the Earth (Wyatt and Whipple 1950). The higher eccentricity makes them harder to trap.

The trapping time scale is proportional to the time it takes for the resonant interaction to significantly affect the eccentricity and libration amplitude of the particle. When the planet has a circular orbit, these time scales are on the order of the local Poynting-Robertson decay time (Liou and Zook 1997), which scales as r_0^2/β, where β is the ratio of the Sun's radiation-pressure force on a particle to the Sun's gravitational force on the particle. Compared to the P-R drag at the heliocentric distance of the Earth, the Poynting-Robertson drag force at the orbit of Mars is less for a given particle by a factor of $1.52^2 = 2.31$. The small mass of Mars and the higher eccentricities of the orbits of the incoming particles work against the formation of a dense ring, but the greater heliocentric distance of Mars compared to the Earth works in favor of the formation of the ring.

So far our discussion has assumed that the trapping is adiabatic— that the orbital elements of the particles change on time scales much longer than the orbital period. This approximation may not be as good for trapping by Mars as it is for trapping by the Earth. Mars has a greater orbital eccentricity ($e = 0.093$) than the Earth ($e = 0.017$). This increases the widths of the zones of resonance overlap, and makes a larger fraction of dust orbits chaotic (Murray and Holman 1997).

Predicting the density of the Mars wake is another step more complex than predicting the density of the Mars ring. Compared to the Earth wake, the Mars wake may form closer to the planet and have a smaller σ_θ. Since Mars is less massive than the Earth, a given particle would need to have a closer interaction with Mars than with the Earth to receive an impulse from the planet's gravity that would balance the Poynting-Robertson drag on the particle (Weidenschilling and Jackson 1993). For this reason, we expect the trapped particles which form the Mars wake to prefer resonant orbits with higher j and lower ϕ than similar particles trapped by the Earth, where ϕ is the angle between the perihelion of the orbit of a particle and the longitude of conjunction of the particle and the planet. Our upper limit shows that the Mars wake is less dense than the Earth wake by more than the simple factor of the mass ratio times the square of the ratio of the semimajor axes $= 0.11 \times 2.31 = 0.25$. However, a thorough numerical simulation which includes the effects we mentioned and others such as resonant interactions with Jupiter may be the only good way to relate our upper limit to the dynamical properties of the dust near Mars.

6.4 Trojan Dust

While the Earth and Mars can collect abundant low eccentricity particles from all different orbital phases spiraling in from the asteroid belt, Jupiter orbits in a distinctly different dust environment. Outside the asteroid belt, the dust background probably consists mainly of small particles with high orbital eccentricities: submicron particles released by asteroids or comets that are kicked by radiation pressure into more eccentric orbits than their parent bodies (Berg and Grün 1973; Mann and Grün 1995b). There is also a stream of submicron particles from the interstellar medium (Grün et al. 1994; Grogan et al. 1996) and there are probably a few particles near Jupiter that originated in the Kuiper belt (Liou et al. 1996). Jupiter probably traps many of the small particles in 1:1 mean motion resonances (Liou and Zook 1995). However, these small trapped particles should occupy both "tadpole" and "horseshoe" orbits, without a strong preference for either, and the locations of their Lagrange points vary with β (Murray 1994). They probably form large, diffuse ring-like clouds which are difficult for us to detect.

But there is another potential source of dust that could form concentrated clouds we could hope to detect against the asteroid bands in the DIRBE data: the Trojan asteroids. This population of asteroids orbits the Sun at $\sim$ 5.2 AU in 1:1 resonances with Jupiter, librating about Jupiter's L4 and L5 Lagrange points, roughly 60° before and behind the planet. They number about as many as the main-belt asteroids.

Marzari et al. (1997) have simulated the collisional evolution of the Trojan asteroids, and concluded that collisions in the L4 swarm produce on the order of 2000 fragments in the 1–40 km diameter range every million years. If we simplisticly assume an equilibrium size distribution for the produced particles, $dn \propto a^{-3.5} da$, where a is the particle radius (Dohnanyi 1969), we find that there are roughly 10^{23} particles in the 10–100 micron diameter size range produced every million years. These large particles are likely to stay in roughly the same orbits as their parent bodies, trapped by Jupiter in "tadpole" orbits—orbits that librate around a single Lagrange point. They could conceivably form detectable clouds at L4 and L5.

Liou and Zook (1995) calculated that 2-micron diameter particles will stay trapped in 1:1 resonances for $\sim$ 5000 years. A 20-micron diameter particle at Jupiter's orbit experiences 1/10 of the Poynting-Robertson acceleration of 2 micron particles, and will typically stay trapped for 10 times as long (Schuerman 1980). Assuming a trapping time of 5000 years × the dust grain diameter/2 microns, and emissivity appropriate for amorphous icy grains (e.g., Backman and Paresce 1993), the 10–100 micron diameter particles in the Trojan cloud will emit a total flux, as viewed from the Earth, of $\sim 3 \times 10^{-4}$ MJy at 60 microns, a few orders of magnitude below our detection limit.

However, this is a drastic extrapolation and probably a poor guess at the actual cloud brightness; the size-frequency distribution of the Trojan asteroids is not well known and dust cloud is probably not near collisional equilibrium. Moreover, the total amount of trapped dust is subject to severe transients, such as the events that produced the dust bands associated with main belt asteroid families (Sykes and Greenberg 1986). For example, a 20-km diameter Trojan asteroid ground entirely into 10-micron diameter dust corresponds to a transient cloud which, as viewed from the Earth, would produce a 60 micron flux of $\sim$ 6 MJy. A similarly enhanced cloud might be visible a few percent of the time.

Unfortunately, Jupiter's Lagrange points do not move far with respect to the galactic background during the COBE mission; L4 moves 10 degrees and L5 moves 50 degrees, as shown in Figure 6.2. Only L5, the trailing Lagrange point, moves far enough during the mission to make subtracting the galactic background feasible. There are about half as many L5 Trojans known as L4 Trojans, but this is probably because the L5 region has been searched less intensely than the L4 region, not because the L4 and L5 populations are significantly different (Shoemaker et al. 1989).

To make a background-subtracted image of the L5 region, we chose two subsets from the zodi-subtracted data set. The first, subset A, is from the beginning of the mission (weeks 5-10) when L5 was in the viewing swath and approximately stationary on the sky. The second, subset B, is the same region of sky, but contains data from half a year later in the mission (weeks 33–38), when L5 has moved 45 degrees away, out of the viewing swath. The average distance from Earth to L5 is approximately the same during each time period. These data sets are depicted in Figure 6.2. We focused on data in the 60 micron band, the band which contains the emission peak for dust at the local blackbody temperature at 5.2 AU. To minimize the residuals from the zodiacal dust model, we used only data from solar elongations between 65 and 115 degrees (or between 245 and 295 degrees). Figure 6.7 shows an image constructed from data set A and an image constructed from data set B, and the difference, A$-$B, which is dominated by residuals from dust bands associated with asteroid bands and shows no obvious evidence of enhanced emission at L5.

We made a simple model for a Trojan cloud of large dust particles by assuming that they occupy the same dynamical space as the Trojan asteroids themselves, following Sykes (1990). Sykes modeled the asteroidal dust bands by showing how particles constrained to orbits with a given inclination, eccentricity, and semimajor axis form a cloud when the remaining three orbital elements are randomized. He then convolved the shapes of these clouds with the distributions of orbital elements of the asteroids. In our case, however, the distribution of orbital elements is much broader and more important in determining the shape of the final distribution of particles. Since there are only about 70 Trojan asteroids whose orbits are well studied, the distributions of Trojan asteroid orbital parameters have severe statistical uncertainties. Therefore, we settle for a simple Gaussian model for the Trojan cloud, using the orbital parameters as a guide to the parameters of the Gaussian. In the following calculations, we will neglect the inclination of Jupiter's orbit relative to the Earth's orbit (1.305°).

