

UC Berkeley

UC Berkeley Electronic Theses and Dissertations

Title

From Dust to Dust: Protoplanetary Disk Accretion, Hot Jupiter Climates, and the Evaporation of Rocky Planets

Permalink

<https://escholarship.org/uc/item/9jq3136f>

Author

Perez-Becker, Daniel Alonso

Publication Date

2013

Peer reviewed|Thesis/dissertation

**From Dust to Dust:
Protoplanetary Disk Accretion, Hot Jupiter Climates,
and the Evaporation of Rocky Planets**

By

Daniel Alonso Perez-Becker

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Physics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Eugene Chiang, Co-chair

Professor Christopher McKee, Co-chair

Professor Eliot Quataert

Professor Geoffrey Marcy

Fall 2013

**From Dust to Dust:
Protoplanetary Disk Accretion, Hot Jupiter Climates,
and the Evaporation of Rocky Planets**

Copyright 2013
by
Daniel Alonso Perez-Becker

Abstract

**From Dust to Dust:
Protoplanetary Disk Accretion, Hot Jupiter Climates,
and the Evaporation of Rocky Planets**

by

Daniel Alonso Perez-Becker

Doctor of Philosophy in Physics

University of California, Berkeley

Professor Eugene Chiang, Co-chair
Professor Christopher McKee, Co-chair

This dissertation is composed of three independent projects in astrophysics concerning phenomena that are concurrent with the birth, life, and death of planets. In Chapters 1 & 2, we study surface layer accretion in protoplanetary disks driven stellar X-ray and far-ultraviolet (FUV) radiation. In Chapter 3, we identify the dynamical mechanisms that control atmospheric heat redistribution on hot Jupiters. Finally, in Chapter 4, we characterize the death of low-mass, short-period rocky planets by their evaporation into a dusty wind.

Chapters 1 & 2: Whether protoplanetary disks accrete at observationally significant rates by the magnetorotational instability (MRI) depends on how well ionized they are. We find that disk surface layers ionized by stellar X-rays are susceptible to charge neutralization by condensates—ranging from μm -sized dust to angstrom-sized polycyclic aromatic hydrocarbons (PAHs). Ion densities in X-ray-irradiated surfaces are so low that ambipolar diffusion weakens the MRI. In contrast, ionization by stellar FUV radiation enables full-blown MRI turbulence in disk surface layers. Far-UV ionization of atomic carbon and sulfur produces a plasma so dense that it is immune to ion recombination on grains and PAHs. Even though the FUV-ionized layer is ~ 10 – 100 times more turbulent than the X-ray-ionized layer, mass accretion rates of both layers are comparable because FUV photons penetrate to lower surface densities than do X-rays. We conclude that surface layer accretion occurs at observationally significant rates at radii $\gtrsim 1$ – 10 AU. At smaller radii, both X-ray- and FUV-ionized surface layers cannot sustain the accretion rates generated at larger distance and an additional means of transport is needed. In the case of transitional disks, it could be provided by planets.

Chapter 3: Infrared light curves of transiting hot Jupiters present a trend in which the atmospheres of the hottest planets are less efficient at redistributing the stellar energy absorbed on their daysides than colder planets. Here we present a shallow water model of the atmospheric dynamics on synchronously rotating planets that explains why heat

redistribution efficiency drops as stellar insolation rises. To interpret the model, we develop a scaling theory which shows that the timescale for gravity waves to propagate horizontally over planetary scales, τ_{wave} , plays a dominant role in controlling the transition from small to large temperature contrasts. This implies that heat redistribution is governed by a wave-like process, similar to the one responsible for the weak temperature gradients in the Earth’s tropics. When atmospheric drag can be neglected, the transition from small to large day-night temperature contrasts occurs when $\tau_{\text{wave}} \sim \sqrt{\tau_{\text{rad}}/\Omega}$, where τ_{rad} is the radiative relaxation time of the atmosphere and Ω is the planetary rotation frequency. Our results subsume the more widely used timescale comparison for estimating heat redistribution efficiency between τ_{rad} and the horizontal day-night advection timescale, τ_{adv} .

Chapter 4: Short-period exoplanets can have dayside surface temperatures surpassing 2000 K, hot enough to vaporize rock and drive a thermal wind. Small enough planets evaporate completely. Here we construct a radiative-hydrodynamic model of atmospheric escape from strongly irradiated, low-mass rocky planets, accounting for dust-gas energy exchange in the wind. Rocky planets with masses $\lesssim 0.1M_{\oplus}$ (less than twice the mass of Mercury) and surface temperatures $\gtrsim 2000$ K are found to disintegrate entirely in $\lesssim 10$ Gyr. When our model is applied to *Kepler* planet candidate KIC 12557548b—which is believed to be a rocky body evaporating at a rate of $\dot{M} \gtrsim 0.1 M_{\oplus}/\text{Gyr}$ —our model yields a present-day planet mass of $\lesssim 0.02 M_{\oplus}$ or less than about twice the mass of the Moon. Mass loss rates depend so strongly on planet mass that bodies can reside on close-in orbits for Gyrs with initial masses comparable to or less than that of Mercury, before entering a final short-lived phase of catastrophic mass loss (which KIC 12557548b has entered). We estimate that for every object like KIC 12557548b, there should be 10–100 close-in quiescent progenitors with sub-day periods whose hard-surface transits may be detectable by *Kepler*—if the progenitors are as large as their maximal, Mercury-like sizes. KIC 12557548b may have lost $\sim 70\%$ of its formation mass; today we may be observing its naked iron core.

A mis padres

Contents

List of Figures	v
List of Tables	viii
Acknowledgments	ix
1 Surface Layer Accretion in Transitional and Conventional Disks: from Polycyclic Aromatic Hydrocarbons to Planets	1
1.1 Introduction	2
1.1.1 The Need for Companions	2
1.1.2 Companions are Not Enough: The Case for the Magnetorotational Instability for the Origin of Disk Viscosity	5
1.1.3 The Threat Posed by Polycyclic Aromatic Hydrocarbons to the MRI	6
1.2 Model for Disk Ionization	8
1.2.1 X-ray and Cosmic-ray Ionization Rates and Gas Densities	8
1.2.2 Gas Temperature	11
1.2.3 Chemical Network	12
1.2.4 Properties and Abundances of Trace Species	14
1.2.5 Collisional Charging of PAHs and Grains	17
1.2.6 Numerical Method of Solution	19
1.3 Results	20
1.3.1 Charge Distributions on PAHs and Grains	21
1.3.2 Ionization Fraction vs. PAH Abundance	21
1.3.3 Degree of Magnetic Coupling: Re and Am	31
1.3.4 Comparison with Previous Work: Ionization Fractions	37
1.4 Summary and Discussion	41
1.4.1 Future Directions	42
2 Surface Layer Accretion in Conventional and Transitional Disks Driven by Far-Ultraviolet Ionization	45
2.1 Introduction	46
2.2 Model for FUV ionization	50

2.2.1	FUV Luminosity	50
2.2.2	Total Gas Columns and Densities	51
2.2.3	Gas Temperature	51
2.2.4	Very Small Grains (VSGs)	52
2.2.5	Polycyclic Aromatic Hydrocarbons (PAHs)	53
2.2.6	H ₂ : Photodissociation and Re-Formation on Grains	54
2.2.7	Carbon: Abundance and Ionization Equilibrium	55
2.2.8	Sulfur: Abundance and Ionization Equilibrium	56
2.2.9	Numerical Method of Solution	57
2.3	Results for MRI-Active Surface Densities	57
2.3.1	Photodissociation and Ionization Fronts	57
2.3.2	MRI-Active Surface Density Σ^*	59
2.3.3	Parameter Survey, Including Sensitivity to PAHs	62
2.3.4	Hall Diffusion	65
2.3.5	The Possibility of Turbulent Mixing: Chemical Equilibration Timescale vs. Dynamical Timescale	66
2.4	Summary and Implications for Disk Accretion	66
2.4.1	Transitional Disks	71
3	Atmospheric Heat Redistribution on Hot Jupiters	74
3.1	Introduction	75
3.2	The Model	80
3.3	Numerical Solutions	87
3.3.1	Basic Behavior of the Solutions	87
3.3.2	Physical Explanation for Forcing Amplitude Dependence	88
3.3.3	Metric for the Day-Night Contrast	89
3.4	Analytic Theory for Day-Night Differences	91
3.4.1	Drag-dominated: Valid for Both Equatorial and Non-equatorial Regions	96
3.4.2	Coriolis-dominated: Valid for Non-equatorial Regions Only	96
3.4.3	Advection- or Vertical-transport-dominated: Valid for $Ro \gtrsim 1$	97
3.5	Interpretation of Theory	98
3.5.1	Timescale Comparison	98
3.5.2	Vertical Advection	101
3.5.3	Wave Adjustment Mechanism	102
3.6	Application to Hot Jupiter Observations	104
3.7	Conclusions	105
4	Catastrophic Evaporation of Rocky Planets	108
4.1	Introduction	109
4.1.1	KIC 12557548b: A Catastrophically Evaporating Planet	109
4.1.2	Plan of This Chapter	111
4.2	The Model	112

4.2.1	Base Conditions of the Atmosphere	112
4.2.2	Equations Solved	114
4.2.3	Relaxation Code	117
4.3	Results	119
4.3.1	Isothermal Solution	119
4.3.2	Full Solution	121
4.3.3	Mass-Loss Histories	127
4.3.4	Possible Mass Loss Histories for KIC 1255b	127
4.4	Discussion	128
4.4.1	Mass Loss is Not Energy-Limited	131
4.4.2	Dust-Gas Dynamics	131
4.4.3	Time Variability	132
4.4.4	Long-Term Orbital Evolution	134
4.4.5	Flow Confinement by Stellar Wind	135
4.4.6	Occurrence Rates of Close-In Progenitors	136
4.4.7	Iron Planets	137
4.4.8	Dust Condensation	138
4.5	Conclusions	141
	Bibliography	143

List of Figures

1.1	Diagram of X-ray ionized disk surface layers	8
1.2	Diagram of our chemical reaction network	13
1.3	Equilibrium charge distribution on PAHs	22
1.4	Equilibrium charge distribution on grains	23
1.5	Average charge state of PAHs ad a function of PAH abundance	24
1.6	Average charge state of grains ad a function of grain abundance	25
1.7	Ionization fraction as a function of PAH abundance for the standard metal abundance ($x_M = 10^{-8}$)	30
1.8	Ionization fraction as a function of PAH abundance for the metal-rich case ($x_M = 10^{-6}$)	32
1.9	Magnetic Reynolds number Re and ambipolar diffusion number Am as a function of surface density Σ at $a = 3$ AU	35
1.10	Magnetic Reynolds number Re and ambipolar diffusion number Am as a function of surface density Σ at $a = 30$ AU	36
1.11	Ratio of the chemical equilibration timescale t_{eq} to the dynamical time $t_{dyn} = \Omega^{-1}$, for $a = 3$ AU and $x_M = 10^{-8}$	38
1.12	Test comparison with the model of Bai & Goodman (2009)	39
1.13	Test comparison with the model of Turner et al. (2010)	40
2.1	H ₂ photodissociation front, and C and S ionization fronts, for our standard model at $a = 3$ AU	58
2.2	Magnetic Reynolds number Re and ambipolar diffusion number Am versus vertical surface density Σ for standard model parameters and $a = 3$ AU	60
2.3	Maximum transport parameter α and mass accretion rate $\dot{M}$ versus vertical surface density Σ for standard model parameters and $a = 3$ AU	61
2.4	Dependence of Σ^* to model parameters	63
2.5	Electron fraction as a function of PAH abundance for FUV and X-ray ionized surface layers	64
2.6	Hall parameter Ha as a function of a in FUV-ionized surface layers	65
2.7	Ratio of the ion-electron recombination timescale t_{rec} to the dynamical time t_{dyn} , as a function of a	67
2.8	Accretion rates for conventional (hole-less) disks as a function of a	69

2.9	Disk accretion rates driven by the combined effects of X-ray and sideways cosmic ray ionization	70
2.10	Accretion rates for transitional disks driven by FUV ionization of gas at the inner rim of the disk	72
3.1	Fractional day-night infrared flux variations A_{obs} vs. global-mean equilibrium temperature T_{eq} for hot Jupiters with measured light curves	76
3.2	Diagram of the two-layer shallow water model	79
3.3	Equirectangular maps of steady-state geopotential (gh) contours for the equilibrated solutions of the shallow-water model for a forcing amplitude of $\Delta h_{\text{eq}}/H = 1$	81
3.4	Equirectangular maps of steady-state geopotential (gh) contours for the equilibrated solutions of the shallow-water model for a forcing amplitude of $\Delta h_{\text{eq}}/H = 0.001$	82
3.5	Absolute values of the zonal components of individual terms in the momentum equation	84
3.6	Contours of day-night height amplitude A and $\tau_{\text{adv}}/\tau_{\text{wave}}$ for the numerical and analytical models	85
3.7	Contours of normalized day-night amplitude A of our full shallow water simulations for different $\Delta h_{\text{eq}}/H$ values	92
3.8	Contours of equatorial day-night amplitude A_{equator} of our full shallow water simulations and our analytical model for different $\Delta h_{\text{eq}}/H$ values	93
3.9	Simplified diagram of the upper layer of the shallow water model	94
3.10	Contours of day-night height amplitude A as would be predicted by comparing τ_{rad} vs. τ_{adv} —the horizontal advective timescale	99
3.11	Comparison between observed fractional day-night infrared flux variations $A_{\text{obs}}(T_{\text{eq}})$ and shallow water model results $A(T_{\text{eq}})$	106
4.1	Equilibrium vapor pressures of pyroxene and olivine vs. temperature	113
4.2	Evaporative mass loss rates $\dot{M}$ vs. planet mass M for isothermal dust-free winds	120
4.3	Radial dependence of wind properties in the full solution for a planet of mass $M = 0.03M_{\oplus}$	122
4.4	Dependence of $\dot{M}$ and τ_{surface} on dust abundance for a planet of mass $M = 0.03M_{\oplus}$	123
4.5	Radial dependence of wind properties in the full solution for a planet of mass $M = 0.01M_{\oplus}$	125
4.6	Radial dependence of wind properties in the full solution for a planet of mass $M = 0.07M_{\oplus}$	126
4.7	Dependence of $\dot{M}$ on dust abundance for a planet of mass $M = 0.006M_{\oplus}$ or $M = 0.1M_{\oplus}$	128
4.8	Maximum mass loss rates $\dot{M}$ vs. planet mass M for the full model	129
4.9	Mass-loss histories $\dot{M}(t)$ for the isothermal and full models	130
4.10	Mass loss rates $\dot{M}$, computed using the isothermal model at $T = 2145$ K, for iron planets and olivine planets of varying M	137

4.11	Mass-loss histories $M(t)$ for an iron planet	139
4.12	Trajectories of the wind in pressure-temperature space, computed using the full model for $M = 0.01, 0.03$, and $0.07M_{\oplus}$	140

List of Tables

1.1	Model Parameters	10
1.2	Chemical Reactions Including Collisional Charging of PAHs and Grains.	14
1.3	Dust Settling Parameter ϵ_{grain} for Some Transitional Disks	16
1.4	Densities of Charged Species for our Standard Model ($x_{\text{M}} = 10^{-8}, L_{\text{X}} = 10^{29} \text{ erg s}^{-1}$)	27

Acknowledgments

I am extremely grateful to my academic advisor, Eugene Chiang, whose guidance was fundamental in my development as a scientist. In addition to the countless hours spent advising my research projects, Eugene has coached me in how to give effective presentations. He also helped me improve my technical writing skills, and perhaps most importantly, instilled me to always make an order-of-magnitude estimation first—fighting against a lifetime of instruction in physics. Eugene, I feel very fortunate and privileged to have had you as my advisor.

I am also very thankful to Adam Showman, with whom I started collaborating at the International Summer Institute for Modeling in Astrophysics (ISIMA) in Beijing. He has since been pouring time into helping me understand atmospheric dynamics and been a great tele-advisor. He treats me as a member of his group and even provided funding to present our work at various conferences.

My graduate career began in experimental astrophysics, working with Steve Boggs. Although not reflected in this dissertation, I am very thankful to have worked on the Nuclear Compton Telescope. Those were very exciting times, especially during the 2009 New Mexico flight campaign. Many thanks to Gloria Koenigsberger, who has been my mentor since my teenage years. I thank George Smoot for his support and advice while working on the Galactic Emission Mapping project before starting graduate school. I thank Chris McKee, Eliot Quataert, and Geoff Marcy for serving on my thesis committee and advice.

While at Berkeley, I made many great friends through the years. They have been a source of support and inspiration. I am especially grateful to Ana Villarreal, Denia Đokić, Ilse Ruiz Mercado, Abel López Jiménez, Mia Ihm, Bear Haynes, Deb Linton, Pierre-Adrien “PD” Itty, Alex Hegai, Thomas van Alsenoy, Brandon Friedman, Emil Zeuthen, Angie Little, Jeremy Mardon, Allen Sussman, Los Titanic, Pablo Ortega, Homero Lara, Paloma Zapata, Marcela Rico, Emilio Martínez, Valeria Brabata, Maria Ávila, Rachel Hochman, Logan Linderman, Brenda Corona, Pedro Reynoso, and Fermin Reygadas. Some of the most memorable times I have had in the last years were with Andy Ehrensberger aboard *Chelsea B* on expeditions in- and outside of the San Francisco Bay.

Over the last year, I have had the fortune to be supported by a wonderful person. My girlfriend Anna Cetti played a key role in my preparation for the qualifying exam, helping me overcome my anxieties of speaking and answering questions in public. Her words of encouragement lift me up even in the bleakest moments.

I cannot thank my parents Inge and Ruy—to whom this dissertation is dedicated—enough for all they have given me. Everything I have accomplished builds upon the foundation they laid. To this day, they have been a constant source of love, support, and encouragement. Thanks to my brothers Baschtl and Ruy, for all the great times and trips. Many thanks to my grandparents Irmgard Montfort and Ruy Pérez Tamayo for showing us so much of the World. My grandfather Klaus Becker kindled my curiosity in science from an early age. He and my high school teacher Hugo Fahrenkrug inspired me pursue a career in physics.

I would like to thank some of the great friends I grew up with and continue to be an important part of my life, even if they are now scattered across the globe: Catalina Pereda, Wolke Tobon, Adriana Ortega, Rocío Ponce, Anne Pöhlmann, Karla Nájera, Geli Martínez, Olga Galván, Pato Weber, Salva Muñoz, Checo Ortega, Abraham Avellana, Chucho Zendejas, Esteibol.V Reyes, Chepi Bustos, Congo Nuñez, Chango Contreras, Rufino Bolado, and the rest of the Miaus.

My doctoral studies were possible by funding from a Berkeley University Fellowship, a Graduate Research Fellowship from The National Science Foundation, and a CONACyT/UC-MEXUS Fellowship. Thanks to Chris Ormel for suggesting the title of this dissertation at dinner. I would also like to thank the staff from the Physics and Astronomy Departments for all their help, especially Anne Takizawa, Donna Sakima, Dexter Stewart, and Nina Ruymaker. This dissertation was typeset using the [ucastrothesis](#) L^AT_EX template.

Chapter 1

Surface Layer Accretion in Transitional and Conventional Disks: from Polycyclic Aromatic Hydrocarbons to Planets

Daniel Perez-Becker & Eugene Chiang (2011), ApJ, 727:2.¹

Abstract

“Transitional” T Tauri disks have optically thin holes with radii $\gtrsim 10$ AU, yet accrete up to the median T Tauri rate. Multiple planets inside the hole can torque the gas to high radial speeds over large distances, reducing the local surface density while maintaining accretion. Thus multi-planet systems, together with reductions in disk opacity due to grain growth, can explain how holes can be simultaneously transparent and accreting. There remains the problem of how outer disk gas diffuses into the hole. Here it has been proposed that the magnetorotational instability (MRI) erodes disk surface layers ionized by stellar X-rays. In contrast to previous work, we find that the extent to which surface layers are MRI-active is limited not by ohmic dissipation but by ambipolar diffusion, the latter measured by Am : the number of times a neutral hydrogen molecule collides with ions in a dynamical time. Simulations by Hawley & Stone showed that $Am \sim 100$ is necessary for ions to drive MRI turbulence in neutral gas. We calculate that in X-ray-irradiated surface layers, Am typically varies from $\sim 10^{-3}$ to 1, depending on the abundance of charge-adsorbing polycyclic aromatic hydrocarbons, whose properties we infer from *Spitzer* observations. We conclude that ionization of H_2 by X-rays and cosmic rays can sustain, at most, only weak MRI turbulence in surface layers 1–10 g/cm² thick, and that accretion rates in such layers are too small compared to observed accretion rates for the majority of disks.

¹The Astrophysical Journal, 727:2, 2011 January 20
 © 2011. The American Astronomical Society. All rights reserved.

1.1 Introduction

On the road from molecular clouds to planetary systems, transitional disks are among the brightest signposts. Encircling T Tauri and Herbig Ae/Be stars having ages of 1–10 Myr, these disks have large inner holes nearly devoid of dust. Identified by spectral energy distributions (SEDs; e.g., [Strom et al. 1993](#); [Calvet et al. 2005](#); [Kim et al. 2009](#)) and imaged directly (e.g., [Ratzka et al. 2007](#); [Hughes et al. 2007](#); [Brown et al. 2009](#)), transitional disk cavities have radii on the order of 3–100 AU. Transitional disks are so named ([Strom et al. 1990](#)) because they might represent an evolutionary link between optically thick disks without holes (e.g., [Watson et al. 2007a](#)) and debris disks containing only rings of optically thin dust (e.g., [Wyatt 2008](#)). They are of special interest not least because their central clearings may harbor nascent planets, potentially detectable against relatively weak backgrounds.

1.1.1 The Need for Companions

The idea that transitional disk holes are swept clean by companions, possibly of planetary mass, is natural. We adhere to this interpretation, although not all our arguments as given below are the ones usually discussed.

Stellar-mass Companions

Roughly half of all transitional disks—6 out of 13 in the sample of [Kim et al. \(2009\)](#)—are already known to contain stellar-mass companions. A prototypical example is CoKu Tau/4: a transitional system whose hole is practically empty, both of dust ([D’Alessio et al. 2005](#)) and gas (G. Blake, private communication, 2007). Inside its hole of radius ~ 10 AU resides a nearly equal mass K-star binary having a projected separation of 8 AU ([Ireland & Kraus 2008](#)). Gravitational torques exerted by the binary can easily counteract viscous torques in the disk ([Goldreich & Tremaine 1980](#)), staving off accretion onto either star ([Artymowicz & Lubow 1994](#); [Ochi et al. 2005](#)). Indeed stellar accretion rates for CoKu Tau/4 are unmeasurably small, $\lesssim 10^{-10} M_{\odot} \text{ yr}^{-1}$ ([Najita et al. 2007](#)).

Not every hole, however, is as empty as that of CoKu Tau/4. In many cases there is a sprinkling of dust: optical depths at $10 \mu\text{m}$ wavelength for many sources range from 0.01–0.1 ([Calvet et al. 2002, 2005](#)). Observations of rovibrational emission from warm, optically thick CO imply that gas fills many disk holes ([Salyk et al. 2007](#)), albeit with surface densities that may be far below those of conventional disks (we will argue below that this is in fact the case). Most germane to our work, the host stars of many transitional systems actively accrete, at rates that on average are somewhat lower than those of conventional disks ([Najita et al. 2007](#)), but which in several instances approach $10^{-8} M_{\odot} \text{ yr}^{-1}$, the median T Tauri rate. The holes of these systems must contain accreting gas.

How can we reconcile the fact that many holes contain gas, accreting at rates approaching those of conventional disks, with the fact that the holes contain only trace amounts of dust? To explain the paradox of simultaneous accretion and hole transparency, appeals

are sometimes made to grain growth, or the filtering of dust out of gas by hydrodynamic mechanisms (Paardekooper & Mellema 2006; Rice et al. 2006) or radiation pressure (Chiang & Murray-Clay 2007).

These proposals, which invoke changes in disk opacity, may be part of the solution. But they cannot alone explain the observations. Ward (2009) has criticized the hydrodynamic filter. The force of radiation pressure depends on uncertain optical constants and grain porosities, and is likely to expel only grains having a narrow range of sizes (Burns et al. 1979). Even if grains have the right properties to be blown out by radiation pressure in vacuum, the inward flow of accreting gas may be strong enough to carry as much as half of the grains that leak from the rim into the hole (Chiang & Murray-Clay 2007). Grain growth does not explain why transitional disk accretion rates $\dot{M}$ tend to be several times smaller than for conventional T Tauri systems (Najita et al. 2007). Finally, none of these proposals predicts gapped (“pre-transitional”) disks, which are optically thick at stellocentric distances $a \lesssim 0.15$ AU (e.g., LkCa 15; Espaillat et al. 2007a).

An alternative explanation for why holes can be simultaneously transparent and still contain accreting gas involves the special way in which disk gas accretes in the presence of companions, particularly those on eccentric orbits. If the hole rim leaks gas—and for some combinations of disk viscosity, disk pressure, and binary parameters the rim can be quite leaky (Artymowicz & Lubow 1996; Ochi et al. 2005)—the gas can suddenly plunge inward at rates approaching freefall velocities. The catastrophic loss of angular momentum is enabled by the non-axisymmetric and time-dependent potential of the eccentric binary, which directs gas streamlines onto radial orbits that may intersect and shock. Artymowicz & Lubow (1996) explained the paradox of simultaneous accretion and transparency:

“... [for some circumbinary disk parameters] gravitational resonant torques are able to open a fairly wide gap [hole], while *concurrently* the accretion flow proceeds through that gap in the form of time-dependent, *well-developed* or *efficient* gas stream(s) carrying virtually all the unimpeded mass flux. The radial velocity of the stream is of order [the Kepler velocity], i.e., $\sim Re$ [Reynolds number] times faster than in the disk. By mass conservation, the axially averaged surface density must differ by a factor of $Re > 10^3$ between the gap and the disk edge region. ... The spectroscopic ramification of this is a deficit of the observed radiation flux emitted at temperatures appropriate for the gap location.” (italics theirs)

Thus reductions in the dust-to-gas ratio by grain growth or dust filtration are not the only processes that can render accreting gas transparent in transitional disk holes. Companions can accelerate disk gas to such high radial speeds that, by mass continuity, the surface density in *both* gas and dust is reduced by orders of magnitude. According to this explanation, the reduction in total surface density is not necessarily due to consumption of gas by companions, but is rather due to gravitational forcing.

Multiple Planetary Mass Companions

Though a stellar-mass companion can exert torques strong enough to maintain holes of large size, a single planet-mass companion on a circular orbit cannot do the same job. For observationally reasonable values of the disk viscosity, a single companion having of order ~ 1 Jupiter mass embedded within a disk at least a few times more massive carves out only a narrow gap (e.g., [Lubow & D’Angelo 2006](#); [Crida & Morbidelli 2007](#)). Disk gas funnels past the planet by traveling on horseshoe-like orbits ([Lubow et al. 1999](#)). A single planet may siphon off some of the gas that flows past it, but the disk accretion rate inside the planet’s orbit is reduced from that outside by a factor of $\lesssim 10$ ([Lubow & D’Angelo 2006](#)). This modest reduction in $\dot{M}$, combined with the narrowness of the gap seen in simulations ($\Delta r/r \sim 0.1$), implies inner disks far too extensive and optically thick to explain transitional systems.

How, then, do planets fit in? In those cases where the central stars of transitional disks lack stellar-mass companions, the same tasks of maintaining hole rims and increasing accretion velocities v (but not accretion rates $\dot{M}$) can be performed, not by a single planet, but by a system of multiple planets. We imagine a series of planets, with the outermost lying just interior to and shepherding the hole rim. Gas that leaks from the rim is torqued from planet to planet, all the way down to the central star, its optical depth decreasing inversely as its radial speed. The more massive the planets and the more eccentric their orbits, the fewer of them should be required.

Such a picture is supported by numerical simulations of Jupiter and Saturn embedded within a viscous disk ([Masset & Snellgrove 2001](#); [Morbidelli & Crida 2007](#)). In these simulations the two planets were close enough that their gaps overlapped. Gas outside Saturn’s orbit executed half a horseshoe turn relative to Saturn, and then another half-horseshoe turn relative to Jupiter, thereby crossing from the outer disk through the Jupiter-Saturn common gap into the inner disk. [Morbidelli & Crida \(2007\)](#) found that the surface density in the gap region was reduced by 1–2 orders of magnitude, at least near Jupiter.

As our paper was being reviewed, we became aware of planet-disk simulations by [Zhu et al. \(2010, submitted\)](#) which included as many as 4 Jupiter-mass planets and whose results supported those of [Morbidelli & Crida \(2007\)](#). Depending on the assumed efficiency with which planets consumed disk gas, a set of four planets was found to reduce surface densities in their vicinity by up to 2 orders of magnitude, while disk accretion rates were reduced by factors $\lesssim 10$ (see their run P4A10). However, such surface density suppressions are not by themselves large enough to explain the observed low optical depths of disk holes. [Zhu et al. \(2010\)](#) concluded that reductions in gas opacity by some means of dust depletion (e.g., grain growth) are still required.

Companions can also accommodate gapped or “pre-transitional” disks in which optically thin holes contain optically thick annuli. As inferred from spatially unresolved spectra, these annuli are narrow and abut their host stars, extending mere fractions of an AU in radius ([Espaillet et al. 2007a](#); but see also [Eisner et al. 2009](#) who showed using spatially resolved observations that the gapped disk interpretation of SR 21 is incorrect). In regions far removed

from secondary companions—in particular, in those regions closest to the primary star where the potential is practically that of a point mass—the infall speeds of accreting gas must slow back down to the normal rate set by disk viscosity. By continuity, the surface density must rise back up, and optical thickness is thus restored.

1.1.2 Companions are Not Enough: The Case for the Magnetorotational Instability for the Origin of Disk Viscosity

Our case for companions presumes a source of disk viscosity. While a stellar-mass companion or a system of multiple planets can transport gas quickly, effectively generating an enormous viscosity in their vicinity (i.e., inside the hole), they cannot cause the outer disk to diffuse in the first place. An inviscid outer disk will not leak. Another source of viscosity has to act in the outer disk, causing it bleed inward and supply the observed accretion rates $\dot{M}$. We now turn to the main subject of this paper, the possibility that the magnetorotational instability is the source of viscosity in the outer disk.

The magnetorotational instability (MRI) amplifies magnetic fields in outwardly shearing disks and drives turbulence whose Maxwell stresses transport angular momentum outward and mass inward (for a review, see [Balbus 2009](#)). Gas must be sufficiently well ionized for the MRI to operate. For the most part, T Tauri and Herbig Ae disks are too cold at their midplanes for thermal ionization to play a role there. The hope instead is that X-rays emitted by host stars can provide the requisite ionization in irradiated disk surface layers ([Glassgold et al. 1997](#)). The basic picture was conceived by [Gammie \(1996\)](#), who proposed that disk surface layers ionized by some non-thermal means may accrete, leaving behind magnetically “dead” midplane gas. Like other workers (e.g., [Bai & Goodman 2009](#), hereafter BG; and [Turner et al. 2010](#), hereafter TCS), we focus in this study on ionization of H_2 by X-rays. Ionization of trace species by ultraviolet (UV) radiation is also potentially important—we discuss this topic briefly at the close of our paper.

The exposed rim of a transitional disk constitutes a kind of surface layer. X-rays may penetrate the rim wall, activate the MRI there, and dislodge a certain radial column of gas every diffusion time ([Chiang & Murray-Clay 2007](#), hereafter CMC). Within the MRI-active column, both the magnetic Reynolds number

$$Re \equiv \frac{c_s h}{D} \approx 1 \left(\frac{x_e}{10^{-13}} \right) \left(\frac{T}{100 \text{ K}} \right)^{1/2} \left(\frac{a}{\text{AU}} \right)^{3/2} \quad (1.1)$$

and the ion-neutral collision rate (normalized to the orbital frequency)

$$Am \equiv \frac{x_i n_{\text{H}_2} \beta_{\text{in}}}{\Omega} \approx 1 \left(\frac{x_i}{10^{-8}} \right) \left(\frac{n_{\text{H}_2}}{10^{10} \text{ cm}^{-3}} \right) \left(\frac{a}{\text{AU}} \right)^{3/2} \quad (1.2)$$

must be sufficiently large for magnetic fields to couple well to the overwhelmingly neutral disk gas. Here T is the gas temperature, c_s is the gas sound speed, $h = c_s/\Omega$ is the gas scale height, Ω is the Kepler orbital frequency, $D = 234 (T/\text{K})^{1/2} x_e^{-1} \text{ cm}^2 \text{ s}^{-1}$ is the magnetic

diffusivity, $x_{e(i)}$ is the fractional abundance of electrons (ions) by number, n_{H_2} is the number density of hydrogen molecules, $\beta_{in} \approx 1.9 \times 10^{-9} \text{ cm}^3 \text{ s}^{-1}$ is the collisional rate coefficient for ions to share their momentum with neutrals (Draine et al. 1983), and a is the disk radius.

Dimensionless number (1.1) governs how well magnetic fields couple to plasma, while (1.2) assesses how well plasma couples to neutral gas. Both these numbers must be large for good coupling between magnetic fields and neutral gas. Numerical simulations have suggested critical values Re^* of $\sim 10^2$ – 10^4 (Fleming et al. 2000),² depending on the initial field geometry, and Am^* of $\sim 10^2$ (Hawley & Stone 1998, hereafter HS). Some studies (e.g., TCS) assumed $Am^* \sim 1$ in their determination of the thicknesses of MRI-active surface layers, but numerical simulations of marginally coupled ion-neutral systems indicated Am^* may be 2 orders of magnitude higher (HS). The value of Am^* is critical to our work.

For typical T Tauri parameters, CMC found active radial column densities $N^* \sim 5 \times 10^{23} \text{ cm}^{-2}$ or equivalently mass columns of $\Sigma^* \sim 2 \text{ g cm}^{-2}$ —essentially the stopping column for 3 keV X-rays. When they combined their derived value for N^* with an assumed value for the dimensionless disk viscosity $\alpha \sim 10^{-2}$, the accretion rates of many transitional systems were successfully reproduced. According to this model, the maximum accretion rate $\dot{M}$ inside the hole is set by conditions at the rim wall, i.e., by how large a radial column N^* the MRI can draw from the rim. Stellar or planetary companions, known or suspected to be present (Section 1.1.1), regulate how quickly this leaked material spirals onto the host star—these companions modulate the radial inflow speed $v(a)$ and thus the surface density $\Sigma(a) = \dot{M}/(2\pi va)$. But the companions inside the hole do not initiate disk accretion. They may reduce $\dot{M}$ by exerting repulsive torques to keep material in the rim wall from leaking in, or by accreting material that flows past (e.g., Lubow & D’Angelo 2006; Najita et al. 2007). But they do not generate a non-zero $\dot{M}$ in the first place. That fundamental task is left to the MRI operating at the rim—or whatever source of anomalous viscosity must be present in the outer disk to make it bleed.

1.1.3 The Threat Posed by Polycyclic Aromatic Hydrocarbons to the MRI

One concern raised by CMC but left quantitatively unaddressed is the degree to which ultra-small condensates—macromolecules whose sizes are measured in angstroms—may thwart the MRI. In planetary atmospheres, aerosols can strongly damp electrical conductivities (e.g., Schunk & Nagy 2004; Borucki & Whitten 2008).³ Most studies of active layers neglect aerosols and fixate on roughly micron-sized grains, despite the fact that in many particle

²The magnetic Reynolds number as we define it is not as accurate a predictor of MRI turbulence as the Elsasser (a.k.a. Lundquist) number, which is given by (1.1) with c_s replaced by the vertical Alfvén speed v_{Az} . Self-consistent resistive MHD simulations by Turner et al. (2007) found that MRI-active regions coincide with Elsasser numbers greater than unity (see also Sano & Inutsuka 2001; Sano & Stone 2002). Our criterion $Re \gtrsim 10^2$ – 10^4 offsets some of the inaccuracy because $v_{Az} \lesssim 10^{-1} c_s$ in simulations of MRI turbulence. In any case we will find that the limiting factor for active surface layers is not Re but rather Am .

³Some fire alarms work on this principle. A radioactive source inside the alarm drives ionization currents in the air which normally complete an electrical circuit. When smoke particles from a fire reduce the density

size distributions, the smallest particles collectively present the greatest geometric surface area and therefore the greatest cross section for electron adsorption and ion recombination. Exceptions include [Sano et al. \(2000\)](#), who in one model considered a grain size distribution extending down to $0.005 \mu\text{m} = 50 \text{ \AA}$, and BG, who considered grain sizes as small as $0.01 \mu\text{m} = 100 \text{ \AA}$. Both studies found that in principle small grains can be deadly to the MRI.

Notwithstanding their possibly decisive role, small grains are sometimes wishfully dismissed as being depleted in number by grain growth, i.e., assimilated into larger grains. Undeniably grains grow ([Blum & Wurm 2008](#); [Chiang & Youdin 2010a](#)), so much so that their collective mass may be concentrated in particles millimeters in size. But the question relevant for ionization chemistry is not where the mass is weighted in the size spectrum of particles, but rather where the collective surface area for charge neutralization is weighted. Determining the grain size distribution in disks seems a problem that cannot be forward modeled with confidence. [Sano et al. \(2000\)](#) and BG instead parameterized the population of small grains and studied the effects of varying their numbers, leaving undecided the question of whether their parameter choices were favored by observation or theory.

Like [Sano et al. \(2000\)](#) and BG, this paper considers the effects of small condensates on the MRI. What is new about our contribution is that we consider the smallest imaginable condensates that are still accessible to observation: polycyclic aromatic hydrocarbons (PAHs). These molecules, typically containing several dozens of carbon atoms, are excited electronically by ultraviolet radiation and fluoresce vibrationally at 3.3, 6.2, 7.7, 8.6, 11.3, and $12.7 \mu\text{m}$, the signature bands of their constituent C-C and C-H bonds (e.g., [Li & Draine 2001](#); [Pendleton & Allamandola 2002](#)). *Spitzer* satellite spectra and ground-based adaptive optics imaging reveal PAHs to be fluorescing strongly in Herbig Ae/Be and T Tauri disk surface layers directly exposed to stellar ultraviolet radiation ([Geers et al. 2006, 2007](#); [Goto et al. 2009](#)). Thus PAHs help to constrain the aerosol abundance where magnetically driven accretion is thought to occur: in disk surface layers.

In this work we incorporate PAHs into a simple chemical network to assess the proposal that X-ray driven MRI operates in disk surface layers, either on the top and bottom faces of conventional hole-less disks, or at the rims of transitional disks. We make as realistic an estimate as we can of the PAH abundance based on observations, to gauge how deep the X-ray-irradiated, MRI-active layer might actually be.

To summarize this introduction: companions—either stars or a system of multiple planets, but not a single Jupiter-mass planet—can help clear the extensive holes of transitional disks. The outermost companion serves to establish the location of the rim where viscous torques in the disk and gravitational torques from the companion seek balance. If gas leaks inward from the outer disk, it is driven onto the host star so quickly by gravitational torques from companions that its optical depth may be reduced by orders of magnitude. Companions, together with reductions in disk opacity by grain growth, thus maintain the transparency of the hole while still permitting stars to accrete gas. But companions do not, in and of themselves, cause gas in the outer disk to diffuse inward. That responsibility may be reserved

of free ions and electrons in air, the circuit is broken and the alarm is triggered.

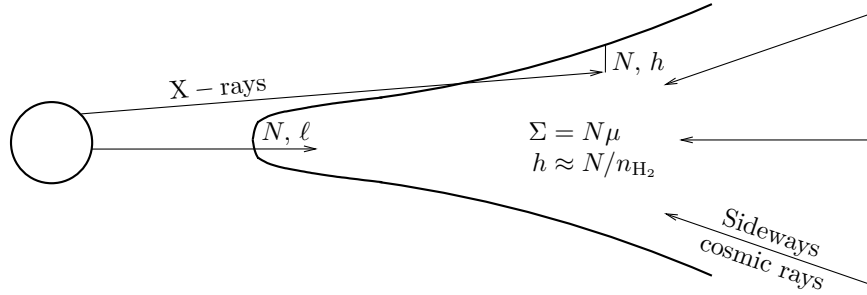


Figure 1.1 X-ray ionized surface layers, located either on the top and bottom faces of a flared disk, or at its inner rim. Our calculations apply to both situations, although they are more accurate for the former. We assume in this work that the lengthscale ℓ over which gas is distributed radially at the hole rim is equal to h , the vertical gas scale height. Other sources of ionization are interstellar cosmic-rays and ultraviolet radiation from the star. At large stellocentric distances ($a \gtrsim 30$ AU), cosmic rays may penetrate the disk from the side. For most of our paper, we neglect ionization of trace species by far ultraviolet radiation, but in Section 1.4.1 we briefly discuss this important topic.

for the MRI—whose ability to operate despite the presence of charge-neutralizing PAHs is the subject of this paper.

Our paper is organized as follows. The ingredients of our numerical model for X-ray-driven ionization chemistry in disk surface layers are laid out in Section 1.2. There we gauge what PAH abundances in disks may be. Results—principally, how Am and Re vary with the column density penetrated by X-rays, and the extent to which PAHs reduce these numbers—are presented in Section 1.3. Analytic interpretations of our numerical results, and direct comparison with previous calculations (BG, TCS), are given there as well. We discuss our main results for X-ray driven MRI in Section 1.4, and close by discussing the possibility of UV-driven MRI.

1.2 Model for Disk Ionization

In this paper we are interested in the degree to which stellar X-rays and Galactic cosmic-rays can ionize H_2 gas in T Tauri disks. In this respect our study is similar to many others, and we make direct comparisons of our work to BG and TCS in Section 1.3.4. For simplicity our model neglects ionization of trace species like C and S by stellar UV radiation. Omitting UV-driven chemistry renders our model inconsistent because our model also includes PAHs, whose abundances we constrain in Section 1.2.4 by using observed PAH emission lines excited by stellar UV radiation. We will discuss the critical issue of UV ionization in Section 1.4.1.

1.2.1 X-ray and Cosmic-ray Ionization Rates and Gas Densities

Chandra spectra of pre-main-sequence stars in the Orion Nebula can be fitted by a pair of thermal plasmas with characteristic temperatures $kT_X \sim 1$ and 3 keV, where k is Boltzmann's constant, and comparable luminosities $L_X \sim 10^{28}$ – 10^{31} erg s $^{-1}$ (Wolk et al. 2005; Preibisch et al. 2005). The softer component is believed to be emitted by shock-heated accreting gas (Stelzer & Schmitt 2004), and the harder component by a strongly magnetized and active stellar corona (Wolk et al. 2005). X-ray luminosities tend to increase with increasing stellar mass and decreasing accretion rate (see Figure 1 of Telleschi et al. 2007, and Figure 17 of Preibisch et al. 2005). These correlations are relevant for transitional disks because some transitional disks are hosted by higher mass Herbig Ae stars, and accretion rates for transitional disks tend to be lower than for conventional disks (Najita et al. 2007). For our standard model we will adopt $L_X = 10^{29}$ erg s $^{-1}$, but we will also experiment with $L_X = 10^{31}$ erg s $^{-1}$. We fix the temperature of the X-ray emitting plasma at $kT_X = 3$ keV, an assumption that ignores how X-ray spectra harden with increasing L_X (Preibisch et al. 2005). Although none of our numerical models explicitly considers $kT_X > 3$ keV, we will discuss quantitatively in Section 1.3.3 how our results scale with the higher ionization rates afforded by a harder ($kT_X = 8$ keV) X-ray spectrum; we will see there that the effects are not large.⁴ See Table 1.1 for a list of all model parameters.

We derive X-ray ionization rates ζ_X as a function of penetration column N from Igea & Glassgold (1999, hereafter IG), who constructed a Monte Carlo radiative transfer model that accounts for Compton scattering and photoionization. Compton scattering enables X-ray photons to penetrate to deeper columns than would otherwise be possible. For our standard model we use IG's Figure 3 for a thermal plasma of $L_X = 10^{29}$ erg s $^{-1}$ and $kT_X = 3$ keV. Their ionization rates were computed for stellocentric distances of 5 and 10 AU; we scale these rates to the stellocentric distances of our model, $a = 3$ and 30 AU, using the geometric dilution factor a^{-2} . We test the accuracy of this approach by scaling their results internally using this dilution factor, finding (by necessity) excellent agreement at low columns where material is optically thin to X-rays, and agreement better than a factor of two at the highest columns calculated. For the case $L_X = 10^{31}$ erg s $^{-1}$, we increase all ionization rates from our standard values by a factor of 100.

Interstellar cosmic-rays can also ionize disk gas, but are attenuated by magnetized stellar winds blowing across disk surface layers. Even the contemporary solar wind, characterized by a mass loss rate of $\sim 10^{-14} M_\odot$ yr $^{-1}$, modulates the cosmic-ray flux at Earth by as much as $\sim 10\%$ with solar cycle (Marsh & Svensmark 2003). T Tauri winds, having mass loss rates up to 5 orders of magnitude higher than that of the solar wind today, seem likely to shield disk surfaces from cosmic-rays directed normal to the disk plane (cf. Turner &

⁴ A minority of sources surveyed by *Chandra* exhibited superhot X-ray flares with peak $L_X \sim 10^{32}$ erg s $^{-1}$ and $kT_X \sim 15$ keV (Getman et al. 2008a; Getman et al. 2008b). The degree to which superhot flares enhance ionization rates depends on the uncertain flare duty cycle. Because only $\sim 10\%$ of the *Chandra* sources flared once or twice over a 15-day observing period, and because each flare lasted less than ~ 1 day, the extra ionization from superhot flares may be modest.

Drake 2009). Nevertheless, cosmic-rays may reach disk gas from the “side,” striking the disk edge-on from the outside. At $a = 3$ AU we estimate that these “sideways cosmic-rays” are too strongly attenuated by intervening disk gas to be significant. The same is not true on the outskirts of the disk at $a = 30$ AU, where column densities measured radially outward may be smaller than the cosmic-ray stopping column of 96 g cm^{-2} (Umebayashi & Nakano 1981). Thus we consider another model at $a = 30$ AU where in addition to our standard X-ray source we include sideways cosmic-rays with a constant, column-independent ionization rate of $\zeta_{\text{CR}} \sim (1/4) \times 10^{-17} \text{ s}^{-1}$ (Caselli et al. 1998, hereafter C98). The factor of $1/4$ is approximately the fraction of the celestial sphere (centered on disk gas at 30 AU) that is not shielded by stellar winds. The total ionization rate $\zeta = \zeta_{\text{X}} + \zeta_{\text{CR}}$.

In all our simulations we neglect ionization by energetic protons emitted by the stars. As discussed by Turner & Drake (2009), estimates of the stellar proton flux rely on extrapolated scaling relations, and the ability of particles to reach the disk surface in the face of strong stellar magnetic fields is uncertain. Moreover, protons are emitted in flares which may occur too infrequently to sustain disk ionization. In one of their models, Turner & Drake (2009) used a time-steady stellar particle luminosity whose ionization rate exceeded, by a factor of 40 at a mass column of $\Sigma = 8 \text{ g cm}^{-2}$, that of an X-ray source having $L_{\text{X}} = 2 \times 10^{30} \text{ erg s}^{-1}$ and $kT_{\text{X}} = 5 \text{ keV}$. This model probably yields a hard upper limit on the stellar proton ionization rate, derived under a set of generous assumptions. Our $L_{\text{X}} = 10^{31} \text{ erg s}^{-1}$ case produces ionization rates ζ that approach those of the aforementioned model to within an order of magnitude. In any case we will see in Section 1.3.3 how our results can be scaled to any ζ .

Figure 1.1 depicts schematically how X-rays irradiate disk surface layers, usually pictured in the vertical direction as ensheathing the disk on its top and bottom faces. But in a transitional disk, a surface layer may also be present in the radial direction, at the rim of the central hole. We consider each of these environments in turn, estimating local number densities $n_{\text{H}_2} [\text{H}_2 \text{ cm}^{-3}]$ from the column density $N [\text{H}_2 \text{ cm}^{-2}]$ penetrated by X-rays.⁵

Surface Layers I: Top and Bottom Faces of a Conventional Flared Disk

When considering the surface layers of a conventional non-transitional disk, we describe our results as a function of the vertical column density N of hydrogen molecules, measured perpendicular to and toward the disk midplane. Thus our N coincides with $N_{\perp}$ of IG, save for a factor of 2 because IG count hydrogen nuclei whereas we count hydrogen molecules. An equivalent measure of vertical column density N is the mass surface density $\Sigma \equiv N\mu$, where $\mu \approx 4 \times 10^{-24} \text{ g}$ is the mean molecular weight of gas.

To good approximation, the local number density

$$n_{\text{H}_2} \approx N/h \quad (1.3)$$

⁵In this paper, ionization rates ζ , column densities N , and fractional densities x are referred to hydrogen molecules, not hydrogen nuclei.

