Circumstellar and Circumplanetary Disks

Thesis by

Eugene Chiang

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2000

(Submitted May 11, 2000)

Acknowledgments

Thanks go first to Peter Goldreich for being a process-oriented and holistic advisor. Chapters 2–4, 7, and 8 have all been blessed by his touch. It has been a privilege and joy to learn from someone I admire as an inimitable artist and fearless explorer of the first rank. I am extremely grateful for the care and advice he has generously imparted over the years. Whatever confidence (both on and off the squash court) I have gained in graduate school is due largely to him. Though I could fill many pages with entertaining and instructive stories about Peter, I believe that some of the more valuable lessons are best left carried in the heart. Here the only words which seem appropriate to write down are those once written by Cardinal Barberini in a letter to Galileo Galilei: "I pray the Lord God to preserve you, for men of great value like you deserve to live a long time to the benefit of the public." The same must be said of Peter Goldreich.

A big thanks goes to Mike Brown for helping me along in my first foray into optical observing. Chapter 6 is the fruit of our collaboration. I am grateful for his astuteness and coolness as a collaborator and as a friend.

Professionally and personally, I have benefitted from the support, encouragement, advice, and scientific know-how of Steve Beckwith, Geoff Blake, Harold Connolly, David Koerner, Doug Lin, Sterl Phinney, Anneila Sargent, and Frank Shu. Saul Rappaport deserves special mention as my principal research advisor during my undergraduate career. The lessons I learned from him have served me well in graduate school and will continue to do so henceforth.

It has been a unique pleasure to supervise Ryan Moo Kwang Joung on his undergraduate senior thesis. It was educational for us both. Chapter 5 is the result of our combined efforts.

I thank all TAPIR (Theoretical AstroPhysics Including Relativity) members past and present, for creating a happy and egalitarian atmosphere in which to work. Special thanks goes to Yoram Lithwick, for being a considerate officemate and serving reliably as the sober

voice of reason; Jason Maron, for introducing me to the concept of the "warrior spirit"; Glenn Soberman, for his infectious eclectic tastes in all things scientific; and Jeremy Heyl, who cured my "belly-aching" by reminding me that, in fact, "everyday is summer."

Chapter 7 concerning the dynamics of planetary rings is the piece of work for which I am most grateful to the good and subtle Lord. I acknowledge as a source of inspiration the book, *Longitude*, elegantly scripted by Dava Sobel, which details the construction of the first sea-worthy clock. I spent a lot of time in the dark before Chapter 7 materialized, and reading about the dramatic assembly of an exquisite mechanical contraption gave me hope that I was also working on a soluble problem.

I am grateful for my family, for their support and encouragement.

A final special thanks goes to Inn, for being fully present with unwavering support and strength when it counted. Though I like to say that BB makes it happen, in fact it is your love which makes it happen.

Abstract

This thesis studies disks in three astrophysical contexts: (1) protoplanetary disks; (2) the Edgeworth-Kuiper Belt; and (3) planetary rings.

We derive hydrostatic, radiative equilibrium models of passive protoplanetary disks surrounding T Tauri and Herbig Ae/Be stars. Each disk is encased by an optically thin layer of superheated dust grains. This layer is responsible for up to $\sim$70% of the disk luminosity at wavelengths between $\sim$5 and 60 μm. The heated disk flares and absorbs more stellar radiation at a given stellocentric distance than a flat disk would. Spectral energy distributions are computed and found to compare favorably with the observed flattish infrared excesses of several young stellar objects. Spectral features from dust grains in the superheated layer appear in emission if the disk is viewed nearly face-on.

We present the results of a pencil-beam survey of the Kuiper Belt using the Keck 10-m telescope. Two new objects are discovered. Data from all surveys are pooled to construct the luminosity function from $m_R = 20$ to 27. The cumulative number of objects per square degree, $\Sigma(< m_R)$, is such that $log_{10}\Sigma(< m_R) = 0.52(m_R - 23.5)$. The luminosity function is consistent with a power-law size distribution for which the smallest objects possess most of the surface area but the largest bodies contain most of the mass. To order-of-magnitude, $0.2M_\oplus$ and 1×10^{10} comet progenitors lie between 30 and 50 AU. The classical Kuiper Belt appears truncated at a distance of 50 AU.

We propose that rigid precession of narrow eccentric planetary rings surrounding Uranus and Saturn is maintained by a balance of forces due to ring self-gravity, planetary oblateness, and interparticle collisions. Collisional impulses play an especially dramatic role near ring edges. Pressure-induced accelerations are maximal near edges because there (1) velocity dispersions are enhanced by resonant satellite perturbations, and (2) the surface density declines steeply. Remarkably, collisional forces felt by material in the last $\sim$100 m of a $\sim$10 km wide ring can increase equilibrium masses up to a factor of $\sim$100. New ring surface densities are derived which accord with Voyager radio measurements.

Contents

List of Figures

List of Tables

Chapter 1

Introduction and Preview

It has been remarked that objects in astronomy divide roughly into two types: spheres and disks.[1] This thesis studies disks. We consider three astrophysical contexts: (1) protoplanetary disks and the radiation they emit; (2) the Kuiper Belt and an observational search for its faintest members; and (3) planetary rings and their dynamics. The variety of these subjects reflects our broad interest in the Solar System and extrasolar planetary systems, and their origins. In this chapter we provide introductory material on each of these topics to orient the reader; at the end of each section, we specify the problems whose solutions comprise this thesis.

1.1 Protoplanetary Disks

In the Solar System, the near co-planarity and circularity of planetary orbits argues that the planets coalesced from dust and gas within a disk—the "solar nebula"—surrounding the young Sun. The minimum mass of the solar nebula is $\sim 0.015 M_\odot$ (Weidenschilling 1977); it is obtained by augmenting the masses of the present-day planets up to solar composition.

Analogues of the early Sun may be found in T Tauri stars—optically visible, pre-main-sequence stars of which T Tauri in the Taurus molecular cloud is the prototype. They were discovered by Joy (1942, 1945, 1949) to be variable, Balmer emission-line stars of spectral type K/M situated near nebulosity. A comprehensive review of T Tauri stellar properties is supplied by Bertout (1989). Their optical luminosities ($L_* \sim 0.1$–$3\, L_\odot$) and effective temperatures ($T_* \sim 3000$–5000 K) situate them above the zero-age main sequence on a Hertzprung-Russell (H-R) diagram. Theoretical models of how stars descend along Hayashi convective tracks in this region of the H-R diagram indicate stellar masses of

[1] Attributed to Scott Tremaine, as part of the opening remarks of his acceptance speech for the 1983 American Astronomical Society Warner Prize (F. Shu, private communication).

$\sim$0.4–1$M_\odot$ and ages of $\sim 10^5$–10^7 yr (see, e.g., Tout, Livio, & Bonnell 1999, and references therein).

The higher mass ($\sim$1–2$M_\odot$) counterparts of T Tauri stars—the Herbig Ae/Be (HAeBe) stars—are also of interest, and are reviewed by Waters & Waelkens (1998).[2] For our purposes, the principal difference between T Tauri stars and HAeBe stars is that the latter are hotter ($T_* \sim$ 7000–10000 K) and consequently more luminous ($L_* \sim$ 3–30 $L_\odot$). For conciseness, unless otherwise noted, we restrict ourselves in this introduction to reviewing T Tauri systems, keeping in mind that our statements apply qualitatively to HAeBe stars as well.

Infrared fluxes from T Tauri stars substantially exceed those expected from stellar photospheres (Mendoza 1966, 1968; Rydgren et al. 1984; Cohen, Emerson, & Beichman 1989; Strom et al. 1989a,b; Kenyon et al. 1990; Sylvester et al. 1996). Mendoza (1966) first suggested that these infrared excesses represent thermal emission from circumstellar dust grains heated by stellar radiation. Spectral energy distributions (SEDs) between 12 and 100 μm of several hundred young stellar objects in the Taurus-Auriga and LYNDS 1641 clouds were first measured by IRAS (Infrared Astronomy Satellite) and compiled in the IRAS Point Source Catalog (Cohen et al. 1989; Strom et al. 1989a,b). In "classical" T Tauri systems (T Tauri stars having an equivalent width in Hα emission greater than 10 $\AA$), the infrared excess typically accounts for $\sim$50% of the bolometric luminosity (Cohen et al. 1989).

Millimeter wavelength excesses were detected in 53% of the classical T Tauri stars surveyed by Beckwith et al. (1990) using the Max-Planck-Institute für Radioastronomie Helium-cooled bolometer at the IRAM 30 m telescope, and the Owens Valley Radio Observatory. The fact that visual extinctions to these sources are modest ($A_V \lesssim 2$ mag) suggests that the dust grains responsible for the long wavelength excess are distributed in flattened circumstellar disks and not in spherical envelopes (Beckwith et al. 1990). This expectation is confirmed by Hubble Space Telescope images of the Orion star-forming cloud, in which circumstellar disks are clearly seen in silhouette against the background flourescent

[2]Herbig Be stars have masses $\sim 5M_\odot$, and temperatures and luminosities higher than those cited above; however, a Herbig Be star has yet to be discovered.

emission of the H II region (e.g., McCaughrean & O'Dell 1996).

The standard disk is modeled as a flat blackbody. In the simplest scenario it lacks intrinsic luminosity and passively reradiates the energy it absorbs from the central star. This blackbody model of a "passive disk" yields an infrared spectral energy distribution (SED) of the form $\nu F_\nu \propto \nu^n$, with spectral index $n = 4/3$ (Adams, Lada, & Shu 1987; and references therein). By coincidence, the steady-state accretional luminosity from an opaque Keplerian disk yields an SED with identical spectral index (Lynden-Bell & Pringle 1974). In contrast to these predictions, SEDs measured between 3 and 100 μm are usually well fitted by flatter power laws for which $0 \leq n \leq 4/3$, with most sources exhibiting $n \leq 3/4$ (see, e.g., Rydgren et al. 1984; Rucinski 1985; Rydgren & Zak 1987; Strom et al. 1989a,b; Beckwith et al. 1990).

The failure of the standard disk model spawned a number of alternative proposals. Kenyon & Hartmann (1987) considered a blackbody disk whose surface flares outward with increasing radius as a consequence of vertical hydrostatic equilibrium. Flared disks intercept more stellar radiation than flat ones, especially at large distances from the star. Other models invoke either an "active" disk having a high intrinsic luminosity, or a dusty component in addition to the disk. The dissipation of a one-armed spiral density wave as an intrinsic disk heating mechanism falls in the first category (Adams, Ruden, & Shu 1989; Shu et al. 1990; Ostriker, Shu, & Adams 1992). The second category includes the proposal that a tenuous dusty envelope surrounds the star/disk system and either scatters stellar radiation back onto the disk (Natta 1993), or absorbs and re-emits it in the infrared (Calvet et al. 1994).

This thesis studies the thermodynamics of passive disks. We argue that the observed flattish spectra are naturally reproduced by passive disks in hydrostatic and radiative equilibrium. Elementary considerations of radiative transfer demonstrate that the blackbody approximation for passive disks is a poor one; in fact, the entire disk is ensheathed in an optically thin layer of hot dust grains that radiates as a dilute blackbody. We refer to this layer as the "superheated surface." This outer layer is responsible not only for most of the continuum excess between $\sim$5 and 60 μm, but also for solid-state emission features from silicates and ices. The basic elements of our theory are set forth in Chapter 2.

Our work finds a close cousin in models by Calvet et al. (1991, 1992). We were notified of this connection after the publication of the first paper in our series (Chiang & Goldreich 1997). In their pioneering work, Calvet et al. (1991) numerically solve in detail for vertical temperature gradients in disks that are both externally illuminated by their central stars and internally heated at their midplanes by viscous dissipation. Their "temperature inversion" characterizing the irradiated disk atmosphere is equivalent to our "superheated surface"; furthermore, they also find that the $10\,\mu$m vibrational resonance in silicates can appear in emission (Calvet et al. 1992). While the simple, analytic, 2-layer disk model that we describe in this thesis is comparatively crude, we believe it better captures both the essence of the relevant physics and the significance of the irradiated atmosphere with regards to the overall SED. We forge ahead of the trail blazed by Calvet et al. (1991, 1992) in several directions; new developments include (1) computation of SEDs for arbitrary disk inclinations using a three-dimensional ray-tracing code, (2) elucidation of the microphysics of energy balance within passive disk atmospheres, (3) delineation of the numerous timescales over which passive disks respond to perturbations, (4) consideration of a variety of grain compositions and sizes, (5) detailed comparison of theory with observations for several individual T Tauri and Herbig Ae/Be systems, and (6) relaxation of the assumption of a constant flaring index. Point (1) is the subject of Chapter 3; points (2)–(3) comprise Chapter 4; and points (4)–(6) receive attention in Chapter 5.

1.2 The Edgeworth-Kuiper Belt

Edgeworth (1943) and Kuiper (1951) independently postulated the existence of a circumsolar disk of planetesimals extending beyond the orbit of Neptune. Each suspected that the surface density of primordial material at heliocentric distances greater than $\sim$30 AU was too low for the accretion of a major planet, but that debris from the formation of the Solar System could well remain at these distances. With remarkable prescience, Edgeworth (1943) also suggested that members of this belt might be perturbed into the inner Solar System and appear as short-period comets (having orbital periods shorter than 200 yr). Dynamical simulations by Fernandez (1980), Duncan, Quinn & Tremaine (1988), and Quinn,

Tremaine & Duncan (1990) quantitatively substantiate this speculation. These simulations demonstrate that the relatively low inclinations and short periods of Jupiter-family comets imply the existence of a reservoir of trans-Neptunian bodies in a flattened disk. This disk is now referred to most commonly as the Kuiper Belt.

The first Kuiper Belt Object (KBO) aside from Pluto was discovered by Jewitt & Luu (1993) in a dedicated optical search for such objects. The object was seen in reflected sunlight and identified by its parallax motion across the sky (of order a few arcseconds per hour). Since then, more than $\sim$200 KBOs have been discovered in searches using reflected light, and these objects have also had their orbital elements measured (see the Minor Planet Center database for trans-Neptunian objects at http://cfa-www.harvard.edu/iau/Ephemerides/Distant/index.html). As outlined by Jewitt, Luu, & Trujillo (1998, hereafter JLT98), KBOs divide into three dynamical classes.

1. Classical KBOs reside in low eccentricity, low inclination orbits beyond 40 AU (JLT98). They are not associated with mean motion resonances with Neptune.

2. Resonant KBOs, of which Pluto is the largest known member, have orbital periods commensurate with that of Neptune and are protected against close encounters with that planet (Malhotra 1996). They possess moderately high eccentricities and inclinations, possibly excited by Neptune during a transient period of orbital migration (JLT98; Malhotra 1995; Malhotra, Duncan, & Levison 1999, hereafter MDL99; and references therein).

3. Scattered KBOs, of which 1996 TL_{66} is one member (Luu et al. 1997), occupy large, highly eccentric and inclined orbits, the result of close encounters with Neptune (Duncan & Levison 1997).

At first glance, of the three populations, the classical Kuiper Belt appears the most untouched by Neptune. However, the record of primordial conditions preserved by the classical Belt, as observed today out to 50 AU, is probably heavily weathered. Recent surveys estimate the mass of the observable Kuiper Belt within 50 AU to be a few $\times$ $0.1 M_{\oplus}$ (e.g., Luu & Jewitt 1998; this thesis). Hamid, Marsden, & Whipple (1968) used

the trajectories of short-period comets to set an upper mass limit of $1.3M_{\oplus}$ on a smooth ring within 50 AU. The mass of the nearby Belt is therefore $\sim$10–100 $\times$ smaller than the $\sim 15M_{\oplus}$ extrapolated from the condensable material of the outer giant planets. Mass depletion since the time of formation is also suggested by the existence of bodies as large as Pluto. To build bodies of this size at 36 AU within 100 Myr (the estimated formation time of Neptune), the standard model of pairwise planetesimal accretion requires the primordial disk to have at least $\sim$1–10 $M_{\oplus}$ from 29 to 41 AU, depending on the assumed sizes of seed planetesimals (Kenyon & Luu 1998; cf. Stern & Colwell 1997).

The cause of the presumed depletion is unclear, but Neptune is considered a prime suspect. Duncan, Levison, & Budd (1995) calculate that Neptune, when fixed in its present orbit, can gravitationally eject more than 90% of the Belt mass inside 39 AU over the age of the Solar System. About 50–90% of the mass between 39 and 50 AU may be depleted by gravitational perturbations alone. These simulations are sensitive to assumed initial eccentricities and inclinations of test particles. Collisions are also proposed to explain the missing mass, either by nudging objects into unstable orbits (Davis & Farinella 1997), or by grinding bodies down to dust to be transported by radiation pressure (Stern & Colwell 1997). In both cases, however, only the smallest KBOs may be significantly depleted. Collisionally relaxed populations place most of their mass in the largest bodies, but most of their geometric cross section in the smallest members. Collisions might therefore preferentially grind down the smallest objects, leaving the largest bodies undisrupted and the total mass mostly intact. This expectation is borne out in computations by Davis & Farinella (1997). Important caveats for all collisional simulations of the Kuiper Belt include oversimplified prescriptions for the impact strengths of KBOs, reflecting our ignorance of their internal structure. Erosive velocities are thought to be gravitationally stirred by Neptune within 50 AU, but physically motivated estimates of the velocity dispersion have yet to be made in these simulations.

However large or small, the destructive influence of Neptune on the Kuiper Belt may be limited to within the location of its outermost 2:1 resonance at 48 AU. This idea has led to speculation that the surface density of Belt material rises by $\sim$2 orders of magnitude to its assumed primordial value somewhere beyond this radius (Stern 1996; Stern & Colwell

1997; MDL99). However, no classical KBO beyond 50 AU has yet been discovered. Assuming the shape of the KBO size distribution does not change with distance, JLT98 find by Monte Carlo simulation that their observations are consistent with an edge to the classical Kuiper Belt at 50 AU—a "Kuiper Cliff." The first theoretical constraints on the classical Belt mass beyond 50 AU are provided by Ward & Hahn (1998). They find under certain conditions that a Belt containing $1.6 M_{\oplus}$ from 48 to 75 AU damps Neptune's eccentricity to its current observed value of 0.009 by the action of density waves excited by Neptune. They calculate that the addition of 10 times more mass in this region (masses comparable to those expected in the minimum-mass solar nebula) would reduce Neptune's eccentricity to less than 10^{-20} over the age of the Solar System. In these computations, the outer Belt is assumed to consist predominantly of small bodies (diameters $\ll$ 140 km) so that velocity dispersions are sufficiently low to sustain wave action.

In the absence of any direct observations of the classical Kuiper Belt beyond 50 AU, we undertook a pilot survey utilizing the Keck 10-m telescope. The methods and results of our search comprise Chapter 6. There, we describe how we combine individual exposures to search (successfully) for KBOs of unprecedentedly faint magnitudes. We also present several order-of-magnitude calculations regarding the structure of the Belt that are based partly on our observations and partly on speculations.

1.3 Dynamics of Planetary Rings

1.3.1 Apse Alignment of Narrow Eccentric Rings

A ring system surrounding Uranus was discovered serendipitously during stellar occultation observations of the Uranian atmosphere (Elliot, Dunham & Mink 1977). Encircling the planet at the distance of $\sim$2 Uranian radii are nine narrow rings having geometric optical depths of order unity and radial widths of order 10 km (Elliot, Dunham & Mink 1977; Millis & Wasserman 1978; Elliot et al. 1978; Nicholson et al. 1978). French et al. (1991) review the observations of the Uranian ring system, including the Voyager fly-by data.

The radial confinement of these narrow rings is threatened by interparticle collisions,

Poynting-Robertson radiation drag, and gas drag from the distended exosphere of Uranus (see, e.g., Goldreich & Tremaine 1979a; Broadfoot et al. 1986). Interparticle collisions in a normally shearing ring transfer angular momentum from the inner half of the ring to the outer half, thereby causing the ring to spread radially (Lynden-Bell & Pringle 1974). Drag forces cause all ring particles to lose angular momentum and spiral in towards the planet. Goldreich & Tremaine (1979a) proposed that each ring is confined in radius by repulsive gravitational torques exerted by a pair of small satellites which bracket the ring on resonant orbits to either side. Analytic expressions for satellite-induced torques applied at Lindblad and corotation resonances within rings are derived by Goldreich & Tremaine (1978; 1979b; 1980). These pioneering studies also allow for non-zero eccentricities in satellite and ring orbits. Their standard shepherding theory posits that the net torque on the ring due to satellites, drag, and interparticle collisions is zero. The inner shepherd moon delivers exactly enough angular momentum per time to the inner half of the ring to cancel the rate of its removal by drag and by viscous torques. Similarly, the outer shepherd moon and drag forces extract enough angular momentum from the outer half of the ring to balance the rate of its accumulation by viscous torques. The discovery by Voyager of the Uranian ϵ ring shepherds, Cordelia and Ophelia, dramatically confirmed the basic predictions of Goldreich & Tremaine (see Porco & Goldreich 1987; Goldreich & Porco 1987). Cordelia establishes a 24:25 outer eccentric resonance at the inner edge of the ϵ ring, and Ophelia maintains a 13:14 inner eccentric resonance at the outer ring edge.

Shepherding theory underwent substantial revision after attempts were made to better understand the role of dissipation. The transfer of energy from satellite to ring leads to the following paradox within the standard theory (Borderies, Goldreich & Tremaine 1982, 1984). Resonant satellite perturbations enhance velocity dispersions near ring edges. The local enhancement in velocity dispersion appears to increase local viscosities. The consequently larger viscous torques overwhelm shepherding torques, prohibiting the ring from being confined. The resolution to this paradox lies in abandoning the (incorrect) assumption that increased velocity dispersions always lead to enhanced outward transport of angular momentum. Ring streamlines that are resonantly perturbed by an external satellite potential are not axisymmetric. Borderies, Goldreich & Tremaine (1983) find that the radial flux

of angular momentum across perturbed streamlines can, over certain ranges in azimuth, be inward (negative). In the hydrodynamic (high optical depth) limit, the direction of local angular momentum transport is given by the direction of orbital shear: outward transport results if the angular velocity decreases outwards, and vice versa. In this new theory of ring confinement, the flux of angular momentum across the ring decreases from its normal (positive) value in the unperturbed ring interior to zero at the perturbed ring edge. The vanishing of the angular momentum flux at the ring edge renders possible the extremely sharp edges of the ϵ ring; optical depths are observed to decrease from unity in the ring interior to zero at the ring edge over a radial lengthscale that is shorter than a few kilometers and is probably of the order of a hundred meters.

Stellar occultation observations reveal that the inner and outer edges of the Uranian ϵ ring can be fitted by two nested ellipses sharing the same line of apsides (Nicholson et al. 1978). Other narrow, eccentric, apse-aligned rings include the Uranian α and β rings, and the Saturnian Maxwell, Titan, and Huygens ringlets (Borderies, Goldreich & Tremaine 1983; Porco 1983; Porco et al. 1983, 1984). Apse alignment is surprising because the oblateness of Uranus causes orbits of particles with different semi-major axes to precess differentially. Timescales for differential precession in the absence of other forces are extremely short—shorter even than the duration of the observations. For example, in the absence of other forces, the inner edge of the ϵ ring would precess and collide into the outer edge in 3.4 yrs. Rigid precession of a narrow eccentric ring has remained a problem in ring dynamics since the discovery of the Uranian ring system more than 20 years ago.

In Chapter 7, we offer a "proof-of-concept" solution to the problem of rigid precession. We propose that rigid precession is maintained by a balance of forces due to ring self-gravity, planetary oblateness, and interparticle collisions. The first two kinds of forces were considered by Goldreich & Tremaine (1979c). Their predictions for the masses of the Uranian ϵ, α, and β rings are too low—in some cases by $\sim$2 orders of magnitude—compared to the lower limits obtained from Voyager radio occultations [see, e.g., French et al. (1991)]. Goldreich & Tremaine (1979c) suspected that interparticle collisions might exert significant stresses within the ring because the specific impulses required to substantially alter precession rates are only of the order of 1 mm/s; velocity dispersions of this

magnitude obtain naturally in the Uranian rings. However, the precise way in which interparticle collisions play a role was left unspecified. In this thesis, we shall illuminate the subtle and surprising way in which interparticle collisions complete the dynamical equilibrium manifested by these remarkably rich systems.

1.3.2 Convection and Angular Momentum Transport

When turbulence in unmagnetized Keplerian disks is driven by thermal convection, angular momentum appears to be transported inwards (Ryu & Goodman 1992; Kley, Papaloizou, & Lin 1993; Stone & Balbus 1996; Cabot 1996). Despite appearances, this result does not contradict the fact that an isolated Keplerian disk relaxes to its minimum energy state by transporting angular momentum outwards (Lynden-Bell & Pringle 1974). The reason is that simulated convective disks are not completely isolated. They tap a source of free energy other than their intrinsic differential rotation: an externally imposed, superadiabatic entropy gradient.

A more mechanistic understanding of the behavior of fluid disks can be obtained by considering disks composed of discrete macroscopic particles. Transport properties and stability criteria obtained under one picture often carry over to the other, to within dimensionless factors of order unity. The fluid/particle dualism has been employed with considerable success in studies of planetary rings (see, e.g., Goldreich & Tremaine 1982) and galactic structure (see, e.g., Chapter 6 of Binney & Tremaine 1987).

In Chapter 8, we briefly examine the transport properties of axisymmetric particle disks whose velocity dispersions may be externally enhanced ("stirred"). We ask whether a particle disk that is externally stirred in the direction perpendicular to the midplane can furnish the mechanical analogue of Stone & Balbus's (1996) vertically convective fluid disks.

Bibliography

[1] Adams, F. C., Lada, C. J., & Shu, F. H. 1987, ApJ, 312, 788

[2] Adams, F. C., Ruden, S. P., & Shu, F. H. 1989, ApJ, 347, 959

[3] Bertout, C. 1989, ARA&A, 27, 351

[4] Binney, J. & Tremaine, S., 1987, *Galactic Dynamics*, Princeton University Press: Princeton

[5] Borderies, N., Goldreich, P., & Tremaine, S. 1982, Nature, 299, 209

[6] Borderies, N., Goldreich, P., & Tremaine, S. 1983, AJ, 88, 1560

[7] Borderies, N., Goldreich, P., & Tremaine, S. 1984, in *Planetary Rings*, ed. R. Greenberg & A. Brahic (University of Arizona Press), 728

[8] Broadfoot, A.L. et al. 1986, Science, 233, 74

[9] Cabot, W. 1996, ApJ, 465, 874

[10] Calvet, N., Patiño, A., Magris, G., & D'Alessio, P. 1991, ApJ, 380, 617

[11] Calvet, N., Patiño, A., Magris, G., & D'Alessio, P. 1992, Rev. Mexicana Astron. Astrof., 24, 27

[12] Calvet, N., Hartmann, L., Kenyon, S. J., & Whitney, B. A. 1994, ApJ, 434, 330

[13] Chiang, E.I. & Goldreich, P. 1997, ApJ, 490, 368

[14] Cohen, M., Emerson, J.P., & Beichman, C.A. 1989, ApJ, 339, 455

[15] Davis, D. R. & Farinella, P. 1997, Icarus, 125, 50

[16] Duncan, M. J. & Levison, H. F. 1997, Science, 276, 1670

[17] Duncan, M. J., Levison, H. F., & Budd, S. M. 1995, AJ, 110, 373

[18] Duncan, M., Quinn, T., & Tremaine, S. 1988, ApJ, 328, L69

[19] Edgeworth, K.E. 1943, J. British Astron. Assoc., XX, 181

[20] Elliot, J.L., Dunham, E., & Mink, D. 1977, Nature, 267, 328

[21] Elliot, J.L., et al. 1978, AJ, 83, 980

[22] Fernandez, J.A. 1980, MNRAS, 192, 481

[23] French, R.G., Nicholson, P.D., Porco, C.C., & Marouf, E.A. 1991, in *Uranus*, ed. J.T. Bergstrahl, E.D. Miner, & M.S. Matthews (University of Arizona Press), 327

[24] Goldreich, P. & Porco, C.C. 1987, AJ, 93, 730

[25] Goldreich, P. & Tremaine, S. 1978, Icarus, 34, 24

[26] Goldreich, P. & Tremaine, S. 1979a, Nature, 277, 97

[27] Goldreich, P. & Tremaine, S. 1979b, ApJ, 233, 857

[28] Goldreich, P. & Tremaine, S. 1979c, AJ, 84, 1638

[29] Goldreich, P. & Tremaine, S. 1980, ApJ, 241, 425

[30] Goldreich, P. & Tremaine, S., 1982, ARA&A, 20, 249

[31] Jewitt, D. & Luu, J. 1993, Nature, 362, 730

[32] Jewitt, D., Luu, J., & Trujillo, C. 1998, AJ, 115, 2125 (JLT98)

[33] Joy, A. H. 1942, PASP, 54, 15

[34] Joy, A. H. 1945, ApJ, 102, 168

[35] Joy, A. H. 1949, ApJ, 110, 424

[36] Hamid, S. E., Marsden, B. G., & Whipple, F. L. 1968, AJ, 73, 727

[37] Kenyon, S.J., Hartmann, L.W., Strom, K.M., & Stron, S.E. 1990, AJ, 99, 869

[38] Kenyon, S. J. & Luu, J. X. 1998, AJ, 115, 2136

[39] Kley, W., Papaloizou, J., & Lin, D., 1993, ApJ, 416, 679

[40] Kuiper, G. 1951, in *Astrophysics: A Topical Symposium*, ed. J.A. Hynek (McGraw Hill: New York), 357

[41] Luu, J., Jewitt, D., Trujillo, C. A., Hergenrother, C. W., Chen, J., & Offutt, W. B. 1997, Nature, 287, 573

[42] Luu, J. X. & Jewitt, D. C. 1998, ApJ, 502, L91

[43] Lynden-Bell, D. & Pringle, J.E. 1974, MNRAS, 168, 603

[44] Malhotra, R. 1995, AJ, 110, 420

[45] Malhotra, R. 1996, AJ, 111, 504

[46] Malhotra, R., Duncan, M., & Levison, H. 1999, in *Protostars and Planets IV*, ed. V. Mannings, A. Boss, & S. Russell (University of Arizona Press: Tucson), preprint (MDL99)

[47] McCaughrean, M. J. & O'Dell C. R. 1996, AJ, 111, 1977

[48] Mendoza, E.E. 1966, ApJ, 143, 1010

[49] Mendoza, E.E. 1968, ApJ, 151, 977

[50] Millis, R.L. & Wasserman, L.H. 1978, AJ, 83, 993

[51] Natta, A. 1993, ApJ, 412, 761

[52] Nicholson, P.D., et al. 1978, AJ, 83, 1240

[53] Ostriker, E. C., Shu, F. H., & Adams, F. C. 1992, ApJ, 399, 192

[54] Porco, C. 1983, Ph.D. Thesis, California Institute of Technology

[55] Porco, C., et al. 1983, in I.A.U. Colloquium No. 75, Planetary Rings, ed. A. Brahic & R.A. Greenberg

[56] Porco, C., et al. 1984, Icarus, 60, 1

[57] Porco, C.C. & Goldreich, P. 1987, AJ, 93, 724

[58] Quinn, T., Tremaine, S., & Duncan, M.J. 1990, ApJ, 355, 667

[59] Rucinski, S. M. 1985, AJ, 90, 2321

[60] Rydgren, A. E., Schmelz, J. T., Zak, D. S., & Vrba, F. J. 1984, Broad Band Spectral Energy Distributions of T Tauri Stars in the Taurus-Auriga Region, Pub. US Naval Obs., 2nd Ser. Vol. 25, 1

[61] Rydgren, A. E. & Zak, D. S. 1987, PASP, 99, 141

[62] Ryu, D. & Goodman, J., 1992, ApJ, 338, 438

[63] Shu, F. H., Adams, F. C., & Lizano, S. 1987, ARA&A, 25, 23

[64] Shu, F. H., Tremaine, S., Adams, F. C., & Ruden, S. P. 1990, ApJ, 358, 495

[65] Stern, S. A. 1996, AJ, 112, 1203

[66] Stern, S. A. & Colwell, J. E. 1997, ApJ, 490, 879

[67] Stone, J. M. & Balbus, S. A., 1996, ApJ, 464, 364

[68] Strom, K. M., Strom, S. E., Edwards, S., Cabrit, S., & Skrutskie, M. F. 1989a, AJ, 97, 1451

[69] Strom, K.M., et al. 1989b, ApJS, 71, 183

[70] Sylvester, R.J., Skinner, C.J., Barlow, M.J., & Mannings, V. 1996, MNRAS, 279, 915

[71] Tout, C.A., Livio, M., & Bonnell, I.A. 1999, MNRAS, 310, 360

[72] Ward, W. R. & Hahn, J. M. 1998, Science, 280, 2104

[73] Waters, L.B.F.M. & Waelkens, C. 1998, ARA&A, 36, 233

[74] Weidenschilling, S. J. 1977, Ap&SS, 51, 153

Part I

PASSIVE DISKS SURROUNDING T TAURI AND HERBIG Ae/Be STARS

Chapter 2

Spectral Energy Distributions: Analytic Theory

This chapter draws from

Chiang & Goldreich 1997, ApJ, 490, 368

Abstract

We derive hydrostatic, radiative equilibrium models for passive disks surrounding T Tauri stars. Each disk is encased by an optically thin layer of superheated dust grains. This layer re-emits directly to space about half the stellar energy it absorbs. The other half is emitted inward and regulates the interior temperature of the disk. The heated disk flares. As a consequence, it absorbs more stellar radiation, especially at large radii, than a flat disk would. The portion of the SED contributed by the disk is fairly flat throughout the thermal infrared. At fixed frequency, the contribution from the surface layer exceeds that from the interior by about a factor 3 and is emitted at more than an order of magnitude greater radius. Spectral features from dust grains in the superheated layer appear in emission if the disk is viewed nearly face-on.

2.1 Introduction

In this chapter, we describe in basic terms how stellar radiation is reprocessed by passive disks. Spectral energy distributions are computed for face-on viewing aspects. The basic ingredients of our model are set forth in §2.2. In §2.3, models of pure blackbody disks are derived with either flat or flared (hydrostatic) geometries. In §2.4, we advance our 2-layer, radiative equilibrium models, treating again the two types of geometries. Section 2.4.2 contains our standard model: a passive disk in radiative and hydrostatic equilibrium. In §2.5, we demonstrate how spatially resolved, broadband observations can distinguish

between emission from disk surface layers and from the disk interior. We close in §2.6 by examining how spectral dust resonances naturally appear in emission when our standard model disk is viewed face-on.

2.2 Model Assumptions

We consider a T Tauri star surrounded by a passive disk. Table 2.1 summarizes the parameters describing our basic theory. The star is modeled as a spherical black body of temperature $T_* = 4000\,\mathrm{K}$, mass $M_* = 0.5 M_\odot$, and radius $R_* = 2.5 R_\odot$ (see, e.g., Table II in Beckwith et al. 1990). Our fiducial disk has surface mass density similar to that of the minimum-mass solar nebula, $\Sigma = a_{\mathrm{AU}}^{-3/2}\Sigma_0$, where a_{AU} is the disk radius measured in AU and $\Sigma_0 = 10^3\,\mathrm{g\,cm^{-2}}$ (Weidenschilling 1977). Dust, which is uniformly mixed with the gas, comprises about one percent of the total mass. Dust grains dominate the continuum opacity from visible through millimeter wavelengths. We take the grains to be spheres with radius $r = 0.1\mu\mathrm{m}$, mass density $\rho_d = 2\,\mathrm{g\,cm^{-3}}$, and negligible albedo. Their emissivity, which is unity for $\lambda \leq 2\pi r$, decreases as $\varepsilon_\lambda \approx (2\pi r/\lambda)^\beta$ for $\lambda \geq 2\pi r$. We denote by $\varepsilon \approx (8\pi r k T/hc)^\beta \approx (T/T_*)^\beta$ the average dust emissivity at temperature T. For our fiducial disk, $\beta = 1$. The dust opacity at visual wavelengths is $\kappa_V \approx 400\,\mathrm{cm^2\,g^{-1}}$, which implies an optical depth $\tau_V \approx 4 \times 10^5\, a_{\mathrm{AU}}^{-3/2}$ for our fiducial disk.

SEDs are computed for disks viewed pole-on. We choose an inner cutoff radius, $a_i \approx 6R_* \approx 0.07\mathrm{AU}$, to mark the condensation boundary of common silicates. The outer cutoff radius for flat disks, $a_o \approx 2.3 \times 10^4 R_* \approx 2.7 \times 10^2\mathrm{AU}$, is fixed to facilitate comparisons among different disk models; a_o is comparable to the size of the largest disks seen in silhouette against the Orion nebula (McCaughrean & O'Dell 1996).

Our model assumptions are chosen to make the telling of our story as simple and direct as possible. Issues regarding these choices and possible alternatives are dealt with in Chapters 4 and 5.

2.3 Blackbody Disk

To begin, we review standard relations for the temperature and SED of a blackbody disk. The flux of stellar radiation incident upon the disk is $(\alpha/2)\ (R_*/a)^2\ \sigma T_*^4$ for $a \gg R_*$, where α is the grazing angle at which the starlight strikes the disk. Equating the emitted and absorbed fluxes yields the disk temperature[1]

$$T_e \approx \left(\frac{\alpha}{2}\right)^{1/4} \left(\frac{R_*}{a}\right)^{1/2} T_* \,. \tag{2.1}$$

The SED is computed as

$$L_\nu \equiv 4\pi d^2 \nu F_\nu = 8\pi^2 \nu \int_{a_i}^{a_o} da\, a B_\nu(T_e)\,, \tag{2.2}$$

where $B_\nu(T)$ is the Planck function, and d is the distance to the source. A scaling relation for L_ν at wavelengths between those that characterize the disk at a_i and a_o is derived as follows. We approximate the integral in equation (2.2) by $8\pi^2 a^2 \nu B_\nu(T_e)$, replace $\nu B_\nu(T_e)$ by $\sigma T_e^4/\pi$, relate a to T_e by equation (2.1), and express T_e in terms of ν by $3kT_e \sim h\nu$.

[1]The subscript e is used to denote effective temperature.

Table 2.1: Input Parameters

Symbol	Meaning	Value
M_*	Stellar Mass	$0.5\,M_\odot$
R_*	Stellar Radius	$2.5\,R_\odot$
T_*	Stellar Effective Temperature	4000 K
Σ	Disk Surface Density	$10^3\,a_{\rm AU}^{-3/2}$ g cm^{-2}
a_i	Inner Disk Radius	$6\,R_* = 0.07$ AU
a_o	Outer Disk Radius	$23000\,R_* = 270$ AU
κ_V	Dust Opacity at Visual Wavelengths[a]	400 cm^2 g^{-1}
ε_ν	Grain Emissivity	$\begin{cases} 1 & \text{if } \lambda \le 2\pi r \\ (\frac{2\pi r}{\lambda})^\beta & \text{otherwise} \end{cases}$
r	Grain Radius	$0.1\,\mu$m
β	Grain Emissivity Index	1

[a] Absorption by dust grains is assumed to dominate the continuum opacity from visible through millimeter wavelengths.

These steps yield

$$L_\nu \sim 8\pi a^2 \sigma T_e^4 \sim \alpha L_* \,. \tag{2.3}$$

In other words, the fraction of the stellar luminosity, L_*, that is reprocessed to frequencies in an octave centered on ν is approximately equal to the grazing angle α at the location where $3kT_e = h\nu$.

2.3.1 Flat Geometry

A flat disk is one whose aspect ratio (opening angle) is independent of a; we assume it to be much less than unity. The grazing angle appropriate to this geometry is $\alpha \approx 0.4R_*/a \ll 1$ for $a/R_* \gg 1$ (Kusaka, Nakano, & Hayashi 1970, Ruden & Pollack 1991). Thus T_e from equation (2.1) takes the form

$$T_e \approx \left(\frac{2}{3\pi}\right)^{1/4} \left(\frac{R_*}{a}\right)^{3/4} T_* \,. \tag{2.4}$$

Application of the scaling relation given by equation (2.3) to the flat disk gives $L_\nu \sim 0.01(\nu/10^{13}\,\mathrm{Hz})^{4/3} L_*$. Figure 2.1 confirms that L_ν obeys this relation for $30 \lesssim \lambda \lesssim 1000\,\mu\mathrm{m}$.

2.3.2 Hydrostatic Equilibrium

We retain the assumption that the disk radiates as a blackbody and consider the consequences of vertical hydrostatic equilibrium in a gravitational field $g = \Omega^2 z$. This is the case investigated by Kenyon & Hartmann (1987).

The disk temperature T_e is still given by equation (2.1), but now the grazing angle α takes the more general form

$$\alpha \approx \frac{0.4R_*}{a} + a\frac{d}{da}\left(\frac{H}{a}\right) , \tag{2.5}$$

where H is the height of the visible photosphere above the disk midplane.

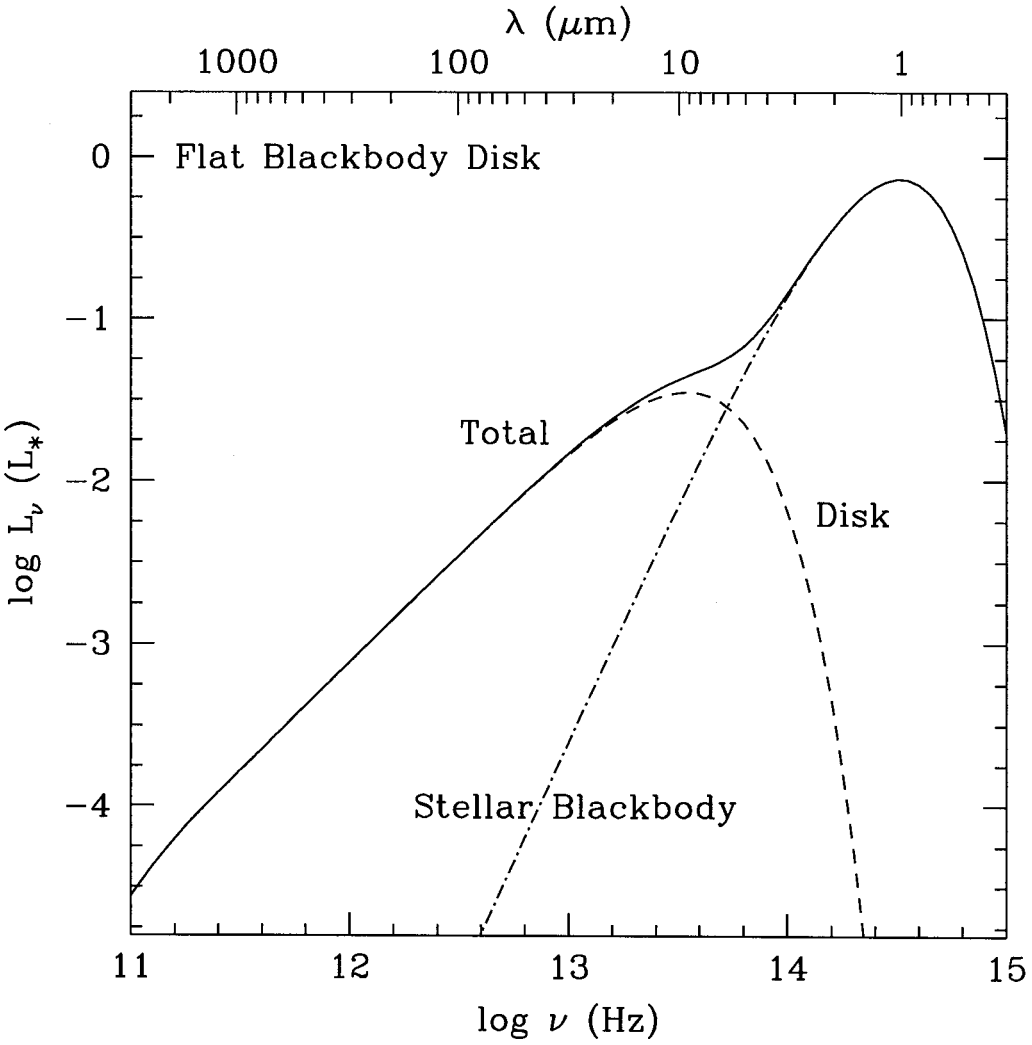

Figure 2.1: *SED for the flat blackbody disk.*

Contributions from the star and the disk are identified. The $n = 4/3$ *law is evident between* $30\,\mu$m *and* 1 mm. *The turnover near* 1 mm *is due to our truncation of the disk at* $a_o \approx 270$ AU.

Taking the gas to be isothermal at temperature T_e yields a Gaussian vertical density profile

$$\frac{n}{n_0} = \exp\left(-\frac{z^2}{2h^2}\right), \tag{2.6}$$

where

$$\frac{h}{a} = \left(\frac{T_e}{T_c}\right)^{1/2}\left(\frac{a}{R_*}\right)^{1/2}. \tag{2.7}$$

The temperature T_c is a measure of the gravitational potential at the surface of the central star;

$$T_c \equiv \frac{GM_*\mu_g}{kR_*} \approx 8\times 10^6\,\mathrm{K}, \tag{2.8}$$

with μ_g the mean molecular weight of the gas. Under the assumption that the dust to gas ratio is uniform throughout the disk, we find

$$\frac{H}{h} = \left[2\ln\left(\frac{n_0}{n_{ph}}\right)\right]^{1/2}, \tag{2.9}$$

where n_{ph} is the number density in the photosphere.

Armed with equations (2.5)–(2.9), we are ready to evaluate H/a. In the limit of large radius where α is dominated by the flaring term,

$$\frac{H}{a} \approx 4\left(\frac{T_*}{T_c}\right)^{4/7}\left(\frac{a}{R_*}\right)^{2/7} \approx 0.18\,a_{\mathrm{AU}}^{2/7}. \tag{2.10}$$

In writing equation (2.10) we have set $H/h = 4$; in reality, this factor declines from about 5 at $a_{\mathrm{AU}} = 3$ to 4 at $a_{\mathrm{AU}} = 10^2$.

It follows from equations (2.5) and (2.10) that $\alpha \approx 0.005a_{\mathrm{AU}}^{-1} + 0.05a_{\mathrm{AU}}^{2/7}$. Thus α is minimal at the transition radius $a_{tr} \approx 0.4\mathrm{AU}$ with $\alpha_{min} \approx 0.05$. Beyond the transition radius the disk flares until $H \approx a$ at $a_o \approx 270\,\mathrm{AU}$.