A typical Trojan asteroid librates around its Lagrange point with a period of 148 days. The mean longitude of the asteroid with respect to Jupiter, ϕ, oscillates within

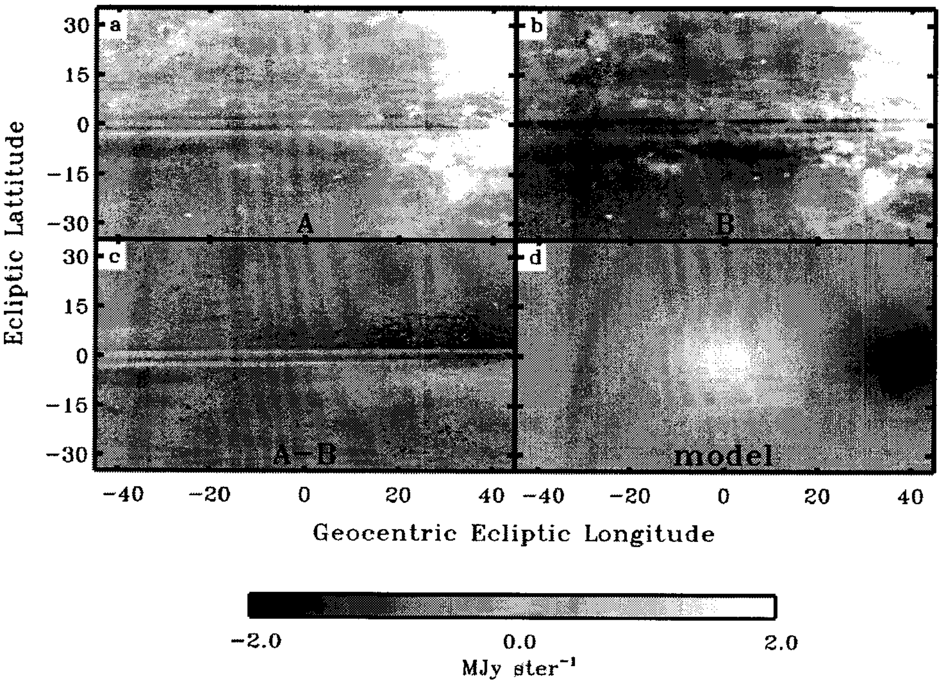

Figure 6.7: Raw DSZA images in the ecliptic plane at 60 microns. L5 is at the center of image A, but it has moved 45 degrees to the right of center in image B. The difference, A−B cancels most of the galactic emission, but is dominated by residuals from dust bands associated with the asteroid belt and does not reveal any Trojan dust. The model shows what we would expect the difference A−B to look like, given some simple assumptions about the Trojan clouds.

limits ϕ_{min} and ϕ_{max}, which can be calculated, according to Yoder et al. (1983), from

$$\sin\frac{\phi_{min}}{2} = \frac{\sin(\alpha/3)}{B} \qquad \sin\frac{\phi_{max}}{2} = \frac{\sin(\alpha/3 + 120^\circ)}{B} \tag{6.3}$$

where $B = \eta_0(3\mu/2E)^{1/2}$, $\sin\alpha = B$, $\mu = \mathrm{M_{Jupiter}/M_\odot} = 0.000955$ and η_0 = mean motion of Jupiter = 0.01341 rad yr^{-1}. These limits are set by E, which is a constant of the motion in the absence of Poynting Robertson drag:

$$E = -\frac{1}{6}\left(\frac{d\phi}{dt}\right)^2 - \frac{\mu\eta_0^2}{2x}(1+4x^3) \tag{6.4}$$

where $x = |\sin(\phi/2)|$. The libration amplitude, D, is $\phi_{max} - \phi_{min}$. We find that the energy constant is approximately

$$E \approx -\frac{3}{2}\mu\eta_0^2(1 - 0.0133D + 0.2266D^2 - 0.0392D^3) \qquad (6.5)$$

for $D \leq 1.3$.

The fraction of time a particle spends at a given phase, or equivalently, the distribution in phase of an ensemble of particles, is given by

$$P_\phi \propto \frac{1}{d\phi/dt}. \qquad (6.6)$$

We can evaluate this as a function of D with the aid of equations (4) and (5). For the distribution of dust libration amplitudes, P_D, we used a simple analytic function that approximates the distribution of libration amplitudes for Trojan asteroids shown in Figure 5 of Shoemaker et al. (1989). When we average P_ϕ over P_D, we find that the L5 dust cloud is distributed in orbital phase roughly as a Gaussian centered at $\theta_0 = 59.5°$ behind Jupiter with a dispersion $\sigma_\theta = 10°$.

The distribution of the dust in heliocentric latitude can be approximated in a similar way. If particle in an orbit of given inclination, i, with small eccentricity, spends a fraction of its time, f, at latitude, β, an ensemble of particles with small eccentricities and evenly distributed ascending nodes will have a distribution, at a fixed orbital phase, of

$$P_\beta \propto f \propto (\cos^2\beta - \cos^2 i)^{-1/2}. \qquad (6.7)$$

We take the inclination distribution of the particles, P_i, to be a simple analytic function that approximates the data for "independently discovered Trojans" shown in Figure 3 of Shoemaker et al. (1989). When we average P_β over P_i, we find the distribution in latitude is roughly a Gaussian with dispersion $\sigma_\beta = 10°$, and the distribution in height above the ecliptic has a dispersion $\sigma_z = 0.94$ AU.

The radial distribution of Trojans is more complicated to model, since both librations and epicycles include radial excursions. The average L5 Trojan eccentricity is 0.063; a particle with this eccentricity orbits at a range of heliocentric distances, $\Delta r \approx 0.66$ AU. In the course of its librations, a particle with a typical Trojan libration amplitude, $D = 29^\circ$, oscillates in semi-major axis over a range of $\Delta a \approx 0.14$ AU. We are not sensitive to the radial structure of the Trojan clouds, so we simply model the radial distribution as a Gaussian with a full width at half maximum of 0.66 AU, or a dispersion $\sigma_r = 0.24$ AU.

Our final model has the form:

$$n = n_0 \exp\left[-\frac{(r - r_0)^2}{2\sigma_r^2} - \frac{z^2}{2\sigma_z^2} - \frac{(\theta - \theta_0)^2}{2\sigma_\theta^2}\right] \tag{6.8}$$

where r, z, and θ are cylindrical coordinates in the plane of the orbit of Jupiter, and the parameters are: $r_0 = 5.203$ AU, $\sigma_r = 0.24$ AU, $\sigma_z = 0.94$ AU, $\theta_0 = -59.5^\circ$, and $\sigma_\theta = 9.7^\circ$. We calculated the surface brightness in the same way as we calculated the surface brightness of the model Mars wake, using an emissivity $E_{60\mu\mathrm{m}} = 1$ because we are not considering small grains. The shaded region at L5 in Figure 6.1 represents this model as viewed from above the ecliptic plane.

In Figure 6.7, we compare the difference image A$-$B to a synthesized image of our model cloud. For this image, n_0 is $3.4 \times 10^{-8}\mathrm{AU}^{-1}$, corresponding to an effective emitting surface area at 60 microns of 3.3×10^{18} cm^2, or one 3-km diameter asteroid ground entirely into 10-micron diameter dust. Figure 6.8 compares the difference image A$-$B and the model image in a different way; it shows the region within $\pm 10^\circ$ of the ecliptic plane averaged in ecliptic latitude. The 1-σ noise in the data in Figure 6.8 is 0.09 MJy ster^{-1}. Based on this, we can place a rough 3-σ upper limit on the effective surface area of the large dust grains at L5 of $\sim 6 \times 10^{17}\mathrm{cm}^2$.