Table 1.1. Model Parameters

Parameter	Variable	Value	Reference
Disk radius	a	3, 30 AU	...
X-ray source luminosity ^a	L_X	$10^{29} (10^{31}) \text{ erg s}^{-1}$	Section 1.2.1
X-ray source temperature	kT_X	3 keV	Section 1.2.1
Cosmic-ray ionization rate	ζ_{CR}	$0, (0.25 \times 10^{-17}) \text{ s}^{-1}$	Caselli et al. (1998)
Initial CO abundance ^b	x_{CO}	10^{-4}	Aikawa et al. (1996)
Total metal abundance ^{a,b}	x_M	$10^{-8} (0, 10^{-6})$	Section 1.2.4
Total grain abundance ^b	x_{grain}	$6 \times 10^{-15} \epsilon_{\text{grain}}$	Section 1.2.4
Grain settling (depletion) factor	ϵ_{grain}	$10^{-3} \leq \epsilon_{\text{grain}} \leq 10^{-1}$	Section 1.2.4
Total PAH abundance ^b	x_{PAH}	$10^{-6} \epsilon_{\text{PAH}}$	Section 1.2.4
PAH depletion factor	ϵ_{PAH}	$10^{-5} \leq \epsilon_{\text{PAH}} \leq 10^{-2}$	Section 1.2.4
Central stellar mass	M_*	$1 M_\odot$	...
Gas temperature	T	80, 30 K	Section 1.2.2

^aValues in parentheses correspond to test cases different from our standard model.

^bAll abundances are relative to H_2 by number.

where the vertical scale height $h = c_s/\Omega = (kT/\mu)^{1/2}/\Omega$. For gas temperature $T \approx 80(30)$ K at $a = 3(30)$ AU (see Section 1.2.2 for how we derive these temperatures), we find $h = 0.09(1.8)$ AU.

Equation (1.3) underpins all our calculations of chemical equilibrium. For a typical $N \sim 10^{23} \text{ H}_2 \text{ cm}^{-2}$, we have $n \sim 7 \times 10^{10} (4 \times 10^9) \text{ H}_2 \text{ cm}^{-3}$ at $a = 3(30)$ AU.

Surface Layers II: Gap Rim of Transitional Disk

We assume that the gap rim is not shadowed from the star by gas interior to the rim. We cannot prove that the rim is not shadowed, but disk models based on the infrared SED suggest it is not (e.g., Calvet et al. 2005). Possibly gas at the rim wall “puffs up” because it is heated by X-rays and can maintain a larger vertical height than gas inside the hole (e.g., Dullemond et al. 2001).

For the case of the rim of a transitional disk, we reinterpret N (equivalently Σ) as the radial column of hydrogen molecules traversed by X-rays (Figure 1). To estimate the local number density n_{H_2} , we need to know the radial lengthscale ℓ over which material at the rim wall is distributed. Plausibly $h \lesssim \ell \lesssim a$. Chiang & Murray-Clay (2007) take $\ell \sim a$, but models based on SEDs and images suggest the rim is much sharper. Here we assume that $\ell \sim h$ so that Equation (1.3) applies equally well to transitional disks as to conventional disks—keeping in mind that N should be measured radially for the former and vertically for the latter.

To calculate ionization rates in the rim, we still use the results of IG, reinterpreting their $N_\perp$ in their Figure 3 as our radial column N . Clearly the scattering geometry differs between the case of a transitional disk rim and the case of the top and bottom faces of IG’s conventional disk. Where material is optically thin to stellar X-rays, the two cases match in ionization rate (per molecule), but where it is optically thick, we underestimate the ionization rate in transitional disk rims by using IG because more X-rays escape by scattering vertically out of conventional disk surface layers than from the rim. Another reason we

underestimate the ionization rate at high column density is because the total X-ray flux per unit surface area of the disk is lower for conventional surface layers—which are illuminated at grazing incidence—than for the rim, which is illuminated at normal incidence. Nevertheless we estimate that these errors are of the order of unity, insofar as the columns that might possibly be MRI-active are not too optically thick to X-rays (Section 1.3.3), because the Thomson scattering phase function is fairly isotropic, and the cross-section for scattering is only comparable to that for photoionization at the relevant photon energies. In any case we will explore the effects of higher ionization rates by running a model with higher $L_X = 10^{31}$ erg s $^{-1}$ (Section 1.3.3).

1.2.2 Gas Temperature

The surface layers of protoplanetary disk atmospheres vary widely in temperature, from ~ 5000 K at the lowest columns where stellar X-rays heat the gas, down to $\lesssim 100$ K at the highest columns where dust reprocesses optical starlight. We draw our temperatures from the thermal model of Glassgold et al. (2004), which in turn is based on the dust temperature model of D’Alessio et al. (1999). At $a = 1$ AU at column densities of interest ($N \gtrsim 10^{22}$ cm $^{-2}$), thermal balance is controlled primarily by reprocessing of starlight by dust, and gas and dust temperatures are nearly equal at ~ 130 K (Glassgold et al. 2004, their Figure 2). We adjust this result for the disk radii of our standard model using the dust temperature scaling law for the midplane of a passive flared disk, $T \propto a^{-3/7}$ (e.g., Chiang & Goldreich 1997). Thus at $a = 3$ AU we have $T = 80$ K, and at $a = 30$ AU we have $T = 30$ K. Note that these temperatures are lower—and arguably more realistic—than those assumed by BG and TCS, who invoked temperatures of the traditional Hayashi nebula without justification.

1.2.3 Chemical Network

Following Ilgner & Nelson (2006a, hereafter IN), CMC, and BG, we apply a simple network of chemical reactions based on that designed for molecular clouds by Oppenheimer & Dalgarno (1974, hereafter OD). Ilgner & Nelson (2006a) and BG compared the results of OD-based schemes to those of more complex networks extracted from the UMIST (University of Manchester Institute of Science and Technology; Woodall et al. 2007, hereafter W07; Vasyunin et al. 2008) database. Fractional electron abundances derived by IN using the simple network were greater than those derived using the complex network, whereas BG, who used a more recent version of the UMIST database, found that the sign of the difference varied from case to case. The magnitude of the difference ranged up to a factor of 10, but was often $\lesssim 3$. Using the simple network seems the most practical approach, if we are content with order-of-magnitude answers. In Section 1.3.4, we test the results of our code against those of BG and TCS.

All reactions in our OD-based network are listed in Table 1.2 and shown schematically in Figure 1.2. Rate coefficients and their temperature dependences are taken from the UMIST database. The chain of events basically proceeds as follows. X-rays ionize H_2 to H_2^+ , which

Table 1.2. Chemical Reactions Including Collisional Charging of PAHs and Grains.

Number	Reaction			Rate Coefficient ^a	Value	Reference
1 ^b	$\text{H}_2 + h\nu$	$\rightarrow$	$\text{H}_2^+ + e^-$	ζ_X	Taken from radiative transfer model	IG
2	$\text{H}_2 + \text{Cosmic-ray}$	$\rightarrow$	$\text{H}_2^+ + e^-$	ζ_{CR}	$(1/4) \times 10^{-17} \text{ s}^{-1}$	C98
3	$\text{H}_2^+ + \text{H}_2$	$\rightarrow$	$\text{H}_3^+ + \text{H}$	$\alpha_{\text{H}_2^+, \text{H}_2}$	2.1×10^{-9}	W07
4	$\text{H}_3^+ + \text{CO}$	$\rightarrow$	$\text{HCO}^+ + \text{H}_2$	$\alpha_{\text{H}_3^+, \text{CO}}$	1.7×10^{-9}	W07
5 ^c	$\text{M} + \text{H}_3^+$	$\rightarrow$	$\text{M}^+ + \text{H}_2 + \text{H}$	$\alpha_{\text{M}, \text{X}^+}$	1.0×10^{-9}	W07
6	$\text{M} + \text{HCO}^+$	$\rightarrow$	$\text{M}^+ + \text{H} + \text{CO}$	$\alpha_{\text{M}, \text{X}^+}$	2.9×10^{-9}	W07
7	$\text{H}_3^+ + e^-$	$\rightarrow$	$\text{H}_2 + \text{H}$	$\alpha_{\text{H}_3^+, e^-}$	$2.3 \times 10^{-8} (T/300\text{K})^{-0.52}$	W07
8	$\text{HCO}^+ + e^-$	$\rightarrow$	$\text{H} + \text{CO}$	$\alpha_{\text{HCO}^+, e^-}$	$2.4 \times 10^{-7} (T/300\text{K})^{-0.69}$	W07
9	$\text{M}^+ + e^-$	$\rightarrow$	$\text{M} + h\nu$	α_{M^+, e^-}	$2.8 \times 10^{-12} (T/300\text{K})^{-0.86}$	W07
10	$\text{PAH}(\text{Z}) + e^-$	$\rightarrow$	$\text{PAH}(\text{Z}-1)$	α_{PAH, e^-}	Section 1.2.5	DS
11 ^d	$\text{PAH}(\text{Z}) + \text{X}^+$	$\rightarrow$	$\text{PAH}(\text{Z}+1)$	$\alpha_{\text{PAH}, \text{X}^+}$	Section 1.2.5	DS
12	$\text{grain}(\text{Z}) + e^-$	$\rightarrow$	$\text{grain}(\text{Z}-1)$	$\alpha_{\text{grain}, e^-}$	Section 1.2.5	DS
13	$\text{grain}(\text{Z}) + \text{X}^+$	$\rightarrow$	$\text{grain}(\text{Z}+1)$	$\alpha_{\text{grain}, \text{X}^+}$	Section 1.2.5	DS
14	$\text{PAH}(\text{Z}=-1) + \text{PAH}(\text{Z}=1)$	$\rightarrow$	$2 \times \text{PAH}(\text{Z}=0)$	$\alpha_{\text{PAH}, \text{PAH}}$	Section 1.2.5	DS
15	$\text{H} + \text{H} + \text{grain}$	$\rightarrow$	$\text{H}_2 + \text{grain}$	$\alpha_{\text{physisorption}}$	Section 1.2.3	BG

^a ζ has units of s^{-1} and α has units of cm^3s^{-1} .

^b $h\nu$ denotes a photon.

^cM represents a gas-phase atomic metal, e.g., Mg.

^d X^+ can be either H_3^+ , HCO^+ , or M^+ .

rapidly reacts with H_2 to produce H_3^+ . The H_3^+ ion combines with CO to form HCO^+ . Most HCO^+ ions dissociatively recombine with free electrons, but some transfer their charge to gas-phase metal atoms such as Mg. Free metals tend to be abundant positive charge carriers, as they recombine with free electrons only by a slow radiative channel. Charged particles in the network (e^- , H_3^+ , HCO^+ , metal^+) can neutralize by collisionally transferring their charge to PAHs and grains. The collisional charging process is described in Section 1.2.5.

The reaction loop is closed by the formation of H_2 on grain surfaces. To compute the rate of this reaction, we take neutral H atoms to collide with grains using the geometrical cross section for grains, and adopt from BG the uniform probability $\eta = 10^{-3}$ for a pair of adsorbed hydrogen atoms to form a hydrogen molecule (see their Equation 27). The precise rate of this reaction is not important for us, as it only sets the equilibrium abundance of H, which is irrelevant for the ionization fraction, as long as $n_{\text{H}} \ll n_{\text{H}_2}$.

In their original study OD included ionization of He and reactions involving atomic and molecular oxygen. We neglect these for simplicity. Most reactions involving oxygen initiate with the formation of the hydroxyl ion ($\text{H}_3^+ + \text{O} \rightarrow \text{OH}^+$), which proceeds at a rate only comparable to the formation of HCO^+ , which we do account for. Thus our neglect of oxygen within the OD framework is not expected to alter our results for the fractional ionization by more than a factor of 2. In any case, in Section 1.3.4 we will compare our results with those of more complex networks considered by BG and TCS.

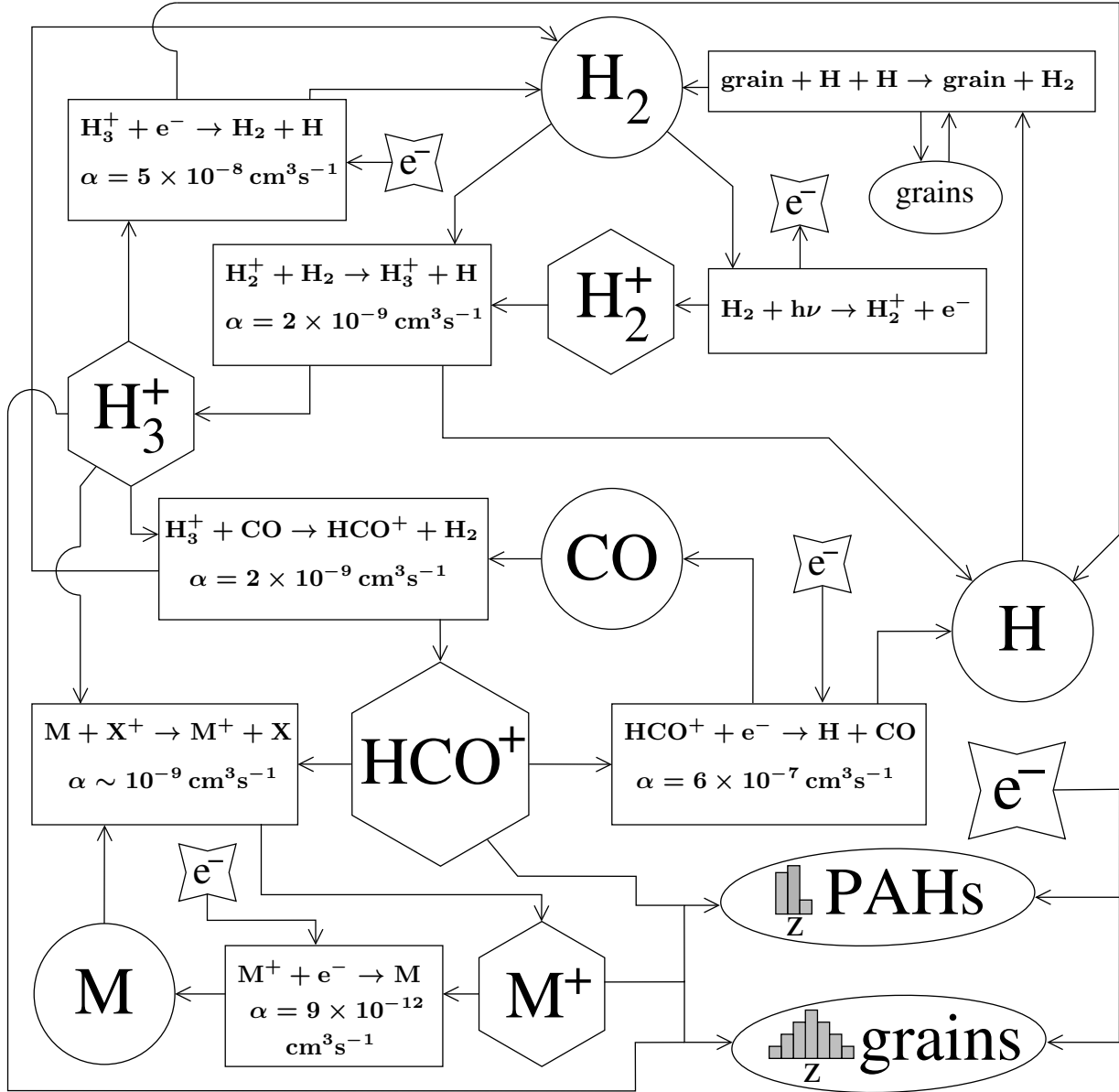


Figure 1.2 Our chemical reaction network, derived from [Oppenheimer & Dalgarno \(1974\)](#). Rate coefficients as shown in this Figure are evaluated at $T = 80$ K. See Table 1.2 for a comprehensive list of all modeled reactions and precise rate coefficients.

1.2.4 Properties and Abundances of Trace Species

The trace ingredients of our model include gas-phase metals (Section 1.2.4), a monodispersion of micron-sized grains (Section 1.2.4), and PAHs (Section 1.2.4).

Gas-phase Metals (Magnesium)

For gas-phase metals which serve importantly as electron donors (OD; Fromang et al. 2002), we are guided by Mg, whose solar abundance is 3.5×10^{-5} atoms per hydrogen nucleus (Lodders 2003). The fraction of Mg that is in the gas phase—neither incorporated into grain interiors nor adsorbed onto grain surfaces—might be at most 3–30% by number, its value in the diffuse interstellar medium (Jenkins 2009). In the dense environments of protoplanetary disks, the gas-phase fraction should be much smaller because magnesium is used toward building grains.

Nominally, our model temperatures of 30–80 K are so low that almost all of the Mg not incorporated into grain interiors should be adsorbed onto grain surfaces, leaving behind only a tiny fraction in the gas phase (Turner et al. 2007, their Section 2.2; see also Equation 26 of BG). Just how tiny is uncertain, given how sensitive the adsorption fraction is to gas temperature, and how steep temperature gradients can be in disk surface layers (Glassgold et al. 2004). Turbulent mixing of hot, high altitude, normally metal-rich layers with cold, low altitude, normally metal-poor layers can also complicate matters (Turner et al. 2007; TCS).

We adopt a standard metal abundance of $x_M = 10^{-8}$ metal atoms per H_2 , which corresponds to a gas-phase fraction of $\sim 10^{-4}$ by number relative to solar. Our choice is similar to those of IN and BG. We also experiment with a metal-free case in which all metals have been adsorbed onto grain surfaces ($x_M = 0$), and a metal-rich case for which $x_M = 10^{-6}$. Although the metal-rich case is not especially realistic and is not justified by our model parameters—in particular our low gas temperatures—we consider it anyway because we would like to understand the effects of metals in principle, and to connect with other studies that consider similarly large metal abundances (CMC; Turner et al. 2007; TCS).

Grains

The number of grains per H_2 molecule is

$$x_{\text{grain}} = \frac{\mu}{\frac{4}{3}\pi s^3 \rho_s} \frac{\rho_{\text{dust}}}{\rho_{\text{gas}}}, \quad (1.4)$$

where $\mu \approx 4 \times 10^{-24}$ g is the mean molecular weight of gas and $\rho_{\text{dust}}/\rho_{\text{gas}}$ is the dust-to-gas mass ratio. For simplicity we consider grains of a single radius $s = 1 \mu\text{m}$ and internal density $\rho_s = 2 \text{ g cm}^{-3}$. There is ample evidence that micron-sized grains abound in surface layers, both from mid-infrared spectra of silicate emission lines (e.g., Natta et al. 2007) and from scattered light images at similar wavelengths (e.g., McCabe et al. 2003).

Table 1.3 Dust Settling Parameter ϵ_{grain} for Some Transitional Disks

Source	ϵ_{grain}	Reference
LkCa 15	10^{-3}	Espaillet et al. (2007a) ; Chiang et al. (2001)
UX Tau A	10^{-2}	Espaillet et al. (2007a)
CS Cha	10^{-2}	Espaillet et al. (2007b)
GM Aur	10^{-1}	Calvet et al. (2005)
DM Tau	10^{-1}	Calvet et al. (2005)

The dust-to-gas ratio in surface layers may differ considerably from its value in the well-mixed diffuse interstellar medium (ISM):

$$\frac{\rho_{\text{dust}}}{\rho_{\text{gas}}} \equiv \epsilon_{\text{grain}} \left. \frac{\rho_{\text{dust}}}{\rho_{\text{gas}}} \right|_{\text{ISM}}, \quad (1.5)$$

where for the ISM of solar abundance $\rho_{\text{dust}}/\rho_{\text{gas}}|_{\text{ISM}} = 0.015$ ([Lodders 2003](#)). Based on model fits to observed far-infrared SEDs ([Chiang et al. 2001](#); [D'Alessio et al. 2006](#); [Dullemond & Dominik 2004](#)), there is consensus that surface layer grains directly illuminated by optical light from their host stars have settled toward the midplane into regions of denser gas. Thus $\epsilon_{\text{grain}} < 1$, but actual values are not known with certainty, because small changes in the SED resulting from small changes in the height of the dust photosphere imply large changes in gas density in a near-Gaussian atmosphere. For example, Figure 15 of [D'Alessio et al. \(2006\)](#) shows that changing the far-infrared SED by less than a factor of 2 changes ϵ_{grain} by a factor of 10.

Table 1.3 lists fitted values of ϵ_{grain} for some transitional disks, drawn from the literature. At best they are accurate to order of magnitude. For our calculations we consider $10^{-3} \leq \epsilon_{\text{grain}} \leq 10^{-1}$ (Table 1.1).

PAHs

For simplicity we model PAHs as spheres, each having a radius $s = 6 \text{ \AA}$ and internal density $\rho_s = 2 \text{ g cm}^{-3}$. Although in reality carbon atoms in PAHs are arranged in sheets and not spheres (e.g., [Allamandola et al. 1999](#)), the difference in cross section arising from geometry is only on the order of unity. Each of our model PAHs has about as much mass as a real PAH containing $N_C = 100$ carbon atoms. A PAH of this size is estimated to be just large enough to survive photo-destruction around Herbig Ae stars ([Visser et al. 2007](#)).

The central wavelengths of PAH emission lines from Herbig Ae/Be (HAe/Be) disks are observed to trend with the effective temperatures of their host stars ([Sloan et al. 2005](#); [Keller et al. 2008](#)). This correlation indicates that PAHs in disks are not merely PAHs from the diffuse ISM transported unadulterated into circumstellar environments. Rather, PAHs in disks have been photo-processed, their chemical bonds altered by radiation from host stars. Possibly PAHs are continuously created and destroyed by local processes, e.g., sublimation of grain mantles and photodestruction ([Keller et al. 2008](#)). In this paper we do not account

explicitly for such processes, i.e., we do not attempt to calculate the abundance of PAHs from first principles. Rather we fix the abundance of PAHs using observations, as detailed in the remainder of this subsection.

Emission from PAHs is detected in an order-unity fraction of HAe/Be stars, but is rarely seen in T Tauri stars (e.g., [Geers et al. 2006](#), hereafter G06). In principle this could mean that PAHs are less abundant in T Tauri disks, but the more likely explanation is that this is an observational selection effect: Herbig Ae/Be stars are more luminous in the ultraviolet (UV) and therefore cause their associated PAHs to fluoresce more strongly (see, e.g., Figure 9 of G06, which shows how the PAH intensity drops below the *Spitzer* detection threshold with decreasing stellar effective temperature). Another clue that PAHs are just as abundant in T Tauri disks as in HAe/Be disks is that those few T Tauri stars with positive PAH detections tend to have unusually low mid-infrared continua, allowing PAH emission lines to stand out more clearly (G06). In other words, those T Tauri systems where PAHs have been detected are transitional systems, and their PAH abundances seem no different than in their HAe/Be counterparts.

[Geers et al. \(2006\)](#) used radiative transfer models to fit the intensities of the $11.2\ \mu\text{m}$ PAH fluorescence line in eight Herbig Ae and T Tauri disks, concluding that the PAH abundance is 10^{-7} – 10^{-8} per H_2 (see their Figure 9). This result is highly model dependent. Perhaps the chief source of uncertainty lies in the grain opacity. Inferred PAH abundances relative to gas are sensitive to assumptions about the local grain size distribution and dust-to-gas ratio because the soft ultraviolet radiation (~ 1000 – $3000\ \text{\AA}$ wavelength) which causes PAHs to fluoresce is also absorbed by ambient grains. Thus the intensity of PAH emission depends on how many grains are competing with PAHs for the same illuminating photons. The grain opacity, in turn, decreases by orders of magnitude as dust settles (Section 1.2.4). Because G06 did not account for dust sedimentation and instead assumed the dust-to-gas ratio in disks was similar to that of the well-mixed ISM, the PAH abundances relative to gas that they computed are actually upper limits. Surface layer grains have likely settled toward the midplane into regions of high gas density, and were the local PAH abundance to remain as inferred at 10^{-7} – 10^{-8} per H_2 , the ratio of PAH line intensity to dust continuum would be larger than observed ([Dullemond et al. 2007](#)).

In our model the number of PAHs per H_2 is

$$x_{\text{PAH}} \equiv \epsilon_{\text{PAH}} \times 10^{-6}, \quad (1.6)$$

where $\epsilon_{\text{PAH}} < 1$ measures how depleted PAHs are in disk surface layers relative to PAHs in the diffuse ISM ([Li & Draine 2001](#)). Based on the considerations above, we should combine the G06 depletion factor of 0.01–0.1 with the grain depletion factor $\epsilon_{\text{grain}} \sim 0.001$ –0.1, which parameterizes the reduction of dust opacity due to dust sedimentation. We thus estimate that $10^{-5} \lesssim \epsilon_{\text{PAH}} \lesssim 10^{-2}$ (Table 1.1). In our calculations we select the parameter combinations $(\epsilon_{\text{grain}}, \epsilon_{\text{PAH}}) = (10^{-1}, 10^{-2})$ and $(\epsilon_{\text{grain}}, \epsilon_{\text{PAH}}) = (10^{-3}, 10^{-5})$ which bracket the range of possibilities.

A final point to consider is whether the observed PAHs are present at the same column depths that are relevant for X-ray driven MRI. The X-ray stopping column should be compared

with the column that presents optical depth unity to the soft UV radiation driving PAH emission. In the model of G06 in which dust has not settled, photons at wavelengths of 1000–3000 Å are stopped by submicron-sized silicate/carbonaceous grains within a hydrogen column of $\sim 0.005 \text{ g cm}^{-2}$ (V. Geers, private communication, 2010; see also [Habart et al. 2004](#) who used similar dust opacities). After we account for grain settling (ϵ_{grain}), the UV absorption column increases to $\sim 0.05\text{--}5 \text{ g cm}^{-2}$. Although model-dependent, our estimate of the UV absorption column corresponds well to X-ray stopping columns, and thus to columns that might possibly be MRI-active.

1.2.5 Collisional Charging of PAHs and Grains

Grains and PAHs are modeled as conducting spheres for simplicity. Electrons and ions collide with and stick to grains and PAHs, charging them. When the total electron capture rate by grains and PAHs matches the total ion capture rate, the distribution of charges carried by PAHs and grains reaches dynamical equilibrium. The average charge state on a PAH/grain $\langle Z \rangle < 0$ because in thermal equilibrium electrons move more quickly than ions. Convenient and readily derived approximations for $\langle Z \rangle$ in various limits were given by [Draine & Sutin \(1987, hereafter DS\)](#). We will find that $\langle Z \rangle$ ranges between -1 and 0 for our PAHs of radius 6 Å , while for our micron-sized grains $\langle Z \rangle \approx -22$. The remainder of this subsection details how we compute the electron and ion capture rates.

The rates at which ions or electrons collide with PAHs or grains are enhanced by Coulomb focusing between static charges, as well as by the induced dipole force ([Natanson 1960; Robertson & Sternovsky 2008](#)). The cross sections can be derived from kinetic theory by considering the potential between a conducting sphere of radius s and charge Ze , located at a distance r from a charge q :

$$\phi(Z, r) = \frac{qZe}{r} - \frac{q^2 s^3}{2r^2(r^2 - s^2)} \quad (1.7)$$

(e.g., [Jackson 1975](#)). The first term is the usual monopole interaction, while the second arises from the induced dipole (image charges). For a neutral sphere, the velocity-dependent cross section derives from applying conservation of energy and momentum to the second term of (1.7). Multiplying this cross section by either the electron or ion velocity, and averaging over a Maxwellian speed distribution at temperature T , yields the rate coefficient (units of $\text{cm}^3 \text{ s}^{-1}$)

$$\alpha = \pi s^2 S c \left(1 + \sqrt{\frac{\pi q^2}{2skT}} \right) \quad \text{for } Ze/q = 0 \quad (1.8)$$

where k is the Boltzmann constant, $c = \sqrt{8kT/\pi m}$ is the mean speed for either electrons of mass $m = m_e$ or ions of mass $m = m_{X+}$, and S is the probability that the electron/ion sticks to the PAH/grain. We will discuss the sticking coefficient S shortly.

For a charged sphere, the cross section is enhanced by both terms in (1.7). There is no analytical solution for this case, but DS provided the following approximate formulae:

$$\alpha = \pi s^2 S c [1 - Ze/(q\tau)] (1 + \sqrt{2/(\tau - 2Ze/q)}) \text{ for } Ze/q < 0 \quad (1.9)$$

$$\alpha = \pi s^2 S c [1 + (4\tau + 3Ze/q)^{-1/2}]^2 \exp(-\beta/\tau) \text{ for } Ze/q > 0 \quad (1.10)$$

where $\tau \equiv skT/q^2$,

$$\beta \equiv \frac{Ze}{qg} - \frac{1}{2g^2(g^2 - 1)}, \quad (1.11)$$

and g is the solution to the transcendental equation

$$\frac{2g^2 - 1}{g(g^2 - 1)^2} = \frac{Ze}{q}. \quad (1.12)$$

Upon colliding with a PAH or grain, the electron or ion sticks with probability S . For ions, we set $S = S_{X^+} = 1$ (DS; IN; BG). For electrons colliding with PAHs, $S = S_e$ depends on the detailed molecular structure of the PAH. Allamandola et al. (1989) calculated how the electron sticking coefficient increases with both the number of carbon atoms and the electron affinity. The dependence on electron affinity is especially strong. A PAH having $N_C = 32$ and an electron affinity of 0.7 eV has $S_e \approx 3 \times 10^{-5}$ (see their Figure 25), while the same-sized PAH with an electron affinity of 1 eV has $S_e \approx 10^{-2}$ (see page 769 of their paper). Estimated electron affinities of real $N_C = 32$ PAHs (e.g., ovalene and hexabenzocoronene) exceed 1 eV. Allamandola et al. (1989) stated that “only for pericondensed PAHs [which are more stable than catacondensed PAHs] containing considerably more than 20 C atoms will the electron sticking coefficient approach unity.” Based on these considerations, we take $S_e = 0.1$ for our $N_C = 100$ PAHs. For the much larger grains we set $S_e = 1$.

In our code, the range of charges a grain can possess extends from $Z = -200$ to $+200$. We have verified that this range is large enough to accommodate the entire equilibrium charge distribution, which for our μm -sized grains peaks at -22 (see Figure 1.4). Accounting only for a few charges—up to $|Z| = 3$ as did Sano et al. (2000), IN, BG, and TCS—is not necessarily adequate for micron-sized grains which have fairly large capacitances. For PAHs we consider charges Z between -16 and $+16$. Most PAHs will turn out to have either $Z = 0$ or -1 . Because of their smaller size, a single PAH will be less charged than a single grain; electrons collide less frequently with a negatively charged sphere as the radius of the sphere decreases and the Coulomb potential steepens.

We neglect adsorption of neutral gas-phase species onto grain surfaces, and any mass increase of grains and PAHs from collisions with ions. Grain-grain and PAH-grain collisions are negligible and ignored. We do account for the possibility that a PAH with a single negative charge can neutralize by colliding with a PAH with a single positive charge (reaction 14 in Table 1.2), though in practice this reaction is not significant.

1.2.6 Numerical Method of Solution

The time-dependent rate equations for the abundances of species are readily constructed from the reactions listed in Table 1.2. For example, the number density of electrons n_e obeys

$$\begin{aligned} \frac{dn_e}{dt} = & n_{\text{H}_2}\zeta - n_e \sum_{\text{X}^+} \alpha_{\text{X}^+,e} n_{\text{X}^+} \\ & - n_e \sum_{Z=-16}^{16} \alpha_{\text{PAH},e} n_{\text{PAH}} \\ & - n_e \sum_{Z=-200}^{200} \alpha_{\text{grain},e} n_{\text{grain}} \end{aligned} \quad (1.13)$$

where the index X^+ runs over reactions 7, 8, and 9 in Table 1.2. The first sum over Z occurs over the charge states of PAHs, while the second sum occurs over the charge states of grains, with rate coefficients α given in Section 1.2.5.

The charge distributions of PAHs and grains are governed by recurrence equations (Parthasarathy 1976; Whitten et al. 2007), e.g., for PAHs:

$$\begin{aligned} \frac{dn_{\text{PAH},Z}}{dt} = & (n_{\text{PAH}} \alpha_{\text{PAH},e})_{Z+1} n_e \\ & + \sum_{\text{X}^+} (n_{\text{PAH}} \alpha_{\text{PAH},\text{X}^+})_{Z-1} n_{\text{X}^+} \\ & - (n_{\text{PAH}} \alpha_{\text{PAH},e})_Z n_e \\ & - \sum_{\text{X}^+} (n_{\text{PAH}} \alpha_{\text{PAH},\text{X}^+})_Z n_{\text{X}^+} . \end{aligned} \quad (1.14)$$

The right-hand side of Equation (1.14) accounts for all the ways in which PAHs of charge Z can be created or destroyed by collisions with electrons and ions (reactions 10 and 11 in Table 1.2). When $Z = \pm 1$, Equation (1.14) is supplemented by an extra loss term accounting for reaction 14.

All rate equations are discretized to first order and advanced simultaneously using a forward Euler algorithm with a fixed timestep $\Delta t \leq 1 \times 10^{-3}$ s.

At $t = 0$, all PAHs and grains have $Z = 0$ and all hydrogen is in the form of H_2 . In principle we could simply advance the network forward until the system equilibrates, i.e., until the time rates of change of the abundances fall below some specified tolerance. However the reaction rates in our network span almost 5 orders of magnitude. Thus, our equations are stiff and a brute-force integration would require an inordinate number of timesteps. Metals are typically the slowest constituent to reach equilibrium because they react with electrons only slowly by radiative recombination (reaction 9).

To circumvent the bottleneck posed by metals, we proceed as follows. We run R versions of the code having R evenly spaced initial abundances for charged metals n_{M^+} . For each run,

we initially set $n_e = n_{M+}$ to ensure charge neutrality. We run each code until the abundances of all species drift only because of slow changes in n_{M+} . We evaluate dn_{M+}/dt at the end of each run. The equilibrium value of n_{M+} is bracketed by the two runs having opposing signs for dn_{M+}/dt . We then start a new iteration with R runs having initial metal abundances evenly spaced between the two bounding runs of the previous iteration. In this way we refine our initial guesses for n_{M+} until we arrive at two sets of initial conditions that differ by less than 30%. The equilibrium value of n_{M+} we report lies at the intersection of the two curves for $n_{M+}(t)$, linearly extrapolated forward in time. Other variables (n_e , n_{HCO+} , and $n_{H_3^+}$) are also extrapolated. The number of runs R at each iteration varies from 2 to 5.

We use the time t_{eq} at which the two extrapolated curves for n_{M+} intersect as an estimator of the equilibration time of the chemical network. For t_{eq} to be a robust estimator, it should be independent of initial conditions. We found that the value of t_{eq} remained constant to within a factor of 3 when initial conditions varied over 2 orders of magnitude. Our values for t_{eq} will be compared to dynamical timescales Ω^{-1} in Section 1.3.3.

Errors are estimated by monitoring conservation of charge and conservation of the total number density of PAHs + grains. Over 10^{10} timesteps, the charge remains constant (at zero) to better than one part in 10^8 , with similar results for the number density of PAHs + grains. As a test of our code, we reproduced the normalized charge distribution on PAHs computed by Jensen & Thomas (1991, see their Figure 2a).

1.3 Results

In Section 1.3.1, we describe how charges distribute themselves on PAHs and grains in dynamical equilibrium. In Section 1.3.2, we explore how the free electron and ion abundances vary with increasing PAH abundance. In Section 1.3.3, we show what all this implies for the degree of magnetic coupling in disk surface layers, interpreting our numerical results whenever possible with simple analytic estimates. In that section we also compute timescales for the chemical network to equilibrate, and compare to the dynamical timescales over which the MRI may act. In Section 1.3.4, we test the validity of our simple network/code by seeing how closely we can reproduce the results of more complex networks/codes by BG and TCS.

1.3.1 Charge Distributions on PAHs and Grains

Figures 1.3 and 1.4 show the charge distributions on PAHs and grains, respectively, for the case $a = 3$ AU, $\Sigma = 0.3 \text{ g cm}^{-2}$, $x_M = 10^{-8}$ (standard metal abundance), $\epsilon_{PAH} = 10^{-5}$ (low PAH abundance), and $\epsilon_{grain} = 10^{-3}$ (low grain abundance). Most of the PAHs either have $Z_{PAH} = 0$ or $Z_{PAH} = -1$. For grains, the average charge state (the peak of the distribution) is $\langle Z_{grain} \rangle \approx -22$. The shape of the charge distribution for grains approaches the Gaussian given by Equation (4.15) of DS.

We may understand $\langle Z \rangle$ simply. Consider the PAHs; identical considerations apply to grains. We take the limit that the dominant ions are metals and the limit that the total

charge carried by PAHs is much less than the free charge. Together these limits imply that $x_e \approx x_{M+}$. Then detailed balance between forward and reverse reaction rates dictates that (cf. Equation 1.14):

$$(n_{\text{PAH}} \alpha_{\text{PAH,e}})_{Z+1} = (n_{\text{PAH}} \alpha_{\text{PAH,M+}})_Z . \quad (1.15)$$

From this equation it is evident that if ever the rate coefficients $(\alpha_{\text{PAH,e}})_{Z+1}$ and $(\alpha_{\text{PAH,M+}})_Z$ were to be equal, the densities $(n_{\text{PAH}})_{Z+1}$ and $(n_{\text{PAH}})_Z$ would be equal, i.e., the charge distribution would be at an extremum. Thus we may estimate the average charge $\langle Z \rangle$ by merely plotting the rate coefficients $\alpha_{\text{PAH,M+}}$ and $\alpha_{\text{PAH,e}}$ against Z and seeing where the curves intersect. This exercise is performed in Figures 1.3 and 1.4. Indeed what the full numerical model gives for $\langle Z \rangle$ is close to the Z for which the curves for the rate coefficients intersect. (Of course, perfect agreement cannot be obtained because it is never strictly true that $(\alpha_{\text{PAH,e}})_{Z+1} = (\alpha_{\text{PAH,M+}})_Z$.)

There is another, even simpler limit where $\langle Z \rangle$ may be estimated. In the extreme case that the gas is so saturated with grains or PAHs that practically no free charges are left, we must have $\langle Z \rangle \rightarrow 0$. Figure 1.5 shows the results of an experiment using our full code in which we increase ϵ_{PAH} until this regime is reached. For this Figure, the grain abundance is set to zero to isolate the effects of PAHs. Figure 1.6 is analogous; ϵ_{grain} is increased while the PAH abundance is held fixed at zero. Both figures follow the transition from $\langle Z \rangle \neq 0$ to $\langle Z \rangle \rightarrow 0$. Observationally inferred values for ϵ_{grain} (see the shaded region of Figure 1.6) are never so high as to cross into the $\langle Z_{\text{grain}} \rangle \rightarrow 0$ regime. By contrast, Figure 1.5 shows that PAHs may be sufficiently abundant in disks that they impact the density of free charges. The critical PAH abundance x_{PAH}^* dividing the $\langle Z_{\text{PAH}} \rangle \neq 0$ limit from the $\langle Z_{\text{PAH}} \rangle \rightarrow 0$ limit is the one for which an electron attaches itself to a PAH as frequently as it recombines with an ion. This critical abundance is discussed further in Section 1.3.2.

1.3.2 Ionization Fraction vs. PAH Abundance

Figure 1.7 plots the fractional electron and ion densities, x_e and x_i , against the PAH abundance x_{PAH} , for $a = 3$ AU, $\Sigma = 0.3$ g cm⁻², $x_M = 10^{-8}$ (standard metal abundance), and $\epsilon_{\text{grain}} = 10^{-3}$ – 10^{-1} (see the figure caption for how ϵ_{grain} is assigned to each ϵ_{PAH}). Figure 1.8 is identical except that it considers the metal-rich case $x_M = 10^{-6}$. The primary ions in both cases are atomic metals and HCO⁺ molecules. At low PAH abundances, charged metal ions are the most abundant. As the number of PAHs is increased, HCO⁺ becomes the dominant ion. See Table 1.3.2 for a precise breakdown of component ion densities for our standard metal abundance case.

According to Figures 1.7 and 1.8, the electron and ion densities are nearly equal and constant with x_{PAH} as long as x_{PAH} is not too large. In going from the standard metal abundance of $x_M = 10^{-8}$ to the metal-rich case of $x_M = 10^{-6}$, the free charge abundance increases by an order of magnitude. Once x_{PAH} exceeds some critical abundance x_{PAH}^* , the electron and ion densities diverge—the ion density is higher, and the balance of negative charges is carried by PAHs. In the limit $x_{\text{PAH}} \gg x_{\text{PAH}}^*$, both the electron and ion densities

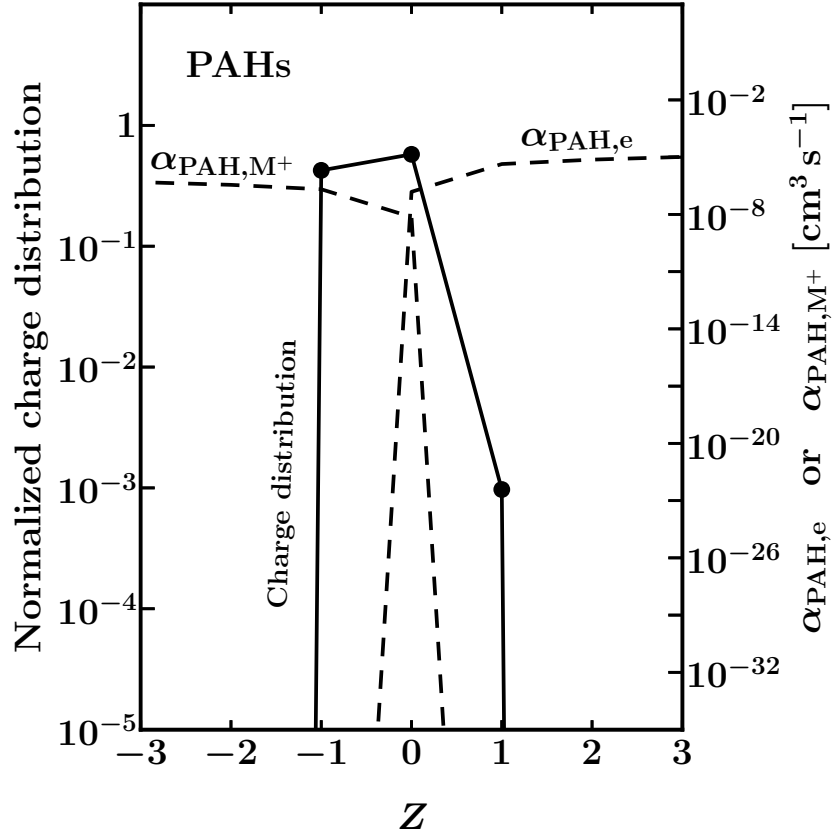


Figure 1.3 Equilibrium charge distribution on PAHs (solid circles, left axis) for $a = 3$ AU, $\Sigma = 0.3 \text{ g cm}^{-2}$, $x_{\text{M}} = 10^{-8}$ (standard metal abundance), $\epsilon_{\text{PAH}} = 10^{-5}$ (low PAH abundance), and $\epsilon_{\text{grain}} = 10^{-3}$ (low grain abundance). The distribution peaks at $Z = 0$, approximately where the attachment coefficients (dashed lines, right axis) for electrons with PAHs and metal ions with PAHs cross.

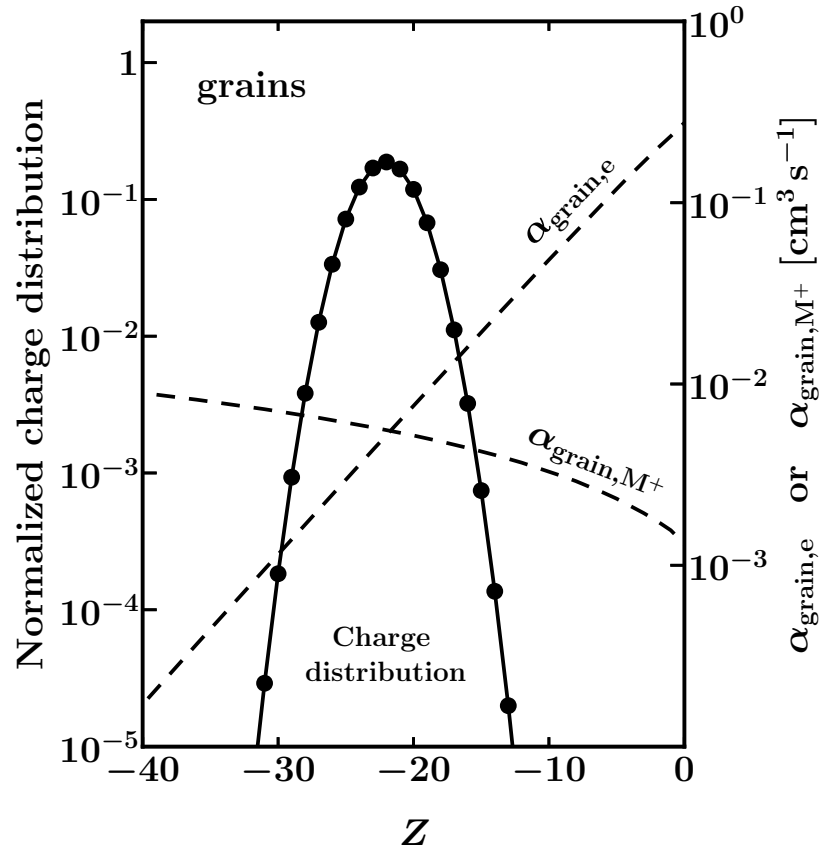


Figure 1.4 Same as Figure 1.3 but for grains.

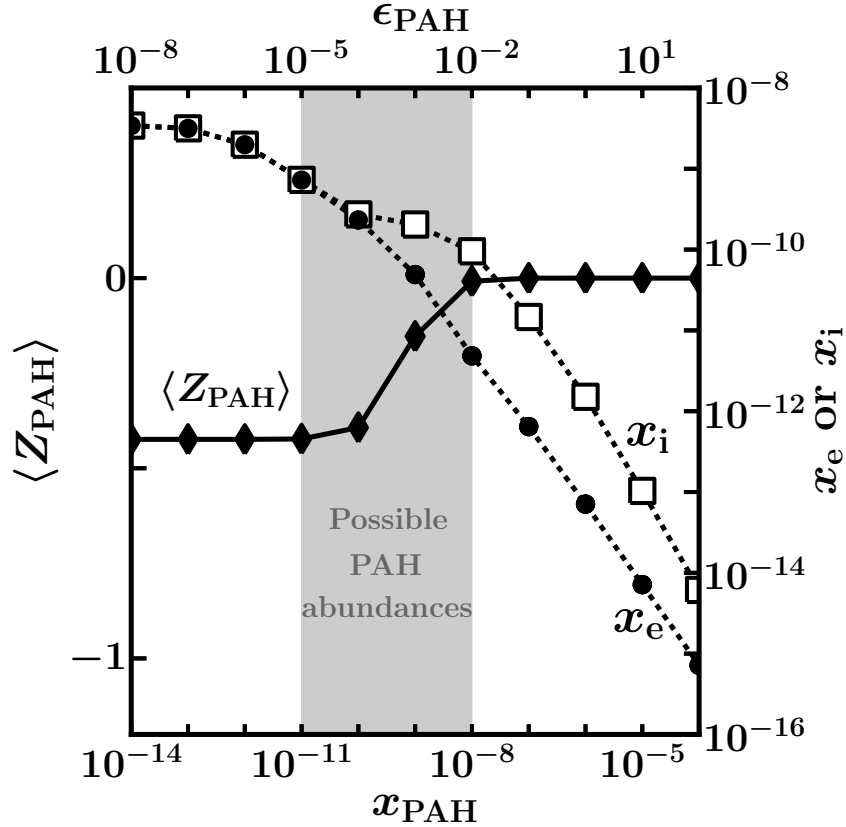


Figure 1.5 Average charge state of PAHs as a function of PAH abundance (solid diamonds, left axis). Dashed lines show simulation results for fractional electron abundance x_e (solid circles, right axis) and fractional ion abundance x_i (open squares, right axis). The shaded region marks observationally inferred PAH abundances, measured by number either relative to H_2 (x_{PAH} , bottom axis) or relative to the PAH abundance in the diffuse ISM (depletion factor $\epsilon_{\text{PAH}} \equiv x_{\text{PAH}}/10^{-6}$, top axis). Parameters for this run are $a = 3$ AU, $x_{\text{M}} = 10^{-8}$, $\Sigma = 0.3 \text{ g cm}^{-2}$, and $\epsilon_{\text{grain}} = 0$ (kept at zero to isolate the effect of PAHs). The shift to $\langle Z_{\text{PAH}} \rangle = 0$ occurs when there are so many PAHs that they begin to adsorb most of the free charge. At this point x_i and x_e diverge; see also Figure 1.7.