The SED for the flared blackbody disk truncated at a_o, as computed from equation (2.2), is shown in Figure 2.2. At wavelengths between 10 and 100 μm, the disk emission

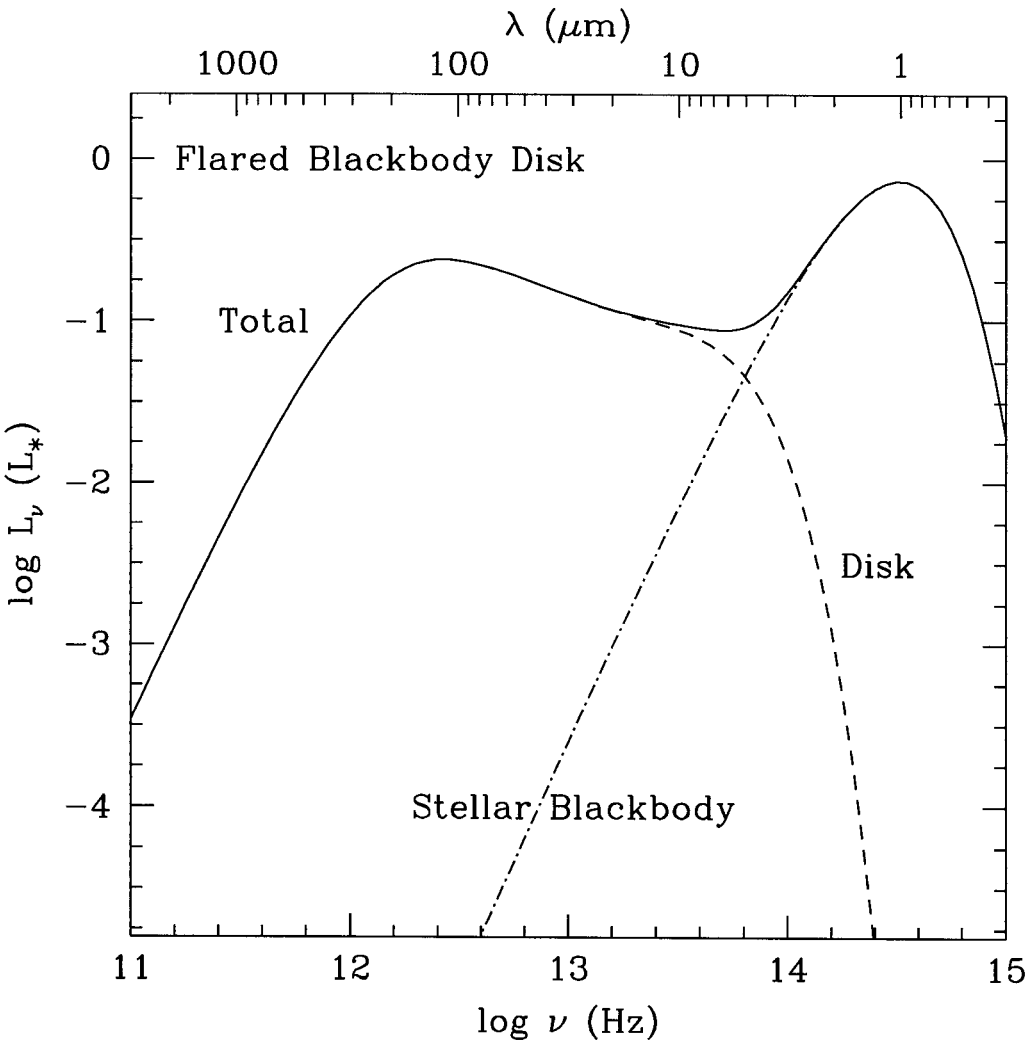

Figure 2.2: *SED for the flared blackbody disk.*
At mid-IR wavelengths, $L_\nu \propto \nu^{-2/3}$. *At longer wavelengths,* $L_\nu \propto \nu^3$.

follows the scaling relation [cf. eq. (2.3)] $L_\nu \sim \alpha L_* \sim 0.1(\nu/10^{13}\,\mathrm{Hz})^{-2/3} L_*$. Longward of 300 μm in the Rayleigh-Jeans regime, the SED varies as ν^3.

2.4 Radiative Equilibrium Disk

Here we drop the blackbody assumption and determine the SED by application of the techniques of radiative transfer. The definition of the effective temperature, T_e, remains as given by equation (2.1). A simplified description of the disk distinguishes two regions.

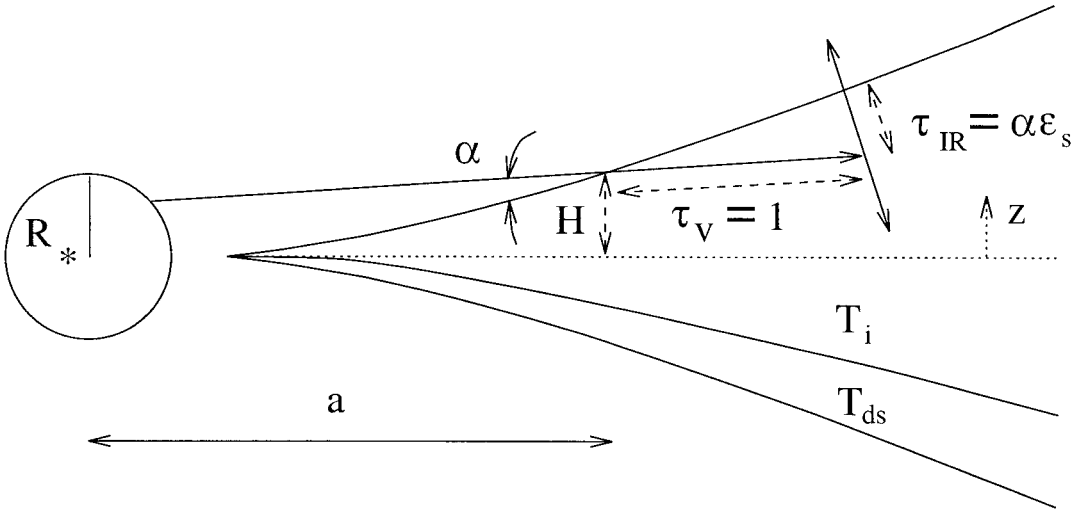

Figure 2.3: *Radiative transfer in the passive disk.*

Stellar radiation strikes the surface at an angle α and is absorbed within visible optical depth unity. Dust particles in this first absorption layer are superheated to a temperature T_{ds}. About half of the emission from the superheated layer emerges as dilute blackbody radiation. The remaining half heats the interior to a temperature T_i.

The surface layer contains grains which are directly exposed to light from the central star. Variables evaluated there are denoted by a subscript s. We allow for the possibility that the gas temperature in this layer, T_{gs}, may be smaller than the dust temperature, T_{ds}. The rest of the disk is lumped together as the interior. Variables evaluated there carry a subscript i. We assume that in steady-state the interior gas and dust temperatures are equal; $T_{gi} = T_{di} = T_i$.

A schematic of how the stellar radiation is reprocessed is illustrated in Figure 2.3. The radiation penetrates the disk to an optical depth of order unity along a slant path inclined by an angle α to the surface. Dust particles along the slant path are "superheated" to a temperature

$$T_{ds} \approx \frac{1}{\varepsilon_s^{1/4}} \left(\frac{R_*}{2a}\right)^{1/2} T_* \approx \frac{550}{a_{\rm AU}^{2/5}}\,{\rm K}\,, \tag{2.11}$$

which is greater than the blackbody temperature $T_{BB} \approx (R_*/a)^{1/2} T_*$ because their absorp-

tivity in the visible exceeds their emissivity in the infrared. The normal optical depth of the superheated dust layer at visual wavelengths is $\tau_V \approx \alpha$. The infrared optical depth is smaller still; $\tau_{IR} \approx \alpha\varepsilon_s$.

The superheated dust radiates equal amounts of IR radiation into the inward and outward hemispheres. The outward directed radiation is similar to that of a dilute blackbody. Where the disk is opaque to blackbody radiation at

$$T_{i1} \approx \frac{T_e}{2^{1/4}} \approx \left(\frac{\alpha}{4}\right)^{1/4} \left(\frac{R_*}{a}\right)^{1/2} T_* \,, \tag{2.12}$$

the inward directed radiation is thermalized at that temperature. The outer boundary of this region is denoted by a_{th}. Just outside a_{th}, the interior is optically thin to its own radiation but still opaque to radiation from the superheated surface; in these regions the interior temperature is determined by thermal balance to be

$$T_{i2} \approx \left(\frac{\alpha}{4\varepsilon_i \kappa_V \Sigma}\right)^{1/4} \left(\frac{R_*}{a}\right)^{1/2} T_* \,. \tag{2.13}$$

At still greater radii, the encased material is transparent to both its own radiation and to radiation from the surface; the internal temperature here is given by

$$T_{i3} \approx \left(\frac{\alpha\varepsilon_s^2}{\varepsilon_i}\right)^{1/4} T_{ds} \approx \left(\frac{\alpha\varepsilon_s}{4\varepsilon_i}\right)^{1/4} \left(\frac{R_*}{a}\right)^{1/2} T_* . \tag{2.14}$$

The SED for the radiative equilibrium disk is computed from

$$L_\nu = 8\pi^2 \nu \int_{a_i}^{a_o} da\, a \int_{-\infty}^{\infty} dz \, \frac{d\tau_\nu}{dz} \, e^{-\tau_\nu} B_\nu(T) \,, \tag{2.15}$$

where τ_ν measures optical depth from z to ∞ along the axis perpendicular to the disk midplane.

2.4.1 Flat Geometry

Once again we consider the flat disk, but now under conditions of radiative equilibrium. The appropriate expression for the effective temperature is given by equation (2.4). For our fiducial flat disk $a_{th} \approx 50$ AU. Runs of T_{ds} and T_i as functions of a are displayed in Figure

2.4.

The SED for the flat, radiative equilibrium disk as calculated from equation (2.15) is displayed in Figure 2.5. Its appearance is similar to that of the SED for the flat blackbody disk. Over most of the IR, it is dominated by radiation from the optically thick interior. The surface layer radiates more than the interior shortward of $6\,\mu\mathrm{m}$; however, there its contribution is hidden by that from the central star. Most of the radiation longward of a millimeter comes from the outer, optically thin part of the disk. This accounts for the drop of the SED below the extrapolation of the $n = 4/3$ power law.

2.4.2 Hydrostatic Equilibrium

Now we investigate the disk model in which both vertical hydrostatic equilibrium and radiative transfer are treated in a self-consistent fashion. The flaring geometry is governed by equations (2.5) through (2.9). In the limit where the grazing angle is dominated by the flaring term and the disk is opaque to its thermal radiation, the expression for H/a is nearly identical to that given by equation (2.10). In other, more optically thin regimes, the flaring geometry changes slightly as T_i takes successively different forms.

A plot of T_i for the flared, radiative equilibrium disk is included in Figure 2.4. Approximate fitting formulae for T_i and H/a in the three regions of the flared, radiative equilibrium disk are as follows. For $0.4 < a_{\mathrm{AU}} < 84$,

$$T_{i1} \approx \frac{150}{a_{\mathrm{AU}}^{3/7}}\ \mathrm{K} \tag{2.16}$$

$$H/a \approx 0.17\, a_{\mathrm{AU}}^{2/7}\,; \tag{2.17}$$

for $84 < a_{\mathrm{AU}} < 209$,

$$T_{i2} \approx 21\ \mathrm{K} \tag{2.18}$$

$$H/a \approx 0.59 \left(\frac{a_{\mathrm{AU}}}{84}\right)^{1/2}\,; \tag{2.19}$$

and for $209 < a_{\mathrm{AU}} < 270$,

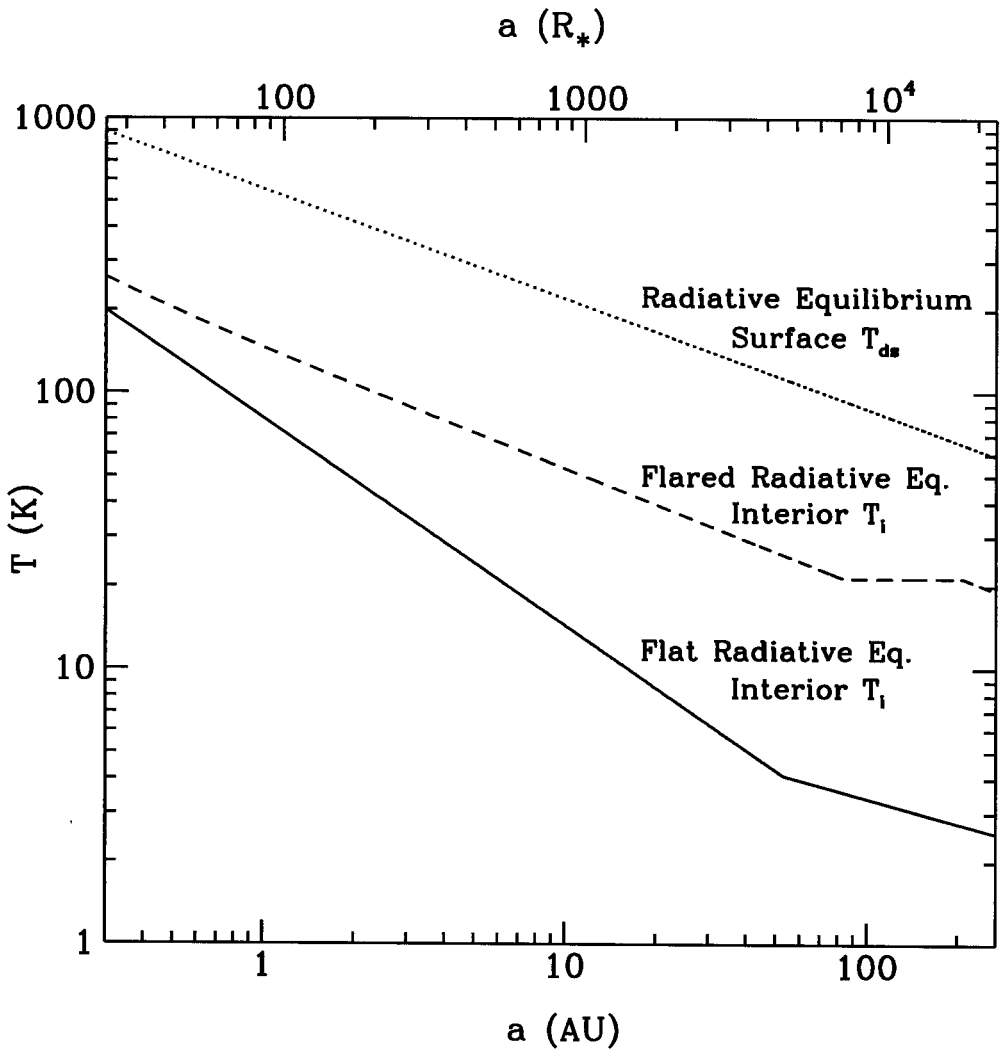

Figure 2.4: *Temperature profiles of the flat and flared radiative equilibrium models. The dust temperature T_{ds} of the superheated layer is independent of disk geometry. Expressions for T_{ds} and T_i are provided in the text. The flat disk is truncated at $a_o = 270$ AU (to facilitate comparison with the flared models), before the third temperature regime is reached.*

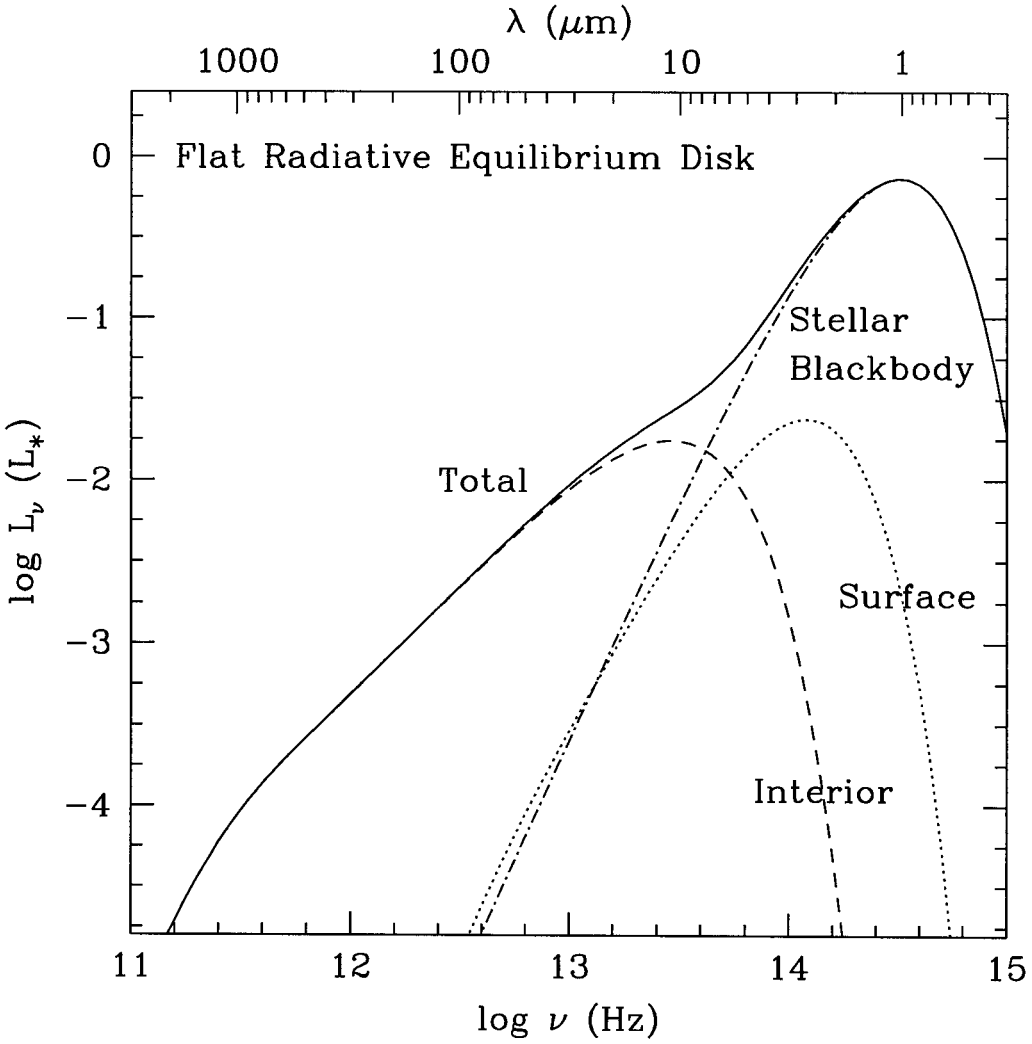

Figure 2.5: *SED for the flat radiative equilibrium disk.*

Emission from the superheated surface is hidden by central starlight. For $30 \lesssim \lambda(\,\mu\mathrm{m}) \lesssim 1000$, *optically thick emission from the disk interior resembles that from the flat blackbody disk (see Figure 2.1), but is reduced by a factor of 2. Emission longward of* $1\,\mathrm{mm}$ *is optically thin.*

$$T_{i3} \approx 21 \left(\frac{209}{a_{\rm AU}}\right)^{19/45} \ {\rm K} \tag{2.20}$$

$$H/a \approx 0.92 \left(\frac{a_{\rm AU}}{209}\right)^{13/45} . \tag{2.21}$$

The SED for the flared disk truncated at $a_o \approx 270\,{\rm AU}$, as computed from equation (2.15), is shown in Figure 2.6.[2] Radiation emitted by the surface mirrors that of the interior, but is shifted towards and dominates at shorter wavelengths. The near power law rises of both disk components of the SED in the mid-infrared are readily obtained following the procedure used in the derivation of equation (2.3). The superheated layer obeys the scaling relation $L_{\nu,s} \sim 8\pi a^2 \tau_{IR}\, \sigma T_{ds}^4 \sim \alpha L_*/2 \sim 0.06(\nu/10^{13}\,{\rm Hz})^{-5/7} L_*$. Similarly, the opaque interior obeys $L_{\nu,i} \sim 8\pi a^2 \sigma T_i^4 \sim \alpha L_*/2 \sim 0.02(\nu/10^{13}\,{\rm Hz})^{-2/3} L_*$. That the former exceeds the latter at fixed wavelength reflects the fact that compared to $L_{\nu,i}$, most of the contribution to $L_{\nu,s}$ arises from larger radii and hence larger α.

2.5 Spatially Resolved Broadband Observations

Pioneering observations which marginally resolve disks at mm wavelengths are reported in Lay et al. (1994) and Mundy et al. (1996). Fits of elliptical Gaussians to the $\lambda = 0.87\,{\rm mm}$ brightness distributions of HL Tau and L1551 IRS 5 yield semi-major radii at half-maximum brightness of 60 and 80 AU, respectively (Lay et al. 1994). Similar results were obtained at $\lambda = 2.7\,{\rm mm}$ for HL Tau, with the semi-minor radius marginally resolved to be $\sim 30\,{\rm AU}$ (Mundy et al. 1996).

Interferometric observations of the nearest T Tauri disks at $\lambda \approx 10\,\mu{\rm m}$ with 10 milliarcsecond resolution might separate the contributions to the SED from the superthermal dust layer and the disk interior. Running integrals for $L_{\nu,s}$ and $L_{\nu,i}$ against a are displayed in Figure 2.7. Most of the $10\,\mu{\rm m}$ emission inwards of 1 AU originates from the disk interior, whereas the contribution from the superthermal layer is localized in an annulus centered on $\sim 10\,{\rm AU}$. Similar results hold for the $20\,\mu{\rm m}$ flux. At wavelengths longer than $\sim 100\,\mu{\rm m}$, the disk interior dominates the emission.

[2]The disk may terminate in a slow thermal wind since $H/a \approx 1$ at $a_{\rm AU} = 270$.

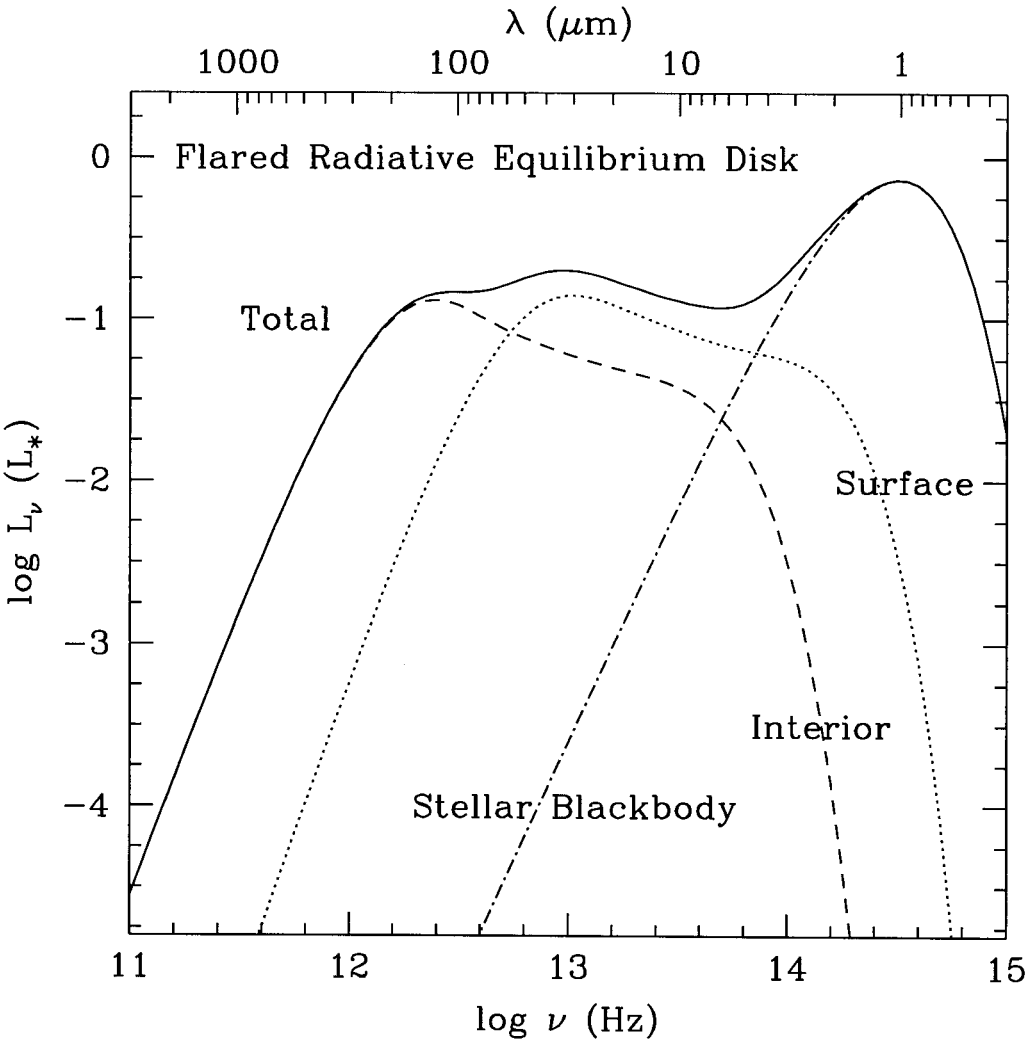

Figure 2.6: *SED for the hydrostatic, radiative equilibrium disk.*

At mid-IR wavelengths, the superheated surface radiates approximately 2–3 *times more power than the interior. Longward of* 300 μm, n *gradually steepens from about* 3 *to* $3 + \beta$ *as the disk becomes increasingly optically thin.*

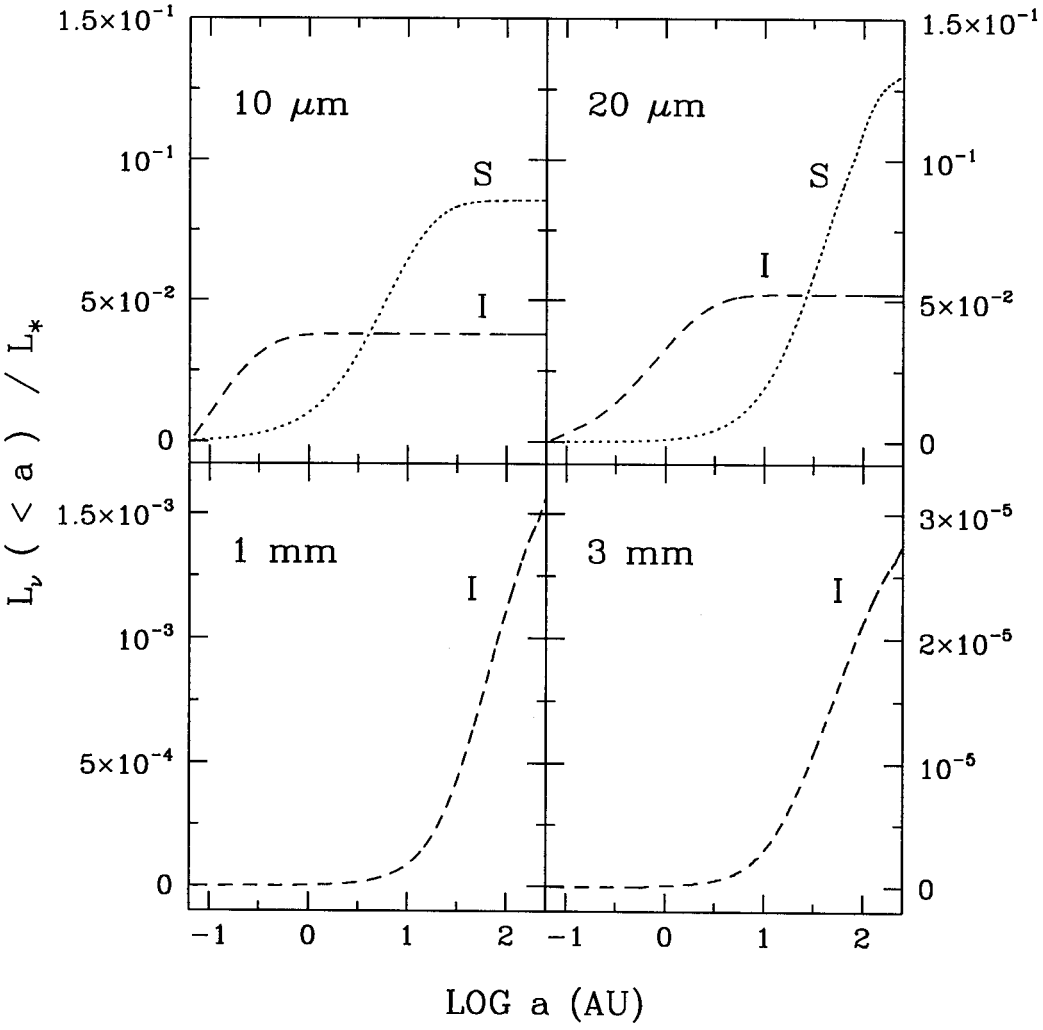

Figure 2.7: *Running contributions to the SED from the superthermal dust layer (S) and the disk interior (I) at wavelengths for which the disk may be increasingly resolved by interferometers.*

For a given disk component, $L_\nu(< a)$ *is given by equation (2.15), with* a_o *replaced by* a*. At* $\lambda = 1$ *and* 3 mm*, the contributions from the surface layer are negligible and not shown.*

By contrast, a flared blackbody disk concentrates all of its mid-infrared flux at small radii. It would only generate curves similar to those marked "I" in the panels for 10 and 20 μm. Infrared observations of face-on disks with upcoming nulling interferometers would provide an excellent discriminant between the conventional blackbody hypothesis and our conception of an optically thin superheated layer that emits substantial power at large radius; this test seems quite feasible (Kuchner 2000).

2.6 Spectral Dust Features

Dust grains possess spectral resonances. Among the best-known are those at $\lambda = 9.7\,\mu$m and $18\,\mu$m which are thought to arise from silicates (see, e.g., Draine 1995); the former has been observed in both emission and absorption in the spectra of T Tauri stars. For example, Cohen & Witteborn (1985) identified $10\,\mu$m emission features in 24 stars and $10\,\mu$m absorption features in another 7 in a spectrophotometric survey of 32 T Tauri stars. The emission features evince line to continuum ratios of about 1.2–3.0 to 1.

Emission features are a natural consequence of the superheated dust layer for disks viewed nearly face-on (see also Calvet et al. 1992). Line to continuum ratios should be of the same order as the percentage increase in dust emissivity associated with the spectral feature. To illustrate the effects of spectral resonances, we have employed an emissivity profile motivated by the data for "outer-cloud dust" as given by Mathis (1990). The resulting (face-on) SED displayed in Figure 2.8 exhibits emission features at 10 and 20 μm.

Absorption features may be associated with disks viewed nearly edge-on. Support for this hypothesis comes from two observational findings. These features appear most prominently among the "extreme" flat spectrum sources for which L_ν at infrared wavelengths matches or exceeds L_ν in the visible. There is a positive correlation between the strengths of these features and the fractional linear polarization at optical wavelengths. Both findings are plausibly the consequence of the extinction of optical radiation in a flared disk viewed at large inclination angle. This expectation is realized in the next chapter.

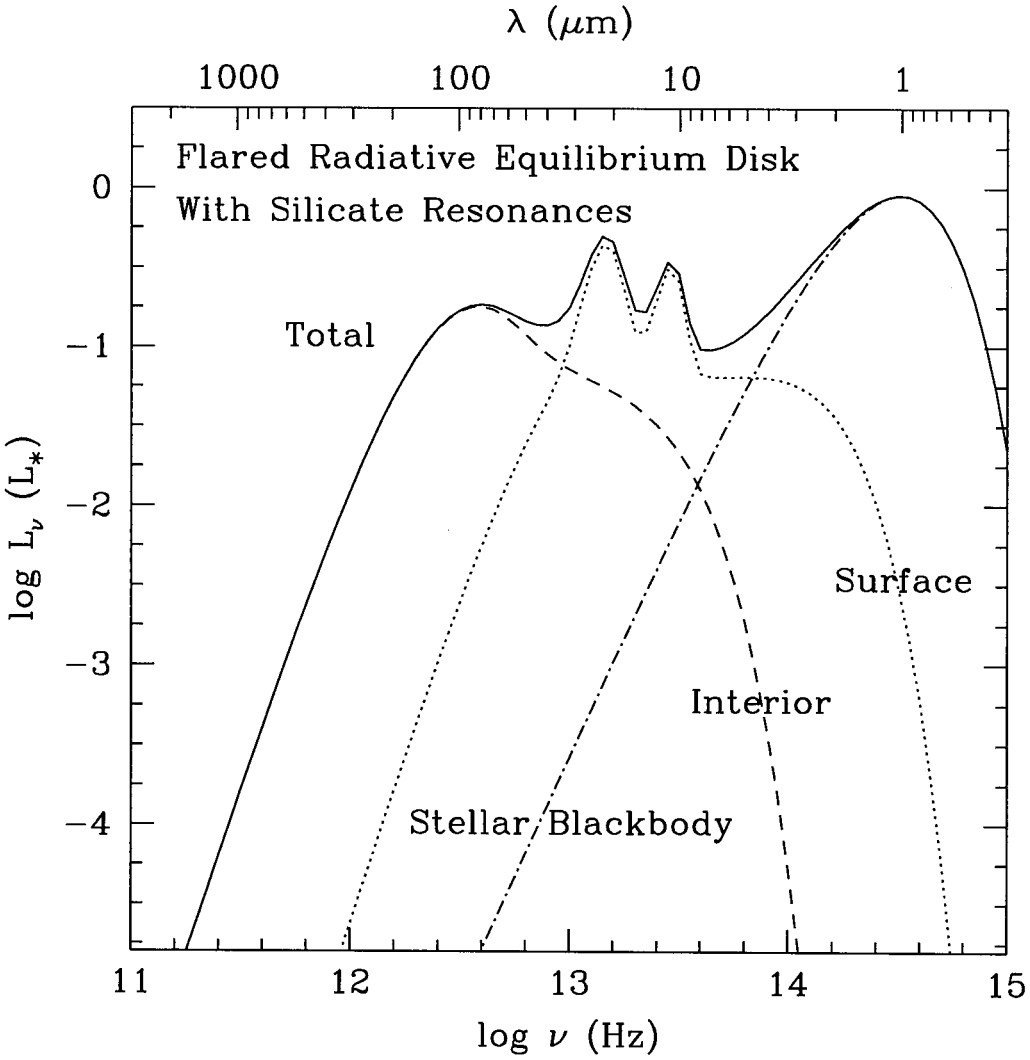

Figure 2.8: *SED for the hydrostatic, radiative equilibrium disk using a grain emissivity profile motivated by data from Mathis (1990).*

For wavelengths shorter than $0.3\ \mu$m, *our assumed emissivity is unity; longward of* $0.3\ \mu$m, *it obeys a (single) power law relation* $\varepsilon_\lambda = (0.3\ \mu\mathrm{m}/\lambda)^{1.4}$, *on which are superposed two Gaussians centered on* 10 *and* $20\ \mu$m, *having amplitudes that are* 3 *times their local continuum emissivities and FWHM equal to* 3 *and* $9\ \mu$m, *respectively.*

Acknowledgments

We thank Steve Beckwith for a thorough and thoughtful referee's report, and also for bringing to our attention the possibility of silicate emission lines from the superheated surface. Financial support for this research was provided by NSF Grant 94-14232 and an NSF Graduate Fellowship.

Bibliography

[1] Beckwith, S. V. W., Sargent, A. I., Chini, R. S., & Gusten, R. 1990, AJ, 99, 924

[2] Calvet, N., Patiño, A., Magris, G., & D'Alessio, P. 1992, Rev. Mexicana Astron. Astrof., 24, 27

[3] Cohen, M. & Witteborn, F. C. 1985, ApJ, 294, 345

[4] Draine, B. T. 1995, in *The Physics of the Interstellar Medium and Intergalactic Medium*, ed. A. Ferrara, C. F. McKee, C. Heiles, & P. R. Shapiro. A. S. P. Conf. Series, 80, 133

[5] Kenyon, S. J. & Hartmann, L. 1987, ApJ, 323, 714

[6] Kusaka, T., Nakano, T., & Hayashi, C. 1970, Prog. Theo. Phys., 44, 1580

[7] Lay, O. P., Carlstrom, J. E., Hills, R. E., & Phillips, T. G. 1994, ApJ, 434, L75

[8] Mathis, J. M. 1990, ARA&A, 28, 37

[9] McCaughrean, M. J. & O'Dell C. R. 1996, AJ, 111, 1977

[10] Mundy, L. G., Looney, L. W., Erickson, W., Grossman, A., Welch, W. J., Forster, J. R., Wright, M. C. H., Plambeck, R. L., Lugten, J., & Thornton, D. D. 1996, ApJ, 464, L169

[11] Ruden, S. P. & Pollack, J. B. 1991, ApJ, 375, 740

[12] Weidenschilling, S. J. 1977, Ap&SS, 51, 153

Chapter 3

Spectral Energy Distributions: Dependence on Viewing Angle Θ

This chapter draws from

Chiang & Goldreich 1999, ApJ, 519, 279

Abstract

We compute spectral energy distributions (SEDs) for passive T Tauri disks viewed at arbitrary inclinations. Semi-analytic models of disks in radiative and hydrostatic equilibrium are employed. Over viewing angles for which the flared disk does not occult the central star, the SED varies negligibly with inclination. For such aspects, the SED shortward of $\sim 80\ \mu\mathrm{m}$ *is particularly insensitive to orientation, since short wavelength disk emission is dominated by superheated surface layers which are optically thin. The SED of a nearly edge-on disk is that of a class I source. The outer disk occults inner disk regions, and emission shortward of* $\sim 30\ \mu\mathrm{m}$ *is dramatically extinguished. Spectral features from dust grains may appear in absorption. However, millimeter wavelength fluxes decrease by at most a factor of 2 from face-on to edge-on orientations.*

We present illustrative applications of our SED models. The class I source 04108+2803B is considered a T Tauri star hidden from view by an inclined circumstellar disk. Fits to its observed SED yield model-dependent values for the disk mass of $\sim 0.015\ M_\odot$ *and a disk inclination of* $\sim 65^\circ$ *relative to face-on. The class II source GM Aur represents a T Tauri star only slightly obscured by its circumstellar disk. Fitted parameters include a disk mass of* $\sim 0.050\ M_\odot$ *and an inclination of* $\sim 60^\circ$*, where the viewing angle is chosen to reproduce the observed visual extinction of* $A_V = 0.5$ *mag.*

3.1 Introduction

In the previous chapter, we computed SEDs of passive, flared, radiative equilibrium disks viewed face-on. The calculated SED was fairly constant over the thermal infrared (IR). Spectral features from dust grain resonances appeared in emission.

This chapter relaxes the assumption of a face-on viewing geometry. In §3.2, we compute the continuum spectrum of a passive disk whose midplane is inclined at an arbitrary angle, θ, to the plane of the sky. In §3.3, we study how spectral signatures of dust grain resonances change from emission to absorption as the disk is viewed increasingly edge-on. In §3.4, we apply our semi-analytic models to observations of the continuum SEDs of GM Aur (a class II system) and 04108+2803B (a class I system).[1] In particular, we assess the possibility that the relatively low 12 μm flux of GM Aur might be caused by a non-zero disk inclination, as opposed to a central AU-sized gap. We also discuss the extent to which differences between class I and class II SEDs might reflect differences in viewing angle rather than evolutionary status. Fits to more class II spectra using more sophisticated, numerical models are presented in Chapter 5.

3.2 Continuum SED vs. Θ

3.2.1 Model Ingredients

We employ the flared, radiative equilibrium disk model described in §2.4.2. The occulting angle,

$$\theta_0 \equiv \arctan \frac{a_o}{H(a_o)} = 45^\circ , \tag{3.1}$$

is a natural angle with which to compare the viewing angle θ. For $\theta \lesssim \theta_0$, both the central star and most of the disk surface are in direct view. For $\theta \gtrsim \theta_0$, star and inner disk are blocked from view by the flared outer "wall."

[1] In the classification scheme of Lada & Wilking (1984; see also Lada 1987), class I sources exhibit SEDs which rise from 2 to 10 μm, i.e., $-3 \lesssim n \lesssim 0$ where $\nu F_\nu \propto \nu^n$. For class II sources, $0 \lesssim n \lesssim 2$.

3.2.2 Interior SED

For clarity, we first consider only the disk interior. Figure 3.1 depicts the viewing geometry and coordinate system. Define x and y to be orthogonal spatial axes in the plane of the sky, centered on the star. The SED is the integral of specific intensity over projected disk area, viz.,

$$L_\nu \equiv 4\pi d^2 \nu F_\nu = 8\pi\nu \int_0^{a_o} dx \int_{-y(x)}^{y(x)} dy\, I_\nu\,, \tag{3.2}$$

where

$$y(x) = \sqrt{a_o^2 - x^2}\cos\theta + H(a_o)\sin\theta \tag{3.3}$$

traces the disk boundary on the sky, d measures distance to the source from Earth, and $I_\nu(x,y)$ is the specific intensity;

$$I_\nu = I_{\nu,i} = \kappa_V \varepsilon_\nu \int_0^\infty dl\, \rho_{gi}\, B_\nu(T_i)\, \exp\left(-\tau_\nu\right), \tag{3.4}$$

with

$$\tau_\nu = \kappa_V \varepsilon_\nu \int_0^l d\tilde{l}\, \rho_{gi}\,. \tag{3.5}$$

Here B_ν is the Planck function, and l, $\tilde{l}$ both measure line-of-sight distance from the observer. Since temperature, density, and opacity are specified by our model, calculating the SED at arbitrary inclination is primarily an exercise in multidimensional integration. We perform these integrals numerically using standard techniques (Press et al. 1992). Optical depths are evaluated using a Romberg integrator with a mandated fractional accuracy of 10^{-3}. Specific intensities are computed on an adaptive stepsize grid, where the source function (B_ν) changes by no more than 20% between adjacent grid points. Integrals over y are performed using either a Romberg integrator or, in cases where the integrand is very sharply peaked and the Romberg integrator takes prohibitively many steps, a 60-point Gaussian quadrature routine. The final integration over x employs the trapezoidal rule on a 60-point logarithmic grid. Answers are routinely checked for robustness against changes

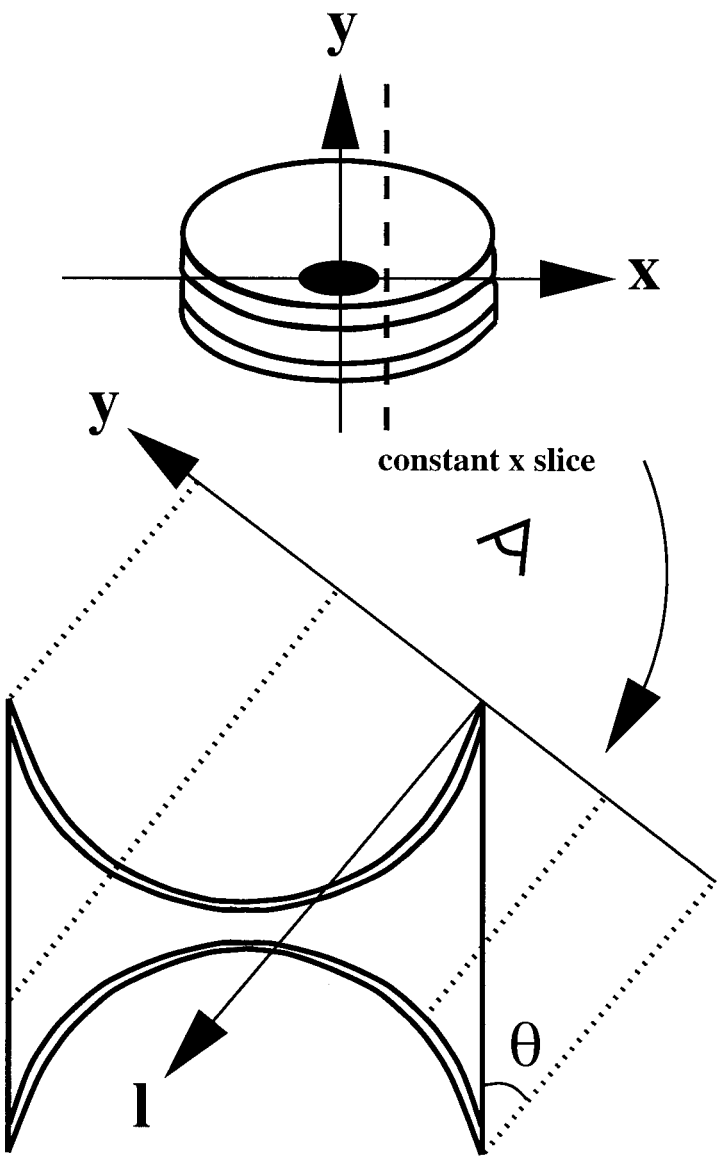

Figure 3.1: *Schematic of viewing geometry and coordinate system for inclined disks. The disk interior is capped above and below by the superheated surface. To calculate the SED, specific intensities along lines-of-sight (dotted lines parallel to l) are summed over the projected disk area in x and y. Here θ is slightly less than the occulting angle θ_0.*

in stepsize. We verify that the SED for $\theta = 0°$ computed using the multi-dimensional integration code matches the SED for face-on disks computed using the one-dimensional integrator employed in §2.4.2.

Figure 3.2 displays a family of interior SEDs for different θ. First consider SEDs for angles $\theta \leq \theta_0\,(\theta = 0°, 30°, 45°)$. Radiation at wavelengths shorter than $\sim$100 μm comes mainly from optically thick regions; consequently, L_ν is approximately proportional to the areal projection factor $\cos\theta$. At longer wavelengths, the radiation is emitted by increasingly transparent material, so that L_ν tends to be independent of orientation.

For angles $\theta > \theta_0\,(\theta = 65°, 85°)$, short wavelength emission from the inner disk is strongly extincted by the intervening outer disk. However, millimeter (mm) wavelength fluxes only drop by a factor of 1.8 as the disk is viewed increasingly edge-on, since about half of the outer disk remains visible in that limit.

3.2.3 Superthermal + Interior SED

Next we compute the total SED. The superheated layer is treated as a plane-parallel atmosphere having visual optical depth α normal to the surface. Whenever a line-of-sight intersects the surface, we increment the specific intensity from the interior by[2]

$$\Delta I_\nu = B_\nu(T_{ds}) \left[1 - \exp\left(-\frac{\alpha\varepsilon_s}{|\hat{n}\cdot\hat{l}|}\right)\right] \exp(-\tau_{\nu,l})\,. \tag{3.6}$$

Here, $\hat{n}$ and $\hat{l}$ are unit vectors normal to the surface and parallel to the line-of-sight, respectively, ε_s is the Planck-averaged dust emissivity at the surface, and $\tau_{\nu,l}$ is the intervening optical depth between disk surface and observer.

Figure 3.3 displays a family of total disk SEDs labeled by θ. The behavior of the SED for $\lambda \gtrsim 80\,\mu$m is similar to that described in §3.2.2, since the interior dominates emission at those wavelengths.

Shortward of $\sim$80 μm, emission from the superheated surface makes a qualitative dif-

[2]This prescription fails for lines-of-sight which graze the surface tangentially. However, their fractional contribution to the integrated flux is of order $\sqrt{h_s/R}$, where $h_s \approx (\partial \ln \rho_{gi}/\partial z)^{-1}|_{z=H}$ is the geometrical thickness of the surface, and R is the radius of curvature of H. Since this fraction is less than 10%, for computational simplicity we let the intersected optical depth of the surface be $\alpha\varepsilon_s \times \min(|\hat{n}\cdot\hat{l}|^{-1}, 20)$.