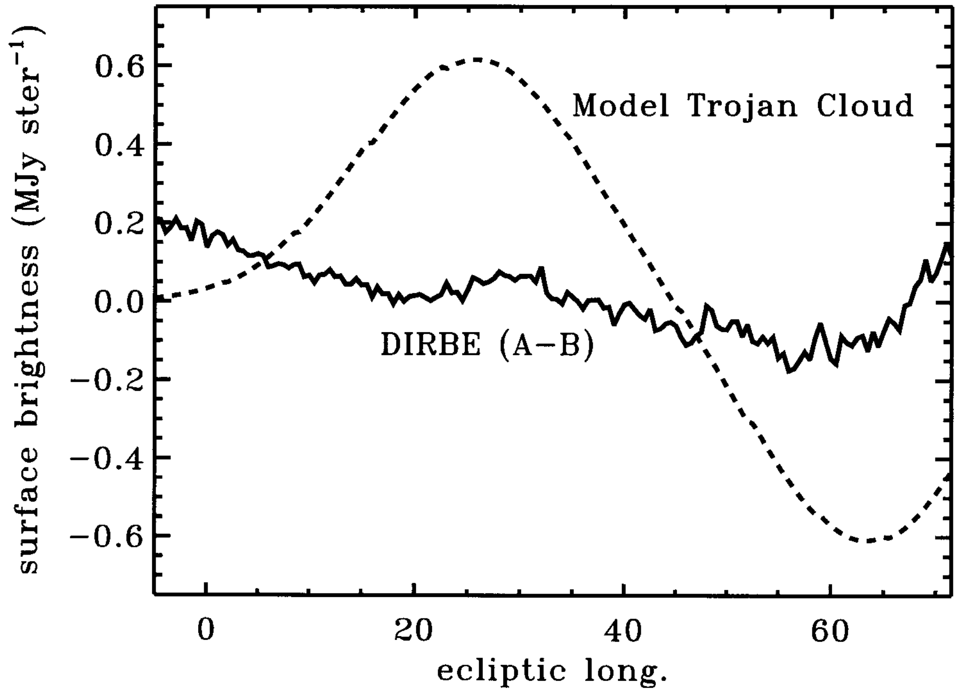

Figure 6.8: The difference A−B compared to the model for the L5 cloud. This plot shows a region of the 60-micron maps from Figure 6.7 within ±10° of the ecliptic plane that has been averaged in latitude. Based on this comparison, we place a 3-σ upper limit on the surface area of the L5 cloud of $6 \times 10^{17}\text{cm}^2$.

6.5 Conclusions

The zodiacal cloud near the ecliptic plane is a complex tapestry of dynamical phenomena. We could not detect the Mars wake or Jupiter's Trojan clouds among the asteroidal dust bands in the DIRBE maps, despite the efforts of the DIRBE team to subtract these bands from the maps. We would have detected the Mars wake if it had 18% of the overdensity of the Earth wake, based on our empirical model for the Earth wake. This upper limit illustrates the complexity of relating resonant structures in circumstellar dust disks to the properties of perturbing planets. For instance, we would have detected the Mars wake if the surface area of the dust in the wake scaled simply with the mass of the planet times the Poynting-Robertson time scale.

The Trojan clouds, by our crude estimation, would have been a few orders of magnitude too faint to detect if the dust concentration in these clouds were at its mean levels. However, a transient cloud created by a recent collision of Trojan asteroids might have been detectable. We measured that the total 60-micron flux from large (10–100 micron diameter) dust particles trapped at Jupiter's L5 Lagrange point is less than ~ 30 kJy.

We thank Antonin Bouchez, Eric Gaidos, Peter Goldreich, Renu Malhotra and Ingrid Mann for helpful discussions, and our referees for their thoughtful comments.

Chapter 7

A Search for Resonantly Trapped Kuiper Belt Dust

7.1 Introduction

[1]Dust released in the Kuiper Belt (KB dust) has never been directly observed. The major obstacle to detecting this potentially vast cloud is that the Kuiper Belt must be viewed through the local zodiacal cloud, which is much hotter than dust beyond the orbit of Neptune could be. Kelsall et al. (1998) subtracted a model for the local zodiacal cloud from the DIRBE maps, and some upper limits have been computed based on the residuals (Backman et al. 1995; Teplitz et al. 1999). However, those upper limits all share a common struggle; it is difficult to say how much emission from KB dust was subsumed by the zodiacal cloud model. Vega-excess dust—analogous dust around other stars—has proven easier to observe than KB dust in our own solar system!

We explored a new way of detecting dust in the Kuiper Belt that is largely independent of models of the local zodiacal cloud. The outer planets probably trap KB dust in mean motion resonances, just as the Earth traps asteroidal dust. Trapped dust may create azimuthal asymmetries in the KB dust cloud which could stand out from the backgrounds in the ecliptic plane because they would appear to orbit the sun at the same rate as the planet that creates them—just as the Mars wake might have stood out by virtue of its apparent motion. Detecting the orbital motion of an outer planet requires observations taken over a long time baseline. To obtain this baseline, we compared DIRBE 60 micron maps with 60 micron maps of the whole sky made with data from the IRAS satellite.

---

[1] This work was done in collaboration with William Reach and Jer-Chyi Liou.

7.2 IRAS and DIRBE

IRAS surveyed 98% of the sky from low-earth orbit from January to November, 1983, in four bands, with effective central wavelengths of 12, 25, 60 and 100 microns. The mission was designed to create a catalog of point sources, but the detectors were found to be stable enough to accurately map extended sources. Subsequent to the release of the IRAS Point Source Catalog, the Infrared Processing and Analysis Center (IPAC) reanalyzed the data to create evenly calibrated maps of the large scale structure of the whole surveyed portion of the infrared sky. They fit a 13-parameter model of the zodiacal cloud to the recalibrated data and subtracted the model from the data, to create the IRAS Sky Survey atlas (ISSA), described in detail in The Iras Sky Survey Atlas Explanatory Supplement (Wheelock et al. 1994). The ISSA first revealed the asteroidal dust bands and the Earth's ring and wake; no attempt was made to remove these structures from the maps. The portion of the IRAS Sky Survey Atlas that contains the ecliptic plane is called the "ISSA Reject Set" because of the presence of the asteroid bands in the data.

We compared the ISSA Reject Set to the DIRBE Sky and Zodi Atlas (DZSA), the same data set we used to search for the Mars wake and Jupiter's Trojan clouds. DIRBE mapped the entire sky at large angular scales over a period from December 1989 to September 1990. During the interval between the DIRBE and IRAS observations, Neptune moved from an ecliptic longitude of 268° to an ecliptic longitude of 283°. Resonant structures are apt to be tens of degrees in size, so this baseline is just enough to make the search worthwhile; a longer time difference between surveys might improve the results of this sort of search by a factor of a few. The COBE DIRBE Explanatory supplement gives a recalibration of the ISSA that matches it to the DIRBE maps.

$$\mathrm{I(DIRBE)}_{60} = 0.87 \pm 0.05 \times \mathrm{I(IRAS)}_{60} + 0.13 \pm 0.65 \tag{7.1}$$

We applied this correction to the ISSA to allow us to directly compare it with the

DSZA.

Figure 7.1a shows the DIRBE map of the whole sky within 45° of the ecliptic plane, and Figure 7.1b shows the ISSA map of the same region. Figure 7.1c shows the difference of these two maps. The vertical grey strips in the IRAS map at $\sim 170^\circ$ and $\sim 350^\circ$ are regions of the sky not surveyed by IRAS. The grey triangles at the top and bottom borders of the DIRBE map are the edges of the COBE data cube; these regions were surveyed by DIRBE, but they were not included in this image for convenience. Neptune itself is not visible in the two maps, but black trident symbols indicate its position.

7.3 Models of Trapped KB Dust

The first perturber dust encounters as it spirals in from the Kuiper Belt is Neptune. Liou and Zook (1999; LZ) numerically integrated the orbits of dust particles started on orbits at 45 AU with eccentricities of 0.1 and inclinations of 10°. These initial orbits correspond to the region called the classical Kuiper Belt. LZ find that many particles become trapped in resonances with Neptune, and the result is a ring with a gap, like the Earth ring and like Figure 2.2. When particles leave these resonances, they are either scattered from the solar system, or they typically have much higher eccentricities and inclinations, and can't easily be trapped again. Most of the particles that do remain in the solar system are subsequently scattered out of the solar system by Saturn. The result is that in the LZ model, the Neptune ring is the most visible azimuthal structure in the outer solar system.

Holmes et al. (1999) have another concept of resonant structure in the Kuiper Belt. They find in unpublished simulations that dust produced by Plutinos, Kuiper Belt Objects locked in Neptune's 2:3 resonance at 39 AU, often remains in the 2:3 resonance, or escapes the resonance, but is soon retrapped. The result is a population of dust in orbits similar to the Plutinos themselves, a cloud with a very different azimuthal structure than the LZ model.

We compared the satellite observations to both of these dust cloud models. LZ integrated orbits for four different groups of dust particles. We chose the group with the largest value of β, $\beta = 0.05$. For spherical particles with density 2 g cm^{-3} this corresponds to a radius of 5.7 microns. The simulation contained 100 particles, half released at $a = 45$ AU, half released at $a = 50$ AU. LZ converted the output of the orbit integrations into a model of the cloud density by accumulating the locations of the particles in 1 AU × 1 AU × 1 AU bins at regular intervals to. Figures 4a and 5a of LZ (1999) show images of these clouds viewed at a distance face-on.