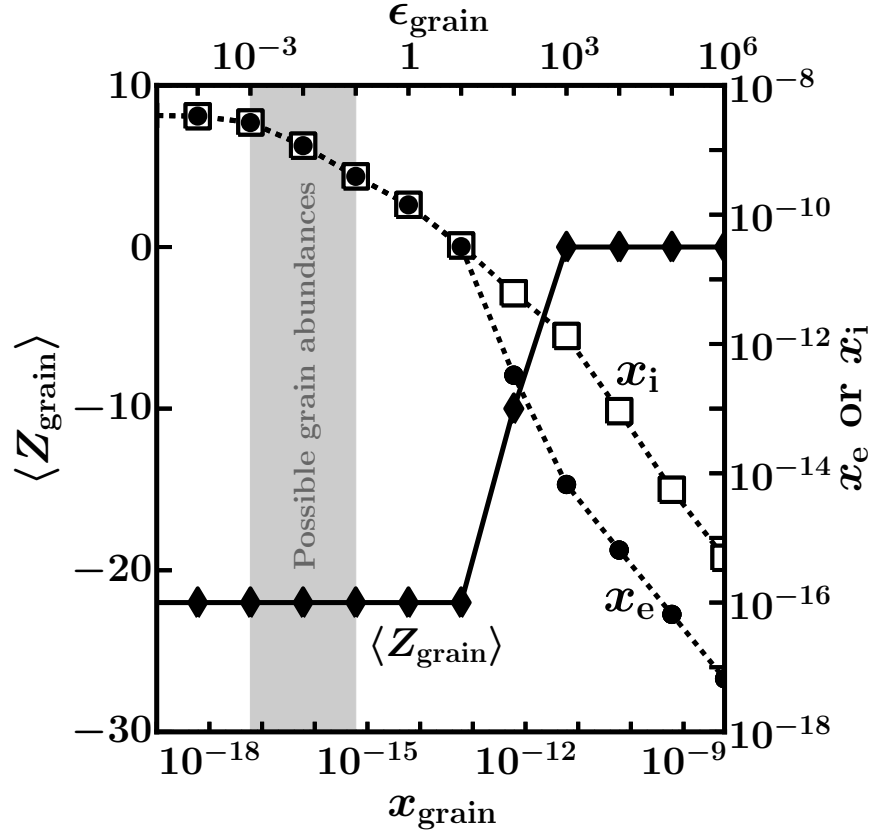


Figure 1.6 Same as Figure 1.5 but for grains. Parameters for this run are $a = 3$ AU, $x_{\text{M}} = 10^{-8}$, $\Sigma = 0.3 \text{ g cm}^{-2}$, and $\epsilon_{\text{PAH}} = 0$ (kept at zero to isolate the effect of grains). Our grains all have radii of $1 \mu\text{m}$, which is so large that their corresponding abundance as inferred from observation (shaded region) is too low to significantly affect the amount of free charge.

decrease with increasing PAH abundance in an approximately inverse linear way. All of this behavior can be understood analytically as follows.

Analytical Model for Ionization Fraction vs. PAH Abundance

Our code's results for $x_e(x_{\text{PAH}})$, $x_i(x_{\text{PAH}})$, and x_{PAH}^* may be understood using the following simple model. The model consists only of X-rays, molecular hydrogen, electrons, PAHs, and one ion species—either HCO^+ for our standard metal abundance case, or ionized metals M^+ for the metal-rich case. In the simplified model, X-ray ionization of a hydrogen molecule produces a free electron and—skipping the entire reaction chain—one ion. The system reduces to the rate equations

$$\frac{dn_e}{dt} = \zeta n_{\text{H}_2} - n_e n_i \alpha_{i,e} - n_e x_{\text{PAH}} n_{\text{H}_2} (\alpha_{\text{PAH},e})_{\langle Z \rangle} \quad (1.16)$$

$$\frac{dn_i}{dt} = \zeta n_{\text{H}_2} - n_i n_e \alpha_{i,e} - n_i x_{\text{PAH}} n_{\text{H}_2} (\alpha_{\text{PAH},i})_{\langle Z \rangle} \quad (1.17)$$

for the electron and ion densities, n_e and n_i . The subscript i denotes either HCO^+ or M^+ .

In the simplified model, all PAHs with abundance $x_{\text{PAH}} n_{\text{H}_2}$ are assumed to be identically charged. We set this common charge equal to the average charge state $\langle Z \rangle$, results for which were given in Section 1.3.1.

We exclude our large, micron-sized grains from the analytic model. Although these grains were useful for inferring PAH abundances from observations (Section 1.2.4), their collective surface area is too low to significantly influence the electron chemistry in any of our model runs. Of course, because PAHs and grains are both modeled the same way, i.e., as spherical conductors, all of the equations below would still be valid were we to replace PAHs with grains.

Standard metal abundance. As stated above, we assume for this case that all ions are HCO^+ molecules and neglect M^+ . We may solve for x_e and $x_i = x_{\text{HCO}^+}$ in the limits of low and high PAH abundance. In the limit of low x_{PAH} , the rightmost terms in Equations (1.16) and (1.17) can be ignored, yielding in steady state:

$$x_e = x_{\text{HCO}^+} = \sqrt{\zeta / n_{\text{H}_2} \alpha_{\text{HCO}^+,e}} \quad (1.18)$$

$$\begin{aligned} &\sim 10^{-10} \left(\frac{L_X}{10^{29} \text{ erg s}^{-1}} \right)^{1/2} \\ &\times \left(\frac{\Sigma}{0.3 \text{ g cm}^{-2}} \right)^{-1/2} \left(\frac{a}{3 \text{ AU}} \right)^{-0.5} \end{aligned} \quad (1.19)$$

for standard metals and low PAHs.

Table 1.4. Densities of Charged Species for our Standard Model
 $(x_{\text{M}} = 10^{-8}, L_{\text{X}} = 10^{29} \text{ erg s}^{-1})$

$a = 3 \text{ AU}, \epsilon_{\text{PAH}} = 10^{-5}, \epsilon_{\text{grain}} = 10^{-3}$											
Σ	n_{H_2}	n_{e}	n_{M^+}	n_{HCO^+}	$n_{\text{H}_3^+}$	x_{e}	x_{i}	x_{PAH}	x_{grain}	$\langle Z_{\text{PAH}} \rangle$	$\langle Z_{\text{grain}} \rangle$
6×10^{-3}	1×10^9	2×10^1	1×10^1	8×10^0	5×10^{-1}	2×10^{-8}	2×10^{-8}	1×10^{-11}	6×10^{-18}	-4×10^{-1}	-22
4×10^{-2}	7×10^9	4×10^1	4×10^1	4×10^0	8×10^{-2}	6×10^{-9}	6×10^{-9}	1×10^{-11}	6×10^{-18}	-4×10^{-1}	-22
3×10^{-1}	5×10^{10}	3×10^1	3×10^1	1×10^0	3×10^{-3}	6×10^{-10}	6×10^{-10}	1×10^{-11}	6×10^{-18}	-4×10^{-1}	-22
2×10^0	4×10^{11}	1×10^1	1×10^1	4×10^{-1}	1×10^{-4}	4×10^{-11}	4×10^{-11}	1×10^{-11}	6×10^{-18}	-4×10^{-1}	-22
1×10^1	2×10^{12}	6×10^{-1}	3×10^0	3×10^{-2}	6×10^{-6}	2×10^{-13}	1×10^{-12}	1×10^{-11}	6×10^{-18}	-1×10^{-1}	-15
6×10^1	1×10^{13}	7×10^{-4}	5×10^{-3}	2×10^{-5}	3×10^{-9}	6×10^{-17}	4×10^{-16}	1×10^{-11}	6×10^{-18}	-4×10^{-5}	-14
$a = 3 \text{ AU}, \epsilon_{\text{PAH}} = 10^{-2}, \epsilon_{\text{grain}} = 10^{-1}$											
Σ	n_{H_2}	n_{e}	n_{M^+}	n_{HCO^+}	$n_{\text{H}_3^+}$	x_{e}	x_{i}	x_{PAH}	x_{grain}	$\langle Z_{\text{PAH}} \rangle$	$\langle Z_{\text{grain}} \rangle$
6×10^{-3}	1×10^9	1×10^1	5×10^{-1}	1×10^1	5×10^{-1}	9×10^{-9}	1×10^{-8}	1×10^{-8}	6×10^{-16}	-4×10^{-1}	-21
4×10^{-2}	7×10^9	5×10^0	1×10^0	2×10^1	8×10^{-2}	7×10^{-10}	2×10^{-9}	1×10^{-8}	6×10^{-16}	-2×10^{-1}	-17
3×10^{-1}	5×10^{10}	2×10^{-1}	7×10^{-1}	3×10^0	3×10^{-3}	4×10^{-12}	8×10^{-11}	1×10^{-8}	6×10^{-16}	-8×10^{-3}	-9
2×10^0	4×10^{11}	1×10^{-2}	7×10^{-2}	2×10^{-1}	1×10^{-4}	4×10^{-14}	8×10^{-13}	1×10^{-8}	6×10^{-16}	-8×10^{-5}	-9
1×10^1	2×10^{12}	6×10^{-4}	4×10^{-3}	1×10^{-2}	6×10^{-6}	3×10^{-16}	6×10^{-15}	1×10^{-8}	6×10^{-16}	-1×10^{-7}	-9
6×10^1	1×10^{13}	3×10^{-7}	2×10^{-6}	5×10^{-6}	3×10^{-9}	2×10^{-20}	6×10^{-19}	1×10^{-8}	6×10^{-16}	-1×10^{-10}	-3×10^{-3}
$a = 30 \text{ AU}, \epsilon_{\text{PAH}} = 10^{-5}, \epsilon_{\text{grain}} = 10^{-3}$											
Σ	n_{H_2}	n_{e}	n_{M^+}	n_{HCO^+}	$n_{\text{H}_3^+}$	x_{e}	x_{i}	x_{PAH}	x_{grain}	$\langle Z_{\text{PAH}} \rangle$	$\langle Z_{\text{grain}} \rangle$
6×10^{-3}	6×10^7	5×10^{-1}	4×10^{-1}	1×10^{-1}	7×10^{-3}	9×10^{-9}	9×10^{-9}	1×10^{-11}	6×10^{-18}	-3×10^{-1}	-9
4×10^{-2}	4×10^8	9×10^{-1}	9×10^{-1}	4×10^{-2}	8×10^{-4}	2×10^{-9}	2×10^{-9}	1×10^{-11}	6×10^{-18}	-3×10^{-1}	-9
3×10^{-1}	3×10^9	4×10^{-1}	4×10^{-1}	2×10^{-2}	2×10^{-5}	2×10^{-10}	2×10^{-10}	1×10^{-11}	6×10^{-18}	-3×10^{-1}	-9
2×10^0	2×10^{10}	1×10^{-1}	2×10^{-1}	6×10^{-3}	1×10^{-6}	8×10^{-12}	1×10^{-11}	1×10^{-11}	6×10^{-18}	-3×10^{-1}	-9
1×10^1	1×10^{11}	7×10^{-3}	5×10^{-2}	4×10^{-4}	7×10^{-8}	7×10^{-14}	4×10^{-13}	1×10^{-11}	6×10^{-18}	-4×10^{-2}	-6
6×10^1	6×10^{11}	3×10^{-6}	5×10^{-5}	2×10^{-7}	3×10^{-11}	5×10^{-18}	1×10^{-16}	1×10^{-11}	6×10^{-18}	-8×10^{-6}	-4
$a = 30 \text{ AU}, \epsilon_{\text{PAH}} = 10^{-2}, \epsilon_{\text{grain}} = 10^{-1}$											
Σ	n_{H_2}	n_{e}	n_{M^+}	n_{HCO^+}	$n_{\text{H}_3^+}$	x_{e}	x_{i}	x_{PAH}	x_{grain}	$\langle Z_{\text{PAH}} \rangle$	$\langle Z_{\text{grain}} \rangle$
6×10^{-3}	6×10^7	2×10^{-1}	9×10^{-3}	3×10^{-1}	7×10^{-3}	3×10^{-9}	5×10^{-9}	1×10^{-8}	6×10^{-16}	-2×10^{-1}	-8
4×10^{-2}	4×10^8	5×10^{-2}	2×10^{-2}	3×10^{-1}	8×10^{-4}	1×10^{-10}	8×10^{-10}	1×10^{-8}	6×10^{-16}	-7×10^{-2}	-6
3×10^{-1}	3×10^9	2×10^{-3}	9×10^{-3}	3×10^{-2}	2×10^{-5}	7×10^{-13}	2×10^{-11}	1×10^{-8}	6×10^{-16}	-2×10^{-3}	-4
2×10^0	2×10^{10}	1×10^{-4}	8×10^{-4}	2×10^{-3}	1×10^{-6}	7×10^{-15}	2×10^{-13}	1×10^{-8}	6×10^{-16}	-2×10^{-5}	-4
1×10^1	1×10^{11}	8×10^{-6}	5×10^{-5}	1×10^{-4}	7×10^{-8}	7×10^{-17}	2×10^{-15}	1×10^{-8}	6×10^{-16}	-1×10^{-9}	-3
6×10^1	6×10^{11}	3×10^{-9}	2×10^{-8}	5×10^{-8}	3×10^{-11}	5×10^{-21}	1×10^{-19}	1×10^{-8}	6×10^{-16}	-1×10^{-10}	-2×10^{-3}

Note. — The surface density Σ has units of g cm^{-2} ; number densities n have units of cm^{-3} ; and fractional densities x are measured per H_2 . The ion density $x_{\text{i}} = x_{\text{M}^+} + x_{\text{HCO}^+}$.

In going from Equations (1.18) to (1.19) we account for the distance dependence of temperature but assume material is optically thin to X-rays. The square-root law of Equation (1.18) is often used by other workers (e.g., Gammie 1996; Glassgold et al. 1997). It is plotted as a horizontal dashed line in Figure 1.7, and should be compared with the curves for x_e and x_i from our code, plotted as solid lines. In the limit of low x_{PAH} , the electron and ion abundances computed from the code are nearly constant with x_{PAH} , as predicted by the analytic model. However, the results from the code sit above the line for Equation (1.18) by about an order of magnitude. The factor of 10 offset arises because Equation (1.18) ignores ionized metals, which recombine with electrons much more slowly than does HCO^+ and which remain abundant compared to HCO^+ in our standard model. The offset also implies that reducing the total metal abundance below that of our standard model ($x_{\text{M}} = 10^{-8}$) can only decrease x_e and x_i by at most a factor of ~ 10 . In this sense our uncertainty in the metal abundance (Section 1.2.4) has only a limited impact on the ionization fraction, assuming $x_{\text{M}} < 10^{-8}$. See also Section 1.3.3 where we consider the case $x_{\text{M}} = 0$.

In the limit of high PAH abundance, electron recombination on PAHs dominates electron recombination with HCO^+ . Low equilibrium abundances of free electrons imply the average charge on PAHs $\langle Z \rangle \rightarrow 0$ (Section 1.3.1). Then the steady-state solutions to Equations (1.16) and (1.17) are, respectively,

$$x_e = \frac{\zeta}{x_{\text{PAH}} n_{\text{H}_2} (\alpha_{\text{PAH},e})_{\langle Z \rangle=0}} \quad (1.20a)$$

$$x_{\text{HCO}^+} = \frac{\zeta}{x_{\text{PAH}} n_{\text{H}_2} (\alpha_{\text{PAH},\text{HCO}^+})_{\langle Z \rangle=0}} \quad (1.20b)$$

for high PAHs.

Note that in this limit of high PAH abundance, $x_e < x_{\text{HCO}^+}$ because $\alpha_{\text{PAH},e} > \alpha_{\text{PAH},\text{HCO}^+}$; electrons move faster than ions. The remaining negative charge required to maintain charge neutrality is carried by PAHs. That $x_e \neq x_{\text{HCO}^+}$ is relevant for the computation of Am and Re (Section 1.3.3) because Am depends on x_{HCO^+} (ions carry the bulk of the momentum in a plasma) while Re depends on x_e (electrons are the most mobile charge carriers). Expressions (1.20a) and (1.20b) are plotted as diagonal dashed lines in Figure 1.7; they compare well with the full numerical results, shown as solid lines, in the limit of high x_{PAH} .

The critical PAH abundance dividing these limits is estimated by equating Equations

(1.20a) to (1.18):

$$\begin{aligned}
 x_{\text{PAH}}^* &= \sqrt{\frac{\zeta \alpha_{\text{HCO}^+, \text{e}}}{n_{\text{H}_2} (\alpha_{\text{PAH}, \text{e}})^2 \langle Z \rangle = 0}} \\
 &\sim 5 \times 10^{-10} \left(\frac{L_{\text{X}}}{10^{29} \text{ erg s}^{-1}} \right)^{1/2} \\
 &\quad \times \left(\frac{\Sigma}{0.3 \text{ g cm}^{-2}} \right)^{-1/2} \left(\frac{a}{3 \text{ AU}} \right)^{-0.2}
 \end{aligned} \tag{1.21}$$

for standard metals.

The critical value x_{PAH}^* marks the abundance at which PAHs start to reduce significantly the number of free charges, i.e., the abundance at which electron recombination on PAHs becomes competitive with electron recombination with molecular ions. It is plotted in Figure 1.7 as a vertical line and does reasonably well at delineating the regime where the ionization fraction does not depend on PAHs from the regime where it does. Note that possible PAH abundances as inferred from observations (Section 1.2.4) happen to straddle x_{PAH}^* .

Metal-rich case. Analogous results are obtained for the metal-rich case as shown in Figure 1.8, with the only difference that M^+ replaces HCO^+ as the dominant ion. In the limit of low PAH abundance, the abundance of free charges is

$$\begin{aligned}
 x_{\text{e}} = x_{\text{M}^+} &= \sqrt{\zeta / n_{\text{H}_2} \alpha_{\text{M}^+, \text{e}}} \\
 &\sim 3 \times 10^{-8} \left(\frac{L_{\text{X}}}{10^{29} \text{ erg s}^{-1}} \right)^{1/2} \\
 &\quad \times \left(\frac{\Sigma}{0.3 \text{ g cm}^{-2}} \right)^{-1/2} \left(\frac{a}{3 \text{ AU}} \right)^{-0.5}
 \end{aligned} \tag{1.22}$$

for high metals and low PAHs.

In going from Equations (1.22) to (1.23) we account for the distance dependence of temperature but assume material is optically thin to stellar X-rays. Just as assuming all ions took the form of HCO^+ in the standard model gave a lower limit (Equation 1.18) for the ionization fraction, assuming that all ions take the form of metals gives an upper limit (Equation 1.22) because fast recombination of electrons with HCO^+ is neglected.

The analogous asymptotic solutions in the high PAH limit are practically unchanged from Equations (1.20a) and (1.20b) because the charging rates of PAHs by HCO^+ and M^+ are similar; the mass of the HCO^+ molecule and that of a metal ion like Mg^+ are similar. The critical PAH abundance at which PAHs begin to reduce the number of free charges is

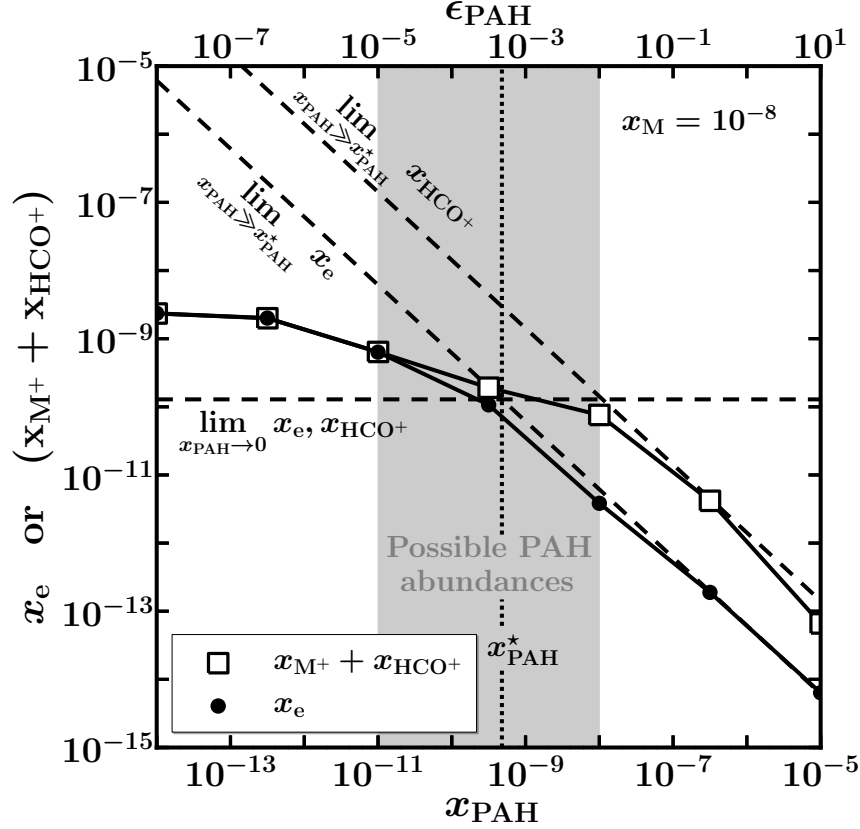


Figure 1.7 Ionization fraction as a function of PAH abundance for $x_M = 10^{-8}$ (standard metal abundance), $a = 3$ AU, and $\Sigma = 0.3$ g cm $^{-2}$. Dashed lines: asymptotic values for x_e and x_{HCO^+} of the simplified model of Section 1.3.2. Solid lines: simulation results for fractional electron abundance x_e (solid circles) and fractional ion abundance $x_{M^+} + x_{HCO^+}$ (open squares). The dotted vertical line marks x_{PAH}^* (Equation 1.21), which roughly divides the regime of “low PAH abundance” where electron and ion densities are equal and insensitive to PAH abundance, from the regime of “high PAH abundance” where the ion density exceeds that of electrons and both decrease approximately as $1/x_{PAH}$. The behavior at high PAH abundance is independent of the metal abundance; compare with Figure 1.8. The shaded region marks observationally inferred PAH abundances and happens to span the transition from low to high PAH regimes. Simulation data use $\epsilon_{grain} = 10^{-3}$ for $\epsilon_{PAH} \leq 10^{-5}$; $\epsilon_{grain} = 10^{-2}$ for $\epsilon_{PAH} = 10^{-3.5}$; and $\epsilon_{grain} = 10^{-1}$ for $\epsilon_{PAH} \geq 10^{-2}$.

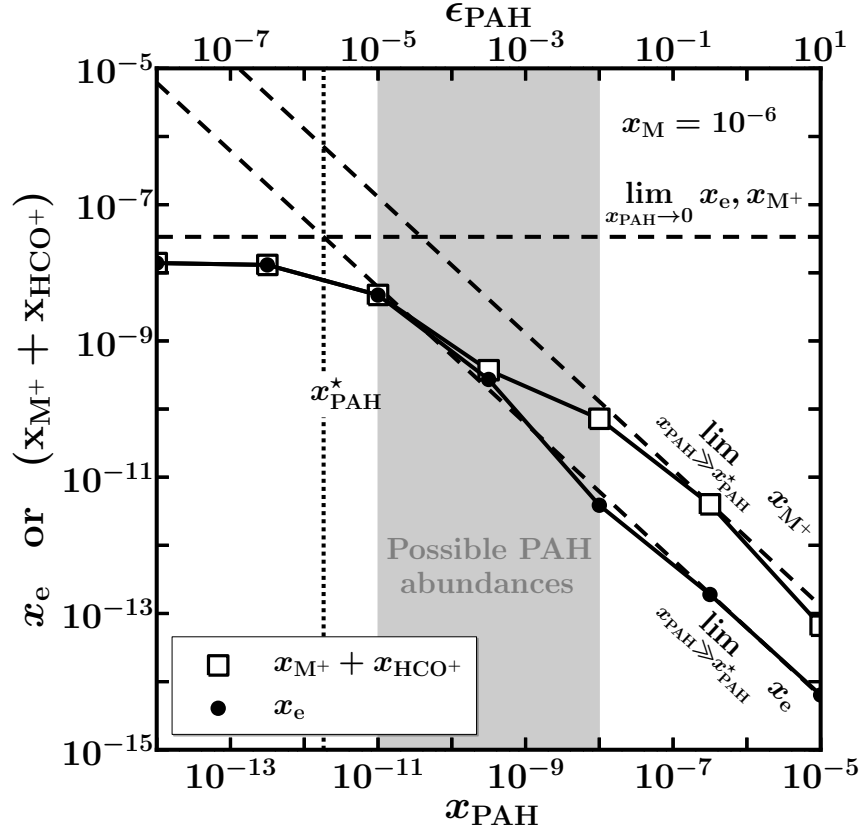


Figure 1.8 Same as Figure 1.7 but for the metal-rich case ($x_M = 10^{-6}$). The curves for x_e and x_i at high PAH abundance ($x_{\text{PAH}} > x_{\text{PAH}}^*$, where x_{PAH}^* is now given by Equation 1.24) are essentially the same as in Figure 1.7: when PAHs dominate charge balance, the metal abundance ceases to matter.

$$\begin{aligned}
 x_{\text{PAH}}^* &= \sqrt{\frac{\zeta \alpha_{M^+,e}}{n_{\text{H}_2} (\alpha_{\text{PAH},e})_{\langle Z \rangle=0}^2}} \\
 &\sim 2 \times 10^{-12} \left(\frac{L_X}{10^{29} \text{ erg s}^{-1}} \right)^{1/2} \\
 &\quad \times \left(\frac{\Sigma}{0.3 \text{ g cm}^{-2}} \right)^{-1/2} \left(\frac{a}{3 \text{ AU}} \right)^{-0.2} \\
 &\quad \text{for high metals}
 \end{aligned} \tag{1.24}$$

and is confirmed by the code.

1.3.3 Degree of Magnetic Coupling: Re and Am

We measure the extent of the MRI-active column by means of the magnetic Reynolds number Re (Equation 1.1) and the ion-neutral collisional frequency Am (Equation 1.2). Figures 1.9 and 1.10 show both dimensionless numbers as a function of the surface density $\Sigma = N\mu$ penetrated by X-rays at $a = 3$ and 30 AU, respectively, over the range of observationally inferred PAH abundances. We overplot for comparison the solution obtained when we omit PAHs completely.

In both Figures 1.9 and 1.10, the middle panels display results for our standard model parameters: $x_M = 10^{-8}$ per H_2 , $L_X = 10^{29}$ erg s $^{-1}$, and $\zeta_{CR} = 0$. In each of the panels on the left and on the right, we vary one of these parameters. We describe here results for our standard model and compare with other test cases in Section 1.3.3.

The middle top panels of Figures 1.9 and 1.10 show that if Re were the only discriminant, MRI-active surface layers could well exist, even with PAHs present. At surface densities $\Sigma \sim 0.3$ g cm $^{-2}$ (column densities $N \sim 10^{23}$ cm $^{-2}$), Re lies comfortably above the critical values of 10^2 – 10^4 (Section 1.1.2) required for plasma to couple to the magnetic field, for a wide range of possible PAH abundances. If the critical $Re \sim 10^2$ (as assumed by BG), and if PAHs are at their lowest possible abundance as inferred from observation ($\epsilon_{PAH} = 10^{-5}$ relative to the ISM; inverted triangles), then the MRI-active layer could extend as far as $\Sigma \sim 20$ g cm $^{-2}$ —if ohmic dissipation were the only limiting factor for the MRI.

But ohmic dissipation is not the only factor. The same margin of safety enjoyed by Re does not at all apply to the ambipolar diffusion number Am , for any surface density. Even in the unrealistic case that there are no PAHs, Am stays < 10 in the middle bottom panels of Figures 1.9 and 1.10. By comparison, values of Am exceeding 10^2 are reported by Hawley & Stone (1998) as necessary for the MRI to excite turbulence in predominantly neutral gas. When PAHs are present, Am barely exceeds 1, and then only for the low end of possible PAH abundances. Compared with ohmic dissipation, ambipolar diffusion seems the much greater concern for the viability of the MRI in disk surface layers.

Values of $Am(\Sigma)$ and $Re(\Sigma)$ vary only slightly as the stellocentric distance increases from $a = 3$ AU (Figure 1.9, middle) to 30 AU (Figure 1.10, middle); the former decreases while the latter increases, each typically by factors of a few. This behavior is readily understood. First recognize that $x_i \propto a^{-0.6}$ approximately; this is an average scaling between the low PAH limit, which implies $x_i \propto a^{-0.5}$ according to Equation (1.18), and the high PAH limit, which implies $x_i \propto a^{-5/7} \approx a^{-0.7}$ according to Equation (1.20b). Combining this result with $n_{H_2} \propto h^{-1} \propto \Omega/T^{1/2} \propto a^{-9/7}$, we find that $Am = x_i n_{H_2} / \Omega \propto a^{-0.4}$. Similarly, $Re = c_s h / D \propto x_e T^{1/2} / \Omega \propto a^{0.7}$.

In computing Am , we have omitted the contribution from collisions between neutral H_2 and negatively charged PAHs. The latter are as well coupled to magnetic fields as molecular ions are—see, e.g., the ion and grain Hall parameters calculated in Section 2.2 of BG. Thus, collisions between H_2 and charged PAHs should increase Am . However, we find that in practice the gain is negligible. We estimate that the collisional rate coefficient β_{in} that enters into Am is about the same for charged PAHs as for ions; in both cases a collision with an H_2

molecule is mediated by the induced dipole in H_2 , and the relative velocity is dominated by the thermal speed of H_2 . At $\Sigma \gtrsim 10 \text{ g cm}^{-2}$, charged PAHs are about as abundant as ions and thus raise Am by a factor of 2—but at these Σ 's, Am is already too low for the MRI to be viable. At $\Sigma \lesssim 10 \text{ g cm}^{-2}$ —i.e., at those columns where Am peaks—charged PAHs are much less abundant than ions and thus hardly affect Am .

Higher L_X and T_X , Higher and Lower x_M , and Cosmic-ray Ionization

In each of the leftmost and rightmost panels of Figures 1.9 and 1.10, we vary one model parameter away from its standard value. We begin with the case of higher L_X . Increasing L_X certainly raises Re and Am , but as the left panels of Figure 1.9 show, even a fairly high $L_X = 10^{31} \text{ erg s}^{-1}$ only causes Am to just exceed 10 at the lowest PAH abundance. In the limit of low PAH abundance, $x_e = x_i \propto L_X^{1/2}$, as predicted by Equation (1.18). In the limit of high PAH abundance, x_e and x_i scale linearly with L_X , according to Equations (1.20a) and (1.20b). Thus the space of possible values of Re and Am narrows with increasing L_X , as the lower envelope increases as L_X while the upper envelope increases as $L_X^{1/2}$.

These same scaling relations, with L_X replaced by the ionization rate ζ at fixed distance, enable us to estimate the effects of a higher T_X , a case we did not explicitly compute using our numerical model. According to IG, raising kT_X from 3 keV (our standard value) to 8 keV increases the ionization rate ζ by factors of 2–4 at $\Sigma = 1\text{--}30 \text{ g cm}^{-2}$. Thus in the extreme case that $L_X = 10^{31} \text{ erg s}^{-1}$ and $kT_X = 8 \text{ keV}$ —parameters appropriate only for a small minority of young stars (Telleschi et al. 2007; Preibisch et al. 2005)—we apply the low-PAH scaling relation $x_i \propto \zeta^{1/2}$ to the bottom left panel of Figure 1.9 to find that the largest possible value of Am is ~ 20 , obtained if PAHs are at their lowest plausible abundance.

The rightmost panels of Figure 1.9 display the case of a higher metal abundance $x_M = 10^{-6}$ per H_2 . As discussed in Section 1.2.4, the higher metal abundance is not especially realistic and is considered primarily as an exercise. Comparing the middle and right panels of Figure 1.9, we see that increasing the metal abundance by a factor of 100 raises Re and Am at low PAH abundance by a factor of ~ 10 . At high PAH abundance, Re and Am also increase with increasing metal abundance, but the gain is less. This same behavior is reflected in the solid curves of Figures 1.7 and 1.8: at low PAH abundance, increasing x_M by a factor of 100 leads to a factor of ~ 10 increase in $x_e = x_i$, but at high PAH abundance, the now divergent curves for x_e and x_i are essentially independent of metal abundance. Equations (1.20a) and (1.20b) from our analytic analysis reflect this insensitivity to metal abundance at high PAH abundance.

At $\Sigma \gtrsim 1 \text{ g cm}^{-2}$, gas temperatures may be so low that all of the metals condense onto grains. The case $x_M = 0$ is shown in the leftmost panels of Figure 1.10. Here $Am \lesssim 0.1$ for $\Sigma \gtrsim 1 \text{ g cm}^{-2}$, and it seems safe to conclude that X-ray driven MRI is unviable under these conditions.

Additional ionization by “sideways cosmic-rays” at $a = 30 \text{ AU}$ is considered in the rightmost panels of Figure 1.10. These cosmic-rays, which we have imagined enter the disk edge-on from the outside, dominate stellar X-rays at large Σ . At $\Sigma \sim 10 \text{ g cm}^{-2}$ —comparable to the

full surface density of the disk at $a = 30$ AU—sideways cosmic-rays raise the maximum value of Am to ~ 2 . We have verified that the gains in Am afforded by cosmic-rays are consistent with our scalings of x_e and x_i with ζ as derived above.

Chemical Equilibration Timescales vs. Dynamical Timescales

In assessing whether disk surface layers are MRI-active, we have relied on the critical value $Am^* \sim 10^2$ reported by HS. As a simplifying assumption, HS held fixed the global (box-integrated) ion abundance in each of their simulations. A fixed ion abundance would apply if the chemical equilibration timescale t_{eq} exceeds the dynamical timescale $t_{\text{dyn}} = \Omega^{-1}$ over which HS’s simulations ran. A fixed ion abundance would also apply if ion recombination occurs predominantly on condensates, regardless of $t_{\text{eq}}/t_{\text{dyn}}$ (e.g., [Mac Low et al. 1995](#)). This last statement follows from our Equation (1.20b), which shows $x_i n_{\text{H}_2}$ does not depend on n_{H_2} in the high condensate limit.

The high condensate limit applies for PAH abundances near the high end of those inferred from observation (Figure 1.7). For this high PAH case we expect the assumption of constant ion abundance, and by extension the results of HS, to hold. For high PAH abundance and our standard X-ray luminosity, $Am < 1$ for all Σ and a (Figure 1.9), and our conclusion that X-ray-driven MRI shuts down everywhere seems safe.

For PAH abundances at the low end of those inferred from observation, Equation (1.20b) for the high condensate limit does not apply. Moreover, as shown in Figure 1.11, $t_{\text{eq}}/t_{\text{dyn}} < 1$ —but only by a factor of 10 at most for the low PAH case and for $\Sigma \lesssim 10 \text{ g cm}^{-2}$. Because $t_{\text{eq}}/t_{\text{dyn}} \gtrsim 0.1$ under these conditions, the assumption of constant ion abundance, although not strictly valid, might still be good enough that the results of HS hold to order unity. But even if they do not, we would argue that the sign of any correction for a dynamically variable ion abundance would only hurt the prospects for MRI turbulence. In the simulations of HS—see also [Brandenburg & Zweibel \(1994\)](#) and [Mac Low et al. \(1995\)](#)—ions became concentrated in thin filaments within magnetic nulls. Recombination rates inside the dense filaments were higher than those outside. Were such simulations to account for ion recombination, lower ion densities within the filaments would result, and neutrals would be even less coupled to ions. Consequently, Am^* would be even higher than the reported value of $\sim 10^2$.

1.3.4 Comparison with Previous Work: Ionization Fractions

The ionization chemistry in disks remains inherently uncertain, with rate coefficients for many reactions in the UMIST database determined to no better than factors of ~ 3 ([Vasyunin et al. 2008](#)). A measure of the uncertainty in the ionization fraction in disks is given by the differences between the simple and complex networks computed by BG, which amount to factors of 2–10 for the electron fraction x_e . We should reproduce their results by at least this margin, as a validation of our code. In the following we directly compare our results to those of BG and TCS, adjusting the input parameters of our code to match theirs. Once we match these input parameters, any difference in our codes’ outputs should result primarily from our

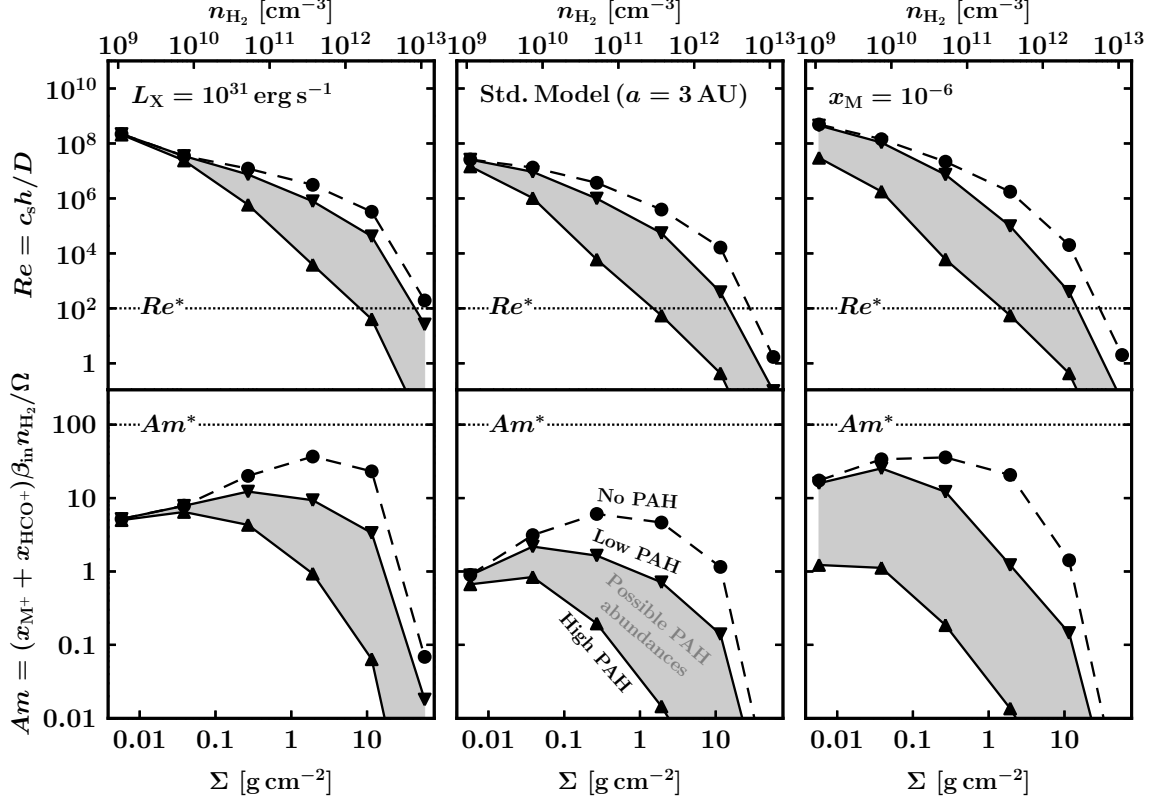


Figure 1.9 Magnetic Reynolds number Re and ambipolar diffusion number Am as a function of surface density Σ at $a = 3$ AU. The middle panels show results for our standard model ($x_M = 10^{-8}$, $L_X = 10^{29}$ erg s $^{-1}$, $\zeta_{CR} = 0$). The side panels have the same parameters as our standard model, except for a $100\times$ more luminous X-ray source (left panels), and a $100\times$ greater metal abundance (right panels). Oppositely pointing triangles bracket values for Am and Re corresponding to possible PAH abundances. These abundances were inferred in Section 1.2.4 from observations. The dashed curve refers to the case with no PAHs and is shown for comparison only. Polycyclic aromatic hydrocarbons reduce ionization fractions and thus the degree of magnetic coupling by an order of magnitude or more. The dotted lines mark the critical values Am^* and Re^* above which coupling between magnetic fields and neutral gas is sufficient to drive the MRI (Section 1.1.2). The curves for Am first rise as Σ increases—a consequence of the increasing number density—and then fall as the ion fraction decreases, never reaching Am^* . Ambipolar diffusion threatens the MRI more than ohmic dissipation does. If the critical Am required for good collisional coupling between ions and neutrals is $Am^* = 10^2$, as evidenced in simulations by Hawley & Stone (1998), then even a disk without PAHs cannot sustain X-ray-driven MRI at any Σ .

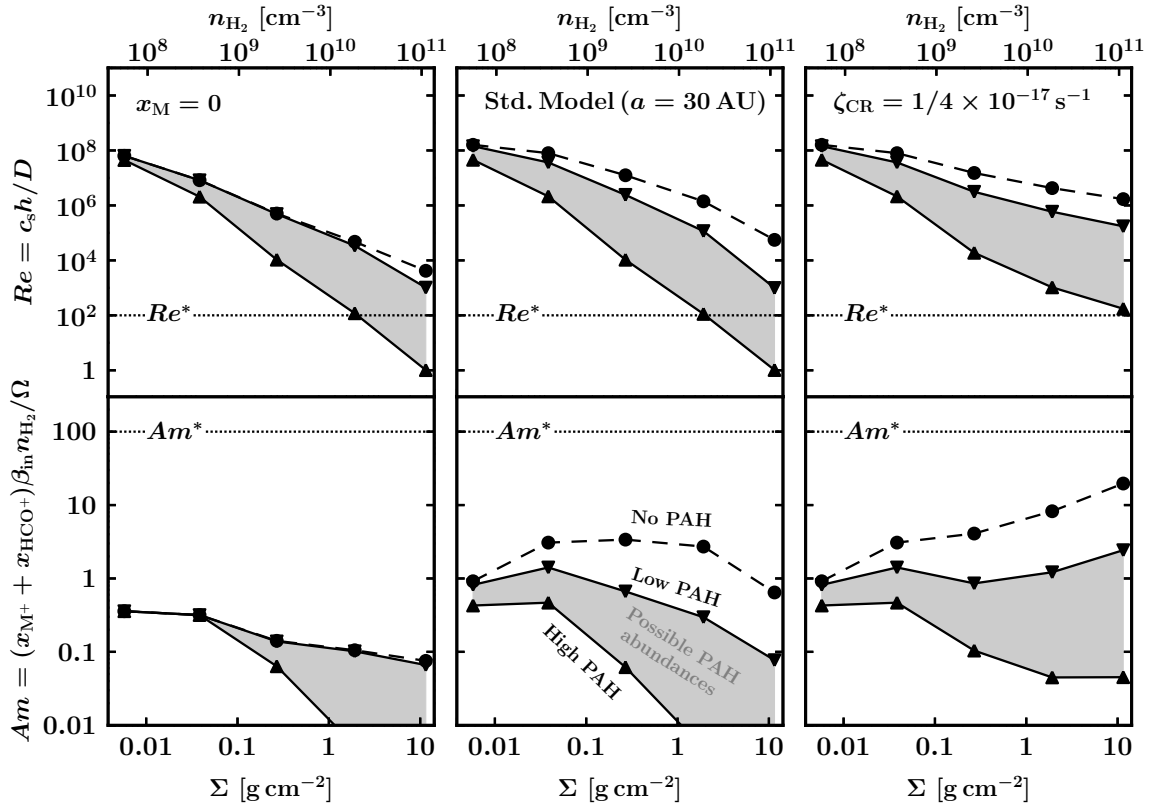


Figure 1.10 Same as Figure 1.9, but at a stellocentric distance of $a = 30$ AU. The middle panels show results for our standard model ($x_M = 10^{-8}$, $L_X = 10^{29}$ erg s $^{-1}$, $\zeta_{CR} = 0$). The side panels have the same parameters as our standard model, except that gas-phase metals are omitted in the left panels ($x_M = 0$), and sideways cosmic-rays are added in the right panels ($\zeta_{CR} = 1/4 \times 10^{-17}$ s $^{-1}$).

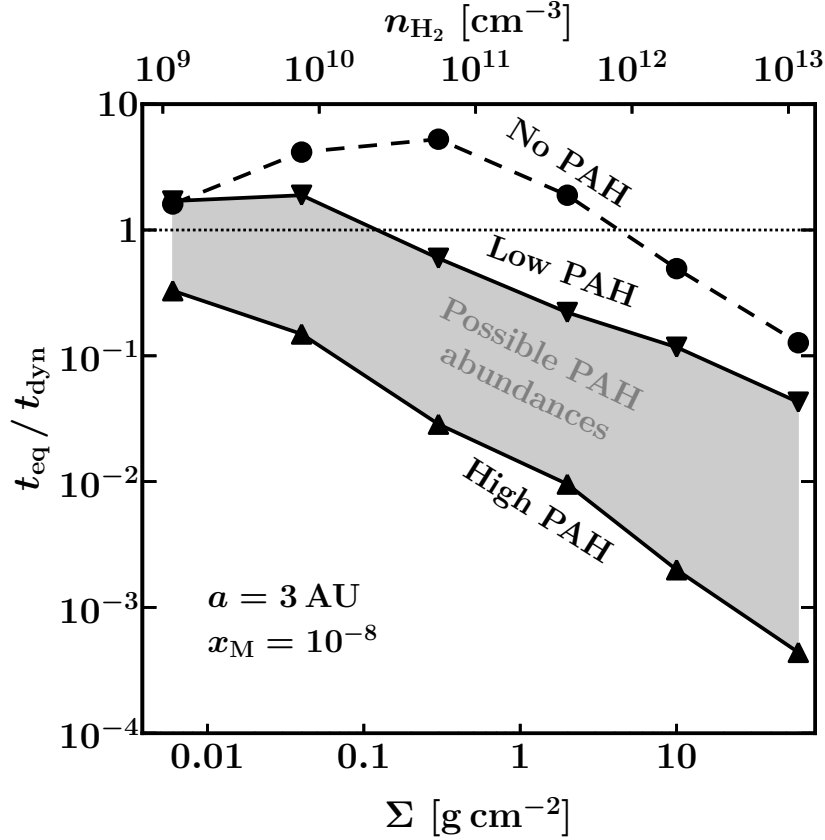


Figure 1.11 Ratio of the chemical equilibration timescale t_{eq} , computed according to the procedure described in Section 1.2.6, to the dynamical time $t_{\text{dyn}} = \Omega^{-1}$, for $a = 3 \text{ AU}$ and $x_{\text{M}} = 10^{-8}$. Values of $t_{\text{eq}}/t_{\text{dyn}}$ at $a = 30 \text{ AU}$ are typically lower than those shown here by factors of 3 or less. Simulations by HS assumed a constant (volume-integrated) abundance of ions, a condition satisfied if $t_{\text{eq}}/t_{\text{dyn}} > 1$. The ion abundance is also constant if ion recombination occurs primarily on condensates (see Equation 1.20b), a situation that obtains for PAH abundances near the high end of those inferred from observations. For the lowest possible PAH abundances, $0.1 < t_{\text{eq}}/t_{\text{dyn}} < 1$ at $\Sigma = 0.1\text{--}10 \text{ g cm}^{-2}$. In this low PAH case, the results of HS might still be expected to apply to order unity. Even if they do not, we argue in the main text that when $t_{\text{eq}}/t_{\text{dyn}} < 1$ the effects of ambipolar diffusion should be even stronger than reported by HS.

different chemical networks (ours is the simplest of the three), and not from differences in radiative transfer, as all our codes rely on the ionization rates calculated by Igea & Glassgold (1999).

We start with BG by computing x_e as a function of density n_{H_2} at a fixed ionization rate $\zeta = 10^{-17} \text{ s}^{-1}$, following their Figure 3. We reset $T = 280 \text{ K}$, $x_{\text{M}} = 2.5 \times 10^{-8}$ per H_2 , and the electron-grain sticking coefficient $S_e = 0.03$ to match their standard parameters. To compare to their “grain-free” case, we run our code without any PAHs or grains. To compare to their standard monodispersion of grains, we run our code with a single population of grains having $s = 0.1 \mu\text{m}$, internal density $\rho_s = 3 \text{ g cm}^{-3}$, and a mass fraction of 1% relative to gas. Figure 1.12 shows the comparison. Our results for the condensate-free case track those of BG, but are higher by factors of 3–10 depending on whether the comparison is made with their simple or complex network. For the case with grains, the agreement with the simple model is excellent and that with the complex model is good to a factor of 2.

In Figure 1.13, we make a similar comparison with TCS, computing x_e as a function of N at a distance of $a = 5 \text{ AU}$ from an X-ray source of $L_X = 2 \times 10^{30} \text{ erg s}^{-1}$ and $kT_X = 5 \text{ keV}$, for $T = 125 \text{ K}$ and a metal abundance of $x_{\text{M}} = 6.8 \times 10^{-7}$ per H_2 . We consider the two cases of their Figure 1, one without any grains or PAHs, and another with a single population of grains having $s = 1 \mu\text{m}$, $\rho_s = 5 \text{ g cm}^{-3}$, and a mass fraction of 1%. For both cases our computed electron abundances are higher, but only by factors of 2 or less.

These comparisons with BG and TCS give us confidence that we have computed ionization fractions about as well as they did. Where our ionization fractions differ, ours are often higher. Our higher values will only bolster the conclusion we make in Section 1.4 that thicknesses of X-ray-ionized MRI-active surface layers have been overestimated by them and others.