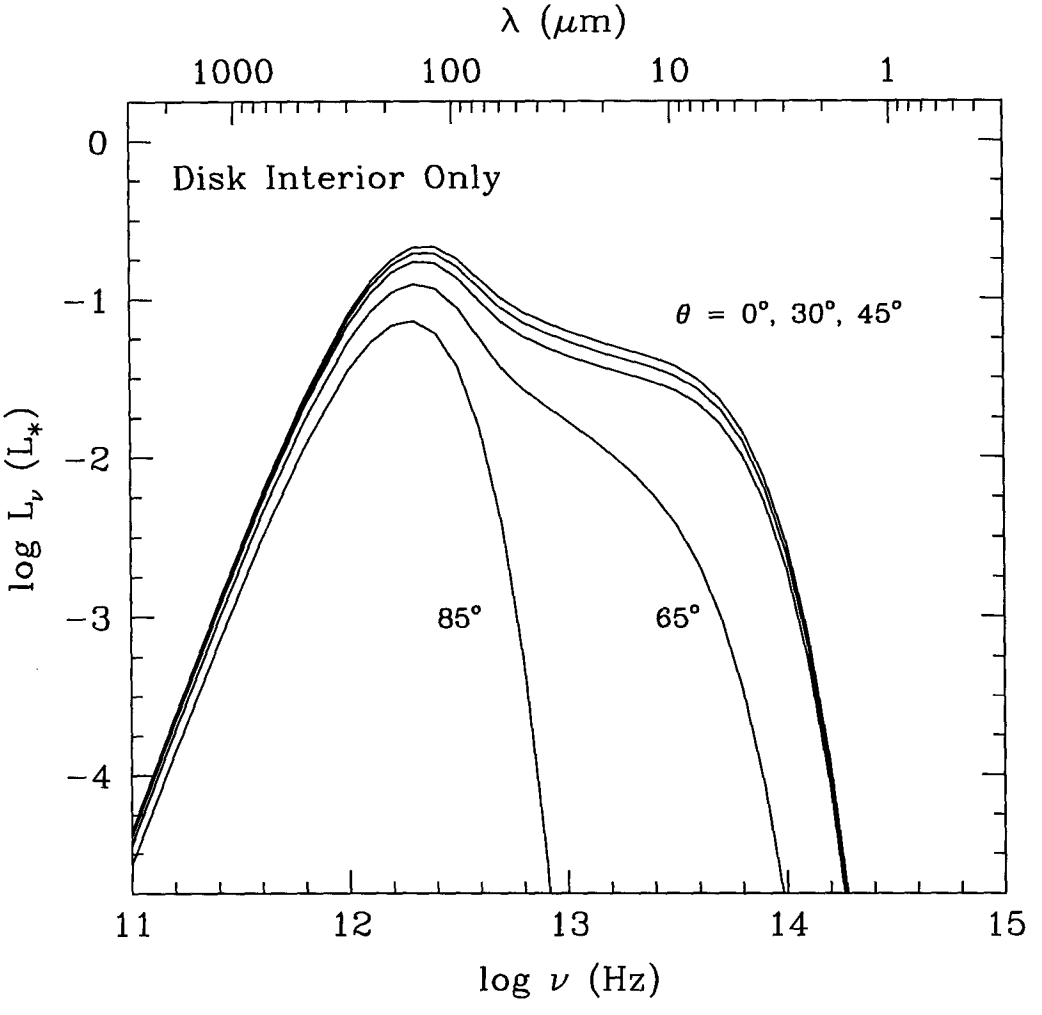

Figure 3.2: *SED of the disk interior only, as a function of viewing angle.*

Short wavelength emission from the inner disk is extinguished by the intervening outer disk for $\theta > \theta_0$. L_* denotes the luminosity of the central star.

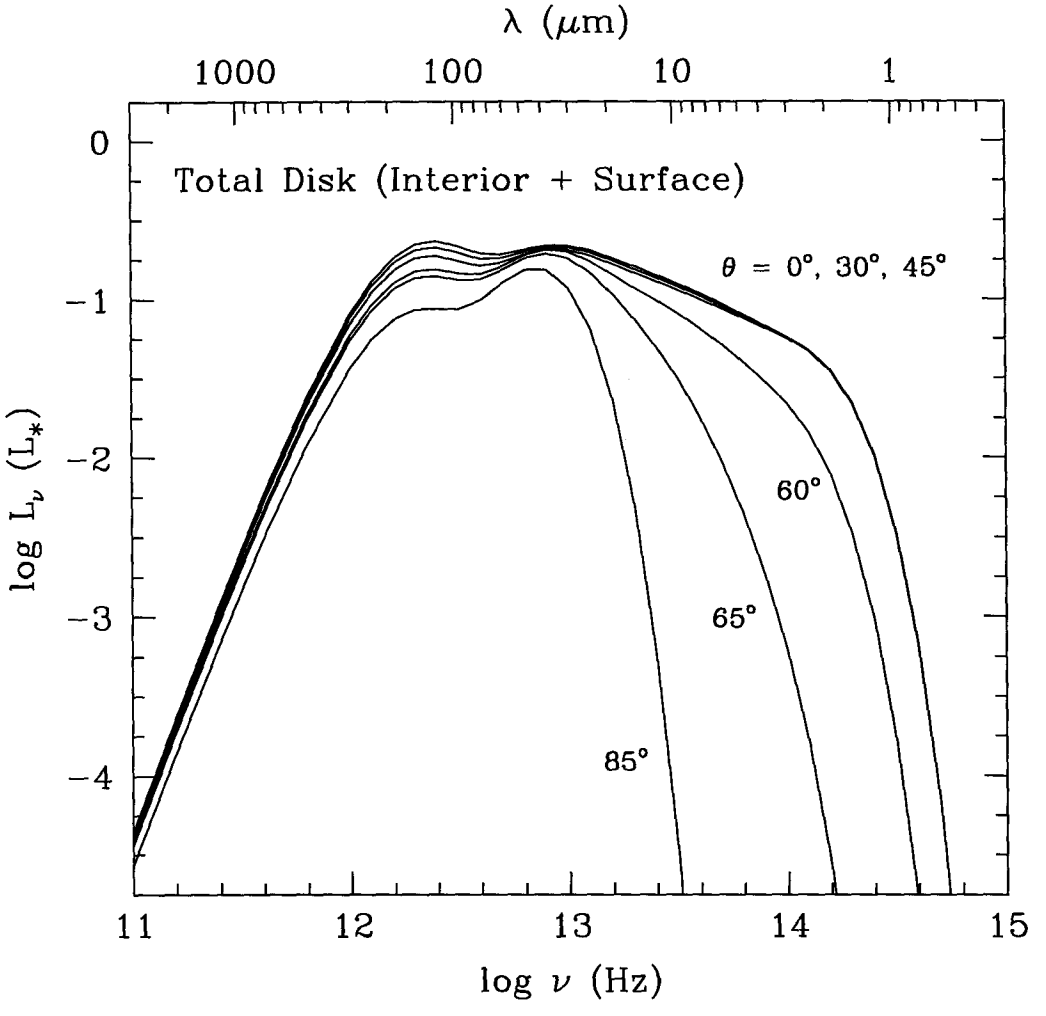

Figure 3.3: *SED of the entire disk (interior + surface), as a function of viewing angle. For $\theta < \theta_0$, the SED is insensitive to viewing aspect since short wavelength radiation originates from the optically thin surface, while long wavelength radiation is emitted by the optically thin interior. For $\theta > \theta_0$, the intervening outer disk dramatically extinguishes surface radiation at the shortest wavelengths.*

ference to the appearance of the SEDs. One difference is that for angles $\theta \leq \theta_0$, the flux is remarkably insensitive to viewing geometry.[3] This is because shorter wavelength radiation emerges primarily from the superheated surface layers, which are optically thin along most lines of sight.

For angles $\theta > \theta_0$, the intervening outer disk dramatically extinguishes surface radiation at the shortest wavelengths. However, the total disk SED between 30 and 80 μm is relatively robust to changes in inclination; it decreases by less than a factor of 3 between face-on and edge-on orientations. Radiation at these wavelengths originates from surface layers at large radii and is only slightly obscured by the tenuous interior at still larger radii.

Finally, in Figure 3.4, we add the contribution from the central star to the SED. The stellar flux fades rapidly with increasing θ beyond $\sim 55°$.

3.3 Spectral Dust Features vs. Θ

When viewed nearly face-on, passive disks exhibit emission lines associated with dust grain resonances (Calvet et al. 1992; Chiang & Goldreich 1997; see Chapter 2, §2.6). These features arise from material along the line-of-sight which is optically thin: line-to-continuum flux ratios are of the same order as the percentage enhancement in line opacity. Emission lines may originate from both the optically thin superheated surface and the optically thin interior.

How do such lines vary as the disk is viewed increasingly edge-on? Might they appear in absorption instead, as Cohen & Witteborn (1985) find for 7 of the 32 T Tauri stars they survey spectrophotometrically? To address these questions, we add 6 Gaussian-shaped spectral resonances to our grain emissivity law. In choosing resonant wavelengths and strengths, we aim to illustrate the range of effects which occur with variable viewing angle. Roughly, our 6 lines can be associated with crystalline H_2O ice at 60 and 100 μm (Waters & Waelkens 1998), amorphous silicates at 10 and 20 μm (Mathis 1990), and small carbon-rich grains or PAHs (polycyclic aromatic hydrocarbons) at 3.3 and 6.2 μm (Waelkens et

[3] Slight reductions in mid-infrared flux still occur for $\theta \leq \theta_0$, both because surface layers intermediate in radius are seen through the outer disk's atmosphere and because emission from the optically thick interior is proportional to the projected area.

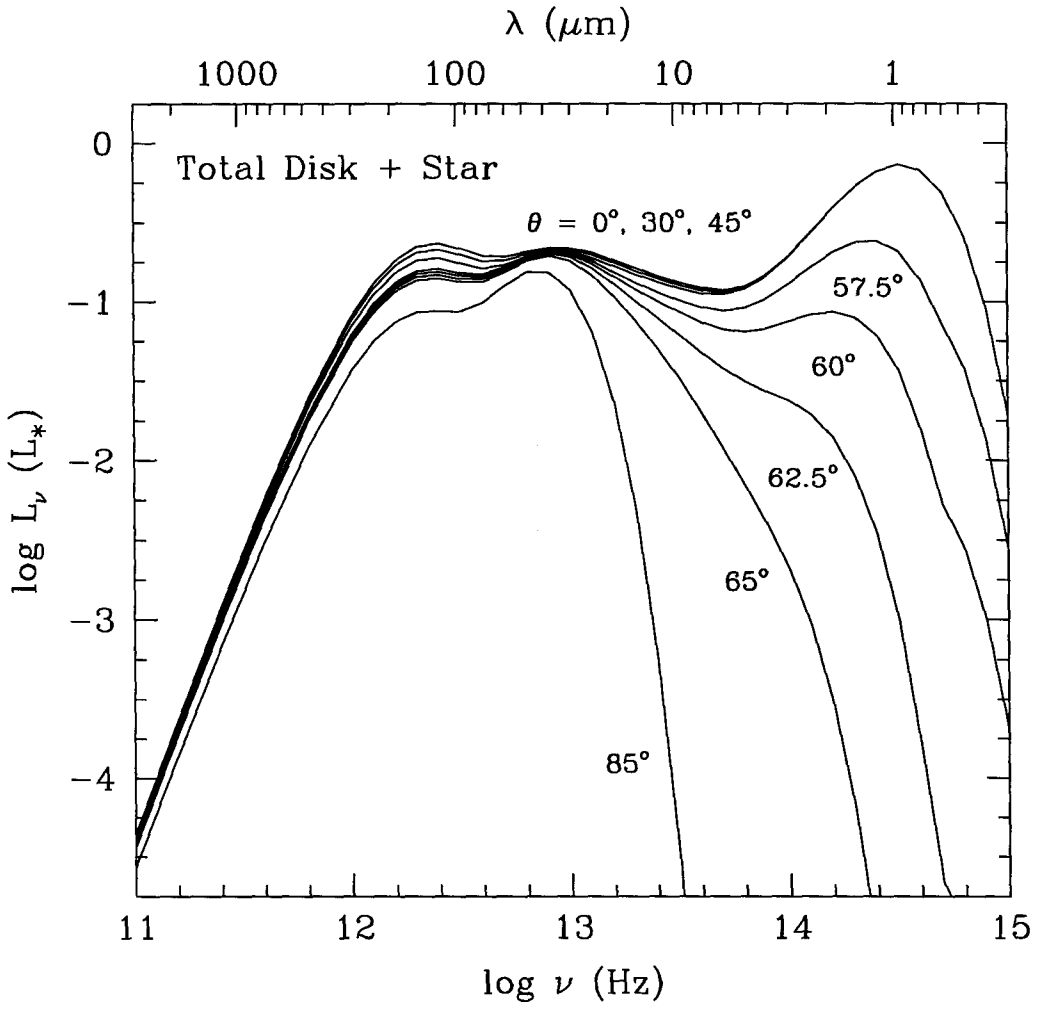

Figure 3.4: *SED of the entire disk + central star, as a function of viewing angle. At* $\theta = 60°$, *the visual optical depth to the star is* $\tau_{*,V} = 4.5$. *At* $\theta = 65°$, $\tau_{*,V} = 22$. *At* $\theta = 85°$, $\tau_{*,V} = 7.6 \times 10^5$.

al. 1996).

Solid-state line spectra at various disk inclinations are displayed in Figure 3.5. Emission at all wavelengths varies little with viewing geometry for angles $\theta \lesssim \theta_0$. This is true particularly at the shortest wavelengths where lines arise largely from the optically thin surface layers. At wavelengths progressively longer than $\sim$30 μm, the optically thin interior at large radius contributes increasingly to line emission. The drop in the long wavelength continuum with increasing inclination reflects the $\cos\theta$ dependence of radiation emitted by the optically thick interior.

Emission lines at $\sim$60 and 100 μm persist at angles $\theta \gtrsim \theta_0$, although their fluxes are lower by a factor $\sim$2 for edge-on as compared to face-on disks. For edge-on orientations, only about half of the optically thin interior remains visible. Emission around 20 μm is more robust to changes in inclination; most of this radiation emerges from the superheated surface at large radii and passes through the rarefied disk interior at yet larger radii with little attenuation.

At the shortest wavelengths, where occultation of the inner disk is significant, lines indeed appear in absorption, though only over a limited range of inclinations between θ_0 and 90°. For a dust grain resonance to appear in absorption, disk inclinations must be sufficiently high that line emission from the superheated surface is extinguished. However, θ cannot be so high that line and continuum optical depths both exceed unity at large radius where the disk interior is nearly isothermal. Only over an intermediate range of inclinations does radiation on and off the resonant wavelength probe a variety of interior disk temperatures. For the 3.3 μm resonance to appear in absorption, $60° \lesssim i \lesssim 65°$. At 6.2 μm, opacities are lower, and inclinations needed for absorption higher: $65° \lesssim i \lesssim 70°$. In principle, near-infrared line spectra may provide a sensitive diagnostic of disk inclination.

While the trends outlined above apply qualitatively to all passive disks, quantitative conclusions are model-dependent. For example, in our standard model, no absorption line at 10 μm appears because line-of-sight optical depths are too small to extinguish emission from the superthermal surface. However, such an absorption line would appear in the SED of a sufficiently massive, inclined disk.

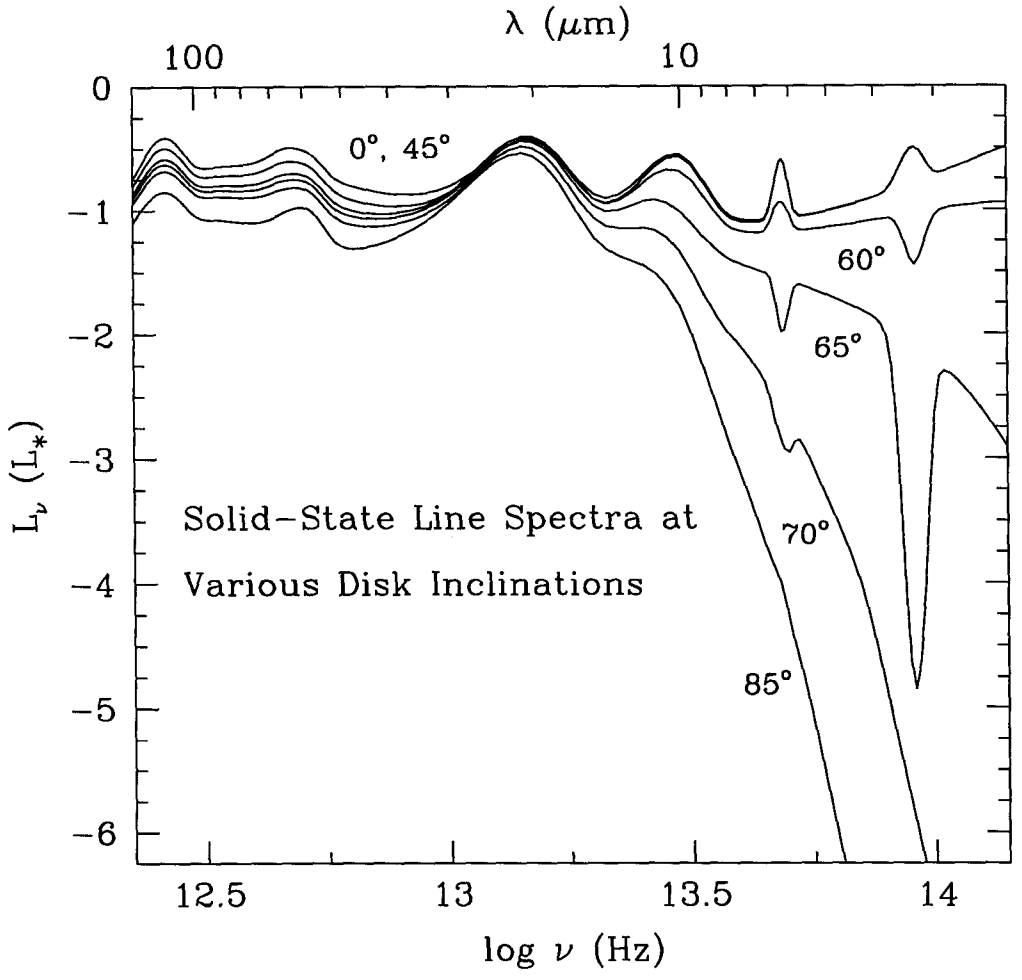

Figure 3.5: *Solid-state line spectra at various disk inclinations.*

Resonances at the shortest wavelengths where opacities are highest appear in absorption over a limited range of viewing angles between θ_0 and 90°. Longer wavelength features persist in emission.

3.4 Two Instructive Applications to Observations

3.4.1 GM Aurigae and Central Holes

GM Aur is a particularly clean, well-studied T Tauri system to which we can apply our SED models. The single central star has mass $0.72M_{\odot}$, luminosity $0.7L_{\odot}$, and age 2×10^6 yr (Beckwith et al. 1990). Aperture synthesis maps in $^{13}CO(2 \rightarrow 1)$ evince a rotating circumstellar gas disk having approximate projected dimensions 950×530 AU, inclined at $\theta \approx 30°$ (Koerner, Sargent, & Beckwith 1993). Hubble Space Telescope (HST) images in scattered visible light reveal the disk surface to be flared, and suggest i is closer to $\sim 60°$ (Koerner 1997; Stapelfeldt et al. 1995). These same observations detect no outflow or remnant envelope surrounding this relatively evolved system.

Figure 3.6 displays our fit to the SED of GM Aur with a disk inclined at $\theta = 59°$. Fluxes at IRAS (Infrared Astronomical Satellite) wavelengths emerge mainly from the optically thin surface layers, and those at mm wavelengths emerge mostly from the optically thin interior. Neither of these fluxes is sensitive to viewing angles θ near or less than θ_0.[4] Therefore, our fitted inclination does not reflect the shape of the infrared SED; rather, it is chosen to reproduce the observed modest visual extinction to the central star of $A_V = 0.5$ magnitudes and to conform with the imaging data.

Matching the visual extinction comes at the cost of requiring a large central gap having a radius ~ 60 times greater than the dust sublimation radius: $a_i = 4.8 \pm 2.7$AU, where the uncertainty reflects that of the $12\,\mu$m IRAS point. There is no inclination for which the outer edge of our model disk attenuates the $12\,\mu$m flux without also extinguishing central starlight by several magnitudes. Line-of-sight column densities to surface regions inside ~ 5 AU (where most of the $12\,\mu$m emission originates) are only $\sim 20\%$ greater than the column density to the star. This small difference in the amount of obscuring material is completely insignificant; the opacity at $12\,\mu$m is 13 times lower than at visible wavelengths. Thus, the low visual extinction to the central star implies that we should have a clear view of disk regions inside a few AU at near-infrared wavelengths.

In spite of these considerations, we remain skeptical of the existence of such a large cen-

[4] θ_0 for this disk equals $51°$.

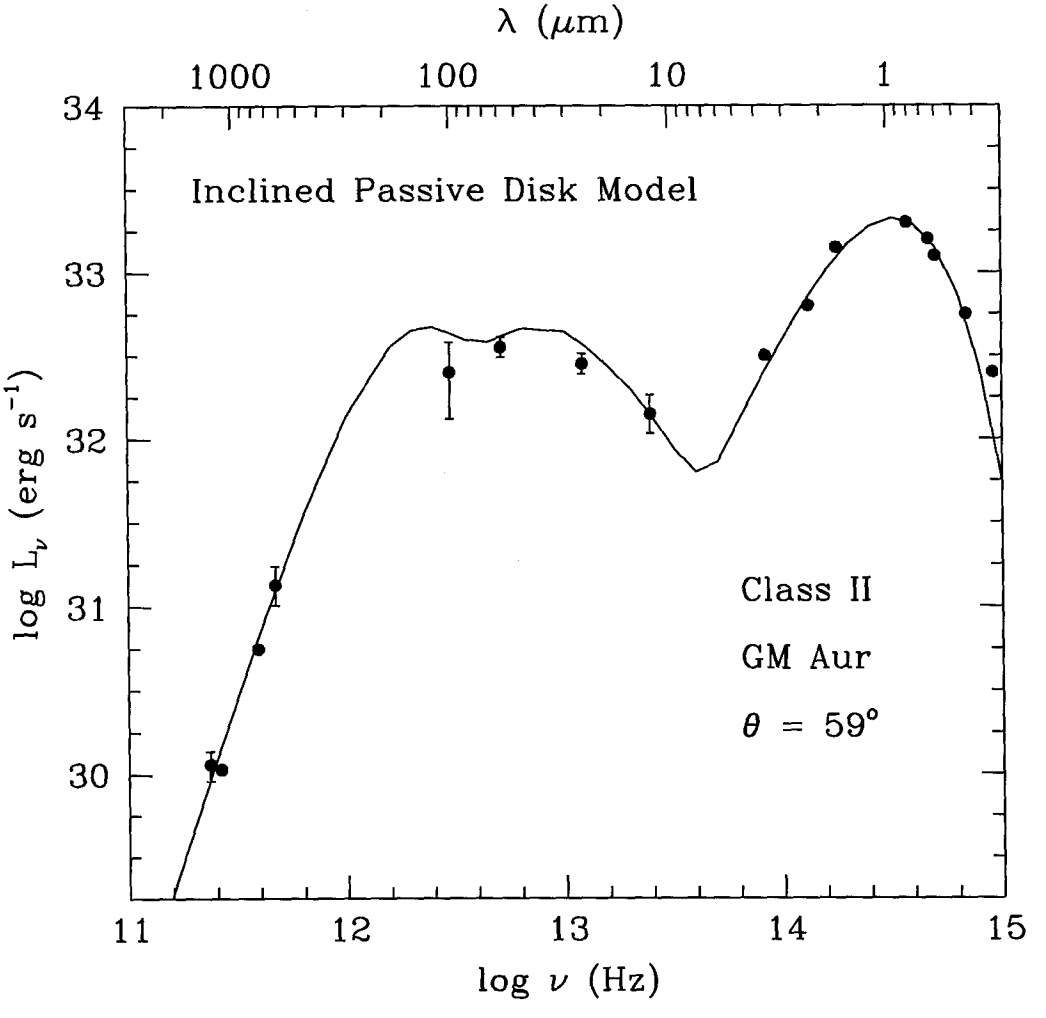

Figure 3.6: *Observed SED (filled circles) of GM Aur and accompanying inclined passive disk model.*

Fit parameters are as follows: $\theta = 59°$, $\Sigma = 2 \times 10^3 \, a_{\rm AU}^{-3/2}$ g cm^{-2}, $\beta = 1.4$, $r = 0.3\,\mu$m, $a_o = 311$ AU *(which implies* $\theta_0 = 51°$*), and* $a_i = 4.8$ AU. *The formal requirement of a large AU-sized central gap persists despite accounting for a non-zero viewing angle.*

tral gap devoid of dust. A more palatable alternative might be that the inner disk's aspect ratio does not increase monotonically with radius. Such undulations in the height of the surface would shadow annular regions from the central star and lower their temperatures. In addition, those ripples nearest the observer would hide their warmer, starlit sides when viewed at non-zero inclination. Accretional heating may be responsible for such changes in surface geometry, as discussed in the next chapter (§4.6); as estimated there, enhancements in disk thickness due to heating of the midplane may become significant inside a few AU. For an accretion disk which derives its luminosity solely from local viscous dissipation, the aspect ratio indeed decreases with radius whenever opacities increase steeply with temperature, as demonstrated by Bell et al. (1997).

3.4.2 Class I Sources as Inclined Class II Sources

Might some class I spectra represent T Tauri stars obscured by inclined disks? The possibility should be entertained for sources such as 04108+2803B; Figure 3.7 demonstrates that a passive disk inclined at 65° provides a reasonable fit (to within factors of 2) to this class I SED. Fit parameters are similar to those of our standard model, and are listed in the figure caption. For this particular model, near-IR emission at J, H, and K is interpreted as central starlight extincted by the disk's outer edge; inclusion of starlight scattered off the disk surface would imply a larger inclination. To highlight the contribution from the superheated surface, we also plot the SED with the surface emission removed. Fluxes between 3 and 60 μm arise primarily from the superthermal surface, seen both through and over the lip of the outer disk atmosphere (refer to Figure 3.1).

Clearly, a class I SED does not imply a unique distribution of circumstellar material. High angular resolution images provide additional clues. Many class I sources exhibit near-IR nebulosity on scales ranging from 1500 to 3000 AU (Kenyon et al. 1993; and references therein); for these sources, an additional, non-disk component of dust is needed. Kenyon, Calvet, & Hartmann (1993) model class I SEDs using rotationally flattened, infalling envelopes that are passively heated by central stars. These envelopes extend out to larger radii than do our disks, typically 3000 AU, but contain roughly the same amount of

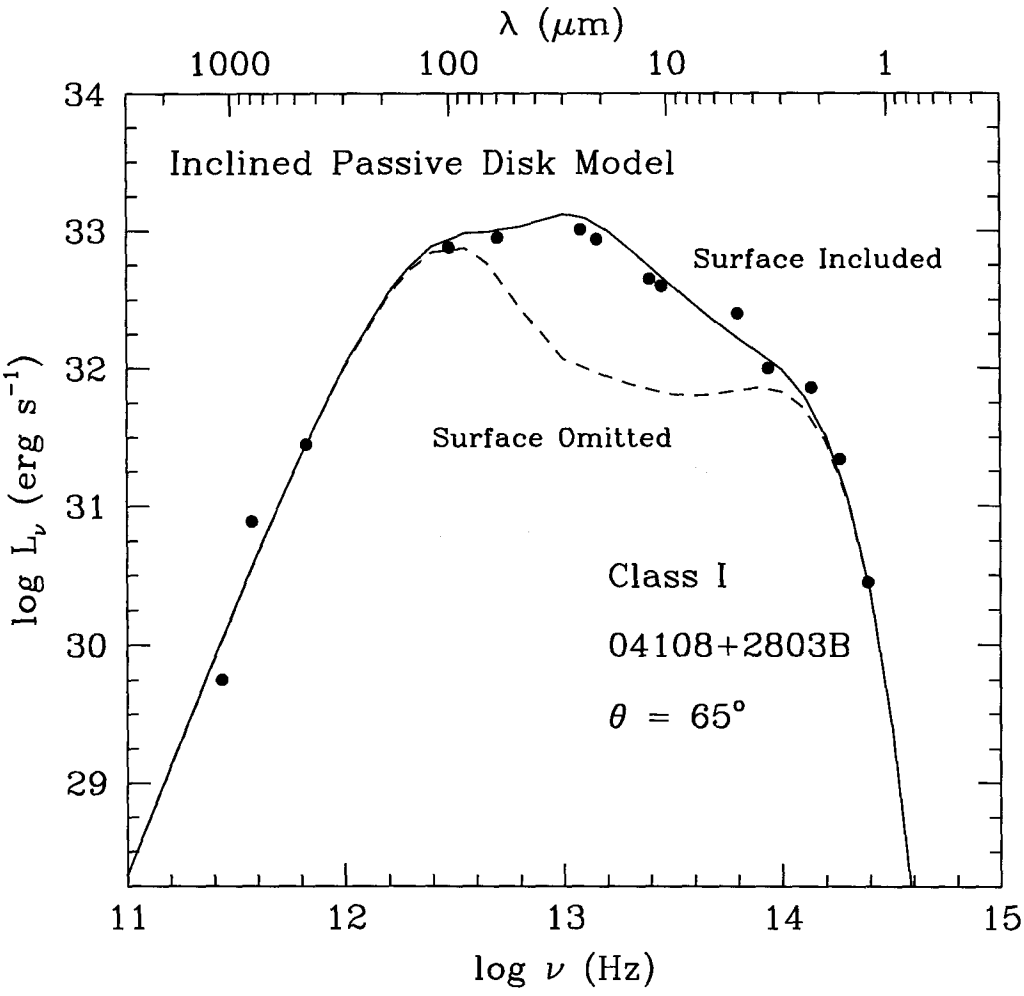

Figure 3.7: *Observed SED (filled circles) of class I source 04108+2803B (Kenyon, Calvet, & Hartmann 1993) and accompanying inclined passive disk model.*

Fit parameters for the disk are as follows: $\theta = 65°$, $\Sigma = 0.3 \times 10^3\, a_{\rm AU}^{-3/2}$ g cm^{-2}, $\beta = 1.2$, $r = 0.1\ \mu$m, $a_o = 270$ AU *(which implies* $\theta_0 = 46°$*), and* $a_i = 0.07$ AU. *Stellar parameters are identical to those of our standard model T Tauri star.*

mass, about $0.1 M_{\odot}$. Bipolar holes empty of material, presumably evacuated by outflows, are invoked so that central starlight may scatter off cavity walls towards Earth, thereby explaining the observed near-IR fluxes. In imaging observations of three class I sources using HST/NICMOS (Near Infrared Camera and Multi-Object Spectrometer), Padgett et al. (1998) discover nearly edge-on, flared circumstellar disks having diameters 300–750 AU. Near-IR emission is observed to be scattered not only off upper and lower disk surfaces, but also off "dusty material within or on the walls of outflow cavities."

In general, a combination of an inclined, passively heated disk, and a dusty bipolar outflow or partially evacuated envelope may best describe class I sources. Since 04108+2803B reveals no optical or near-IR emission beyond $\sim$140 AU from its central star (Kenyon, Calvet, & Hartmann 1993), it is consistent with being a limiting example of a simple inclined disk. Two other examples for which this limiting case scenario may also apply include the embedded sources 04295+2251 and 04489+3042. Both sources exhibit nearly flat excesses between $\sim$3 and 100 μm which may arise from flared circumstellar disks. Neither source betrays extended near-IR emission on scales greater than a few hundred AU, or high-velocity molecular gas from an outflow (Kenyon et al. 1993).

Near-IR polarimetry also addresses the possible presence of envelopes. Whitney, Kenyon, & Gómez (1997) determine that linear polarizations greater than $\sim$20% at J, H, and K characterize the largest reflection nebulae associated with class I sources. Such large polarizations are interpreted as arising from the scattering of central starlight off cavity walls in envelopes. In contrast, 04108+2803B evinces fractional polarizations of 5.1% and 1.6% at H and K, respectively (Whitney, Kenyon, & Gómez 1997)—levels more comparable to those in class II sources than in truly younger, more embedded protostars, and which may arise from starlight scattered off the disk surface. Resolved polarimetry maps of 04108+2803B can test our hypothesis.

Degeneracies inherent in models deduced from a SED may also be broken by kinematic studies. Interferometric radial velocity maps in molecular lines may distinguish between an infalling envelope (e.g., Hayashi, Ohashi, & Miyama 1993; Ohashi et al. 1997), and a rotating disk (e.g., Koerner & Sargent 1995).

Terquem & Bertout (1996) also independently model 04108+2803B as a T Tauri star

embedded within a dynamically warped disk. The antisymmetric warp considered by these authors is tidally induced by a hypothetical stellar companion in an orbit inclined to the primary's disk. High spatial resolution images (e.g., from speckle interferometry) can test for possible multiplicity and telltale signs in disk geometry.

Finally, we note that the inverse problem to 04108+2803B is presented by the star HK Tau B: a system observed to possess an edge-on ($i \approx 85°$) circumstellar disk and no observable envelope, but whose infrared SED is not well measured because of confusion from an infrared-bright companion (Stapelfeldt et al. 1998; Koresko 1998).

Acknowledgments

We thank Anneila Sargent for providing data on GM Aur and other T Tauri stars, David Koerner and Karl Stapelfeldt for informative discussions regarding GM Aur, and an anonymous referee for a careful reading of our paper and for providing suggestions that helped to improve its presentation. Financial support for this research was provided by NSF grant 94-14232, NASA grant NAG5-7008, and an NSF Graduate Fellowship.

Bibliography

[1] Beckwith, S. V. W., Sargent, A. I., Chini, R. S., & Gusten, R. 1990, AJ, 99, 924

[2] Bell, K. R., Cassen, P. M., Klahr, H. H., & Henning, Th. 1997, ApJ, 486, 372

[3] Calvet, N., Patiño, A., Magris, G., & D'Alessio, P. 1992, Rev. Mexicana Astron. Astrof., 24, 27

[4] Calvet, N., Hartmann, L., Kenyon, S. J., & Whitney, B. A. 1994, ApJ, 434, 330

[5] Chiang, E. I. & Goldreich, P. 1997, ApJ, 490, 368 (CG97)

[6] Cohen, M. & Witteborn, F. C. 1985, ApJ, 294, 345

[7] Hayashi, M., Ohashi, N., & Miyama, S. M. 1993, ApJ, 418, L71

[8] Kenyon, S. J., Calvet, N., & Hartmann, L. 1993, ApJ, 414, 676

[9] Kenyon, S. J., Whitney, B. A., Gomez, M., & Hartmann, L. 1993, ApJ, 414, 773

[10] Koerner, D. W. 1997, Origins of Life and Evolution of the Biosphere, 27, 157

[11] Koerner, D. W. & Sargent, A. I. 1995, Ap&SS, 223, 169

[12] Koerner, D. W., Sargent, A. I., & Beckwith, S. V. W. 1993, Icarus, 106, 2

[13] Koresko, C. D. 1998, ApJ, 507, L145

[14] Kusaka, T., Nakano, T., & Hayashi, C. 1970, Prog. Theo. Phys., 44, 1580

[15] Lada, C. J. 1987, in *IAU Symp. 115, Star Forming Regions*, ed. M. Peimbert & J. Jugaku (Dordrecht:Reidel), 1

[16] Lada, C. J. & Wilking, B. A. 1984, ApJ, 287, 610

[17] Mathis, J. M. 1990, ARA&A, 28, 37

[18] Ohashi, N., Hayashi, M., Ho, P.T.P., & Momose, M. 1997, ApJ, 475, 211

[19] Padgett, D., Brandner, W., Stapelfeldt, K., Koerner, D., Terebey, S., Strom, S., & Kenyon, S. 1998, American Astronomical Society Meeting, 192, #10.04

[20] Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 1992, *Numerical Recipes in Fortran* (Cambridge University Press)

[21] Stapelfeldt, K., Burrows, C. J., Koerner, D., Krist, J., Watson, A. M., Trauger, J. T., & WFPC2 IDT 1995, American Astronomical Society Meeting, 187, #113.04

[22] Stapelfeldt, K. R., Krist, J. E., Ménard, F., Bouvier, J., Padgett, D. L., & Burrows, C. J. 1998, ApJ, 502, L65

[23] Terquem, C. & Bertout, C. 1996, MNRAS, 279, 415

[24] Waelkens, C., et al. 1996, A&A, 315L, 245

[25] Waters, L.B.F.M. & Waelkens, C. 1998, ARA&A, 36, 233

[26] Whitney, B. A., Kenyon, S. J., & Gómez, M. 1997, ApJ, 485, 703

Chapter 4

Physics of Passive Disks: Extensions and Refinements

This chapter draws from

Chiang & Goldreich 1997, ApJ, 490, 368

Abstract

The underpinnings of our theory of passive disks are probed. Our principal findings are as follows. (1) The existence of an optically thin surface layer containing superheated dust grains depends on the the coupling between grains and molecular coolants being sufficiently weak. This criterion is normally satisfied in passive disk atmospheres for $a \gtrsim 1$ AU. *(2) Vertical settling of dust grains over estimated disk lifetimes causes the superheated layer to sit between* $\sim$*1 and* $\sim$*5 gas scale heights above the midplane. (3) Grain albedos are unlikely to be high enough to significantly affect the thermally emitted portion of the spectral energy distribution. (4) Passive disks appear to be unstable to perturbations in height and temperature in the dual limit that (a) the vertical thermal diffusion time is shorter than the local dynamical time and (b) the horizontal scale of the perturbation is longer than the vertical thickness of the perturbed region. Passive disks appear to be stable otherwise. The details and consequences of our newly discovered "ripple instability" remain to be worked out.*

4.1 Introduction

This chapter critically examines several of the assumptions made in Chapter 2.

In §4.2–§4.3, we study the microscopic thermodynamics of passive disk atmospheres. We explain in qualitative terms why irradiated disks are unlikely to behave as perfect blackbodies. A simple model which accounts for the exchange of energy between dust grains

and gas molecules in disk surface layers is set forth. The model elucidates the relationship between dust grain temperatures, gas kinetic temperatures, and molecular excitation temperatures.

In §4.4, we question the assumption of uniform mixing of dust and gas in interstellar proportions within circumstellar disks. In a quiescent disk, dust grains settle gravitationally towards the midplane, and the height of the superheated surface layer is likely to be lower than what our previous assumption of uniform mixing implies.

In §4.5, we consider the effects of non-zero grain albedo on the thermally emitted SED.

In §4.6, we briefly discuss the effects of internal disk heating if shear viscosity were present.

In §4.7, we delineate the various timescales over which passive disks respond to perturbations. This section furnishes the basis for our preliminary study of disk stability in §4.8.

Finally, unresolved issues are summarized in §4.9.

4.2 The Planetary Atmosphere Analogy

Planetary atmospheres are generally transparent to visible radiation above their uppermost clouds decks.[1] Visible radiation is absorbed either by atmospheric particulates (dust grains), or by the planetary surface itself in the case of the terrestrial planets. The absorbed energy is either re-emitted as thermal infrared radiation, or conducted by collisions to ambient gas molecules. At infrared wavelengths, the atmospheric gas is optically thick, even above the cloud tops; pressure broadened molecular lines are the main source of infrared opacity.[2] The total infrared luminosity of the planet—i.e., the sum of absorbed sunlight and internal heat loss—is emitted as approximate black body radiation. Thus the energy balance of a planetary atmosphere resembles that of the blackbody disk.[3] This analogy

[1]Column densities above the cloud tops are similar to that of the minimum mass solar nebula at $a_{\rm AU} \approx 1$, $\Sigma \sim 10^3\,{\rm g\,cm^{-2}}$.

[2]Collision induced dipole transitions of molecular hydrogen dominate in the atmospheres of the outer planets.

[3]The analogy even extends to the association of the planet's intrinsic luminosity with the accretion luminosity of the disk.

might account for the casual acceptance of the blackbody disk model.

However compelling, this analogy is flawed. T Tauri disks differ from planetary atmospheres in a crucial parameter, the vertical gravitational acceleration. For disks, $g_{disk} \approx \Omega^2 z$, with Ω the orbital angular velocity, and z the distance from the midplane. For our standard parameters, $g_{disk} \approx 0.4\, z_{\rm AU}\, a_{\rm AU}^{-3}\, {\rm cm\,s^{-2}}$, much smaller than $g_{planet} \sim 10^3\, {\rm cm\,s^{-2}}$ for $a \gtrsim a_i$. Three consequences of the low vertical gravitational acceleration in disks which invalidate the analogy with planetary atmospheres are described below.

Dust grains settle slowly in disks. The gravitational settling time to the midplane,

$$t_{settle} \approx \frac{\Sigma}{r \rho_d \Omega} \approx 10^7 \left(\frac{0.1\,\mu{\rm m}}{r}\right) \left(\frac{\Sigma_0}{10^3\,{\rm g\,cm^{-2}}}\right) {\rm yr}\,, \tag{4.1}$$

is independent of a. It is longer than the disk lifetime for particles which most efficiently absorb stellar radiation. Moreover, the settling may be slowed by turbulent mixing or vertical circulation of the gas.

The gas pressure in disks is so low that dust grains which are directly exposed to stellar radiation reradiate almost all of the stellar energy they absorb. Thus the temperature of the superheated grains is determined by radiative balance and is independent of the ambient gas temperature. The regulation of the gas temperature is a more subtle issue (see §4.3).

The low pressure in disks also means that molecular lines suffer negligible pressure broadening. Thus the gas opacity is concentrated in narrow, Doppler broadened lines. Important molecular sources of opacity must have permitted dipole transitions near the peak of the blackbody spectrum corresponding to the ambient temperature T. Furthermore, these transitions must connect energy levels which do not lie too far above kT. Water is a prime candidate for a coolant at temperatures greater than 100 K. It is abundant, and as a consequence of being both a hydride and an asymmetric rotor, it possesses a rich rotation spectrum shortward of 300 μm. However, our crude estimates suggest that even if much of the cosmic abundance of oxygen is tied up in water, water lines would not cover more than a small fraction of the infrared spectrum at any temperature.

4.3 Energy Balance in the Superthermal Dust Layer

Stellar energy absorbed by dust grains is lost through emission of infrared radiation and collisional transfer to gas molecules. The gas molecules, mostly H_2 and He, gain energy in collisions with dust grains and lose it in collisions with molecular coolants such as H_2O. The energy added to the internal degrees of freedom of the coolants in collisions with gas molecules and by absorption of dust radiation is lost in radiation to space. Thus $T_{ds} \gtrsim T_{gs} \gtrsim T_x$, where the subscripts $j = \{ds, gs, x\}$ denote quantities pertaining to the dust, the inert gas molecules, and the molecular coolants, respectively.[4]

Balancing the rates of energy gain and loss per unit surface area of a dust grain yields

$$\sigma T_*^4 \left(\frac{R_*}{a}\right)^2 \approx \varepsilon_s \sigma T_{ds}^4 + n_g v_g k (T_{ds} - T_{gs}) \,, \tag{4.2}$$

with v_g the sound speed. For $n_g \ll n_{c1}$, where the critical number density

$$n_{c1} \approx \frac{\varepsilon_s \sigma T_{ds}^3}{v_g k} \approx \frac{\sigma \mu_g^{1/2} T_*^{5/2}}{k^{3/2}} \left(\frac{R_*}{a}\right)^{7/5} \approx \frac{10^{14}}{a_{\rm AU}^{7/5}} \, {\rm cm}^{-3} \,, \tag{4.3}$$

gas-grain collisions have a negligible effect on T_{ds} and equation (4.2) reduces to equation (2.11).[5] On the other hand, for $n_g \gg n_{c1}$, $T_{ds} \approx T_{gs}$. The gas density in the superheated dust layer, n_{gs}, is unknown. It could be as large as n_0, the value at the disk midplane, or as low as n_{min}, the value obtained by assuming uniform mixing of dust and gas up to the tenuous outer layers of the disk. Runs of n_{c1}, n_0, and n_{min} as functions of a are displayed in Figure 4.1. It is seen that even n_0 falls below n_{c1} beyond $a_{\rm AU} \approx 2$.

Balancing the rates at which inert gas molecules gain and lose energy in collisions with dust grains and molecular coolants yields

$$n_{ds} \sigma_d (T_{ds} - T_{gs}) \approx n_x \sigma_x (T_{gs} - T_x) \,, \tag{4.4}$$

where n is a number density, and σ denotes a collisional cross section. The compositions and mixing ratios of molecular coolants are uncertain. Photodissociation and incorpora-

[4]The symbol T_x refers to the excitation temperature of the coolant transition.

[5]We evaluate v_g on the right-hand side of equation (4.3) at temperature T_{ds}.

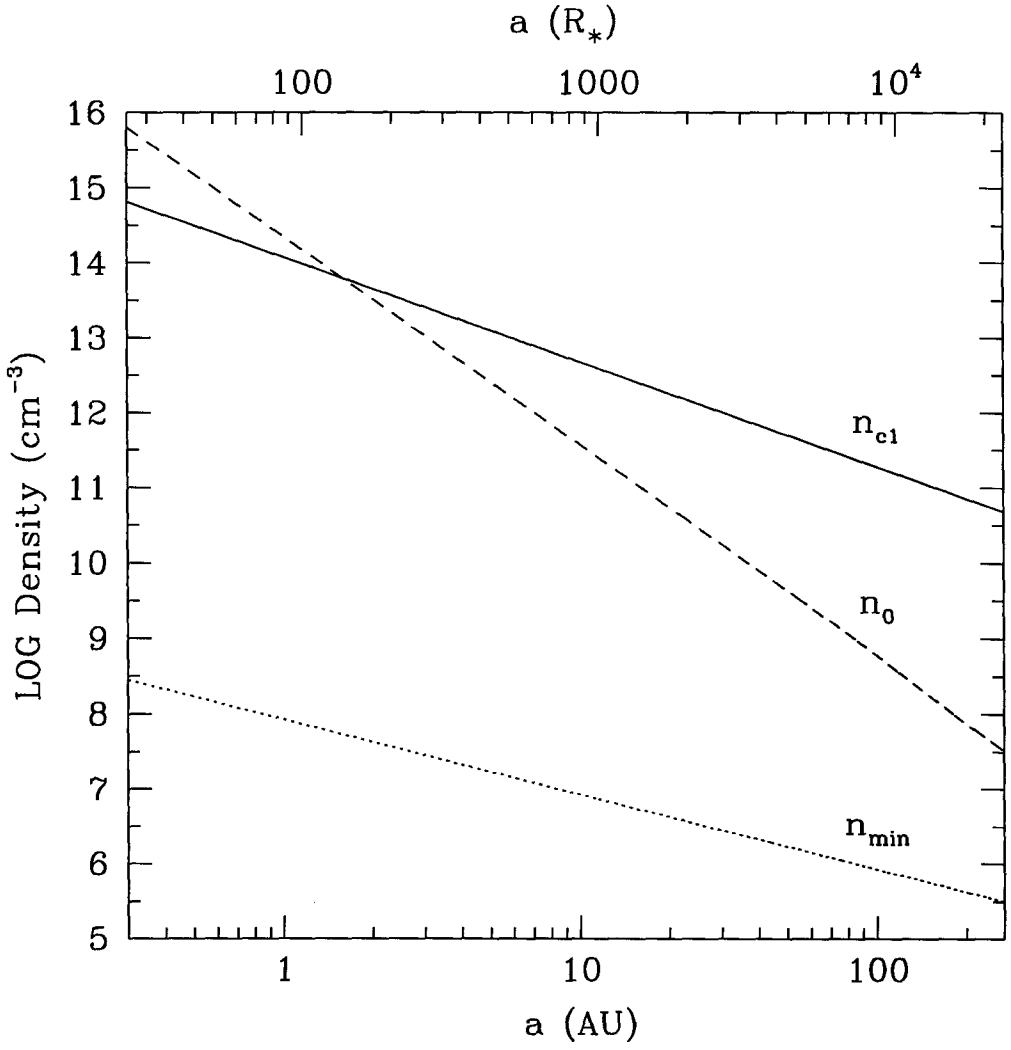

Figure 4.1: *Comparison of possible gas densities in the superheated layer to the critical density n_{c1} above which gas-grain collisions might lower the dust temperature.*

The actual gas density could be as high as n_0, the value at the disk midplane, or as low as n_{min}, the value for a disk in which gas and dust are uniformly mixed everywhere. Except in the innermost 2AU, *it is evident that the dust likely remains superheated at the temperature given equation (2.11).*

tion into dust grains represent loss mechanisms that can reduce abundance ratios far below values based on cosmic abundances. Nevertheless, suppose that the mass in coolant molecules is comparable to that in dust grains. Then $n_{ds}\sigma_d/n_x\sigma_x \sim 10^{-3}$, the ratio of the linear size of a gas molecule to a dust grain. Thus we would not be surprised to discover that $T_{gs} - T_x \ll T_{ds} - T_{gs}$.