We set the overall density of the model using measurements from Pioneer 10 Meteoroid Experiment. Pioneer 10 was sensitive to particles as small as 10^{-9} g. Since smaller particles far outnumber large particles in a collisional size distribution, most of the particles Pioneer 10 detected are probably not much more massive than that detection threshold, which corresponds to spherical particles with radii 5 microns with densities of 2 g cm^{-3}, similar to the $\beta = 0.05$ particles used in the simulation. Pioneer 10 recorded about 20 penetrations in the region from 6 to 18 AU from the sun (Humes 1980). The penetrations occurred at a roughly constant rate. If we divide this measured penetration flux of 3×10^{-6} penetrations m^{-2} s^{-1} by the circular velocity at 14 AU, 8000 m s^{-1}, we get an estimate of the density at this radius of 4×10^{-10} particles m^{-3}. We scaled the model density at 14 AU to match this density. Naturally, a proper analysis would take into account the effects of the angles of incidence of the particles, the velocity of the spacecraft, and so on. This estimate of the particle density is only meant to be used as a reference point. The typical brightness of the scaled model in the ecliptic plane is 0.6 MJy ster^{-1} at 60 microns.

For the plutino dust, we made a simple Gaussian model for the density. We assumed that the interesting aspects of the cloud are the two blobs that would appear at the perihelia of the orbits, leading and trailing Neptune by 90°. Any azimuthally symmetric aspect of the cloud would be cancelled in a difference of images taken at two different epochs. We modeled the blob densities as

$$n = n_0 \exp\left[-\frac{(r-r_0)^2}{2\sigma_r^2} - \frac{z^2}{2\sigma_z^2} - \frac{(\theta \pm \pi/2)^2}{2\sigma_\theta^2}\right] \tag{7.2}$$

where n is the local average of particle number density times particle cross section ($< n\sigma >$), and r,z, and θ are cylindrical coordinates in the plane of the orbit of Neptune centered on the sun, where Neptune is at $\theta = 0$. We started with this concept of a typical plutino orbit: semimajor axis 39.0 AU, inclination 10°, eccentricity 0.25, libration amplitude of 23° (Pluto's). We translated these values into the dust cloud parameters as follows:

r_0 = semimajor axis × (1 - eccentricity/2)

σ_r = semimajor axis × eccentricity/2

σ_z = semimajor axis × inclination

σ_θ = libration amplitude

We set the central $< n\sigma >$ of the plutino blobs to $n_0 = 10^{-7}$ AU^{-1}, similar to the value at $r = 1$ AU in the zodiacal cloud, where $< n\sigma >= 1.13 \times 10^{-7}$ AU^{-1} in the Kelsall et al. (1998) model. The total cross section is $\approx n_0 \sigma_r \sigma_z r_0 \sigma_\theta = 4.5 \times 10^{-5}$ AU^2 $= 10^{22}$ cm^2. This might correspond to a ring of plutino dust with about 5 times that total cross section. The maximum brightness of this model is 1.54 MJy ster^{-1} at 60 microns.

To compare the models with the data, we synthesized images of how the models would appear as viewed from the sun using this relation:

$$I_\lambda = \int n(r, z, \theta) B_\lambda(T) ds \tag{7.3}$$

where $B_\lambda(T)$ is the Planck function, and $T = 278$ K $(r/1\ \mathrm{AU})^{0.5}$, the local Blackbody temperature. We created images of how the clouds would appear during the IRAS mission, then images of how they would appear during the COBE mission, and differenced the images, just as we differenced the observed maps. Figure 7.1d shows the time-differenced LZ model, and Figure 7.1e shows the time-differenced plutino dust model. The LZ model image has some graininess due to the finite number of particles in the simulation.

The primary residuals in the DIRBE - IRAS image are in the Galactic plane, caused mostly by non-linearity and hysteresis in the IRAS detectors. Unfortunately, because Neptune was near the galactic plane during both the IRAS and COBE missions, most of the signal in the differenced LZ model is near the galactic plane, where there is the most noise. However, because Plutinos are in the 2:3 resonance with Neptune, the perihelia of their orbits lead or trail Neptune by 90°, far from the Galactic plane during the IRAS and DRIBE observations. In the LZ simulations, some dust populates this resonance, but dust populates other Neptune mean-motion resonances too, the 3:4, the 4:5, the 5:6 and so on, so the 2:3 resonance does not nearly dominate the appearance of the cloud. Instead, in the LZ model, the gap near Neptune dominates the azimuthal structure. The net result is that our search is more sensitive to plutino dust. The plutino dust model is also less azimuthally symmetric than the LZ model.

A secondary noise source is the yearly variation of the zodiacal background. IRAS and DIRBE did not map the entire sky in one day; they mapped only certain longitudes at a time. During the year, the Earth moves with respect to the warped asteroidal dust bands and with respect to the Earth's ring and wake, so the signal from these sources varies. When the whole sky maps are summed, this time variation translates into vertical stripes which appear in the IRAS and DIRBE images of Figure 7.1. They are more prominent in the IRAS map because the DIRBE zodiacal dust model corrects for much of the flux from the dust bands and from the Earth ring and wake. This effect was the limiting noise source in our search for the Mars Wake.

The model for the zodiacal cloud that has been subtracted from the DSZA is different from the model that has been subtracted from the ISSA. The DIRBE model is the 88-parameter Kelsall et al. (1998) model, which incorporates three pairs of asteroidal bands and the Earth's ring and wake. The ISSA model does not account for any of those structures. The result is that the ISSA map in Figure 7.1b has severe asteroidal dust band residuals, the horizontal stripes near the ecliptic plane, and subtracting the DIRBE map does not remove these residuals.

However, we are interested only in azimuthal structure, so we can reduce the effects of those residuals by performing an average in ecliptic latitude. Figure 7.2 shows another view of the differenced satellite images and the differenced model images, panels c, d, and e of Figure 7.1. It shows the region within ±23° of the ecliptic, averaged in latitude. We masked out the region within ±12° of the galactic plane before performing the averages to make Figure 7.2. Still the noise that remains in the latitudinal average of the DIRBE-IRAS map is primarily residuals from subtraction of the galaxy. In this figure, the LZ model has been multiplied by 10, and the plutino model has been scaled by a factor of 0.45, so $n_0 = 4.5 \times 10^{-8}$. We do not see any sign of resonantly trapped Kuiper belt dust in the IRAS and DIRBE images. But Figure 7.2 shows that these scalings are upper limits on the amount of dust that could exist in a ring like the LZ model or the plutino dust model. We can translate these upper limits into limits on the typical ecliptic plane brightnesses for the two cloud models; they correspond to brightnesses of 6 MJy ster^{-1} for the LZ model, and 0.7 MJy ster^{-1} for the Plutino model.

7.4 Discussion

How do our upper limits on the dust in the outer solar system compare to other known clouds of dust? If the outer solar system hosted a dust cloud like the dust clouds we have observed around other stars, could we have detected it? Could we detect our Kuiper Belt if it orbited a nearby G star instead of the sun? Table 7.1 offers some answers to these questions by comparing our upper limit models to the dust clouds around Vega and Epsilon Eridani and to the solar zodiacal cloud using two measures for each cloud: the total dust surface area, and the infrared excess. The infrared excesses for Vega and Epsilon Eridani, the surface area of Vega's dust cloud and much more information can be found in similar tables comparing various Vega-excess stars which appear in Backman and Gillett (1987), Backman and Paresce (1993), Backman et al. (1997), and Dent et al. (2000). We derived the numbers we quote for the solar zodiacal cloud from the model by Kelsall et al. (1998).

Table 7.1: Comparison of dust clouds

Dust Cloud	60-micron excess	dust surface area (cm^{-2})
Solar Zodiacal Cloud	0.04%	4×10^{20}
Plutino Dust Model	$< 3.5\%$	$< 2 \times 10^{22}$
LZ Model	$< 20\%$	$< 10^{23}$
ϵ Eridani	400%	5×10^{25}
Vega	750%	10^{27}

The total effective surface area of the dust is a good measure to use for comparing two clouds around different stars since it can be inferred from the images of the cloud taken at a few different wavelengths by making some assumptions about the dust radiative properties. It might be preferable to compare the total mass of the dust, for instance, but inferred masses are much more sensitive to the size of the dust grains, so they are poorly constrained by observations. The surface areas quoted in Table 7.1 assume that all the dust is at blackbody temperature. This approximation is very good for zodiacal dust and in general for large grains—larger than the wavelength range where blackbody dust would emit most of its energy. The effective emitting area of a cloud generally decreases with wavelength, however, so the surface area we quote for Epsilon Eridani, which Greaves et al. 1998 derived from images at 850 microns, is probably lower than the effective surface area of the same dust cloud at 60 microns. For the plutino dust model, we assume that total cross section area of the trapped dust ring is five times the total cross section of the model blobs.