1.4 Summary and Discussion

In Section 1.1, we presented the evidence that holes and gaps of transitional disks are cleared by companions to their host stars. Residing within the hole, these companions could either be stars—already observed in about half of all transitional systems—or multi-planet systems. A single Jupiter-mass planet on a circular orbit carves out too narrow a gap to explain the large cavities inferred from observations. But multiple planets can shuttle gas quickly from one planet to the next, all the way down to the central star. Surface densities fall in inverse proportion to radial infall speeds, and radial infall speeds can approach freefall speeds for sufficiently many and massive planets. In this way, multi-planet systems might help to clear holes and simultaneously sustain stellar accretion rates that approach those in disks without holes. The more eccentric the planets’ orbits, the fewer of them may be required to explain a given hole size. Accretion in the presence of multi-planet systems has not received much attention and seems an interesting area for future simulation (e.g., Zhu et al. 2010, submitted).

Stellar or planetary companions regulate accretion velocities v but do not give rise to mass accretion rates $\dot{M}$ in the first place. A planet orbiting just inside the circumference

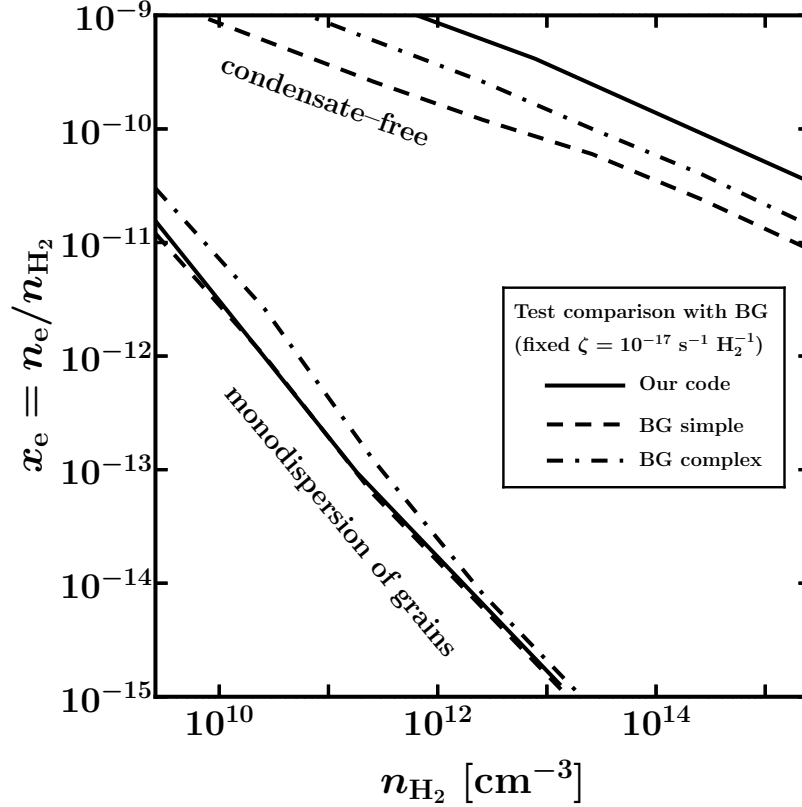


Figure 1.12 Test comparison with BG: electron abundance as a function of gas density at a fixed ionization rate of $\zeta = 10^{-17} \text{ s}^{-1} \text{ H}_2^{-1}$. The upper set of lines are for a condensate-free system, and the lower set are for a monodispersion of grains. Results from BG were drawn from their Figure 3. To generate our results, the parameters of our code were reset to those of BG: temperature $T = 280 \text{ K}$, metal abundance $x_{\text{M}} = 2.5 \times 10^{-8}$ per H_2 , electron-grain sticking coefficient $S_{\text{e}} = 0.03$, grain radius $s = 0.1 \mu\text{m}$, internal grain density $\rho_s = 3 \text{ g cm}^{-3}$, and a mass fraction in grains relative to gas of 1%.

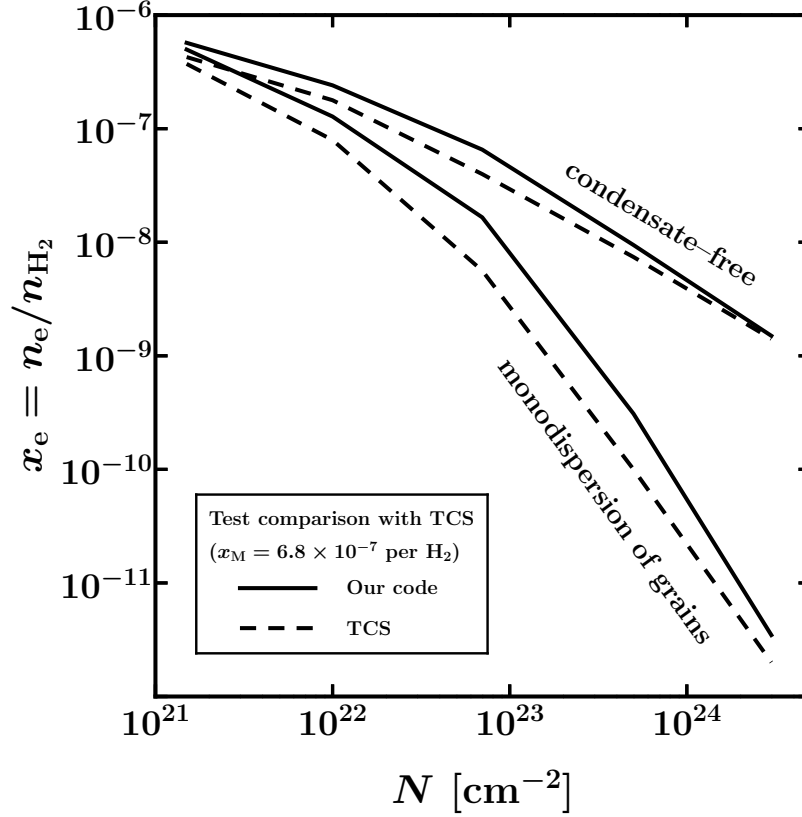


Figure 1.13 Test comparison with TCS: electron abundance as a function of column depth penetrated by X-rays. The upper pair of lines are for a condensate-free system, and the lower pair are for a monodispersion of grains. Ionization rates are from IG. Results from TCS are taken from their Figure 1. To generate our results, we reset the parameters of our code to match those of TCS: stellocentric distance $a = 5$ AU, X-ray luminosity $L_X = 2 \times 10^{30}$ erg s $^{-1}$, temperature $T = 125$ K, metal (magnesium) abundance $x_M = 6.8 \times 10^{-7}$ per H $_2$, electron-grain sticking coefficient $S_e = 0.03$, grain radius $s = 1 \mu\text{m}$, internal grain density $\rho_s = 5$ g cm $^{-3}$, and a mass fraction in grains relative to gas of 1%.

of a disk hole exerts torques to repel gas in the hole’s rim away from the star. Thus, the shepherding planet may reduce $\dot{M}$ —and indeed accretion rates in transitional systems tend to be smaller than those in conventional disks (Najita et al. 2007)—but the planet does not initiate disk accretion. A separate mechanism must act to pull or diffuse gas inward from the hole rim to supply the stellar accretion rates that are observed. That mechanism may be turbulence driven by the magnetorotational instability (MRI), activated by stellar radiation ionizing rim gas. Whether the MRI can operate depends on how well ionized the gas is. The greater the free electron fraction, the greater the magnetic Reynolds number Re , and the less ohmic dissipation dampens the MRI. The greater the atomic and molecular ion densities, the greater the collisional rate Am between neutral particles and ions, and the less ambipolar diffusion weakens the MRI.

A principal threat to the MRI is posed by dust grains, which adsorb electrons and ions. The smallest grains may present the biggest danger, because in many particle size distributions the smallest grains have the greatest surface area for attachment. The smallest grains that can also be detected observationally are polycyclic aromatic hydrocarbons (PAHs), each several angstroms across and containing of order a hundred carbon atoms. These macromolecules may reside in the very disk surface layers that promise to be MRI-active. Excited by soft ultraviolet radiation, PAHs fluoresce in a distinctive set of infrared emission lines detectable from *Spitzer* and from the ground. The hydrocarbon molecules are probably generated locally, photo-sputtered off larger particles exposed to hard UV and X-ray radiation from host stars.

To assess the impact of PAHs on the MRI, we need to know PAH abundances relative to gas. These can be inferred from observed PAH emission lines. Unfortunately such inferences are model dependent; they depend on knowing the local grain opacity, because the soft UV radiation that excites PAHs is also absorbed by grains. In other words, observed PAH line intensities depend on PAH-to-dust ratios. It follows that the quantity of interest to us—the PAH-to-gas ratio—depends on knowing the dust-to-gas ratio. The latter can vary widely with the degree to which grains settle toward disk midplanes. The more grains have settled, the lower are local dust-to-gas ratios, and the lower the PAH-to-gas abundance that is needed to explain a given set of PAH emission spectra. By compiling a few lines of model-dependent evidence from the literature, we estimated that disk PAHs have abundances anywhere from 10^{-11} – 10^{-8} per H_2 , with lower values corresponding to a greater degree of dust settling.

Such PAH abundances, although depleted relative to the ISM by 10^{-5} – 10^{-2} , are still large enough to significantly weaken the MRI in disk surface layers. In fact, they might even shut off X-ray-driven MRI altogether, everywhere. For stellar X-ray luminosities of $L_X = 10^{29}$ – 10^{31} erg s $^{-1}$ and X-ray stopping columns of $\Sigma \sim 1$ – 10 g cm $^{-2}$, PAHs reduce electron and ion densities—which are not equal when PAHs are present—by factors of ~ 10 or more. At these surface densities, the collisional coupling frequency $Am \approx 10^{-3}$ – 10 , depending on PAH abundance and X-ray luminosity. These values fall short, by 1–5 orders of magnitude, of the critical value $Am^* \sim 10^2$ required for good coupling between ions and neutrals, as measured in simulations by Hawley & Stone (1998, HS). The potentially catastrophic effect that small grains can have on the MRI was highlighted by Bai & Goodman (2009, BG). Our study grounds their concern in real-life observations.

Other studies reported X-ray-driven MRI-active surface layers to be alive and well (e.g., Chiang & Murray-Clay 2007, CMC; Turner, Carballido, & Sano 2010, TCS; and BG, in many of whose models the active layer extended to $\sim 1 \text{ g cm}^{-2}$, even with grains present). We should understand why our conclusions differ from theirs. In part, the difference arises because previous studies neglected PAHs. A further difference with TCS is that they assumed a metal abundance of $x_{\text{M}} = 6.8 \times 10^{-7}$ per H_2 , nearly 2 orders of magnitude higher than our standard model value, and one that we find difficult to justify. Still another difference, as significant as any of the ones just mentioned, is the criterion used for whether ambipolar diffusion defeats the MRI. Turner et al. (2010) assumed $Am^* \sim 1$ (see their Equation 9). Bai & Goodman (2009) did not present results for Am . Using their data, we computed the Am values characterizing their claimed active layers. At $a \gtrsim 1 \text{ AU}$, the Am values of BG’s grain-free active layer are at most on the order of unity. For BG’s standard models containing grains, $Am \approx 0.0004\text{--}0.4$, with the lowest value corresponding to $a = 50 \text{ AU}$ and a population of grains having two sizes, and the highest value corresponding to $a = 1 \text{ AU}$ and a single-sized grain population (we computed both limits using results from their complex chemical network). Hawley & Stone (1998) showed that when $Am \lesssim 0.01$, ions and neutrals were effectively decoupled. Even when $Am \sim 1$, HS showed that the MRI saturation amplitude scaled with the ion and not the neutral density, with the neutrals acting to damp out MRI turbulence in the ions. If the MRI drives turbulence only in the ions of protoplanetary disks, it might as well not operate at all, given how overwhelmingly neutral such disks are.

1.4.1 Future Directions

We have shown in this paper that the MRI cannot drive surface layer accretion under typical circumstances in protoplanetary disks, either transitional or conventional, if the critical $Am^* \sim 10^2$ and if stellar X-rays and Galactic cosmic-rays are the dominant source of ionization. These two “if”s are subject to further investigation. We discuss each in turn.

We are not aware of more modern estimates of Am^* apart from that given by HS. As these authors cautioned, numerical resolution is a greater concern for two-fluid simulations than for single-fluid ones, and HS did not demonstrate convergence of their results with resolution. In addition, the value of Am^* was not as precisely determined by HS for toroidal field geometries as for vertical ones—although $Am^* \sim 10^2$ did seem to apply equally well to the cases of uniform vertical field and zero net vertical field. Perhaps higher resolution simulations will reveal that $Am^* < 10^2$ —although accounting for ion recombination in these simulations should only increase Am^* (Section 1.3.3).

The second possibility is that our model has neglected a significant source of ionization. Stellar radiation just longward of the Lyman limit—so-called far ultraviolet (FUV) radiation at photon energies between ~ 6 and 13.6 eV —can ionize trace species such as C, S, Mg, Si, and Al (e.g., Cruddace et al. 1974). Of these elements, C and S may be the most important, as they are the most abundant and possibly least depleted onto grains. For example, if the full solar abundance of C were singly ionized, the ion fraction x_{i} would be a few $\times 10^{-4}$, or

5 orders of magnitude higher than the largest values of x_i reported in this paper! At disk midplanes which are shielded from photodissociating radiation, an order-unity fraction of the full solar abundance of C is expected to take the form of CO ($x_{\text{CO}} = 10^{-4}$; Aikawa et al. 1996). As computed in chemical models by Gorti & Hollenbach (2004; see also Tielens & Hollenbach 1985a and Kaufman et al. 1999), CO near disk surfaces photodissociates nearly entirely by FUV radiation into a layer of neutral C. At the highest altitudes, nearly all of this carbon is photoionized by FUV radiation. The column density of C^+ depends on how many small grains having sizes $\lesssim 0.1\mu\text{m}$ are present, as grains compete to absorb the same FUV photons that photodissociate CO and photoionize C.

We may estimate maximum FUV-ionized column densities by neglecting such dust extinction, and by neglecting shielding of FUV radiation by molecular hydrogen. Consider a trace species T whose total number density regardless of ionization state is $f_{\text{T}}n_{\text{H}_2}$. Take all of T to be singly ionized within a Strömgren slab at the disk surface: $n_{\text{T}^+} = n_{\text{e}} = f_{\text{T}}n_{\text{H}_2}$. Per unit surface area of slab, the rate of photoionizations balances the rate of radiative recombinations:

$$\begin{aligned} \frac{L_{\text{FUV}}}{E_{\text{FUV}}4\pi a^2} &\sim n_{\text{T}^+}n_{\text{e}}\alpha_{\text{T}^+,e}h \\ &\sim f_{\text{T}}^2n_{\text{H}_2}^2\alpha_{\text{T}^+,e}h, \end{aligned} \quad (1.25)$$

where the FUV luminosity capable of ionizing T is $L_{\text{FUV}} \sim 10^{30} \text{ erg s}^{-1}$ (Gorti et al. 2009), the photon energy $E_{\text{FUV}} \sim 10 \text{ eV}$, the rate coefficient $\alpha_{\text{T}^+,e} \sim 4 \times 10^{-12} \text{ cm}^3 \text{ s}^{-1}$ at an FUV-heated gas temperature of 300 K, and the slab thickness $h \sim 0.1a$. Solve for the hydrogen column

$$\begin{aligned} N_{\text{FUV}} &= n_{\text{H}_2}h \\ &\sim 2 \times 10^{22} \left(\frac{L_{\text{FUV}}}{10^{30} \text{ erg s}^{-1}} \right)^{1/2} \\ &\quad \times \left(\frac{3 \text{ AU}}{a} \right)^{1/2} \left(\frac{10^{-4}}{f_{\text{T}}} \right) \text{ cm}^{-2}, \end{aligned} \quad (1.26)$$

or equivalently

$$\begin{aligned} \Sigma_{\text{FUV}} &= N_{\text{FUV}}\mu \\ &\sim 0.07 \left(\frac{L_{\text{FUV}}}{10^{30} \text{ erg s}^{-1}} \right)^{1/2} \\ &\quad \times \left(\frac{3 \text{ AU}}{a} \right)^{1/2} \left(\frac{10^{-4}}{f_{\text{T}}} \right) \text{ g cm}^{-2}. \end{aligned} \quad (1.27)$$

In Equation (1.27), we have normalized f_{T} to its highest plausible value, appropriate for C. An MRI-active surface density $\Sigma_{\text{FUV}} \sim 0.07 \text{ g cm}^{-2}$ is modest, and would drive mass accretion rates only barely observable. Lowering f_{T} would increase Σ_{FUV} . But accounting for extinction of FUV radiation by dust and molecular hydrogen would decrease Σ_{FUV} . We are currently undertaking a more careful study of FUV ionization to quantify these effects.

Chapter 2

Surface Layer Accretion in Conventional and Transitional Disks Driven by Far-Ultraviolet Ionization

Daniel Perez-Becker & Eugene Chiang (2011), ApJ, 735:8.¹

Abstract

Whether protoplanetary disks accrete at observationally significant rates by the magnetorotational instability (MRI) depends on how well ionized they are. Disk surface layers ionized by stellar X-rays are susceptible to charge neutralization by small condensates, ranging from $\sim 0.01\text{-}\mu\text{m}$ -sized grains to angstrom-sized polycyclic aromatic hydrocarbons (PAHs). Ion densities in X-ray-irradiated surfaces are so low that ambipolar diffusion weakens the MRI. Here we show that ionization by stellar far-ultraviolet (FUV) radiation enables full-blown MRI turbulence in disk surface layers. Far-UV ionization of atomic carbon and sulfur produces a plasma so dense that it is immune to ion recombination on grains and PAHs. The FUV-ionized layer, of thickness $0.01\text{--}0.1\text{ g/cm}^2$, behaves in the ideal magnetohydrodynamic limit and can accrete at observationally significant rates at radii $\gtrsim 1\text{--}10\text{ AU}$. Surface layer accretion driven by FUV ionization can reproduce the trend of increasing accretion rate with increasing hole size seen in transitional disks. At radii $\lesssim 1\text{--}10\text{ AU}$, FUV-ionized surface layers cannot sustain the accretion rates generated at larger distance, and unless turbulent mixing of plasma can thicken the MRI-active layer, an additional means of transport is needed. In the case of transitional disks, it could be provided by planets.

¹The Astrophysical Journal, 735:8, 2011 July 1
 © 2011. The American Astronomical Society. All rights reserved.

2.1 Introduction

T Tauri stars, Herbig Ae/Be stars, and young brown dwarfs accrete (e.g., [Hartmann et al. 2006](#)). Perhaps the most direct evidence for accretion comes from observations of radiation at blue to ultraviolet wavelengths, emitted in excess of the Wien tail of the stellar photosphere. The ultraviolet excess arises in large part from gas that is shock heated upon free-falling onto the stellar surface ([Calvet & Gullbring 1998](#); [Johns-Krull et al. 2000](#); [Calvet et al. 2004](#)). From these observations, and others of optical emission lines (e.g., [Hartmann et al. 1994](#); [Herczeg & Hillenbrand 2008](#)), it is inferred that young solar mass stars accrete at rates $\dot{M} \approx 10^{-9}$ – $10^{-7} M_{\odot} \text{ yr}^{-1}$ (e.g., [Muzerolle et al. 2005](#)). Any individual system may exhibit order-unity variations in $\dot{M}$ with time ([Eisner et al. 2010](#)).

Strictly speaking, these observations imply only that gas in the immediate stellar vicinity—i.e., within a few stellar radii, in and around the star’s magnetosphere—is accreting. Gas and dust orbit at much greater stellocentric distances, residing in disks that in many cases extend continuously from ~ 0.1 to ~ 100 AU (e.g., [Watson et al. 2007b](#)), but there is no direct indication that any of this material is actually accreting (but see [Hughes et al. 2011](#) for an attempt at measuring non-Keplerian motions of circumstellar gas). Nevertheless the standard assumption is that infalling gas near the star was transported there through the disk, traveling decades in radius over the system age of $t_{\text{age}} \sim 10^6$ yr. This assumption is at least consistent with the fact that disk masses at large radii—as inferred from continuum dust emission at radio wavelengths (e.g., [Andrews & Williams 2005](#))—are similar to $\dot{M} t_{\text{age}}$.

One of the most intensively studied mechanisms for transporting material is turbulence driven by the magnetorotational instability (MRI; for a review, see [Balbus 2009](#)). The MRI requires that gas be ionized enough to couple dynamically to disk magnetic fields. [Gammie \(1996\)](#) appreciated that for protoplanetary disks, which are predominantly neutral, the MRI may not operate everywhere. Over a wide range of radii at the disk midplane, the magnetic Reynolds number

$$Re \equiv \frac{c_s h}{D} \approx 1 \left(\frac{x_e}{10^{-13}} \right) \left(\frac{T}{100 \text{ K}} \right)^{1/2} \left(\frac{a}{\text{AU}} \right)^{3/2} \quad (2.1)$$

is so low that MRI turbulence cannot be sustained against Ohmic dissipation. The magnetic Reynolds number compares the strength of magnetic induction (flux freezing) to Ohmic dissipation. Contemporary numerical simulations suggest that the MRI requires $Re > Re^* \approx 10^2$ – 10^4 , the threshold depending on the initial field geometry (e.g., [Fleming et al. 2000](#)). In equation (2.1), a is the disk radius; T is the gas temperature; c_s is the gas sound speed; $h = c_s/\Omega$ is the gas scale height; Ω is the Kepler orbital frequency; $D = 234 (T/\text{K})^{1/2} x_e^{-1} \text{ cm}^2 \text{ s}^{-1}$ is the magnetic diffusivity; and x_e is the fractional abundance of electrons by number.

[Gammie \(1996\)](#) proposed further that MRI-dead midplane gas is encased by MRI-active surface layers which satisfy the Reynolds number criterion because they are ionized sufficiently by interstellar cosmic rays. He estimated that the MRI-active layer extends down to a surface density $\Sigma^* \approx 100 \text{ g/cm}^2$ —essentially the stopping column for GeV-energy cosmic rays ([Umebayashi & Nakano 1981](#)). For comparison, the total surface density at $a = 1$ AU in the minimum-mass solar nebula is about 2200 g/cm^2 (e.g., [Chiang & Youdin 2010b](#)). Subsequent

work, discussed below, has shown that a fully MRI-active surface density of $\Sigma^* \approx 100 \text{ g/cm}^2$ is a gross overestimate.

Glassgold et al. (1997) pointed out that X-rays emitted by active stellar coronae (e.g., Preibisch et al. 2005) are more important than cosmic rays in determining the ionization fraction, in part because disk surface layers are likely shielded from cosmic rays by magnetized stellar outflows. Such shielding is almost certainly present, as its effects are appreciable even for the comparatively feeble Solar wind. At cosmic ray energies of 0.1–1 GeV, the present-day Solar wind lowers the Galactic cosmic ray flux at Earth by roughly an order of magnitude compared to estimated fluxes outside the heliosphere (see Figure 1 of Reedy 1987). Moreover, the anti-correlation between the terrestrial cosmic ray flux and Solar activity has been clearly observed over many Solar cycles (Svensmark 1998). By comparison with the Solar wind, pre-main-sequence winds have mass loss rates that are five orders of magnitude larger, and thus we expect the shielding of disk surface layers from Galactic cosmic rays to be essentially complete for young stellar systems. Nevertheless cosmic rays might still leak in from the “side,” entering the disk edge-on from the outside. The impact of “sideways” cosmic rays is difficult to assess because whether cosmic rays can penetrate radially inward depends on the radial distribution of matter and on the magnetic field distribution, both of which are unknown. Cosmic rays might spiral toward the disk along interstellar field lines only to be mirrored back. If we assume that this does not happen, and are willing to be guided by the surface density profile of the minimum-mass solar nebula, we estimate that ionization by sideways cosmic rays might be significant on the outskirts of the disk at $a \gtrsim 30 \text{ AU}$. Under these generous assumptions, cosmic rays can sustain MRI turbulence in the entire outer disk—from the disk surface to the midplane, at $a \gtrsim 30 \text{ AU}$. We will return briefly to this scenario in Section 2.4.

Bai & Goodman (2009), Turner et al. (2007), and Turner et al. (2010) studied how disk surface layers could be made MRI active by stellar X-rays. They found $\Sigma^* \sim 1\text{--}30 \text{ g/cm}^2$ —the range over which photons of energy 1–10 keV are stopped. In their models, the precise extent of the MRI-turbulent region depended on the abundance of charge-neutralizing grains. X-ray ionized layers, though less deep than the layer envisioned by Gammie (1996), were considered by Turner et al. (2010) to be capable of transporting mass at observationally significant rates at $a = 1 \text{ AU}$.

All of the aforementioned studies of layered accretion assumed that Ohmic dissipation limits Σ^* , as gauged by Re or the related Elsasser number (Turner et al. 2007). This assumption, widespread in the literature, overestimates Σ^* and the vigor of MRI turbulence in X-ray ionized surface layers. In a recent attempt to estimate Σ^* , Perez-Becker & Chiang (2011; Chapter 1 of this dissertation; PBC11 hereafter) employed an X-ray ionization model similar to those used by previous workers, but accounted for two additional factors. First, following Chiang & Murray-Clay (2007), they tested whether ambipolar diffusion—the decoupling of neutral matter from plasma—limits Σ^* . They gauged the importance of ambipolar diffusion by computing at a given depth the number of times a neutral H_2 molecule collides with

charged species in a dynamical time Ω^{-1} :

$$Am \equiv \frac{n_{\text{charge}}\beta_{\text{in}}}{\Omega} \approx 1 \left(\frac{x_{\text{charge}}}{10^{-8}} \right) \left(\frac{n_{\text{tot}}}{10^{10} \text{ cm}^{-3}} \right) \left(\frac{a}{\text{AU}} \right)^{3/2}. \quad (2.2)$$

Here $n_{\text{charge}} \equiv x_{\text{charge}}n_{\text{tot}}$ is the total number density of singly charged species (assuming multiply charged species are negligible); x_{charge} is the fractional charge density; n_{tot} is the number density of hydrogen nuclei (mostly in the form of neutral H_2); and $\beta_{\text{in}} \approx 2 \times 10^{-9} \text{ cm}^3 \text{ s}^{-1}$ is the collisional rate coefficient for singly charged species to share their momentum with neutrals (Draine et al. 1983; Millar et al. 1997). Whereas Re measures how well plasma is tied to magnetic fields, Am assesses how well neutral gas is coupled to plasma. No matter how large Re may be, unless Am is also large, the magnetic stresses felt by the plasma will not be effectively communicated to the bulk of the disk gas, which is composed overwhelmingly of neutral H_2 (Blaes & Balbus 1994; Kunz & Balbus 2004; Desch 2004; Wardle 2007). Numerical simulations by Hawley & Stone (1998, hereafter HS) suggested that unless $Am \gtrsim Am^* \approx 10^2$, MRI turbulence would not be sustained in neutral gas. Ambipolar diffusion is of especial concern in disk surface layers because Am depends on the absolute number density of charged particles, and not just the relative density as in Re . At a fixed ionization fraction, the rapidly decreasing gas density with increasing height above the midplane renders the MRI increasingly susceptible to ambipolar diffusion.

The second factor considered by PBC11 concerned polycyclic aromatic hydrocarbons (PAHs). Like grains, these macromolecules act as sites of ion recombination. Their abundances in disk surface layers were constrained by PBC11 using *Spitzer* satellite observations of PAH emission lines, detected in a large fraction of Herbig Ae stars surveyed (Geers et al. 2006).² Disk PAHs were found to reduce Am and Re by factors $\gtrsim 10$ (PBC11). For typical stellar X-ray spectra, Am reached maximum values on the order of unity at $\Sigma \sim 1 \text{ g/cm}^2$. Such maximum values were attained only for the lowest plausible PAH abundances. Comparing $\max Am \sim 1$ with HS's determination that $Am^* \sim 10^2$, PBC11 concluded that ambipolar diffusion, abetted by PAHs, reduced accretion rates in X-ray ionized surface layers to values too small compared to observations.

The conclusion of PBC11 has since been tempered by two new developments. First, PBC11 underestimated Am because they omitted the contribution from momentum-coupling collisions between H_2 and charged PAHs. This neglect is potentially significant because charged PAHs are as well coupled to the magnetic field as are atomic/molecular ions (HCO^+ and metal ions like Mg^+ in their simple model). They argued that because the abundance of charged PAHs is less than that of molecular ions, Am should not be significantly increased. This statement applies for the lowest PAH abundance considered in PBC11 of $x_{\text{PAH}} \sim 10^{-11}$

²Some workers wish to ignore PAHs when studying X-ray-driven chemistry on the grounds that they are rarely detected in T Tauri stars. It is hard for us to follow this argument. Herbig Ae disks commonly evince PAHs and exhibit accretion rates similar to if not higher than those of T Tauri disks. Treating the problem of T Tauri disk accretion separate from the problem of Herbig Ae disk accretion seems unjustified. As discussed by Geers et al. (2006), PAHs usually go undetected in T Tauri stars not because they are absent but because they fluoresce less luminously in the weaker ultraviolet fields of their host stars.

(here measured per H), but [Bai \(2011\)](#) and [Mohanty et al. \(2013\)](#) have pointed out that PAHs can dominate the charge budget at higher x_{PAH} . In this high-PAH limit, although PAHs decrease the free electron abundance and therefore Re , they can boost Am —and in principle $\dot{M}$ —in comparison to the low-PAH case. In this paper we correct for this effect by taking

$$n_{\text{charge}} = n_{\text{i}} + n_{\text{PAH}(Z=-1)} + n_{\text{PAH}(Z=1)} \quad (2.3)$$

when computing Am , where the various densities refer to atomic/molecular ions, PAHs with a single negative charge, and PAHs with a single positive charge, respectively.

In another recent development, [Bai & Stone \(2011; hereafter BS11\)](#) undertook new simulations of the MRI with ambipolar diffusion to update those of HS. They confirmed the result of HS that when $Am \gtrsim 10$ – 10^2 , neutrals and ions behave essentially as one fluid with a Shakura-Sunyaev transport parameter α that can be as high as ~ 0.1 – 0.5 . Their new result is that MRI turbulence can still be sustained in the neutrals for $Am \lesssim 1$, albeit in a weakened form with $\max \alpha$ decreasing with decreasing Am . When $Am \sim 1(0.1)$, BS11 find that $\max \alpha \approx 0.01(0.0007)$. Whether α attains its maximum value for a given Am depends on the assumed strength and geometry of the background magnetic field.

In this paper we consider a third, often neglected source of ionization: far ultraviolet (FUV $\equiv$ photon energies between ~ 6 and 13.6 eV) radiation emitted by the central star. As measured with the *Hubble Space Telescope* and *FUSE*, typical FUV luminosities from young stars are $\gtrsim 10^{30}$ erg/s, much of it in atomic lines (e.g., [Bergin et al. 2007](#)). This radiation originates from the stellar accretion shock ([Calvet & Gullbring 1998; Johns-Krull et al. 2000; Calvet et al. 2004](#)) and from the active stellar chromosphere ([Alexander et al. 2005](#)). Photoionization/photodissociation regions generated by FUV radiation have been studied extensively (e.g., [Gorti & Hollenbach 2008](#), and references therein; [Tielens & Hollenbach 1985b](#)), but their implications for the MRI have not been well quantified. [Semenov et al. \(2004\)](#) carried out detailed calculations of the disk ionization fraction due to X-ray and FUV radiation in order to estimate the extent of the MRI-active layer, but they did not consider the charge-adsorbing effects of PAHs or ambipolar diffusion. We seek here to give a more comprehensive treatment, including new estimates of $\dot{M}$ driven by FUV ionization, made possible in part by the powerful new results of BS11.

High in the disk atmosphere, molecules are photodissociated and elements take their atomic form. We focus on FUV photons (having energies less than the H Lyman limit) because they are not stopped by atomic hydrogen and as such penetrate more deeply into the disk than do Lyman continuum photons. Far-UV photons ionize trace atoms. Some elements, when nearly fully ionized, may be so abundant that the criteria for MRI accretion are easily satisfied. Carbon, for example, has a first ionization energy of 11.26 eV and a cosmic number abundance relative to hydrogen of approximately 2.9×10^{-4} ([Lodders 2003](#), her Table 2). The Strömgren layer of CII generated by FUV radiation is characterized by ionization fractions—and, by extension, values of Am and Re —up to 10^5 times larger than those reported by PBC11 for X-ray ionized layers of H_2 . Thus ambipolar diffusion may not pose the same threat in FUV-ionized layers that it does in X-ray ionized layers. Moreover,

electron densities in FUV-ionized layers may be so large that they are little impacted by PAHs or other small grains.

Of the various elements that can be ionized by FUV radiation, we model in this paper only carbon and sulfur. Their cosmic abundances are relatively high, and they are among the least likely elements to be depleted onto grains (as evidenced in the diffuse interstellar medium; e.g., [Jenkins 2009](#); [Savage & Sembach 1996](#)). Line emission from ionized carbon ([CII] 158 μm) has been detected by *Herschel* in protoplanetary disks ([Pinte et al. 2010](#); [Sturm et al. 2010](#)). The observed line fluxes can be reproduced by models that assume a cosmic gas phase abundance of carbon ([Woitke et al. 2010](#)), although our own analysis shows that lower abundances are also possible because the [CII]-emitting layer is optically thick to its own emission. The abundance of sulfur has not been so constrained because no lines from sulfur have been unambiguously detected in disks ([Lahuis et al. 2007](#); [Watson et al. 2007b](#); [Gorti & Hollenbach 2008](#)). Nevertheless [Meijerink et al. \(2008\)](#) found that a sulfur abundance of up to 7×10^{-6} per H, or $\sim 2/5$ that of cosmic ([Lodders 2003](#)), could be reconciled with the non-detection of the [SI] 25 μm line. The true upper limit on the sulfur abundance may be even higher because [Meijerink et al. \(2008\)](#) did not account for FUV ionization of SI.

Our paper is structured as follows. The equations governing the ionization balance of carbon and sulfur in disk surface layers are presented in Section 2.2. Account is made of PAHs; sub-micron sized grains which can attenuate FUV radiation; and of H_2 which can absorb photons having energies > 11.2 eV. Results for the carbon and sulfur ionization fronts, and the values for A_m , Re , and Σ^* they imply, are given in Section 2.3. The impact of the Hall effect (e.g., [Wardle & Salmeron 2012](#)) on FUV-ionized layers is also considered in Section 2.3. A summary is supplied in Section 2.4, together with estimates of $\dot{M}$ as a function of disk radius, as driven by FUV radiation, X-rays, or cosmic rays. We compare our results to stellar accretion rates observed in conventional T Tauri disks and transitional disks with inner optically thin holes.

2.2 Model for FUV ionization

At a given radius a from a star of mass $M = 1M_\odot$, we compute the equilibrium abundances of H_2 , HI, CI, CII, SI, and SII as a function of the vertical column penetrated by FUV radiation. The key equations governing the ratios of $\text{H}_2:\text{H}$, CI:CII, and SI:SII are, qualitatively,

$$\begin{aligned} \text{Rate of dissociation of } \text{H}_2 \text{ by FUV photons} \\ = \text{Rate of formation of } \text{H}_2 \text{ on grains} \end{aligned} \quad (2.4)$$

$$\begin{aligned} \text{Rate of ionization of CI by FUV photons} \\ = \text{Rate of recombination of CII} \\ \text{with electrons, grains, and PAHs} \end{aligned} \quad (2.5)$$

and

$$\begin{aligned}
 &\text{Rate of ionization of SI by FUV photons} \\
 &= \text{Rate of recombination of SII} \\
 &\quad \text{with electrons, grains, and PAHs.}
 \end{aligned} \tag{2.6}$$

For our standard model we take $a = 3$ AU. This parameter and others will be varied to explore their impact on the MRI-active column Σ^* .

2.2.1 FUV Luminosity

Two ingredients of our model, H_2 and C, are dissociated and ionized, respectively, by FUV photons having practically the same wavelength range: $\lambda \approx 912\text{--}1109$ Å. A third ingredient, S, is ionized by photons having $\lambda < 1198$ Å. In the combined wavelength interval $\lambda \approx 912\text{--}1198$ Å we take the stellar luminosity of our standard model to be $L_{\text{FUV}} = 10^{30}$ erg/s, in continuum photons. We take L_{FUV} to be distributed uniformly over this wavelength interval so that approximately $(2/3)L_{\text{FUV}}$ can be absorbed by H_2 , C, and S, while the remaining $(1/3)L_{\text{FUV}}$ can be absorbed by S but not H_2 or C.

Our choice for the continuum luminosity is compatible with *FUSE* observations of T Tauri stars (Bergin et al. 2003, see their Figure 1; admittedly their spectra are extrapolated at $\lambda < 950$ Å for TW Hydra and $\lambda < 1150$ Å for BP Tau). Our standard L_{FUV} underestimates the true ionizing luminosity because we neglect commonly observed FUV emission lines (but note that our wavelength range does not include the powerful H Lyman-alpha line at 1216 Å). It also neglects the large contribution from the hotter photospheres of more massive stars (e.g., Martin-Zaïdi et al. 2008). We make some account for these effects by varying L_{FUV} up to 10^{32} erg/s in our parameter survey (Section 2.3.3).

2.2.2 Total Gas Columns and Densities

We present our results as a function of the total vertical hydrogen column density $N_{\text{tot}} = N_{\text{H}} + 2N_{\text{H}_2}$, measured perpendicular to and toward the disk midplane. An equivalent measure is the mass surface density $\Sigma \equiv N_{\text{tot}}m_{\text{H}}$, where $m_{\text{H}} = 1.7 \times 10^{-24}$ g is the mass of the hydrogen atom. The local number density of hydrogen nuclei is approximated by

$$n_{\text{tot}} \approx N_{\text{tot}}/h. \tag{2.7}$$

Stellar radiation enters the disk at a grazing angle $\theta \sim 3h/a$ measured from the flared disk surface (e.g., Chiang et al. 2001). We assume in this work that photons travel in straight paths, i.e., we neglect scattering of FUV radiation into directions other than that of the original beam entering the disk. Thus a beam of radiation that penetrates a vertical column N_{tot} has traversed a larger column parallel to the incident beam direction of $\sim N_{\text{tot}}/\theta$, and is attenuated according to the latter quantity, not the former. To order of magnitude, $\theta \sim 0.3$.³

³We have verified that our solution does not sensitively depend on the choice of θ .

2.2.3 Gas Temperature

Gas temperatures in disk surface layers are set by a host of heating and cooling processes. Models in the literature account differently for these processes and yield temperatures which disagree. At a vertical column of $N_{\text{tot}} \sim 10^{22} \text{ cm}^{-2}$ —roughly where the deepest of the fronts we compute, the SI/SII ionization front, may lie—Glassgold et al. (2007, hereafter GNI07) reported a gas temperature of $T \approx 60 \text{ K}$ at $a \approx 5\text{--}10 \text{ AU}$ (see their Figure 2). By comparison, Gorti & Hollenbach (2008, hereafter GH08) found that at $a = 8 \text{ AU}$, the S ionization front occurs at $T \approx 600 \text{ K}$ (see their Figures 1 and 5). One reason for this difference seems to be that GNI07 omitted photoelectric heating from PAHs, which dominates thermal balance according to GH08. But the PAH abundance assumed by GH08 seems too high, exceeding by about an order of magnitude the highest plausible abundance inferred by PBC11. Reducing the heating rate from PAHs by more than a factor of 10 from its value in GH08 (see their Figure 4) would imply that X-ray heating dominates thermal balance—as it tends to do by default in GNI07.

In this paper we do not solve the thermal balance equations, but instead approximate the disk surface layer as vertically isothermal at a temperature drawn from GNI07. This is similar to the choice we made in PBC11, except that there we were interested in X-ray ionized columns $N_{\text{tot}} \gtrsim 10^{22} \text{ cm}^{-2}$, whereas here we are interested in FUV-ionized columns $N_{\text{tot}} \lesssim 10^{22} \text{ cm}^{-2}$. As N_{tot} decreases from 10^{22} to 10^{21} cm^{-2} , gas temperatures in GNI07 rise by factors of $\sim 2\text{--}3$. At still higher altitudes for which $N_{\text{tot}} < 10^{21} \text{ cm}^{-2}$, temperatures rise steeply by an order of magnitude or more. As a simple compromise we multiply the temperature profile of PBC11 by a factor of 3 but otherwise keep the same scaling:

$$T \approx 240 \left(\frac{a}{3 \text{ AU}} \right)^{-3/7} \text{ K}. \quad (2.8)$$

Our assumption that the gas is vertically isothermal ignores the steep temperature gradient at $N_{\text{tot}} < 10^{21} \text{ cm}^{-2}$ where dust-gas cooling is not effective. This neglect should not be serious as the amount of mass contained in these uppermost layers is small compared to the mass in the MRI-active layer, which we will find is concentrated near the C and S ionization fronts, located within $N_{\text{tot}} \sim 10^{21}\text{--}10^{23} \text{ cm}^{-2}$. Moreover, the locations of the fronts are not especially sensitive to temperature, as we have verified with a few test cases.

2.2.4 Very Small Grains (VSGs)

Very small grains (VSGs) having sizes of order $0.01 \mu\text{m}$ can attenuate FUV radiation. We parameterize this absorption with the variable $N_{\text{VSG}}^{\tau=1}$, the total gas column for which VSGs present unit optical depth in the FUV.⁴ We consider $N_{\text{VSG}}^{\tau=1}$ between 10^{21} and 10^{24} cm^{-2} , taking as our standard value $N_{\text{VSG}}^{\tau=1} = 10^{22} \text{ cm}^{-2}$. In the diffuse interstellar medium, $N_{\text{VSG}}^{\tau=1} \approx 10^{21} \text{ cm}^{-2}$; its value in disk surface layers may be much higher—by 1–3 orders of

⁴Equivalent to the variable σ_{H} used by GH08.

magnitude as judged by disk infrared spectral energy distributions (see Table 3 of PBC11 and references therein)—because of grain growth and sedimentation.

Colliding with ions and faster moving electrons, VSGs carry on average a net negative charge. They therefore act as sites of ion recombination. To account for such recombination, we model VSGs as conducting spheres of radius $s_{\text{VSG}} = 0.01 \mu\text{m}$. Their number abundance relative to hydrogen is

$$x_{\text{VSG}} \equiv n_{\text{VSG}}/n_{\text{tot}} = \frac{1}{N_{\text{VSG}}^{\tau=1} \pi s_{\text{VSG}}^2}. \quad (2.9)$$

In PBC11, we solved for the detailed charge distributions of grains using a set of recurrence equations. Solving the recurrence equations to obtain self-consistent solutions for the electron density, the ion density, and the total charge carried by grains/PAHs was necessary because in the lightly ionized regions irradiated by X-rays, the charge carried by PAHs could be significant compared to the amount of charge in free electrons. By comparison, FUV ionized regions are simpler to treat because free electrons are orders of magnitude more abundant than grains/PAHs. Thus our analysis here is simplified: we assume that every grain has an equilibrium charge $\bar{Z}_{\text{VSG}}$ such that it collides with ions as frequently as it collides with electrons. That is, $\bar{Z}_{\text{VSG}}$ is such that the rate coefficients (units of $\text{cm}^3 \text{s}^{-1}$) $\alpha_{\text{VSG,ion}}$ and $\alpha_{\text{VSG,e}}$ —as given by equations 9 and 10 of PBC11—are equal (see Figures 3 and 4 of PBC11). Under this approximation, the rate at which an ion (CII or SII) neutralizes by colliding with VSGs is simply $n_{\text{VSG}}\alpha_{\text{VSG}}$, where $\alpha_{\text{VSG}} = \alpha_{\text{VSG,ion}} = \alpha_{\text{VSG,e}}$. The values of α_{VSG} so derived differ by 50–70% depending on whether the dominant ion is CII (atomic weight 12) or SII (atomic weight 32); for simplicity we take a mean value. Given our choices for T and s , and assumed sticking coefficients between VSGs and electrons/ions of 1, we find that for our standard model $\bar{Z}_{\text{VSG}} \approx -1$ in units of the electron charge,⁵ and the corresponding $\alpha_{\text{VSG}} \approx 2.5 \times 10^{-6} \text{ cm}^3/\text{s}$.

A few parting comments about our treatment of VSGs: first, in deciding how VSGs attenuate FUV radiation, we do not make any explicit assumption about the grain size distribution. Our proxy for flux attenuation $N_{\text{VSG}}^{\tau=1}$ is the column at which FUV radiation is absorbed by grains of all sizes. Flux attenuation by VSGs will play a significant role in determining the extent of MRI activity (Section 2.3.3). Second, in deciding how ions recombine on VSGs, we do make an explicit assumption about grain size, namely we take $s_{\text{VSG}} = 0.01 \mu\text{m}$. Ignoring larger grain sizes maximizes the number abundance and geometric surface area of VSGs, and thus maximizes their relevance for charge balance. With this choice for grain size, ion recombination on VSGs is competitive with other recombination pathways for our standard model ($N_{\text{VSG}}^{\tau=1} = 10^{22} \text{ cm}^{-2}$), but is negligible for more dust-depleted models ($N_{\text{VSG}}^{\tau=1} > 10^{22} \text{ cm}^{-2}$).

⁵Our $\bar{Z}$ differs slightly from the true average $\langle Z \rangle$ calculated by PBC11.

2.2.5 Polycyclic Aromatic Hydrocarbons (PAHs)

Ion recombination on PAHs is treated analogously to ion recombination on VSGs. Each PAH has radius 6 Å; an electron (ion) sticking coefficient of 0.1 (1); and thus an average charge $\bar{Z}_{\text{PAH}} \approx -0.05$ at $a = 3$ AU. The corresponding ion-PAH rate coefficient $\alpha_{\text{PAH}} \approx 1 \times 10^{-8} \text{ cm}^3/\text{s}$.

For our standard (fiducial) model, we take the number abundance of PAHs relative to hydrogen nuclei to be $x_{\text{PAH}} \equiv n_{\text{PAH}}/n_{\text{tot}} = 10^{-8}$. This is their maximum plausible abundance, as inferred by PBC11. For this PAH abundance, the ion recombination rate on PAHs will turn out to be only comparable to the ion recombination rate with free electrons in FUV-ionized layers (the contribution from PAHs was much more significant for the deeper and more poorly ionized layers considered in the X-ray model of PBC11). Thus we expect the thickness of the FUV-irradiated, MRI-active layer to be insensitive to PAHs, a finding we highlight at the end of Section 2.3.3.

Note that under our simplified treatment of recombination on condensates, the net charge of our system is not zero: VSGs and PAHs carry a net negative charge, which together with the contribution from free electrons is not balanced by positive CII and SII ions. However the deviation from charge neutrality (which was not present in PBC11 because that study accounted explicitly for collisional charging of condensates by ions and electrons) is small. That is, we have verified that $\bar{Z}_{\text{PAH}}n_{\text{PAH}} + \bar{Z}_{\text{VSG}}n_{\text{VSG}}$ is much less than the electron number density over the domain of our calculation.

2.2.6 H₂: Photodissociation and Re-Formation on Grains

Molecular hydrogen is photodissociated by FUV photons in two steps: a photon of wavelength $912 \text{ Å} < \lambda < 1109 \text{ Å}$ sends the molecule into the first-excited electronic state, and then in the subsequent radiative decay, there is a $p = 23\%$ probability that the molecule lands in a vibrationally excited state where it becomes unbound (Dalgarno & Stephens 1970).⁶ Our calculation of the FUV dissociation rate of H₂ molecules follows that of de Jong et al. (1980), according to which there are $\eta_{\text{L}} = 60$ vibrationally split line transitions connecting the ground and first-excited electronic states. These lines constitute the so-called Lyman band.

The dissociation rate is

$$R_{\text{diss}} = pn_{\text{H}_2}\eta_{\text{L}} \frac{\pi e^2 f / (m_e c)}{\nu} \frac{(2/3)L_{\text{FUV}}}{4\pi a^2 h \nu} \beta_{\text{VSG}} \beta_{\text{H}_2}. \quad (2.10)$$

All lines of frequency ν (energy $h\nu \approx 12 \text{ eV}$) in the Lyman band are assumed to be equally spaced and to have equal oscillator strengths $f = 4.6 \times 10^{-3}$. The factor $\pi e^2 f / (m_e c)$ is the frequency-integrated cross section of a single line, where e and m_e are the charge and

⁶After the publication of this work, we were alerted that our value for p is incorrect. Running our model with the currently accepted value of $p = 11\%$ (e.g., Browning et al. 2003) will not affect the extent of the active layer, because the depths of the Carbon and Sulfur ionization fronts are controlled either by self-shielding or FUV absorption by grains—never by H₂-shielding.

mass of an electron, and c is the speed of light. Thus $\eta_L[\pi e^2 f / (m_e c)] / \nu \equiv \sigma_{\text{eff}}$ approximates the effective cross section of the entire Lyman band, averaged over the incident broadband spectrum of frequency width $\sim \nu$. The factor of $2/3$ accounts approximately for the fraction of L_{FUV} (as we have defined it for $912 \text{ \AA} < \lambda < 1198 \text{ \AA}$) that covers the Lyman band (see Section 2.2.1).