Estimating excitation temperatures for the transitions in molecular coolants is a daunting task. We proceed by making a number of drastic approximations. Each transition is assumed to give rise to an optically thick, Doppler broadened line having excitation temperature T_x. For clarity, we begin by considering the pure radiative excitation of these levels. It is not difficult to show that the excitation comes mainly from radiation emitted by the superheated dust grains as opposed to that from the inner portion of the disk. We specialize further to transitions whose frequencies lie on the Rayleigh-Jeans part of the blackbody spectrum at temperature T_{ds}. Applying the basic concepts of the escape probability formalism for radiative transfer to a two level molecule, we obtain

$$T_x \approx \alpha\varepsilon_s T_{ds} \qquad \text{for} \quad h\nu \lesssim kT_x\,, \tag{4.5}$$

$$T_x \approx \frac{h\nu}{k}\ln\left(\frac{1}{\alpha\varepsilon_s}\frac{h\nu}{kT_{ds}}\right) \qquad \text{for} \quad h\nu \gtrsim kT_x\,. \tag{4.6}$$

Pure radiative excitation sets a floor on the value of T_x.

Next we consider the effects of collisions while neglecting radiative excitation. Here we suppose that lines of brightness temperature T_x cover a fraction f of the blackbody spectrum at temperature T_x. We equate the emission line flux to the rate (per unit disk surface area) at which the gas gains energy in collisions with the superheated dust grains:

$$f\sigma T_x^4 \approx n_{gs}v_g k\alpha(T_{ds} - T_{gs})\,. \tag{4.7}$$

For $n_{gs} \ll n_{c2}$, where the second critical gas density

$$n_{c2} \approx \frac{f}{\alpha\varepsilon_s}n_{c1}\,, \tag{4.8}$$

we find

$$T_x \approx \left(\frac{n_{gs}}{n_{c2}}\right)^{1/4} T_{ds} \,. \tag{4.9}$$

In the high density limit $n_g \gg n_{c1}$, $T_x \approx T_{gs} \approx T_{ds} \approx T_s$ with

$$T_s \approx \left(\frac{\alpha}{f + \alpha\varepsilon_s}\right)^{1/4} \left(\frac{R_*}{a}\right)^{1/2} T_* \,. \tag{4.10}$$

If $f \ll \alpha\varepsilon_s$, equation (4.10) reduces to (2.11), the appropriate expression for the dust temperature in the absence of cooling by molecular line emission. On the other hand, if $f \gg \alpha\varepsilon_s$, (4.10) collapses to the expression appropriate for a blackbody disk as given by equation (2.1), magnified by a factor of about $1/f^{1/4}$.

4.4 Location of the Superthermal Layer

We have been assuming that the dust and gas are uniformly mixed throughout the disk. Consequently, for our model the gas density is much smaller in the superthermal dust layer than in the disk midplane (cf. Fig. 4.1). Although the timescale for dust grains to settle to the disk midplane may be long compared to the disk lifetime [cf. equation (4.1)], the timescale for grains to fall through the superthermal layer (in the absence of advection by the gas) is considerably shorter. The latter is given by

$$t_{settle,s} \sim \frac{v_g}{ar\kappa_V \rho_d \Omega^2} \sim a_{\rm AU}^2 \left(\frac{0.1\ \mu{\rm m}}{r}\right) \left(\frac{400\ {\rm cm^2\ g^{-1}}}{\kappa_V}\right) {\rm yr} \,, \tag{4.11}$$

where κ_V is the dust opacity in the surface layer at visual wavelengths. The timescale for dust collisions/coagulation to occur in the superthermal layer is $1/\alpha$ times larger than $t_{settle,s}$. Thus, the superthermal dust layer is likely to be found closer to the disk midplane and to have a lower κ_V than our simple assumption of uniform dust-gas mixing implies.

How might we learn about the level at which the superthermal dust layer sits? One possibility is through the observation of emission lines that would form in the gas associated with this layer if the gas temperature is comparable to the dust temperature. Perhaps most

revealing would be observations of the rotation lines of H_2. Because these arise from quadrupole transitions and are therefore weak, they would only be seen if the dust layer were located where the gas density is much larger than the minimum value given by the assumption of uniform mixing of dust and gas.

4.5 Dust Albedo

A nonzero dust grain albedo, A_d, at visible wavelengths would reduce the absorbed stellar flux by a factor of order $(1 - A_d)^{1/2}$. This would lower the disk thickness, but by a less impressive factor of order $(1 - A_d)^{1/16}$.

How small might $(1 - A_d)^{1/2}$ be? Values of $A_d \sim 0.4 - 0.5$ are typical of interstellar grains (Draine & Lee 1984). Visual albedos of pure ice particles are much closer to unity. And ices are likely to comprise a significant fraction of the mass in grains in the outer disk. Starting inside $a_{\rm AU} = 10^2$, the partial pressures of H_2O, then NH_3, and finally CH_4 are expected to exceed their vapor pressures. However, taking the outer planet atmospheres as a guide, it seems unlikely that any ice so formed would be pure enough to yield $(1-A_d)^{1/2} \lesssim 0.3$.

4.6 Accretional Heating

A star which derives its radiant energy from accretion has $T_* \approx (GM_*\dot{M}_*/R_*^3\sigma)^{1/4}$, where $\dot{M}_*$ is the mass accretion rate.[6] The stellar flux impinging on the circumstellar disk, $\sigma T_*^4(R_*/a)^2\alpha$, is to be compared with the local viscous dissipation rate per unit area, $\sigma T_*^4(R_*/a)^3$. Thus in the extreme case that accretion accounts for the entire stellar luminosity, viscous dissipation is only competitive with stellar heating for $a \lesssim a_{tr} \approx 1\,{\rm AU}$ [see the discussion following equation (2.10)].

The bulk of the viscous dissipation is commonly thought to occur deep inside the disk. Provided the disk is sufficiently opaque, this could result in midplane temperatures being higher than $T_*(R_*/a)^{3/4}$ by a factor of order $\tau^{1/4}$, where $\tau \approx \varepsilon_i \kappa_V \Sigma$ is the effective vertical

[6]Provided the accretional energy is radiated over the entire stellar surface.

optical depth. A concomitant increase in the disk thickness by a factor of $\tau^{1/8}$ changes the surface geometry and may cause some portions of the disk to be shadowed from direct exposure to starlight. Such effects are likely to be important only for the innermost regions of the disk where radiation temperatures, surface densities, and consequently optical depths are highest. For example, around an accretion powered star, the midplane temperature of our standard disk is about 2000 K at $a_{\rm AU} = 1.5$. This yields about a factor 2 increase in disk thickness relative to that of the passive disk. For smaller radii, the enhanced thickness would be buffered somewhat by the reduction of opacity associated with the vaporization of silicates.

Bell et al. (1997) have calculated the geometry of a protoplanetary accretion disk which derives its luminosity solely from viscous dissipation at the midplane. They find that the aspect ratio, h/a, decreases with radius, a, at stellocentric distances of a few AU because the dominant opacities are molecular in origin and consequently increase steeply with temperature. Thus, to the extent that the height of the directly exposed surface, H, is proportional to the gas scale height, h, portions of the disk at distances of $\sim$1–10 AU might indeed be cast into shadow.

4.7 Temporal Behavior

Temporal variations would provide another diagnostic of circumstellar disks. Of relevance here is the claim by Moriarty-Schieven & Butner (1997) that the submillimeter and millimeter fluxes from the T Tauri binary GG Tau increased by factors of order 2 between 1992 and 1994. The cause of this "radio-wave flare" has not been identified. A plausible hypothesis is that it resulted from enhanced disk heating associated with a burst in luminosity originating near one or both components of the central binary. This leads us to consider relevant timescales for the radiative and hydrostatic response of the disk.

Seven different timescales come into play. They are: (1) the timescale over which superheated dust grains in the surface layer equilibrate with the ambient stellar radiation field,

$$t_{ds} \sim \frac{r\rho_d k T_{ds} a^2}{\mu_d R_*^2 \sigma T_*^4} \sim 0.02 a_{\rm AU}^{8/5}\,{\rm s}\,, \tag{4.12}$$

where $\mu_d \approx 10\mu_g$ is the mean molecular weight per degree of freedom in a dust grain; (2) the light-travel timescale from star to disk;

$$t_{lt} \sim \frac{a}{c} \sim 5 \times 10^2 a_{\rm AU}\,{\rm s}\,; \tag{4.13}$$

(3) the photon diffusion timescale through the disk,[7]

$$t_{diff} \sim \frac{\varepsilon_i \kappa_V \Sigma h}{c} \sim \frac{2 \times 10^5}{a_{\rm AU}^{9/14}}\,{\rm s}\,; \tag{4.14}$$

(4) the thermal timescale for the dust if it is decoupled from the gas,[8]

$$\begin{aligned} t_{di} &\sim \frac{\Sigma_d k}{\mu_d \sigma T_i^3}\left(\varepsilon_i \kappa_V \Sigma + \frac{1}{\varepsilon_i \kappa_V \Sigma}\right) \\ &\sim \frac{1 \times 10^9}{a_{\rm AU}^{3/14}}\,{\rm s} \times \begin{cases} (a_{\rm AU}^{-27/14} + 4 \times 10^{-8} a_{\rm AU}^{27/14}) & \text{for } a_{\rm AU} \lesssim 84 \\ 2 \times 10^{-4}[1 + \left(\frac{a_{\rm AU}}{84}\right)^{3/2}] & \text{for } 84 \lesssim a_{\rm AU} \lesssim 270 \end{cases}, \end{aligned} \tag{4.15}$$

where $\Sigma_d \approx 10^{-2}\Sigma$ is the surface density of dust; (5) the thermal timescale for the gas,

$$\begin{aligned} t_{gi} &\sim \frac{\Sigma k}{\mu_g \sigma T_i^3}\left(\varepsilon_i \kappa_V \Sigma + \frac{1}{\varepsilon_i \kappa_V \Sigma}\right) \\ &\sim \frac{1 \times 10^{12}}{a_{\rm AU}^{3/14}}\,{\rm s} \times \begin{cases} (a_{\rm AU}^{-27/14} + 4 \times 10^{-8} a_{\rm AU}^{27/14}) & \text{for } a_{\rm AU} \lesssim 84 \\ 2 \times 10^{-4}[1 + \left(\frac{a_{\rm AU}}{84}\right)^{3/2}] & \text{for } 84 \lesssim a_{\rm AU} \lesssim 270 \end{cases}; \end{aligned} \tag{4.16}$$

(6) the timescale for the dust temperature to relax to the gas temperature,

$$t_{relax} \sim \frac{r\rho_d}{\mu_d n_g v_g} \sim \frac{\mu_g r \rho_d}{\mu_d \Sigma \Omega} \sim 10^{-2} a_{\rm AU}^3\,{\rm s}\,; \tag{4.17}$$

(7) the dynamical timescale over which the disk adjusts to departures from hydrostatic

[7]This timescale is only relevant where the disk interior is opaque to its thermal radiation.

[8]In evaluating the timescale at the outermost radii, we have set $T_i = 21K$.

equilibrium,

$$t_{dyn} \sim 1.4 a_{\rm AU}^{3/2} \, {\rm yr} \, . \tag{4.18}$$

How rapidly might the SED vary in response to changes in the luminosity of the central star? Since $t_{ds} \ll t_{lt}$, contributions from the surface layer are limited by t_{lt}. Those from the interior are limited by t_{gi}, since $t_{relax} \lesssim t_{di}$ and $t_{diff} \ll t_{gi}$.[9] Figure 4.2 displays the reverberation timescales most relevant for our standard model. To relate these response times to timescales for variation at a fixed wavelength, λ, consult Figure 2.7.

4.8 Disk Stability

Two timescales characterize the response of the passive disk to perturbations in height and temperature. The first is the dynamical timescale (orbital period), $t_{dyn} = \Omega^{-1}$, over which the disk height adjusts to departures from hydrostatic equilibrium. The second is the thermal time, t_T, over which gas equilibrates thermally with incident radiation. For the full vertical extent of the disk interior, $t_T = t_{gi}$ as given by equation (4.16).

Cunningham (1976) ascertains that irradiated accretion disks surrounding black holes are stable to height-integrated perturbations in the limit $F \equiv \Omega t_T \gg 1$. His methods are applied to protostellar accretion disks by D'Alessio et al. (1998), who reach the same conclusion. The latter also claim that when $F \ll 1$, pure temperature perturbations simply decay away on the thermal time.

Here we uncover the existence of an instability triggered by initial perturbations in height and temperature. The physics underlying the instability can be understood simply in the limit $F \ll 1$. Imagine introducing a positive, initially static, axisymmetric, pure height perturbation ("bump") on the disk surface. The side facing the star experiences a positive perturbation in intercepted flux, and the underlying disk temperature rises on the short thermal timescale. The hotter gas expands vertically on the longer dynamical timescale. In doing so, it intercepts yet more light from the central star, and the process runs away.[10]

[9] The former inequality is only marginally satisfied at a_o.

[10] Grains which absorb central stellar radiation must be entrained in rising gas for the instability to proceed.

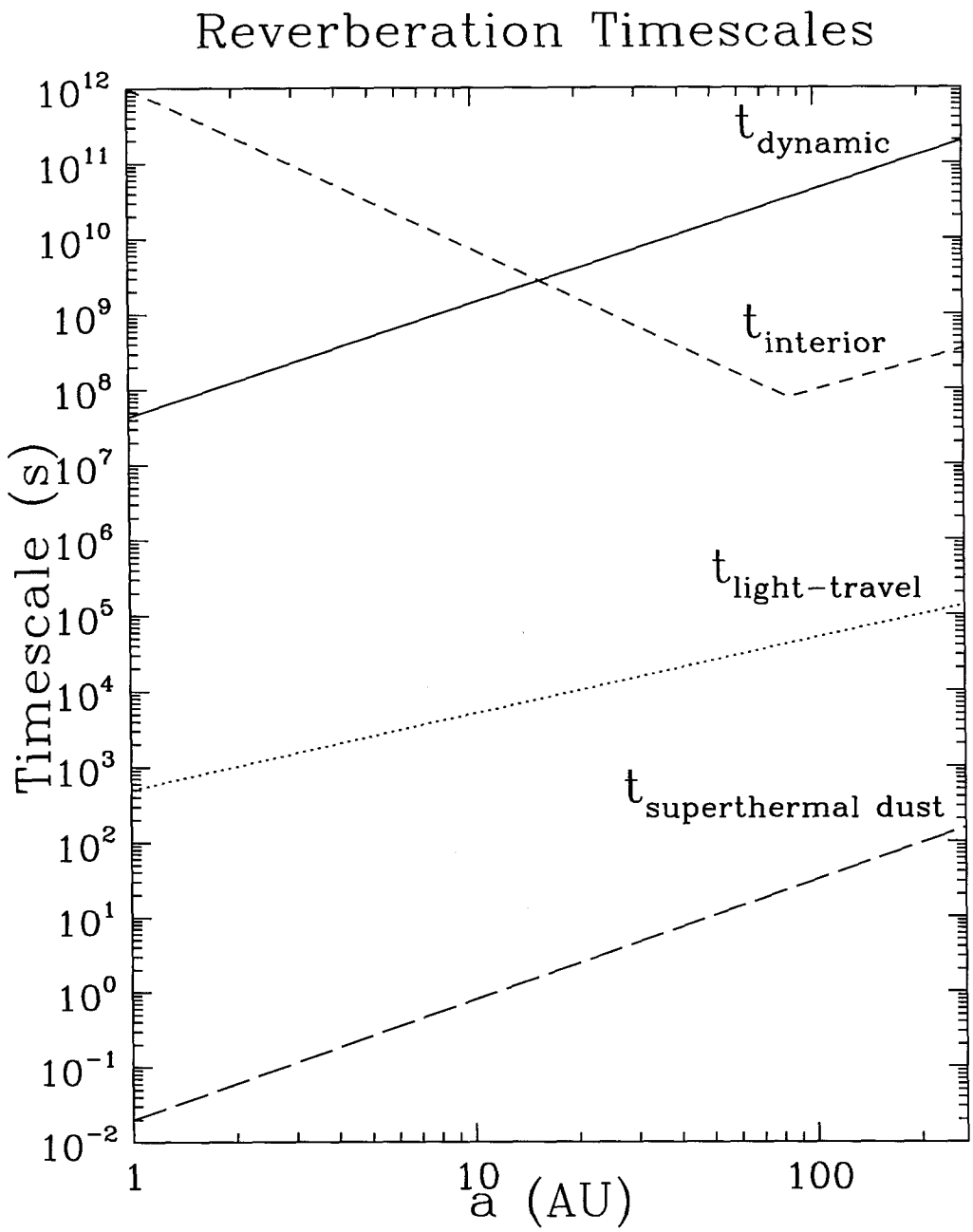

Figure 4.2: *Reverberation timescales for our standard passive disk.*

The timescale over which superheated dust grains in the surface layer equilibrate with ambient stellar radiation is given by $t_{superthermaldust}$. *It is much shorter than the light-travel time from star to disk,* $t_{light-travel}$. *Outside of* $\sim$10AU, *the time it takes a thermal wave to propagate through the disk interior,* $t_{interior}$, *is shorter than the timescale over which hydrostatic equilibrium is established,* $t_{hydrostatic}$.

By the same reasoning, the side facing away from the star collapses in thickness. Naively, we might expect disturbances in thickness to thereby travel towards the central star as they amplify on a dynamical timescale.

Such behavior was suspected by Cunningham (1976) who considered a similar thought experiment. However, the formal discovery of an instability eluded him because his analysis was restricted to the limit $F \gg 1$. D'Alessio et al. (1998) briefly consider the opposite limit, but also overlook the instability because they restrict themselves to perturbations in temperature only, not in height.

In §4.8.1, we devise a toy model that not only captures the behavior found previously by Cunningham (1976) and D'Alessio et al. (1998), but also reveals the existence of the long-suspected "ripple instability." A more realistic linear perturbation analysis is then presented in §4.8.2, in which these ripples will appear to amplify as they propagate away from, not towards, the central star. We discuss these findings in §4.8.3.

4.8.1 Toy Model for the Ripple Instability

We derive a dispersion relation governing waves which propagate within the following toy disk. Simplified versions of the equations of radiative and hydrostatic balance read

$$(\tau_T + 1/\tau_T)\frac{\Sigma k}{\mu_g \sigma}\frac{\partial T}{\partial t} = T_*^4 \left(\frac{R_*}{a}\right)^2 a\frac{d}{da}\left(\frac{H}{a}\right) - T^4 \tag{4.19}$$

$$\frac{\partial^2 H}{\partial t^2} = \frac{kT}{\mu H} - \Omega^2 H\,, \tag{4.20}$$

where τ_T is the vertical optical depth of the disk to its own radiation. A number of approximations have been invoked in the writing of these height-integrated equations. The disk is modeled as a flared blackbody having a single temperature T and height H at every radius a. We take the grazing angle of stellar radiation to be $\alpha = a(d/da)(H/a) \ll 1$. The vertical pressure acceleration, $\rho^{-1}dP/dz$, in equation (4.20) is replaced by its order-of-magnitude counterpart, $\sim kT/\mu_g H$.

This requires the stopping time of grains in gas, $\frac{r}{v_g}\frac{\rho_d}{\rho_g}$, to be shorter than Ω^{-1}. Here, ρ_d is the mass density of a dust grain, and v_g is the sound speed in gas. Our standard model disk satisfies this requirement inside $\sim$100 AU.

Consider small perturbations away from equilibrium: $\eta \equiv \Delta H/H$, $\theta \equiv \Delta T/T \ll 1$. Define $t_T = (\tau_T + 1/\tau_T)\Sigma k/\mu_g \sigma T^3$ as the time required for a thermal wave to diffuse vertically through the interior. Linearized versions of (4.19) and (4.20) read

$$\begin{aligned} t_T \frac{\partial \theta}{\partial t} &= \left(\frac{T_*}{T}\right)^4 \left(\frac{R_*}{a}\right)^2 a \frac{d}{da}\left(\eta \frac{h}{a}\right) - 4\theta \\ &= \frac{1}{\gamma - 1}\frac{d\eta}{d\ln a} + \eta - 4\theta \end{aligned} \tag{4.21}$$

and

$$\frac{1}{\Omega^2}\frac{\partial^2 \eta}{\partial t^2} = \theta - 2\eta\,, \tag{4.22}$$

where $\gamma \equiv d\ln H/d\ln a \approx 9/7$ and we have used the fact that $(T_*/T)^4(R_*/a)^2 h/a = 1/(\gamma - 1)$ in hydrostatic equilibrium. Solve equation (4.22) for θ and substitute into equation (4.21) to find

$$\left(F\frac{\partial^3}{\partial y^3} + \frac{\partial^2}{\partial y^2} + F\frac{\partial}{\partial y} - \frac{\partial}{\partial x} + 1\right)\eta \approx 0\,, \tag{4.23}$$

where $x \equiv \ln a$, $y \equiv \Omega t$, and we have omitted numerical coefficients of order unity for clarity. Normal modes of the form $\eta \propto \exp i(kx - \omega y)$ obey the resultant dispersion relation

$$F\omega^3 + i\omega^2 - F\omega - k - i \approx 0\,. \tag{4.24}$$

Equation (4.24) is a cubic equation for the wave frequency, ω, in terms of the wavenumber, k, for every $F \equiv \Omega t_T$.

Case I: $F \gg 1$. In the limit of long thermal times, all of the oscillatory modes are stable. Two of the three roots of our dispersion relation are

$$\omega^I_{1,2} \approx \pm 1\,. \tag{4.25}$$

These correspond to adiabatic vertical oscillations of the disk. The third root is

$$\omega_3^I \approx -(k+i)\,/\,F\,. \tag{4.26}$$

This root corresponds to waves which spread radially inwards as they damp over the local thermal diffusion time. It was first discovered by Cunningham (1976).

Case II: $F \ll 1$. In the limit of short thermal times, two modes are stable, but the third is unstable. The eigenfrequency of the first mode is

$$\omega_1^{II} \approx -i/F\,, \tag{4.27}$$

and represents a stationary perturbation which decays in place. This is the mode whose existence was deduced by D'Alessio et al. (1998); it corresponds to pure temperature perturbations which damp quickly over the short thermal time. The other 2 modes have eigenfrequencies

$$\omega_{2,3}^{II} \approx \pm\,(\sqrt{1+\sqrt{1+k^2}} - ik/\sqrt{1+\sqrt{1+k^2}}) \tag{4.28}$$

and represent waves which amplify (decay) over approximately the same timescale that they propagate inwards (outwards). The amplifying mode is the long-suspected "ripple instability."

While this analysis succeeds in recovering previous results and uncovering new ones, it is nonetheless extremely crude. It neglects (1) radial and azimuthal velocity fluctuations, (2) effects associated with gas compressibility (e.g., buoyancy forces), and (3) non-inertial forces associated with the mean rotation. Points (1) and (2) are addressed below.

4.8.2 An Irradiated Non-Rotating Sheet

Here we undertake a more rigorous analysis that relaxes several of the restrictions alluded to above. We consider a non-rotating sheet of gas in a mirror-symmetric, constant gravitational field $\vec{g} = -sgn(z)\,g\,\hat{z}$. For conciseness, we shall refer only to the upper half of the sheet ($z > 0$), keeping in mind that the modes examined here have mirror-symmetric dy-

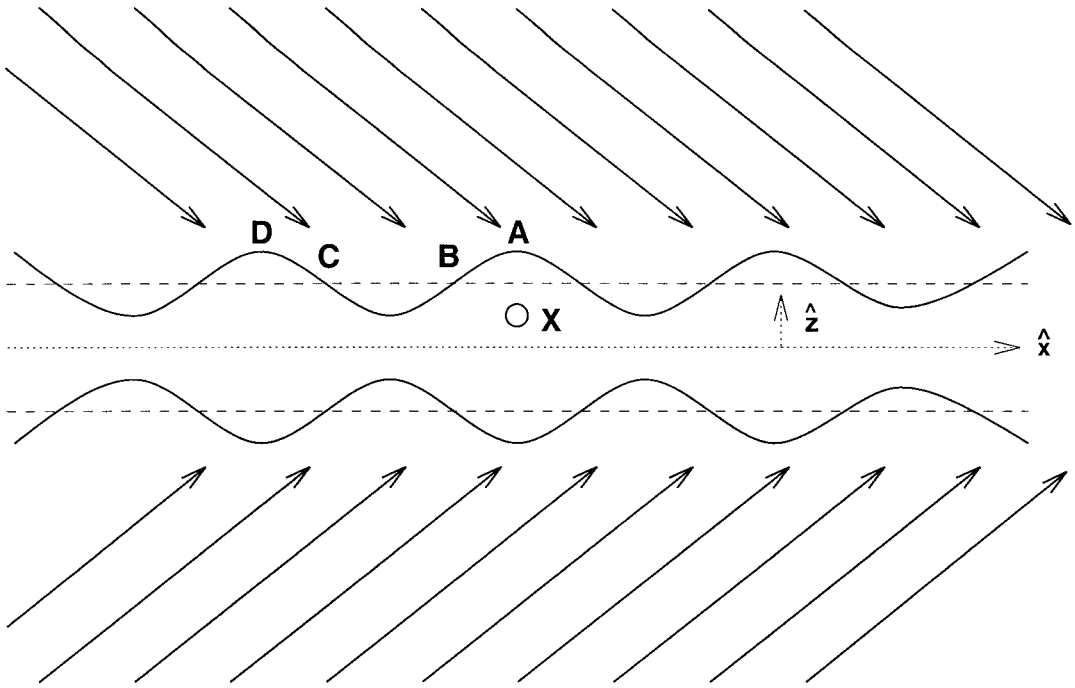

Figure 4.3: *Schematic of Irradiated, Rippled Sheet.*

Plane waves of radiation from an external source strike the gas sheet from above and below at a constant, finite inclination. Dashed lines indicate positions of unperturbed photospheres at $|z| = H$. Solid wavy lines represent perturbed photospheres. In the upper half of the sheet, material beneath crests (troughs) undergo positive (negative) pressure perturbations. Material beneath positive (negative) slope perturbations undergo positive (negative) temperature perturbations by external heating.

namics. Plane waves of radiation from an external source strike the unperturbed gas sheet at a constant, finite inclination. In static equilibrium, the gas is isothermal at temperature T; its density, ρ, and pressure, P, decrease as $\exp(-z/h)$. We define a photospheric surface to the sheet at $z = H \sim h$. A schematic of our system is given in Figure 4.3.

The Lagrangian perturbation equations read[11]:

$$\frac{\partial^2 \xi_x}{\partial t^2} = -\frac{1}{\rho}\frac{\partial(\delta P)}{\partial x} - g\frac{\partial \xi_x}{\partial x} \tag{4.29}$$

$$\frac{\partial^2 \xi_z}{\partial t^2} = -\frac{1}{\rho}\frac{\partial(\delta P)}{\partial z} - g\frac{\partial \xi_z}{\partial z} - g\frac{\delta \rho}{\rho} \tag{4.30}$$

[11]These same equations govern perturbations in rotating disks if (1) mean specific angular momentum gradients are zero, (2) perturbations are axisymmetric, and (3) perturbations involve no azimuthal motions.

$$\frac{\delta\rho}{\rho} + \frac{\partial\xi_x}{\partial x} + \frac{\partial\xi_z}{\partial z} = 0 \tag{4.31}$$

$$\frac{\delta P}{P} = \frac{\delta\rho}{\rho} + \frac{\delta T}{T} \tag{4.32}$$

$$\frac{\delta T}{T} = \frac{1}{4}\frac{\partial\xi_z}{\partial x}|_{z=H}\,. \tag{4.33}$$

Equations (4.29) and (4.30) are the equations of motion in the x (horizontal) and z directions, respectively; $\vec{\xi}$ is the Lagrangian displacement, and δQ is the Lagrangian perturbation in quantity Q. Equations (4.31) and (4.32) are the continuity equation and the equation of state, respectively. Finally, equation (4.33) is our thermal equation. It states that the temperature perturbation at a given position depends on the spatial slope of the photosphere directly above. The two assumptions contained in (4.33) are that $F \ll 1$ and $kH \ll 1$, where k is the horizontal wavenumber of the disturbance. In other words, horizontal diffusion of heat is neglected and the sheet is vertically isothermal at a temperature set instantly by the flux of radiation striking the blackbody surface. As found in the previous section (§4.8.1), the limit $F \ll 1$ is the one appropriate to the ripple instability.

We solve for the dispersion relation governing normal modes of the form $\vec{\xi}, Q \propto \exp i(kx - \omega t)$. Combine equations (4.29)–(4.33) to yield

$$(\omega^2 - ghk^2)\,\xi_x = ighk\,(\frac{\partial\xi_z}{\partial x}|_{z=H} + \frac{\xi_z}{h} - \frac{\partial\xi_z}{\partial z})\,. \tag{4.34}$$

Impose the boundary condition $\delta P|_{z=H} = 0$ to write equation (4.29) as

$$\omega^2\,\xi_x|_{z=H} = igk\,\xi_z|_{z=H}\,. \tag{4.35}$$

Insert (4.35) into (4.34) to find the dispersion relation:

$$\omega^2 = \frac{ghk^2\xi_z|_{z=H}}{h\frac{\partial\xi_z}{\partial z}|_{z=H} - \frac{ikh}{4}\xi_z|_{z=H}}\,. \tag{4.36}$$

We replace the derivative in (4.36) by its order-of-magnitude counterpart, $\sim(\xi_z|_{z=H})/H$,

to re-cast the dispersion relation as

$$\omega^2 \approx ghk^2\,(1 + i\,\frac{kh}{4})\,. \tag{4.37}$$

Not surprisingly, this resembles the dispersion relation for shallow water waves. The roots are

$$\omega_{1,2} \approx \pm\sqrt{ghk^2}(1 + i\,\frac{kh}{8})\,. \tag{4.38}$$

The positive root, ω_1, represents the eigenfrequency of the ripple instability for this system. The most important difference between the instability found here and that found in our previous toy model is that they yield opposite directions of propagation for amplifying waves; here, only waves which travel away from the external radiation source amplify. We may recover the sign of the previous result by returning to equations (4.29)–(4.33), discarding equation (4.29), and setting ξ_x and $\delta\rho$ to 0. Clearly, our present analysis is more realistic.

The reason why amplification occurs for waves which propagate away from the radiation source can be seen from the work integral:

$$W \equiv -\oint \delta P \frac{\partial}{\partial t}\left(\frac{\delta\rho}{\rho}\right) dt \tag{4.39}$$

$$= \oint \delta P \frac{\partial}{\partial t}\left(\frac{\delta T}{T}\right) dt \tag{4.40}$$

where the integration can be performed at any fixed position over one mode period. Mode driving and damping correspond to $W > 0$ and $W < 0$, respectively (see, e.g., Unno, Osaki, & Ando 1989). The relevant eigenfunctions read:

$$\xi_z \approx A\frac{e^{z/h} - 1}{e^{H/h} - 1}\,\cos(kx - \omega t) \tag{4.41}$$

$$\delta P \approx A\rho g\,[1 - e^{(z-H)/h}]\,\cos(kx - \omega t) \tag{4.42}$$

$$\delta T \approx -(kAT/4)\,\sin(kx - \omega t)\,, \tag{4.43}$$

where $A \ll h$ is the maximum amplitude of ξ_z. For waves which propagate away from the radiation source ($k, Re(\omega) > 0$), equations (4.42) and (4.43) yield $W > 0$. In the absence of an external radiation field, pressure gradients caused by displacement of material drive wave motion; equations (4.41) and (4.42) indicate that positive (negative) pressure perturbations occur beneath wave crests (troughs), just as they do in the case of water waves. The introduction of a directed radiation field generates temperature perturbations which perform work. Consider the work done by a fluid parcel at a fixed position, X, in Figure 4.3 as a wave passes by in the positive $\hat{x}$ direction. Initially, the parcel sits beneath a local maximum in wave height (A), so that $\delta P > 0$ and $\delta T = 0$. The wave propagates forward until the parcel sits beneath point B on the wave. During the A–B interval, $\partial(\delta T)/\partial t > 0$ while δP remains positive; the increasingly hot parcel expands and does positive work on its environment. Over the interval B–C, the parcel cools and consequently contracts; it does so beneath a trough, so that $\delta P < 0$; the work done is again positive. Finally, from C–D, $\delta P > 0$ and $\partial(\delta T)/\partial t > 0$, so that $W > 0$. To summarize, the radiation field adds energy to forward-propagating waves at every phase of the mode cycle.

4.8.3 Discussion of Disk Stability

We have discovered a "ripple instability" afflicting externally irradiated, non-rotating sheets of gas in the dual limit that (1) the vertical thermal diffusion time is shorter than the local dynamical time ($F \equiv \Omega t_T \ll 1$), and (2) the horizontal wavelength of the perturbation is longer than the vertical thickness of the sheet ($kH \ll 1$). The instability is triggered by perturbations in height and temperature; those disturbances which propagate away from an external, directed source of radiation are amplified by that source. In the limit $kH \gg 1$, we expect horizontal diffusion of heat to rapidly dampen perturbations.

The significance of the ripple stability for actual passive disks is difficult to assess within the confines of our crude and preliminary study. The principal shortcomings of our analysis are the neglect of rotational effects and possible non-axisymmetries, and our implicit use of the WKB approximation ($ka \gg 1$, where a is the stellocentric location of

the wave).

In our standard model disk, the thermal time for a column extending through the entire disk interior is shorter than the local dynamical time outside $a_F \sim 10\,\mathrm{AU}$ (see Figure 4.2). This transition radius, a_F, is parameter-dependent, since $t_T \propto \Sigma^2$ for the optically thick interior. Outside a_F, it seems possible that outwardly propagating waves that are coherent across the full vertical extent of the disk can amplify by incident starlight. If we make the order-of-magnitude identification $g \sim \Omega^2 h$, then equation (4.38) shows that these waves propagate non-dispersively with group speed $\Omega h \sim c_s$, the gas sound speed. Inside a_F, the instability might afflict surface layers down to depths for which the thermal time of the material above is comparable to the dynamical time.

By equation (4.38), the number of times a wave e-folds as it propagates from radius a to $\sim$2a is $N \sim (ka)\,(kh_F)/8$, where h_F is the vertical depth over which instability can occur ($h_F \sim h$ when $a > a_F$). Further evolution of the wave cannot be followed by our WKB methods. Since $max(kh_F) \sim 1$, then $max(N) \sim a/8h_F$. Outside a_F in our standard model disk, $max(N) \sim 1$; amplification is modest and perturbations here are probably of minor importance. Inside a_F, the instability might affect only the uppermost, geometrically thin layers of the disk; here, N might be quite large. What impact these surface waves have on the disk at $a \lesssim a_F$ and on the resultant SED at wavelengths $\lambda \lesssim 10\,\mu\mathrm{m}$ are open questions. We reserve further exploration of disk stability to a future study.

4.9 Unresolved Issues

Our investigation leaves many unresolved issues.

How much of the thermally emitted spectrum is covered by molecular lines? Do the lines appear in absorption or emission? Which molecules are they associated with?

How high does the layer of superthermal dust grains sit above the disk midplane? What is the gas temperature in this layer? Is gravitational settling of dust grains seriously impeded by turbulent mixing or by vertical circulation? Or does radiation pressure from the central star dominate all the forces, so that the superthermal grains are driven into denser environments and the disk geometry flattens?

What is the albedo of the circumstellar dust? How does scattered stellar radiation compare to thermal emission in the near IR?

Does active accretion significantly affect the SEDs, either by preventing the settling of dust grains or by thickening the inner regions of the disk?

What are the consequences of the ripple instability, particularly in the case of disk surface waves at small radius? How do these disturbances couple to the rotation of the disk? Might localized high/low pressure zones couple to Coriolis forces to spawn vortices? Does the instability drive turbulence? Might such turbulence transport significant angular momentum?

What impact do ripples have on the SED? Are enough regions cast into shadow to cause deficits in emission like those at 12 μm for GM Aur? Or do enhanced fluxes from the ripples' starlit faces compensate for such deficits, leaving the SED unchanged?

Acknowledgments

We thank Shri Kulkarni for suggesting that temporal variability of disk emission can provide a further means of probing disk characteristics, Zeltjko Ivecic for prompting us to examine the stability of passive disks, and Steve Beckwith for his detailed referee's report. Financial support for this research was provided by NSF Grant 94-14232 and an NSF Graduate Fellowship.

Bibliography

[1] Bell, K. R., Cassen, P. M., Klahr, H. H., & Henning, Th. 1997, ApJ, 486, 372

[2] Chiang, E. I. & Goldreich, P. 1997, ApJ, 490, 368 (CG)

[3] Cunningham, C. 1976, ApJ, 208, 534

[4] D'Alessio, P., Cantó, J., Hartmann, L., Calvet, N., & Lizano, S. 1998, ApJ, 511, 896

[5] Draine, B.T. & Lee, H.M. 1984, ApJ, 285, 89

[6] Moriarty-Schieven, G. H. & Butner, H. M. 1997, ApJ, 474, 768

[7] Unno, W., Osaki, Y., & Ando, H. 1989, *Nonradial Oscillations of Stars* (University of Tokyo)

Chapter 5

Spectral Energy Distributions: Ice and Silicate Mineralogy

Abstract

More sophisticated models of passive disks are constructed in which we (1) account for a range of particle sizes, (2) employ laboratory-based optical constants of water ice, olivine, and iron to compute the emissive properties and temperatures of disk grains and (3) solve numerically the equations of radiative and hydrostatic equilibrium within our 2-layer approximation. We explore how the SED of a face-on disk depends on grain size distributions, disk surface densities, and stellar photospheric temperatures. Observed SEDs of 4 Herbig Ae/Be and 2 T Tauri stars, some of which include data from the Long Wavelength Spectrometer (LWS) aboard the Infrared Space Observatory (ISO), are fitted with our models. Silicate and water ice emission from superheated disk surfaces appear in several sources. The magnitude of the infrared excesses in some sources is consistent with significant vertical settling of superheated dust grains.

5.1 Introduction

In this chapter we further refine our theory of passive disks by (1) explicitly accounting for a range of particle sizes, (2) employing laboratory-based optical constants of a representative suite of circumstellar grain materials, and (3) numerically solving the equations of radiative and hydrostatic equilibrium within the 2-layer approximation. We return to the restriction of computing SEDs for face-on viewing aspects; the analysis in Chapter 3 proves that the SED varies negligibly with disk inclination provided the flared disk does not occult the central star. Hence, we shall deploy these more sophisticated models only against observations of systems for which the central star is largely unobscured (i.e., "class II

spectra").

Input parameters for our refined standard model are detailed in §5.2. The basic equations describing our disk are set forth in §5.3. Results, including a systematic exploration of how the SED varies in input parameter space, are presented in §5.4. Finally, model fits to observed SEDs of 4 Herbig Ae/Be stars and 2 T Tauri stars are supplied in §5.5.

5.2 Input Parameters

5.2.1 Grain Compositions and Size Distribution

Table 5.1 lists the input parameters of our refined model. Where local dust temperatures (= gas temperatures) fall below $T_{sub}^{ice} \approx 150$ K, the grains are taken to be spheres of amorphous olivine ($MgFeSiO_4$, bulk density = 3.71 g cm^{-3}) mantled with water ice (H_2O, bulk density = 1 g cm^{-3}). Where local temperatures fall between T_{sub}^{ice} and $T_{sub}^{sil} \approx 1500$ K, only the pure olivine cores are assumed to remain. In innermost disk regions where local temperatures fall between T_{sub}^{sil} and $T_{sub}^{iron} \approx 2000$ K, the grains are taken to be spheres of metallic iron (Fe, bulk density = 7.87 g cm^{-3}). The iron or silicate cores in the disk surface (interior)

Table 5.1: Input Parameters of Refined Model

Symbol	Meaning	Standard Value
M_*	Stellar Mass	$0.5 M_\odot$
R_*	Stellar Radius	$2.5 R_\odot$
T_*	Stellar Effective Temperature	4000 K
Σ_0	Surface Density at 1 AU	10^3 g cm^{-2}
p	$-$ d log Σ/d log a	1.5
a_o	Outer Disk Radius	$8600 R_* = 100$ AU
H/h	Visible Photospheric Height / Gas Scale Height	4.0
q_i	$-$ d log N/d log r in Interior	3.5
q_s	$-$ d log N/d log r in Surface	3.5
$r_{max,i}$	Maximum Grain Radius in Interior	1000 μm
$r_{max,s}$	Maximum Grain Radius in Surface	1 μm
T_{sub}^{iron}	Iron Sublimation Temperature	2000 K
T_{sub}^{sil}	Silicate Sublimation Temperature	1500 K
T_{sub}^{ice}	H_2O Ice Sublimation Temperature	150 K

possess a power law distribution of radii r between r_{min} and $r_{max,s}$ ($r_{max,i}$):

$$dN \propto r^{-q_{i(s)}}\, dr\,, \tag{5.1}$$

where dN is the number density of grains having radii between r and $r + dr$. Our standard value of $q_i = q_s = 3.5$ is such as to place most of the geometric surface area in the small grains and most of the mass in the large grains. In practice, r_{min} is fixed at $10^{-2}\,\mu$m, while $r_{max,i}$, $r_{max,s}$, q_i, and q_s are free to vary. Generally $r_{max,s} < r_{max,i}$ since large grains tend to settle quickly out of tenuous surface layers (see §4.4). All of the cosmically abundant iron is assumed to be locked within grains. Following Pollack et al. (1994), we take 50% of the cosmically abundant oxygen to be locked in H_2O ice. Values for all cosmic abundances are obtained from Lang (1980). Together, these assumptions yield a fractional thickness, $\Delta r/r$, for the water ice mantle equal to 0.5. For simplicity, this ratio is assumed to be independent of grain size.

Optical constants for amorphous olivine are obtained from the University of Jena Database [http://www.astro.uni-jena.de; see also Jäger et al. (1994)]. Longward of $500\,\mu$m where optical data for silicates are not available, the complex refractive index $(n + ik)$ for glassy olivine is extrapolated such that $n\,(\lambda \geq 500\,\mu\text{m}) = n\,(500\,\mu\text{m})$ and $k\,(\lambda \geq 500\,\mu\text{m}) = k\,(500\,\mu\text{m})\left(\frac{\lambda}{500\,\mu\text{m}}\right)^{-1.01}$. Optical constants for pure crystalline H_2O ice are taken from the NASA ftp site (climate.gsfc.nasa.gov / pub / wiscombe / Refrac_Index / ICE /; see also Warren 1984). Though employing the constants for a cosmic mixture of amorphous ices (H_2O : CH_3OH : CO : NH_3; see Hudgins et al. 1993) would be more appropriate, we nonetheless adopt the data for pure H_2O ice because the latter are available over all wavelengths of interest, from the ultraviolet to the radio, whereas the former are not. One consequence of using the constants for crystalline (213–272 K) water ice as opposed to amorphous (∼100 K) ice is that spectral features due to translational lattice modes at 45 and 62 μm are slightly underestimated in width and overestimated in amplitude (see, e.g., Hudgins et al. 1993). Optical constants for metallic Fe are obtained from Pollack et al. (1994).

The inner cutoff radius of the disk, a_i, is fixed at $2R_*$. For T Tauri stellar parameters, this

radius coincides with the distance at which iron grains in the surface layer attain their sublimation temperature. For the hotter HAeBe stars, the iron condensation boundary occurs at $a \approx 20$–$100 R_*$. Inside the iron condensation line, the disk may still be optically thick at stellar wavelengths even if dust is absent. Opacity sources include pressure-broadened molecular lines and Rayleigh scattering off hydrogen atoms [see the appendix of Bell & Lin (1994)]. For simplicity, when modelling HAeBe stars we employ a 1-layer blackbody disk inside the iron boundary.

The outer cutoff radius, a_o, remains a free parameter, though it is not allowed to exceed the value for which $H(a_o) = a_o$.

5.2.2 Grain Absorption Efficiencies and Opacities

The grain emissivity, $\varepsilon(r, \lambda)$, is equal to its absorption efficiency and is calculated using Mie-Güttler theory (Bohren & Huffman 1983; see their subroutine BHCOAT.F).[1] Figure 5.1 displays absorption efficiencies for ice-silicate, silicate, and iron particles having a few representative sizes. The emissivity index in the Rayleigh-limit, β, equals 1.64, 2.00, and 0.50 for the three types of particles, respectively. Noteworthy resonances include the O-H stretching (3.1, 4.5 μm) and H-O-H bending (6.1 μm) modes in water ice; the Si-O stretching (10 μm) and O-Si-O bending (18 μm) modes in silicates; and the intermolecular translational (45, 62 μm) modes in water ice (Bohren & Huffman 1983). Oscillatory behavior near the onset of the Rayleigh limit ($2\pi r/\lambda \approx 1$) reflects the so-called "ripple structure" that arises from our use of perfectly spherical particles (Bohren & Huffman 1983); we expect real-world deviations from sphericity to smooth out this artificial behavior.

In the disk interior, the opacity is given by

$$\kappa_i(\lambda) = \frac{\pi}{\rho_t} \int_{r_{min}}^{r_{max,i}} \frac{dN}{dr}\, r^2\, \varepsilon(\lambda, r)\, dr\,, \tag{5.2}$$

where ρ_t is the total density of gas and dust. Figure 5.2 displays κ_i for our distribution of ice-silicate and silicate particles in cosmic abundance gas.

[1] Our grain emissivity, ε, equals Q_{abs} in the notation of Bohren & Huffman (1983).

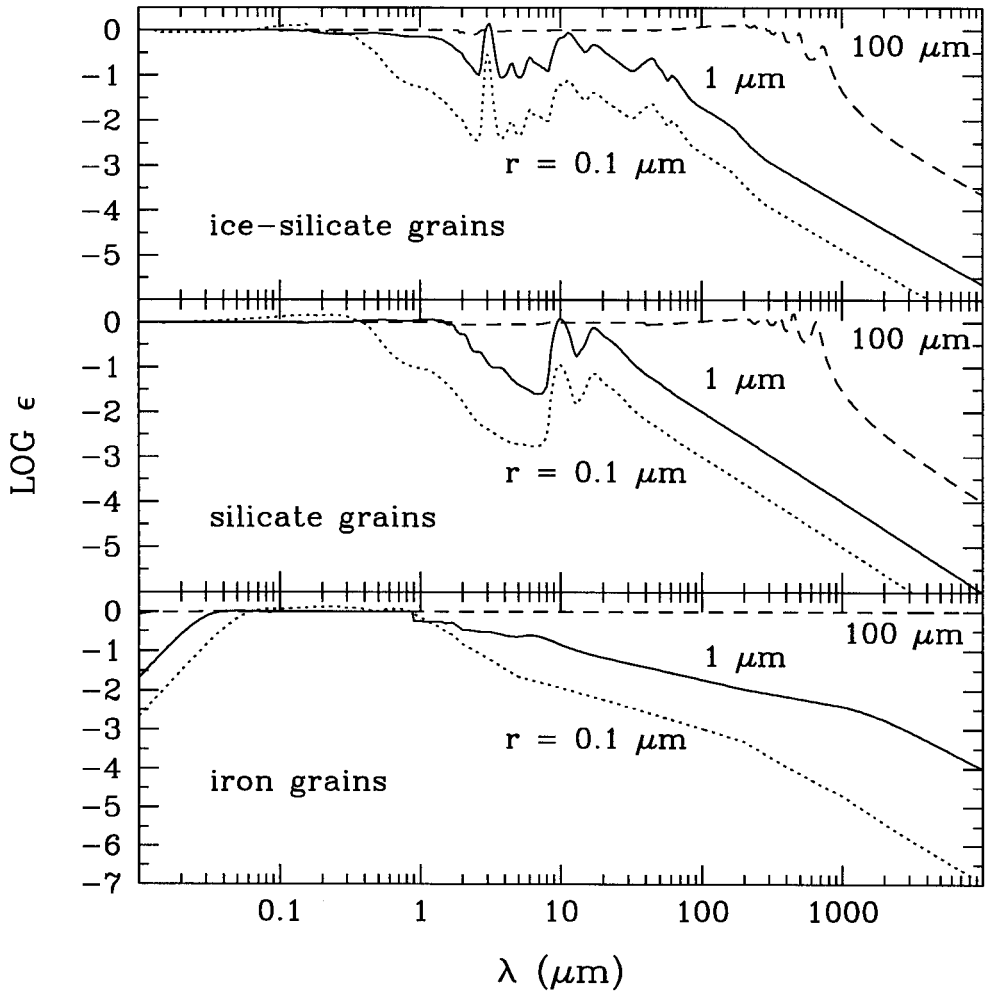

Figure 5.1: *Emissivities of ice-silicate (H_2O / amorphous olivine), silicate (amorphous olivine), and iron grains having three representative core sizes.*

The fractional thickness of the water ice mantle is $\Delta r/r = 0.5$. Resonant features include the O-H stretching (3.1, 4.5 μm) and H-O-H bending (6.1 μm) modes in water ice; the Si-O stretching (10 μm) and O-Si-O bending (18 μm) modes in silicates; and the intermolecular translational (45, 62 μm) modes in water ice (Bohren & Huffman 1983). Oscillatory behavior near the onset of the Rayleigh limit ($2\pi r/\lambda \approx 1$) reflects "ripple structure" arising from our use of perfectly spherical particles.