Table 7.1 shows that the surface area in our upper-limit models is midway between the surface area of the zodiacal cloud and the Vega-excess disks in a logarithmic sense. Our upper limits seem so poor compared to the zodiacal cloud because the zodiacal dust near the earth is roughly seven times as hot as blackbody dust at the heliocentric distance of pluto. However, if the sun were surrounded by as much dust as Epsilon Eridani or Vega, the fraction of that dust that became trapped in resonances with Neptune would probably be quite easy to detect by our method.

The infrared excess of a dust disk—the ratio of the total flux emitted by the dust to the photospheric emission of their central star—is a way to compare two different dust clouds that has direct bearing on their detectability in photometric surveys. This number reflects the luminosity of the star, the dust orbits, and the distance of the dust from the star as well as the dust emitting area. The typical measurement error in the IRAS photometry at 60 microns for bright main sequence stars is about 11%. So according to our upper limits, IRAS might have been able to detect our LZ-type upper-limit-model if it orbited a very nearby solar-type star, but not our upper-limit plutino dust model. New telescopes like the planned Space Infrared Telescope Facility (SIRTF) can make more accurate and more sensitive measurements than IRAS. In this case photometric detection of circumstellar dust may be limited by the accuracy of models of stellar spectra in the mid-infrared, which in the best case is about 2% (Cohen et al. 1996). So according to our upper limits SIRTF might be capable of detecting the dust in our outer solar sytem photometrically if it viewed it from a distance of several parsecs at 60 microns no matter which model for the dust configuration is better.

Of course, our upper limits are on the total amount of dust trapped in resonant orbits, not the total amount of outer-solar system dust. There might be much more dust in the Kuiper Belt that would not appear in a difference of DIRBE and IRAS images because of its azimuthal symmetry. If that were the case, the solar system dust cloud would be easier to detect as a photometric excess when viewed from a nearby star than our upper limits suggest.

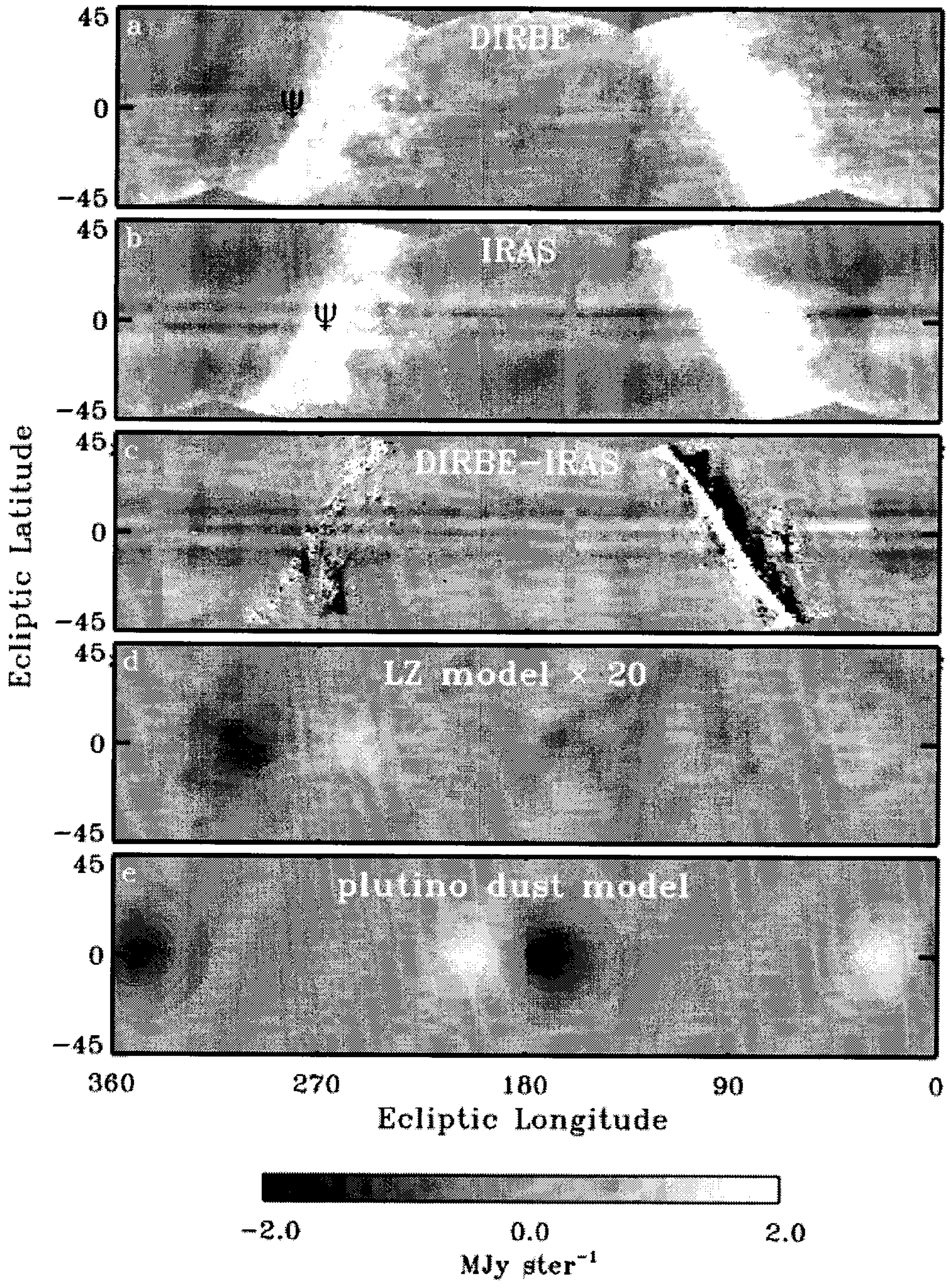

Figure 7.1: Searching for the Neptune Ring By Comparing DIRBE and IRAS 60 micron maps. The trident symbol shows the location of Neptune during the IRAS and COBE missions. a) DIRBE image minus DIRBE zodiacal cloud model. b) IRAS image minus IRAS zodiacal cloud model. c) image a - image b. d) how c would appear if the dust distribution matched the Liou and Zook model for the neptune ring. e) another model where all the dust is in the 2:3 resonance with Neptune.

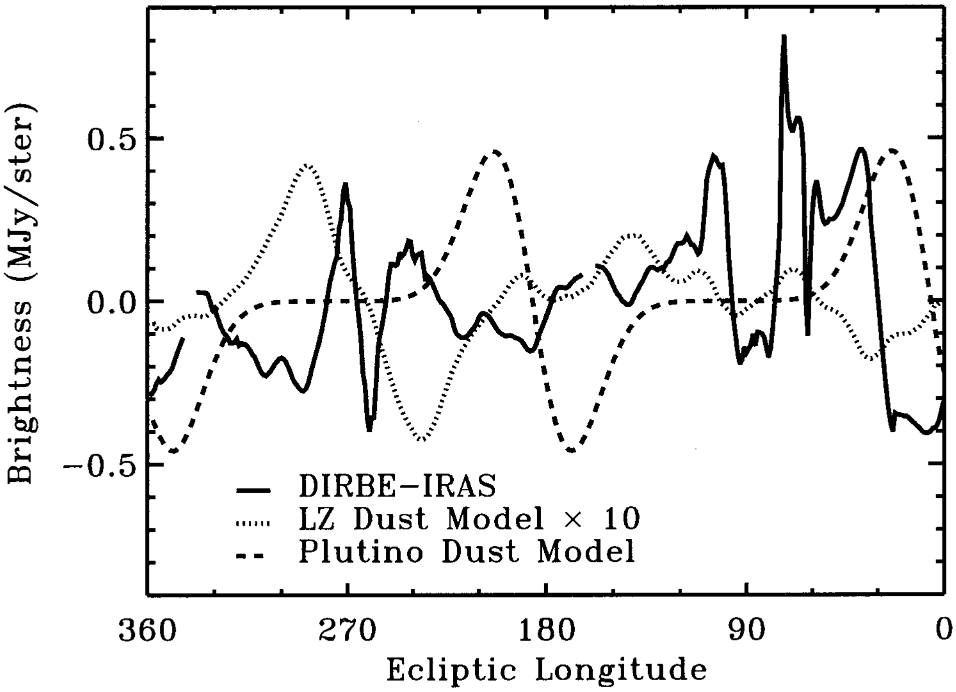

Figure 7.2: Cuts through panels c, d, and e in Figure 7.2.