The factor $\beta_{\text{VSG}} = \exp[-N_{\text{tot}} / (\theta N_{\text{VSG}}^{\tau=1})]$ accounts for attenuation of FUV radiation by VSGs (recall that N_{tot} is the vertical column so that N_{tot}/θ is the column parallel to the incident beam of radiation). The factor β_{H_2} accounts for attenuation of H_2 -dissociating photons by intervening H_2 (de Jong et al. 1980):

$$\beta_{\text{H}_2} = \left(\frac{1}{\tau_{\text{H}_2} [\ln(\tau_{\text{H}_2}/\sqrt{\pi})]^{1/2}} + \sqrt{\frac{b}{\tau_{\text{H}_2}}} \right) \text{erfc} \left(\sqrt{\frac{\tau_{\text{H}_2} b}{\pi \Delta^2}} \right) \quad (2.11)$$

where erfc is the complementary error function. The core of each line is assumed to be Doppler broadened with frequency width $\Delta\nu_{\text{D}} \equiv c_s \nu / c$; then

$$\tau_{\text{H}_2} = \frac{\pi e^2 f}{m_e c \Delta\nu_{\text{D}}} N_{\text{H}_2} / \theta \quad (2.12)$$

is a measure of the optical depth at line center along the incident beam. The Voigt parameter

$$b = \gamma / (4\pi \Delta\nu_{\text{D}}) \quad (2.13)$$

compares the natural line width $\gamma = 1.16 \times 10^9 \text{ s}^{-1}$ to the Doppler width, and the non-dimensional parameter

$$\Delta = \Delta\nu_{\text{L}} / (2\eta_{\text{L}} \Delta\nu_{\text{D}}) \quad (2.14)$$

compares the inter-line spacing to the Doppler width, where $\Delta\nu_{\text{L}} = 6 \times 10^{14} \text{ Hz}$ is the frequency width of the entire Lyman band. Equation (2.11) applies only for $\tau_{\text{H}_2} \gg 1$, a condition which is valid over the domain of our calculation.

Although the photodissociating continuum for H_2 and the photoionizing continuum for CI nearly overlap in the FUV, we can safely neglect shielding of H_2 by CI because in our model the H_2 dissociation front occurs at a higher altitude (lower N_{tot}) than the C ionization front. By the same token, shielding of CI by H_2 cannot be neglected (see Section 2.2.7 where it is accounted for by the factor $\beta_{\text{C,H}_2}$).

Molecular hydrogen forms when two H atoms combine on a grain surface, releasing the heat of formation to the grain lattice (Gould & Salpeter 1963). The rate of formation is given by

$$R_{\text{form}} = \frac{1}{2} n_{\text{H}} n_{\text{VSG}} \pi s_{\text{VSG}}^2 v_{\text{H}} \eta \quad (2.15)$$

where n_{H} and v_{H} are the number density and mean thermal speed of atomic hydrogen, respectively, and $\eta \sim 0.2$ is the formation efficiency on olivine grains (Cazaux & Tielens 2010).

At each depth in our 1D model, we solve for $n_{\text{H}}/n_{\text{H}_2}$ using the equilibrium condition (2.4):

$$R_{\text{diss}} = R_{\text{form}}. \quad (2.16)$$

2.2.7 Carbon: Abundance and Ionization Equilibrium

We take the abundance of atomic carbon in the gas phase to be $x_C \equiv n_C/n_{\text{tot}} = 10^{-4}\epsilon$, where the dimensionless factor ϵ accounts for the sequestering of carbon into grains. For our standard model, $\epsilon = 1$, which gives a gas phase carbon abundance similar to that of the diffuse interstellar medium (Jenkins 2009). By using [CII] 158 μm line fluxes from disks as measured by the *Herschel* satellite (Pinte et al. 2010), we estimate a lower bound on ϵ of $\sim 1/30$.

Equation (2.5) for photoionization equilibrium of carbon is quantified as

$$\begin{aligned} & \frac{(2/3)L_{\text{FUV}}}{4\pi a^2 h\nu} n_{\text{CI}} \sigma_{\text{bf,C}} \beta_C \beta_{\text{C,H}_2} \beta_{\text{VSG}} \\ &= n_{\text{CII}} (n_e \alpha_{\text{rec,C}} + n_{\text{PAH}} \alpha_{\text{PAH}} + n_{\text{VSG}} \alpha_{\text{VSG}}), \end{aligned} \quad (2.17)$$

where $h\nu \approx 12$ eV is the typical FUV photon energy, n_{CI} ($n_{\text{CII}} = n_C - n_{\text{CI}}$) is the neutral (ionized) carbon number density, n_e is the electron number density, $\sigma_{\text{bf,C}} = 10^{-17} \text{ cm}^2$ is the bound-free cross section for CI (Cruddace et al. 1974), and $\alpha_{\text{rec,C}} = 5.3 \times 10^{-12} (T/240 \text{ K})^{-0.6} \text{ cm}^3/\text{s}$ is the radiative recombination rate coefficient for CII (Pequignot et al. 1991). The factor of 2/3 accounts for the fraction of L_{FUV} (as we have defined it in Section 2.2.1) covered by the photoionizing continuum for CI. The factor $\beta_C = \exp(-N_{\text{CI}} \sigma_{\text{bf,C}}/\theta)$ accounts for self-shielding of carbon, where N_{CI} is the vertical column of CI. The factor $\beta_{\text{C,H}_2}$ accounts for shielding of carbon by the Lyman band transitions of molecular hydrogen (de Jong et al. 1980):

$$\beta_{\text{C,H}_2} = \left(1 + \frac{\tau_{\text{H}_2} b}{\pi \Delta^2}\right)^{-1} \exp\left(-\frac{\tau_{\text{H}_2} b}{\pi \Delta^2}\right). \quad (2.18)$$

This factor drops significantly when the H_2 column $N_{\text{H}_2}/\theta \gtrsim 10^{22} \text{ cm}^{-2}$ —large enough for the Lorentzian wings of each Lyman line to blacken the entire Lyman band. As noted previously, all the FUV photons that can ionize CI ($1104 \text{ \AA} > \lambda > 912 \text{ \AA}$) can also dissociate H_2 ($1109 \text{ \AA} > \lambda > 912 \text{ \AA}$). This is true for C but not for other elements like S.

We neglect shielding of CI by SI because the S ionization front lies at a greater depth than that of the C ionization front (because S is a factor of 10 less abundant than C everywhere).

2.2.8 Sulfur: Abundance and Ionization Equilibrium

The abundance of gas-phase atomic sulfur is $x_S \equiv n_S/n_{\text{tot}} = 10^{-5}\epsilon$. For our standard model, $\epsilon = 1$, which yields a sulfur abundance similar to that of the diffuse interstellar medium (Jenkins 2009). For simplicity, we assume that the depletion factor ϵ for sulfur is the same as that for carbon, so that $x_S/x_C = 0.1$ for all our models. The lower abundance of sulfur relative to carbon implies that S-ionizing radiation will penetrate to greater depths than C-ionizing radiation (see, e.g., the order-of-magnitude estimate in PBC11), implying a deeper and more massive MRI-active layer.

Equation (2.6) for photoionization equilibrium of sulfur is quantified as

$$\begin{aligned} & \frac{(1/3)L_{\text{FUV}}}{4\pi a^2 h\nu} n_{\text{SI}} \sigma_{\text{bf,S}} \beta_{\text{S}} \beta_{\text{VSG}} \\ &= n_{\text{SII}} (n_{\text{e}} \alpha_{\text{rec,S}} + n_{\text{PAH}} \alpha_{\text{PAH}} + n_{\text{VSG}} \alpha_{\text{VSG}}), \end{aligned} \quad (2.19)$$

where $h\nu \approx 11$ eV, $\sigma_{\text{bf,S}} \approx 5.5 \times 10^{-17} \text{ cm}^2$ is the photoionizing cross-section for S (Cruddace et al. 1974), $\alpha_{\text{rec,S}} = 4.9 \times 10^{-12} (T/240 \text{ K})^{-0.63} \text{ cm}^3/\text{s}$ is the radiative recombination rate coefficient for SII (Aldrovandi & Pequignot 1973; Gould 1978), and n_{SI} ($n_{\text{SII}} = n_{\text{S}} - n_{\text{SI}}$) is the neutral (ionized) sulfur number density. The factor $\beta_{\text{S}} = \exp(-N_{\text{SI}} \sigma_{\text{bf,S}}/\theta)$ accounts for self-shielding of sulfur, where N_{SI} is the vertical column of SI.

The factor of 1/3 accounts for the fraction of L_{FUV} (as we have defined it for $912 \text{ \AA} < \lambda < 1198 \text{ \AA}$ in Section 2.2.1) that can ionize S but not be absorbed by CI and H₂. Technically there should be a second term on the left-hand side of (2.19) that is proportional to $(2/3)L_{\text{FUV}}$, to account for radiation that can ionize S and be absorbed by CI and H₂. But this term can be dropped because it does not significantly influence the location of the S ionization front; the term is negligibly small there, at a greater depth than the C ionization front. Note that the governing equation for sulfur is coupled to the other equations for C and H₂ only via the electron density $n_{\text{e}} = n_{\text{CII}} + n_{\text{SII}}$.

2.2.9 Numerical Method of Solution

Our calculation starts at $N_{\text{tot}} = 10^{18} \text{ cm}^{-2}$, under the assumption that the incident flux $L_{\text{FUV}}/4\pi a^2$ is not attenuated at this column. We take 10^3 logarithmically spaced steps to $N_{\text{tot}} = 10^{24} \text{ cm}^{-2}$. At each step, we solve for the ratio $n_{\text{H}}/n_{\text{H}_2}$ using the balance condition (2.16). The self-shielding factor β_{H_2} depends on N_{H_2} and so we keep a running integral of n_{H_2} , i.e., $N_{\text{H}_2} = \sum_i (n_{\text{H}_2}/n_{\text{tot}})_i (\Delta N_{\text{tot}})_i$.

By approximating $n_{\text{e}} = \max(n_{\text{CII}}, n_{\text{SII}})$, we may solve analytically for the ionization states of C and S. At low N_{tot} , nearly all C and S are ionized and $n_{\text{e}} = n_{\text{CII}}$ with an error of $x_{\text{C}}/x_{\text{S}} \approx 10\%$. Equation (2.17) becomes a quadratic for $n_{\text{CII}}/n_{\text{CI}}$, and $n_{\text{SII}}/n_{\text{SI}}$ is determined by substituting n_{CII} for n_{e} in equation (2.19).

At N_{tot} past the C ionization front, SII becomes the dominant ion and we set $n_{\text{e}} = n_{\text{SII}}$. Now equation (2.19) becomes a quadratic for $n_{\text{SII}}/n_{\text{SI}}$. The resultant value for n_{SII} is substituted for n_{e} in equation (2.17) to solve for $n_{\text{CII}}/n_{\text{CI}}$.

At each N_{tot} , we keep track of the running columns $N_{\text{CI}} = \sum_i (n_{\text{CI}}/n_{\text{tot}})_i (\Delta N_{\text{tot}})_i$ and $N_{\text{SI}} = \sum_i (n_{\text{SI}}/n_{\text{tot}})_i (\Delta N_{\text{tot}})_i$ to compute the self-shielding factors β_{C} and β_{S} , respectively.

2.3 Results for MRI-Active Surface Densities

In Section 2.3.1 we describe how the H₂:H, CI:CII, and SI:SII ratios vary with depth. In Section 2.3.2 we estimate the surface density Σ^* of the layer that can be MRI-active. In Section 2.3.3 we perform a parameter study of how Σ^* varies with system properties.

The MRI-active surface densities Σ^* reported in Section 2.3.3 are those resulting from FUV ionization only. In Section 2.3.4 we check that the Hall effect is not significant for FUV-ionized surface layers. All our calculations neglect turbulent mixing of plasma from disk surface layers into the disk interior. In Section 2.3.5, we try to assess whether this neglect is justified, by comparing the dynamical timescale with the timescale over which the FUV-irradiated layer equilibrates chemically.

2.3.1 Photodissociation and Ionization Fronts

In the top panel of Figure 2.1, we show the relative abundances of species as a function of vertical mass column Σ . High in the atmosphere, hydrogen is mostly in atomic form and nearly all of the carbon and sulfur are ionized. Deeper down, starting at $\Sigma \sim 3 \times 10^{-5} \text{ g cm}^{-2}$, atomic H gives way to H_2 . At the H_2 :H front the Doppler broadened cores of the various H_2 Lyman transitions have become optically thick. Far-UV photons in the wings of the Lyman transitions can stream past the H_2 :H front and continue to ionize C and S at greater depths.

The CI:CII ionization front is located at $\Sigma \sim 3 \times 10^{-3} \text{ g cm}^{-2}$, where the optical depth to carbon ionizing photons is of order unity. At this front, both shielding of carbon by H_2 (quantified by $\beta_{\text{C,H}_2}$) and carbon self-shielding (β_{C}) are becoming significant; see the bottom panel of Figure 2.1.

Far-UV photons having $1198 \text{ \AA} > \lambda > 1109 \text{ \AA}$ interact with neither H_2 nor C, but can ionize S. These photons penetrate both H_2 and C fronts, and ionize S at greater depths. The SI:SII ionization front is located at $\Sigma \sim 1 \times 10^{-2} \text{ g cm}^{-2}$. In this fiducial model, the location of the S ionization front is determined more by flux attenuation by very small grains than by self-shielding by sulfur; at the front, $\beta_{\text{VSG}} \sim 0.1$ and $\beta_{\text{S}} \sim 1$. Lower but still observationally realistic grain abundances are considered in Section 2.3.3, where we will find that the S ionization front can be as deep as $\Sigma \sim 1 \times 10^{-1} \text{ g cm}^{-2}$.

2.3.2 MRI-Active Surface Density Σ^*

We measure the extent of the MRI-active column by means of the magnetic Reynolds number Re and the ion-neutral collision rate Am . Figure 2.2 shows both dimensionless numbers as a function of Σ for our standard model at $a = 3 \text{ AU}$.

Above the CI:CII ionization front, at $\Sigma \lesssim 3 \times 10^{-3} \text{ g cm}^{-2}$, the value of Re remains constant at $\sim 10^{10}$ because x_e saturates at 10^{-4} where nearly all the carbon is singly ionized. In this region, $Am \propto \Sigma$ because $Am \propto n_{\text{tot}}$. The value of Am peaks at $\sim 3 \times 10^3$ near the C ionization front. Both Am and Re fall just past the C ionization front, but not by more than an order of magnitude: S, which is $10\times$ less abundant than C, remains fully ionized out to the S ionization front at $\Sigma \sim 10^{-2} \text{ g/cm}^2$. Past the S ionization front, virtually all FUV photons are used up; the charge density, and by extension both dimensionless numbers, drop sharply.

We define the MRI-active surface density Σ^* as follows. Bai & Stone (2011) find from an extensive series of numerical simulations that for a given Am , the maximum value of α is

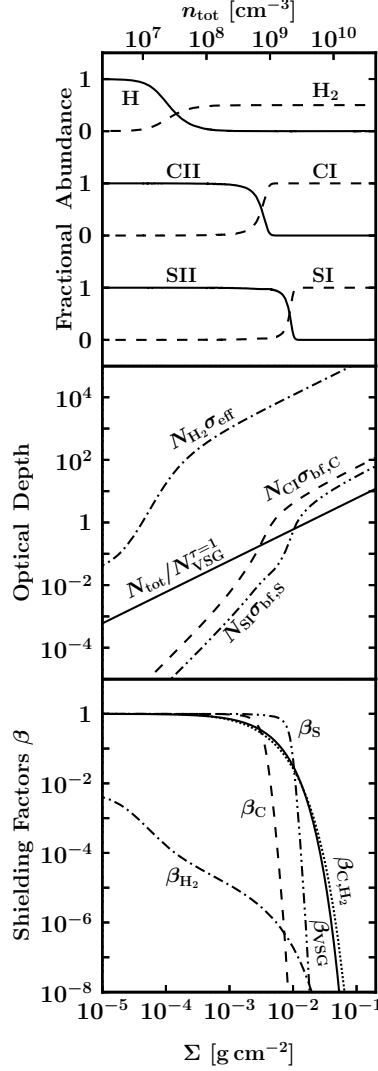


Figure 2.1 H_2 photodissociation front, and C and S ionization fronts, for our standard model at $a = 3$ AU. Results are given as functions of the vertical surface density Σ , measured perpendicular to the disk midplane (the column actually penetrated by radiation parallel to the incident beam direction is Σ/θ). The corresponding number density of H nuclei is given on the top x-axis. *Upper panel:* Normalized abundances of atomic H, H_2 , CII, CI, SII, and SI. The H_2 photodissociation front occurs at the highest altitude above the midplane, and the S ionization front occurs at the lowest. *Middle panel:* Optical depths to FUV radiation presented by H_2 (averaged over the Lyman band), CI, SI, and dust (VSGs). *Lower panel:* Shielding factors by which FUV radiation is attenuated: self-shielding of H_2 (β_{H_2}); self-shielding of C (β_{C}); shielding of C by H_2 ($\beta_{\text{C,H}_2}$); self-shielding of S (β_{S}); and shielding by very small grains (β_{VSG}).

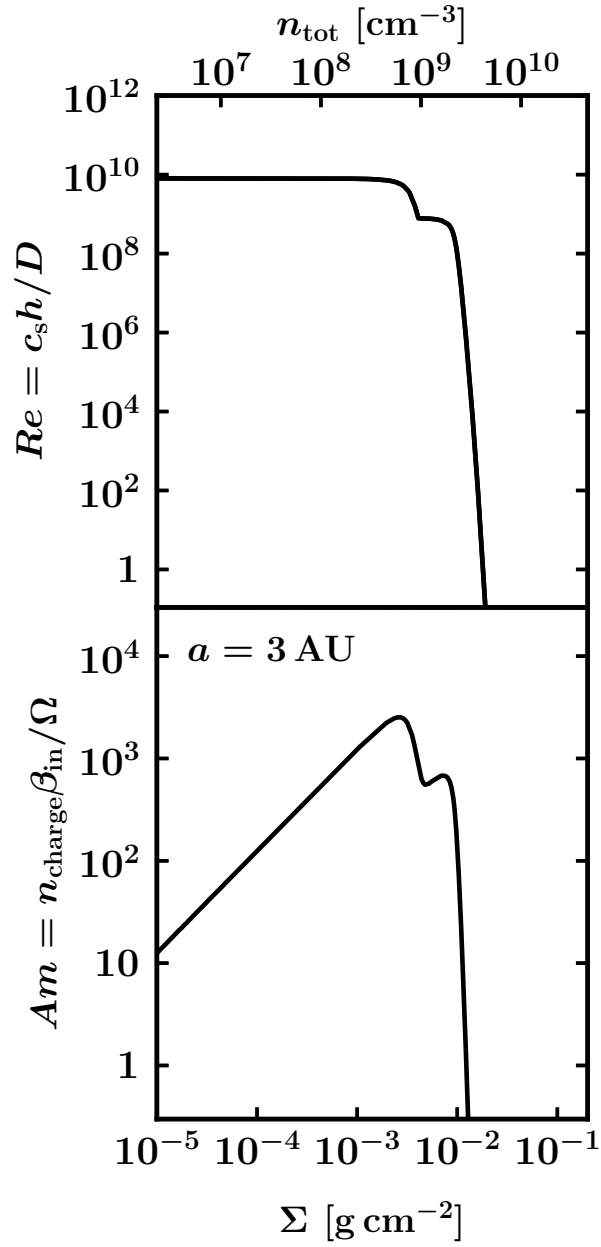


Figure 2.2 Magnetic Reynolds number Re and ambipolar diffusion number Am versus vertical surface density Σ for standard model parameters and $a = 3 \text{ AU}$.

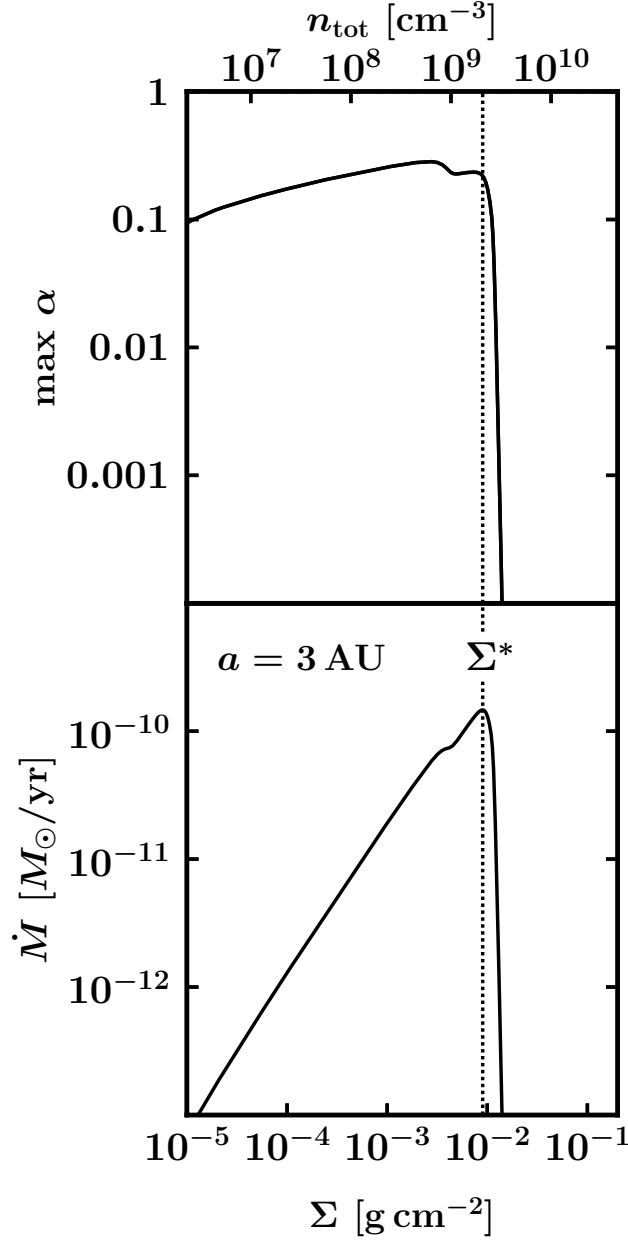


Figure 2.3 Maximum transport parameter α and mass accretion rate $\dot{M}$ versus vertical surface density Σ for standard model parameters and $a = 3 \text{ AU}$. The value of $\max \alpha$ is computed using the empirical relation (2.20) of BS11. A vertical dotted line marks the MRI-active surface density Σ^* at which, by definition, $\dot{M}$ peaks. The value of Σ^* in FUV-ionized layers is essentially set by the SI:SII ionization front. It can be as large as $\sim 0.1 \text{ g/cm}^2$ for parameter choices other than those assumed here; see Figure 2.4.

given by

$$\max \alpha = \frac{1}{2} \left[\left(\frac{50}{Am^{1.2}} \right)^2 + \left(\frac{8}{Am^{0.3}} + 1 \right)^2 \right]^{-1/2}. \quad (2.20)$$

From the value of $\max \alpha$ so computed, we evaluate the mass accretion rate

$$\dot{M} = 2 \times 3\pi \Sigma \nu = 6\pi \Sigma \times \max \alpha \times \frac{kT}{\mu \Omega} \quad (2.21)$$

as a function of Σ , where $\nu = \alpha c_s h$ is the viscosity, $\mu = 4 \times 10^{-24}$ g is the mean molecular weight, k is Boltzmann's constant, and the prefactor of 2 accounts for the top and bottom surfaces of the disk. Note that $\max \alpha$ is a function of Am , which in turn is a function of Σ that we have computed (Figure 2.2). We define Σ^* as that value of Σ for which $\dot{M}$ peaks; see Figure 2.3, which plots equation (2.21). According to this definition, $\Sigma^* \approx 1 \times 10^{-2}$ g/cm². At this column, according to Figure 2.2, Re remains larger than Re^* , demonstrating that ambipolar diffusion limits MRI activity more than Ohmic dissipation does, although the effect is slight (the dominance of ambipolar diffusion over Ohmic dissipation was much more pronounced for the X-ray ionized layers analyzed by PBC11). Referring back to Figure 2.1, we see that Σ^* corresponds essentially to the SI:SII ionization front.

2.3.3 Parameter Survey, Including Sensitivity to PAHs

In Figure 2.4 we explore the sensitivity of Σ^* to various parameters. For a given solid curve in a given panel, only the one parameter on the x-axis is varied while all others are held fixed. Labels on a curve indicate the values to which certain parameters are fixed when computing that curve. If a parameter is not labeled, its value was set to the standard value.

Over most of parameter space, the MRI-active surface density Σ^* in FUV-ionized layers is of order 10^{-2} – 10^{-1} g/cm². The MRI-active surface density is set by the depth of the sulfur ionization front. Very small grains (VSGs) reduce the depth of this front by absorbing FUV radiation (and also by enhancing the ion recombination rate, although this is not a dominant effect for FUV-ionized layers where the electron density and thus the radiative recombination rate are high). At their standard abundance—which is probably near their maximum abundance (Section 2.2.4)—VSGs shield sulfur from ionizing radiation more than sulfur shields itself; compare β_S and β_{VSG} in Figure 2.1.

As $N_{VSG}^{\tau=1}$ increases from 10^{22} to 10^{24} cm⁻² (i.e., as the grain abundance decreases), Σ^* increases by a factor of 3 (Figure 2.4b). Now most of the shielding of sulfur from FUV radiation is provided by sulfur itself, not by VSGs. When grains are insignificant ($N_{VSG}^{\tau=1} = 10^{24}$ cm⁻²), Σ^* depends on L_{FUV} and the carbon/sulfur abundance ϵ more strongly than in the case when grains are significant absorbers of FUV radiation (Figure 2.4c, 2.4d). The dependences in the case of low grain abundance resemble those calculated by PBC11 in their simple FUV-Strömberg model: $\Sigma^* \propto L_{FUV}^{1/2}$ and $\Sigma^* \propto 1/\epsilon$.

The MRI-active thickness Σ^* in FUV-ionized layers hardly varies with disk radius a (Figure 2.4a), because the decrease in ionizing flux with increasing a is compensated by the lengthening dynamical time Ω^{-1} (to which Am , the factor determining Σ^* , is proportional).

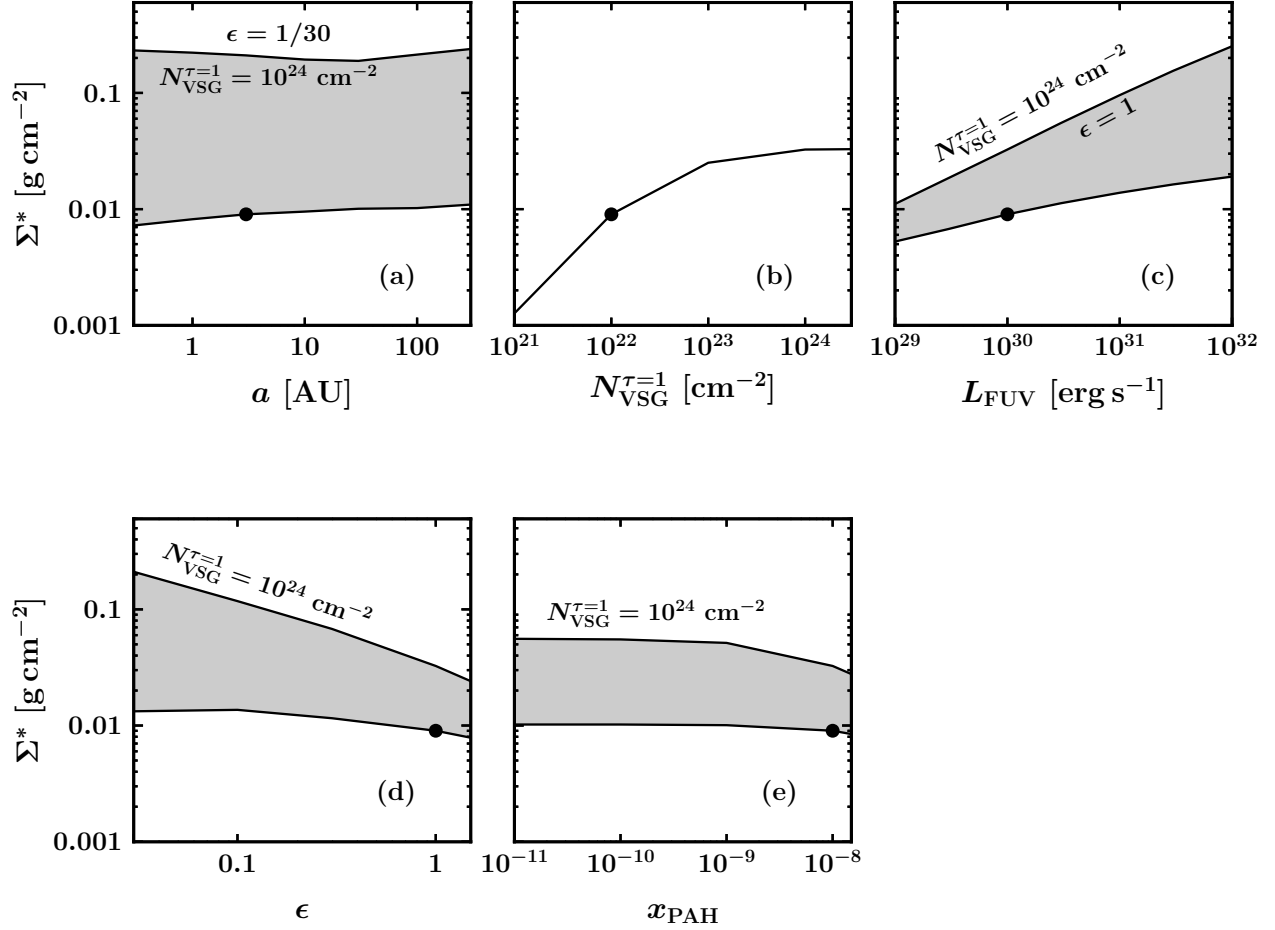


Figure 2.4 How Σ^* varies with (a) stellocentric distance a ; (b) our proxy $N_{\text{VSG}}^{\tau=1}$ for the abundance of very small grains (see equation 2.9); (c) FUV luminosity L_{FUV} ; (d) the gas-phase abundance of carbon and sulfur ϵ (normalized so that $\epsilon = 1$ corresponds to near-solar abundances); and (e) the abundance of charge-absorbing PAHs. For a given solid curve in a given panel, only the one parameter on the x-axis is varied while all other parameters, unless otherwise annotated on the figure, are held fixed at their standard values. The upper bounding curves in panels (a), (c), (d), and (e) are calculated assuming $N_{\text{VSG}}^{\tau=1} = 10^{24} \text{ cm}^{-2}$, the lowest grain abundance considered. In panel (a), the upper bounding curve assumes $\epsilon = 1/30$, which we estimate to be the lowest possible value still consistent with observations of far-infrared [CII] emission from disks (Pinte et al. 2010; Sturm et al. 2010). A dot in each panel marks the standard (fiducial) model.

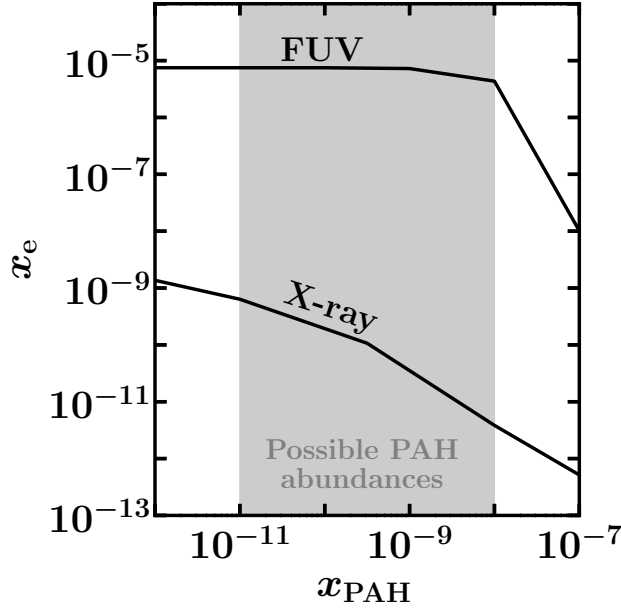


Figure 2.5 Electron fraction as a function of PAH abundance for FUV and X-ray ionized surface layers. The FUV curve is computed for our standard model parameters ($a = 3$ AU, $L_{\text{FUV}} = 10^{30}$ erg/s, $N_{\text{VSG}}^{\tau=1} = 10^{22}$ cm $^{-2}$, $\epsilon = 1$) at $\Sigma \approx 0.01$ g cm $^{-2}$. The large abundance of electrons generated in FUV-ionized layers is immune to the effects of charge recombination on PAHs, over the range of plausible PAH abundances (shaded in grey; PBC11). We contrast this behavior with the X-ray curve taken from Figure 7 of PBC11 ($a = 3$ AU, $L_X = 10^{29}$ erg/s, metal abundance $x_M = 10^{-8}$, $\Sigma = 0.3$ g cm $^{-2}$), which shows that PAHs can reduce electron fractions in X-ray-irradiated layers by two orders of magnitude.

Finally, Σ^* is not sensitive to the abundance of PAHs (Figure 2.4e). For the large electron fractions $x_e \approx 10^{-5}$ – 10^{-4} generated by FUV ionization of carbon and sulfur, the primary channel for charge neutralization is radiative recombination with free electrons. Only at the highest PAH abundances does the ion recombination rate with PAHs ($x_{\text{PAH}}\alpha_{\text{PAH}}$) approach the ion recombination rate with electrons ($x_e\alpha_{\text{rec,S}}$). By contrast, X-ray ionized gas has much lower electron fractions and is much more susceptible to charge recombination on small condensates like PAHs (PBC11).

The last point is echoed in Figure 2.5, which shows how the free electron fraction varies with PAH abundance in both FUV-ionized and X-ray-ionized surface layers. Each curve is computed at a fixed surface density characteristic of each layer: $\Sigma \approx 0.01$ g/cm 2 for the FUV-irradiated layer, and $\Sigma \approx 0.3$ g/cm 2 for the X-ray-irradiated layer. For the FUV-ionized layer, x_e remains constant over the range of plausible PAH abundances $10^{-11} \lesssim x_{\text{PAH}} \lesssim 10^{-8}$. For the X-ray-ionized layer, x_e varies by about two orders of magnitude over the same range of PAH abundance; $x_e \propto 1/x_{\text{PAH}}$ for $x_{\text{PAH}} \gtrsim 5 \times 10^{-10}$ (see Section 3.2.1 of PBC11).

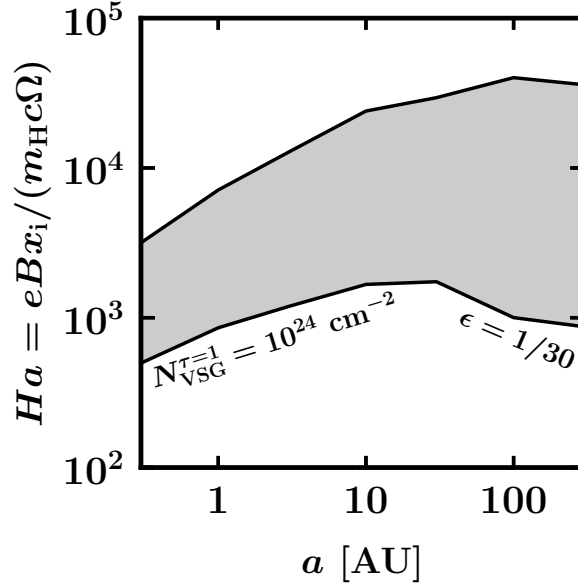


Figure 2.6 Hall parameter Ha as a function of a in FUV-ionized surface layers. Values of B are computed using equations (2.23) and (2.24). The upper bounding curve is computed for our standard (fiducial) parameters, at $\Sigma^* \approx 0.01$ g/cm². The lower bounding curve is computed for a model depleted in dust, carbon, and sulfur, for which $\Sigma^* \approx 0.2$ g/cm² (see also Figure 2.4a). Hall diffusion is not significant for FUV-irradiated surface layers.

2.3.4 Hall Diffusion

When computing Σ^* , we have ignored the effects of Hall diffusion (Wardle 1999; Balbus & Terquem 2001; Sano & Stone 2002). When Hall diffusion dominates, only electrons are coupled to magnetic fields, and ions are de-coupled from magnetic fields by collisions with neutrals. Under these circumstances, the evolution of the MRI depends on the direction of the magnetic field $\mathbf{B}$ with respect to the angular frequency $\mathbf{\Omega}$. Hall diffusion can increase/decrease Σ^* by an order of magnitude or more when $\mathbf{B}$ and $\mathbf{\Omega}$ are parallel/anti-parallel (Wardle & Salmeron 2012).

To assess the importance of Hall diffusion in FUV-ionized surface layers, we evaluate the Hall parameter

$$Ha \equiv \frac{v_A^2}{\eta_{Ha} \Omega} = \frac{eBx_i}{m_H c \Omega}. \quad (2.22)$$

Here $x_i \equiv n_i/n_{\text{tot}}$ is the fractional abundance of ions of density n_i relative to hydrogen nuclei of density n_{tot} ; $v_A = B/\sqrt{4\pi n_{\text{tot}} m_H}$ is the Alfvén velocity; e is the electron charge; c is the speed of light; and $\eta_{Ha} = cB/(4\pi e n_i)$ is the Hall diffusivity (e.g., Wardle & Salmeron 2012, but note that their Hall parameter is the inverse of ours, and they assume $n_i = n_e$). The dimensionless Hall parameter is the ratio of the inductive term to the Hall term in the magnetic induction equation. If $Ha \gg 1$, then Hall diffusion is not important.

To evaluate Ha , we estimate B from $\dot{M}$ by assuming that the inequality

$$B^2 \geq B_r^2 + B_\phi^2 \geq 2B_r B_\phi \quad (2.23)$$

saturates, and that the Maxwell stress which drives accretion is given by

$$B_r B_\phi \approx \frac{\dot{M}\Omega}{2h} \quad (2.24)$$

(see, e.g., [Bai & Goodman 2009](#); the factor of 2 in equation 2.24 arises because the disk has a top and bottom face).

In Figure 2.6 we show that $Ha \gg 1$ at Σ^* for all a , over the entire parameter space that we have explored. Thus we conclude that Hall diffusion is not a concern in FUV-ionized surface layers (the same may not be true for the more poorly ionized X-ray-irradiated layers; [Wardle & Salmeron 2012](#)).

2.3.5 The Possibility of Turbulent Mixing: Chemical Equilibration Timescale vs. Dynamical Timescale

Throughout this paper, we have computed ionization fractions in a static atmosphere. But the MRI-active layer is not static; it is turbulent. We might have underestimated Σ^* because turbulence can mix plasma vertically toward the midplane, deeper into the disk interior ([Inutsuka & Sano 2005](#); [Ilgner & Nelson 2006b](#); [Turner et al. 2007](#)). Mixing would be effective if the vertical mixing time is short compared to the timescale over which ionized layers achieve chemical equilibrium. Without a full-out simulation of MRI turbulence, we approximate the vertical mixing time as the dynamical time $t_{\text{dyn}} = \Omega^{-1}$. For the chemical equilibration time, we substitute the radiative recombination time $t_{\text{rec}} = 1/(n_e \alpha_{\text{rec,S}})$. The ratio $t_{\text{rec}}/t_{\text{dyn}}$ is plotted in Figure 2.7 (cf. Figure 11 of PBC11). Because $t_{\text{rec}}/t_{\text{dyn}} > 1$ over some portion of parameter space, mixing might well be significant in FUV-ionized layers, and might increase Σ^* above the values we have computed in this paper. Whether the increase would be large or small is hard to say. On the one hand, $t_{\text{rec}}/t_{\text{dyn}}$ is never far above unity, suggesting that turbulent mixing will merely introduce order-unity corrections to our estimates. On the other hand, as turbulent mixing dilutes the electron density n_e , the timescale ratio $t_{\text{rec}}/t_{\text{dyn}}$ might increase with increasing depth, and the process might run away.

2.4 Summary and Implications for Disk Accretion

Circumstellar disk material must be sufficiently ionized if it is to accrete by the magnetorotational instability (MRI). We have considered in this paper ionization by stellar far-ultraviolet (FUV) radiation. Although FUV radiation cannot penetrate the disk as deeply as can X-rays, it generates ionization fractions orders of magnitude larger. Carbon and sulfur may be the principal sources of free electrons and ions, as these elements are cosmically abundant and

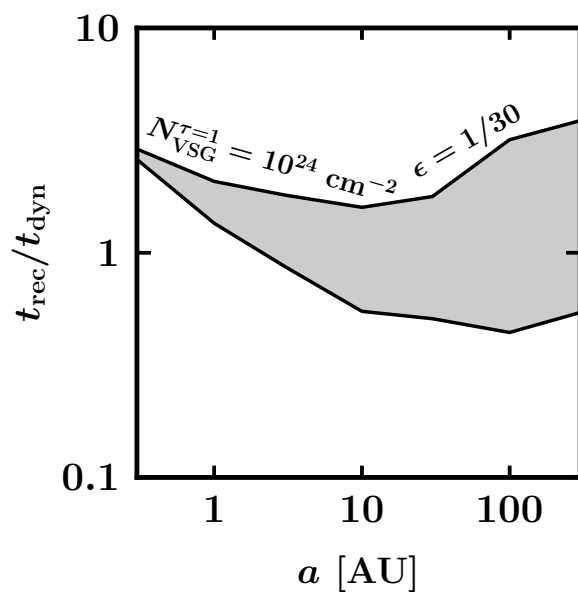


Figure 2.7 Ratio of the ion-electron recombination timescale $t_{\text{rec}} = (n_e \alpha_{\text{rec,S}})^{-1}$ to the dynamical time $t_{\text{dyn}} = \Omega^{-1}$, as a function of a . The lower bounding curve is computed for our standard (fiducial) parameters, at $\Sigma^* \approx 0.01 \text{ g/cm}^2$. The upper bounding curve is computed for a dust-depleted and carbon/sulfur-depleted model, for which $\Sigma^* \approx 0.2 \text{ g/cm}^2$ (see also Figure 2.4).

least likely to be depleted onto grains. Far-infrared searches for carbon and sulfur emission from disks are so far consistent with gas phase abundances for both elements within a factor of ~ 2 of solar. Because the electron and ion abundances generated from ionized carbon and sulfur are so large, ionization fractions in FUV-irradiated layers are little impacted by PAHs. In FUV-ionized layers, fractional ion abundances $x_i \approx 10^{-5}$ – 10^{-4} , dwarfing PAH abundances of $x_{\text{PAH}} \sim 10^{-11}$ – 10^{-8} . More to the point, ion recombination on PAHs in FUV-ionized layers is at most competitive with ion recombination with free electrons. This immunity to PAHs does not apply to X-ray ionized layers where $x_i \lesssim 10^{-9}$ and where PAHs or very small ($0.01 \mu\text{m}$ sized) grains can suppress the MRI (Bai & Goodman 2009; Perez-Becker & Chiang 2011, PBC11). It is refreshing that FUV ionization is robust against the usual difficulties plaguing other, weaker sources of ionization.

In Figure 2.8 we compute the possible ranges of disk accretion rate $\dot{M}$ due to either FUV or X-ray ionization, and compare them against the range of observed stellar accretion rates. The accretion rate $\dot{M}$ driven by FUV ionization is derived using equations (2.20) and (2.21) for $\Sigma = \Sigma^*$ (taken from Figure 2.4a) and the corresponding value of $\max \alpha$. The accretion rate $\dot{M}$ driven by X-ray ionization is computed similarly, using the abundances of charged species of the standard model of PBC11, their temperature of $T = 80(a/3\text{AU})^{-3/7}$ K, and Equations (2.2)–(2.3) of this work to compute $Am(\Sigma)$. Although our estimates of $\dot{M}$ driven by X-ray ionization may have to be revised because of the Hall effect (Wardle & Salmeron 2012), our estimates of $\dot{M}$ driven by FUV ionization should not be: FUV-ionized, MRI-active layers behave in the ideal magnetohydrodynamic limit.

Modulo the impact of the Hall effect on X-ray-driven MRI, from Figure 2.8 we conclude that accretion rates from FUV-ionized surface layers are of the same order of magnitude as accretion rates from X-ray ionized layers. At large radii $a \gtrsim 10$ AU, FUV-driven accretion rates may exceed X-ray-driven rates. At $a \lesssim 10$ AU, their contributions may be more nearly equal. At $a \gtrsim 10$ AU, the FUV-irradiated surface layer can sustain accretion rates similar to those observed. At $a \lesssim 10$ AU, surface layer accretion rates, driven either by FUV or X-ray ionization, can still be observationally significant, but they diminish with decreasing radius. At $a \lesssim 1$ AU, surface layer accretion rates fall below the range typically observed for young solar-type stars.

The problem of too low an accretion rate at small radius might be alleviated by turbulent mixing of plasma from FUV-irradiated disk surface layers into the disk interior, as such mixing would enhance the thickness of the MRI-active layer. Our crude estimate of the timescales involved (Section 2.3.5) suggests that this possibility is worth further consideration.

Note in Figure 8 how the X-ray-driven $\dot{M}$ actually increases from the low-PAH case to the high-PAH case. This surprising effect arises because increasing the PAH abundance (over this particular range from $x_{\text{PAH}} = 10^{-11}$ to 10^{-8}) increases the abundances of charged PAHs themselves, which more than offsets the decreased abundances of free electrons and ions. In fact, in the high-PAH case, positively and negatively charged PAHs are approximately equal in number and represent by far the most abundant charged species. The primary recombination pathway is positively charged PAHs colliding with negatively charged PAHs (reaction 14 in PBC11). Thus as the overall PAH abundance increases, charged PAHs can

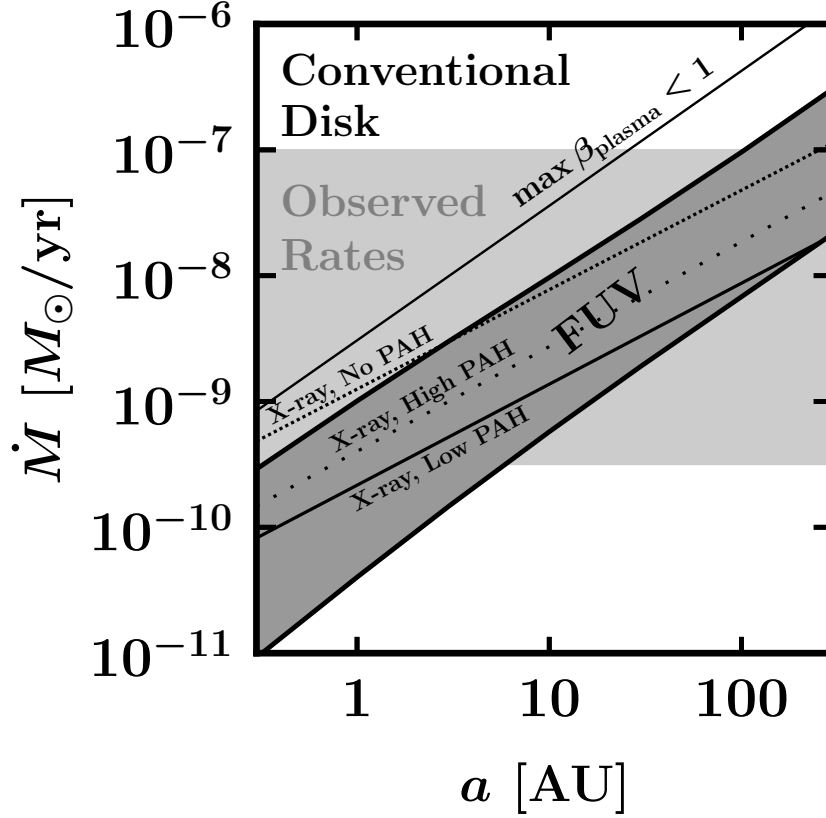


Figure 2.8 Accretion rates for conventional (hole-less) disks. The shaded region labeled “FUV” (bounded by heavy solid lines) corresponds to the range of accretion rates possibly driven by the MRI in FUV-ionized surface layers. The range of FUV-driven accretion rates corresponds to the range of MRI-active surface densities Σ^* shown in Figure 2.4a, which in turn corresponds to a range of possible abundances for grains, carbon, and sulfur (PAHs, whose abundances for all our FUV models are set to their maximum value, are not significant for FUV-ionized layers). We compute X-ray-driven accretion rates using our corrected definition for Am (Equations 2.2–2.3), for the low-PAH and high-PAH cases considered by PBC11. The case with no PAHs is shown for comparison only. Note how little the X-ray-driven $\dot{M}$ varies, despite the PAH abundance having changed by three orders of magnitude; the small variation in $\dot{M}$ follows from Equation (2.3), in which reductions in n_i from charge-neutralizing PAHs are offset by the increased abundances of charged PAHs themselves. In fact, the high-PAH $\dot{M}$ even exceeds slightly the low-PAH $\dot{M}$ (see also Bai 2011). At $a \lesssim 10$ AU, accretion rates in FUV-ionized layers may be comparable to those in X-ray ionized layers. At $a \gtrsim 10$ AU, FUV-driven accretion rates tend to be larger than X-ray-driven rates. The light shaded region labeled “Observed Rates” brackets the range of stellar accretion rates for stars of mass $M \approx 0.3\text{--}1M_\odot$ as shown in Figure 5 of Muzerolle et al. (2005).