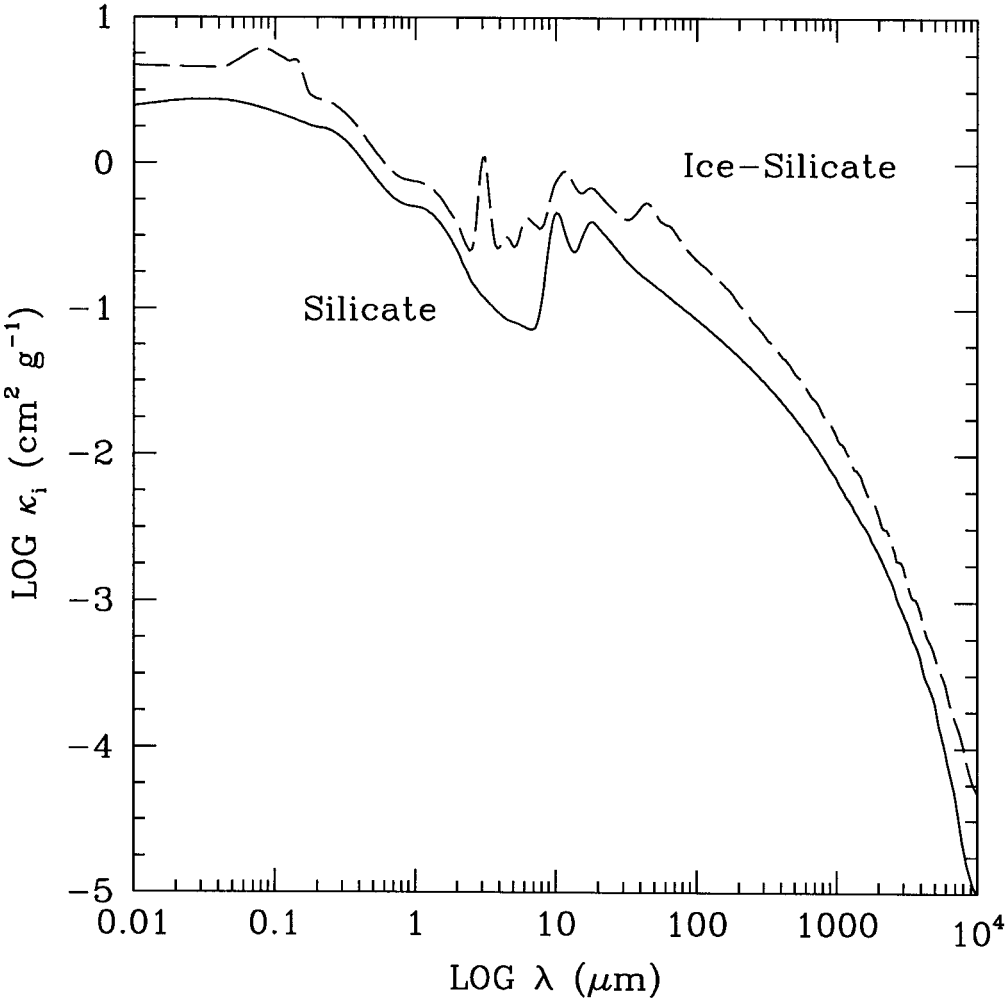

Figure 5.2: *Mass absorption coefficients for our distributions of ice-silicate and silicate particles in solar abundance gas.*

The same resonant features found in Figure 5.1 are seen here. We have smoothed these curves to remove the so-called "ripple structure" at $\lambda \gtrsim 1000\,\mu$m *that arises from the sphericity of our particles having* $r \approx 1000\,\mu$m.

5.3 Basic Equations

In the surface layer, dust grains are directly exposed to stellar radiation. The grains attain an equilibrium temperature

$$T_{ds}(a,r) = T_* \left(\frac{R_*}{a}\right)^{1/2} \left(\frac{\phi\langle\varepsilon(r)\rangle_{T_*}}{4\langle\varepsilon(r)\rangle_{T_{ds}}}\right)^{1/4}, \tag{5.3}$$

where $\langle\varepsilon(r)\rangle_T$ is the emissivity of a grain of radius r averaged over the Planck function at temperature T, and $\phi \approx 1/2$ is the fraction of the stellar hemisphere that is seen by the grain.

In our two-layer formalism, exactly half of the radiation emitted by the surface layer is directed towards the disk interior. Of this half, a fraction, $1 - e^{-\Sigma\langle\kappa_i\rangle_s}$, is absorbed by the disk interior, where $\langle\kappa_i\rangle_s$ is the opacity of the disk interior averaged over the spectrum of radiation from the surface.[2] In radiative balance,

$$\frac{\phi}{2}\left(1 - e^{-\Sigma\langle\kappa_i\rangle_s}\right)\left(\frac{R_*}{a}\right)^2 T_*^4 \sin\alpha = \left(1 - e^{-\Sigma\langle\kappa_i\rangle_i}\right) T_i^4, \tag{5.4}$$

where $\langle\kappa_i\rangle_i$ is the interior opacity averaged over the spectrum of radiation from the disk interior.[3] Here,

$$\alpha \approx \arctan\left(\frac{d\ln H}{d\ln a}\frac{H}{a}\right) - \arctan\frac{H}{a} + \arcsin\left(\frac{4}{3\pi}\frac{R_*}{a}\right) \tag{5.5}$$

is the angle at which stellar radiation strikes the surface [cf. equation (2.5)]. The height of the disk photosphere, H, is assumed to be proportional to h, the vertical gas scale height, with a fixed constant of proportionality.[4] In hydrostatic equilibrium,

$$\frac{H}{a} = \frac{H}{h}\frac{h}{a} = \frac{H}{h}\frac{c_s}{\Omega a} = \frac{H}{h}\sqrt{\frac{T_i}{T_c}}\sqrt{\frac{a}{R_*}}. \tag{5.6}$$

[2] In practice, we perform this average over a Planck function evaluated at the temperature of the most luminous grains in the surface. For our assumed size distribution, these dominant grains typically have radii $r \approx 0.5\,\mu$m; surface grains having $r \approx 0.2$–$1\,\mu$m are responsible for absorbing ~50% of the incident stellar radiation.

[3] We perform this average over a Planck function evaluated at temperature T_i.

[4] In reality, when dust and gas are well-mixed, the ratio H/h decreases slowly from ~5 at 1 AU to ~4 at 100 AU.

where H/h is an input parameter equal to 4 for our standard model.

Equations (5.4) and (5.6) are two equations for the two unknown functions, $H(a)$ and $T_i(a)$. Substitution of (5.6) into (5.4) yields an algebraic equation for T_i and the slowly-varying flaring index, $\gamma \equiv d\ln H/d\ln a$. We solve this equation numerically on a logarithmic grid in stellocentric distance, beginning at $a = a_i$ and stepping incrementally forward. At every step, we employ Brent's root-finder (Press et al. 1992) to solve for T_i for a given value of γ. The initial value of γ at $a = a_i$ is guessed; subsequent values at $a > a_i$ are computed with every step. Only one initial value of γ generates a smooth evolution of all quantities with distance; this is our desired solution.

The SED of the disk equals the sum of emission from the disk interior,

$$L_{\lambda,i} = 8\pi^2\lambda \int_{a_i}^{a_o} B_\lambda(T_i)\,(1 - e^{-\Sigma\kappa_i})\,a\,da\,, \tag{5.7}$$

and from the surface layers (above and below the disk midplane),

$$L_{\lambda,s} = 8\pi^2\lambda\,(1 + e^{-\Sigma\kappa_i}) \int_{a_i}^{a_o} S_\lambda\,(1 - e^{-\tau_s})\,a\,da\,. \tag{5.8}$$

Here

$$S_\lambda = \frac{2\int_{r_{min}}^{r_{max,s}} B_\lambda(T_{ds})\,\frac{dN}{dr}\,r^2\,\varepsilon(\lambda,r)\,dr}{\int_{r_{min}}^{r_{max,s}} \frac{dN}{dr}\,r^2\,\varepsilon(\lambda,r)\,dr} \tag{5.9}$$

and

$$\tau_s(\lambda,a) = \frac{\int_{r_{min}}^{r_{max,s}} \frac{dN}{dr}\,r^2\,\varepsilon(\lambda,r)\,dr}{\int_{r_{min}}^{r_{max,s}} \frac{dN}{dr}\,r^2\,\langle\varepsilon(r)\rangle_{T_*}\,dr}\,\sin\alpha\,. \tag{5.10}$$

In equation (5.9), the factor of 2 is inserted so that exactly half of the incident radiation is reprocessed by the surface layer.

5.4 Results

5.4.1 Flaring Index

Figure 5.3 displays the behavior of the flaring index, γ. We conceptually divide the disk into three annular regions, as was done in §2.4.2. In the region marked "I," the disk interior is opaque to both its own reprocessed radiation and to radiation from the surface. Here γ increases from its flat disk value of $1.125 \approx 9/8$ to its asymptotic value of $1.275 \approx 9/7$ as the first two terms on the right-hand side of equation (5.5) gradually dominate the last term. In region "II," the disk interior remains opaque to radiation from the surface, but is optically thin to its own reprocessed radiation. Here γ steeply rises with a because grains in the disk interior equilibrate at relatively high temperatures to compensate for the relative inefficiency with which they re-radiate the incident energy. Finally, in region "III," the interior is transparent to radiation from the surface (i.e., $\Sigma\langle\kappa_i\rangle_s \lesssim 1$); the inability of the interior to absorb the incident energy causes γ to decrease.

5.4.2 Disk Temperatures

Figure 5.4 exhibits temperature profiles for the surface (T_{ds}) and for the interior (T_i). The temperatures of grains in the surface layer vary slightly with their sizes; in Figure 5.4, we have chosen to plot T_{ds} for the size bin containing the most luminous grains, i.e., the size bin which absorbs most of the incident stellar radiation. For our choices of grain composition and size distribution, these dominantly absorbing grains have radii $r \approx 0.1$–$0.7\,\mu\mathrm{m}$. For reference, we also overlay in Figure 5.4 the temperature of an imaginary blackbody sphere which is naked before half of the central star (T_{BB}).

These temperature profiles are largely similar to those found in the simpler model presented in Chapter 2, and help to justify the approximations made there. At a given distance, the surface is hotter than the interior by a factor of $\sim$3. Though the interior in region "II" (see §5.4.1 above) is not radially isothermal as was found in the cruder analysis of §2.4.2, $T_i(a)$ does flatten slightly at these distances, as expected. Deviations from a single power-law behavior for $T_{ds}(a)$ arise from structure in $\varepsilon(\lambda)$. For example, T_{ds} declines slightly

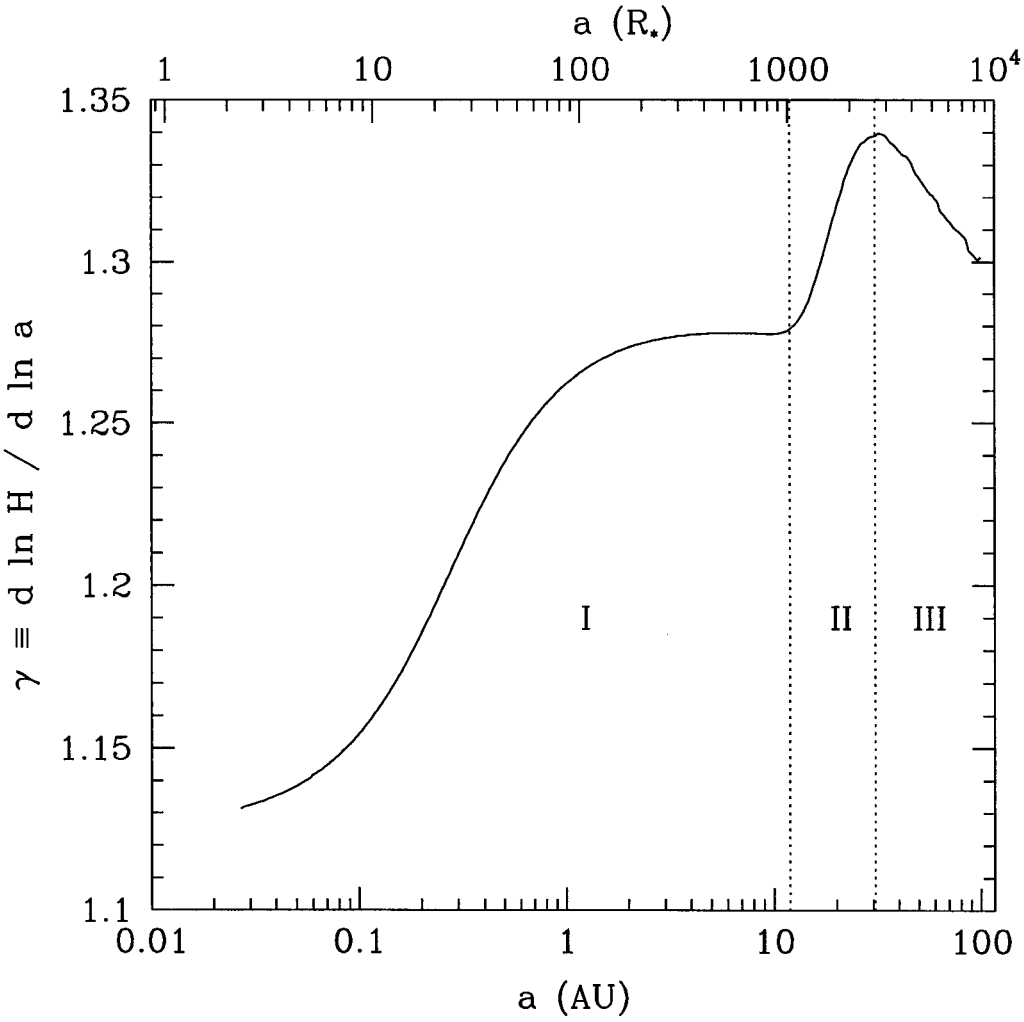

Figure 5.3: *The flaring index, $\gamma \equiv d \ln H / d \ln a$, for our refined standard model.*

As was done in §2.4.2, we divide the disk into three annular regions depending on the optical depth of the disk interior. In region "I," the interior behaves as a blackbody; γ increases from its flat disk value of $1.125 \approx 9/8$ to its asymptotic flared value of $1.275 \approx 9/7$ as the disk thickness becomes increasingly larger than the stellar radius. In region "II," the disk interior becomes optically thin to its own reprocessed radiation; γ increases as interior grains enhance their temperatures to compensate for the inefficiency with which they re-radiate. In region "III," the interior is transparent to radiation from the surface and cools quickly with increasing distance, causing γ to decrease.

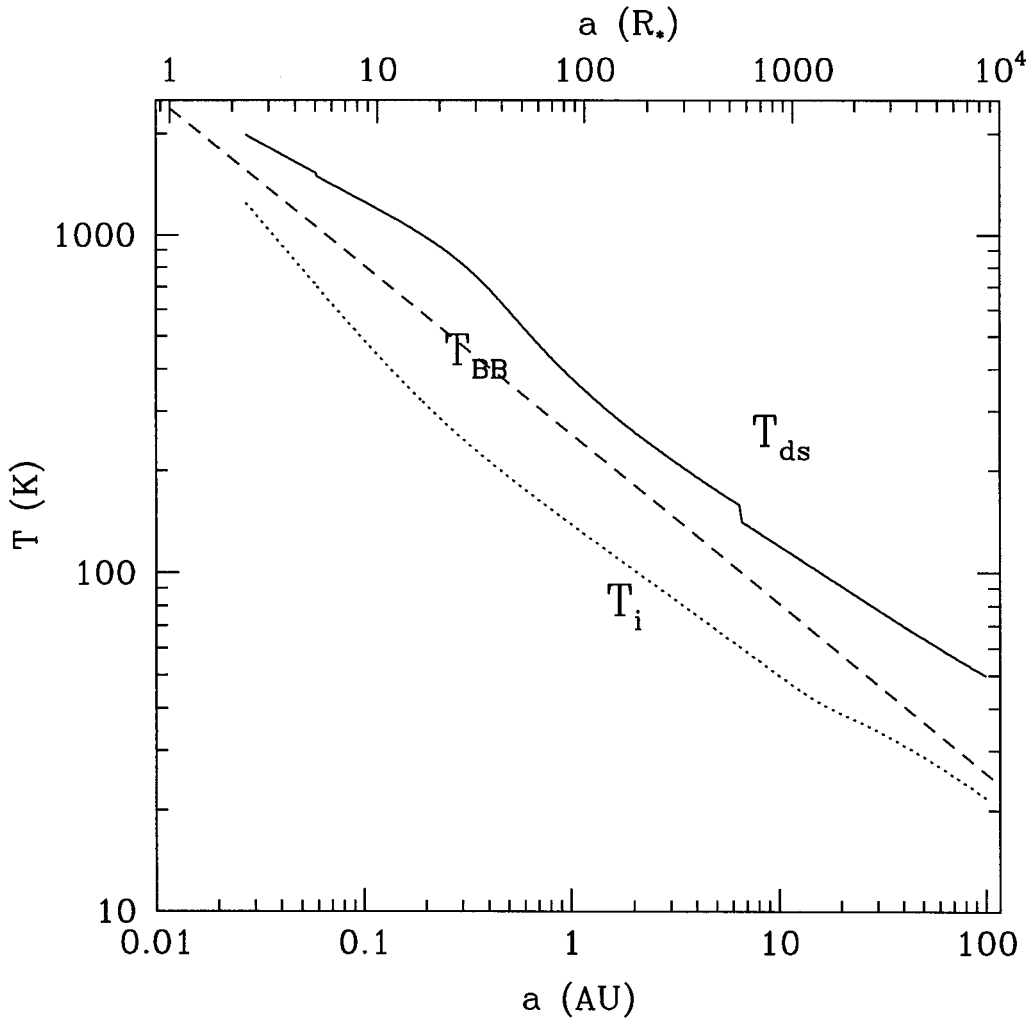

Figure 5.4: *Temperature profiles for the surface and for the interior in our refined standard model.*

The temperatures of grains in the surface layer depend on their sizes; here, the curve marked T_{ds} represents the size bin containing the most luminous grains. The discontinuity in T_{ds} at $a \approx 10\mathrm{AU}$ marks the water condensation boundary, outside of which H_2O ice coats silicate cores in the surface. For reference, the temperature of a spherical blackbody which is naked before the central star is shown as a dashed line.

more steeply with a for $0.5 \lesssim a_{\rm AU} \lesssim 2$ because grains at these distances cool relatively efficiently through silicate resonances at 10–20 μm.

5.4.3 Dependence of SED on Input Parameters

The SED for our refined standard model is displayed in Figure 5.5. The result shown here and that for our simpler model in Chapter 2 (see Figure 2.6) differ primarily in the emission from the superheated surface. Here we have accounted in detail for the optical properties of a few likely circumstellar grain materials. Solid-state resonances of superheated dust grains residing in disk surface layers appear in emission. These include vibrational modes in silicates at 10 and 18 μm, and lattice translational modes in water ice at 45 and 62 μm. Emission lines from vibrational resonances in ice shortward of 10 μm are absent; ice is not present in the disk surface inside 10 AU. Note the dearth of emission from the surface between 1 and 8 μm; amorphous olivine ($MgFeSiO_4$) grains having $r \lesssim 1$ μm are relatively transparent at these wavelengths (see Figure 5.1). The signature of this "silicate transparent region", however, is largely masked by emission from the optically thick interior at small radius. The broad peak in disk surface emission near $\lambda \sim 1.5$ μm arises from pure iron particles and iron impurities in our silicate particles.

In Figures 5.6–5.7, we explore the dependence of the SED on our input parameters. All the variations are easily understood. We observe the following behavior:

1. Millimeter-wave fluxes are most sensitive to Σ_0, $p \equiv -d\ln\Sigma/d\ln a$, $r_{max,i}$, and $q_i \equiv -d\ln N/d\ln r$. The former two variables determine the amount of mass in the cool disk interior at large radius; the latter two variables affect the millimeter-wave opacity in the disk interior.

2. Millimeter-wave SEDs for $r_{max,i} = 1$ and 10 μm are identical; these two cases are appropriate to the Rayleigh limit, where absorptive cross sections are proportional to grain volume (Bohren & Huffman 1983). In this limit, millimeter-wave opacities are independent of how the total condensable mass in water and silicates is distributed with particle size.

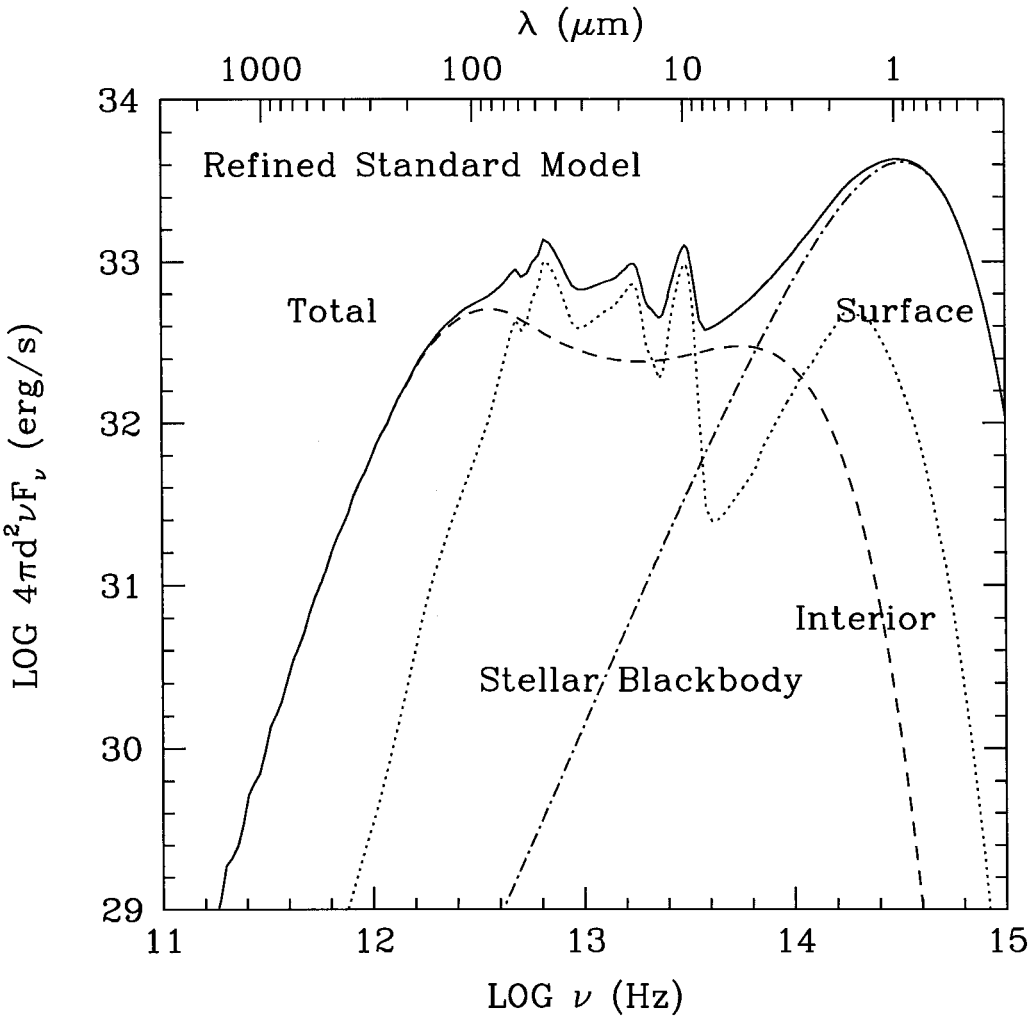

Figure 5.5: *Spectral energy distribution for our refined standard model.*

Resonant bands of superheated dust grains in disk surface layers appear in emission. These include vibrational modes in silicates at 10 *and* 18 μm, *and lattice translational modes in water ice at* 45 *and* 62 μm. *Absent are emission lines from vibrational resonances in ice shortward of* 10 μm; *ice is not present in the disk surface inside 10 AU. Note the dearth of emission from the surface between* 1 *and* 8 μm; *amorphous olivine* ($MgFeSiO_4$) *grains having* $r \lesssim 1$ μm *are relatively transparent at these wavelengths (see Figure 5.1).*

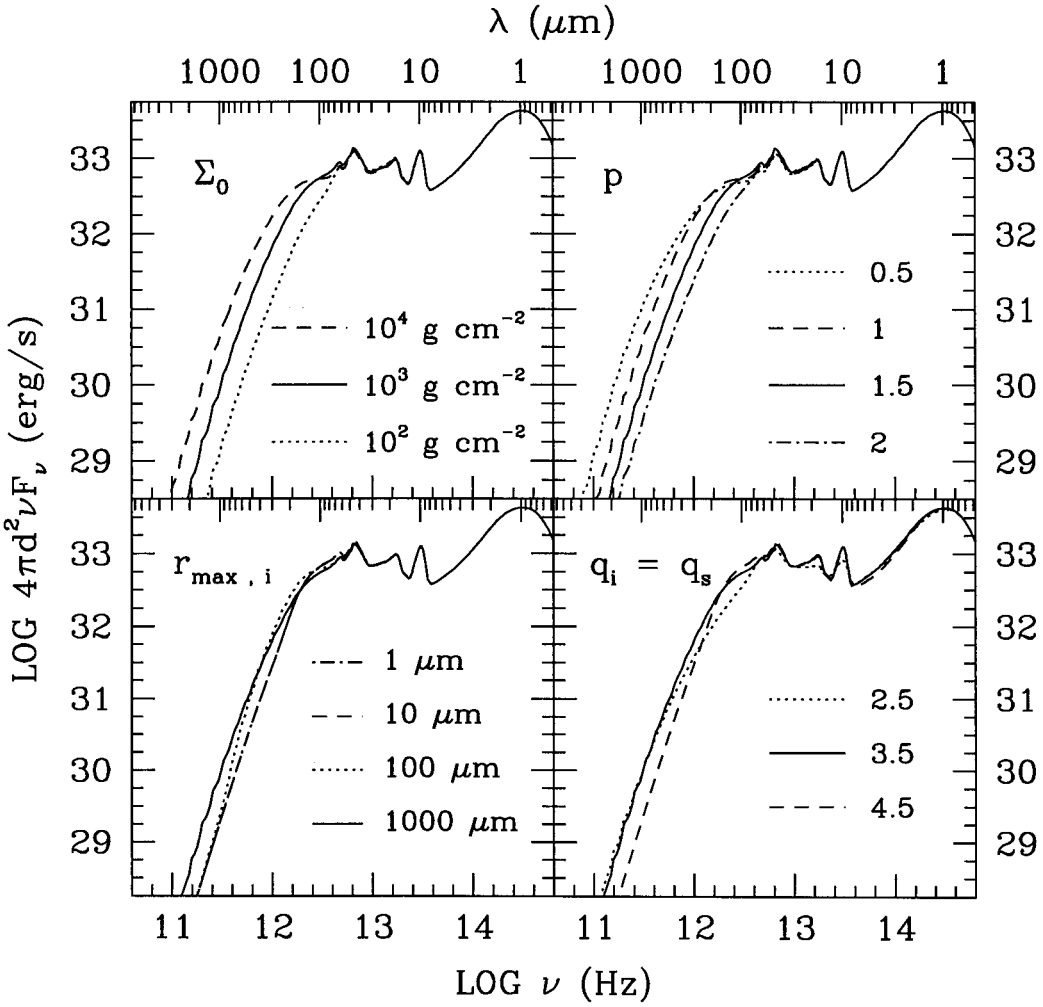

Figure 5.6: *Dependence of SED on input parameters* Σ_0, p, $r_{max,i}$, *and* $q_i = q_s$.

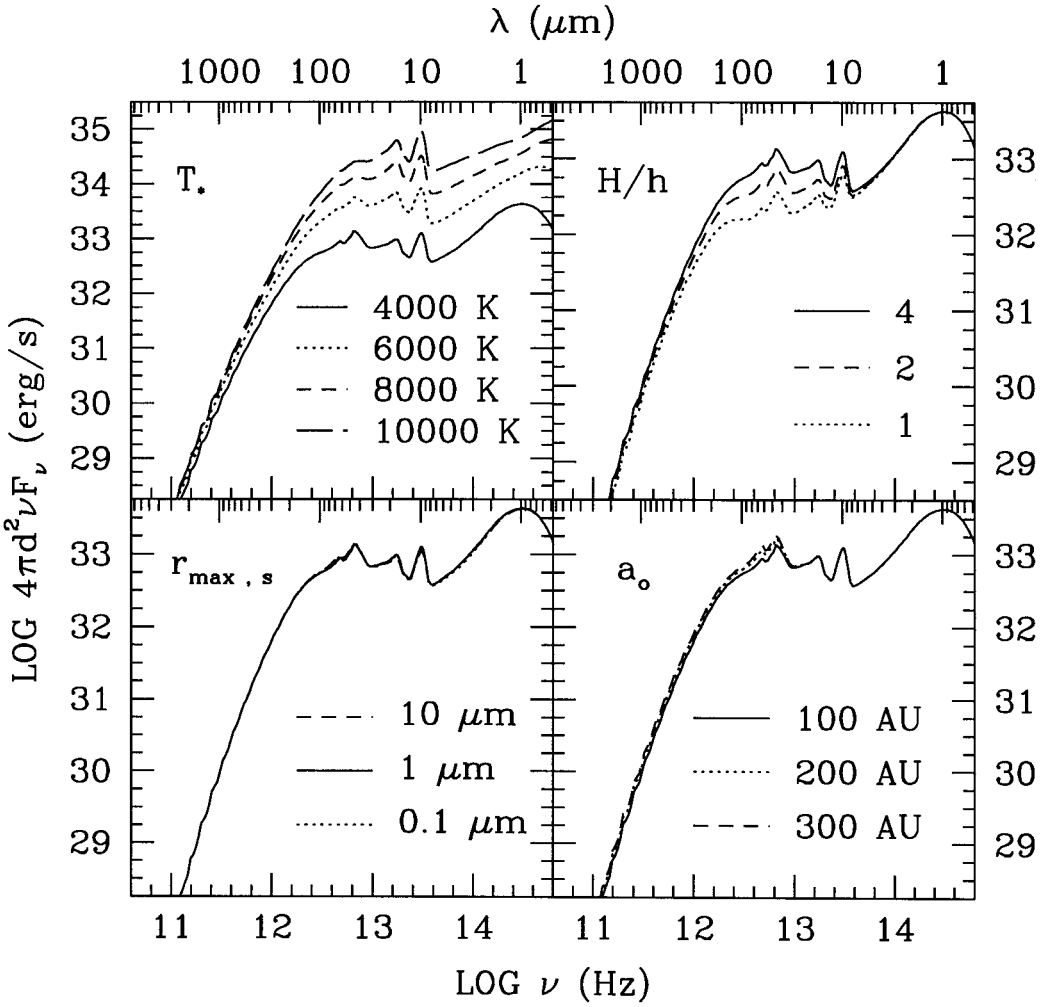

Figure 5.7: *Dependence of SED on input parameters T_*, H/h, $r_{max,s}$, and a_o.*

3. For all values of $r_{max,i}$ considered, the spectral index of the SED at $\lambda = 2$–4 mm equals $n_{2-4} \equiv d\ln(\nu F_\nu)/d\ln\nu \equiv 3 + \beta_{\rm eff} = 4.6$. The value of $\beta_{\rm eff} = 1.6$ equals the value of β for our ice-silicate grains, indicating that radiation at these wavelengths emerges from optically thin material.

4. For $q_s = q_i = 2.5$, most of the geometric cross section in our dust size distribution is concentrated in the largest grains ($r_{max,s} \approx 1\,\mu$m in the surface and $r_{max,i} \approx 1000\,\mu$m in the interior). Compared to the standard model, silicate emission features from the surface are weaker because the Rayleigh limit does not obtain for these large grains. Furthermore, $n_{2-4} = 1.3 < \beta = 1.6$; a substantial fraction of the mm-wavelength emission arises from disk material made optically thick by the increased number of mm-sized particles.

5. For $q_s = q_i = 4.5$, most of the condensable mass is concentrated in the smallest grains ($r_{min} \approx 0.01\,\mu$m). Millimeter-wave opacities and therefore fluxes are lower than for our standard model. Emission at $\lambda \lesssim 100\,\mu$m does not differ from standard model results because the Rayleigh limit still obtains for the most luminous grains in the surface; shapes of the emissivity curves for silicate and ice-silicate grains are independent of r in the Rayleigh limit (see Figure 5.1).

6. The radial locations of condensation boundaries in disk surface layers move outward approximately as T_*^2. Consequently, as T_* increases, surface emission from water ice noticeably diminishes.

7. Reducing the height of the disk photosphere by reducing the scaling parameter H/h (thereby crudely modelling the effects of vertical settling of dust) lowers the amount of stellar radiation intercepted and reprocessed by the disk. Emission at $\lambda \lesssim 200\,\mu$m scales nearly linearly with H/h. Radiation at these wavelengths arises from the optically thick interior and from the optically thin surface and scales as $\sin\alpha$ (see §2.4.2). Radiation at $\lambda \gtrsim 200\,\mu$m—the Rayleigh-Jeans regime—is less sensitive to H/h; here νF_ν is proportional to T_i which is approximately proportional to $(\sin\alpha)^{0.25}$.

8. Surface SEDs vary negligibly with $r_{max,s}$. For our standard slope of the size distribution

($q_s = 3.5$), silicate and ice-silicate particles having radii $r_* \approx 0.5\,\mu$m absorb the bulk of the radiation from the T Tauri star; surface grains having $r \approx 0.2$–$1\,\mu$m are responsible for absorbing ~50% of the incident stellar radiation. If $r_{max,s} > r_*$, the SED is unchanged because those grains having $r \gg r_*$ are insufficiently numerous to be significant absorbers of radiation. If $r_{max,s} < r_*$, the SED remains unaltered because the Rayleigh limit still obtains.

5.5 Fitting SEDs of T Tauri and HAeBe Stars

Observations and fitted theoretical models are displayed in Figures 5.8–5.13 for 4 HAeBe stars and 2 T Tauri stars. For five of our sources (AB Aur, MWC 480, HD 36112, CQ Tau, AA Tau), high-resolution ($\Delta\lambda = 0.2\,\mu$m) ISO LWS spectra between 43 and 195 μm are available. Visual extinctions for our sample range from $A_V = 0.3$ mag (MWC 480) to 1.6 mag (CQ Tau). These modest values imply that the central stars are not significantly occulted by the flared outer edges of their disks and that the simplifying assumption of a face-on viewing geometry is adequate for computation of SEDs.

For each source, a model SED is fitted to the ISO LWS scan (if available), millimeter wavelength fluxes, and ~3–25 μm photometric data. In 3 of the sources (AA Tau, CQ Tau, and MWC 480), ISO fluxes are greater than corresponding IRAS (Infrared Astronomical Satellite) fluxes at 60 and 100 μm by factors of ~2–3. The origin of the discrepancies is not known. One possible explanation is that IRAS fluxes at 60 and 100 μm are underestimated due to subtraction of overestimated background fluxes in the large IRAS beam. Where there are discrepancies, preference is given to the ISO LWS data for which the beam area is ~2 times smaller than that of IRAS. For HD 36112 and AB Aur, there is excellent agreement between ISO and IRAS. In fitting the spectra, preference is given also to the central portions of the ISO scans between 50 and 170 μm where individual detectors overlap in wavelength coverage and measured fluxes are consequently more reliable.

In practice, we choose to fix $q_s = 3.5$; smaller values ($q_s < 3$) seem unlikely since they would imply that the largest grains, which tend to settle out of surface layers most quickly, dominate the geometric cross section; the SED is insensitive to larger values ($q_s > 4$). We