Chapter 8

Conclusion

8.1 Detecting Exozodiacal Dust

The nature of the Poynting-Robertson force dictates that the equilibrium configuration of a collisionless dust cloud interior to the dust sources approaches a constant surface-density fan shape; this model describes our zodiacal cloud well. Viewed at thermal wavelengths, the brightness distribution of a cloud like this falls off roughly exponentially with separation from the star. Unfortunately, this brightness distribution is difficult to distinguish from the typical point-spread-function of a real telescope with a circular aperture. This problem limits the power of direct imaging with a single-dish telescope or imaging with the aid of a coronagraph to detect exozodiacal clouds. Table 8.1 summarizes the upper limits from our single-dish surveys.

Table 8.1: Summary of upper limits on exozodiacal clouds

Star	Spectral Type	Telescope/Instrument	Log Disk Density Solar Disk = 0
Vega	A0V	Keck/LWS	≤ 4.0
Altair	A7V	Keck/LWS	≤ 3.2
Sirius	A1V	HST/NICMOS Coronagraph	≤ 5.1
Procyon	F5IV	HST/NICMOS Coronagraph	≤ 5.5

Nulling interferometers are fundamentally better suited to detecting exozodiacal dust. The soon-to-be completed Keck Interferometer promises the ability to detect exozodiacal clouds with only 10 times the optical depth of the solar zodiacal cloud, even though the nulled beam will transmit starlight comparable to the signal from a 30 × solar cloud. But even a nulling interferometer as powerful as the soon-to-be completed Keck Interferometer can only be used to constrain models of exozodiacal clouds.

Ultimately we will want to combine data from a few nulling interferometers, such as the Keck Interferometer and the Large Binocular Telescope, to provide the (u,v)-plane coverage needed to decide the orientation and radial-structure of exozodiacal clouds.

8.2 Signatures Of Planets

One might hope that a good understanding of how planets trap dust in resonances could be used to decode maps of clumpy circumstellar clouds to deduce the orbits and masses of the perturbing planets. My experiences studying the detailed structure of our zodiacal cloud suggest that this technique may often be impractical for the following reasons:

1) The eccentricities of the dust orbits, the eccentricity of the planet's orbit, and the β's of the dust grains all affect the dynamics. These free parameters burden models of the effect of a planet on a dust cloud.

2) Not all planets massive enough to form rings and wakes do. In models of dust in the outer solar system by Liou and Zook (1999), Neptune traps many dust particles, but Uranus doesn't. Resonances of Neptune overlap with the resonances of Uranus and destabilize them, and many of the best particles for trapping, those with the right orbital eccentricity and β, get trapped by Neptune and kicked into orbits that makes them harder for Uranus to trap. Similar effects may occur in the inner solar system, where the Earth's resonances spoil the effects of Venus's resonances. These phenomena have yet to be confirmed observationally, but this effect may allow planets as massive as Uranus to hide in dust clouds, with no ring of trapped dust to signal their presence.

3) The debris from a single asteroid or comet can remain localized in dynamical space and produce a vivid structure in a zodiacal cloud. I searched for dust trapped in resonances with Mars, and I found that any signal from structures associated with Mars was swamped by signal from the asteroidal dust bands. Each of these bands probably comes from a family of asteroids which all share a single common progenitor

asteroid.

On the other hand, there are a few other signatures of planets that have not yet received much attention that we should watch out for in exozodiacal clouds.

1) The secular perturbations of Jupiter-mass planets may be important sculptors of exozodiacal clouds. The pericenter shift of the zodiacal dust orbits induced by Jupiter is about 0.01 AU. At a distance of 10 parsecs, this corresponds to an angle of 1 milliarcsecond—not much less than the resolution of a 100-meter baseline interferometer in the near and mid-infrared. Moreover, in the limit that the dust orbits near the circumstellar distance of the planet, the pericenter shift is proportional to the eccentricity of the planet times its semimajor axis, and the extra-solar giant planets that appear in radial velocity surveys often have orbital eccentricities much higher than Jupiter's.

2) Massive extra-solar planets may engender clouds of trojan asteroids that could release exozodiacal clouds with dramatic azimuthal structures.

3) Finally, the asteroid belt itself may have formed where it did because of the presence of strong mean motion and secular resonances with Jupiter inhibited the formation of a terrestrial planet near 3 AU. In this sense, the outer cutoff of the zodiacal cloud at the asteroid belt is a signature of the presence of Jupiter. If we can measure the radial structure of exozodiacal clouds to look for similar outer edges, perhaps by combining data from the Keck Interferometer and The Large Binocular Telescope, we might use this information to infer the presence of massive planets with orbital periods too long or masses too small for radial-velocity techniques to detect.

8.3 The Future

A zodiacal cloud is a potential sign of fertile planetary system. A massive zodiacal cloud could indicate a planetary system undergoing a period of heavy bombardment akin to the time when comets may have delivered water to the Earth's oceans (Chyba 1987; 1990). A less massive cloud as smooth as the solar cloud indicates a region of dynamical stability where a terrestrial planet might survive the dynamical perturbations of massive planets long enough for biogenesis to occur. Even though zodiacal dust can easily outshine planets, making them harder to detect, stars with evolved zodiacal clouds may be the most likely stars to harbor extraterrestrial life.

But beyond the quest for extra-solar terrestrial planets and extraterrestrial life, the science of exozodiacal dust has a place at the center of the comparative study of planetary systems. When we find zodiacal dust around another star, we reason by analogy that the dust comes from either extra-solar comets or asteroids. One might say that, in general, an asteroid is a chunk of rock that releases dust when it is struck by another asteroid, and a comet is a clod of rocks and dirt soaked in volatiles that releases dust when the volatiles start to sublimate during a close approach to a star. The best reason to study exozodiacal dust may be to investigate the roles of comets and asteroids in planetary systems.

Massive planets can create belts of small bodies by sweeping them into resonances, as Neptune probably swept the plutinos into its 3:2 and 2:1 resonances (Malhotra 1995) as it migrated outward. It is not understood yet what process halts the migration of massive planets, why some giant planets fall into the stars they orbit or become hot Jupiters and others stop their inward spirals at 5 AU like Jupiter. But the parking of massive planet $\sim$ 5 AU from a star may be a prerequisite for terrestrial planet formation. Cataloging belts of small bodies around other stars—particularly belts at terrestrial temperatures—could help unravel the mysteries of planet migration.

Comets must originate where it is cold, far from a star, and somehow get injected into the terrestrial-temperature zone to release dust. We might imagine that all planetary systems have reservoirs of cold comets, Oort clouds and Kuiper Belts, the residue of planet formation. But do all planetary systems inject comets into their habitable zones? Are comets everywhere important for delivering volatiles to terrestrial planets?

Observations of exozodiacal clouds could begin to answer these questions.

One possible way to distinguish cometary dust from asteroidal dust is spectroscopy. The dust around β Pictoris shows a strong silicate emission feature at 9-11 microns similar to the silicate feature seen in spectra of comets (Telesco and Knacke 1991; Knacke et al. 1993). Exozodiacal clouds may reveal similar features. The radial structure of zodiacal clouds is another clue to their origin. By combining data from a range of baselines, using the Keck Interferometer, the Large Binocular Telescope, the Very Large Telescope Interferometer, and the Terrestrial Planet Finder, we may be able to decide whether a cloud terminates abruptly at an asteroid belt, or has a surface-density that increases with proximity to a star, as the sublimation rate of comets would. We can also compare observations of exozodiacal clouds with observations of Vega-excess dust at longer wavelengths made with the Space Infrared Telescope Facility (SIRTF) and submillimeter interferometers like the Submillimeter Array (SMA) and the Atacama Large Millimeter Array (ALMA) to investigate the transport of dust from the outer reaches of planetary systems to the inner regions.

Bibliography

Angel, R. and Woolf, N. 1997, in *Planets Beyond the Solar System and the Next Generation of Space Missions* (Baltimore: ASP Conference Series, D. Soderblom ed.)