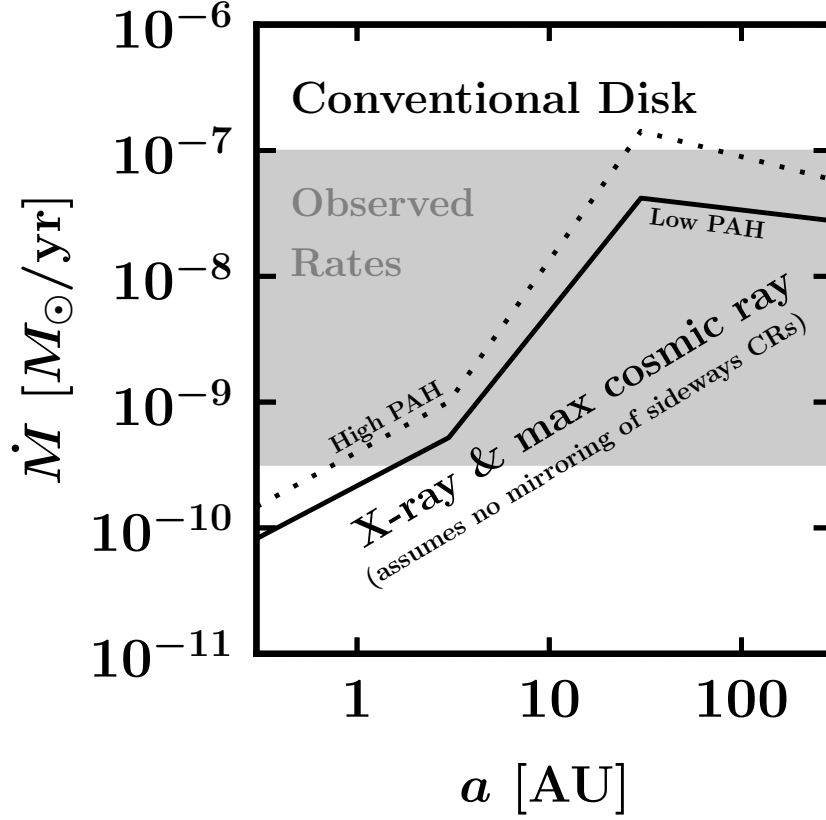


Figure 2.9 Same as Figure 2.8, but showing disk accretion rates driven by the combined effects of X-ray and cosmic ray ionization. Galactic cosmic rays cannot reach the disk through its top and bottom faces because of shielding by the magnetized stellar wind. They can, however, penetrate the disk “sideways” through the outermost portions of the disk, parallel to the midplane—assuming they are not magnetically mirrored away. To generate this plot we assume a total (vertical) surface density of $2200(a/1 \text{ AU})^{-3/2} \text{ g/cm}^2$, and a cosmic ray ionization rate $\zeta_0 \exp(-\Sigma_a/\Sigma_0)$, where $\zeta_0 = 1/4 \times 10^{-17} \text{ s}^{-1}$ (Caselli et al. 1998) and $\Sigma_0 = 96 \text{ g/cm}^2$ is the stopping column for GeV-energy cosmic rays (Umebayashi & Nakano 1981). The factor of $1/4$ in ζ_0 is our estimate for the fraction of the celestial sphere (centered at the midplane) not shielded by stellar winds. The attenuating column Σ_a is obtained by radially integrating the volume mass density at the midplane from a to infinity. For $a \gtrsim 30 \text{ AU}$, sideways cosmic rays may provide enough ionization to render the entire disk MRI active (PBC11). Sideways cosmic rays cannot penetrate the disk at $a \lesssim 10 \text{ AU}$ because of intervening disk material, and so the accretion rates there are practically identical to those for the X-ray-only case (plotted in Figure 2.8). The solid and dotted curves correspond to the cases of low and high PAH abundances, respectively. The main assumption underlying this plot is that sideways cosmic rays are not magnetically mirrored away; this is an uncertain prospect.

increase Am at a given column Σ . This effect was missed by PBC11, and is highlighted by Bai (2011). The consequent enhancement in $\dot{M}$ is mitigated by the decrease in the free electron fraction and thus in Re . In computing the X-ray-driven $\dot{M}$ for the high-PAH case in Figure 8, we accounted for the mitigating effects of a lower Re by evaluating Am at the Σ for which $Re \approx 100$, below which Ohmic dissipation would weaken the MRI.

One concern is whether gas pressures in FUV-ionized layers are lower than the magnetic field pressures required to drive our computed accretion rates. If the pressure ratio

$$\beta_{\text{plasma}} \equiv \frac{\Sigma^* kT / (\mu h)}{B^2 / 8\pi} \quad (2.25)$$

is < 1 , magnetic tension defeats the MRI. An upper bound on β_{plasma} is obtained by substituting the lower bound on B^2 inferred from equations (2.23) and (2.24). For a representative case $\Sigma^* = 0.1 \text{ g/cm}^2$, equations (2.23)–(2.25) combine to yield the topmost slanted line in Figure 2.8, above which $\max \beta_{\text{plasma}} < 1$ and the MRI cannot operate. The range of FUV-driven accretion rates we have computed sits safely below this line.

In sum, surface layer accretion by FUV ionization can, by itself, solve the problem of protoplanetary disk accretion at large radius, but not at small radius (unless turbulent mixing of plasma can substantially thicken the MRI-active layer). This statement remains one of principle and not of fact, because MRI accretion rates depend on the strength and geometry of the background magnetic field threading the disk (e.g., Fleming et al. 2000; Pessah et al. 2007; Bai & Stone 2011), and unfortunately the field parameters are not known for actual disks.⁷

For completeness, we show in Figure 2.9 how sideways cosmic rays can, in principle, enhance accretion rates in the outermost portions of disks (see the Introduction). For an assumed radial surface density profile resembling that of the minimum-mass solar nebula, accretion rates at $a \gtrsim 30 \text{ AU}$ can be strongly increased by sideways cosmic rays. Although cosmic-ray induced accretion rates at large radius can be competitive with FUV-induced rates, it is not clear that cosmic rays are not mirrored away from disks by ambient magnetic fields. Far-UV ionization does not suffer from this uncertainty.

2.4.1 Transitional Disks

What is the relevance of MRI accretion to transitional disks, i.e., disks with inner holes (e.g., Espaillat et al. 2010; Hughes et al. 2007; Calvet et al. 2005)? Chiang & Murray-Clay (2007, hereafter CMC) proposed that MRI-active surface layers at the rim of the hole—i.e., in the rim “wall” oriented perpendicular to the disk midplane—could supply the observed

⁷The rotation measures of magnetized FUV-ionized layers could be large, on the order of 10^4 rad/m^2 at $a \sim 30 \text{ AU}$ for a face-on disk (assuming an untangled $B \sim 10 \text{ mG}$ and an electron column density of $N_e \sim 3 \times 10^{18} \text{ cm}^{-2}$). Perhaps measurements of Faraday rotation at radio wavelengths using polarized background active galaxies, or polarized emission from the central star itself, could be used to constrain disk magnetic fields.

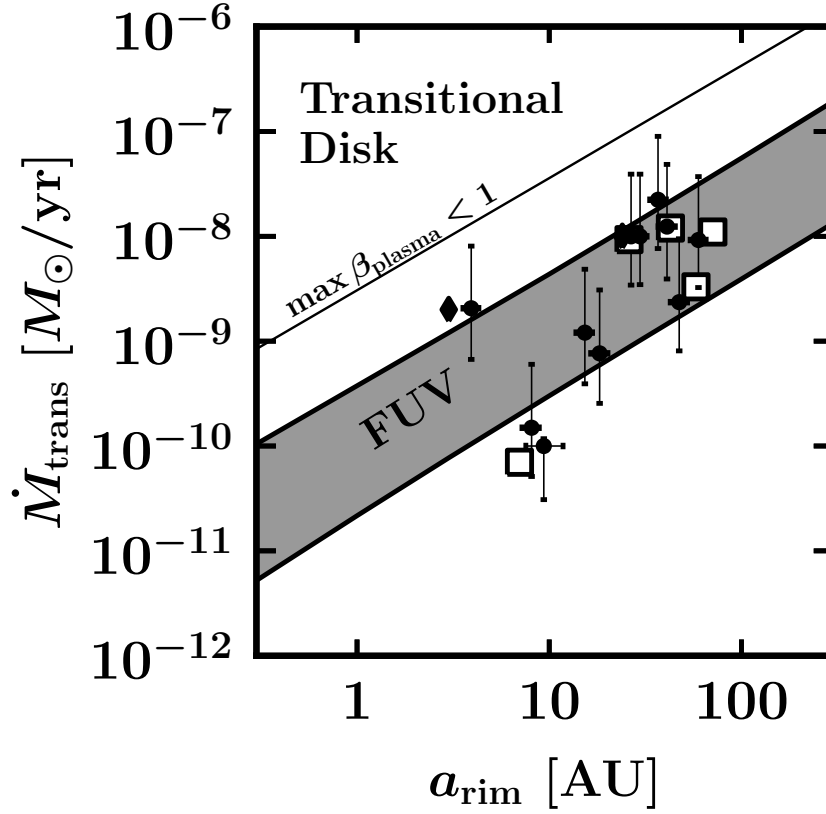


Figure 2.10 Same as Figure 2.8 but for transitional disks and showing only our FUV model. Observed rates are taken from Calvet et al. (2005, diamonds), Espaillat et al. (2007b, 2008, 2010, squares), and Kim et al. (2009, circles with error bars). Theoretical rates for FUV-ionized layers are computed with equation (2.26), with Σ^* recomputed using $\theta \sim 1$. The observed trend of increasing $\dot{M}_{\text{trans}}$ with increasing a_{rim} is reproduced (see, however, Section 4.2 of Kim et al. 2009 for how the observed trend could reflect other correlations between stellar mass M and a_{rim} , and M and $\dot{M}$).

accretion rates of transitional disks. Following CMC, we compute the accretion rate at the rim according to the formula

$$\dot{M}_{\text{trans}} \sim \frac{3M_{\text{rim}}}{t_{\text{diff}}} \sim 12\pi\Sigma^* \times \max \alpha \times \left(\frac{kT}{\mu}\right)^{3/2} \frac{a_{\text{rim}}^2}{GM} \quad (2.26)$$

where $M_{\text{rim}} \approx 4\pi ah\Sigma^*$ (h is the disk half-thickness), $t_{\text{diff}} \sim a_{\text{rim}}^2/\nu$ is the time for material to diffuse from a_{rim} to $a_{\text{rim}}/2$, and Σ^* is reinterpreted for transitional disk rims as the radial, not vertical, column of MRI-active material. We re-compute Σ^* for transitional disk rims by taking the grazing angle $\theta \sim 1$, as is appropriate for radiation which penetrates the rim wall at normal incidence.

Figure 2.10 shows accretion rates for transitional disk rims computed according to (2.26), overlaid with observations taken from Calvet et al. (2005), Espaillat et al. (2007b, 2008, 2010), and Kim et al. (2009). We have chosen these references and not others because they utilize disk models similar enough to each other to yield reasonably consistent hole radii (cf. Merín et al. 2010). Rim accretion by FUV ionization seems capable of reproducing the trend of increasing $\dot{M}_{\text{trans}}$ with increasing a_{rim} . However, there remains the problem of transporting the material that is dislodged from the rim over the decades in disk radius between the rim and the host star. Inside the rim, a conventional disk geometry applies, and we have already noted in that context that surface layer accretion rates at small radii are too small compared to those observed. For example, according to Figure 2.10, an accretion rate of $\dot{M}_{\text{trans}} \sim 10^{-9}M_{\odot}/\text{yr}$ may be initiated at $a_{\text{rim}} = 30$ AU, but according to Figure 2.8, this same accretion rate cannot be sustained by the MRI inside a radius of ~ 3 AU. Multiple planets could solve this problem by shuttling gas inward (PBC11; Zhu et al. 2011).

Chapter 3

Atmospheric Heat Redistribution on Hot Jupiters

Daniel Perez-Becker & Adam P. Showman (2013), ApJ, 776:134.¹

Abstract

Infrared light curves of transiting hot Jupiters present a trend in which the atmospheres of the hottest planets are less efficient at redistributing the stellar energy absorbed on their daysides—and thus have a larger day-night temperature contrast—than colder planets. To this day, no predictive atmospheric model has been published that identifies which dynamical mechanisms determine the atmospheric heat redistribution efficiency on tidally locked exoplanets. Here we present a shallow water model of the atmospheric dynamics on synchronously rotating planets that explains why heat redistribution efficiency drops as stellar insolation rises. Our model shows that planets with weak friction and weak irradiation exhibit a banded zonal flow with minimal day-night temperature differences, while models with strong irradiation and/or strong friction exhibit a day-night flow pattern with order-unity fractional day-night temperature differences. To interpret the model, we develop a scaling theory which shows that the timescale for gravity waves to propagate horizontally over planetary scales, τ_{wave} , plays a dominant role in controlling the transition from small to large temperature contrasts. This implies that heat redistribution is governed by a wave-like process, similar to the one responsible for the weak temperature gradients in the Earth’s tropics. When atmospheric drag can be neglected, the transition from small to large day-night temperature contrasts occurs when $\tau_{\text{wave}} \sim \sqrt{\tau_{\text{rad}}/\Omega}$, where τ_{rad} is the radiative relaxation time and Ω is the planetary rotation frequency. Alternatively, this transition criterion can be expressed as $\tau_{\text{rad}} \sim \tau_{\text{vert}}$, where τ_{vert} is the timescale for a fluid parcel to move vertically over the difference in day-night thickness. These results subsume the more widely used timescale comparison for estimating heat redistribution efficiency between τ_{rad} and the horizontal day-night advection

¹The Astrophysical Journal, 776:134, 2013 October 20

timescale, τ_{adv} . Only because $\tau_{\text{adv}} \sim \tau_{\text{vert}}$ for hot Jupiters does the commonly assumed timescale comparison between τ_{rad} and τ_{adv} yield approximately correct predictions for the heat redistribution efficiency.

3.1 Introduction

Stellar radial velocity surveys have discovered a class of extrasolar planets whose masses are comparable to Jupiter’s, but that orbit their host stars at distances less than 0.1 AU (Lovis & Fischer 2010). In contrast to Jupiter, which roughly receives as much power from stellar irradiation as it releases from its interior, these “hot Jupiters” have power budgets dominated by external irradiation. In addition, the strong tidal torques are presumed to lock them into a state of synchronous rotation (Guillot et al. 1996), forcing them to have permanent daysides and nightsides. The extreme insolation, fixed day-night thermal forcing pattern, and slow rotation rates of hot Jupiters provide a unique laboratory to explore the atmospheric dynamics of giant planets under conditions not present in the Solar System.

Continuous photometric observations of eclipsing (transiting) hot Jupiters, over half an orbit or longer, allow for precise determination of the flux emitted from the planet as it presents different phases to us. Emission from the nightside is visible around the time the planet transits the stellar disk, while the dayside is visible just before and after the planet passes behind the star (secondary eclipse). The planetary contribution to the total flux received is up to $\sim 0.1\%$ – 0.3% at infrared wavelengths and can be isolated because the precise stellar flux is known from observations during secondary eclipse. The longitudinal atmospheric temperature profile of the planet is inferred from the orbital flux variations. These light-curve observations are currently one of the most powerful tools to constrain the atmospheric dynamics of these planets (for a review, see Deming & Seager 2009). So far, visible and infrared light curves for at least 11 hot Jupiters have been published.

Figure 3.1 presents the fractional day-night flux differences, obtained from such light curves, for transiting hot Jupiters on near-circular orbits.² Specifically, we plot the flux difference between the dayside and nightside, divided by the dayside flux, as a function of the planet’s global-mean equilibrium temperature calculated assuming zero albedo. Interestingly, these light-curve observations suggest an emerging trend wherein planets that receive greater stellar insolation—and therefore have hotter mean temperatures—exhibit greater fractional flux contrasts between the dayside and the nightside. Cool planets like HD 189733b (Knutson et al. 2007, 2009b, 2012) and HD 209458b (Cowan et al. 2007) have only modest day-night flux differences. Planets receiving intermediate flux, such as HD 149026b (Knutson et al. 2009a), exhibit intermediate day-night flux differences. And the most strongly irradiated

²In practice, we are dividing the amplitude of the phase curve variation by the depth of the secondary eclipse relative to the maximum in the phase curve, a procedure that only works for transiting hot Jupiters. Light curves of non-transiting hot Jupiters such as Ups And b (Harrington et al. 2006; Crossfield et al. 2010), 51 Peg b, and HD 179949b (Cowan et al. 2007) are therefore not included in Figure 3.1. We also exclude planets on highly eccentric orbits such as HD 80606b (Laughlin et al. 2009) and HAT-P-2b (Lewis et al. 2013).

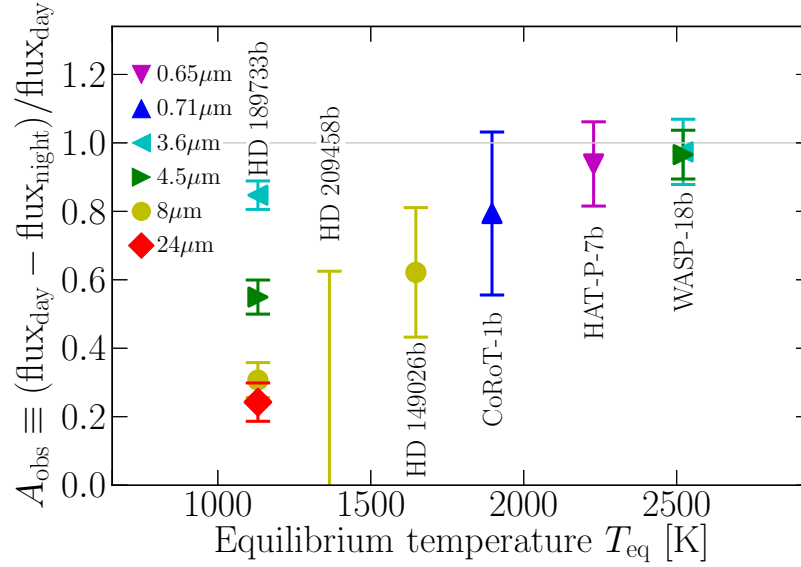


Figure 3.1 Fractional day-night infrared flux variations A_{obs} versus global-mean equilibrium temperature T_{eq} for hot Jupiters with measured light curves. Here, A_{obs} is defined, at a particular wavelength, as the flux difference between dayside and nightside divided by the dayside flux. The equilibrium temperature is defined as $T_{eq} = [F_*/(4\sigma)]^{1/4}$, where F_* is the stellar flux received by the planet and σ is the Stefan-Boltzmann constant. Planets with hotter mean temperatures have larger day-night flux variations, indicating larger longitudinal temperature gradients at the photosphere. Colored symbols with error bars are from published observations (Knutson et al. 2007, 2009a,b, 2012; Cowan et al. 2007, 2012; Borucki et al. 2009; Maxted et al. 2013). The error bar without a data point for HD 209458b indicates that only an upper limit was published (Cowan et al. 2007).

planets, such as HAT-P-7b (Borucki et al. 2009), WASP-18b (Maxted et al. 2013), and WASP-12b (Cowan et al. 2012) exhibit fractional day-night flux differences close to unity.³ This trend is consistent with a comparison of secondary eclipse depths with predicted dayside equilibrium temperatures in a broad sample of ~ 24 systems performed by Cowan & Agol (2011). Together, Figure 3.1 and the analysis of Cowan & Agol (2011) suggest that the day-night temperature difference at the photosphere (expressed as a ratio to the dayside temperature) increases with effective temperature, and approaches unity for hot Jupiters with effective temperatures of ~ 2200 K or greater.

What causes the trend observed in Figure 3.1? The commonly invoked explanation has been that the efficiency with which hot Jupiters redistribute heat is determined by the extent to which atmospheric winds transport hot gas across planetary scales faster than it takes the gas to radiate its heat into space. This balance is typically described by a comparison between two characteristic timescales, the timescale for winds to advect gas horizontally across the planet, τ_{adv} , and the timescale for gas to reach local radiative equilibrium, τ_{rad} . Showman & Guillot (2002) first suggested that hot Jupiters will exhibit large fractional day-night temperature differences when $\tau_{\text{rad}} \ll \tau_{\text{adv}}$ and small fractional day-night temperature differences when $\tau_{\text{rad}} \gg \tau_{\text{adv}}$. They pointed out that the radiative time constant decreases strongly with increasing temperature, and they presented a heuristic theory suggesting that planets with greater characteristic radiative heating rates will exhibit larger fractional day-night temperature differences. Subsequently, a wide range of authors proposed that this timescale comparison could describe the pressure-, opacity-, and insolation-dependence of the day-night temperature differences on hot Jupiters (Cooper & Showman 2005; Showman et al. 2008; Fortney et al. 2008; Rauscher & Menou 2010; Heng et al. 2011; Cowan & Agol 2011; Perna et al. 2012). This picture is based on the expectation, from both three-dimensional (3D) circulation models and observations, that hot Jupiters should develop fast atmospheric jets capable of transporting heat over planetary scales.

However, this timescale comparison is neither self-consistent nor predictive, as τ_{adv} is not known a priori and depends on many atmospheric parameters, including τ_{rad} . In particular, one cannot even evaluate the criterion under given external forcing conditions unless one already has a theory for (or simulations of) the atmospheric flow itself. No such theory for the atmospheric circulation generally, or the day-night temperature difference specifically, has ever been presented. Moreover, the comparison between τ_{adv} and τ_{rad} neglects any role for other important timescales, including timescales for wave propagation, frictional drag, and planetary rotation. These timescales almost certainly influence the day-night temperature difference, among other aspects of the circulation. More generally, it is crucial to emphasize that the ultimate goal is not simply to obtain a timescale comparison for the transition between small and large day-night temperature difference, but rather to obtain a predictive

³The HAT-P-7b point in Figure 3.1 lies in the visible, so there remains ambiguity about whether the dayside is bright due to thermal emission or due to reflected starlight. However, an unpublished full-orbit light curve of HAT-P-7b at $3.6 \mu\text{m}$ from Spitzer exhibits a phase-curve amplitude comparable to the secondary-eclipse depth (Knutson 2011), where emission is predominantly thermal rather than reflected starlight. This demonstrates that HAT-P-7b indeed exhibits large fractional day-night temperature contrasts at the photosphere.

theory for the day-night temperature difference itself, valid across the full parameter space.

To reiterate the role that other timescales can play, consider the Earth’s tropics. Over most of the tropics, the horizontal temperature gradients are weak, and the radiative cooling to space is balanced not by horizontal advection but by *vertical* advection (e.g., [Sobel et al. 2001](#))—raising questions about the relevance of the horizontal advection time for this system. In fact, the longitudinal variation of the temperature structure in the tropics is primarily regulated by adjustment due to gravity waves (see [Showman et al. 2013b](#), for a review). Moist convection, gradients in radiative heating, and other processes can lead to horizontal temperature and pressure variations that in turn cause the generation of gravity waves. These waves induce horizontal convergence/divergence that, via mass continuity, causes air columns to stretch or contract vertically. This coherent vertical motion pushes isentropes up or down, and if the Coriolis force is weak (as it is in the tropics) and the waves are able to radiate to infinity, the final state is one with flat isentropes. A state with flat isentropes is a state with constant temperature on isobars; therefore, this wave-adjustment process acts to erase horizontal temperature differences. As emphasized by [Bretherton & Smolarkiewicz \(1989\)](#) and others, this adjustment process occurs on characteristic timescales comparable to the time for gravity waves to propagate over the length scale of interest. A key point is that, in many cases, this wave-propagation timescale is much shorter than the timescales for horizontal advection or mixing (e.g., between a cumulus cloud and the surrounding environment). Such wave propagation is not only local but can also occur over planetary scales, on both Earth (e.g., [Matsuno 1966](#); [Gill 1980](#); [Bretherton & Sobel 2003](#)) and exoplanets ([Showman & Polvani 2011](#)). Therefore, it is reasonable to expect that this wave-adjustment process will play a key role in the regulation of horizontal temperature differences on exoplanets—and, as a corollary, that the horizontal wave-propagation timescale will be important.

Motivated by these issues, our goals are to (1) quantify in numerical simulations how the day-night temperature difference on synchronously rotating exoplanets depends on external forcing parameters, (2) develop a quantitative theory for this behavior, and (3) isolate the dynamical mechanisms responsible for controlling the day-night temperature differences. A natural spin-off of this undertaking will be a criterion, expressed in terms of the various timescales, for the transition between small and large day-night temperature differences.

Because our emphasis is on developing a basic understanding, we adopt perhaps the simplest dynamical model that can capture the key physics: a shallow-water model representing the flow in the observable atmosphere. This means that our model will exclude many details important on hot Jupiters, but it will allow us to identify the key dynamical processes in the cleanest possible environment. We view this as a prerequisite to understanding more realistic systems.

Section [3.2](#) introduces our dynamical model. Section [3.3](#) presents numerical solutions of the atmospheric circulation and explores the dependence on external forcing parameters. Section [3.4](#) presents an analytic scaling theory of the day-night differences explaining the behavior of our numerical solutions. Section [3.5](#) provides a dynamical interpretation of the behavior in our theory and simulations. Section [3.6](#) applies this understanding to observations, and in particular we show that our model can explain the trend observed in Figure [3.1](#).

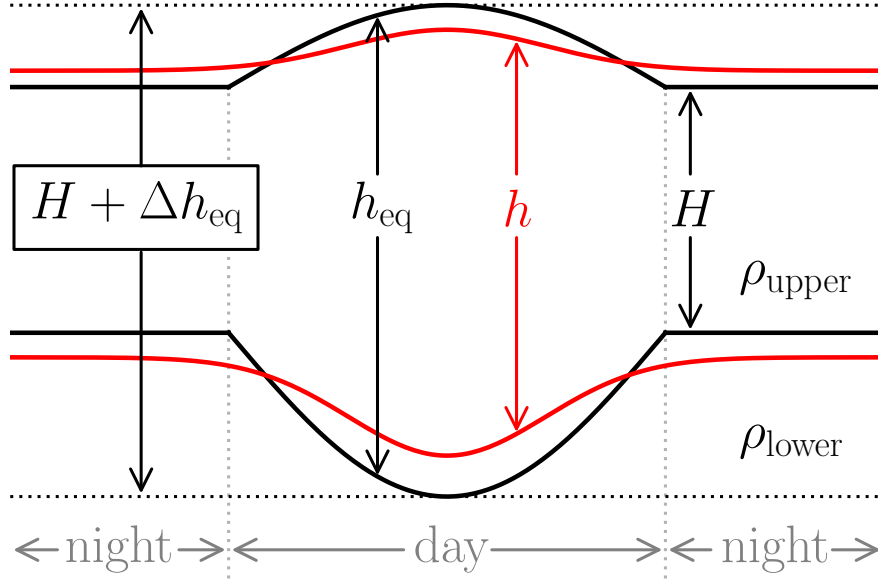


Figure 3.2 The two-layer shallow water model. A layer of fluid with constant density ρ_{upper} and variable thickness h floats above an infinitely deep layer of fluid with higher constant density ρ_{lower} . The fluid thickness h (red solid lines) represents the atmospheric mass column that has an entropy above a given reference value. Atmospheric heating thus will locally increase h , while cooling will locally reduce h . Radiative relaxation tends to restore the fluid to a temporally invariant profile, h_{eq} (black solid lines), over a radiative relaxation timescale τ_{rad} . The radiative equilibrium profile, defined in equation (3.3), varies in height from H on the nightside to $H + \Delta h_{\text{eq}}$ at the substellar point.

Finally, Section 3.7 concludes.

3.2 The Model

Global, 3D circulation models (GCMs) of planetary atmospheres involve many interacting processes, which makes it difficult to identify which dynamical mechanisms are dominating the solution. Simplified models have therefore played an important role in atmospheric dynamics of giant planets, both solar (Dowling & Ingersoll 1989; Cho & Polvani 1996; Scott & Polvani 2007; Showman 2007) and extrasolar (Cho et al. 2003, 2008; Showman & Polvani 2011; Showman et al. 2013a).

We study atmospheric heat transport on hot Jupiters with an idealized two-layer shallow-water model (see Figure 3.2). The buoyant upper layer of the model, having constant density ρ_{upper} and variable thickness h , represents the meteorologically active atmosphere of the planet. The infinitely deep bottom layer has a higher constant density ρ_{lower} and represents

the convective interior of the planet. We assume isostasy: the total mass column above a given depth in the lower layer is constant. For an infinitely deep lower layer, isostasy is guaranteed for baroclinic waves—those where the upper free surface and the interfacial boundary bow in opposite directions (Figure 3.2). Isostasy captures baroclinic modes but screens out barotropic ones (see Section 6.2 of Gill 1982); isostasy implies that there are no horizontal pressure gradients in the lower layer and therefore no horizontal velocities there. However, there can be vertical velocities; indeed mass will be transferred between the two layers.

The equations governing the upper layer are

$$\frac{d\mathbf{v}}{dt} + g\nabla h + f\mathbf{k} \times \mathbf{v} = \mathbf{R} - \frac{\mathbf{v}}{\tau_{\text{drag}}}, \quad (3.1)$$

$$\frac{\partial h}{\partial t} + \nabla \cdot (\mathbf{v}h) = \frac{h_{\text{eq}}(\lambda, \phi) - h}{\tau_{\text{rad}}} \equiv Q, \quad (3.2)$$

where $\mathbf{v}(\lambda, \phi, t)$ is the horizontal velocity, $h(\lambda, \phi, t)$ is the local thickness, t is time, g is the reduced gravity,⁴ $f = 2\Omega \sin \phi$ is the Coriolis parameter, $\mathbf{k}$ is the vertical unit vector, ∇ is the horizontal gradient operator, Ω is the planetary rotation frequency, and (λ, ϕ) are the longitudinal and latitudinal angles. The term $\mathbf{R}$ in equation (3.1) accounts for the momentum advected with mass transfer between layers and is quantified below. Here $d/dt \equiv \partial/\partial t + \mathbf{v} \cdot \nabla$ is the time derivative following the flow (this includes curvature terms in spherical geometry).

Our equations apply to a two-layer system with a free upper surface. For a derivation, see equations (3.37) and (3.38) in Vallis (2006), which are identical to our equations with the exception of the source terms, which we explain below. Note that their momentum equation is written in terms of the height of the lower layer (their η_1) above some reference level, whereas our equations are expressed in terms of the upper layer thickness (h). Our momentum equation can be derived from theirs by noting that isostasy in the lower layer requires $\nabla \eta_1 = -\nabla h \rho_{\text{upper}}/\rho_{\text{lower}}$.

In local radiative equilibrium, the height field $h(\lambda, \phi, t) = h_{\text{eq}}(\lambda, \phi)$, with

$$h_{\text{eq}} = \begin{cases} H + \Delta h_{\text{eq}} \cos \lambda \cos \phi & \text{on the dayside} \\ H & \text{on the nightside,} \end{cases} \quad (3.3)$$

where H is the (flat) nightside thickness and Δh_{eq} is the difference in radiative-equilibrium thickness between the substellar point and the nightside (see Figure 3.2). The expression adopts a substellar point at $(\lambda, \phi) = (0^\circ, 0^\circ)$. The planet is assumed to be synchronously rotating so that $h_{\text{eq}}(\lambda, \phi)$ remains stationary. Note that h_{eq} represents the two-dimensional

⁴The reduced gravity is the local gravitational acceleration times the fractional density difference between the upper and lower layers, $g \equiv GM_{\text{planet}}/a^2 \times (\rho_{\text{lower}} - \rho_{\text{upper}})/\rho_{\text{lower}}$, where G is the gravitational constant, M_{planet} is the planet mass, and a is the planet radius. Note that our definition of g differs from g' in Vallis (2006, see their Section 3.2), with $g = g' \times \rho_{\text{lower}}/\rho_{\text{upper}}$. The reduced gravity measures the restoring force at the interface between the two layers. When both layers have the same density, no pressure gradient forces exist and $g = 0$. Similarly, when $\rho_{\text{upper}} \ll \rho_{\text{lower}}$, g will equal the full gravitational acceleration.

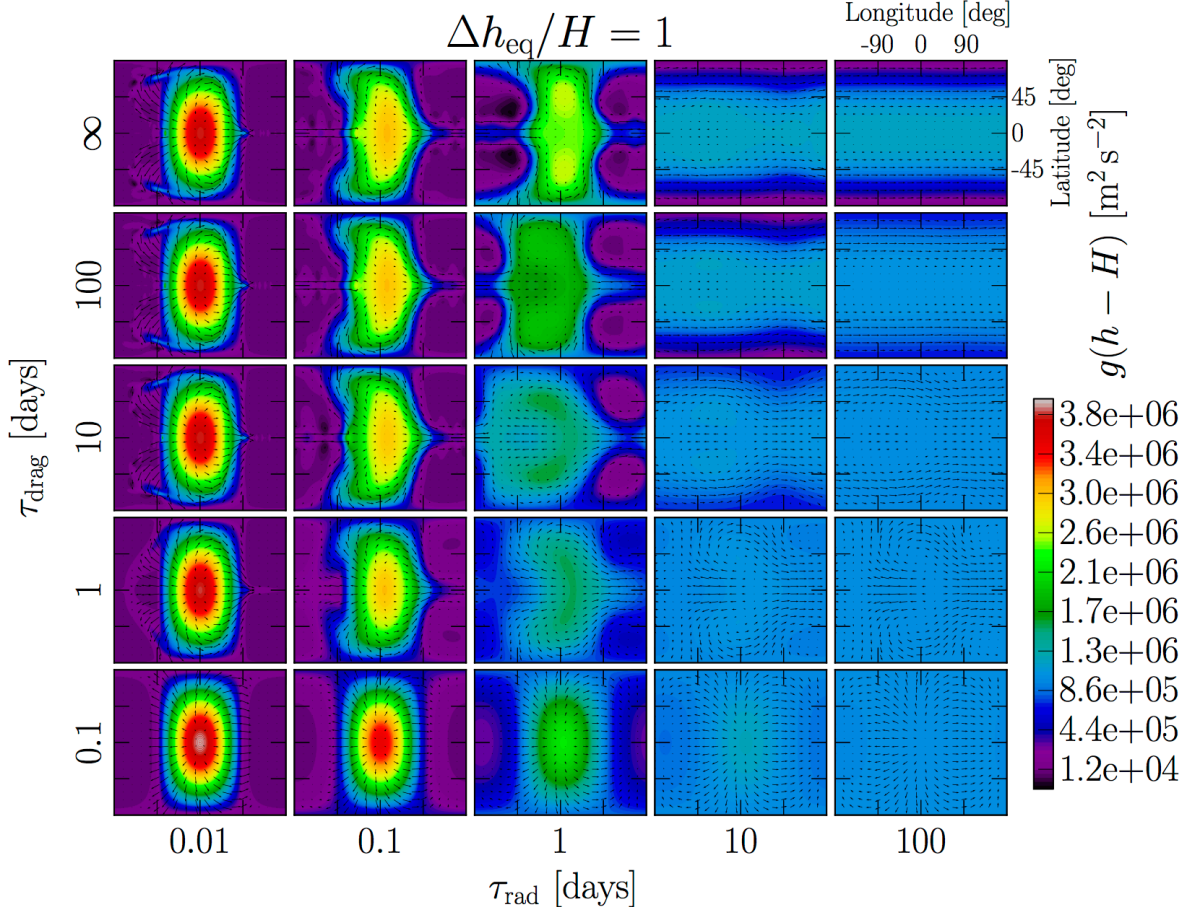


Figure 3.3 Equiarectangular maps of steady-state geopotential (gh) contours for the equilibrated solutions of the shallow-water model for a fractional forcing amplitude of $\Delta h_{\text{eq}}/H = 1$. We have subtracted the constant value of $gH = 4 \times 10^6 \text{ m}^2 \text{ s}^{-2}$ from each panel. Overplotted are vector fields of the steady-state winds. Each panel in the grid was computed for a different combination of radiative and drag timescales, τ_{rad} and τ_{drag} , expressed in Earth days. Panels share the same scale for the geopotential, but wind speeds are normalized independently in each panel. The substellar point is located at the center of each panel, at a longitude and latitude of $(0^\circ, 0^\circ)$. Short radiative timescales result in steady-state gh profiles dominated by stellar forcing with a hot dayside and a cold nightside (see equation (3.3)), while the atmosphere relaxes to a constant gh for long values of τ_{rad} . In contrast, the dependence of gh on τ_{drag} is weaker. The atmospheric circulation shifts from a zonal jet pattern at long τ_{rad} and τ_{drag} to day-to-night flow when either τ_{rad} or τ_{drag} is reduced, as explained in [Showman et al. \(2013a\)](#).

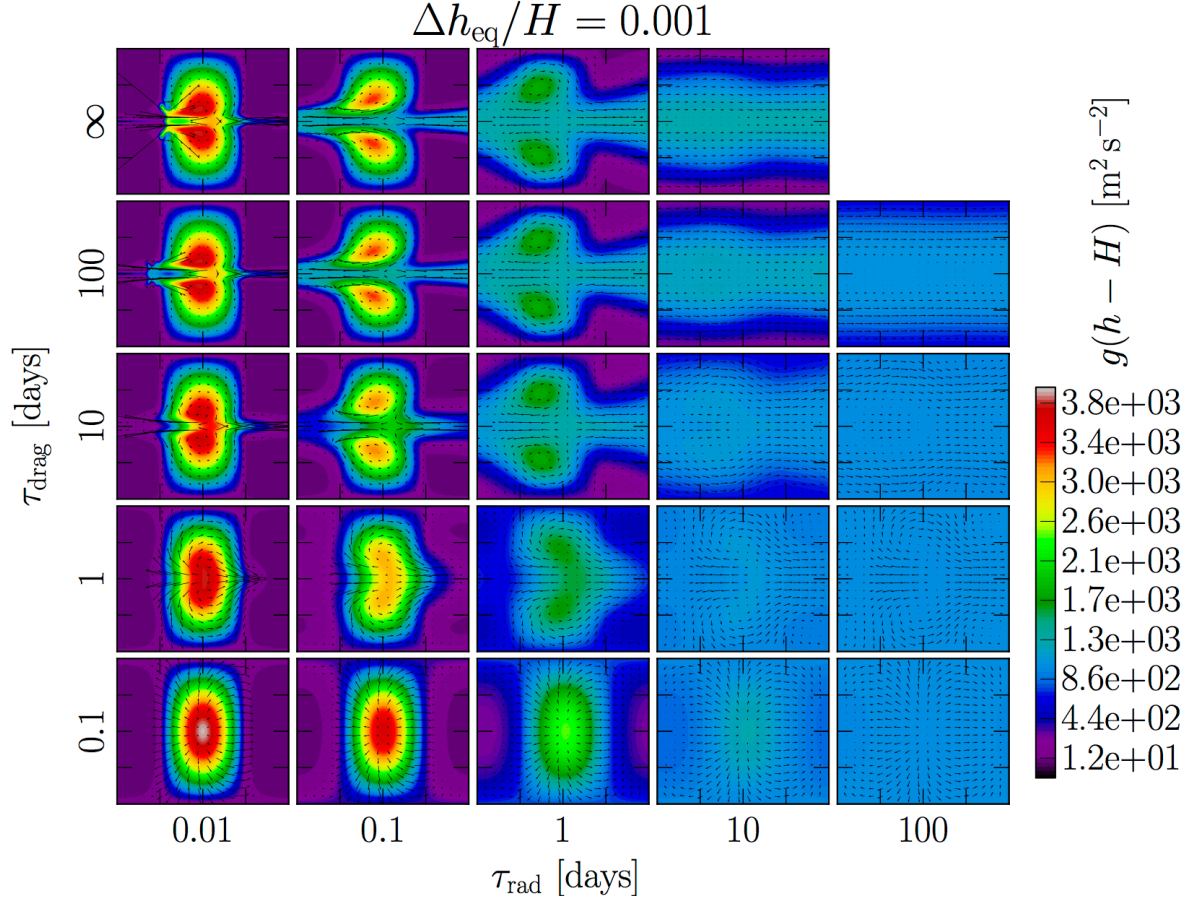


Figure 3.4 Same as Figure 3.3, but at a low forcing amplitude of $\Delta h_{\text{eq}}/H = 0.001$. Note that the dynamic range of the gh -contour color-scale has been reduced by a factor of 1000 compared to Figure 3.3. Wind speeds have also scaled down between figures. Comparing individual panels between Figures 3.3 and 3.4 reveals that even though the details of the solutions depend on $\Delta h_{\text{eq}}/H$, the characteristic value of the fractional day-night geopotential difference (normalized to the radiative equilibrium difference) remains largely unchanged.

height field set by local radiative equilibrium, whereas Δh_{eq} is its maximum deviation with respect to the layer thickness H at the nightside of the planet.

Equation (3.2) indicates that when the upper layer is not in radiative equilibrium, mass will be transferred between the layers, increasing or decreasing h . The relaxation toward equilibrium occurs over a radiative timescale τ_{rad} —a free parameter in the model. We can understand this mass transfer over a radiative timescale in the context of a 3D atmosphere. In a 3D context, h represents the amount of fluid having a specific entropy greater than a certain reference value—i.e., h is a proxy for the mass column above an isentrope. In regions that are heated to return to local radiative equilibrium ($Q > 0$), fluid acquires entropy and rises above the reference isentrope, increasing h . Likewise, in regions that cool ($Q < 0$), fluid sinks below the reference isentrope, and h decreases.

Mass transfer between the horizontally static lower layer and the active upper layer will affect not only the local height of the upper layer but also its momentum. The momentum advected with mass transfer between layers is accounted for by $\mathbf{R}$ —the vertical transport term. In regions where gas cools ($Q < 0$), mass is locally transferred from the upper layer to the lower layer. This process should not affect the specific momentum of the upper layer, so $\mathbf{R} = 0$ when $Q < 0$. In regions where gas is heated ($Q > 0$) mass is transferred from the lower layer into the upper layer. The added mass carries angular momentum with it and changes the specific angular momentum of the upper layer. In our case the added mass has no horizontal velocity so that the column-integrated value of $\mathbf{v}h$ should remain unchanged. As explained by [Showman & Polvani \(2011\)](#), one can obtain an expression for $\partial(\mathbf{v}h)/\partial t$ by adding $\mathbf{v}$ times the continuity equation to h times the momentum conservation equation. When $\partial(\mathbf{v}h)/\partial t = 0$, terms involving heating and cooling ($h\mathbf{R} + \mathbf{v}Q$) also have to vanish, which yields $\mathbf{R} = -\mathbf{v}Q/h$ in regions of heating. This expression for $\mathbf{R}$ is also used in [Shell & Held \(2004\)](#), [Showman & Polvani \(2010, 2011\)](#), and [Showman et al. \(2013a\)](#).

Finally, we parameterize atmospheric drag with Rayleigh friction, $-\mathbf{v}/\tau_{\text{drag}}$, where τ_{drag} is a specified characteristic drag timescale. Potential sources of atmospheric drag are hydrodynamic shocks and Lorentz-force drag. The latter is caused by ion-neutral collisions induced by magnetic deflection of thermally ionized alkali metals. In this work we keep a general prescription for drag that is proportional to the flow velocity, at the cost of missing details of the individual physical processes (e.g., Lorentz-force drag only affects the flow component moving orthogonally to the local planetary magnetic field and strongly depends on the local gas temperature ([Rauscher & Menou 2013](#))).

We solve equations (3.1) and (3.2) in global, spherical geometry with the following parameter choices. We choose g and H such that the gravity wave speed is $\sqrt{gH} = (10 \text{ m s}^{-2} \times 400 \text{ km})^{1/2} = 2 \text{ km s}^{-1}$.⁵ For our fiducial hot Jupiter model, we set the relative

⁵Our model is specified by the product gH and does not require separate specification of g and H , as can be seen by multiplying equation (3.2) by g . Gravity waves in our shallow water model are a proxy for internal gravity waves that exist in stratified atmospheres. In contrast to shallow water—where waves propagate exclusively in the plane perpendicular to gravity—gravity waves in a stratified atmosphere can propagate in any direction. Nevertheless, the latitudinal forcing on hot Jupiters will primarily excite horizontally traveling gravity waves whose propagation speeds will be similar to $\sqrt{g_{\text{actual}}H_{\text{atm}}}$, where $g_{\text{actual}} \approx 10 \text{ m s}^{-2}$ is the full

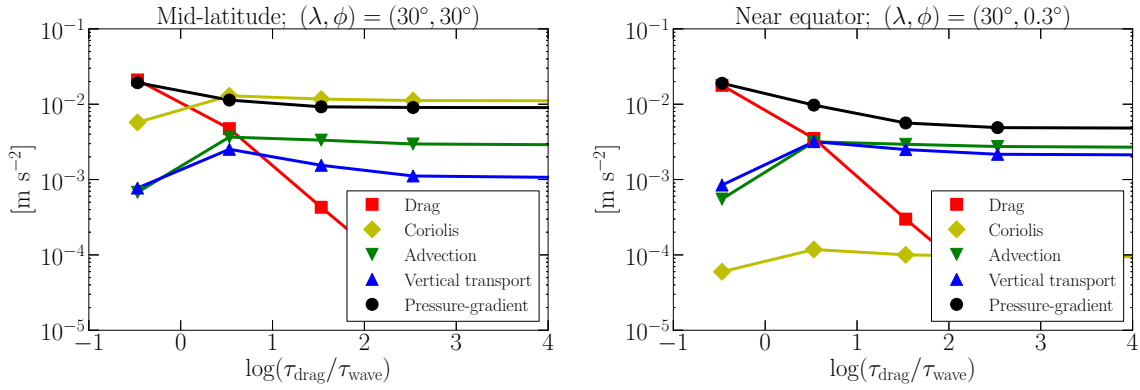


Figure 3.5 Absolute values of the zonal components of individual terms in the momentum equation (3.1) as a function of τ_{drag} . For both panels, $\Delta h_{\text{eq}}/H = 1$ and τ_{rad} is held constant at 0.1 days [$\log(\tau_{\text{rad}}/\tau_{\text{wave}}) = -0.5$]. The left panel is computed at a typical mid-latitude of the planet with coordinates $(\lambda, \phi) = (30^\circ, 30^\circ)$, while the right panel is computed near the equator $(\lambda, \phi) = (30^\circ, 0.3^\circ)$. Far away from the equator (left panel), the primary balance in the momentum equation is between the horizontal component of the pressure-gradient force and either the Coriolis force (weak-drag regime) or drag (drag-dominated regime). The transition from the weak-drag regime to the drag-dominated regime in the momentum balance occurs when $\tau_{\text{drag}} \sim 1/f$ (equation (3.20)). Near the equator (right panel) the primary term balancing the pressure-gradient force is either advection (together with the vertical transport term $\mathbf{R}$) or drag.