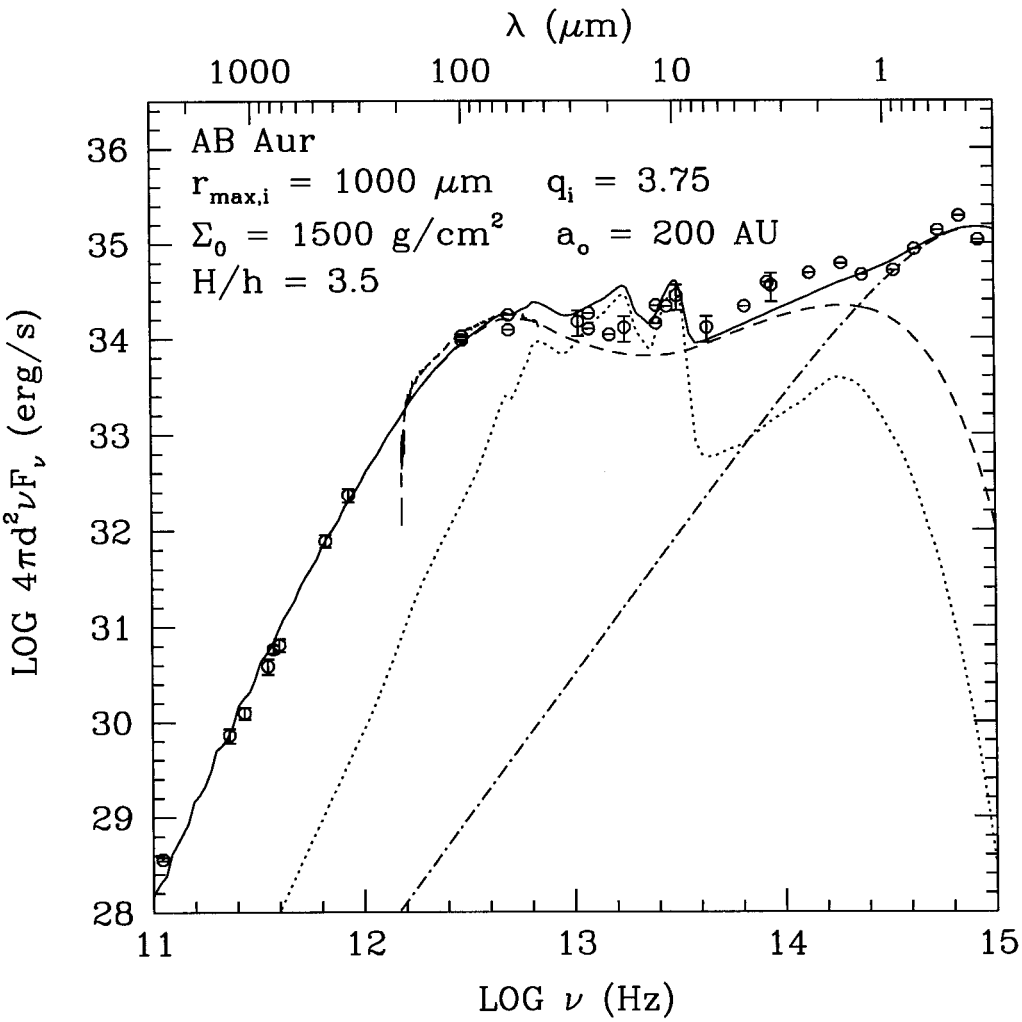

Figure 5.8: *Refined 2-layer model fitted to data for AB Aur.*

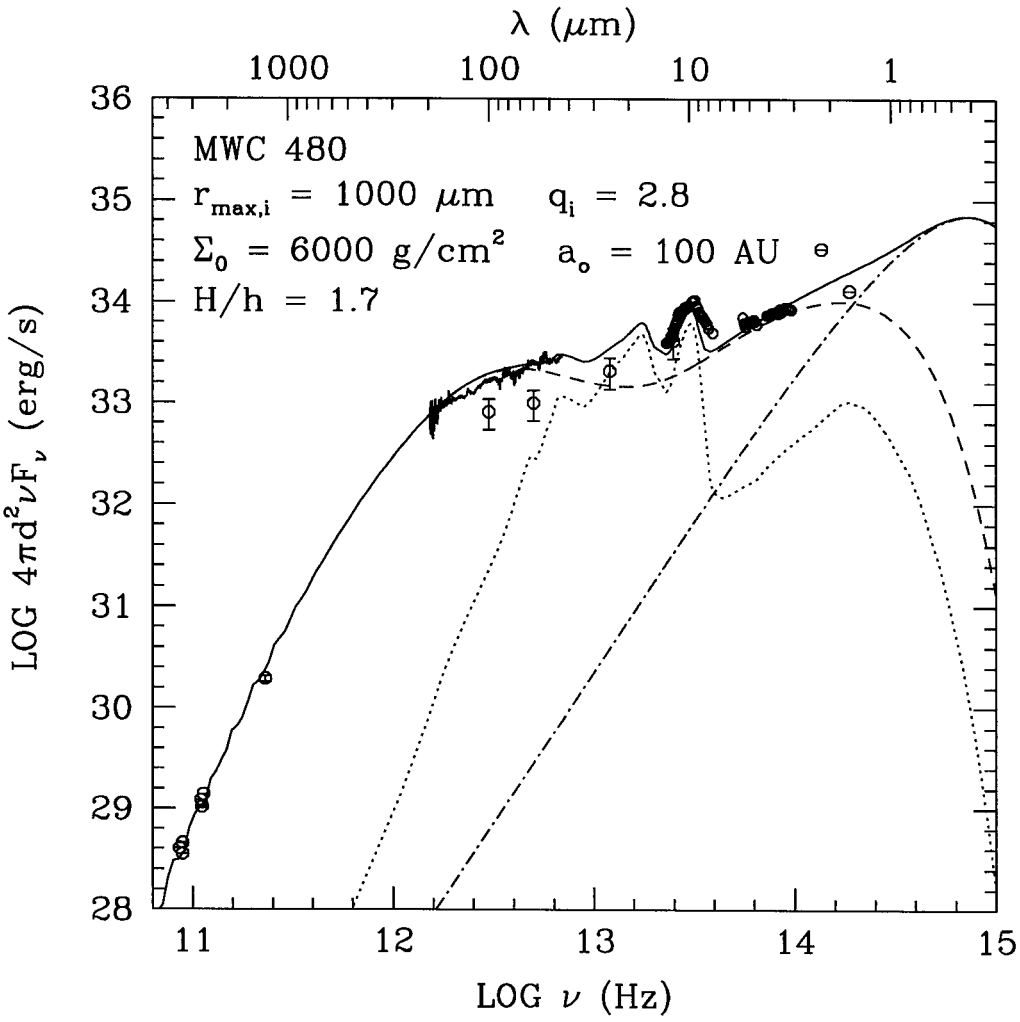

Figure 5.9: *Refined 2-layer model fitted to data for MWC 480.*

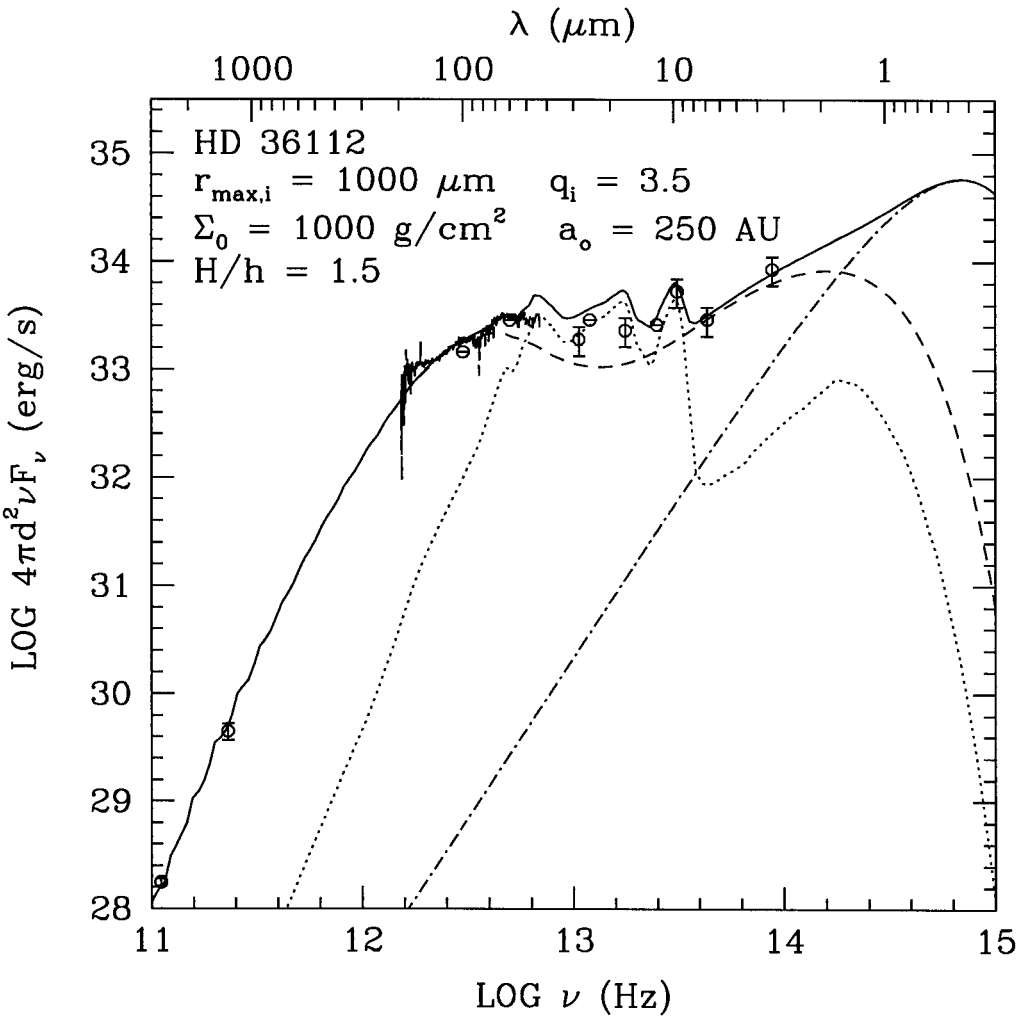

Figure 5.10: *Refined 2-layer model fitted to data for HD 36112.*

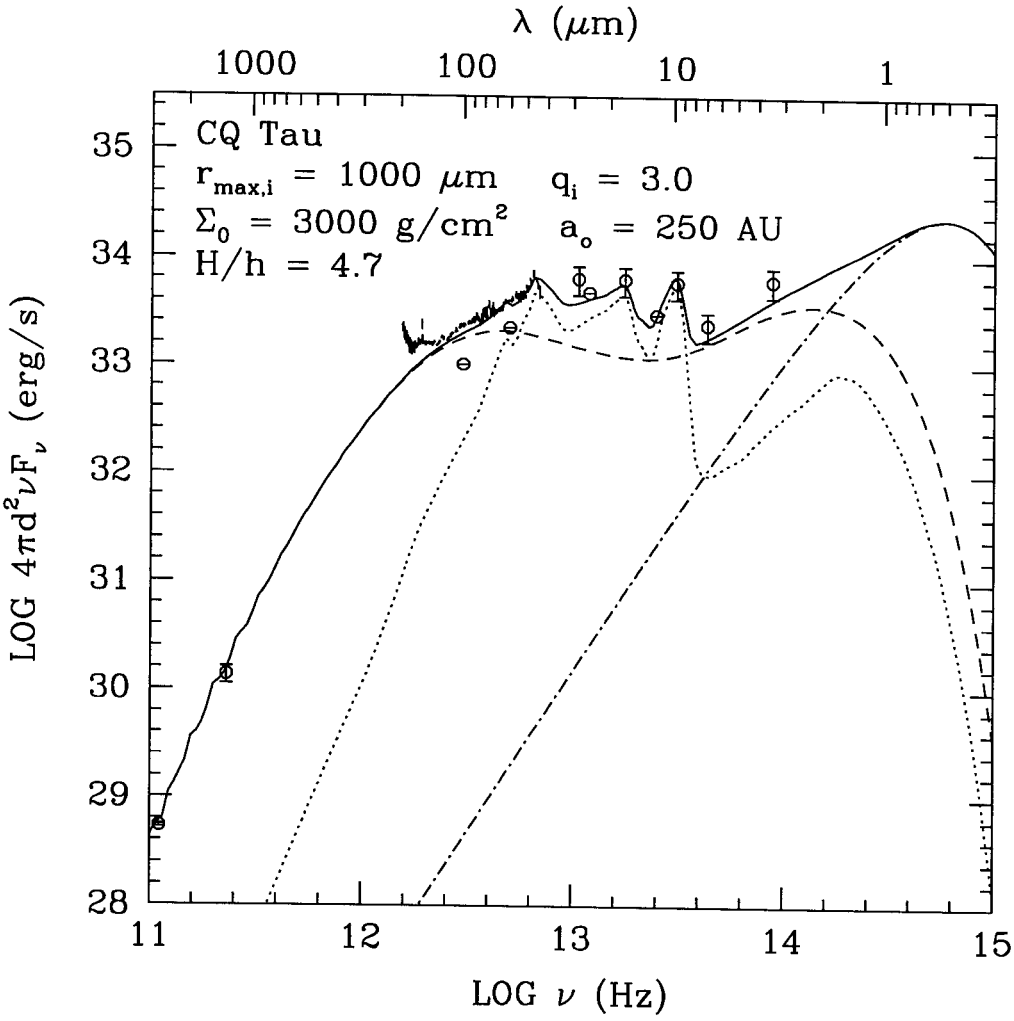

Figure 5.11: *Refined 2-layer model fitted to data for CQ Tau.*

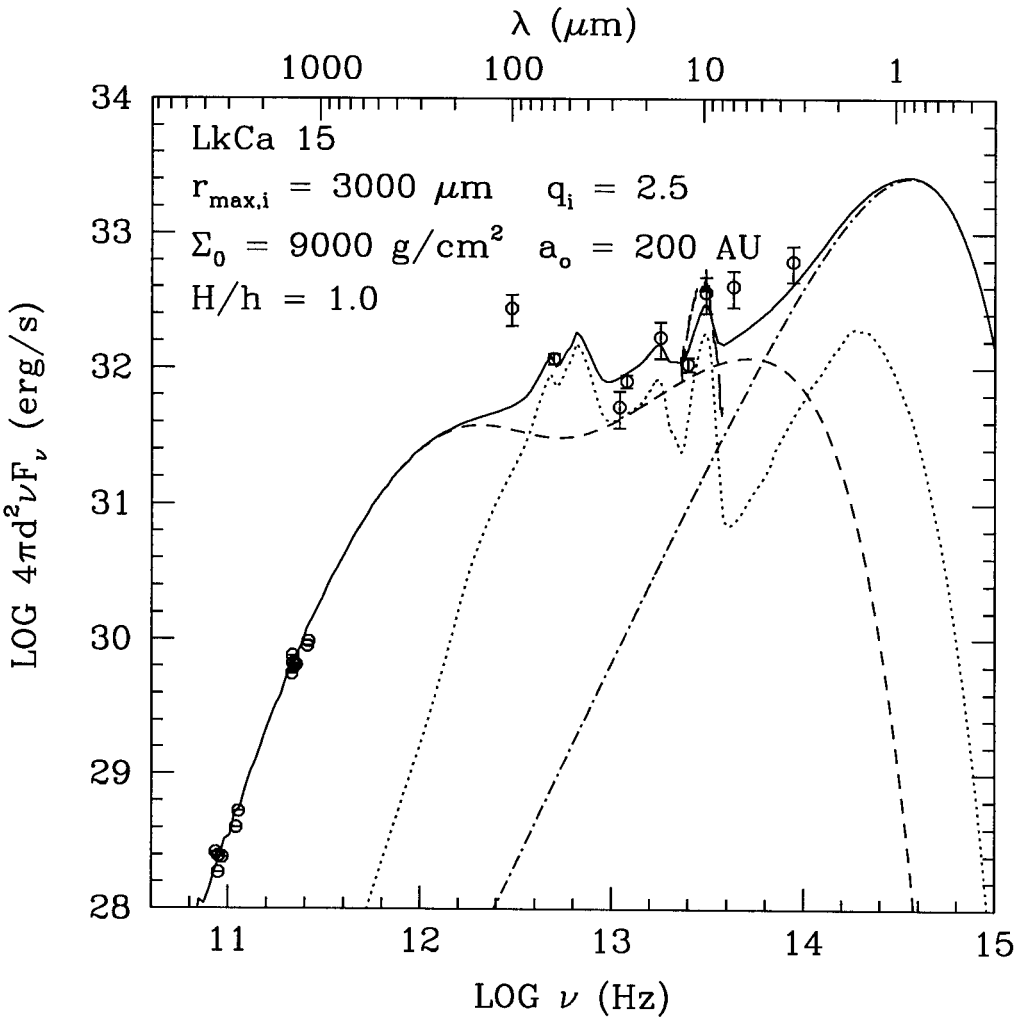

Figure 5.12: *Refined 2-layer model fitted to data for LkCa 15.*

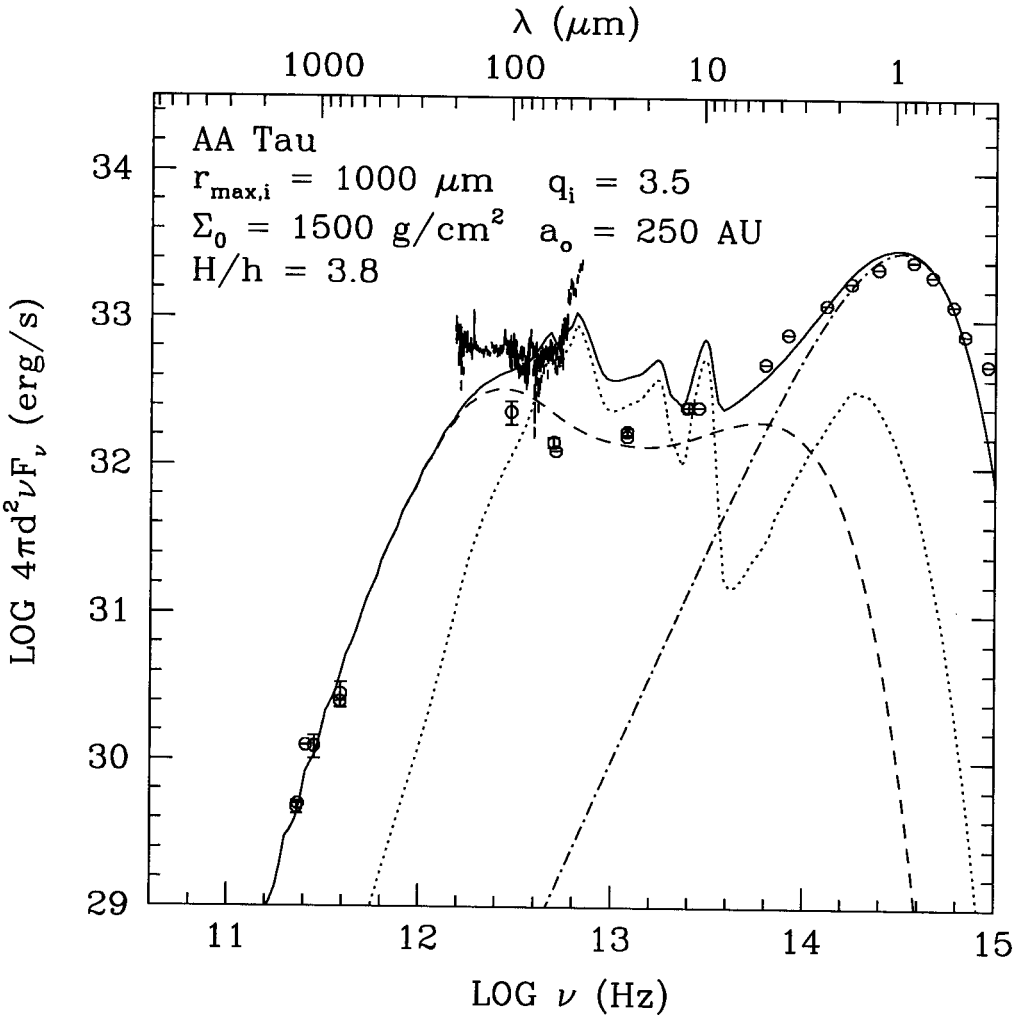

Figure 5.13: *Refined 2-layer model fitted to data for AA Tau*

Table 5.2: Fitted Parameters[a] of Herbig AeBe Star/Disk Systems[b]

Parameter	AB Aur	MWC 480	HD 36112	CQ Tau
T_* (K)	10210	8890	8465	7130
$R_*(R_\odot)$	2.3	2.1	2.1	1.9
$M_*(M_\odot)$	2.5	2.3	2.2	1.7
d (pc)	160	140	150	150
$\Sigma_0(\mathrm{g\,cm^{-2}})$[c]	1000	8000	1000	3000
a_o (AU)[c]	200	100	250	250
H/h	3.5	1.7	1.5	4.7
q_i[c]	3.75	2.8	3.5	3.0
$r_{max,i}$ (μm)[c]	1000	1000	1000	1000

[a] For all sources, we fix $q_s = 3.5$, $r_{max,s} = 1\,\mu$m, and $p = 1.5$. See text for rationale.

[b] Stellar parameters and distances for HAeBe stars are taken from Mannings & Sargent (1997).

[c] The continuum SED is largely degenerate with respect to simultaneous changes in Σ_0, $r_{max,i}$, q_i, and a_o. The values shown here are not uniquely constrained.

also fix $r_{max,s} = 1\,\mu$m and $p = 1.5$ for all models. The discussion in the previous section demonstrates that the SED is insensitive to $r_{max,s}$ once q_s is fixed at 3.5, and that Σ_0 and p affect the SED in similar ways.

Tables 5.2 and 5.3 contain the fitted parameters for our sample. The fits are intended to be illustrative; no attempt is made to minimize fit deviations in a formal, statistical sense. The results of such an analysis would not be very meaningful anyway, since the SED tends to be degenerate with respect to simultaneous changes in several of the parameters.

5.5.1 Degeneracy between Disk Mass and Grain Size Distribution

As might be gleaned from Figures 5.6 and 5.7, the values for Σ_0, p, $r_{max,i}$, q_i, and a_o presented in Tables 5.2 and 5.3 cannot be uniquely constrained by the continuum SED alone. We display one degenerate combination in Figure 5.14, where 2 models using 2 different sets of parameters are fitted to the observed data for HD 36112. We feel that the fits are of comparable quality, given the crudeness of our 2-layer model. It follows that the total disk mass (in dust) is highly uncertain; in fact, our two models for HD 36112 differ in their

total dust mass by a factor of $\sim$6.[5] Upcoming spatially resolved observations at millimeter wavelengths will help to break this degeneracy between grain size (i.e., millimeter-wave opacity) and disk surface density (Beckwith, Henning, & Nakagawa 2000).

5.5.2 Evidence for Dust Settling

The one disk parameter which appears to be most uniquely constrained is H/h, the height of the disk photosphere in units of the gas scale height. Its value is roughly proportional to the overall level of infrared excess at $\lambda \lesssim 100\,\mu$m. A few sources exhibit $H/h \approx 4$–5, the value appropriate for gas and dust that are well-mixed in interstellar proportions. However, many of our sources are fitted with significantly lower values between 1 and 4. We interpret these low values to mean that dust in disk surface layers has settled vertically towards the midplane. This was expected from our analysis in Chapter 4.

[5] Of course, a substantial amount of mass can hidden in the optically thick disk interior without any telltale signature in the continuum SED.

Table 5.3: Fitted Parameters[a] of T Tauri Star/Disk Systems[b]

Parameter	LkCa 15	AA Tau
T_* (K)	4395	4000
$R_*(R_\odot)$	1.64	2.1
$M_*(M_\odot)$	1.0	0.67
d (pc)	140	140
Σ_0(g cm^{-2})[c]	9000	1500
a_o (AU)[c]	200	250
H/h	1.0	3.8
q_i[c]	2.5	3.5
$r_{max,i}$ (μm)[c]	3000	1000

[a] For all sources, we fix $q_s = 3.5$, $r_{max,s} = 10\,\mu$m, and $p = 1.5$. See text for rationale.

[b] Stellar parameters and distances from Beckwith et al. (1990) and Webb et al. (1999).

[c] The continuum SED is largely degenerate with respect to simultaneous changes in Σ_0, $r_{max,i}$, q_i, and a_o. The values shown here are not uniquely constrained.

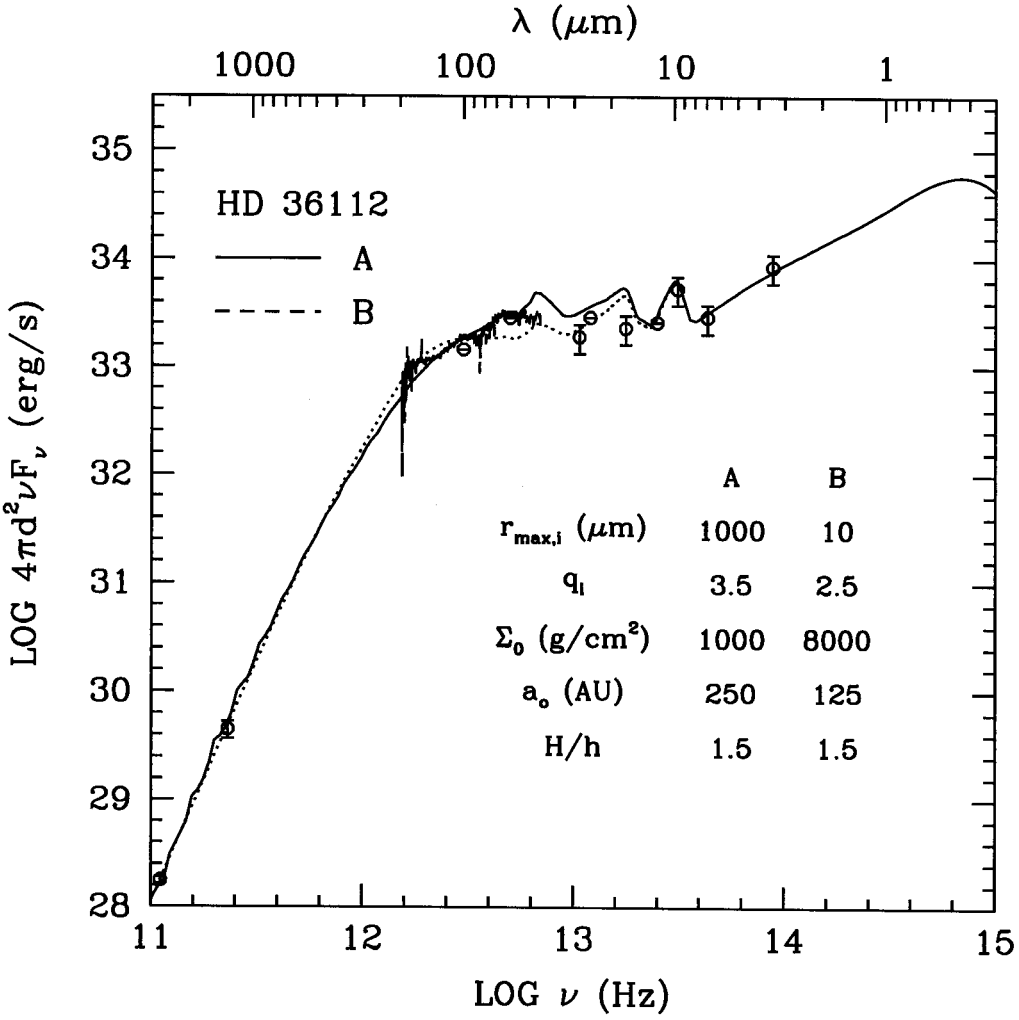

Figure 5.14: *Degeneracy between surface density and grain size for HD 36112.*

Two models having different sets of parameters are fitted to the same dataset. Smaller grain sizes (i.e., smaller millimeter-wave opacities) may be traded for larger disk masses to achieve fits of comparable quality. Note, however, that both models require $H/h = 1.5$ *to match the overall level of infrared excess at* $\lambda \lesssim 100\,\mu$m.

5.5.3 Ice and Silicate Emission Lines

Evidence for silicate emission at 10 μm from the superheated surface exists for most of our sources. In cases where high resolution spectra exist (MWC 480, LkCa 15), the exact shape of the emission feature is imperfectly fitted by our model; this indicates that surface layer silicates have allotropic states (crystalline vs. amorphous) and compositions (pyroxene vs. olivine) slightly different from the amorphous olivine that we employ. Somewhat puzzling is the lack of observed emission from the O-Si-O bending mode at 18 μm in some sources, especially those (AB Aur, HD 36112) which otherwise exhibit emission from the Si-O stretch at 10 μm.

Evidence for water ice emission from the superheated surface is present in ISO scans of sources containing the coolest central stars. The sources CQ Tau and AA Tau exhibit prominent emission bands at $\sim$45 μm, corresponding to the translational mode in water ice having the highest oscillator strength. The lattice mode at 60 μm also appears to be weakly in emission in the sources MWC 480 and HD 36112, though the relative absence of emission at 45 μm in these two sources is puzzling. The hottest star in our sample, AB Aur, exhibits no far-infrared emission bands. This is in qualitative accord with the trend noted in §5.4.3, whereby the amount of ice present in disk surface layers decreases rapidly with increasing stellar temperature.

5.5.4 Near-Infrared Excesses

Near-infrared fluxes shortward of 10 μm appear systematically underestimated in AB Aur, MWC 480, LkCa 15, and AA Tau. The relative dearth of (surface layer) emission is partly because silicate particles are particularly transparent in this wavelength regime that exists between the vibrational resonances near 10 μm and absorption due to iron impurities near 1 μm. It seems possible that thermally-spiked emission from polycyclic aromatic hydrocarbons in disk surface layers may help to fill in this transparent region of the SED. The strongest resonances are due to C-H and C-C stretching and bending modes at 3.3, 6.2, 7.7, 8.6 and 11.3 μm (Draine 1995). Upcoming high spectral resolution observations with SOFIA (Stratospheric Observatory For Infrared Astronomy) will test this hypothesis.

Other neglected but possibly relevant contributions to near-IR excesses include active accretion, which is likely to play an important role at disk radii inside a few AU (see §4.6), and reflected starlight.

Acknowledgments

This work constitutes the undergraduate senior thesis of Ryan Moo Kwang Joung, completed under the supervision of E.C. and P. Goldreich. The computer code which solves the equations of our refined model is written largely by him. We thank Ted Roush and Cornelia Jäger for providing optical constants of olivine and iron, Michelle Creech-Eakman, Charlie Qi, and Jackie Kessler for providing the observed SEDs, and Bruce Draine for suggesting that thermal spiking of PAHs might be responsible for the observed near-infrared excesses.

Bibliography

[1] Beckwith, S.V.W., Henning, Th., & Nakagawa, Y. 2000, in *Protostars and Planets IV*, preprint, http://phobos.caltech.edu/~vgm/ppiv/preprints.html

[2] Beckwith, S.V.W., Sargent, A.I., Chini, R.S., & Güsten, R. 1990, AJ, 99, 924

[3] Bell, K.R. & Lin, D.N.C. 1994, ApJ, 427, 987

[4] Bohren, C.F. & Huffman, D.R. 1983, *Absorption and Scattering of Light by Small Particles* (Wiley-Interscience)

[5] Draine, B.T. 1995, in *The Physics of the Interstellar Medium and Intergalactic Medium*, ASP Conference Series, 80, 133

[6] Jäger, C., Mutschke, C., Begemann, B., Dorschner, J., & Henning, Th. 1994, A & A, 292, 641

[7] Hudgins, D.M., Sandford, S.A., Allamandola, L.J., & Tielens, A.G.G.M. 1993, ApJS, 86, 713

[8] Lang, K. R. 1980, *Astrophysical Formulae* (Springer-Verlag)

[9] Mannings, V. & Sargent, A.I. 1997, ApJ, 490, 792

[10] Pollack, J.B., et al. 1994, ApJ, 421, 615

[11] Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 1992, *Numerical Recipes in Fortran* (Cambridge University Press)

[12] Warren, S.G. 1984, Applied Optics, 23, 1206

[13] Webb, R.A., et al. 1999, ApJ, 512, L63

Part II

FAINT KUIPER BELT OBJECTS

Chapter 6

Keck Pencil-Beam Survey for Faint Kuiper Belt Objects

This chapter draws from

Chiang & Brown 1999, AJ, 118, 1411

Abstract

We present the results of a pencil-beam survey of the Kuiper Belt using the Keck 10-m telescope. A single 0.01 square degree field is imaged 29 times for a total integration time of 4.8 hr. Combining exposures in software allows the detection of Kuiper Belt Objects (KBOs) having visual magnitude $m_V \lesssim 27.9$. *Two new KBOs are discovered. One object having* $m_V = 25.5$ *lies at a probable heliocentric distance* $R \approx 33\,\mathrm{AU}$. *The second object at* $m_V = 27.2$ *is located at* $R \approx 44\,\mathrm{AU}$. *Both KBOs have diameters of about 50 km, assuming comet-like albedos of 4%.*

Data from all surveys are pooled to construct the luminosity function from $m_R = 20$ *to 27. The cumulative number of objects per square degree,* $\Sigma(< m_R)$, *is fitted to a power law of the form* $\log_{10}\Sigma = \alpha(m_R - 23.5)$, *where the slope* $\alpha = 0.52 \pm 0.02$. *Differences between slopes reported in the literature are due mainly to which survey data are incorporated in the fit, and not to the method of analysis. The luminosity function is consistent with a power-law size distribution for objects having diameters* s = *50–500 km;* $dN \propto s^{-q}\,ds$, *where the differential size index* $q = 3.6 \pm 0.1$. *The distribution is such that the smallest objects possess most of the surface area, but the largest bodies contain the bulk of the mass. We estimate to order-of-magnitude that* $0.2 M_\oplus$ *and* 1×10^{10} *comet progenitors lie between 30 and 50 AU. Though our inferred size index nearly matches that derived by Dohnanyi (1969), it is unknown whether catastrophic collisions are responsible for shaping the size distribution. Impact strengths may increase strongly with size from 50 to 500 km, whereas the derivation by Dohnanyi (1969) assumes impact strength to be independent of size. In*

the present-day Belt, collisional lifetimes of KBOs having diameters 50–500 km exceed the age of the Solar System by at least two orders of magnitude, assuming bodies consist of solid, cohesive rock. Implications of the absence of detections of classical KBOs beyond 50 AU are discussed.

6.1 Introduction

As discussed in detail in Chapter 1, of the 3 dynamical families of Kuiper Belt Objects (KBOs) known today, the classical Kuiper Belt appears the most untouched dynamically. These objects reside in low eccentricity ($e \lesssim 0.1$), low inclination ($i \lesssim 0.5$ rad) orbits beyond 40 AU (Jewitt, Luu, & Trujillo 1998, hereafter JLT98). In the absence of any direct observations of the classical Kuiper Belt beyond 50 AU, we undertook a pilot survey utilizing the Keck 10-m telescope. A single 600s exposure on Keck can achieve a depth $m_V \approx 26$, allowing objects 100 km in diameter with comet-like albedos to be seen out to distances just beyond 50 AU. Combining exposures in software enables the detection of such bodies inside 70 AU. Our primary motivation is to observationally characterize the extent and structure of the primordial circumsolar disk. In practical terms, we constrain the KBO luminosity function out to $m_V \approx 28$, and make several inferences about the Belt based on this observable.

Observations are described in §6.2. Methods of data reduction and search strategies are set forth in §6.3. Results, including actual detections and our construction of the luminosity function from $m_R = 20$ to 28, are presented in §6.4. Implications of our results on the size, mass, and distance distributions of KBOs are discussed in §6.5. Our principal findings are summarized in §6.6.

6.2 Observations

Data were taken on 31 August 1997 UT using the Keck II 10-m telescope atop Mauna Kea in Hawaii. The Low-Resolution Imaging Spectrometer (LRIS; Oke et al. 1995) was mounted at Cassegrain focus and employed in direct imaging mode. The plate scale on

LRIS's Textronix CCD was 0.215 ″/pixel. The camera had a useable (vignetted) field of view of 5.67 × 7.34 square arcminutes (1582 × 2048 square pixels = 0.0115 square degrees). A standard V filter was used. The choice of V over R was motivated by lower sky brightness and greater solar flux at V. While some KBOs have higher reflectances at R, others also appear neutral (Tegler & Romanishin 1998).

We searched for KBOs in a single, relatively star-free field at opposition [$\alpha = 22^h54^m54^s$, $\delta = -6°20'34''$ (J2000)].[1] Twenty-nine exposures, each 600s in duration, were recorded of this field. Data were read out from the CCD through two amplifiers operating simultaneously; this procedure halved the readout time to 60s at the cost of introducing small differences in the amount of noise between chip halves. Each frame was offset in position by ∼5–100″ relative to other frames; our dithering routine enabled the construction of high-fidelity flatfields ("skyflats") from the science data themselves (see §6.3.1). Provided the Keck telescope functioned properly, our duty cycle efficiency was nearly 90%. Unforeseen crashes in the mirror alignment software limited our total effective integration time to 4.8 hr over a 6.2 hr baseline.

A Landolt field (Landolt 1992) provided photometric standards. The seeing ranged from 0.″65 to 1.″0 full-width at half-maximum (FWHM), with the median seeing equal to 0.″75.

6.3 Data Reduction and Search Strategy

Kuiper Belt Object candidates are identified by their parallax motions (of order ″/hr) against the fixed stars. We employed two search methods: a simple blinking of individual frames to visually scan for slow-moving objects, and a deep, recombinative blinking approach which blends the search algorithms of Gladman et al. (1998) and Cochran et al. (1995).

Observations of candidates over a single night are insufficient to constrain orbital parameters and to prove membership in the Kuiper Belt. Candidates might instead be eccentric, near-Earth asteroids whose apparent motions mimic those of true KBOs. However, as discussed by Luu & Jewitt (1998), the possibility of mistaken identity appears remote, since masquerading slow-moving objects have not appeared in their many surveys to date.

[1] The field happened to be located 43° away in ecliptic longitude from Neptune.

We proceed on the assumption that our (small) field is likewise uncontaminated.

6.3.1 Shallow Survey: Basic Blinking

All image processing described in this paper was performed with the Interactive Data Language (IDL) software package. The 29 science frames were first corrected for CCD bias and pixel-to-pixel variations in gain (flat-fielded). For each science frame, a tailored flatfield was constructed from the median of the other 28 dithered science frames.[2] The fact that the images were dithered ensured that each CCD pixel sampled the flat sky several times.

Each flattened frame had its mean sky value subtracted and its flux normalized by scaling eleven bright, unsaturated stars distributed across the entire frame. Position offsets required to align the dithered images were obtained by minimizing stellar residuals of frames subtracted pairwise.

Aligned images were blinked and visually scanned for slow-moving objects. Three images of comparable seeing, spaced about 1 hr apart, were blinked per session. Four triplets were blinked in all, including the first and last frames of the night. Results of this comparatively shallow survey are presented in §6.4.1.

6.3.2 Deep Survey: Forward-Reverse Recombinative Blinking

The basic idea underlying our deep survey is simple. Images are stacked and shifted on top of each other according to a hypothetical KBO proper motion. The shifted stack of images is co-added to form a recombination image. While stationary objects appear smeared in the recombination image, an object whose motion matches that assumed has its signal strengthened and appears as a single seeing disk. Thus, all collected photons are used to identify KBOs too faint to rise above the noise of an individual image.

To reduce confusion and noise in the recombination image, it is desirable to remove non-KBO sources of emission from individual frames before co-adding. Towards this end,

[2]Seven additional frames from other observations during the same night were included in the median flatfield.

we subtracted from each individual frame the median of the other 28 (aligned) frames.[3] Extended, stationary, low surface brightness emission (from resolved galaxies, for example) was thereby mostly removed from individual images. Some pixel positions did not have the full overlap of all 29 frames because of our dithering routine; these were purged to ensure uniform statistics.

Cosmic rays and asteroid streaks remained in the median-subtracted frames. Substantial residuals from stationary point sources were also left behind, a consequence of frame-to-frame seeing variations. All three non-KBO sources of emission were largely eliminated by clipping high-valued pixels from the shifted stack of images. After experimenting with various schemes, we decided to clip the 5 highest values from each column of 29 pixels and average the remaining 24 values.[4] Columns not having the full overlap of all 29 frames due to the shifting process were purged altogether. Finally, to the clipped mean image we added a positive constant frame to restore the average background level to zero. The resulting (rectangular) array constituted our recombination image, which appeared satisfyingly clean aside from a few well-localized and easily recognizable residuals from bloomed stars.

The proper motion vector of a KBO is described by its amplitude, μ, and its apparent inclination angle, θ, relative to the ecliptic as seen on the CCD. Following Gladman et al. (1998), we visually searched for KBO candidates by blinking, in any one session, 4 recombination images corresponding to 4 successively higher amplitudes along one inclination. Objects characteristically came into focus and then smeared as their actual rates of motion were approached and passed. Recombination amplitudes ranged from $\mu = 1.1$ to 6.3 ″/hr in steps of $\Delta\mu = 0.4$″/hr. Inclinations ranged from $\theta = -5$ to 5° in steps of $\Delta\theta = 5°$.[5] These ranges cover proper motions (as seen at opposition) of KBOs moving on prograde, circular, heliocentric orbits with semi-major axes R of 20–120 AU and actual inclinations i of up to 30°.[6]

[3]Image processing for the deep survey began with the sky-subtracted, flux-normalized images from the shallow survey.

[4]Taking the median of all 29 frames as an example of an alternative scheme generated a still noisier background than averaging 24 frames ($\sigma_{median-29}/\sigma_{average-24} \approx \sqrt{\frac{\pi}{2}\frac{24}{29}} \approx 1.14$).

[5]Negative (positive) θ implies motion near a descending (ascending) node.

[6]A small correction term due to the fact that our field was 4°.5 away from opposition in ecliptic longitude was included in calculating these ranges.

Roughly 130 artificial KBOs were implanted at random locations and searched for simultaneously with true KBO candidates. Their magnitudes were spread uniformly between $m_V = 26$ and 29, and their orbital parameters were chosen randomly within the ranges cited above. Their presence in recombination images trained the eye to recognize *bona fide* KBOs, and their rate of recovery provided an estimate of true KBO detection efficiency as a function of magnitude. Differences between artificial objects' given and recovered properties (δm_V, $\delta\mu$, $\delta\theta$) provided estimates of systematic errors in the parameters of true candidates. Recombination spacings were just small enough to detect artificially implanted KBOs in at least two recombination frames.

A list of KBO candidates was made containing objects (including artificial ones) which (1) focussed and de-focussed in the correct manner, (2) appeared in at least 2 adjacent recombination images, (3) did not appear as a single hot pixel in any one image, and (4) were not situated too close to the noisy environs of stellar/asteroidal residuals. Objects in this list had their magnitudes and proper motions subsequently refined on a grid of resolution $\Delta\mu = 0.1''$/hr and $\Delta\theta = 1.^\circ25$. This process involved extracting square subframes 20 pixels wide surrounding each candidate and recombining them on the finer grid. Simply selecting the grid point (μ, θ) for which counts inside a circular sampling aperture were maximized proved too simplistic a procedure. Often the maximum-count image simply pushed hot noise pixels into our sampling aperture. In practice, we selected the best recombination image based on visual appearance, a well-behaved flux profile, and in the few cases where we could not decide, maximal counts.

Without a second night to confirm the reality of our candidates, visual surveys of this kind are more prone to false detections. Even apart from human bias, noisy pixels may still conspire to masquerade as slow-moving objects. To estimate the number of false detections in our candidate list, we repeated our entire deep search on images recombined in the reverse direction. Reverse in this case actually means in the apparent prograde direction, since proper motions of KBOs at opposition are dominated by the Earth's parallax motion and must appear retrograde. Artificial, apparently prograde objects were also randomly inserted in individual frames and searched for in images recombined in the reverse direction. In this reverse survey, which suffered the same kinds of errors as afflicted the forward

survey, objects that fulfilled the four requirements listed above and that turned out not to be artificially inserted were deemed chance alignments of noise. We refer to these as "reverse survey noise objects." Statistical confidence in the detection of real KBOs demands that the number of candidates detected in the forward survey exceed the number detected in the reverse survey plus the uncertainty in the latter number. In practice, we blinked recombination frames without knowing whether they were recombined in the forward or reverse directions, thereby avoiding another potential source of human bias.

6.4 Results

6.4.1 Shallow Survey Results

One KBO was discovered by blinking individual images. The object, hereafter OBJ1, appears at the $\sim 5.5\sigma$ level in 21 out of 29 frames. In the other 8 frames, light from OBJ1 had fallen off the CCD chip as a consequence of our dithering routine. Figure 6.1 displays our newly discovered object in three consecutive exposures, and Table 6.1 summarizes its measured and inferred properties. Its motion over 6.2 hr is consistent with being uniform; a best-fit line through centroid positions yields proper motion parameters $\mu = 3.83''$/hr and $\theta = -1.^{\circ}0$. The corresponding heliocentric distance and inclination for an assumed circular orbit are $R = 32.9$ AU and $i = 4.^{\circ}5$. With a measured visual magnitude of $m_V = 25.5 \pm 0.3$ (1σ dispersion among 21 measurements), the object is 56 ± 6 km in diameter, assuming it has a comet-like visual albedo of 0.04 (Allen 1973).

The area of sky covered by our shallow survey was A_s = 0.0102 square degrees, after correcting approximately for dithering losses (-11%) and area taken up by bright stars and galaxies (-0.5%). A discussion of the cumulative luminosity function is reserved for §6.4.3.

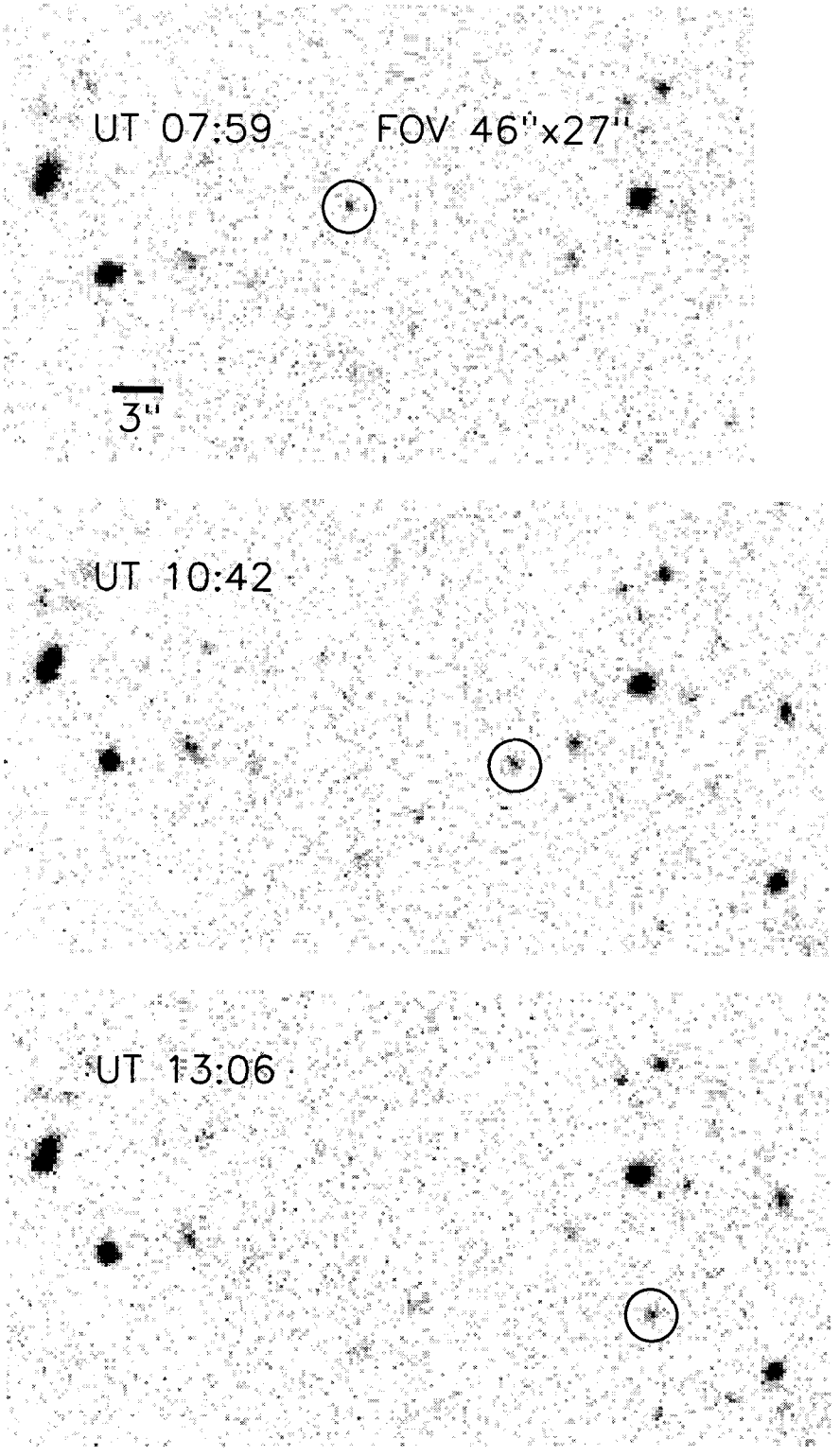

Figure 6.1: *Individual exposures of OBJ1.*

Time and angular scale are indicated. The object appears in a total of 21 frames.

6.4.2 Deep Survey Results

6.4.2.1 Artificial Object Recovery

After refining estimates of candidates' magnitudes and proper motions, we culled artificial objects from the candidate list. Figure 6.2 displays our recovery rate η of artificial objects as a function of their given m_V, for both forward and reverse surveys. In both surveys, the rate of recovery was similar, falling from 100% near $m_V = 27.3$ to 0% at $m_V = 28.4$. The datasets were combined and fitted to the function

$$\eta(m_V) = \frac{1}{2}\left[1 - \tanh\left(\frac{m_V - m_V(50\%)}{W}\right)\right] \tag{6.1}$$

(Gladman et al. 1998). The fit yields a detection efficiency which falls to 50% at $m_V(50\%)$ = 27.94, over a characteristic width $W = 0.38$ mag. Satisfyingly, $m_V(50\%)$ is only 0.04 mag brighter than the nominal 3σ limit obtained by reducing the noise of an individual image by $\sqrt{24}$.

Figure 6.3 plots δm_V, $\delta\mu$, and $\delta\theta$—differences between artificial objects' given and recovered properties—versus their given magnitude. To clarify possible trends with increasing magnitude, we also plot averages and standard deviations within bins of width 0.5 mag; these points are positioned at the centers of each bin.

We note first that there is no significant bias in our estimation of parameters; averages $\overline{\delta m_V}$, $\overline{\delta\mu}$, and $\overline{\delta\theta}$ are consistent with being zero. The scatter, however, is significant. We adopt the scatter in δm_V as our estimate of the uncertainty in true candidates' magnitudes; the 1σ dispersion increases from 0.22 mag near $m_V = 27$ to 0.33 mag near $m_V = 28$. The analogous 1σ uncertainty in μ ranges from 0.05 to 0.17 ″/hr, and the 1σ uncertainty in θ ranges from 2.°3 to 3.°2. These results for μ and θ appear reasonable. A difference of $\Delta\mu = 0.1$ ″/hr over a time $\Delta t = 6.2$ hr smears images by $\Delta\mu \times \Delta t = 0.''6$—about half the value of the worst seeing during our observations. A difference of $\Delta\theta = 2.°5$ at a fixed, typical amplitude of $\mu = 3$ ″/hr smears images over a comparable distance: $\mu \times \Delta\theta \times \Delta t = 0.''8$.

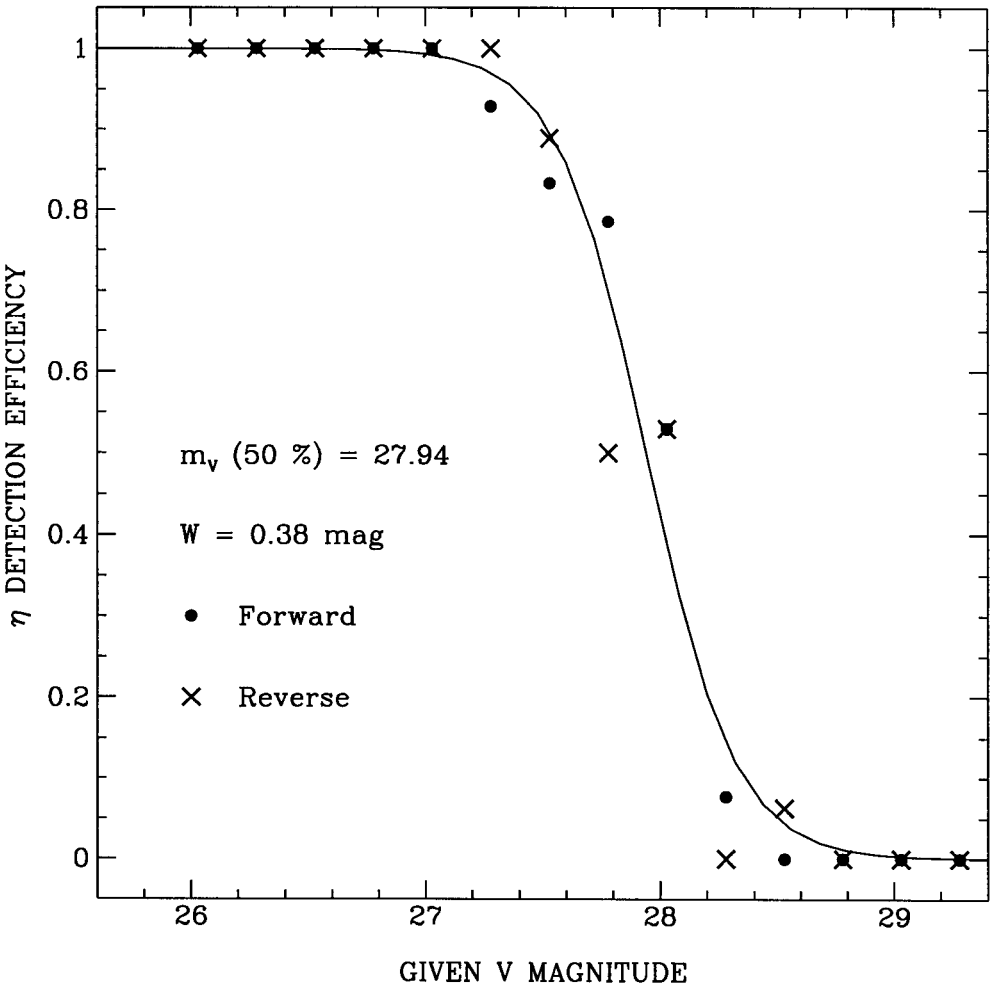

Figure 6.2: *Rate of recovery of artificially implanted objects versus their given magnitude. Detection efficiencies from forward and reverse surveys are averaged and fitted to equation (6.1), shown as a solid line.*

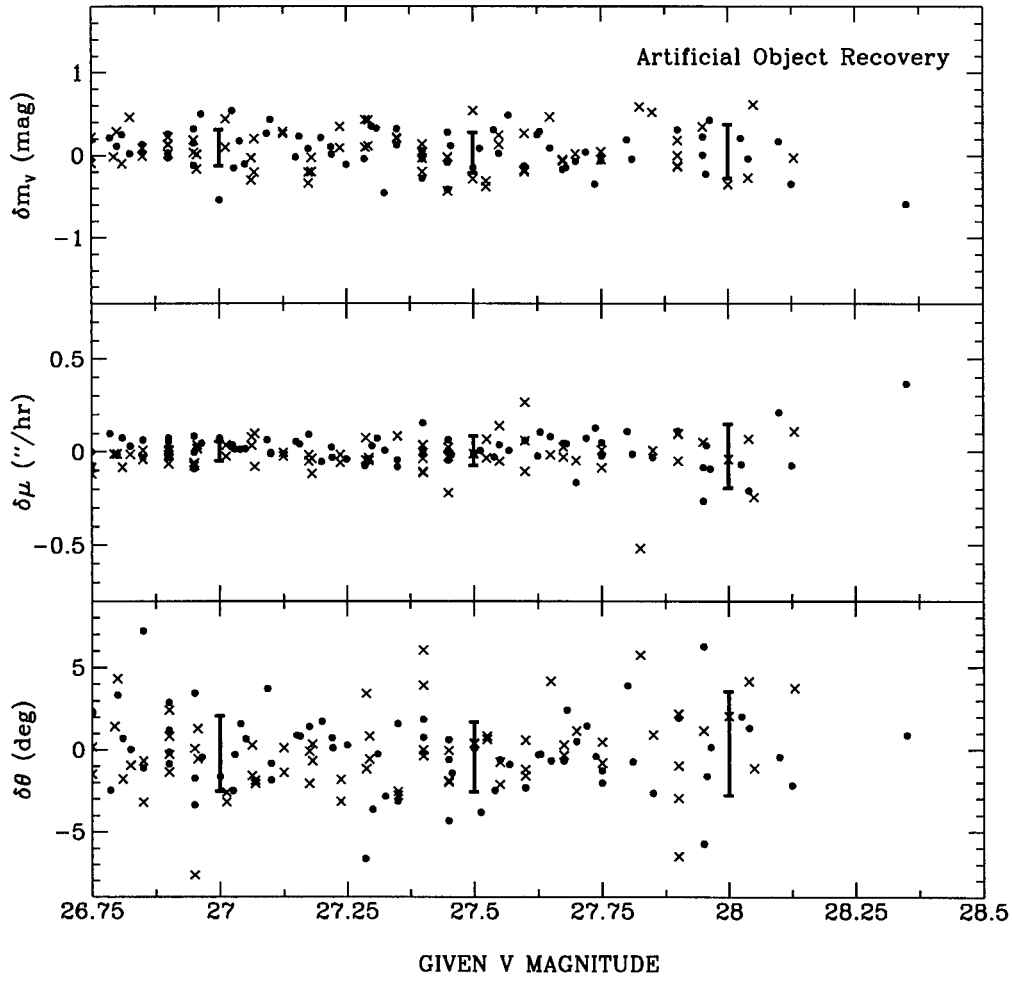

Figure 6.3: *Differences between given and recovered properties of artificial objects. Solid circles indicate objects recovered in the forward survey, and crosses denote objects recovered in the reverse survey. Error bars reflect $\pm 1\sigma$ dispersions in bins of width 0.5 mag. Since average differences are consistent with being zero, we conclude that our measurements of m_V, μ, and θ are not biased. We adopt the dispersions to be our measurement uncertainties.*

6.4.2.2 True Object Discovery and Upper Limits

A second KBO, hereafter OBJ2, was discovered by blinking recombination frames. Figure 6.4 presents the best recombination image of OBJ2, surrounded by images of the same object recombined at adjacent points on the (μ, θ) grid. Its smearing pattern is identical to those of artificially planted objects having similar motions. Properties of OBJ2 are summarized in Table 6.1. Its visual magnitude is $m_V = 27.22 \pm 0.22$, and its proper motion parameters are $\mu = 2.92 \pm 0.05$ ″/hr and $\theta = 0° \pm 2°.3$. Alternative recombination frames for OBJ2 were constructed by clipping the top 1 pixel out of each column of 29 pixels and then averaging the remaining 28 values. Exactly the same parameters for OBJ2 were obtained. If we assume a comet-like visual albedo of $p_V = 0.04$, these measurements are consistent with those of an object 46 ± 6 km in diameter, occupying a circular, uninclined orbit of semi-major axis $R = 43.9 \pm 0.8$ AU.