Aumann, H. H. 1984, ApJ, 278, L23

Backman, D. E., Dasgupta, A. and Stencel, R. E 1995, ApJ, 450, L35

Backman, D. E., Fajardo-Acosta, S. B., Stencel, R. E. and Stauffer, J. R. 1997, *Ap & SS*, 255, 91

Backman, D. E. and Gillett, F. C. 1987, in *Proceedings of the fifth Cambridge workshop on cool stars, stellar systems, and the sun* (Berlin: Springer-Verlag, J. L. Linsky and R. E. Stencel eds.), p. 340

Backman, D. E., Gillette, F. C. and Low, F. J. 1986, *Adv. Space Res.*, 6, 43

Backman, D. E. and Paresce, F. 1993, in *Protostars and Planets III* (Tucson: Univ of Arizona Press, E. H. Levy and J. I. Lunine, eds.)

Baudoz, P., Rabbia, Y., Gay, J., Rossi, E., Petro, L., Casey, S. C., Bely, P. Y., Burg, R., MacKenty, J. W., Fleury, B. and Madec, P.-Y. 1998, *Proc. SPIE*, 3353, 455

Beaugé, C. and Ferraz-Mello, S. 1994, *Icarus*, 110, 239

Beichman, C. 1996, *A Road Map for the Exploration of Neighboring Planetary Systems (ExNPS)* (JPL Publication 96-22, C. A. Beichman ed.)

Beichman, C. A. 1998, in *Exozodiacal Dust Workshop Conference Proceedings* (NASA/CP-1998-10155, D. E. Backman, L. J. Caroff, S. A. Sandford and D. H. Wooden, eds.)

Beichman, C. A., Woolf, N. J. and Lindensmith, C. A. eds. 1999, *The Terrestrial Planet Finder* (Pasadena: JPL Publication 99-3)

Benest, D. 1989, A&A, 223, 361

Benest, D. and Duvent, J. L. 1995, A&A, 299, 621

Berg, O. E. and Grün, E. 1973, in *Space Research XIII* (Berlin: Akademie-Verlag)

Boggess, N. W., Mather, J. C., Weiss, R., Bennett, C. L., Cheng, E. S., Dwek, E., Gulkis, S., Hauser, M. G., Janssen, M. A., Kelsall, T., Meyer, S. S., Moseley, S. H., Murdock, T. L., Shafer, R. A., Silverberg, R. F., Smoot, G. F., Wilkinson, D. T., Wright, E. L. 1992, ApJ, 379, 420

Bonnet-Bidaud, J. M. and Gry, C. 1991, A&A, 252, 193

Bracewell, R. N. 1986, *The Fourier Transform and its Applications* (McGraw-Hill: USA)

Brecher, K. 1979, in *Astronomy of the ancients* (Cambridge: MIT Press, Brecher, Feirtag eds.), p. 91

Briggs, R. E. 1962, AJ, 67, 710

Briotta, D. A. Jr. 1976, Ph.D. Thesis, Cornell Univ.

Brouwer, D. and Clemence, G. M. 1961, *Methods of Celestial Mechanics* (New York: Academic Press)

Chyba, C. F. 1987, *Nature*, 330, 632

Chyba, C. F. 1990, *Nature*, 343, 129

Cohen, M., Witteborn, F. C., Carbon, D. F., Davies, J. K., Wooden, D. H. and Bergman, J. D. 1996, AJ, 112, 2274

Colavita, M. M. et al. 1998, *Proc. SPIE*, 3350, 776

Colavita, M. M. and Wizinowich, P. 2000, *Proc. SPIE*, 4006, 310

Danner, R. and Unwin, S. eds. 1999, *Space Interferometry Mission: Taking the Measure of the Universe* (Pasadena: JPL 400-811)

Dent, W. R. F., Walker, H. J., Holland, W. S. and Greaves, J. S. 2000, MNRAS, 314, 702

Dermott, S. F., Gomes, R. S., Durda, D. D., Gustafson, B. Å. S., Jayaraman, S. and Xu, Y. L. 1992, in *Chaos, Resonance and Collective Phenomena in the Solar System* (Dordrecht: Kluwer, S. Ferraz-Mello, ed.)

Dermott, S.F., Jayaraman, S., Xu, Y. L., Gustafson, B.Å.S. and Liou, J.-C. 1994, *Nature*, 369, 719

Dermott, S.F., Nicholson, P.D., Kim, Y., Wolven, B. and Tedesco E. F. 1988, in *Comets to Cosmology* (Berlin: Springer-Verlag, A. Lawrence, ed.)

Durda, D. D., Dermott, S. F. 1997, *Icarus*, 130,140

Egret, D., Wenger, M. and Dubois, P. 1991, in *Databases & On-line Data in Astronomy* (Kluwer Acad. Publ., Albrecht & Egret eds.), pp. 79–88.

Engelke, C. W. 1990, *LWIR Stellar Calibration: Infrared Spectral Curves for 30 Standard Stars* (Lincoln Lab. Project Rept. SDP-327.)

ESA, 1997, The Hipparcos and Tycho Catalogues, ESA SP-1200

Fajardo-Acosta, S. B., Beichman, C. A., Cutri, R. M. 2000, ApJ, 538, L155

Gatewood, G. D. and Gatewood, C. V. 1978, ApJ, 225, 191

Good, J. 1997, private communication

Greaves, J. S., Holland, W. S., Moriarty-Schieven, G., Jenness, T., Dent, W. R. F., Zuckerman, B., McCarthy, C., Webb, R. A., Butner, H. M., Gear, W. K. and Walker, H. J. 1998, ApJ, 506, L133

Grogan, K., Dermott, S. F. and Gustafson, B. Å. S. 1996, ApJ, 472, 812

Grün, E., Gustafson, B. Å. S., Mann, I., Baguhl, M., Morfill, G. E., Staubach, P., Taylor, A. and Zook, H. A. 1994, A&A, 286, 915

Gustafson, B. Å. S. 1994, *Ann. Rev. Earth Planet Sci.*, 22, 553

Hahn, J. M., Zook, H. A., Cooper, B. and Sunkara, B. 2000, "Mapping the Inner Zodiacal Light with Clementine," in *Lunar and Planetary Science XXXII*, Abstract #1690, Lunar and Planetary Institute, Houston (CD-ROM)

Hanbury Brown, R., Davis, J. and Allen, L. R. 1974, MNRAS, 167, 121

Hardy, J. W. 1998, *Adaptive Optics for Astronomical Telescopes* (Oxford, UK: Oxford Univ. Press)

Hauser, M. G., Kelsall, T., Lesiawitz, D. and Weiland J. eds. 2000, *COBE Diffuse Infrared Background Experiment (DIRBE) Explanatory Supplement* (Greenbelt, MD: NASA/GSFC, COBE Ref. Pub. No. 97-A)

Hinz, P.M., Angel J.R.P., Hoffmann W.F., McCarthy D.W., McGuire P.C., Cheselka, M., Hora J.L., Woolf N.J. 1998, Nature, 395, 251

Holmes, E. K., Dermott, S. F., Grogan, K. 1999, Bulletin of the American Astronomical Society, 195, 108

Hong, S. S. 1985, A&A, 146, 67

Humes, D. H. 1980, *Journal of Geophysical Research*, 85, 5841

Jackson, A. A. and Zook, H. A. 1989, *Nature*, 337, 629

Jones, B. and Peutter, R. 1993, *Proc. SPIE*, 1946, 610

Kelsall, T., Weiland, J. L., Franz, B. A., Reach, W. T., Arendt, R. G., Dwek, E., Freudenreich, H. T., Hauser, M. G., Moseley, S. H., Odegard, N. P., Silverberg, R. F. and Wright, E. L. 1998, ApJ, 508, 44

Kirkpatrick, J. D., Reid, I. N., Liebert, J., Cutri, R. M., Nelson, B., Beichman, C. A., Dahn, C. C., Monet, D. G., Gizis, J. E. and Skrutskie, M., F. 1999, ApJ, 519, 802

Koerner, D. W., Ressler, M. E., Werner, M. W., Backman, D. E. 1998, ApJ, 503, L83

Knacke, R. F., Fajardo-Acosta, S. B., Telesco, C. M., Hackwell, J. A., Lynch, D. K. and Russell, R. W. 1993, ApJ, 418, 440