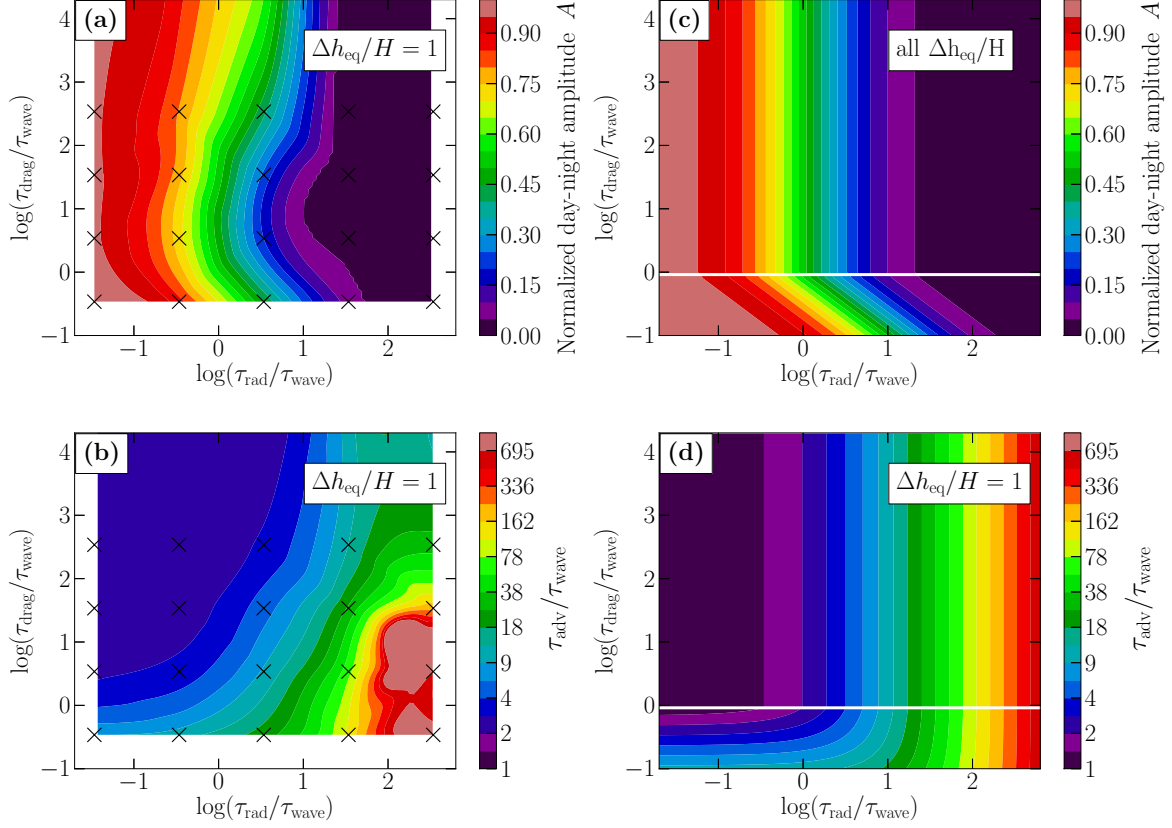


Figure 3.6 **Panel (a)**: contours of day-night height amplitude A (defined by equations (3.5)–(3.7)) as a function of τ_{rad} and τ_{drag} from our full numerical simulations with $\Delta h_{\text{eq}}/H = 1$. Values range from $A = 0$, which corresponds to a planet with a constant height field, to $A = 1$, which corresponds to a planet with $h = h_{\text{eq}}$ everywhere, i.e., zero heat redistribution. Simulations were carried out for a matrix of τ_{rad} and τ_{drag} values as described in Figure 3.3. These runs are marked by the $\times$ symbol; intervening values were calculated with a cubic-spline interpolation. In general terms, when τ_{rad} is short compared to τ_{wave} , atmospheres have a small heat redistribution efficiency. **Panel (b)**: contours of $\tau_{\text{adv}}/\tau_{\text{wave}} \equiv \sqrt{gH}/U$ from these same simulations as a function of τ_{rad} and τ_{drag} with $\Delta h_{\text{eq}}/H = 1$. Here U is the RMS value of atmospheric winds (both zonal and meridional) averaged over the entire planet, and $\sqrt{gH}$ is the gravity wave speed. Throughout the sampled space $\tau_{\text{adv}}/\tau_{\text{wave}} > 1$, with a minimum value of 2.1, colored purple. In the global mean, gravity waves always travel faster than winds. **Panel (c)**: contours of day-night geopotential amplitude A as a function of τ_{rad} and τ_{drag} as predicted by our analytical theory (equation (3.24)). Depending on the strength of atmospheric drag, different terms are being balanced in the momentum equation, yielding two distinct regions in the plot, separated by the solid white line, defined by equation (3.20). Drag is irrelevant in the upper region, but plays a significant role in the lower one. The simple timescale comparisons of our analytical theory broadly reproduce the results of the full numerical shallow-water model shown in panel (a). The scaling theory predicts that A -contours are independent of $\Delta h_{\text{eq}}/H$. **Panel (d)**: contours of $\tau_{\text{adv}}/\tau_{\text{wave}}$ as a function of τ_{rad} and τ_{drag} as predicted by our analytical theory (equation (3.17) for the drag-dominated regime below the white line and equation (3.19) for the weak drag regime above the white line). In contrast to A , contour values of $\tau_{\text{adv}}/\tau_{\text{wave}}$ depend on $\Delta h_{\text{eq}}/H$.

forcing amplitude $\Delta h_{\text{eq}}/H = 1$, implying that in radiative equilibrium temperature differences between day- and nightside vary by order unity. We will vary $\Delta h_{\text{eq}}/H$ down to 0.001 to understand dynamical mechanisms and to verify that our theoretical predictions are valid in the linear forcing limit. Our values for the rotation frequency $\Omega = 3.2 \times 10^{-5} \text{ s}^{-1}$ and a planetary radius of $a = 8.2 \times 10^7 \text{ m}$ are similar to those of HD 189733b. The characteristic wave travel timescale is

$$\tau_{\text{wave}} \sim \frac{L}{\sqrt{gH}} \sim 0.3 \text{ Earth days}, \quad (3.4)$$

where L is the characteristic horizontal length scale of the flow. For typical hot Jupiters, this is comparable to $L_{\text{eq}} \equiv (\sqrt{gH}a/2\Omega)^{1/2} \approx 5.1 \times 10^7 \text{ m}$, the equatorial Rossby deformation radius—a natural length scale that results from the interaction of buoyancy forces and Coriolis forces in planetary atmospheres. Note that for a typical hot Jupiter, L and L_{eq} happen to be of the same order as the planetary radius a (Showman & Guillot 2002; Menou et al. 2003; Showman & Polvani 2011). We explore how the solution depends on the characteristic damping timescales τ_{rad} and τ_{drag} , which are free parameters in the model. In contrast, the characteristic time over which a gas parcel is advected over a global scale, τ_{adv} , depends on the resulting wind profile and cannot be independently varied.

We solve equations (3.1) and (3.2) in spherical geometry with the Spectral Transform Shallow Water Model (STSWM) of Hack & Jakob (1992). The equations are integrated using a spectral truncation of T170, corresponding to a resolution of $0^\circ.7$ in longitude and latitude (i.e., a 512×256 grid in longitude and latitude). A ∇^6 hyperviscosity is applied to each of the dynamical variables to maintain numerical stability. The code adopts the leapfrog time-stepping scheme and applies an Asselin filter at each time step to suppress the computational mode. Models are integrated from an initially flat layer at rest until a steady state is reached. This system has been shown to be insensitive to initial conditions by Liu & Showman (2013). The models described here include those presented in Showman et al. (2013a), as well as additional models that we have performed for the present analysis.

In summary, the model contains three main input parameters: τ_{rad} , τ_{drag} , and $\Delta h_{\text{eq}}/H$. Our main goal in Section 3.3 is to determine the dependence of the equilibrated fractional day-night height difference on these parameters. In Section 3.4 we develop a simple analytic scaling theory that reproduces trends found in the numerical model.

gravitational acceleration (not the reduced gravity) and $H_{\text{atm}} \approx 400 \text{ km}$ is the atmospheric pressure scale height. Our nominal value for gH is chosen to match the propagation speed of the gravest mode of zonal gravity waves at the photosphere of hot Jupiters. Although it may be tempting to do so, we do not think of g as approximating g_{actual} (since the reduced gravity is conceptually distinct from and generally differs from the full gravity), nor do we think of H as approximating H_{atm} , since H is instead supposed to be a proxy for the thickness of material above an isentrope. Vertically propagating gravity waves are not captured by our model, but seem to be of lesser importance for the dynamics as they are not being excited by the longitudinal heating gradient—the driver of the system.

3.3 Numerical Solutions

3.3.1 Basic Behavior of the Solutions

As discussed in Section 3.1, observations indicate that as the stellar insolation increases, atmospheres transition from having small longitudinal temperature variations to having large day-night temperature contrasts. Our model solutions capture this transition, as shown in Figure 3.3. There, we plot the difference between the steady-state geopotential (gh) and the nightside geopotential at radiative equilibrium (gH) for twenty-five models performed at high amplitude ($\Delta h_{\text{eq}}/H = 1$) over a complete grid in τ_{rad} and τ_{drag} . Models are shown for all possible combinations of 0.01, 0.1, 1, 10, and 100 Earth days in τ_{rad} and 0.1, 1, 10, 100, and ∞ days in τ_{drag} . In the context of a 3D atmosphere, h represents the mass column above a reference isentrope; large h represents more material at high specific entropy (high temperature on an isobar). The stellar insolation is varied in the model by adjusting the damping timescales τ_{rad} and τ_{drag} . Generally, the higher the stellar flux (as measured by T_{eq}), the lower τ_{rad} will be. We will quantify the dependence of τ_{rad} on T_{eq} in Section 3.6.

The models in Figure 3.3 capture major transitions in both the structure of the flow and the amplitude of the day-night thickness contrast. When τ_{rad} is longer than one Earth-day, longitudinal gradients of gh are small. If τ_{drag} is also long compared to a day, the circulation primarily consists of east-west-aligned (zonal) flows varying little in longitude (upper right corner of Figure 3.3). Despite the lack of longitudinal variation, such models exhibit an equator-pole gradient in gh , albeit with an amplitude that remains small compared to the radiative-equilibrium gradient. When τ_{rad} is long but τ_{drag} is short, winds flow from the dayside to the nightside over both the eastern and western hemispheres, and gh varies little in either longitude or latitude (lower right corner of Figure 3.3). Intermediate values of τ_{rad} (e.g., ~ 1 day; middle column of Figure 3.3) lead to flows with greater day-night temperature differences and significant dynamical structure, including zonal-mean zonal winds that are eastward at the equator (i.e., equatorial superrotation). In contrast, when τ_{rad} is short (left column of Figure 3.3), the geopotential amplitude and morphology closely match the radiative forcing profile: a spherical bulge on the hot dayside and a flat, cold nightside (see equation 3.3). The circulation consists of strong airflow from day to night along both terminators. Showman & Polvani (2011) and Showman et al. (2013a) showed that much of the wind behavior in Figure 3.3 (and in many published 3D global circulation models of hot Jupiters) can be understood in terms of the interaction of standing, planetary-scale waves with the mean flow.

Many of the characteristics of the full solution can be understood by studying the model under weak forcing ($\Delta h_{\text{eq}}/H \ll 1$). In this limit, the day-night variations in h are much smaller than H , and terms in the shallow water equations exhibit their linear response. For example, the term $\nabla \cdot (\mathbf{v}h)$ in the continuity equation will behave approximately as $H\nabla \cdot \mathbf{v}$. The balance between $H\nabla \cdot \mathbf{v}$ and Q in the continuity equation is linear. If the balance in the momentum equation is also linear, then wind speeds $\mathbf{v}$ and amplitudes of h -variation should scale linearly with the forcing amplitude $\Delta h_{\text{eq}}/H$. In Figure 3.4 we present solutions of the

model forced at the low amplitude of $\Delta h_{\text{eq}}/H = 0.001$ for the same values for τ_{rad} and τ_{drag} as in Figure 3.3. Note that contour values for $g(h - H)$ have been scaled down by a factor of 1000 from those in Figure 3.3. At this low amplitude, the system responds linearly for most of parameter space. A comparison of Figures 3.4 and 3.3 demonstrates that the low-amplitude behavior is extremely similar to the high-amplitude behavior when τ_{drag} is short, but differs when τ_{drag} is long. The amplitude dependence under weak-drag conditions is greatest when τ_{rad} is short: at low amplitude, the mid-and-high latitudes are close to radiative equilibrium, whereas the equator exhibits almost no longitudinal variations in gh .

3.3.2 Physical Explanation for Forcing Amplitude Dependence

We now proceed to examine the reasons for these amplitude differences. Consider the principal force balances that determine the model solutions in the linear limit as a function of both drag and latitude.

For sufficiently strong drag ($\tau_{\text{drag}} \lesssim 1$ day), the balance in the momentum equation is primarily between the horizontal component of the pressure-gradient force—which drives the flow—and drag. This is a linear balance, and because the term balance in the continuity equation is likewise linear, we expect the h -field to scale with $\Delta h_{\text{eq}}/H$. This is indeed the case, as can be appreciated by the similarity of the lower panels of Figures 3.3 and 3.4.

When drag is reduced to the point where it becomes negligible ($\tau_{\text{drag}} \gtrsim 1$ day), other terms in the momentum equation have to balance the horizontal component of the pressure-gradient force. Which term plays the dominant role depends on the latitude of the planet. This is evident in Figure 3.4, where solutions in the upper left corner have a flat h -field at the equatorial region, while the h -field has a day-night amplitude that approaches radiative equilibrium at high latitudes. For $\Delta h_{\text{eq}}/H = 0.001$, winds are weak, and away from the equator, the Rossby number $Ro = U/fL \sim 0.001 \ll 1$, where U is a characteristic horizontal wind speed. As a result, the primary force balance away from the equator is between the Coriolis force and the horizontal component of the pressure-gradient force. This force balance is linear. Because the continuity equation is also linear, both h and $\mathbf{v}$ should scale with $\Delta h_{\text{eq}}/H$ at mid-latitudes. Comparing the upper rows of Figures 3.3 and 3.4 confirms that h -fields away from the equator scale with forcing amplitude in the weak-drag limit. Nevertheless, as the forcing amplitude is raised, wind speeds increase until $Ro \sim 1$ when $\Delta h_{\text{eq}}/H = 1$. Therefore, the advective term becomes comparable to the Coriolis term at high amplitudes. This results in differences, but no fundamental changes, in the flow structure at high latitudes. The linear dynamics in the weak drag regime are described in more detail in Appendix C of [Showman & Polvani \(2011\)](#).

In the weak-drag limit, why does the equator exhibit large fractional height variations at large forcing amplitude but only small fractional height variations at small forcing amplitude? At the equator, the Coriolis force vanishes ($Ro \gg 1$) and, if drag is weak, the force balance is between the horizontal component of the pressure-gradient force and advection. This is an inherently non-linear balance because advection scales with the square of the velocity. Thus there is no linear limit for the dynamical behavior at the equator, and h will not scale

linearly with $\Delta h_{\text{eq}}/H$. At $\Delta h_{\text{eq}}/H = 1$, the advection term is comparable to the Coriolis term at mid-latitudes. As a result, pressure (height) gradients remain as large at the equator as they are at mid-latitudes as can be appreciated in the upper rows of Figure 3.3. However, as the forcing amplitude is reduced to $\Delta h_{\text{eq}}/H = 0.001$, the advection term diminishes at a quadratic rate. To maintain balance with advection, the pressure-gradient force must also weaken as $(\Delta h_{\text{eq}}/H)^2$. Thus the height field becomes flat near the equator, as evidenced in the upper rows of Figure 3.4.

In Figure 3.5 we plot the magnitudes of the zonal components of all terms in the momentum equation as a function of τ_{drag} . The term balances discussed above are apparent. Figure 3.5 is computed for $\Delta h_{\text{eq}}/H = 1$, with τ_{rad} held constant at 0.1 days, while τ_{drag} is varied in the abscissa. Note that Figure 3.5 normalizes τ_{drag} with τ_{wave} —the (constant) wave travel time (equation (3.4)). The relevance of τ_{wave} will be explained in Section 3.4; from this point onward, we will express timescales in terms of τ_{wave} . The left panel plots the terms at a typical mid-latitude with coordinates $(\lambda, \phi) = (30^\circ, 30^\circ)$, while the right panel is for a point near the equator $(\lambda, \phi) = (30^\circ, 0^\circ.3)$. As noted before, the horizontal component of the pressure-gradient force away from the equator is balanced primarily against either the Coriolis force or drag. Near the equator, advection and vertical transport balance the pressure gradient when drag is weak.

3.3.3 Metric for the Day-Night Contrast

To compare our model solutions to the observed fractional flux variations of extrasolar planets, as well as to theoretical predictions, we need a measure similar to A_{obs} (see Figure 3.1). Our proxy for flux variations will be a day-night height difference A representative for the entire planet. We thus reduce each panel in Figures 3.3 and 3.4 to a single A , which we compute as follows. We start by evaluating the root-mean-square variations of h over circles of constant latitude, and normalize them to the values at radiative equilibrium:

$$A(\phi) = \left\{ \frac{\int_{-\pi}^{\pi} [h(\lambda, \phi) - \bar{h}(\phi)]^2 d\lambda}{\int_{-\pi}^{\pi} [h_{\text{eq}}(\lambda, \phi) - \bar{h}(\phi)]^2 d\lambda} \right\}^{1/2}, \quad (3.5)$$

where $\bar{h}(\phi)$ is the zonally averaged height at a given latitude

$$\bar{h}(\phi) = \frac{1}{2\pi} \int_{-\pi}^{\pi} h(\lambda, \phi) d\lambda. \quad (3.6)$$

We then average $A(\phi)$ over a 60° band centered at the equator

$$A = \frac{3}{\pi} \int_{-\pi/6}^{\pi/6} A(\phi) d\phi. \quad (3.7)$$

We find that for bands of width $\gtrsim 60^\circ$, A becomes insensitive to the range in latitudes used for averaging. As defined, A can vary from 0, when $h(\lambda, \phi, t) = \bar{h}(\phi)$ everywhere (corresponding

to an atmosphere without longitudinal height variations), to 1, when the height equals that imposed by radiative forcing, $h(\lambda, \phi, t) = h_{\text{eq}}(\lambda, \phi)$.

Figure 3.6(a) shows how model values of A depend on the choice of damping timescales τ_{rad} and τ_{drag} (which have been normalized to τ_{wave}) for high-amplitude models with $\Delta h_{\text{eq}}/H = 1$. Models with short radiative time constants exhibit large fractional day-night differences. When the drag timescale is short ($\tau_{\text{drag}}/\tau_{\text{wave}} < 1$), friction in the atmosphere starts to play a more important role in controlling the day-night height difference. We define the (normalized) horizontal advection timescale as

$$\tau_{\text{adv}}/\tau_{\text{wave}} \equiv \sqrt{gH}/U, \quad (3.8)$$

where U is the RMS-value of atmospheric winds (both zonal and meridional) averaged over the entire planet, and $\sqrt{gH}$ is the gravity wave speed. We show contours of $\tau_{\text{adv}}/\tau_{\text{wave}}$ as a function of τ_{rad} and τ_{drag} in Figure 3.6(b). Throughout the sampled space $\tau_{\text{adv}}/\tau_{\text{wave}} > 1$, with a minimum value of 2.1, colored purple. Since $\tau_{\text{wave}} < \tau_{\text{adv}}$ everywhere, the characteristic global-mean speed of gravity waves is always faster than the characteristic global-mean wind speed (however, note that, at the highest forcing amplitude, this is not always true locally everywhere over the globe).⁶

We can explain the qualitative behavior of these numerical solutions with an analytic theory in which we substitute dominant terms in the mass and momentum conservation equations with their order-of-magnitude counterparts. These analytical predictions, derived in Section 3.4, are showcased in Figures 3.6(c) and (d)—which are a reasonable match to Figures 3.6(a) and (b) showing the results of the numerical model. Two behaviors become apparent: above the white line of Figure 3.6(c), the contours are vertical, indicating that atmospheric drag does not affect the day-night temperature variation, a prediction that is in agreement with Figure 3.5. Below the white line, both the radiative and drag timescales affect A . Our scaling theory will confirm that τ_{wave} plays a central role in controlling the heat redistribution efficiency.

In Figure 3.7 we show the variation of A (as defined in equations 3.5–3.7) with forcing amplitude $\Delta h_{\text{eq}}/H$. The morphology of A is largely independent of forcing strength. The two upper panels show the solution close to the linear limit, where we expect A to be independent of $\Delta h_{\text{eq}}/H$.⁷ As $\Delta h_{\text{eq}}/H \rightarrow 1$, Ro increases to order-unity values at mid-latitudes. Nevertheless, the Coriolis force remains comparable to or greater than advection in regions away from the equator (see the left panel of Figure 3.5). Thus, even for a large forcing amplitude, A

⁶Supercritical flows with $U > \sqrt{gH}$ occur near the day-night terminator in the long τ_{drag} and short τ_{rad} limit (upper-left panels in Figure 3.3). When supercritical flow rams into slower moving fluid, hydraulic jumps develop which convert some kinetic energy into heat (Johnson 1997). We do not account for this source of heating as it is likely to be modest and is only present for solutions where A is already ~ 1 . Hydraulic jumps can also occur for supercritical flows in a stratified atmosphere. These are distinct from acoustic shocks that can develop in supersonic flow. The importance of acoustic shocks in modifying the photospheric temperature profile of hot Jupiters is still an open question, as most global simulations (including ours) do not capture the relevant physics, i.e., sound waves, and/or lack sufficient spatial resolution (Li & Goodman 2010).

⁷For low $\Delta h_{\text{eq}}/H$, the non-linear equatorial region—defined by $Ro \gtrsim 1$ —is thin and does not significantly contribute to the integral that makes up A in equation (3.7).

remains largely independent of $\Delta h_{\text{eq}}/H$, because the primary force balance is close to linear. In contrast, the day-night difference evaluated solely at the equator, $A_{\text{equator}} \equiv A(\phi = 0)$, will depend on $\Delta h_{\text{eq}}/H$. In the left panels of Figure 3.8 we show A_{equator} -contours from our full numerical simulations as a function of τ_{rad} , τ_{drag} , and $\Delta h_{\text{eq}}/H$. The right panels of Figure 3.8 show the A_{equator} -contours predicted by our scaling theory (Section 3.4). In the strong-drag regime (region below the white line), the behavior of A_{equator} in Figure 3.8 is relatively independent of amplitude, because the balance between drag and pressure-gradient forces is linear. In the weak-drag regime, the force balance is between advection and the horizontal component of the pressure-gradient force. For this non-linear balance there is no linear limit for the dynamical behavior at the equator, and A_{equator} depends on the forcing amplitude.

3.4 Analytic Theory for Day-Night Differences

Here we obtain an approximate analytic theory for the day-night thickness differences and wind speeds in the equilibrated steady states. Our full numerical nonlinear solutions exhibit steady behavior, so the partial time derivatives in both the continuity and momentum equations can be neglected. The mass conservation equation (3.2) can thus be approximated as

$$h(\nabla \cdot \mathbf{v}) + \mathbf{v} \cdot \nabla h \sim \frac{h_{\text{eq}} - h}{\tau_{\text{rad}}}. \quad (3.9)$$

On the right side, the quantity $h_{\text{eq}} - h$ gives the difference between the local radiative-equilibrium height field and the local height field (this difference is generally positive on the dayside and negative on the nightside). As shown in Figure 3.9,

$$|h_{\text{eq}} - h|_{\text{day}} + \Delta h + |h_{\text{eq}} - h|_{\text{night}} \sim \Delta h_{\text{eq}}, \quad (3.10)$$

where $|h_{\text{eq}} - h|_{\text{day}}$ is the characteristic (scalar) difference between $h_{\text{eq}}(\lambda, \phi)$ and $h(\lambda, \phi, t)$ on the dayside and $|h_{\text{eq}} - h|_{\text{night}}$ is the characteristic difference on the nightside. Define $|h_{\text{eq}} - h|_{\text{global}}$ to be the arithmetic average between $|h_{\text{eq}} - h|_{\text{day}}$ and $|h_{\text{eq}} - h|_{\text{night}}$.⁸ Then, to order of magnitude,

$$|h_{\text{eq}} - h|_{\text{global}} \sim \Delta h_{\text{eq}} - \Delta h. \quad (3.11)$$

We thus can approximate $h_{\text{eq}} - h$ in equation (3.9) with $\Delta h_{\text{eq}} - \Delta h$. The left side of equation (3.9) is, to order of magnitude, UH/L , where U is the characteristic horizontal wind speed and L is the characteristic horizontal lengthscale of the flow, which happens to be of order the planetary radius a .⁹ Thus, we have for the continuity equation

$$H \frac{U}{L} \sim \frac{\Delta h_{\text{eq}} - \Delta h}{\tau_{\text{rad}}}. \quad (3.12)$$

⁸The terms $|h_{\text{eq}} - h|_{\text{day}}$ and $|h_{\text{eq}} - h|_{\text{night}}$ are always of the same order because in steady state the rate at which mass is pumped into the active layer on the dayside ($\propto |h_{\text{eq}} - h|_{\text{day}}/\tau_{\text{rad}}$) has to equal the rate at which mass is removed from the nightside ($\propto |h_{\text{eq}} - h|_{\text{night}}/\tau_{\text{rad}}$).

⁹ This scaling is actually subtle. Consider the first term on the left-hand side of equation (3.9). In a flow where the Rossby number $Ro = U/fL \gtrsim 1$, the divergence simply scales as $\nabla \cdot \mathbf{v} \sim U/L$. In a flow where

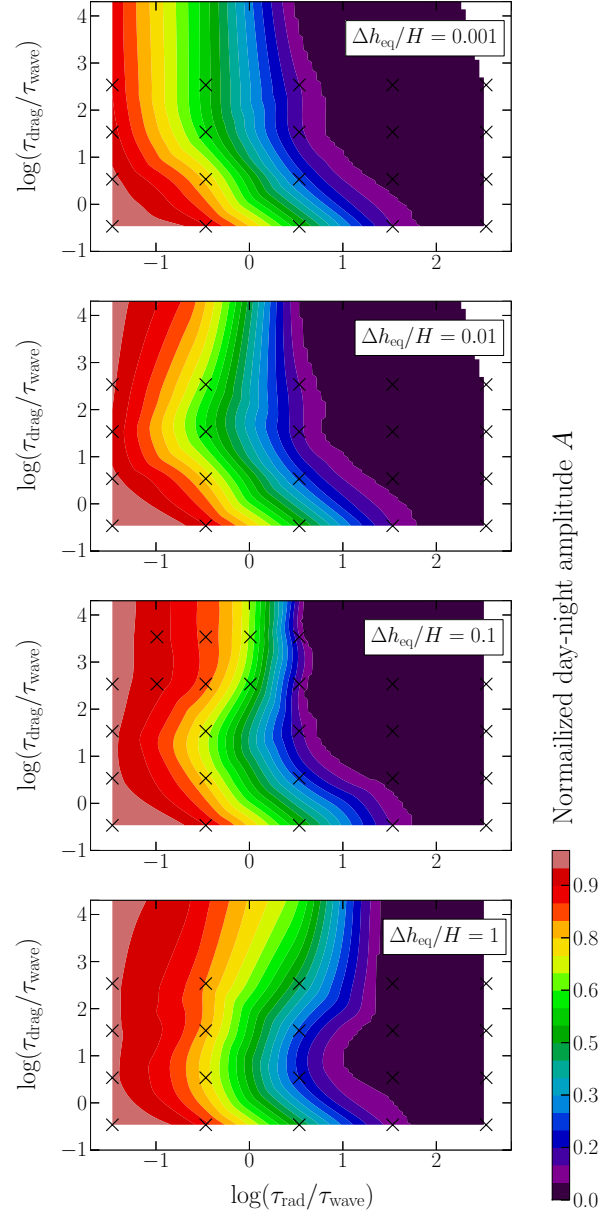


Figure 3.7 Contours of normalized day-night amplitude A (equations (3.5)–(3.7)) from our full shallow water simulations as a function of τ_{rad} and τ_{drag} . Each panel was computed for a different $\Delta h_{\text{eq}}/H$, ranging from 0.001 (top panel) to 1 (bottom panel). The lowermost panel is identical to Figure 3.6(a) and is repeated to facilitate a direct comparison with the other panels. Over the three orders of magnitude in $\Delta h_{\text{eq}}/H$ shown, the morphology of A seems to remain roughly unchanged, with nearly vertical A -contours in the upper half of each panel and slanted A -contours for the lower half, where the drag force dominates the momentum equation. All panels compare well to our analytical scaling theory, shown in Figure 3.6(c).

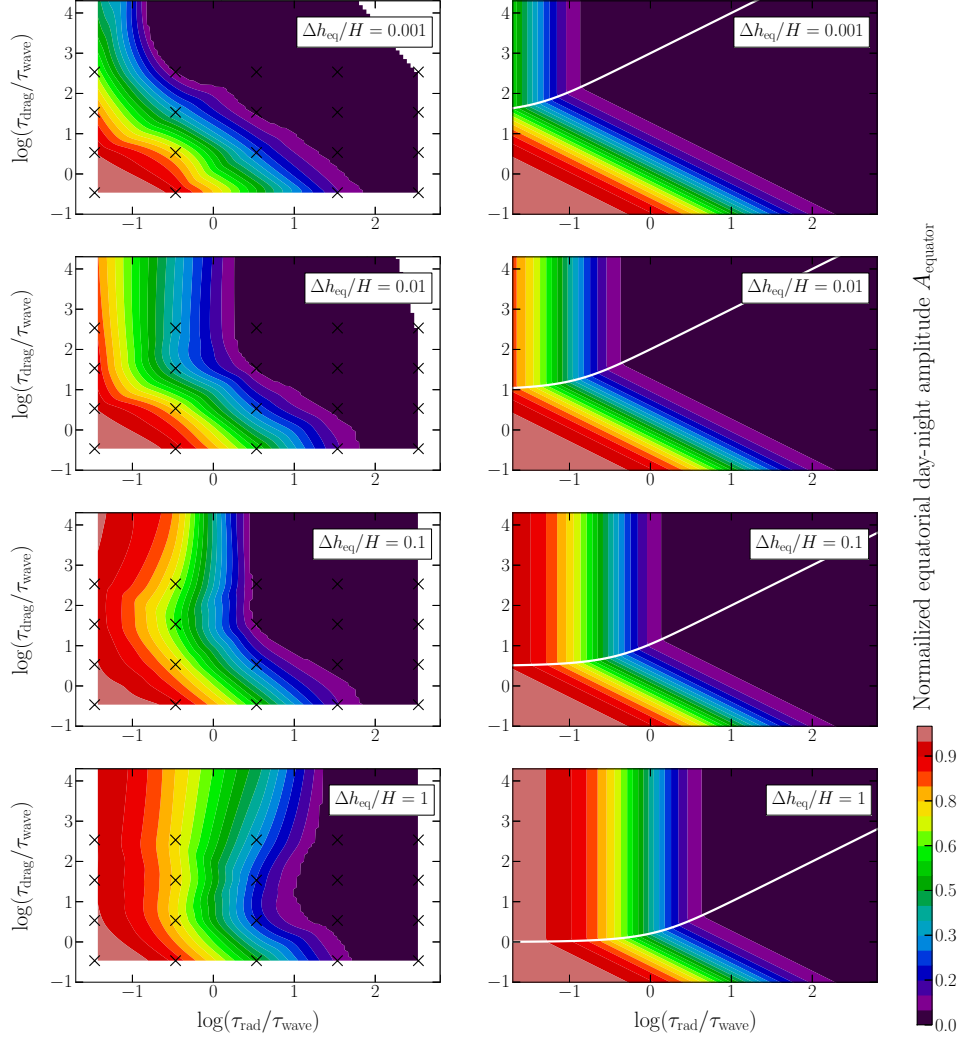


Figure 3.8 **Left-hand panels:** same as Figure 3.7, but for A_{equator} , the normalized day-night amplitude from our full shallow-water simulations evaluated only at the planetary equator. For $\Delta h_{\text{eq}}/H = 1$, A_{equator} contours are similar to those of A (shown in the lowermost panel of Figure 3.7). But as $\Delta h_{\text{eq}}/H$ is reduced, more of the τ_{rad} , τ_{drag} parameter space becomes drag-dominated, characterized by slanted A_{equator} contours. At $\Delta h_{\text{eq}}/H = 0.001$ (uppermost panel), most of the parameter space is drag dominated and the height field becomes flat at the equator when $\tau_{\text{drag}} \rightarrow \infty$ (as can be appreciated directly from model solutions shown in the upper rows of Figure 3.4). **Right-hand panels:** normalized day-night equatorial amplitude A_{equator} as predicted from our scaling theory. The white line marks the transition from the low-drag regime (where A_{equator} is given by equation (3.21)) to the drag-dominated regime (where A_{equator} is given by equation (3.16)). Our theory compares well with the numerical shallow water solutions shown in the left panels of this figure.

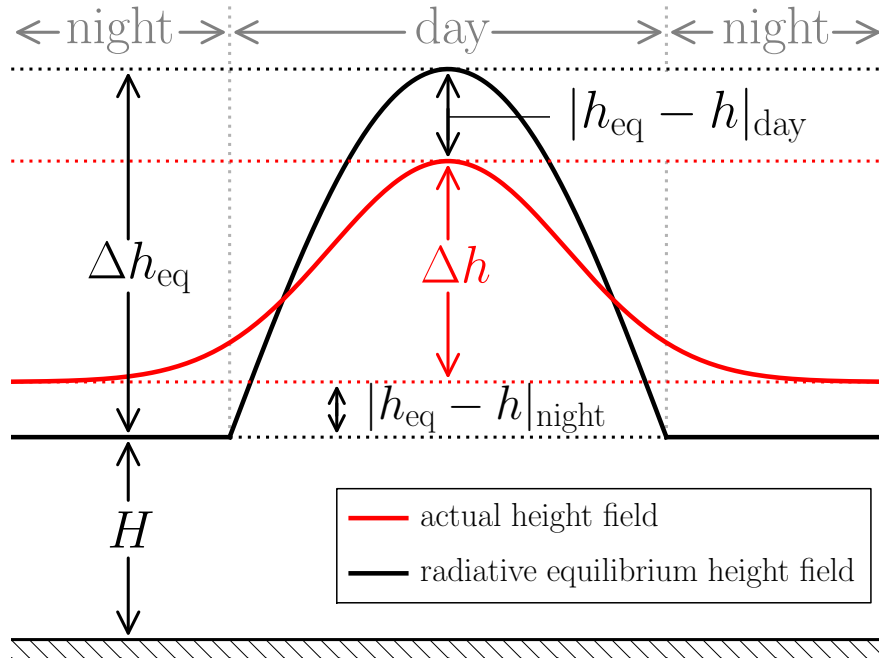


Figure 3.9 Simplified diagram of the upper layer of the shallow water model. We show this diagram as an aid to understanding equation (3.10). The interface between the upper and lower model layers is drawn as a flat floor (cf. Figure 3.2). Both the actual height field h (red solid line) and the radiative equilibrium height field h_{eq} (black solid line) are now measured with respect to this floor. We define a characteristic difference between h_{eq} and h on the dayside ($|h_{\text{eq}} - h|_{\text{day}}$) and on the nightside ($|h_{\text{eq}} - h|_{\text{night}}$) of the planet.

The balance in the momentum equation (3.1) involves several possibilities. Generally, the horizontal component of the pressure-gradient force $-g\nabla h$,¹⁰ which drives the flow, can be balanced by either atmospheric drag ($-\mathbf{v}/\tau_{\text{drag}}$), the Coriolis force ($-f\mathbf{k} \times \mathbf{v}$), horizontal advection ($\mathbf{v} \cdot \nabla \mathbf{v}$), or the vertical transport term ($\mathbf{R}$), which accounts for the momentum transfer from the lower layer. To order of magnitude, the balance is given by

$$g \frac{\Delta h}{a} \sim g \frac{\Delta h}{L} \sim \max \left[\frac{U}{\tau_{\text{drag}}}, fU, \frac{U^2}{L}, \frac{U}{H} \frac{\Delta h_{\text{eq}} - \Delta h}{\tau_{\text{rad}}} \right]. \quad (3.13)$$

We solve equations (3.12) and (3.13) for the dependent variables Δh and U . In the following, we express Δh in terms of the dimensionless amplitude A (compare with equation 3.5):

$$A \sim \Delta h / \Delta h_{\text{eq}}. \quad (3.14)$$

We also non-dimensionalize U in terms of the timescale ratio $\tau_{\text{adv}}/\tau_{\text{wave}}$:

$$\tau_{\text{adv}}/\tau_{\text{wave}} \sim \sqrt{gH}/U. \quad (3.15)$$

There are four possible balances in equation (3.13). Which of the terms is balancing the pressure-gradient force will generally depend on the values of τ_{rad} , τ_{drag} , the planetary latitude (ϕ), and the strength of forcing ($\Delta h_{\text{eq}}/H$). Below we solve for the four possible term balances and describe the conditions under which they operate.

3.4.1 Drag-dominated: Valid for Both Equatorial and Non-equatorial Regions

When drag is the dominant term balancing the pressure-gradient force, the solutions to equations (3.12) and (3.13) are

$$A \sim \left(1 + \frac{\tau_{\text{rad}}\tau_{\text{drag}}}{\tau_{\text{wave}}^2} \right)^{-1}, \quad (3.16)$$

$Ro \ll 1$, the flow is geostrophically balanced; a geostrophically balanced flow has a horizontal divergence $-\beta v/f$, and there will be an additional possible ageostrophic contribution to the divergence up to order RoU/L . Here, v is the meridional (north-south) wind velocity and $\beta = df/dy$ is the gradient of the Coriolis parameter with northward distance y , equal to $2\Omega \cos \phi/a$ on the sphere, where a is the planetary radius. Thus the geostrophic contribution to the horizontal divergence is $v \cot \phi/a$. Because the dominant flows on hot Jupiters have horizontal scales comparable to the Rossby deformation radius (Showman & Polvani 2011), which are comparable to the planetary radius for conditions appropriate to hot Jupiters, we have that $L \sim a$. Thus, the first term in Equation (3.9) scales as UH/L at a typical mid-latitude. Next consider the second term in Equation (3.9). When $Ro \gtrsim 1$, this term scales as $U\Delta h/L$. When $Ro \ll 1$, geostrophic balance implies that the geostrophic component of the flow is perpendicular to ∇h , leaving only the ageostrophic component available to flow along pressure gradients. Thus, in this case, the second term scales as $RoU\Delta h/L$. Because $\Delta h \lesssim H$, the first term generally dominates and the left-hand side of equation (3.9) therefore scales as UH/L .

¹⁰We have loosely been referring to $-g\nabla h$ as the horizontal component of the pressure-gradient force, when in fact it is the acceleration in the horizontal direction due to a horizontal gradient in pressure. Recall that ∇ is the horizontal (2D) gradient operator and that pressure is proportional to gh in a fluid with constant density.

$$\frac{\tau_{\text{adv}}}{\tau_{\text{wave}}} \sim \left(\frac{\tau_{\text{drag}}}{\tau_{\text{wave}}} \right)^{-1} \left(\frac{\Delta h_{\text{eq}}}{H} \right)^{-1} \left(1 + \frac{\tau_{\text{rad}} \tau_{\text{drag}}}{\tau_{\text{wave}}^2} \right). \quad (3.17)$$

Contours of equations (3.16) and (3.17) are shown in the region below the white horizontal line in Figures 3.6(c) and (d) ($\tau_{\text{drag}} \lesssim \tau_{\text{wave}}$) and below the white curved line in the right panels of Figure 3.8. This white line marks the boundary of the drag-dominated regime in our model; we will formally define the white line transition in equation (3.20). The same expression for A as in equation (3.16) results when linearizing the full model about a state of rest, where $\tau_{\text{adv}} \rightarrow \infty$ (this linearization is carried out in detail in Showman & Polvani 2011).¹¹ This strongly suggests that in this region of the $\tau_{\text{rad}}, \tau_{\text{drag}}$ plane, the advection timescale plays a minor role in controlling the day-night thickness and temperature differences. Indeed, Figure 3.6(d) shows that τ_{adv} is always significantly larger than τ_{wave} in this region. Equation (3.16) implies that the transition from small to large A occurs when $\tau_{\text{wave}} \sim \sqrt{\tau_{\text{rad}} \tau_{\text{drag}}}$. Note that the value of A in equation (3.16) is independent of the forcing strength $\Delta h_{\text{eq}}/H$. This is not true for the characteristic wind speed U and by extension τ_{adv} (equation (3.17)).

3.4.2 Coriolis-dominated: Valid for Non-equatorial Regions Only

When atmospheric drag is weak, the dominant balance in the force equation (3.13) depends on the Rossby number, $Ro = U/fL$. For conditions appropriate to hot Jupiters (typical rotation periods of a few Earth days, length scales comparable to a planetary radius, and wind speeds on the order of the wave speed or less), the Rossby number $Ro \lesssim 1$ —except near their equators. For small Ro , the Coriolis force tends to dominate over horizontal advection and vertical transport (the last two terms in equation 3.13). Thus, away from the equator, we can balance the pressure-gradient force against the Coriolis force. In this case, equations (3.12) and (3.13) yield

$$A \sim \left(1 + \frac{\tau_{\text{rad}}}{f \tau_{\text{wave}}^2} \right)^{-1}, \quad (3.18)$$

$$\frac{\tau_{\text{adv}}}{\tau_{\text{wave}}} \sim \left(\frac{\Delta h_{\text{eq}}}{H} \right)^{-1} \left(f \tau_{\text{wave}} + \frac{\tau_{\text{rad}}}{\tau_{\text{wave}}} \right). \quad (3.19)$$

Contours of equations (3.18) and (3.19) are shown in the region above the white line in Figures 3.6(c) and (d), with the Coriolis parameter f evaluated at a latitude of $\phi = 45^\circ$. Notice the similarity between equations (3.16) and (3.18), where the role of τ_{drag} has been replaced by $1/f$. Away from the equator, the Coriolis parameter $f \sim \Omega$, and equation (3.18) implies that the transition between small and large A occurs when $\tau_{\text{wave}} \sim \sqrt{\tau_{\text{rad}}/\Omega}$. The

¹¹To derive equation (3.16) from the shallow water equations, linearize a one-dimensional Cartesian version of equations (3.1) and (3.2), dropping the Coriolis term. Substitute one equation into the other to eliminate velocity. Impose sinusoidal forcing $h_{\text{eq}} = \Delta h_{\text{eq}} \exp(ikx)$ and solve for steady, sinusoidal solutions of the form $h = \Delta h \exp(ikx)$. Finally, solve for $A \sim \Delta h / \Delta h_{\text{eq}}$ to obtain the same expression as in equation (3.16).

transition between the Coriolis-dominated and the drag-dominated regimes occurs when equations (3.16) and (3.18) are equal, a condition which yields

$$\tau_{\text{drag}} \sim 1/f. \quad (3.20)$$

This condition formally defines the white line transition between the Coriolis-dominated and drag-dominated regimes in Figures 3.6(c) and (d).

We combine our expressions for A and $\tau_{\text{adv}}/\tau_{\text{wave}}$ valid away from the equator (equations (3.16) & (3.18) and (3.17) & (3.19), together with boundary condition equation (3.20)) to create Figures 3.6(c) and (d). These analytical results (derived for $Ro \lesssim 1$) are a good representation of the globally averaged numerical results shown in Figures 3.6(a) and (b)—even though the former applies only for non-equatorial regions whereas the latter averages over both non-equatorial and equatorial regions. Nonetheless, the comparison we make between Figures 3.6(a) and (b) and Figures 3.6(c) and (d) is fair because either the equatorial solution shows trends in A and $\tau_{\text{adv}}/\tau_{\text{wave}}$ similar to those of the mid-latitudes, or the equatorial region is small compared to the non-equatorial region (see Section 3.4.3).

Our analytical theory predicts that A is independent of the forcing strength ($\Delta h_{\text{eq}}/H$) for the entire $\tau_{\text{rad}}, \tau_{\text{drag}}$ plane. We test this prediction by running our numerical model at smaller $\Delta h_{\text{eq}}/H$. We show these results in Figure 3.7—indeed all panels exhibit the same general features of Figure 3.6(a).

3.4.3 Advection- or Vertical-transport-dominated: Valid for $Ro \gtrsim 1$

Near the equator (i.e., $Ro \gg 1$), the Coriolis force will vanish and the pressure-gradient force will be balanced by either advection or vertical transport. Both possibilities yield the same solution

$$A \sim 1 + \frac{1}{2} \left(\frac{\Delta h_{\text{eq}}}{H} \right)^{-1} \left(\frac{\tau_{\text{rad}}}{\tau_{\text{wave}}} \right)^2 \times \left\{ 1 - \left[1 + 4 \left(\frac{\Delta h_{\text{eq}}}{H} \right) \left(\frac{\tau_{\text{rad}}}{\tau_{\text{wave}}} \right)^{-2} \right]^{1/2} \right\}, \quad (3.21)$$

$$\frac{\tau_{\text{adv}}}{\tau_{\text{wave}}} \sim \left[\frac{\Delta h_{\text{eq}}}{H} + \frac{1}{2} \left(\frac{\tau_{\text{rad}}}{\tau_{\text{wave}}} \right)^2 - \sqrt{\frac{\Delta h_{\text{eq}}}{H} \left(\frac{\tau_{\text{rad}}}{\tau_{\text{wave}}} \right)^2 + \frac{1}{4} \left(\frac{\tau_{\text{rad}}}{\tau_{\text{wave}}} \right)^4} \right]^{-1/2}. \quad (3.22)$$

We compare these analytical predictions for A at the equator (equations 3.16 and 3.21) with the equatorial day-night contrast of our numerical simulations—that is, A_{equator} —in Figure 3.8 (left panels are the numerical simulations, and right panels are analytical predictions).

Notice how the weak drag solution (3.21) depends on forcing amplitude $\Delta h_{\text{eq}}/H$, whereas the strong drag solution (3.16) does not. The white line—marking the transition between the weak- and strong-drag regimes—is now a parabola, obtained by equating (3.16) and (3.21). As $\Delta h_{\text{eq}}/H$ is reduced, wind speeds are reduced and a greater region of phase space is drag dominated.

For the strong forcing expected on hot Jupiters ($\Delta h_{\text{eq}}/H \sim 1$), the numerical solution for the day-night contrast at the equator A_{equator} (lowermost panels in Figure 3.8) is very similar to the one obtained for mid-latitudes (Figures 3.6(a) and (c)), deviating by at most $\sim 15\%$. For weaker forcing, the solutions valid at mid-latitudes and the equator differ; however, at weak forcing, equations (3.21) and (3.22) are only valid in a narrow range of latitudes centered on the equator. This latitudinal range is delimited by the condition $Ro = 1$ and can be found analytically by solving

$$Ro = \frac{U}{fL} \sim 1 \sim \left(\frac{\sqrt{gH}}{2\Omega a} \right)^{1/2} \frac{1}{\sin \phi} \times \left[\frac{\Delta h_{\text{eq}}}{H} + \frac{1}{2} \left(\frac{\tau_{\text{rad}}}{\tau_{\text{wave}}} \right)^2 - \sqrt{\frac{\Delta h_{\text{eq}}}{H} \left(\frac{\tau_{\text{rad}}}{\tau_{\text{wave}}} \right)^2 + \frac{1}{4} \left(\frac{\tau_{\text{rad}}}{\tau_{\text{wave}}} \right)^4} \right]^{1/2} \quad (3.23)$$

for ϕ . For $\Delta h_{\text{eq}}/H = (1, 0.1, 0.01, 0.001)$ the width of the equatorial region is at most $\phi \sim (\pm 30^\circ, \pm 10^\circ, \pm 3^\circ, \pm 1^\circ)$ and goes to zero in the limit that the forcing amplitude goes to zero.

3.5 Interpretation of Theory

3.5.1 Timescale Comparison

To give an executive summary of Section 3.4: we can reproduce the characteristic day-night difference $A(\tau_{\text{rad}}, \tau_{\text{drag}}, \Delta h_{\text{eq}}/H)$ with a set of simple scaling relations

$$A \sim \begin{cases} \left(1 + \frac{\tau_{\text{rad}} \tau_{\text{drag}}}{\tau_{\text{wave}}^2} \right)^{-1} & \text{when } \tau_{\text{drag}} \lesssim \Omega^{-1} \\ \left(1 + \frac{\tau_{\text{rad}}}{\Omega \tau_{\text{wave}}^2} \right)^{-1} & \text{when } \tau_{\text{drag}} \gtrsim \Omega^{-1}, \end{cases} \quad (3.24)$$

which are valid for all forcing strengths ($\Delta h_{\text{eq}}/H$) and nearly all latitudes (except those closest to the equator where A is given by equation (3.21) in the weak drag limit). We have found both numerically and analytically that a transition from a planet with uniform atmospheric temperature ($A \sim 0$) to one with a large day-night temperature contrast relative to radiative equilibrium ($A \sim 1$) occurs when

$$\begin{cases} \tau_{\text{wave}} \sim \sqrt{\tau_{\text{rad}} \tau_{\text{drag}}} & \text{when } \tau_{\text{drag}} \lesssim \Omega^{-1} \\ \tau_{\text{wave}} \sim \sqrt{\tau_{\text{rad}} / \Omega} & \text{when } \tau_{\text{drag}} \gtrsim \Omega^{-1}. \end{cases} \quad (3.25)$$

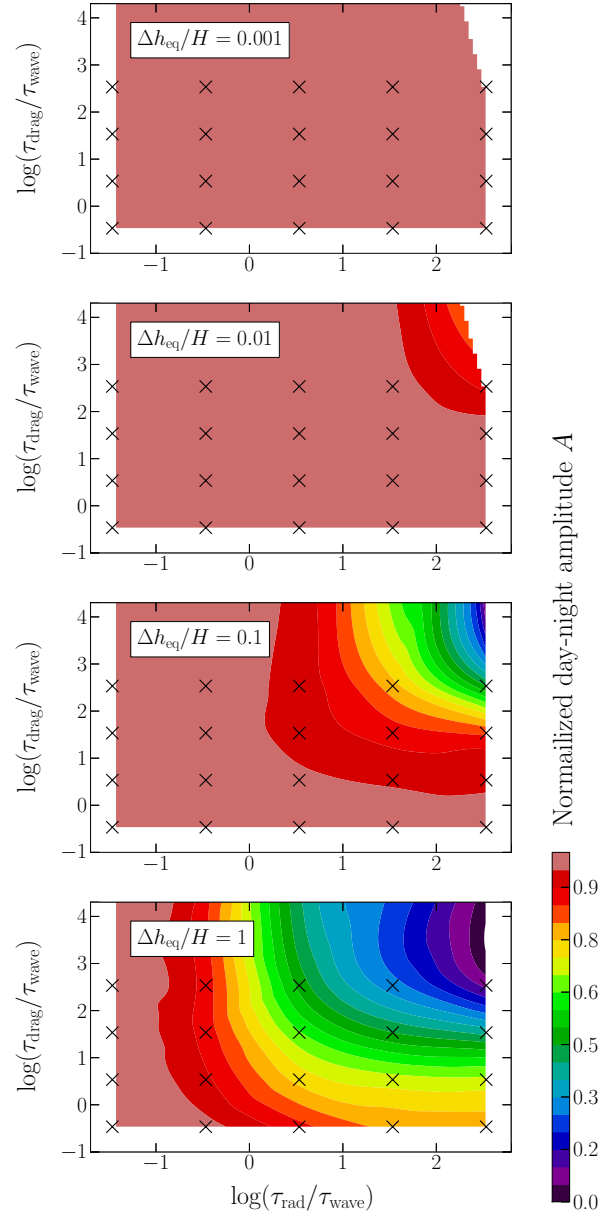


Figure 3.10 Contours of day-night height amplitude A as would be predicted by comparing τ_{rad} vs. τ_{adv} —the horizontal advective timescale. Values for A were computed with equation (3.26), where $\tau_{\text{adv}} \equiv L/U$, and U set to the global RMS value of the wind speed in the shallow water model. At $\Delta h_{\text{eq}}/H = 1$ (lowermost panel), the advective term in the momentum equation is of the same order of magnitude as other terms. As a result, comparison between τ_{adv} and τ_{rad} yields A -contours that show some similarity to the numerical results (Figure 3.6(a)). But as $\Delta h_{\text{eq}}/H$ is reduced, the advective term dies off faster than other terms, making advection less relevant (with the exception of a thin band around the equator). By the time $\Delta h_{\text{eq}}/H = 0.001$, τ_{adv} has become so large compared to τ_{rad} that equation (3.26) predicts $A \rightarrow 1$ over the entire parameter space, clearly contradicting numerical results. The low forcing amplitude cases demonstrate that the height field in the shallow water model is not being predominantly redistributed by planetary-scale horizontal advection.