Confidence in the reality of OBJ2 is further bolstered by Figure 6.5, in which we compare cumulative numbers of objects detected in forward and reverse surveys. No false alarm went off in the reverse survey at the magnitude of OBJ2; the object distinguishes itself as the brightest detection at 5.5σ.

By contrast, we view all candidates in the forward survey fainter than $m_V = 27.4$ as false detections, partly because their numbers do not exceed those in the reverse survey. No object in both surveys is as visually convincing as OBJ2; many other candidates vanished at several (but not all) adjacent recombination gridpoints. Moreover, regarding the last of the 4 search criteria set forth in §6.3.2, it was occasionally unclear when an object was "too close" to a smeared stellar residual. Thus, some of our detections fainter than $m_V = 27.4$ undoubtedly arise from the confusing noise of bloomed stars (OBJ2 is far removed from any such noise). We use the population of noise objects detected in the reverse survey to set upper limits on the cumulative sky density of KBOs fainter than $m_V = 27.4$. Details of this calculation follow in the next section.

The area searched in our deep survey is less than that of our shallow survey because of the shifting process. Areal losses ranged from 7–13% depending on the value of μ. To simplify the analysis, we adopt an average loss of -10%; the error introduced is negligible

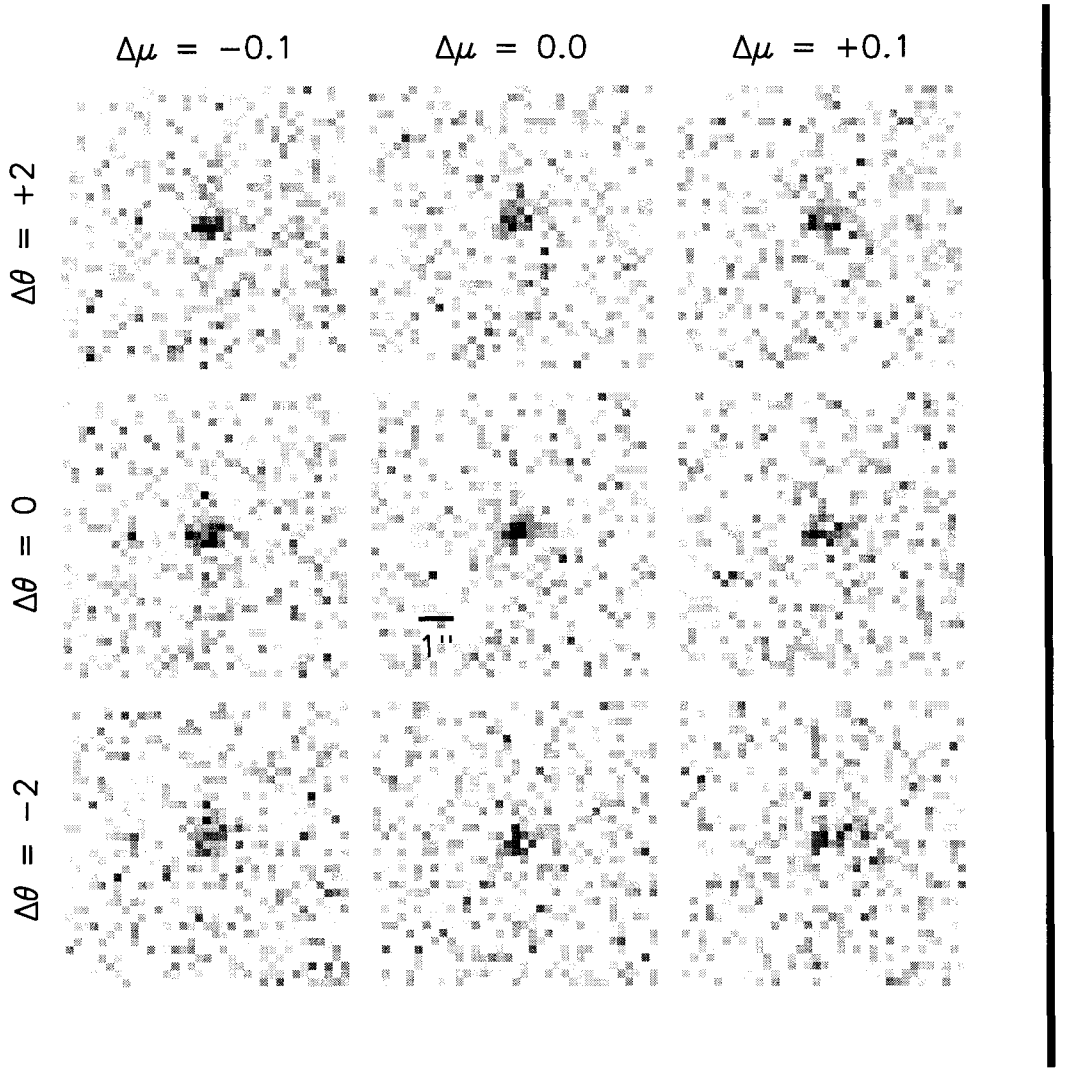

Figure 6.4: *Recombination images of OBJ2.*

The central image is the best recombination image. Surrounding it are images recombined at adjacent points on the (μ, θ) grid. Each panel to the right advances $\Delta\mu = 0.1''/\mathrm{hr}$. Each panel towards the top of the page advances $\Delta\theta = 2°$.

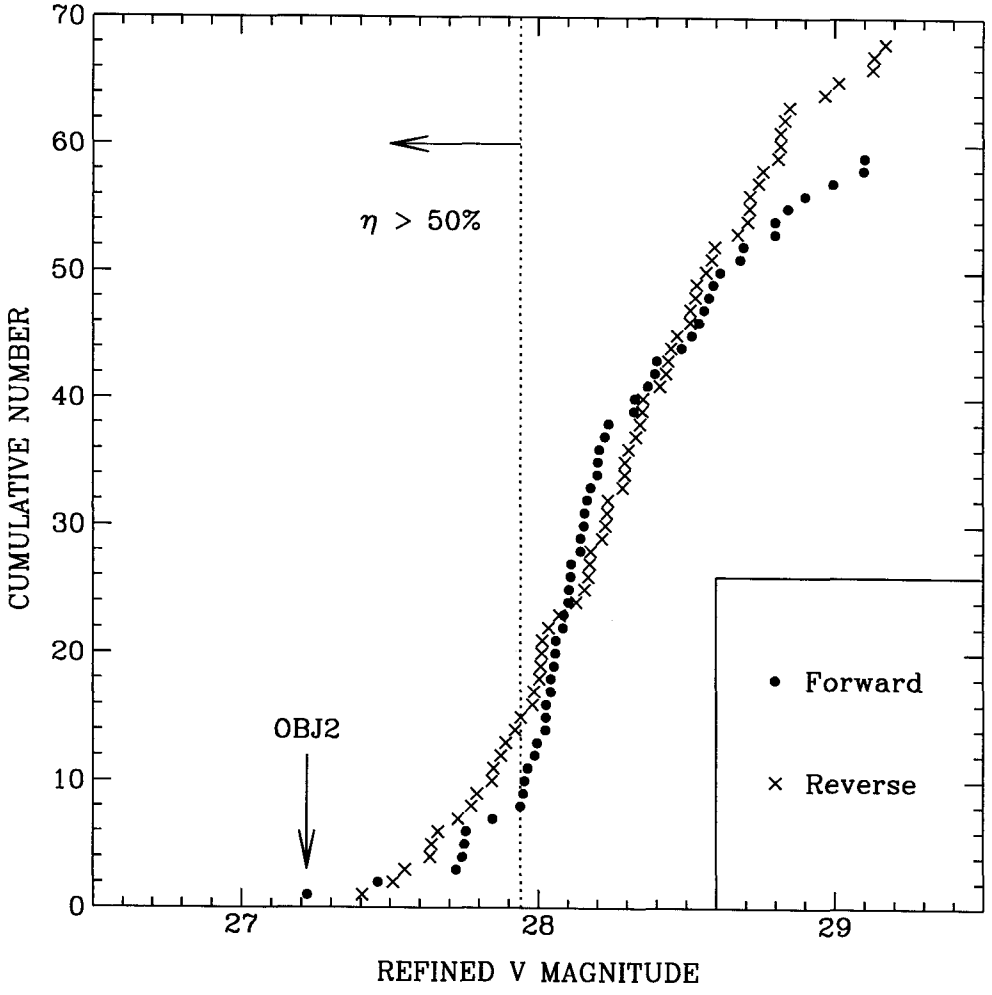

Figure 6.5: *Comparison of the number of KBO candidates found in the deep forward survey and the number of noise objects detected in the reverse survey.*

At $m_V < 27.4$*, the only object detected is OBJ2, and no noise object is bright enough to confuse the identity of OBJ2 as a true KBO. At* $m_V > 27.4$*, the number of KBO candidates never significantly exceeds the number of false alarms, and we can only compute upper limits on the sky density.*

compared to Poisson uncertainties in the number of objects detected. Corrected for additional losses due to stellar/asteroidal residuals (-1%), our deep survey area equals 0.009 square degrees $\equiv A_d$.

6.4.3 Cumulative Luminosity Function

Figure 6.6 displays our estimates of the cumulative KBO sky density, $\Sigma(< m_R)$, together with estimates made by various other groups. We emphasize that each survey's points represent estimates made independently of all other groups; i.e., survey areas have not been added.[7] Surveys conducted in V were included by assuming a solar color, V-R = 0.36, corresponding to a neutrally reflective KBO (red albedo $p_R = p_V$).

From our detection of OBJ1, we independently estimate $\Sigma(m_R < 25.14) = 98 \pm 98(1\sigma)$ objects/deg^2. Combining this result with our detection of OBJ2, which is the faintest KBO detected to date, we estimate $\Sigma(m_R < 26.86) = 209 \pm 149(1\sigma)$ objects/deg^2, where Poisson uncertainties have been added in quadrature.

Upper limits are derived at fainter magnitudes as follows. We assume that in the forward survey, the occurrence of noise objects plus real KBOs is Poissonian. The expected mean number of forward survey candidates brighter than magnitude m in survey area A_d equals $N_{\rm Noise}(< m) + \langle \eta A_d \rangle \Sigma(< m)$, where $N_{\rm Noise}(< m)$ is the mean cumulative number of noise objects, and $\langle \eta A_d \rangle$ is the efficiency-weighted survey area. We take $N_{\rm Noise}(< m) = N_R(< m)$, where $N_R(< m)$ is the cumulative number of reverse survey noise objects found. Given the number of forward survey candidates that we actually detected, $N_F(< m)$, we ask what minimum value of $\Sigma(< m)$ can be ruled out at the 99.99% confidence level ($\sim$"4σ" in Gaussian parlance):

$$\frac{(N_R + \langle \eta A_d \rangle \Sigma)^{N_F} \exp -(N_R + \langle \eta A_d \rangle \Sigma)}{N_F!} = 10^{-4}, \tag{6.2}$$

an implicit equation for Σ where the magnitude dependence has been dropped for compactness. For reverse survey noise objects brighter than $m_V(50\%)$, we take $\langle \eta A_d \rangle = 0.7 \times A_d$.

[7]The one exception to independence involves the points from Luu & Jewitt (1998), which presumably incorporate data from their previous surveys.

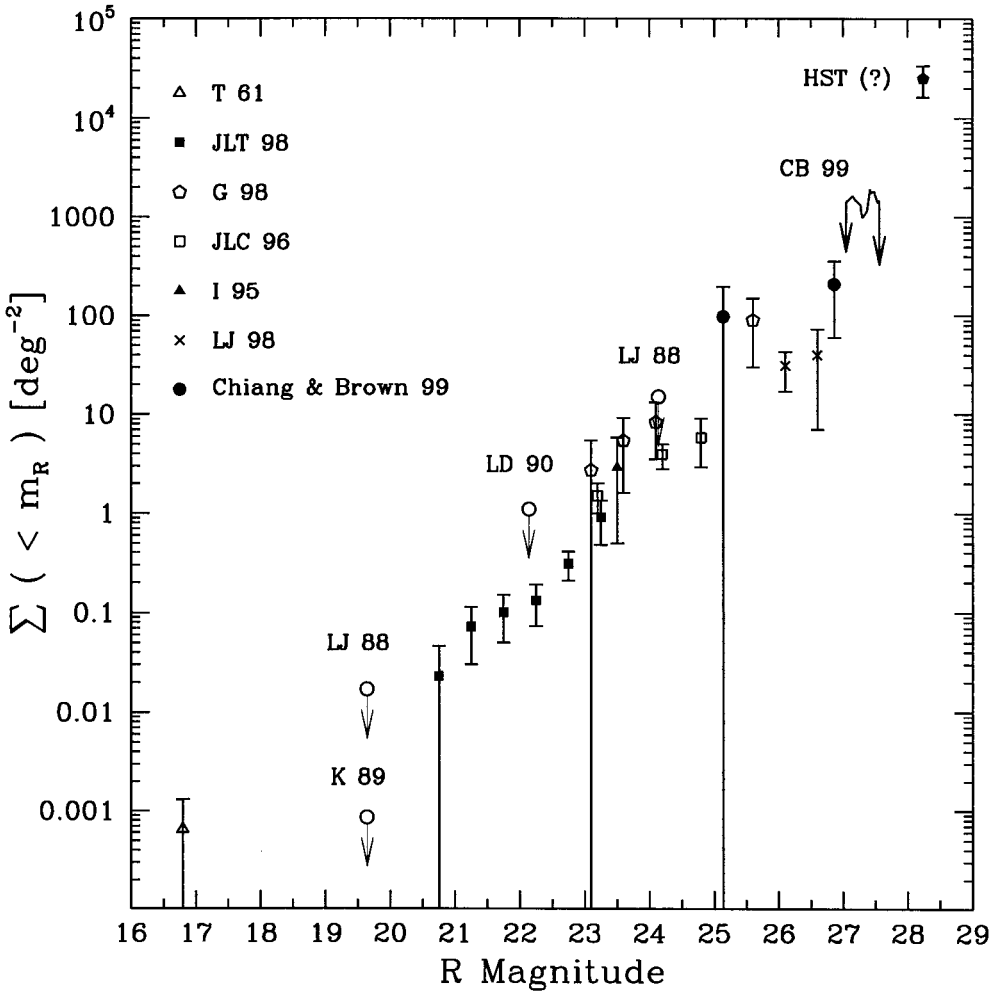

Figure 6.6: *Independent estimates of the cumulative sky density of KBOs as made by various groups.*

Abbreviations for surveys are defined in the text and in the references. Upper limits from this paper (CB99) are computed at the 99.99% ("4σ") confidence level. Upper limits from other surveys are published values at the 99% confidence level.

We do not calculate upper limits for $m_V > m_V(50\%)$, since our detection efficiency falls rapidly to zero past that magnitude (see Figure 6.2). Upper limits on Σ computed using equation (6.2) are plotted in Figures 6.6 and 6.8.

While Figure 6.6 summarizes the history of KBO surveys, quantitative results such as the slope of the luminosity function (or even the degree to which Σ resembles a single-slope power law) are better extracted from a fairer pooling of the data. To this end, we imagine the areas from all surveys as being combined into one giant frame over which the detection efficiency varies. At magnitude m_i of a detected KBO,

$$\Sigma(< m_i) = \sum_{j=1}^{i} \frac{1}{\sum_{k=1}^{n} \eta_k(m_j) A_k}, \tag{6.3}$$

where m_j is the magnitude of the j^{th} brightest KBO, $\eta_k \times A_k$ is the efficiency-weighted area of the k^{th} survey, and n is the total number of surveys. Most surveys have published efficiency functions. Exceptions include the Mauna Kea-Cerro Tololo survey of Jewitt, Luu, & Chen (1996, hereafter JLC96), the Keck survey by Luu & Jewitt (1998, hereafter LJ98), the McGraw-Hill CCD survey by Luu & Jewitt (1988, hereafter LJ88), and the U.S. Naval Observatory survey by Levison & Duncan (1990, hereafter LD90). For data from JLC96, we assume η behaves in a similar manner to that described in their companion paper I (Jewitt & Luu 1995, hereafter JL95); i.e., η is assumed to fall linearly from 100% to 0% over 0.7 mag centered on published values of $m_R(50\%)$. For data from LJ98, we obtained η by consulting the lead author (Luu 1999). For the two remaining older surveys [which detected no KBOs, but which nonetheless contribute slightly to the total survey area in equation (6.3)], we adopted Heaviside step functions centered at $m_R = 24$ (LJ88) and $m_R = 22.14$ (LD90). We have verified that the conclusions of our paper are not affected by how we incorporate the latter two surveys. No photographic survey was included in the pool.

Figure 6.7 displays the results of pooling datasets according to equation (6.3). To clarify the roles played by individual surveys, we pool an incrementally larger number of surveys in Figures 6.7a through 6.7c. In these and subsequent plots, magnitudes of individual

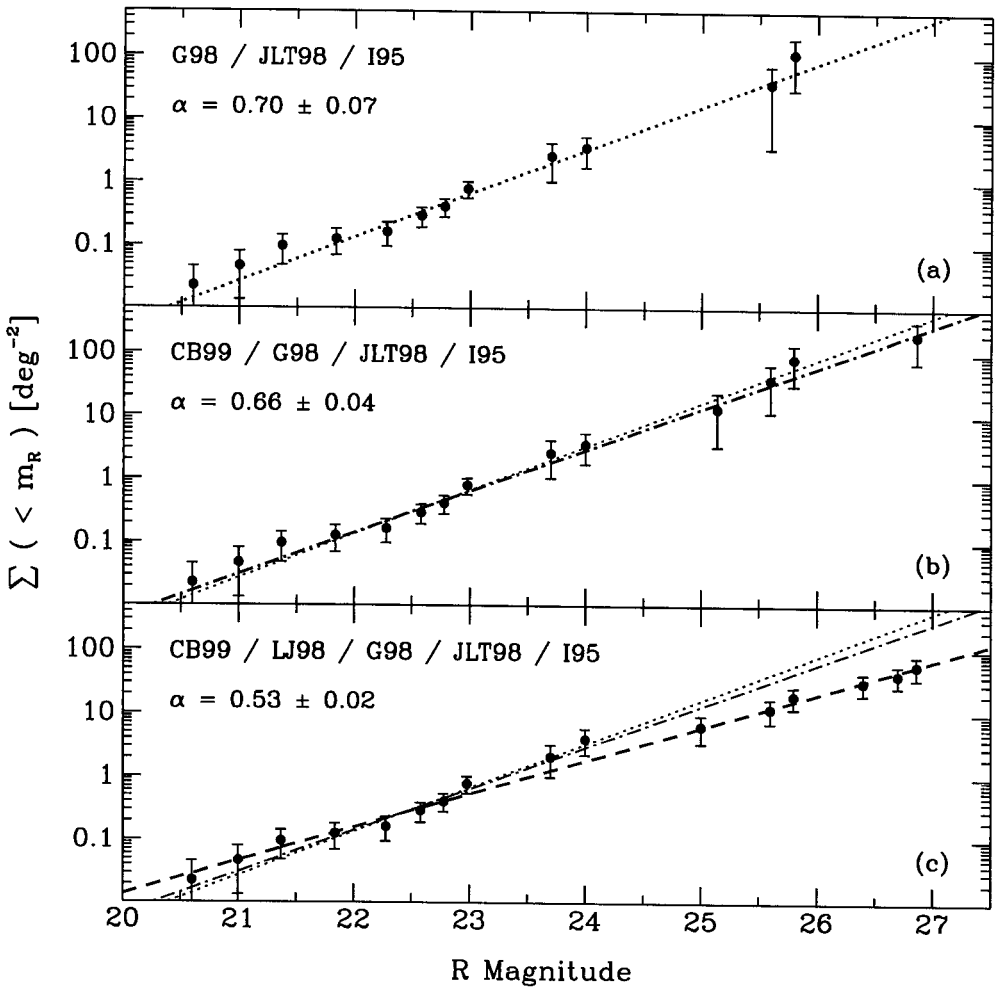

Figure 6.7: *Cumulative sky density obtained by pooling surveys according to equation (6.3).*

From panels (a) through (c), successively more surveys are pooled, as indicated by the growing list of acronyms at the top of each panel. Data are fitted by least-squares, with fits from preceding panels plotted for comparison. In panel (a), the fitted slope α of the luminosity function is identical to that derived using a maximum likelihood analysis by Gladman et al. (1998). In panel (c), the slope α decreases significantly when data from LJ98 are included.

points are identical to magnitudes m_i of individual KBOs. However, only points separated by at least $\sim$0.2 mag are plotted; this represents a minor smoothing of the dataset, but is still preferable to imposing arbitrary bin boundaries. Error bars reflect Poisson counting statistics.

Consider first Figure 6.7a, which incorporates data from Irwin, Tremaine, & Zytkow (1995), Jewitt, Luu, & Trujillo (1998), and Gladman et al. (1998) (hereafter I95, JLT98, and G98, respectively). These constitute the 3 surveys preferred by Gladman et al. (1998), excluding upper limit data. The points are well described by a power law, written in conventional notation as

$$\Sigma(< m_R) = 10^{\alpha(m_R - m_0)}\,, \tag{6.4}$$

where slope α and reference magnitude m_0 are fitted parameters. A least-squares fit to these three surveys alone yields $\alpha = 0.70 \pm 0.07$, $m_0 = 23.3 \pm 0.1$. These values coincide with those derived using a maximum likelihood analysis by Gladman et al. (1998); see their Figure 6c. We realize that least squares is not the preferred statistic for data whose errors are not Gaussian and which are correlated from point to point. However, the agreement between our result and G98's suggests that differences between slopes derived by various groups are due mainly to which surveys are kept and which are neglected, and not to the method of analysis. This will be borne out in what follows.

In Figure 6.7b, we incorporate our survey (hereafter CB99) into the pool. The slope is lowered slightly to $\alpha = 0.66 \pm 0.04$, but the change is negligible over the range of observed magnitudes. On the basis of these 4 surveys alone, our data extend the $\alpha \approx 0.7$ law to $m_R \approx 26.9$.

In Figure 6.7c, we fold in the Keck survey of LJ98. The observed faint end of the luminosity function is suppressed by the weight of their relatively sparsely populated fields. The luminosity function still resembles a single-slope power law, but the refitted slope is substantially shallower; $\alpha = 0.53 \pm 0.02$, coincident with the value given by LJ98. Though the LJ98 fields have a few times fewer objects at $m_R \approx 26.5$ compared to our CB99 fields, discrepancies are at the $\sim 1\sigma$ level or less; uncertainties in our points (see Figure 6.7b) are

large because we detect only two objects.

Finally, in Figure 6.8, the remaining surveys by JL95, JLC96, LJ88, and LD90 are assimilated into the pool. The fitted luminosity function hardly changes; for this final pool, $\alpha = 0.52 \pm 0.02$ and $m_0 = 23.5 \pm 0.06$. We note that shallow values for the fitted slope depend not only on surveys by JL95 and JLC96, but also on recent data from LJ98 (see Table 6.2). Some crude, model-dependent considerations of why values of $\alpha < 0.6$ might be preferred are given in §6.5.1.

The claimed Hubble Space Telescope (HST) detection of 2.5×10^4 objects/deg^2 at $m_R \approx 28.2$ (Cochran et al. 1995) lies $\sim$10 times above the steepest extrapolation, and $\sim$100 times above the extrapolation derived from all surveys combined. Brown, Kulkarni, & Liggett (1997) independently suggest on statistical grounds that the detections are erroneous. The increasing difficulty of reconciling the ground-based observations with the HST claim appears to support this suggestion.

6.5 Discussion

6.5.1 Size, Surface Area, and Mass Distributions

We consider a power-law differential size distribution with index q, $dN(s) \propto s^{-q}\, ds$, where $dN(s)$ is the number density of objects having diameters between s and $s + ds$. If all observed KBOs had the same albedo and were at the same heliocentric distance, the measured slope α of the cumulative luminosity function would imply a unique size index, viz.

$$q = 5.02\alpha + 1\,. \tag{6.5}$$

This relation is straightforward to derive and is first given by I95. The assumption of uniform distance is not a bad one, since KBOs detected to date have present-day heliocentric distances between 30 and 50 AU; adopting a geometric mean distance for all objects mis-estimates sizes by at most a factor of 5/3. This is less than the possible factor of 4 uncertainty in size introduced by the unknown albedo, which might range from 4% (Comet Halley) to 60% (Pluto). JLT98 employ Monte Carlo models which incorporate more real-

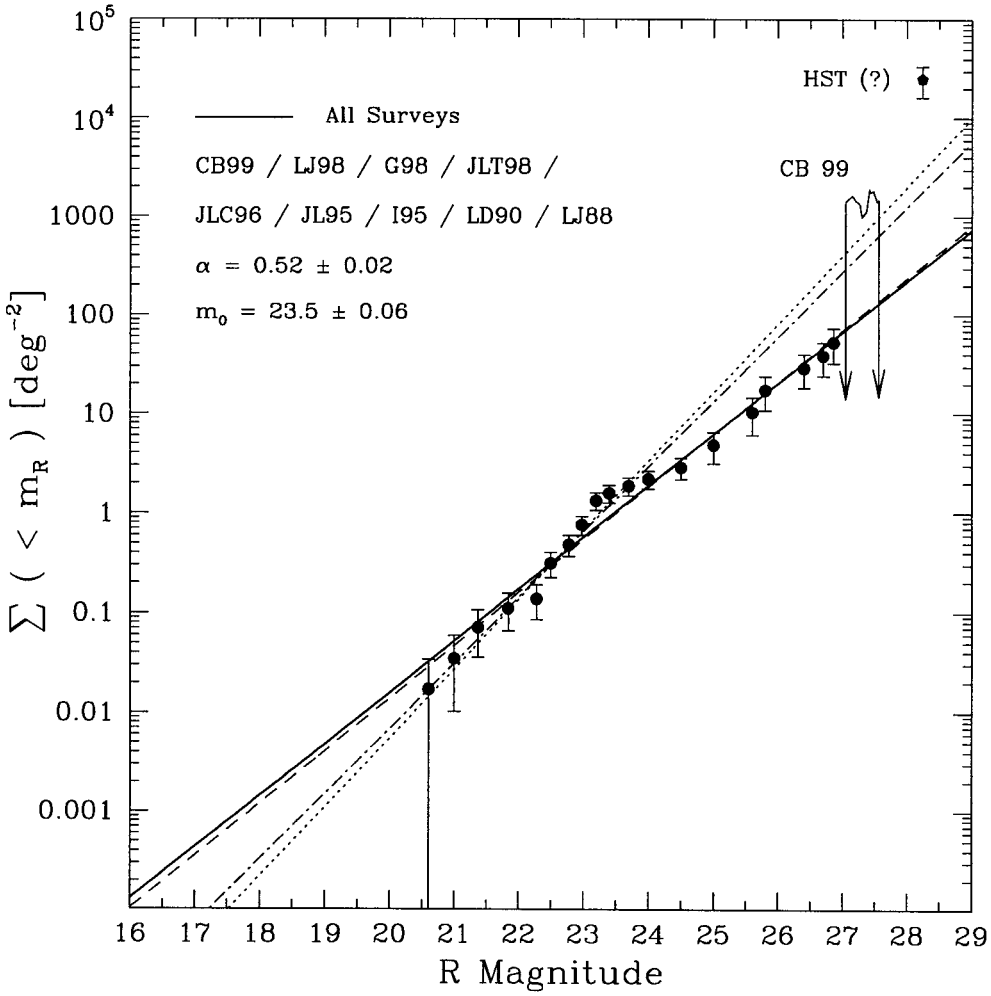

Figure 6.8: *Cumulative sky density obtained by pooling all surveys.*

The solid line is the power law fitted to all survey data. Other lines are fits from Figure 6.7, re-plotted here for comparison. Neither the HST datum nor Tombaugh's (1961) datum is included in any fit. Data from JL95 and JLC96 reinforce the shallow slope forced by data from LJ98.

istic distance distributions to extract the size index from the measured luminosity function. Their best-fit $q = 4.0 \pm 0.5$ agrees with the value obtained by inserting their measured $\alpha = 0.58 \pm 0.05$ into equation (7.13). We shall use equation (7.13) to calculate q from α below, keeping in mind that such q's may be uncertain by ± 0.5.

Table 6.2 summarizes possible values of the sky density slope α and the size index q and their implications. Depending on which surveys are incorporated, q takes values from 3.6 to 4.5. We compare these values to those of erosive disks in our Solar System. Main-belt asteroids are inferred to obey $q \approx 3.3$ in the diameter range 3–30 km (Durda, Greenberg, & Jedicke 1998). A value of $q = 3.5$ corresponds to a quasi-steady-state population for which catastrophic collisions grind as much mass per time into every size bin as they grind out, as first derived by Dohnanyi (1969). The derivation further assumes that critical specific energies for shattering and dispersal are independent of size.[8] This is a fair assumption for solid rocky targets smaller than $\sim$10 km in diameter for which internal compression due to self-gravity is negligible. For asteroids greater than 30 km in diameter, there are significant deviations from the $q \approx 3.5$ law, with slopes ranging from $q \approx 2$ to 4.5 (Durda & Dermott 1997). These deviations likely result from variations of the impact strength with size, as caused by self-gravitational effects (Durda, Greenberg, & Jedicke 1998; Melosh & Ryan 1997). Saturn's ring particles crudely fit $q \approx 3.3$ in the size range of a few centimeters to a few meters based on Voyager radio occultation data (Marouf et al. 1983), though values between 2.8 and 4.0 cannot be completely ruled out (Cuzzi et al. 1984, Weidenschilling et al. 1984). Ring optical depths are sufficiently high that particles have suffered many erosive collisions over the age of the Solar System, so that their size distribution no longer purely reflects initial conditions (Borderies, Goldreich, & Tremaine 1984).

For $q > 3$ ($\alpha > 0.4$), surface areas (geometric optical depths) are dominated by the smallest bodies. All current estimates of q imply that this is the case for the Kuiper Belt.

For $q < 4$ ($\alpha < 0.6$), the total mass is dominated by the largest bodies. If we combine all surveys, we infer a size index $q = 3.6 \pm 0.1$. We use this q to make an order-of-magnitude estimate of the mass in the observable Kuiper Belt. Nearly all KBOs in the surveys we

[8]The critical specific energy for shattering, Q^*_S, is defined as the energy per unit target mass required to produce a fragment with 50% the mass of the original target. It is smaller than Q^*_D, the energy per unit target mass required to disperse such fragments to infinity (Melosh & Ryan 1997).

Table 6.1: Properties of Detected KBOs

Object	OBJ1	OBJ2
m_V [a]	25.5	27.2
μ [b] ($''$/hr)	3.83	2.92
θ [b] (deg)	-1	0
R [c] (AU)	32.9	43.9
i [c] (deg)	4.5	0
s [d] (km)	56	46
η [e]	100%	98%
A [e] (deg^2)	0.0102	0.009

[a] Measured visual magnitude, uncertain by 0.3 (0.22) mag for OBJ1 (OBJ2).

[b] Measured proper motion amplitude and angle relative to ecliptic on CCD, respectively. For OBJ1 (OBJ2), uncertainties are 0.02 (0.05) $''$/hr and 1 (2.3) degrees.

[c] Inferred heliocentric distance and inclination, respectively, for an assumed circular orbit.

[d] Inferred diameter, assuming a visual albedo of 4%.

[e] Detection efficiency and area searched, respectively.

Table 6.2: Measured α and Inferred q

α[a]	Source	q[b]	Implication
0.52	All data (Fig. 6.8)	3.6	Mass in Largest Bodies
0.53	Omit JL95, JLC96 (Fig. 6.7c)	3.7	Mass in Largest Bodies
0.57	Omit LJ98	3.9	Mass in Largest Bodies
0.66	Omit JL95, JLC96, LJ98 (Fig. 6.7b)	4.3	Mass in Smallest Bodies
0.70	Omit JL95, JLC96, LJ98, CB99 (Fig. 6.7a)	4.5	Mass in Smallest Bodies

[a] Power-law slope of cumulative luminosity function; see equation (6.4).

[b] Differential size index derived from $q = 5.02\,\alpha + 1$, which assumes uniform albedo and distance. See §6.5.1 for discussion.

have considered have inferred diameters $s \gtrsim 50\,(0.04/p_R)^{1/2}$ km.[9] At limiting magnitude $m_R = 27$ (the V-R adjusted magnitude above which false alarms prevent additional KBO detections in our deep survey), objects having $s \gtrsim 50$km can be seen out to distances of 48 AU. For values of q and $\Sigma(m_R < 27)$ derived by combining all survey data, the total mass of the Kuiper Belt out to 48 AU is

$$M_{Belt}(a < 48\,\mathrm{AU}) \approx 0.22 \left(\frac{\Sigma(m_R < 27)}{53\,\mathrm{deg}^{-2}}\right) \left(\frac{A_{KB}}{10^4\,\mathrm{deg}^2}\right) \left(\frac{\rho}{2\,\mathrm{g\,cm}^{-3}}\right) \times \left(\frac{0.04}{p_R}\right)^{1.3} \left(\frac{s_{max}}{2000\mathrm{km}}\right)^{0.4} M_\oplus\,. \tag{6.6}$$

Here A_{KB} is the solid angle subtended by the Kuiper Belt (taken to extend $\pm 15°$ in ecliptic latitude), ρ is the internal mass density of KBOs, and s_{max} is the diameter of the largest body in the distribution (taken to be similar to Pluto).[10] Our rough estimate of $\sim$0.2 $M_\oplus$ is consistent with the upper limit of 1.3 $M_\oplus$ within 50 AU derived by Hamid, Marsden, & Whipple (1968) using measured cometary orbits (see Weissman 1995 and Malhotra, Duncan & Levison (1999) for a discussion of upper mass limits). Note that this model predicts the existence of $\sim$10 more Pluto-sized objects in the nearby Kuiper Belt.

The number of 1-10 km sized comet progenitors in the Kuiper Belt may be similarly estimated;

$$N_{Comet}(a < 48\,\mathrm{AU}) \approx 1.4{\times}10^{10} \left(\frac{\Sigma(m_R < 27)}{53\,\mathrm{deg}^{-2}}\right) \left(\frac{A_{KB}}{10^4\,\mathrm{deg}^2}\right) \left(\frac{0.04}{p_R}\right)^{1.3} \left(\frac{1\mathrm{km}}{s_c}\right)^{2.6} \mathrm{comets}\,, \tag{6.7}$$

where s_c is the minimum diameter of a comet. Our order-of-magnitude estimate compares favorably with the population of $\sim 7 \times 10^9$ comets between 30 and 50 AU required to supply the rate of Jupiter-family comets (Levison & Duncan 1997).[11]

[9] The only exception is KBO K3, for which $s = 23$km (Luu & Jewitt 1998).

[10] Our calculation ignores the fact that some surveys observe $\pm 90°$ away in ecliptic longitude from Neptune where Plutinos (KBOs in 3:2 resonance with Neptune) reach perihelion (Malhotra 1996). These surveys might be expected to find an unrepresentatively high sky density of KBOs. In fact these surveys (JL95, JLC96) find lower sky densities than other surveys; see §6.4.3 and section 5.1 of G98.

[11] The scattered KBO disk has also been proposed as an alternative source of short-period comets. Duncan

Omitting data from various surveys while preserving the same magnitude coverage in the luminosity function raises the inferred value of q and places most of the mass of the observable Kuiper Belt into the smallest objects. Gladman et al. (1998) do not incorporate data from JL95, JLC96, LJ98 and CB99. Their maximum likelihood analysis, which can and does assimilate upper limit data from Luu & Jewitt (1988) and Levison & Duncan (1990), concludes that the sky density slope $\alpha = 0.76$. Inserting this value into equation (7.13) yields a size index $q = 4.8$. As a separate example of a shallow slope based on omission of data, a least-squares fit to the luminosity function which omits points from JL95, JLC96 and LJ98, and which does not incorporate upper limit data, yields $q = 4.3$ (see Figure 6.7b). Both size indices would place most of the mass of the observable Kuiper Belt into the smallest objects. Since the size of the smallest object in the distribution is unconstrained, we cannot estimate the mass of the Kuiper Belt using these q's. However, for any $q > 4$, there always exists an s_{min} below which upper limits for the total cometary Belt mass within 50 AU ($\sim$1.3 $M_{\oplus}$) are violated. For values of $q = 4.3$ and $\Sigma(m_R < 27) = 200\,\mathrm{deg}^{-2}$ derived from Figure 6.7b, this minimum value for s_{min} is as large as 2 km, and only increases with increasing q. Explaining the existence of such lower cut-off sizes would be problematic.

Our preferred size index, $q = 3.6 \pm 0.1$, is that of a Dohnanyi-like size distribution for objects having diameters between 50 and 500 km within 50 AU. However, whether the shape of this distribution results from a catastrophic collisional cascade as envisioned in Dohnanyi's (1969) scenario is questionable. The answer depends on impact strengths, relative velocities, and initial populations of KBOs, all of which are poorly constrained. For solid rocky bodies 50–500 km in diameter, critical specific energies for shattering and disruption are expected to increase strongly with size due to self-gravitational compression (Melosh & Ryan 1997, and references therein). The role of self-gravity is magnified yet further if bodies consist predominantly of weaker ices. Whatever their composition, we would not expect Dohnanyi's (1969) derivation to apply to objects as large as those observed, since the derivation assumes that impact strengths are independent of size. Rel-

& Levison (1997) estimate that only 6×10^8 comets are required in the scattered disk to supply the observed rate.

ative velocities required for fragmentation and dispersal of solid rocky bodies $\sim$100 km in size demand KBO eccentricities and inclinations exceeding 0.3; the actual history of the velocity dispersion is unknown. If KBOs consist of solid rock and relative velocities are sufficiently high for disruption and dispersal upon impact, we estimate that lifetimes against catastrophic dispersal of targets $\sim$100 km in diameter exceed the age of the Solar System by a factor of $\sim$150 if projectiles are drawn from the present-day Kuiper Belt. This estimate agrees with that of Stern (1995); see his Figure 2. Shaping the population of objects having sizes 50–500 km by catastrophic collisions would require a primordial Belt orders of magnitude more populous than what is observed today.

6.5.2 A Kuiper Cliff at 50 AU?

To date not one member of the classical Kuiper Belt has been discovered beyond 50 AU, despite observational advances in limiting magnitude and theoretical assurances that the region is dynamically stable. Gladman et al. (1998) have addressed this issue and concluded that the present sample of $\sim$100 KBOs is marginally large enough to expect detection of such bodies. Here we confirm and elaborate upon their results.

We assume the Kuiper Belt begins at an inner edge a_{min}, and that the number density of objects (number per volume) decreases with distance a as a power law with index β:

$$dN(s,a) \propto a^{-\beta}\, s^{-q}\, ds\,. \tag{6.8}$$

For a surface density (number per disk face area) appropriate to the minimum-mass outer solar nebula, the index β may plausibly take values of $\sim$2–3, depending on how quickly random eccentricities e and inclinations i decay with heliocentric distance. In a field of limiting magnitude m, the sky density of objects (number per projected sky solid angle) located beyond distance a_* is proportional to

$$\Sigma\,(a > a_*) \propto \int_{s_m(a_*)}^{s_{max}} \int_{a_*}^{a_m(s)} a^{2-\beta} s^{-q}\, ds\, da\,, \tag{6.9}$$

where $s_m(a_*)$ is the size of the smallest object which can just be seen at a_*, and $a_m(s) =$

$a_* \sqrt{s/s_m(a_*)}$ is the maximum distance out to which an object of size s can be seen. One immediate consequence of a Belt having distance and size indices considered here is that the faint end of the luminosity function is dominated by small nearby objects rather than large distant ones. Extending the limiting magnitude of a visual survey inherently achieves greater dynamic range in observable sizes than in distances because reflected fluxes decrease as s^2/a^4. The greater sensitivity to size is compounded by the shapes of the distributions; $\Sigma \propto s^{1-q}\, a^{3-\beta} \approx s^{-2.6}\, a^{0.5}$. An outer edge to such a Belt at 50 AU (a "Kuiper Cliff") would not significantly flatten the slope of the luminosity function at faint magnitudes, a point which we shall justify more formally below.

The fraction of objects located beyond a_* is

$$f \equiv \frac{\Sigma\,(a > a_*)}{\Sigma\,(a > a_{min})} = \left(\frac{a_{min}}{a_*}\right)^{\gamma} \left\{1 + O\left[\left(\frac{s_m(a_*)}{s_{max}}\right)^{\gamma/2}\right]\right\}, \tag{6.10}$$

where $\gamma = 2q + \beta - 5$ (cf. G98). The order-of-magnitude correction term is valid for $\beta \leq 3$ and is small for surveys and distributions considered here.[12] The fraction f is thus insensitive to the limiting magnitude of the field. This insensitivity justifies our assertion that a Kuiper Cliff would not break the luminosity function at any particular magnitude. It also allows us to easily estimate how many detections beyond 50 AU we might expect. For $a_{min} = 30$ AU, $a_* = 50$ AU, $s_{max} = 2000$km, $q = 3.6$, and $\beta = 3$ (constant dispersion in e and i), the fraction of objects outside 50 AU is $f \approx 8\%$. Decreasing the distance index β to 2 increases f to 13%. In the present total sample of ~100 KBOs, we might therefore expect ~10 to reside beyond 50 AU. Eight of these ten would be located between 50 and 70 AU.

While these rough considerations do not convincingly implicate a Kuiper Cliff, they do argue more strongly against a sudden rise by factors of 3 or more in the surface number density between 50 and 70 AU (a nearby "Kuiper Wall"). Keeping the size distribution fixed and multiplying the surface density by 3 beyond 50 AU would demand that ~25% of all detected classical KBOs reside in such a wall, in contrast to the 0% found to date.

Nonetheless, there are a number of ways the present lack of detections may still accord

[12] For $\gamma \approx 5$ and $s_{max} = 2000$km, the correction term is less than 0.1 for surveys having limiting magnitudes $m_R \gtrsim 22$. All surveys used to construct our luminosity function satisfy the latter requirement.

with a massive outer classical Belt. The size distribution of objects may change dramatically past 50 AU. For instance, it might be that only a few large objects exist between 50 and 70 AU. This may plausibly be the result of runaway accretion unimpeded by the presence of Neptune. Alternatively, if only objects smaller than ~30 km populate the outer Belt, detecting them is a task better suited to occultation surveys than to searches relying on reflected light.

6.6 Summary

Our main results are as follows.

1. We discovered two new Kuiper Belt Objects in a single Keck LRIS field. One object at $m_V = 25.5$ was found by blinking individual frames. It lies at a heliocentric distance of $R \approx 33$ AU and has a diameter $s = 56\,(0.04/p_V)^{1/2}$ km. Another object at $m_V = 27.2$ was discovered by blinking shifted and co-added frames. For this second object, $R \approx 44$ AU and $s = 46\,(0.04/p_V)^{1/2}$ km.

2. We pooled all surveys to construct the cumulative luminosity function from $m_R = 20$ to 27 (Figure 6.8). At the faintest observed magnitude, $\Sigma(m_R \lesssim 26.9) = 53 \pm 20(1\sigma)$ objects/square degree. The best-fit slope is $\alpha = 0.52 \pm 0.02$, where $\log_{10} \Sigma(< m_R) = \alpha(m_R - m_0)$. Differences in α reported in the literature are due mainly to which survey data are incorporated. Values of $\alpha > 0.6$ require the omission of surveys by JL95, JLC96, and LJ98.

3. Our KBO luminosity function is consistent with a power-law size distribution with differential size index $q = 3.6 \pm 0.1$ for objects having diameters 50–500 km within 50 AU. The distribution is such that the smallest objects possess most of the surface area, but the largest bodies contain the bulk of the mass. By extrapolating outside the observed range of sizes, we estimate to order-of-magnitude that $0.2M_\oplus$ and 1×10^{10} comet progenitors lie between 30 and 50 AU. Though our estimated size index is that of a Dohnanyi-like distribution, the interpretation that catastrophic collisions are responsible

is questionable. Impact strengths against catastrophic disruption and dispersal probably increase strongly with size for objects greater than $\sim$10 km in diameter, whereas the derivation by Dohnanyi (1969) assumes impact strength to be independent of size. Lifetimes against catastrophic dispersal of KBOs having diameters 50–500 km exceed the age of the Solar System by at least two orders of magnitude in the present-day Belt, assuming bodies consist of competent rock.

4. A greater than threefold rise in the surface density of the Kuiper Belt just beyond 50 AU would imply that more than 25% of detected objects lie outside that distance, assuming objects are distributed similarly in size at all distances. The absence of detections past 50 AU in the present sample of $\sim$100 KBOs argues against this picture. A massive outer Belt may still be possible if only a few large objects exist between 50 and 70 AU, or if only objects smaller than $\sim$30 km exist in the outer Belt.

Acknowledgments

Data were obtained at the W. M. Keck Observatory, which is operated as a scientific partnership among the California Institute of Technology, the Universities of California, and the National Aeronautics and Space Administration. The observatory was made possible by the generous financial support of the W. M. Keck Foundation. We thank Jane Luu for providing detection efficiencies for the LJ98 data, Peter Goldreich and Sarah Stewart for helpful discussions, and an anonymous referee for a careful reading of this manuscript. E.C. gratefully acknowledges support from an NSF Graduate Research Fellowship.