Kuchner, M. J., Brown, M. E. and Koresko, C. D. 1998, PASP, 110, 1336

Kuchner, M. J. and Brown, M. E. 2000, PASP, 112, 827

Kuchner, M. J., Reach, W. T. and Brown, M. E. 2000, *Icarus*, 145, 44

Labeyrie, A. 1970, A&A, 6, 85

Lienert, C., Richter, I., Pitz, E. and Planck, B. 1981, A&A, 103, 177

Liou, J.-C., Dermott, S. F. and Xu, Y. L. 1995, *Planet. Space Sci.*, 43, 717

Liou, J.-C. and Zook, H. A. 1995, *Icarus*, 113, 403

Liou, J.C. and Zook, H.A. 1996, *Icarus*, 123, 491

Liou, J.-C. and Zook, H. A. 1997, *Icarus*, 128, 354

Liou, J.-C. and Zook, H. A. 1999, AJ, 118, 580

Liou, J.-C., Zook, H. A. and Dermott, S. F. 1996, *Icarus*, 124, 429

Lytle, D., Stobie, E., Ferro, A. and Barg, I. 1999, *ASP Conf. Ser. 172, Astronomical Data Analysis Software and Systems VIII* (San Francisco: ASP, D. Mehringer, R. Plante and D. Roberts eds.), p. 445

MacPhie, R. H. and Bracewell, R. N. 1979, Proc. SPIE, 172, 271

MacQueen, R. M. and Greeley, B. W. 1995, ApJ, 440, 361

Malbet, F. 1996, A&A, 115, 161

Malhotra, R. 1995, AJ, 110, 420

Mann, I. and Grün, E. 1995a, *Planet. Space Sci.* 43,6

Mann, I. and Grün E. 1995b, *Planet. Space Sci.*, 43, 827

Mann, I. and MacQueen, R. M. 1993, A&A, 275, 293

Marcy, G. W. and Butler, R. P. 1998, ARA&A, 36, 57

Marzari, F., Farinella, P., Davis, D. R., Scholl, H. and Campo Bagatin, A. 1997, *Icarus*, 125, 39

Matthews, K., Nakajima, T., Kulkarni, S. R. and Oppenheimer, B. R. 1996, AJ, 112, 1678

McCluskey, S. C. 1987, *Nature*, 325, 87

Moshir, M., et al. 1992, *Explanatory Supplement to the IRAS Faint Source Survey, Version 2* (Pasadena: JPL D-10015 8/92)

Mouillet, D., Larwood, J. D., Papaloizou, J. C. B. and Lagrange, A. M. 1997, MNRAS, 292, 896

Murray, C. 1994, *Icarus*, 112, 465

Murray, C. D. and Dermott, S. F. 1999, *Solar System Dynamics* (Cambridge, U.K: Cambridge University Press)

Murray, N. and Holman, M. 1997, AJ, 114, 1246

Nakajima, T., Oppenheimer, B. R., Kulkarni, S. R., Golimowski, D. A., Matthews, K. and Durrance, S. T. 1995, *Nature*, 378, 463

Noll, R. J. 1976, *JOSA*, 63, 1399

Ozernoy, L. M., Gorkavyi, N. N., Mather, J. C. and Taidakova, T. A. 2000, ApJ, 537, L147

Perryman, M. A. C., Lindegren L., Kovalevsky J., Hog E., Bastian U., Bernacca, P.L., Creze, M., Donati, F., Grenon, M., Grewing, M., Van Leeuwen, F., Van Der Marel, L. H., Mignard, F., Murray, C.A., Le Poole, R.S., Schrijver, H., Turon, C., Arenou, F., Froeschle, M. and Petersen, C.S. 1997, A&A, 323, L49

Provencal, J. L., Shipman, H. L., Wesemael, F., Bergeron, P., Bond, H. E., Liebert, J. and Sion, E. M. 1997, ApJ, 480, 777

Reach, W. T. 1991, ApJ, 369, 529

Reach, W. T., Abergel, A., Boulanger, F., Desert, F.-X., Perault, M., Bernard, J.-P., Blommaert, J., Cesarsky, C., Cesarsky, D., Metcalfe, L., Puget, J.-L., Sibille, F., Vigroux, L. 1996, /aap, 315, L381

Reach, W. T., Franz, B. A., Weiland, J. L. 1997, *Icarus*, 127, 461

Reach, W. T., Franz, B. A., Weiland, J. L., Hauser, M. G., Kelsall, T. N., Wright, E. L., Rawley, G., Stemwedel, S. W. and Spiesman, W. J. 1995, *Nature*, 374, 521

Rieke, M. 1999, in preparation

Robertson, H. P. 1937, MNRAS, 97, 423

Roddier, F. 1986, *Opt. Commun.* 60, 145

Roques, F., Scholl, H., Sicardy, B. and Smith, B. A, 1994, *Icarus*, 108, 37

Schlosser, W., Bergman, W. 1985, *Nature*, 318, 45

Schneider, G., Smith, B. A., Becklin, E. E., Koerner, D. W., Meier, R., Hines, D. C., Lowrance, P. J., Terrile, R. J., Thompson, R. I. and Rieke, M. 1999, ApJ, 513, L127

Schroeder, D. 1987, *Astronomical Optics* (San Diego: Academic Press)

Schuerman, D. W. 1980, ApJ, 238, 337

Serabyn, E., Colavita, M. M. and Beichman, C. A. 2000, Workshop on Thermal Emission, Spectroscopy and Analysis of Dust, Disks and Regoliths Lunar and Planetary Institute, (Houston: LPI, ASP Conference Series, M. L. Sitko, A. L. Sprague and D. K. Lynch, eds.), p.357

Serabyn, E. 2000, in preparation

Shoemaker, E. M., Shoemaker, C. S. and Wolfe, R. F. 1989, in *Asteroids II* (Tucson: Univ. of Arizona Press, R. P. Binzel, T. Gehrels and M. S. Matthews, eds.)

Sivaramakrishnan, A., Koresko, C. D., Makidon, R. B., Berkefeld, T., Kuchner, M. J., in preparation

Schroeder, D. J., Golimowski, D. A., Brukardt, R. A., Burrows, C. J., Caldwell, J. J., Fastie, W. G., Ford, H. C., Hesman, B., Kletskin, I., Krist, J. E., Royle, P. and Zubrowski, R. A. 2000, AJ, 119, 906

Spiesman, W. J., Hauser, M. G., Kelsall, T., Lisse, C. M., Moseley, S. H., Reach, W. T., Silverberg, R. F., Stemwedel, S. W. and Wieland, J. L. 1995, ApJ, 442, 662

Sykes, M. V. 1990, *Icarus*, 84, 267

Sykes, M. V., Greenberg, R. 1986, *Icarus*, 65, 51

Sykes, M. V., Greenberg, R. 1986, *Icarus*, 68, 186

Tang, T. B. 1986, *Nature*, 319, 532

Telesco C. M. and Knacke R. F. 1991, ApJ, 372, L29

Teplitz, V. L., Stern, S. A., Anderson, J. D., Rosenbaum, D., Scalise, R. J. and Wentzler, P. 1999, ApJ, 516, 425

Thompson, A. R., Moran, J. M. and Swenson, G. W. 1998, *Interferometry and Synthesis in Radio Astronomy* (Krieger Publishing Company, Florida)

Van Gent, R. H. 1984, *Nature*, 312, 302

Volet, C. 1932, *Bull. Astr. Paris*, 8, 52

Walbaum, M. and Duvent, J.L. 1983, *l'Astronomie*, 97, 277

Wang, T. and Vaughan, A. H. 1988, *Appl. Opt.*, 27, 27

Weidenschilling, S. J. and Jackson, A. A. 1993, *Icarus*, 104, 244

Wheelock, S. L. et al. 1994, *IRAS Sky Survey Atlas Explanatory Supplement* (Pasadena: JPL Publication 94-11)

Whipple, F. L. 1967, in *Zodiacal Light and the Interplanetary Medium* (NASA SP-150, J.L Weinberg, ed.)

Whittet, D. C. B. 1999, MNRAS, 310, 355

Wisdom, J. 1980, AJ, 85, 1122

Wyatt, M. C., Dermott, S. F., Telesco, C. M., Fisher, R. S., Grogan, K., Holmes, E. K., Pia, R. K. 1999, ApJ, 527, 918

Wyatt, S. P. and Whipple, F. L. 1950, ApJ, 111, 134

Yoder, C. F., Colombo, G., Synnott, S. P. and Yoder, K. A. 1983, *Icarus*, 53, 431