Bibliography

[1] Allen, C. W. 1973, *Astrophysical Quantities* (Athlone:London)

[2] Brown, M. E., Kulkarni, S. R., & Liggett, T. J. 1997, ApJ, 490, L119

[3] Cochran, A. L., Levison, H. F., Stern, S. A., & Duncan, M. J. 1995, ApJ, 455, 342

[4] Borderies, N., Goldreich, P., & Tremaine, S. 1984, in *Planetary Rings*, ed. R. Greenberg & A. Brahic (University of Arizona Press: Tucson), 713

[5] Cuzzi, J. N., et al. 1984, in *Planetary Rings*, ed. R. Greenberg & A. Brahic (University of Arizona Press: Tucson), 73

[6] Durda, D. D. & Dermott, S. F. 1997, Icarus, 130, 140

[7] Durda, D. D., Greenberg, R., & Jedicke, R. 1998, Icarus, 135, 431

[8] Dohnanyi, J. S. 1969, J. Geophys. Res., 74, 2531

[9] Duncan, M. J. & Levison, H. F. 1997, Science, 276, 1670

[10] Gladman, B., et al. 1998, AJ, 116, 2042 (G98)

[11] Irwin, M., Tremaine, S., & Zytkow, A. N. 1995, AJ, 110, 3082 (I95)

[12] Jewitt, D. & Luu, J. 1995, AJ, 109, 1867 (JL95)

[13] Jewitt, D., Luu, J., & Chen, J. 1996, AJ, 112, 1225 (JLC96)

[14] Jewitt, D., Luu, J., & Trujillo, C. 1998, AJ, 115, 2125 (JLT98)

[15] Hamid, S. E., Marsden, B. G., & Whipple, F. L. 1968, AJ, 73, 727

[16] Landolt, A. U. 1992, AJ, 104, 340

[17] Levison, H. F. & Duncan, M. J. 1990, AJ, 100, 1669 (LD90)

[18] Levison, H. F. & Duncan, M. J. 1997, Icarus, 127, 13

[19] Luu, J. 1999, private communication

[20] Luu, J. X. & Jewitt, D. 1988, AJ, 95, 1256 (LJ88)

[21] Luu, J. X. & Jewitt, D. C. 1998, ApJ, 502, L91 (LJ98)

[22] Malhotra, R. 1996, AJ, 111, 504

[23] Malhotra, R., Duncan, M., & Levison, H. 1999, in *Protostars and Planets IV*, ed. V. Mannings, A. Boss, & S. Russell (University of Arizona Press: Tucson), preprint

[24] Marouf, E. A., Tyler, G. L., Zebker, H. A., & Eshleman, V. R. 1983, Icarus, 54, 189

[25] Melosh, H. J. & Ryan, E. V. 1997, Icarus, 129, 562

[26] Oke, J., et al. 1995, PASP, 107, 275

[27] Stern, S. A. 1995, AJ, 110, 856

[28] Tegler, S. C. & Romanishin, W. 1998, Nature, 392, 49

[29] Tombaugh, C. 1961, in *Planets and Satellites*, ed. G. P. Kuiper & B. M. Middlehurst (University of Chicago Press: Chicago), 12

[30] Weidenschilling, S. J., et al. 1984, in *Planetary Rings*, ed. R. Greenberg & A. Brahic (University of Arizona Press: Tucson), 367

[31] Weissman, P. R. 1995, ARA&A, 33, 327

Part III

DYNAMICS OF PLANETARY RINGS

Chapter 7

Apse Alignment of Narrow Eccentric Planetary Rings

This chapter draws from

Chiang & Goldreich 2000, ApJ, accepted

Abstract

The boundaries of the Uranian ϵ, α, and β rings can be fitted by Keplerian ellipses. The pair of ellipses that outline a given ring share a common line of apsides. Apse alignment is surprising because the quadrupole moment of Uranus induces differential precession. We propose that rigid precession is maintained by a balance of forces due to ring self-gravity, planetary oblateness, and interparticle collisions. Collisional impulses play an especially dramatic role near ring edges. Pressure-induced accelerations are maximal near edges because there (1) velocity dispersions are enhanced by resonant satellite perturbations, and (2) the surface density declines steeply. Remarkably, collisional forces felt by material in the last $\sim$100 m of a $\sim$10 km wide ring can increase equilibrium masses up to a factor of $\sim$100. New ring surface densities are derived which accord with Voyager radio measurements. In contrast to previous models, collisionally modified self-gravity appears to allow for both negative and positive eccentricity gradients; why all narrow planetary rings exhibit positive eccentricity gradients remains an open question.

7.1 Introduction

Each narrow eccentric ring surrounding Uranus is composed of particles moving on nested elliptical orbits. The outer and inner edges of a given ring define ellipses having semi-major axes $a \pm \Delta a/2$ and eccentricities $e \pm \Delta e/2$, where $\Delta a \ll a$, $\Delta e \ll e$, and $e \ll 1$. Observed values of a, e, Δa, and Δe for the Uranian ϵ, α, and β rings are listed in Table 7.1.

Table 7.1: Parameters of Eccentric Uranian Rings

Ring[a]	a(km)	Δa(km)	$e\,(\times 10^3)$	$\Delta e\,(\times 10^3)$
ϵ	51149	58.1	7.936	0.711
α	44718	7.15	0.761	0.076
β	45661	8.15	0.442	0.066

[a] All values taken from Tables I and VII of French et al. (1991).

Remarkably, the set of ellipses describing an individual ring share a common line of apsides. As mentioned in Chapter 1, apse alignment is surprising because the quadrupole moment of the central planet induces differential precession on extremely short timescales; in the case of the Uranian ϵ ring, the inner edge would precess a full revolution relative to the outer edge in 175 years. Rigid precession of an eccentric planetary ring has remained a problem in ring dynamics for over 20 years.

Goldreich & Tremaine (1979, hereafter GT) proposed that apse alignment is maintained by self-gravity. Their theory predicts that the eccentricity gradient across the ring,

$$q_e \equiv a\frac{\partial e}{\partial a}\,, \tag{7.1}$$

must be positive. A positive eccentricity gradient in an apse-aligned ring implies that the ring is narrowest at periapse and widest at apoapse. Gravitational forces between particles are therefore greatest near periapse. Material in the inner half of the ring pulls radially inward on the outer half at periapse, generating a differential precession which exactly cancels that due to planetary oblateness.

Though the prediction that $q_e > 0$ accords with observations of all known narrow eccentric rings, the standard self-gravity model (hereafter SSG) predicts Uranian ring masses that are too low compared to those inferred from Voyager radio occultations. Ring masses based on observations exceed predictions by factors of at least $\sim$3 (ϵ ring) to $\sim$50 (α and β rings) (Tyler et al. 1986; Gresh 1990; see also the reviews by Esposito et al. 1991 and French et al. 1991). Low surface densities are particularly problematic for the α and β rings. With SSG surface densities, torques exerted by inner shepherd satellites would be

insufficiently strong to confine the α and β rings against drag from the distended exosphere of Uranus (Goldreich & Porco 1987, hereafter GP). In addition, as discussed by Graps et al. (1995), shapes of the ϵ ring surface density profiles as derived from occultation light curves do not accord with SSG predictions.

This paper points the way towards resolving these problems. In §7.2, a theory of collisionally modified self-gravity (hereafter CMSG) is qualitatively described. A simple quantitative model is set forth in §7.3, in which new surface density profiles are derived for the ϵ and α rings that are in better agreement with observations. In §7.4, implications of our solutions for torque balance, the role of planetary oblateness, and the value of q_e are discussed. Directions for future research are summarized in §7.5.

7.2 Qualitative Solution

For simplicity, consider an apse-aligned eccentric ring having constant, positive q_e across its width. The ring is filled with spherical particles of internal mass density ρ and radius r, and the ring surface density is given by Σ. Let n and $\tilde{\omega}$ be the mean motion and apsidal angle, respectively, of a ring particle. Subscripts i and b denote quantities evaluated in the ring interior and near the ring boundary, respectively. Variables subscripted with p or s are associated with the central planet or shepherd satellite, respectively, and take their usual meanings. The dimensionless strength of the quadrupole moment of the planet is given by J_2. Numerical estimates in this section are made using parameters appropriate for the ϵ ring.

A key ingredient missing in the SSG model is an accounting for interparticle collisions. Since ring optical depths τ measured normal to the orbital plane are typically of order unity, each particle collides, on average, a few times with its neighbors every orbital period. Only modest collisional impulses per unit mass and time, of order 0.1 cm s^{-1} per orbit, are required to generate differential precession rates comparable to those induced by planetary oblateness (GT). Velocity dispersions of order $c_i \sim 0.1 \text{ cm s}^{-1}$ in the ring interior are not unreasonable; both the Keplerian shearing velocity across a particle diameter, $\sim 3nr$, and the escape velocity from the particle surface, $\sim r\sqrt{8G\rho}$, are of that order for the meter-

sized bodies that plausibly compose the ring.

Although a single collision can impart an impulse of dynamically significant magnitude, multiple collisions experienced by a particle in the ring interior leave its precession rate largely unaltered. A particle in the ring interior is struck by its inside neighbors about as frequently and as forcefully as by its outside ones. Differential precession across the ring induced by smooth internal pressure gradients occurs on timescales of order $2\pi\Sigma nae/|\nabla P| \sim 2\pi nae\Delta a/c_i^2 \sim 10^6\,(0.1\text{ cm s}^{-1}/c_i)^2\,\text{yr}$, much longer than misalignment timescales set by planetary oblateness (cf. GT). Here the height-integrated pressure $P \sim \Sigma c_i^2$ is taken to vary over a lengthscale Δa.

Conditions are dramatically different near ring edges. Pressure-induced accelerations are maximal there because (1) velocity dispersions are enhanced by resonant satellite perturbations, and (2) the surface density declines steeply (Borderies, Goldreich, & Tremaine 1982). The velocity dispersion near the ring boundary could be as high as

$$c_b \sim \sqrt{\frac{d}{w_r}}\, c_i \sim 3\,\frac{c_i}{0.1\text{ cm s}^{-1}}\text{ cm s}^{-1}\,, \tag{7.2}$$

where $d \sim 10^3$km is the ring-satellite separation, and w_r is the width of the annulus perturbed by the satellite. To order-of-magnitude, the latter is given by $w_r \sim a\sqrt{M_s/M_p} \sim$ 1km, the distance from the resonant edge at which nested periodic orbits cross. Equation (7.2) is derived by equating the rate of energy dissipation by collisions in the perturbed zone, $\sim \pi\Sigma a w_r c_b^2 n\tau$, to the rate of energy deposition by the satellite, $\sim 3nTd/2a \sim 9\pi\Sigma nac_i^2 d/2$, where T is the satellite-induced confining torque whose magnitude equals that of the viscous torque, $\sim 3\pi\Sigma c_i^2 a^2$, in steady-state.

A particle on the ring edge experiences a radially directed, collisional acceleration

$$C \sim -\frac{\nabla P}{\Sigma} \sim \pm\,\frac{c_b^2}{\lambda}\,\hat{r} \sim \pm\, c_b\, n\,\hat{r}\,, \tag{7.3}$$

where the upper (lower) sign applies to the outer (inner) ring edge. Here $P \sim \Sigma\, c_b^2$ is taken to vary over a radial lengthscale, λ, of order the local ring thickness, c_b/n. In a $q_e > 0$ ring, collision rates are highest near periapse. At the periapsis of a ring boundary, the radial acceleration, C, generates a differential precession rate, $\Delta\langle d\tilde{\omega}/dt\rangle_C \approx -C/nae$, relative

to the precession rate at the ring midline. This collision-induced rate is greater than the local differential rate due to planetary oblateness, $\Delta\langle d\tilde{\omega}/dt\rangle_O$, by a substantial factor:

$$\frac{\Delta\langle d\tilde{\omega}/dt\rangle_C}{\Delta\langle d\tilde{\omega}/dt\rangle_O} \sim \frac{\mp c_b/ae}{\mp(21\Delta a/8a)J_2 n(R_p/a)^2} \sim 40\,\frac{c_b}{1\ \mathrm{cm\ s^{-1}}}\,. \tag{7.4}$$

Self-gravity maintains apse alignment against differential precession caused by planetary oblateness and interparticle collisions. For self-gravity to enforce rigid precession near ring edges, surface densities there must be higher than those predicted by SSG. At ring boundaries, self-gravitational attraction must balance the extra repulsive acceleration due to collisions. To estimate the surface density near the edge, Σ_b, equate the collisional acceleration, $C \sim c_b^2/\lambda$, to the gravitational acceleration from a wire of linear mass density $\Sigma_b\lambda$ located a distance λ away:

$$c_b^2/\lambda \sim 2G\Sigma_b\,. \tag{7.5}$$

Take $\lambda = c_b/n \sim 50\,\mathrm{m}$ to obtain

$$\Sigma_b \sim c_b\, n\,/\,2\,G \sim 10^3 \frac{c_b}{1\ \mathrm{cm\ s^{-1}}}\ \mathrm{g\ cm^{-2}}\,, \tag{7.6}$$

which is greater than corresponding SSG predictions by factors $\gtrsim 40$. Equation (7.6) is equivalent to the condition that Toomre's Q be of order unity at the edge.

These endwires of mass $\sim 2\pi a\Sigma_b\lambda = \pi c_b^2 a/G$ constitute new boundary conditions not found in SSG. Gravitational forces from massive endwires induce substantial differential precession in the ring interior. For self-gravity to maintain apse alignment in the interior, surface densities there must also be greater than those predicted by SSG.

7.3 Quantitative Model

Divide the region occupied by an apse-aligned, constant $q_e = a\Delta e/\Delta a$ ring into an even number N of equally spaced intervals. The center of the j^{th} interval contains an elliptical wire having mass m_j, semi-major axis $a_j = a + [\,j - (N+1)/2\,]\,\Delta a/N$, and eccentricity

$e_j = e + [\, j - (N+1)/2 \,]\, \Delta e/N$. Denote by $\Delta_j \langle d\tilde{\omega}/dt \rangle$ the precession rate of the $j^{\rm th}$ wire relative to the precession rate of a test particle at the ring midline. Uniform precession requires

$$\Delta_j \langle \frac{d\tilde{\omega}}{dt} \rangle = \Delta_j (\langle \frac{d\tilde{\omega}}{dt} \rangle_O + \langle \frac{d\tilde{\omega}}{dt} \rangle_G + \langle \frac{d\tilde{\omega}}{dt} \rangle_C) = 0\,. \tag{7.7}$$

Subscripts O, G, and C denote contributions from planetary oblateness, self-gravity, and interparticle collisions, respectively. The first two terms are given by

$$\Delta_j \langle \frac{d\tilde{\omega}}{dt} \rangle_O = -\frac{21}{4} J_2 n \left(\frac{R_p}{a} \right)^2 \frac{a_j - a}{a} \tag{7.8}$$

and

$$\Delta_j \langle \frac{d\tilde{\omega}}{dt} \rangle_G = \frac{q_e H(q_e^2)}{\pi e} n \frac{a}{M_p} \sum_{k \neq j} \frac{m_k}{a_j - a_k}\,, \tag{7.9}$$

where

$$H(q_e^2) \equiv \frac{1 - \sqrt{1 - q_e^2}}{q_e^2 \sqrt{1 - q_e^2}}$$

(cf. GT).

For $\Delta_j \langle d\tilde{\omega}/dt \rangle_C$, the following simplistic prescription is adopted:

$$\Delta_j \langle \frac{d\tilde{\omega}}{dt} \rangle_C = \begin{cases} +[\, q_e H(q_e^2)\, c_b^2/\lambda n a e \,]\, (1 - x/\lambda) & \text{if } x \equiv (j - 1/2)\, \Delta a/N < \lambda \\ -[\, q_e H(q_e^2)\, c_b^2/\lambda n a e \,]\, (1 - y/\lambda) & \text{if } y \equiv (N - j + 1/2)\, \Delta a/N < \lambda \\ 0 & \text{otherwise.} \end{cases} \tag{7.10}$$

Thus, $\Delta_j \langle d\tilde{\omega}/dt \rangle_C$ is non-zero only within intervals $\delta a = \lambda$ from each edge; there, its magnitude rises linearly from 0 to $q_e H(q_e^2) c_b^2/\lambda n a e$. As a crude justification for this maximum value, approximate the collisional acceleration as

$$
\begin{aligned}
C &\approx -\frac{\nabla P}{\Sigma} \\
&\approx \pm \frac{c_b^2}{\lambda(1 - q_e \cos f)}\,\hat{r}
\end{aligned}
\tag{7.11}
$$

where the upper (lower) sign applies to the outer (inner) edge. Here the pressure gradient is taken to vary inversely as the separation between streamlines. Insert C into Gauss's perturbation equation for $d\tilde{\omega}/dt$ and average over true anomaly:

$$
\begin{aligned}
\langle \frac{d\tilde{\omega}}{dt} \rangle &= -\frac{1}{\pi nae} \int_0^\pi C \cos f \, df \\
&= \mp q_e H(q_e^2)\, c_b^2/\lambda nae \,.
\end{aligned}
\tag{7.12}
$$

For the constant q_e ring models presented here,

$$
\lambda = c_b/n \,, \tag{7.13}
$$

so that the only remaining free parameter is c_b. Equation (7.13) is relaxed for §7.4.2.2 and §7.4.3.2.

Note that this prescription for $\Delta_j \langle d\tilde{\omega}/dt \rangle_C$ ignores the decrease in velocity dispersion from c_b at the ring edge to c_i in the ring interior. The decline in velocity dispersion occurs over a length scale of order w_r. This length scale is large compared to c_b/n so that the gradient of velocity dispersion does not give rise to a significant radial acceleration.

For a given value of c_b, equations (7.7), (7.8), (7.9), (7.10), and (7.13) comprise N equations in N unknowns $\{m_j\}$.[1] Solutions for surface density profiles at quadrature for various values of c_b are displayed in Figure 7.1, for parameters appropriate to the ϵ and α rings; models for the β ring are nearly identical to those of the α ring. In CMSG models, higher surface densities near ring edges are evident, as are higher total ring masses.

[1] Reflection symmetry about the ring midline reduces the number of equations necessary to $N/2$. Typically $N \gtrsim 2000$ wires are needed to converge to within 10% of the solution for $N \to \infty$.

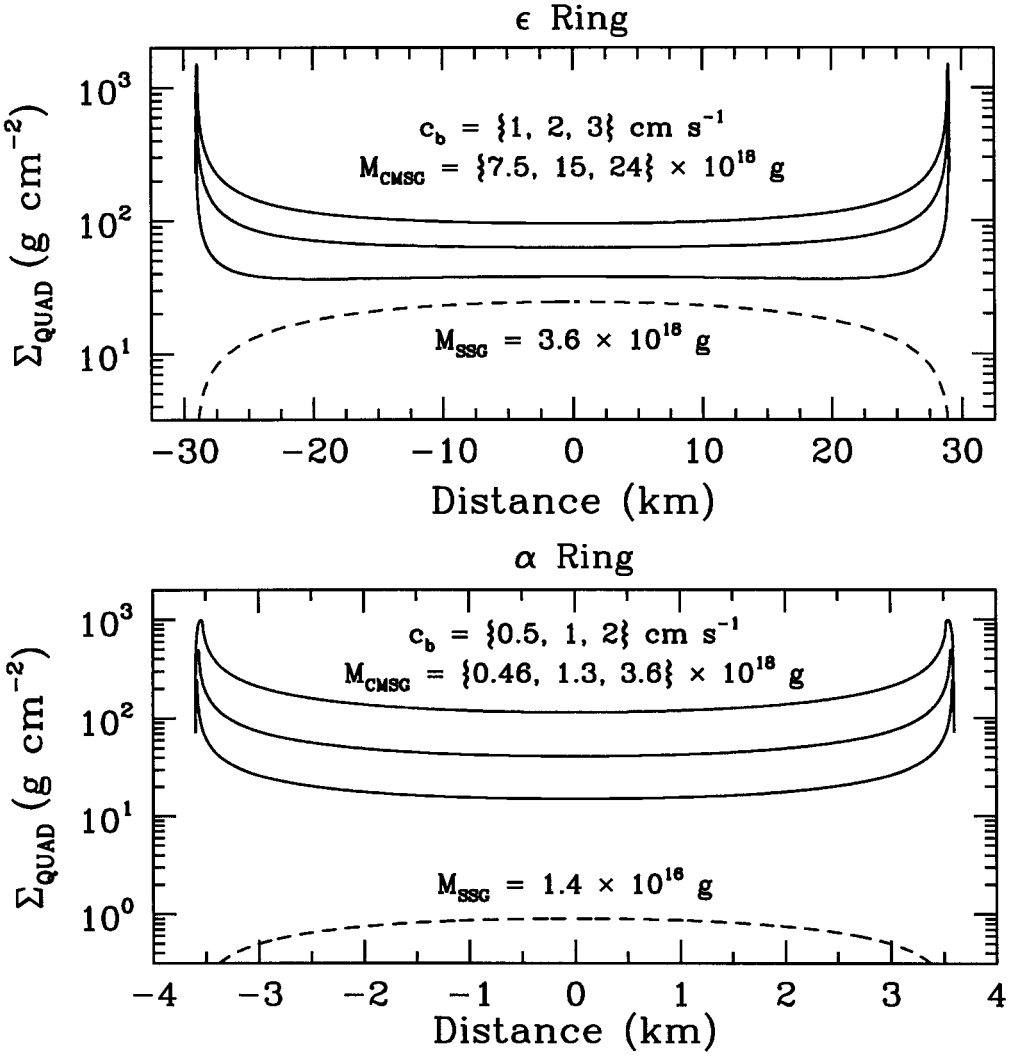

Figure 7.1: *Surface density profiles at quadrature for rings ϵ and α.*

Solid lines represent CMSG ($c_b \neq 0$) models, while dashed lines represent SSG ($c_b = 0$) models. When the total CMSG ring mass, $M_{\rm CMSG}$, greatly exceeds the total SSG ring mass, $M_{\rm SSG}$, it is found empirically that $M_{\rm CMSG} \propto c_b^2/\lambda^{1/2} \propto c_b^{3/2}$.

7.4 Discussion

7.4.1 Surface Density Profiles and Torque Balance

Simple CMSG models, while not fully realistic, demonstrate the existence of a new class of self-gravity solution, that obtained by accounting for the modification of ring boundary conditions by interparticle collisions. Remarkably, forces felt by material in the last $\sim$100 m of a $\sim$10 km wide ring can increase equilibrium masses by factors up to 100.

Large S-band opacities measured by Voyager, which are incompatible with SSG surface densities (see, e.g., the review by Esposito et al. 1991), can be reconciled with average CMSG surface densities of $\sim$75–100 g cm^{-2} for the ϵ, α, and β rings. Moreover, CMSG models predict that surface densities near ring edges are higher than those in the interior. This behavior is reminiscent of the "double-dip" structure seen in occultation light curves for the ϵ and α rings (see, e.g., the review by French et al. 1991).

Greater ring masses as implied by CMSG resolve problems associated with exospheric drag that were pointed out by GP for rings α and β. For the remainder of this subsection, numerical estimates will be made for the α ring; similar conclusions hold for the β and ϵ rings. Surface densities are scaled to a typical CMSG value in the ring interior of $\Sigma = 75\,\mathrm{g\ cm^{-2}}$. An inner shepherd satellite exerts a repulsive, non-linear torque at first-order Lindblad resonances of magnitude

$$T_{\rm nl}^{L} \approx \frac{10\rho_s R_s^3 \Sigma^2 n^2 a^7}{M_p^2 d} \approx 6{\times}10^{17} \left(\frac{\Sigma}{75\,\mathrm{g\ cm^{-2}}}\right)^2 \left(\frac{R_s}{10\mathrm{km}}\right)^3 \left(\frac{\rho_s}{1.5\,\mathrm{g\ cm^{-3}}}\right) \left(\frac{500\mathrm{km}}{d}\right) \mathrm{erg}\,, \tag{7.14}$$

where the satellite radius, R_s, is scaled to the Voyager upper limit of 10 km (Smith et al. 1986). The shepherding torque exceeds the magnitude of the drag torque exerted by the Uranian exosphere,

$$T_d \approx -4\pi m_H n_H v_T n a^3 \Delta a \approx -4 \times 10^{16} \left(\frac{n_H}{10^3\ \mathrm{cm^{-3}}}\right) \mathrm{erg}\,. \tag{7.15}$$

Here $n_H = 7 \times 10^{-6}\, e^{32.4\, R_p/a}\ \mathrm{cm^{-3}}$ is the number density of hydrogen atoms of mass

m_H in the exosphere, and $v_T \approx 1\text{km s}^{-1}$ is their thermal speed normal to the orbital plane (Broadfoot et al. 1986). That $T_{\rm nl}^L > |T_d|$ ensures that the inner shepherd prevents ring particles from spiraling in towards Uranus.

Estimates of viscous torques T_v also require revision. For a ring undergoing Keplerian shear, with minimum kinematic viscosity $n(\Sigma/\rho)^2$, the viscous torque is given by

$$T_v \approx \frac{3\pi n^2 \Sigma^3 a^2}{\rho^2} \approx 2.5 \times 10^{18} \left(\frac{\Sigma}{75\text{ g cm}^{-2}}\right)^3 \left(\frac{1.5\text{ g cm}^{-3}}{\rho}\right)^2 \text{ erg} \qquad (7.16)$$

(GP). That $T_v \gg |T_d|$ ensures that ring particles on the outer edge press against the inner Lindblad resonance established by the outer shepherd.

Conclusions drawn from comparisons between $T_{\rm nl}^L$ and T_v are on less sure footing. For the choice of scaling parameters, the latter exceeds the former, contrary to the requirement of the standard theory of shepherding that the torques be equal. This might be construed as evidence that the angular momentum luminosity in the ring interior is reduced below T_v by the non-Keplerian shear associated with a non-zero q_e (Borderies, Goldreich, & Tremaine 1982; GP). However, the numerical estimates for the two torques differ only by a factor of a few. The shepherding torque should be evaluated using surface densities near the edge, which CMSG predicts are higher than those in the interior; this would increase the estimate of $T_{\rm nl}^L$. Uncertainties in the choice of parameters preclude drawing any conclusion other than that these torques are of the same order of magnitude.

7.4.2 Relative Importance of Planetary Oblateness

7.4.2.1 $J_2 = 0$ vs. $J_2 \neq 0$

What does CMSG predict if $J_2 = 0$? Figure 7.2a displays the answer for the ϵ ring, for $c_b = 2$ and 3 cm s^{-1}. In contrast to SSG, a non-vanishing equilibrium surface density does not require a finite planetary oblateness; self-gravity can be balanced entirely by collisional pressure gradients. For the α and β rings, solutions with and without J_2 are practically indistinguishable for $c_b \geq 0.5$ cm s^{-1}. The influence of J_2 on the equilibrium solution diminishes as Δa decreases or as c_b increases.

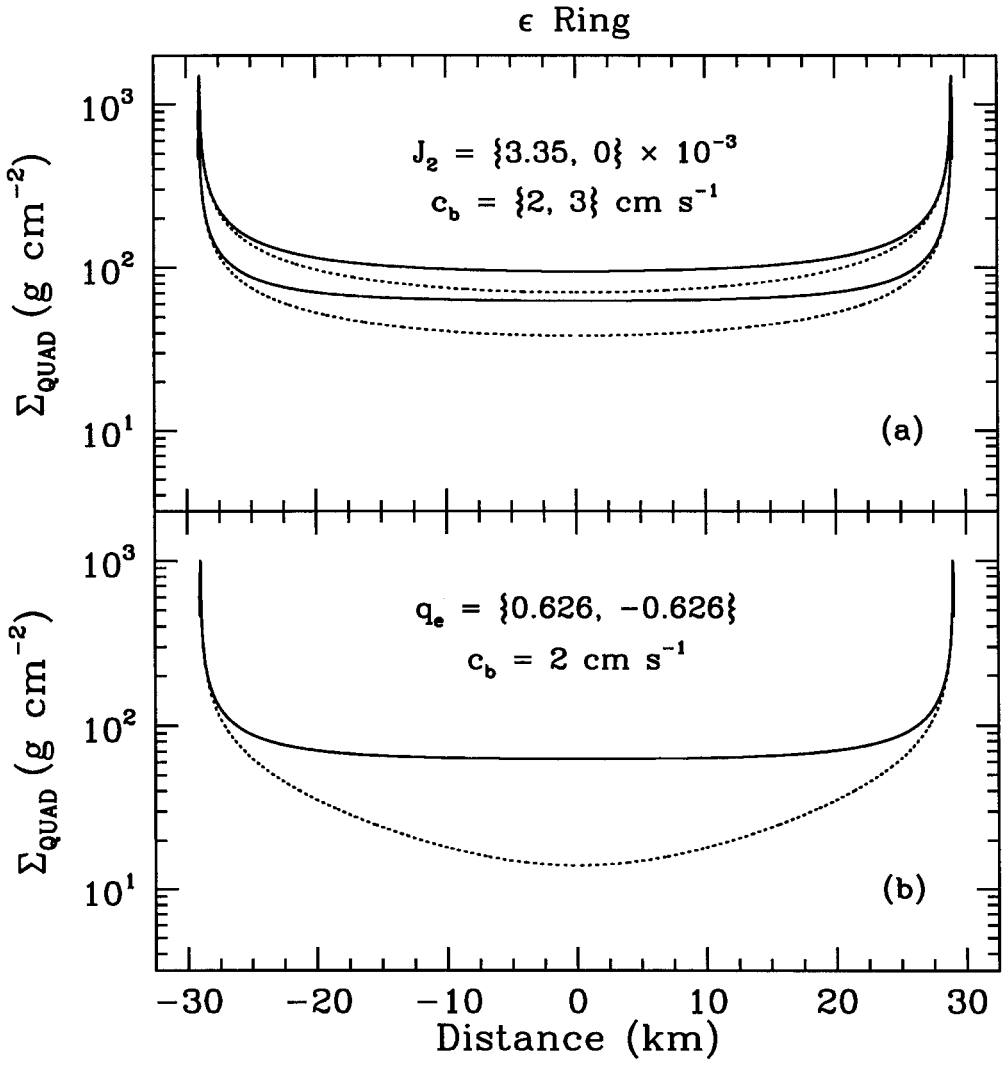

Figure 7.2: *Variations of CMSG ring models with J_2 and the sign of the eccentricity gradient.*

(a) CMSG ϵ ring models for which J_2 is reduced from its nominal value of 3.35×10^{-3} *(Elliot & Nicholson 1984) (solid line) to 0 (dotted line). As c_b is increased from* 2 cm s^{-1} *(lower two curves) to* 3 cm s^{-1} *(upper two curves), the influence of J_2 diminishes. For the α and β rings, CMSG models with and without J_2 are practically indistinguishable for $c_b \geq 0.5$* cm s^{-1} *(data not shown). (b) CMSG ϵ ring models for which $q_e = \pm 0.626$. Contrary to SSG models, a positive q_e is not required to obtain an equilibrium solution.*

7.4.2.2 Empirical Scaling Relations for $J_2 = 0$

For $J_2 = 0$ and fixed ring geometry, the surface density at quadrature near a given edge scales as

$$\Sigma_b \left(0 \leq |x| \lesssim \lambda\right) = \frac{c_b^2}{G\lambda} f(|x|/\lambda)\,, \tag{7.17}$$

where $|x|$ measures distance from the edge, c_b and λ are the same free parameters as in Equation (7.10), and f is a dimensionless function of the similarity variable $|x|/\lambda$. Well away from ring edges, the surface density at quadrature scales as

$$\Sigma_i \left(|x| \gg \lambda\right) = \frac{c_b^2}{G\sqrt{\lambda \Delta a}}\, g(|x|/\Delta a)\,, \tag{7.18}$$

where g is another dimensionless function. The total ring mass scales as

$$M \sim \frac{c_b^2 a}{G} \sqrt{\Delta a/\lambda}\,. \tag{7.19}$$

7.4.3 Value of q_e

7.4.3.1 Sign of q_e

Figure 7.2b displays a CMSG model for the ϵ ring obtained by reversing the sign of q_e. In contrast to SSG, a positive eccentricity gradient is not necessary in CMSG to obtain an equilibrium solution. This resurrects the problem of why all known eccentric planetary rings, including the Titan and Huygens ringlets around Saturn, are narrowest at periapse and widest at apoapse.

It is possible that equilibria obtained using $q_e < 0$ are unstable. To address this issue, a preliminary investigation of ring stability for an $N = 4$ ringlet model has been undertaken. Forces due to pressure gradients are included only for the first and fourth ringlets. Collisional accelerations are treated as if they arise from anti-self-gravity forces (self-gravity with the sign of the acceleration reversed); i.e., collisional shear stresses are ignored. In this crude approximation, equilibria are found to be stable regardless of the sign of q_e; small deviations from equilibrium masses result in apsidal librations (Borderies, Goldreich

& Tremaine 1983). It remains to be seen whether collisional shear stresses alter stability properties.

Another possibility is that initial conditions set the sign of q_e. If the ring were initially uniform in width as a function of azimuth, then planetary oblateness would determine the initial sense of differential precession within the ring. The resultant narrowing of the ring width near a true anomaly of $f = -\pi/2$ would cause a positive eccentricity gradient to grow by self-gravity. Under this hypothesis, an $N = 2$ ringlet model incorporating forces from self-gravity and planetary oblateness yields the following time evolution for the apse and eccentricity differences between outer and inner ringlets:

$$\delta\tilde{\omega} = -A \sin \Omega_{\rm lib} t \tag{7.20}$$

$$\delta e = Ae(1 - \cos \Omega_{\rm lib} t) \tag{7.21}$$

where $A > 0$ and $\Omega_{\rm lib}$ are the amplitude and frequency, respectively, of libration (cf. Borderies, Goldreich, & Tremaine 1983). Note that the time-average of δe is positive. Inelastic collisions would damp librations and the ring would eventually settle into an equilibrium for which $q_e > 0$.

7.4.3.2 Magnitude of q_e Near Ring Boundaries

It has been assumed that the eccentricity gradient, q_e, is finite out to the last $\lambda = c_b/n \sim 50$ meters of ring material. A finite q_e is necessary to generate a non-zero azimuthal average of the collisional acceleration [see equation (7.12)]. The simple quantitative model of §7.3 employed the observed value of q_e averaged over the entire ring width. The true value over the last few hundred meters of ring material is unknown.

In the case of the best-studied ϵ ring, Graps et al. (1995) combined Voyager photopolarimeter and radio occultation measurements to infer the eccentricity gradient as a function of semi-major axis. They found that q_e decreases over the last ~5 km from its nearly constant value of ~0.65 in the interior to ~0.35 near the edge. The radial resolution of their study was between 1 and 2 km.

A decrease in q_e towards ring boundaries is theoretically plausible. Distortions in a circular ring can be described by the change in separation, δr, between neighboring streamlines of the form

$$\delta r \propto \cos m(\phi - \Omega_{\rm pat} t)\,, \tag{7.22}$$

where $\Omega_{\rm pat}$ is the pattern speed of the distortion and m is an integer. A constant q_e ring that precesses rigidly in the quadrupole field of the central planet is equivalent to a distorted circular ring for which $m = 1$ and $\Omega_{\rm pat} = \langle d\tilde{\omega}/dt \rangle_Q$. Resonant satellite perturbations, which enhance velocity dispersions within a distance $w_r \sim 1$km of ring edges, are characterized by much higher values of $m = 2a/3d \gg 1$ and $\Omega_{\rm pat} = \Omega_s$. Satellite-induced disturbances might therefore reduce the local value of q_e. A decrease in q_e over a distance w_r near ring boundaries is roughly equivalent to setting $\lambda = w_r$ in equation (7.10). By the scaling relations (7.18) and (7.19), this would reduce surface densities and total ring masses shown in Figure 7.1 by a factor of $\sqrt{w_r n/c_b} \sim 4$.

7.5 Directions for Future Research

This work is primarily a demonstration that interparticle collisions near ring boundaries play a crucial role in determining ring masses under the self-gravity hypothesis. The nature of ring boundary conditions has not been calculated in rigorous detail; instead a prescription motivated by order-of-magnitude arguments is provided for collision-induced precession rates. Numerical simulations incorporating shepherd satellites will help to determine the actual 3-dimensional collisional stress tensor and eccentricity gradient everywhere within the ring.

Why all narrow eccentric rings surrounding Uranus and Saturn are narrowest at periapse and widest at apoapse remains to be understood. Stability analyses incorporating collisional shear stresses may reveal that rings having $q_e < 0$ are unstable. Alternatively, the sign of q_e may be set by initial conditions. Scenarios for ring formation—e.g., the catastrophic disruption of a small moon—require further elucidation.

Viscous damping gives rise to small differences between apsidal angles of neighboring streamlines (Borderies, Goldreich, & Tremaine 1983). For a given apsidal shift of $\delta\tilde{\omega} \ll 1$, the difference between the azimuth of maximum streamline separation and the azimuth of apoapse is given by the "pinch angle," $\delta\phi = \arctan(e\delta\tilde{\omega}/\delta e) \gg \delta\tilde{\omega}$. The pinch angles calculated by Borderies et al. (1983) for their $N = 2$ streamline models of the Uranian and Saturnian ringlets are suspect, however, because they neglect the boundary effects highlighted in the present work. A careful calculation of $\delta\phi(a)$ that incorporates viscous drag and the global effects of resonant forcing by shepherd satellites has yet to be performed. Upcoming observations of narrow Saturnian ringlets by the Cassini Orbiter might test the predictions of such a calculation, thereby furnishing a powerful diagnostic of stresses within ringlets.

Acknowledgments

Financial support for this research was provided by NSF grant 94-14232 and by a Caltech Kingsley Foundation Fellowship held by E.C.

Bibliography

[1] Borderies, N., Goldreich, P., & Tremaine, S. 1982, Nature, 299, 209

[2] Borderies, N., Goldreich, P., & Tremaine, S. 1983, AJ, 88, 1560

[3] Broadfoot, A.L., et al. 1986, Science, 233, 74

[4] Elliot, J.L. & Nicholson, P.D. 1984, in *Planetary Rings*, ed. R. Greenberg & A. Brahic (University of Arizona Press: Tucson), 25

[5] Esposito, L.W., Brahic, A., Burns, J.A., & Marouf, E.A. 1991, in *Uranus*, ed. J.T. Bergstrahl, E.D. Miner, & M.S. Matthews (University of Arizona Press), 410

[6] French, R.G., Nicholson, P.D., Porco, C.C., & Marouf, E.A. 1991, in *Uranus*, ed. J.T. Bergstrahl, E.D. Miner, & M.S. Matthews (University of Arizona Press: Tucson), 327

[7] Goldreich, P. & Porco, C.C. 1987, AJ, 93, 730 (GP)

[8] Goldreich, P. & Tremaine, S. 1979, AJ, 84, 1638 (GT)

[9] Graps, A.L., Showalter, M.R., Lissauer, J.J., & Kary, D.M. 1995, AJ, 109, 226

[10] Gresh, D. L. 1990, *Voyager Radio Occultation by the Uranian Rings: Structure, Dynamics, and Particle Sizes*. Ph.D. Thesis, Stanford University

[11] Smith, B.A., et al. 1986, Science, 233, 43

[12] Tyler, G.L., et al. 1986, Science, 233, 79

Chapter 8

Angular Momentum Transport in Particle Disks

Abstract

We examine the angular momentum transport properties of disks composed of macroscopic particles whose velocity dispersions are externally enhanced ("stirred"). We prove a general theorem which states that steady inward transport of angular momentum requires that the net effect of collisions and external heating be to further lengthen the principal equatorial long axis of the velocity ellipsoid and to further shorten the principal equatorial short axis. This behavior is not exhibited by particle disks that are vertically stirred; the tendency for interparticle collisions to isotropize velocity dispersions causes transport to be outwards for these disks. A mechanical analogue of Stone & Balbus's (1996) vertically convective fluid disks remains to be found.

8.1 Introduction

Here we briefly examine the transport properties of axisymmetric particle disks whose velocity dispersions may be externally enhanced ("stirred"). We prove a general theorem relating the direction of steady angular momentum transport to the behavior of the velocity dispersion ellipsoid in the presence of external driving. We then ask whether a particle disk that is externally stirred in the vertical direction can furnish the mechanical analogue of Stone & Balbus's (1996) vertically convective fluid disks.

The material in this short chapter comprises a subset of the work by Quataert & Chiang (2000), in which the fluid/particle connection is established in more rigorous detail and more general conditions under which inward transport occurs are illuminated. The theorem which we present below helps to interpret the results found there. I am indebted

to E. Quataert for showing me the utility of Krook's approximation in modelling particle collisions.

8.2 The "Long Axis Must Get Longer" Theorem

THEOREM[1]*:* Steady inward transport of angular momentum requires that the net effect of collisions and external heating be to further lengthen the principal equatorial long axis of the velocity ellipsoid and to further shorten the principal equatorial short axis. This result applies to disks for which the angular velocity gradient, $d\Omega/dr$, is negative.

PROOF: Begin by writing the second moments of the Boltzmann equation which describe the evolution of the viscous stress tensor. For an axisymmetric disk, in cylindrical coordinates, these read (see, e.g., Borderies, Goldreich, & Tremaine 1983):

$$\partial_t p_{rr} - 4\Omega p_{r\phi} = (\partial_t p_{rr})_c + \dot{H}_{rr} \tag{8.1}$$

$$\partial_t p_{r\phi} + (2-S)\Omega p_{rr} - 2\Omega p_{\phi\phi} = (\partial_t p_{r\phi})_c \tag{8.2}$$

$$\partial_t p_{\phi\phi} + 2(2-S)\Omega p_{r\phi} = (\partial_t p_{\phi\phi})_c + \dot{H}_{\phi\phi} \tag{8.3}$$

$$\partial_t p_{zz} = (\partial_t p_{zz})_c + \dot{H}_{zz}\,, \tag{8.4}$$

where we have neglected third moments. Here Ω is the angular frequency, $S \equiv -d\ln\Omega/d\ln r$ parameterizes the shearing profile (e.g., $S = 3/2$ for Keplerian disks, and $S = 2$ for disks of constant specific angular momentum), and $p_{\alpha\beta}$ are the components of the stress tensor. The terms $(\partial_t p_{\alpha\beta})_c$ describe changes in velocity dispersions due to inter-particle collisions. Collisions perform three tasks: they (1) extract energy from the mean shear into random motions, (2) transfer energy in random motions associated with one direction to those associated with another, and (3) convert kinetic energy into heat. The terms $\dot{H}_{\alpha\beta}$ are inserted to simulate external vertical heating, with $\dot{H}_{r\phi} = 0$ so as not to explicitly induce angular momentum transport.

We transform to the principal axis basis of the velocity ellipsoid. Let p_{11} and p_{22} be the

[1]Our name for this theorem was coined by Scott Tremaine, who also derived it independently.

greater and lesser of the two components of the stress tensor, respectively, in the equatorial plane:

$$\begin{aligned} p_{11} &= [\,p_{rr} + p_{\phi\phi} + \sqrt{(p_{rr} - p_{\phi\phi})^2 + 4p_{r\phi}^2}\,] \,/\, 2 \\ p_{22} &= [\,p_{rr} + p_{\phi\phi} - \sqrt{(p_{rr} - p_{\phi\phi})^2 + 4p_{r\phi}^2}\,] \,/\, 2\,, \end{aligned}$$

where we shall always take the positive square root. By symmetry about the midplane, $p_{33} = p_{zz}$. The orientation of the velocity ellipsoid is given by δ, the angle between the longer equatorial axis and the radial direction:

$$\begin{aligned} \tan\delta &\equiv \frac{p_{r\phi}}{p_{11} - p_{\phi\phi}} \\ &= \frac{2p_{r\phi}}{(p_{rr} - p_{\phi\phi}) + \sqrt{(p_{rr} - p_{\phi\phi})^2 + 4p_{r\phi}^2}}\,. \end{aligned} \tag{8.5}$$

Without loss of generality, $|\delta| \leq \pi/2$. Since the denominator in equation (8.5) is positive,

$$\operatorname{sgn} p_{r\phi} = \operatorname{sgn}\delta\,. \tag{8.6}$$

In other words, the direction of angular momentum transport is given by the sign of δ, i.e., by the orientation of the velocity ellipsoid. In the absence of collisions and external heating, $\delta = 0$, $p_{r\phi} = 0$, $p_{11} = p_{rr}$, and $p_{22} = p_{\phi\phi}$.

The transformed viscous stress equations read:

$$\partial_t p_{11} - S\Omega\,(\sin 2\delta)\,p_{11} = (\partial_t p_{11})_c \tag{8.7}$$

$$\partial_t p_{22} + S\Omega\,(\sin 2\delta)\,p_{22} = (\partial_t p_{22})_c \tag{8.8}$$

$$\begin{aligned} (p_{11} - p_{22})\,\partial_t\delta + (2 - S\cos^2\delta)\,\Omega\,p_{11} - (2 - S\sin^2\delta)\,\Omega\,p_{22} = \\ (p_{11} - p_{22})\,(\partial_t\delta)_c - \tan 2\delta\,(\dot{H}_{11} - \dot{H}_{22})\,/\,2 \end{aligned} \tag{8.9}$$

$$\partial_t p_{33} = (\partial_t p_{33})_c + \dot{H}_{33} \,. \tag{8.10}$$

Note that equations (8.7)–(8.10) are slightly more general that those presented by Goldreich & Tremaine (1978).

Equations (8.7) and (8.8) state that for $S > 0$, the steady ($\partial_t = 0$) inward transport ($\delta < 0$) of angular momentum demands that the net effect of collisions and external heating be to further lengthen the principal equatorial long axis of the velocity ellipsoid and to further shorten the principal short axis.

COROLLARY: Particle disks stirred purely in the vertical direction can only transport angular momentum outwards in steady-state.

To see this, set $\dot{H}_{11} = \dot{H}_{22} = 0$ and recognize that the tendency for interparticle collisions to isotropize velocity dispersions in the equatorial plane implies that $(\partial_t \delta)_c = 0$ (Goldreich & Tremaine 1978; Borderies, Goldreich, & Tremaine 1983; see also our appendix). Then Equation (8.9) indicates that in steady-state, the axis ratio p_{11}/p_{22} can only decrease from its collisionless value of $2/(2 - S)$ as δ shifts away from its collisionless value of zero. Thus, $(\partial_t p_{11})_c < 0 < (\partial_t p_{22})_c$ and steady transport must be outwards.

8.3 Discussion

Our corollary states that purely *vertical* stirring in conventional particle disks generates only outward transport in steady-state. This result is a consequence of the tendency for interparticle collisions to "round out" the velocity ellipsoid, i.e., to isotropize velocity dispersions in the equatorial plane. While interesting in itself, our finding means that a mechanical analogue of Stone & Balbus's (1996) vertically convective fluid disks, in which transport was found to be inward, has yet to be discovered.

From the "long axis must get longer" theorem, one might expect *radial* stirring to give rise to inward transport. This is because the velocity ellipsoid ordinarily has a greater projection along the radial direction than the azimuthal direction; in the collisionless, unheated limit, equation (8.9) states that $p_{rr}/p_{\phi\phi} = 2/(2 - S) = 4$ for steady-state Keplerian disks. Thus, radial stirring corresponds to directly heating the long axis of the velocity

ellipsoid and might therefore lead to inward transport. Quataert (Quataert & Chiang 2000) has shown that this expectation is indeed borne out, provided interparticle collisions are sufficiently inelastic.

Acknowledgments

This exercise was originally conceived by Peter Goldreich. Subsequent developments benefited from discussions with him, Eliot Quataert, and Scott Tremaine. E. C. acknowledges financial support from an NSF Graduate Fellowship and a Caltech Kingsley Foundation Fellowship.

Appendix A Interparticle collisions do not directly affect δ

One way to argue that $(\partial_t \delta)_c = 0$ is to employ Krook's approximation for the collisional terms:

$$(\partial_t p_{\alpha\beta})_c = \frac{-p_{\alpha\beta} + \epsilon\sigma^2\delta_{\alpha\beta}}{T}, \tag{8.11}$$

where T is the mean free time between collisions, ϵ is a collisional "elasticity" parameter which can take values between 0 and 1, $\delta_{\alpha\beta}$ is the Kronecker delta function, and

$$\sigma^2 = \frac{p_{rr} + p_{\phi\phi} + p_{zz}}{3} \tag{8.12}$$

is the square of the post-collision velocity dispersion. The case $\epsilon \ll 1$ corresponds to highly inelastic collisions which efficiently damp random motions. Physically, the Krook approximation states that collisions remove particles from phase space at a rate proportional to their phase space density, and returns them with a new, thermalized distribution function. This thermal distribution function is assumed to be isotropic; isotropy of the stress tensor does obtain in the highly collisional (hydrodynamic) limit (Goldreich & Tremaine 1978).

By definition (8.5) of δ,

$$(p_{11} - p_{22})(\partial_t \delta)_c = \cos 2\delta (\partial_t p_{r\phi})_c - \sin\delta\cos\delta(\partial_t p_{rr} - \partial_t p_{\phi\phi})_c\,. \tag{8.13}$$

Insertion of (8.11) into (8.13) yields the desired result that $(\partial_t \delta)_c = 0$.

Bibliography

[1] Borderies, N., Goldreich, P., & Tremaine, S., 1985, Icarus, 63, 406

[2] Goldreich, P. & Tremaine, S., 1978, Icarus, 34, 227

[3] Quataert, E. & Chiang, E.I. 2000, submitted to ApJ (astro-ph/0003368)

[4] Stone, J. M. & Balbus, S. A., 1996, ApJ, 464, 364