Wave adjustment would thus seem to play a key role in controlling whether or not the thermal structure of the day-night contrast is close to radiative equilibrium. In contrast, horizontal advection and radiative damping are usually considered the dominant factors for heat redistribution. The comparison between τ_{adv} and τ_{rad} is intuitive and provides a reasonable estimate for the heat redistribution efficiency on hot Jupiters (e.g., [Perna et al. 2012](#)), which are strongly forced ($\Delta h_{\text{eq}}/H \sim 1$). Nevertheless, the τ_{adv} vs. τ_{rad} comparison is a poor predictor for A in the more general case, where $\Delta h_{\text{eq}}/H$ is not of order unity. We now show this explicitly.

In Figure 3.10 we show A -contours as would be predicted by the τ_{rad} vs. τ_{adv} timescale comparison:

$$A \sim \left(1 + \frac{\tau_{\text{rad}}}{\tau_{\text{adv}}}\right)^{-1}. \quad (3.26)$$

We chose the functional form for A in equation (3.26) because it possesses the correct limiting values (including $A = 1/2$ when $\tau_{\text{rad}} = \tau_{\text{adv}}$) and allows for a direct comparison with our results. Note that $\tau_{\text{adv}} \propto U^{-1}$ is not an input parameter in the model; it has to be estimated either from the numerical solution or by using our scaling solutions (equations (3.17), (3.19), or (3.22)). In Figure 3.10, we set U equal to the global RMS value of the wind speed in the numerical model. At large forcing amplitude ($\Delta h_{\text{eq}}/H = 1$, lowermost panel) the contours of A as predicted by equation (3.26) show some resemblance to the numerical shallow water solution shown in Figure 3.7 (lowermost panel). Nevertheless, as the forcing amplitude is reduced, the values of A predicted by equation (3.26) become increasingly inaccurate. In the limit where $\Delta h_{\text{eq}}/H \rightarrow 0$, equation (3.26) predicts $A \rightarrow 1$ everywhere, because the characteristic wind speed $U \propto \Delta h_{\text{eq}}/H$ (see equations (3.17) and (3.19)). In contrast, our numerical results show that A -contours are largely independent of $\Delta h_{\text{eq}}/H$.

Rather than invoking Equation (3.26), we can demonstrate the breakdown of the τ_{rad} -vs- τ_{adv} prediction in the low-amplitude limit simply by considering the ratio of these two timescales. When $\Delta h_{\text{eq}}/H = 0.001$, values of $\tau_{\text{adv}}/\tau_{\text{rad}}$ vary between 30 and 3×10^5 over our explored parameter space. Therefore, a τ_{rad} -vs- τ_{adv} comparison would predict that the day-night thickness contrast is always very close to radiative equilibrium over the entire explored parameter space. This is inconsistent with our numerical simulations (Figures 3.4 and 3.7, top panel), which clearly show a transition from models close to radiative equilibrium at short τ_{rad} to models with much smaller thermal contrasts at long τ_{rad} .

It is clear that, at least for low forcing amplitudes, the amplitude of the day-night thermal contrast (relative to radiative equilibrium) is not controlled by a comparison between τ_{rad} and τ_{adv} . We now give two physical interpretations of the theory.

3.5.2 Vertical Advection

The timescale comparison in equation (3.25) can be obtained by comparing a *vertical* advection timescale to the radiative timescale. Define the vertical advection time, τ_{vert} , as the time for a fluid parcel to move vertically over a distance corresponding to the day-night

thickness difference Δh . The vertical velocity, by mass continuity, is $\sim H \nabla \cdot \mathbf{v}$ (where ∇ is the horizontal gradient operator and $\mathbf{v}$ is the horizontal velocity). Then

$$\tau_{\text{vert}} \sim \frac{\Delta h}{H \nabla \cdot \mathbf{v}} \sim \frac{\Delta h L}{H U}. \quad (3.27)$$

In the strong drag regime, equation (3.13) becomes $\Delta h/U \sim L/(g\tau_{\text{drag}})$, which when substituted into equation (3.27) implies that

$$\tau_{\text{vert}} \sim \frac{L^2}{gH\tau_{\text{drag}}} \sim \frac{\tau_{\text{wave}}^2}{\tau_{\text{drag}}}. \quad (3.28)$$

If τ_{rad} is of order τ_{vert} , then

$$\tau_{\text{rad}} \sim \tau_{\text{vert}} \sim \frac{\tau_{\text{wave}}^2}{\tau_{\text{drag}}}, \quad (3.29)$$

which is precisely the same comparison in equation (3.25) derived in the strong drag limit. Thus, our solution approaches radiative equilibrium ($\Delta h \rightarrow \Delta h_{\text{eq}}$) when $\tau_{\text{vert}} \gg \tau_{\text{rad}}$, i.e., when $\tau_{\text{wave}} \gg \sqrt{\tau_{\text{rad}}\tau_{\text{drag}}}$. In other words, the atmosphere is close to radiative equilibrium when the vertical advection time (over a distance Δh) is long compared to the radiative time. Conversely, the behavior is in the limit of small thickness variations ($\Delta h \ll \Delta h_{\text{eq}}$) when $\tau_{\text{vert}} \ll \tau_{\text{rad}}$, i.e., when $\tau_{\text{wave}}^2 \ll \tau_{\text{rad}}\tau_{\text{drag}}$. In other words, the day-night thickness difference is small (compared to radiative equilibrium) when the vertical advection time is short compared to the radiative time.

In the Coriolis-dominated regime, the role of τ_{drag} is replaced by $f^{-1} \sim \Omega^{-1}$, as can readily be seen in equation (3.13). Setting τ_{rad} equal to τ_{vert} now yields

$$\tau_{\text{rad}} \sim \tau_{\text{vert}} \sim \Omega \tau_{\text{wave}}^2, \quad (3.30)$$

which is the same timescale comparison in equation (3.25) for the Coriolis-dominated regime.

In either case, the relationship between the vertical and horizontal advection times follows directly from equation (3.27):¹²

$$\tau_{\text{vert}} \sim \tau_{\text{adv}} \frac{\Delta h}{H}. \quad (3.31)$$

Therefore, the vertical advection time, τ_{vert} , is smaller than the horizontal advection time by $\Delta h/H$. The vertical and horizontal advection times become comparable when day-night thickness differences are on the order of unity (which can occur only for large forcing amplitudes $\Delta h_{\text{eq}}/H \sim 1$). Only in that special case, will a comparison between τ_{rad} and the horizontal advection time, τ_{adv} , give a roughly correct prediction for A (as previously mentioned in Section 3.5.1 and shown in the bottom panel of Figure 3.10).

The physical reason for the importance of τ_{vert} over τ_{adv} in controlling the transition stems from the relative roles of vertical and horizontal advection in the continuity equation

¹²This relationship can also be obtained by equating the rate of vertical mass transport $\sim \rho_{\text{upper}} \Delta h L^2 / \tau_{\text{vert}}$ to the rate of horizontal mass transport $\sim \rho_{\text{upper}} H L^2 / \tau_{\text{adv}}$.

(3.9). When $\Delta h/H$ is small, the term $h\nabla \cdot \mathbf{v}$ —which essentially represents vertical advection—dominates over the horizontal advection term $\mathbf{v} \cdot \nabla h$ (see footnote 9). Thus, under conditions of small $\Delta h/H$, the dominant balance is between radiative heating/cooling and *vertical*, rather than horizontal, advection. This is precisely the shallow-water version of the so-called “weak temperature gradient” (WTG) regime that dominates in the Earth’s tropics, where the time-mean balance in the thermodynamic energy equation is between vertical advection and radiative cooling (Sobel et al. 2001; Bretherton & Sobel 2003). Only when $\Delta h/H \sim 1$ does horizontal advection generally become comparable to vertical advection in the local balance.

3.5.3 Wave Adjustment Mechanism

As described in Section 3.1, gravity waves are known to play a central role in regulating the thermal structure in the Earth’s tropics (Bretherton & Smolarkiewicz 1989; Sobel 2002; Showman et al. 2013a). Likewise, our analytic scaling theory (Section 3.4) shows the emergence of a wave timescale in controlling the transition from small to large day-night temperature difference in our global, steady-state solutions. This is not accidental but strongly implies a role for wave-like processes in governing the dynamical behavior. Although our numerical simulations and theory stand on their own, we show here that they can be interpreted in terms of a wave adjustment mechanism.

To illustrate, consider a freely propagating gravity wave in the time-dependent shallow-water system. In such a wave, horizontal variations in the thickness (h) cause pressure-gradient forces that induce horizontal convergence or divergence, which locally changes the height field and allows the wave structure to propagate laterally. In fact, it can easily be shown that this physical process—namely, vertical stretching/contraction of atmospheric columns in response to horizontal pressure-gradient forces—naturally leads to a wave timescale $\tau_{\text{wave}} \sim L/\sqrt{gH}$ for a freely propagating wave to propagate over a distance L .¹³ Now, although our solutions in Section 3.3 lack freely propagating waves (being forced, damped, and steady), the key point is that the *same physical mechanism* that causes wave propagation in the free, time-dependent

¹³Specifically, consider for simplicity a one-dimensional, non-rotating linear shallow-water system in the absence of forcing or damping. If a local region begins with a height Δh different from surrounding regions, the timescale for the wave to propagate over its wavelength L will be determined by the time needed for horizontal convergence/divergence to locally change the height by Δh . This timescale is simply given by

$$\tau \sim \frac{\Delta h}{H\nabla \cdot \mathbf{v}} \sim \frac{\Delta h L}{HU}, \quad (3.32)$$

where U is the wind speed associated with the wave motion. In the absence of forcing or damping, the horizontal momentum equation is the linearized version of Equation (3.1) with the right-hand side set to zero. In the absence of rotational effects, the balance is simply between the time-derivative term and the pressure-gradient term, which to order of magnitude is

$$\frac{U}{\tau} \sim g \frac{\Delta h}{L}. \quad (3.33)$$

Combining these two equations immediately yields $\tau \sim \frac{L}{\sqrt{gH}}$.

case regulates the day-night thickness variation in our steady, forced, damped models. In particular, in our hot-Jupiter models, the dayside mass source and nightside mass sink cause a thickening of the layer on the dayside and a thinning of it on the nightside; in response, the horizontal pressure-gradient forces cause a horizontal divergence on the dayside and convergence on the nightside, which attempts to thin the layer on the dayside and thicken it on the nightside. Although technically no phase propagation occurs in this steady case, this process—being essentially the same process that governs gravity-wave dynamics—nevertheless occurs on the gravity-wave timescale. This is the reason for the emergence of τ_{wave} in our analytic solutions in Section 3.4.

The situation on a rotating planet is more complex, because the Coriolis force significantly modifies the wave behavior. On a rotating planet, freely propagating, global-scale waves within a deformation radius of the equator split into a variety of equatorially trapped wave modes, including the Kelvin wave and equatorially trapped Rossby waves. (For overviews, see Matsuno (1966), Holton (2004, pp. 394–400) or Andrews et al. (1987, pp. 200–208).) The Kelvin wave exhibits pressure perturbations peaking at the equator, with strong zonally divergent east-west winds; these waves propagate to the east. The east-west winds in the Kelvin wave cause strong north-south Coriolis forces that prevent the expansion of the pressure perturbations in latitude; however, nothing resists the pressure perturbations in longitude, and so the Kelvin wave propagates zonally like a gravity wave at a speed $\sqrt{gH}$. In contrast, the Rossby wave exhibits pressure perturbations that peak off the equator and vortical winds that encircle these pressure perturbations; these waves propagate to the west. See Holton (2004, Figure 11.15) and Matsuno (1966, Figure 4(c)), respectively, for visuals of these two wave types.

The link between our solutions and wave dynamics becomes even tighter when one compares the detailed spatial structure of the solutions to these tropical wave modes. Building on a long history of work in tropical dynamics (e.g., Matsuno 1966; Gill 1980), Showman & Polvani (2011) showed that the behavior of steady, forced, damped solutions like those in Figures 3.3 and 3.4 can be interpreted in terms of standing, planetary-scale Rossby and Kelvin waves. Examining, for example, the $\tau_{\text{rad}} = 1$ day, $\tau_{\text{drag}} = 1$ day models in Figure 3.4, the off-equatorial behavior, including the off-equatorial pressure maxima and the vortical winds that encircle them (clockwise in northern hemisphere, counterclockwise in southern hemisphere), is dynamically analogous to the equatorially trapped Rossby wave mentioned above. The low-latitude behavior, with winds that zonally diverge from a point east of the substellar point, is dynamically analogous to the equatorial Kelvin wave mentioned above. See Gill (1980) and Showman & Polvani (2011) for further discussion. Again, here is the key point: the physical mechanisms that cause stretching/contraction of the shallow-water column and wave propagation of Kelvin and Rossby waves in the free, time-dependent case are *the same physical mechanisms* that regulate the spatial variations of the thickness in our forced, damped, steady solutions.

In sum, if the radiative and frictional damping times are sufficiently long, the Kelvin and Rossby waves act efficiently to flatten the layer, and the day-night thickness differences are small. If the radiative and frictional damping times are sufficiently short, the Kelvin and

Rossby waves are damped and cannot propagate zonally; the thermal structure is then close to radiative equilibrium. Although our solutions are steady, this similarity to wave dynamics explains the natural emergence of a wave timescale in controlling the transition between small and large day-night contrast. In particular, because the Kelvin wave propagates at a speed $\sqrt{gH}$, the fundamental wave timescale that emerges is $L/\sqrt{gH}$.

It is important to emphasize that, despite the importance of wave timescales, horizontal advection nevertheless plays a crucial role in the dynamics. Consider an imaginary surface at the terminator dividing the planet into dayside and nightside hemispheres. It is horizontal advection across this surface that ultimately transports heat from day to night, thereby allowing each hemisphere to reach a steady state in the presence of continual dayside heating and nightside cooling. In the linear limit of our shallow-water model, this transport manifests as advection of the mean thickness (i.e., uH integrated around the terminator), although advection of thickness *variations* can also play a role at high amplitude, when these variations are not small relative to H . The importance of horizontal advection does not mean that the flow behavior is controlled by the horizontal advection timescale, and indeed we have shown that it is generally not, particularly when the forcing amplitude is weak.

3.6 Application to Hot Jupiter Observations

Here we compare predictions for the day-night height difference, A , obtained from our shallow water model to the observed fractional infrared flux variations, A_{obs} , on hot Jupiters. Because the shallow water equations do not explicitly include stellar irradiation, we have to express our model's input parameters τ_{rad} and τ_{drag} in terms of T_{eq} (our proxy for stellar irradiation). We find the dependence of the radiative timescale on T_{eq} by approximating τ_{rad} as the ratio between the available thermal energy per unit area within a pressure scale-height and the net radiative flux from that layer (Showman & Guillot 2002),

$$\tau_{\text{rad}} \sim \frac{P c_p}{4g\sigma T_{\text{eq}}^3}, \quad (3.34)$$

where P is the atmospheric pressure at the emitting layer, c_p is the specific heat, and σ is the Stefan-Boltzmann constant. We have chosen to leave τ_{drag} as a free parameter, because the source of atmospheric drag in gas giants remains largely unknown (e.g., Perna et al. 2010; Li & Goodman 2010; Showman et al. 2010), and because our results suggest a weak dependence of A on τ_{drag} (see Figure 3.6(a)).¹⁴

¹⁴In the case of hot Jupiters, the main candidate for atmospheric drag is Lorentz-force breaking of the thermally ionized atmosphere. For the case of HD 209458b, Perna et al. (2010) estimate that τ_{drag} could reach values as low as ~ 0.1 days on the planet's dayside, which matches our lowest considered τ_{drag} . Rauscher & Menou (2013) found that the inclusion of Lorentz drag in their global circulation model of HD 209458b changes the ratio between maximum and minimum flux emission from the planet by up to 5% when compared to a drag-free model (see their Table 2). Although their study excluded the most strongly irradiated hot Jupiters, their simulations are in accordance with our result that A depends only weakly on τ_{drag} .

We plot model values for A together with A_{obs} from hot Jupiter observations in Figure 3.11. Each broken curve shows model results for A for a constant value of τ_{drag} , ranging from values that are marginally in the strong drag regime ($\tau_{\text{drag}} = 0.1$ days) to no drag ($\tau_{\text{drag}} \rightarrow \infty$). Additionally, we show the solution when $\tau_{\text{drag}} = \tau_{\text{rad}}$ with the solid red line. Because τ_{drag} is fixed, the only remaining free variable is τ_{rad} , which we express in terms of T_{eq} using equation (3.34) (compare upper and lower x -axes). All curves follow roughly the same path which reproduces the observational trend of increasing A with equilibrium temperature T_{eq} . The curves run close to each other because the models considered are either in the weak drag regime or barely in the strong drag regime. Solutions with $\tau_{\text{drag}} < 0.1$ days are numerically challenging and were therefore not explored. In any case, for the case of magnetic drag, the temperatures required to reach such low drag timescales are $\gtrsim 1500$ K (Perna et al. 2010), where the corresponding τ_{rad} is already so low that $A \sim 1$ regardless of the strength of drag. In summary, when $\sqrt{\tau_{\text{rad}}/\Omega}$ is shorter than τ_{wave} , the Kelvin and Rossby waves that emerge near the equator are damped before they can propagate zonally, resulting in $A \sim 1$. In contrast, when $\sqrt{\tau_{\text{rad}}/\Omega}$ is long compared to τ_{wave} , the emerging waves can propagate far enough to flatten the fluid layer, resulting in $A \sim 0$.

3.7 Conclusions

We have presented a simple atmospheric model for tidally locked exoplanets that reproduces the observed transition from atmospheres with longitudinally uniform temperatures to atmospheres with large day-night temperature gradients as stellar insolation increases (Figure 3.11). In our model we have parameterized the stellar insolation in terms of a radiative timescale, τ_{rad} , and frictional processes in terms of a drag timescale, τ_{drag} . The shallow water model contains two additional natural timescales: the rotation period of the planet ($\sim \Omega^{-1}$), and τ_{wave} , the timescale over which gravity waves travel horizontally over planetary distances. We have developed an analytical scaling theory to estimate the heat redistribution efficiency in terms of these four timescales. Our scaling theory predicts that for sufficiently weak atmospheric drag, the temperature distribution on the planet can be estimated by the ratio of τ_{wave} and $\sqrt{\tau_{\text{rad}}/\Omega}$. Drag will influence the day-night temperature contrast if it operates on a timescale shorter than Ω^{-1} . In this drag-dominated regime, the heat redistribution efficiency will depend on the ratio of τ_{wave} and $\sqrt{\tau_{\text{rad}}\tau_{\text{drag}}}$. These scaling relations are summarized in equation (3.25). We provide two physical interpretations to understand why these timescales arise from the shallow water model. We derive the first interpretation by noting that the same physical mechanisms that generate equatorially trapped waves in an undamped shallow water model also regulate the steady-state solutions of our forced-damped model. The heat redistribution efficiency is therefore related to the characteristic distance that waves can travel before they are damped. This distance is set by the relative magnitudes of the timescale for waves to travel over planetary distances, τ_{wave} , and the timescale for the waves to damp.

For the second interpretation, we recognize that the timescale comparisons of equation (3.25) can be written as $\tau_{\text{rad}} \sim \tau_{\text{vert}}$, where τ_{vert} is the timescale for a parcel to advect vertically

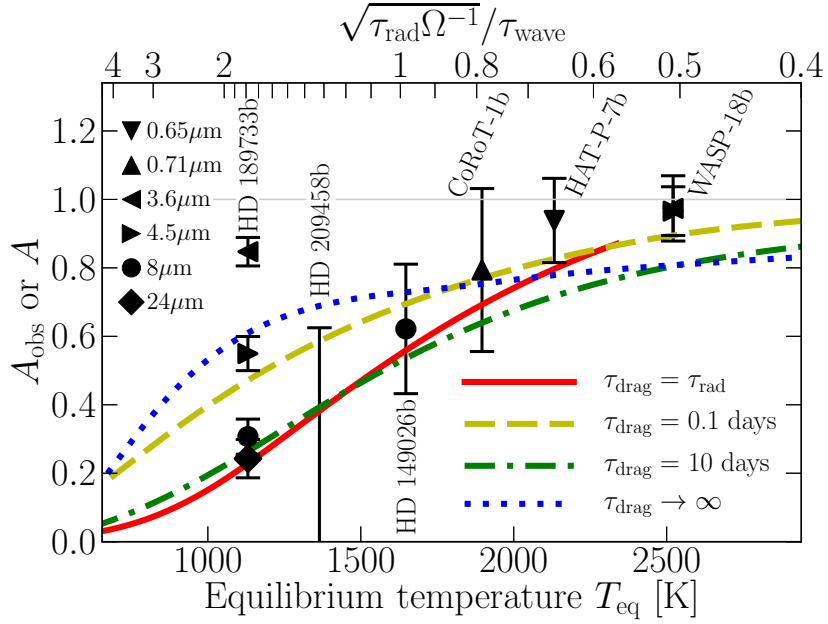


Figure 3.11 Same as Figure 3.1, but including shallow water model results. Black symbols are fractional day-night flux variations (A_{obs}) for hot Jupiters with measured light curves, as explained in detail in Figure 3.1. Colored dashed curves show model results for the normalized day-night height difference A for constant τ_{drag} , expressed in Earth days, while the solid red line shows the solution when $\tau_{\text{drag}} = \tau_{\text{rad}}$. Equilibrium temperatures (T_{eq}) for the model were estimated with equation (3.34) with $c_p = 10^4 \text{ J kg}^{-1} \text{ K}^{-1}$, $g = 10 \text{ m s}^{-2}$, and $P = 0.25 \text{ bar}$, which is approximately the pressure of the layer radiating to space (compare upper and lower x -axes). Because plotted solutions are (mostly) in the weak-drag regime, all curves lie close together and broadly reproduce the observational trend of increasing A_{obs} with increasing T_{eq} .

over a distance equal to the day-night difference in thickness. This criterion emerges from the fact that, as long as forcing amplitudes are not large, it is primarily vertical advection—and not horizontal advection, as commonly assumed—that balances radiative relaxation in the continuity equation. Despite this fact, the timescale comparison between τ_{rad} and the horizontal advection timescale, τ_{adv} , provides reasonable estimates for the heat redistribution efficiency on hot Jupiters. This is because these gas giants have strongly forced atmospheres ($\Delta h_{\text{eq}}/H \sim 1$), where τ_{adv} becomes comparable to τ_{vert} (see equation (3.31)), and where horizontal and vertical thermal advections become comparable. In weakly forced systems, the timescale comparison between τ_{rad} and τ_{adv} is a poor predictor for the heat redistribution efficiency, as we show in Figure 3.10. In contrast, the timescale comparison between τ_{rad} vs. τ_{vert} —derived from the dynamical equations—yields a more accurate estimate of the heat redistribution efficiency at any $\Delta h_{\text{eq}}/H$, including the strongly forced hot Jupiters.

Chapter 4

Catastrophic Evaporation of Rocky Planets

Daniel Perez-Becker & Eugene Chiang (2013), MNRAS, 443, 2294.¹

Abstract

Short-period exoplanets can have dayside surface temperatures surpassing 2000 K, hot enough to vaporize rock and drive a thermal wind. Small enough planets evaporate completely. We construct a radiative-hydrodynamic model of atmospheric escape from strongly irradiated, low-mass rocky planets, accounting for dust-gas energy exchange in the wind. Rocky planets with masses $\lesssim 0.1M_{\oplus}$ (less than twice the mass of Mercury) and surface temperatures $\gtrsim 2000$ K are found to disintegrate entirely in $\lesssim 10$ Gyr. When our model is applied to Kepler planet candidate KIC 12557548b—which is believed to be a rocky body evaporating at a rate of $\dot{M} \gtrsim 0.1 M_{\oplus}/\text{Gyr}$ —our model yields a present-day planet mass of $\lesssim 0.02 M_{\oplus}$ or less than about twice the mass of the Moon. Mass loss rates depend so strongly on planet mass that bodies can reside on close-in orbits for Gyrs with initial masses comparable to or less than that of Mercury, before entering a final short-lived phase of catastrophic mass loss (which KIC 12557548b has entered). Because this catastrophic stage lasts only up to a few percent of the planet’s life, we estimate that for every object like KIC 12557548b, there should be 10–100 close-in quiescent progenitors with sub-day periods whose hard-surface transits may be detectable by *Kepler*—if the progenitors are as large as their maximal, Mercury-like sizes (alternatively, the progenitors could be smaller and more numerous). According to our calculations, KIC 12557548b may have lost $\sim 70\%$ of its formation mass; today we may be observing its naked iron core.

¹Monthly Notices of the Royal Astronomical Society, 443, 2294–2309, 2013 June 12
 © 2013 The Authors. Published by Oxford University Press on behalf of the Royal Astronomical Society.

4.1 Introduction

Atmospheric escape shapes the faces of planets. Mechanisms for mass loss vary across the Solar System. Planets lacking magnetic fields can be stripped of their atmospheres by the Solar wind: the atmospheres of both Mercury and Mars are continuously eroded and refreshed by Solar wind ions (Potter & Morgan 1990, 1997; Lammer & Bauer 1997; Bida et al. 2000; Killen et al. 2004; Jakosky & Phillips 2001). Venus demonstrates the extreme sensitivity of atmospheres to stellar insolation. Although its distance to the Sun is only 30% less than that of the Earth, Venus has lost its water to hydrodynamic escape powered by Solar radiation (Kasting & Pollack 1983; Zahnle & Kasting 1986). Impacts with comets and asteroids can also purge planets of their atmospheres, as has thought to have happened to some extent on Mars (Melosh & Vickery 1989; Jakosky & Phillips 2001) and some giant planet satellites (Zahnle et al. 1992). An introductory overview of atmospheric escape in the Solar System is given by Catling & Zahnle (2009).

Smaller bodies lose their atmospheres more readily because of their lower surface gravities. Bodies closer to their host star also lose more mass because they are heated more strongly. Extrasolar planets on close-in orbits are expected to have highly processed atmospheres. In extreme cases, stellar irradiation might even evaporate planets in their entirety—comets certainly fall in this category.

Hot Jupiters are the best studied case for atmospheric erosion in extrasolar planets (for theoretical models, see Yelle 2004, 2006; Tian et al. 2005; García Muñoz 2007; Holmström et al. 2008; Murray-Clay et al. 2009; Ekenbäck et al. 2010; Tremblin & Chiang 2013). For these gas giants, mass loss is driven by stellar ultraviolet (UV) radiation which photo-ionizes hydrogen and imparts enough energy per proton to overcome the planet’s large escape velocity (~ 60 km/s). Mass loss for hot Jupiters typically occurs at a rate of $\dot{M} \sim 10^{10}$ – 10^{11} g/s so that $\sim 1\%$ of the planet mass is lost over its lifetime (Yelle 2006; García Muñoz 2007; Murray-Clay et al. 2009). Although this erosion hardly affects a hot Jupiter’s internal structure, the outflow is observable. Winds escaping from hot Jupiters have been spectroscopically detected by the *Hubble Space Telescope* (*HST*). During hot Jupiter transits, several absorption lines (H I, O I, C II, Mg II, and Si III) deepen by ~ 2 – 10% (Vidal-Madjar et al. 2003, 2004, 2008; Ben-Jaffel 2007, 2008; Fossati et al. 2010; Lecavelier Des Etangs et al. 2010; Linsky et al. 2010; Lecavelier des Etangs et al. 2012).

Because of their lower escape velocities, lower mass super-Earths should have their atmospheres more significantly sculpted by evaporation. Lopez et al. (2012) found that hydrogen-dominated atmospheres of close-in super-Earths could be stripped entirely by UV photoevaporation (e.g., Kepler-11b).

At even lower masses and stellocentric distances, planets might evaporate altogether. When dayside temperatures surpass ~ 2000 K, rocks, particularly silicates and iron (and to a lesser extent, Ti, Al, and Ca), begin to vaporize. If the planet mass is low enough, thermal velocities of the high-metallicity vapor could exceed escape velocities. Under these circumstances, mass loss is not limited by stellar UV photons, but is powered by the incident bolometric flux.

4.1.1 KIC 12557548b: A Catastrophically Evaporating Planet

In fact, the *Kepler* spacecraft may have enabled the discovery of an evaporating, close-in rocky planet near the end of its life. As reported by [Rappaport et al. \(2012\)](#), the K-star KIC 12557548 dims every 15.7 hours by a fractional amount f that varies erratically from a maximum of $f \approx 1.3\%$ to a minimum of $f \lesssim 0.2\%$. These occultations are thought to arise from dust embedded in the time-variable outflow emitted by an evaporating rocky planet—hereafter KIC 1255b. Because *Kepler* observes in the broadband optical, the obscuring material cannot be gas which absorbs only in narrow lines; it must take the form of dust which absorbs and scatters efficiently in the continuum. As we will estimate shortly, the amount of dust lost by the planet is prodigious.

From the ~ 2000 K surface of the planet is launched a thermal (“Parker-type”) wind out of which dust condenses as the gas expands and cools. The composition of the high-metallicity wind reflects the composition of the evaporating rocky planet (see, e.g., [Schaefer & Fegley 2009](#)): the wind may consist of Mg, SiO, O, and O₂—and Fe if it has evaporated down to its iron core. Stellar gravity, together with Coriolis and radiation-pressure forces, then shape the dusty outflow into a comet-like tail. The peculiar shape of the transit light curve of KIC 1255b supports the interpretation that the occultations are caused by a dusty tail. First, the light curve evinces a flux excess just before ingress, which could be caused by the forward scattering of starlight toward the observer. Second, the flux changes more gradually during egress than during ingress, as expected for a trailing tail. Both features of the light curve and their explanations were elucidated by [Rappaport et al. \(2012\)](#), and further quantified by [Brogi et al. \(2012\)](#) who presented a phenomenological light-scattering model of a dusty tail.²

Spectra of KIC 12557548 taken using the Keck Telescope reveal no radial velocity variations down to a limiting accuracy of ~ 100 m/s (A. Howard and G. Marcy 2012, personal communication). A deep Keck image also does not reveal any background blended stars. These null results are consistent with the interpretation that KIC 1255b is a small, catastrophically evaporating planet.

The characteristic mass loss rate for KIC 1255b—a quantity that we will reference

² Alternative ideas might include Roche lobe overflow from a companion to the star (c.f. [Lai et al. 2010](#)). A problem with this scenario is that the mass transfer accretion stream, funneled through the L1 Lagrange point, would lead the companion in its orbit, not trail it as is required by observation. By comparison, in our dusty wind model, there is no focussing of material at L1 because the wind is not in hydrostatic equilibrium and therefore does not conform to Roche lobe equipotentials. Moreover, stellar radiation pressure, usually neglected in standard Roche-overflow models, is crucial for diverting the flow so that it trails the planet in its orbit (see also [Mura et al. 2011](#)). Another alternative involves bow shocks of planets passing through stellar coronae ([Vidotto et al. 2010, 2011a,b](#); [Llama et al. 2011](#); [Tremblin & Chiang 2013](#)) which may explain the ingress-egress asymmetry observed for the exoplanet WASP-12b ([Fossati et al. 2010](#)). Bow shocks between the planetary wind and the stellar wind are also expected to occur for KIC 1255b, but would be hard to detect through *Kepler*’s broadband filter since the shocked gas alone would absorb only in narrow absorption lines. At the low gas densities characterizing exoplanet winds, only dust can generate the large continuum opacities required by the deep transits observed by *Kepler*. [Rappaport et al. \(2012\)](#) reviews these and additional arguments supporting the dust tail model. A similar interpretation was proposed for the light curve of beta Pictoris ([Lecavelier Des Etangs et al. 1997](#); [Lamers et al. 1997](#)).

throughout this Chapter—can be estimated from the observations as follows. First compute the total mass in grains required to absorb/scatter a fraction f of the starlight, assuming that the dust is optically thin (or marginally so). Out of the total area πR_*^2 presented by the stellar face, take the dust cloud to cover an area A and to have a characteristic line-of-sight optical depth $\Sigma_d \kappa_d$, where Σ_d is the grain mass per unit area (on the sky) and κ_d is the grain opacity. Then

$$f = \frac{A}{\pi R_*^2} \Sigma_d \kappa_d. \quad (4.1)$$

Assume further that grains can be modeled as a monodispersion of spheres with radii s and internal material density ρ_{int} so that

$$\kappa_d = \frac{3}{4\rho_{\text{int}}s}. \quad (4.2)$$

The total mass in dust covering the star is then

$$M_d = \Sigma_d A = \frac{4\pi}{3} f \rho_{\text{int}} R_*^2 s. \quad (4.3)$$

To obtain a mass loss rate, divide this mass by the orbital period of 15.7 hours on the grounds that transit depths change by up to an order of magnitude from orbit to orbit. For $f = 1\%$ —when KIC 1255b is in a state of relatively high mass loss—the resultant mass loss rate in grains is

$$\dot{M}_{1255\text{b}, \text{dust}} \sim 0.4 \left(\frac{R_*}{0.65 R_\odot} \right)^2 \left(\frac{f}{0.01} \right) \left(\frac{s}{0.5 \mu\text{m}} \right) \left(\frac{\rho_{\text{int}}}{1 \text{ g/cm}^3} \right) M_\oplus \text{ Gyr}^{-1}. \quad (4.4)$$

This is the mass loss rate in dust only, by construction. The gas-to-dust ratio by mass, η , must be at least on the order of unity if outflowing gas is to carry, by Epstein drag, embedded grains out of the gravity well of the planet (Rappaport et al. 2012). Henceforth we will adopt as our fiducial total mass loss rate (for the relatively high value of $f = 1\%$):

$$\dot{M}_{1255\text{b}} = (1 + \eta) \dot{M}_{1255\text{b}, \text{dust}} \sim 1 M_\oplus \text{ Gyr}^{-1}. \quad (4.5)$$

The total mass loss rate $\propto \rho_{\text{int}} s (1 + \eta)$. Given the uncertainties in these three factors (some rough guesses: $\rho_{\text{int}} \gtrsim 0.5 \text{ g cm}^{-3}$ for possibly porous grains; $s \gtrsim 0.1 \mu\text{m}$; $\eta \gtrsim 1$), the total mass loss rate could take any value $\dot{M}_{1255\text{b}} > 0.1 M_\oplus / \text{Gyr}$ (but see also our Figures 4.8 and 4.10 which place upper bounds on mass loss rates).

Rappaport et al. (2012) claimed that the present-day mass of KIC 1255b is $\sim 0.1 M_\oplus$. They estimated, to order-of-magnitude, the maximum planet mass that could reproduce the observed mass loss rate given by equation (4.5). Here we revisit the issue of planet mass—and the evaporative lifetime it implies—with a first-principles solution of the hydrodynamic equations for a planetary wind, including an explicit treatment of dust-gas thermodynamics. We will obtain not only improved estimates of the present-day mass, but also full dynamical histories of KIC 1255b.

4.1.2 Plan of This Chapter

We develop a radiative-hydrodynamic wind model to study mass loss of close-in rocky planets with dayside temperatures high enough to vaporize rock. Although our central application will be to understand KIC 1255b—both today and in the past—the model is general and can be easily be modified for parameters of other rocky planets.

The Chapter is structured as follows. In §4.2 we present the model that computes $\dot{M}$ as a function of planet mass M . We give results in §4.3, including estimates for the maximum present-day mass and the maximum formation mass of KIC 1255b. A wide-ranging discussion of the implications of our results—including, e.g., an explanation of why mass loss in this context does not obey the often-used energy-limited mass loss formula (cf. [Lopez et al. 2012](#)), as well as the prospects of detecting close-in Mercuries and stripped iron cores—is given in §4.4. Our findings are summarized in §4.5.

4.2 The Model

We construct a 1D model for the thermally-driven atmospheric mass loss from a close-in rocky planet, assuming that it occurs in the form of a transonic Parker wind. The atmosphere of the planet consists of a high-metallicity gas created by the sublimation of silicates (and possibly iron, as we discuss in §4.4.7) from the planetary surface. We focus on the flow that streams radially toward the star from the substellar point on the planet. The substellar ray is of special interest because the mass flux it carries will be maximal insofar as the substellar point has the highest temperature—and thus the highest rate of rock sublimation—and insofar as stellar gravity will act most strongly to accelerate gas away from the planet.

As the high-metallicity gas flows away from the surface of the planet, it will expand and cool, triggering the condensation of dust grains. Dust grains can affect the outflow in several ways: (a) latent heat released during the condensation of grains will heat the gas; (b) continuously heated by starlight, grains will transfer their energy to the gas via gas-grain collisions; and (c) absorption/scattering of starlight by grains will reduce the stellar flux reaching the planet, reducing the surface temperature and sublimation rate. In this work, we do not account for the microphysics of grain condensation, as cloud formation is difficult to calculate from first principles (see, e.g., the series of five papers beginning with [Helling et al. 2001](#); see also [Helling et al. 2008b](#) and [Helling et al. 2008a](#)). Instead, we treat the dust-to-gas mass fraction x_{dust} as a free function and explore how our results depend on this function. In our 1-fluid model, dust grains are perfectly entrained in the gas flow, an assumption we test *a posteriori* (§4.4.2).

We begin in §4.2.1 by evaluating thermodynamic variables at the surface of the planet—the inner boundary of our model. In §4.2.2 we introduce the 1D equations of mass, momentum, and energy conservation. In §4.2.3 we explain the relaxation method employed to solve the conservation laws.

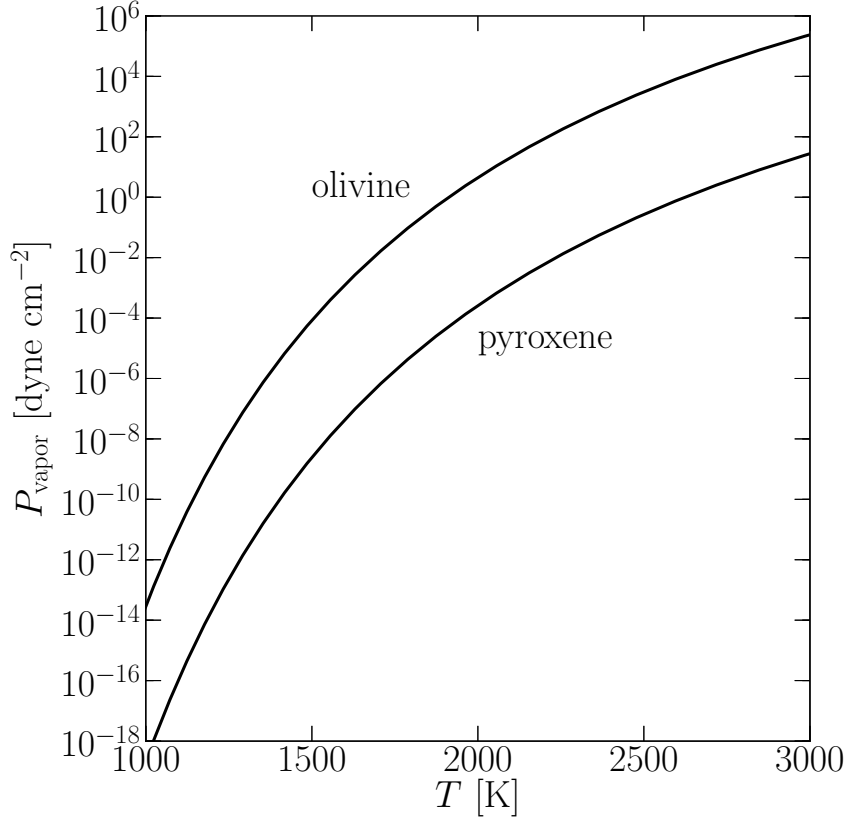


Figure 4.1 Equilibrium vapor pressures of pyroxene and olivine vs. temperature, obtained from equation (4.7).

4.2.1 Base Conditions of the Atmosphere

Dayside planet temperatures peak at approximately

$$T_{\text{surface}} \approx T_* \sqrt{\frac{R_*}{a}} e^{-\tau_{\text{surface}}/4} \approx 2150 e^{-\tau_{\text{surface}}/4} \text{ K}, \quad (4.6)$$

where for our numerical evaluation we have used parameters appropriate to KIC 12557548 (orbital semimajor axis $a \approx 0.013$ AU; effective stellar temperature $T_* \approx 4450$ K; and stellar radius $R_* \approx 0.65 R_{\odot}$; [Rappaport et al. 2012](#)). Here τ_{surface} is the optical depth between the planet's surface and the star due to absorption/scattering by dust grains.

We expect rocky extrasolar planets to consist of the same silicates predominantly found in the mantle of the Earth: pyroxene ($[\text{Fe}, \text{Mg}] \text{SiO}_3$) and olivine ($[\text{Fe}, \text{Mg}]_2 \text{SiO}_4$). At the estimated T_{surface} , these silicates will vaporize and form an atmosphere whose equilibrium

vapor pressure can be described by

$$P_{\text{vapor}} = \exp \left(-\frac{mL_{\text{sub}}}{kT_{\text{surface}}} + b \right), \quad (4.7)$$

where m is the mass of either a pyroxene or olivine molecule, k is Boltzmann's constant, L_{sub} is the latent heat of sublimation, and b is a constant. We follow [Kimura et al. \(2002\)](#) in our choice of parameters for equation (4.7). For olivine, $m = 169 \text{ m}_\text{H}$ ($\text{Mg}_{1.1}\text{Fe}_{0.9}\text{SiO}_4$) where m_H is the mass of atomic hydrogen, $L_{\text{sub}} = 3.21 \times 10^{10} \text{ erg g}^{-1}$, and $e^b = 6.72 \times 10^{14} \text{ dyne cm}^{-2}$; these values were found experimentally by evaporating forsterite (Mg_2SiO_4), an end-member of olivine ([Nagahara et al. 1994](#)). For pyroxene, $m = 60 \text{ m}_\text{H}$, $L_{\text{sub}} = 9.61 \times 10^{10} \text{ erg g}^{-1}$, and $e^b = 3.13 \times 10^{11} \text{ dyne cm}^{-2}$; these values were determined by [Hashimoto \(1990\)](#) for SiO_2 . [Kimura et al. \(2002\)](#) argue that the parameters for SiO_2 are appropriate for pyroxene because evaporation experiments of enstatite (MgSiO_3) show that SiO_2 escapes preferentially, leaving the mineral with a polycrystalline forsterite coating ([Tachibana et al. 2002](#)). We plot equation (4.7) for both olivine and pyroxene in Figure 4.1.

According to [Schaefer & Fegley \(2009\)](#), the atmosphere of a hot rocky planet at temperatures similar to T_{surface} is primarily composed of Mg, SiO, O, and O_2 . In our model we choose a mean molecular weight for the atmospheric gas of

$$\mu = 30 \text{ m}_\text{H}, \quad (4.8)$$

similar to the average molecular mass of Mg and SiO. Additionally, we set the adiabatic index of the gas to be $\gamma \equiv c_\text{P}/c_\text{V} = 1.3$, appropriate for diatomic gases at high temperatures (e.g., $\gamma_{\text{H}_2}^{T=2000\text{K}} = 1.318$, [Lange & Forker 1967](#)). The gas density at the surface of the planet is

$$\rho_{\text{vapor}} = \frac{\mu P_{\text{vapor}}}{kT_{\text{surface}}}. \quad (4.9)$$

4.2.2 Equations Solved

In our 1D hydrodynamic model, we solve the equations for mass, momentum, and energy conservation of the gas, assuming a steady flow. From mass conservation,

$$\frac{\partial}{\partial r}(r^2 \rho v) = 0, \quad (4.10)$$

where r is the radial distance from the center of the planet, and ρ and v are the density and velocity of the gas. In the frame rotating at the orbital angular frequency of the planet, momentum conservation—with the Coriolis force omitted—implies

$$\rho v \frac{\partial v}{\partial r} = -\frac{\partial P}{\partial r} - \frac{GM\rho}{r^2} + \frac{3GM_\star \rho r}{a^3}, \quad (4.11)$$

where G is the gravitational constant, $P = \rho kT/\mu$ is the gas pressure, T is the gas temperature, and M and $M_\star = 0.7M_\odot$ are the masses of the planet and the host star (KIC 12557548),

respectively. The last term on the right-hand side of eq. (4.11), which we refer to as the tidal gravity term, is the sum of the centrifugal force and the gravitational attraction from the star. In deriving eq. (4.11), we have neglected terms of order $(r/a)^2$. We do not include the contribution of the Coriolis force because its magnitude will only be comparable to the gravitational attraction of the planet when the gas moves at bulk speeds comparable to the escape velocity of the planet at its Hill sphere (a.k.a. the Roche lobe). This generally occurs near the outer boundary of our calculation, which is set by the location of the sonic point. Although the sonic point radius and Hill sphere radius are distinct quantities, we will find in practice for our models that they are close to one another, which is not surprising since the wind most easily accelerates to supersonic velocities where the effective gravity is weakest—i.e., near the Hill sphere.

The steady-state energy conservation equation, $\nabla \cdot (\rho u \mathbf{v}) = -P \nabla \cdot \mathbf{v} + \Gamma$, together with equation (4.10) and the specific internal energy $u = kT/[(\gamma - 1)\mu]$, yields

$$\rho v \frac{\partial}{\partial r} \left[\frac{kT}{(\gamma - 1)\mu} \right] = \frac{kTv}{\mu} \frac{\partial \rho}{\partial r} + \Gamma_{\text{col}} + \Gamma_{\text{lat}}. \quad (4.12)$$

The left hand side of equation (4.12) tracks changes in internal energy, while the terms on the right hand side track cooling due to PdV work, heating from gas-grain collisions, and latent heat released from grain condensation. We have omitted conduction because we found it to be negligible compared to all other terms.

For gas heating by dust-gas collisions, we assume that each collision between a gas molecule and a dust grain transfers an energy $k(T_{\text{dust}} - T)$ to the gas molecule. Then the rate of gas heating per unit volume is

$$\Gamma_{\text{col}} = \frac{3x_{\text{dust}}}{4s\rho_{\text{int}}} \frac{k}{\mu} (T_{\text{dust}} - T) \rho^2 c_s, \quad (4.13)$$

where $x_{\text{dust}}(r) \equiv \rho_{\text{dust}}/\rho$ is the spatially varying dust-to-gas mass ratio, $s = 1 \mu\text{m}$ is the assumed grain size, $\rho_{\text{int}} = 3 \text{ g cm}^{-3}$ is the grain bulk density, and

$$c_s = \sqrt{\frac{\gamma kT}{\mu}} \quad (4.14)$$

is the sound speed of the gas. As discussed in §4.4.2, our assumption that micron-sized grains are entrained in the flow is valid for the lowest mass planets we consider (a group that includes, as best we will gauge, KIC 1255b) but not the highest mass planets.

We adopt

$$T_{\text{dust}} = T_* \sqrt{\frac{R_*}{a}} e^{-\tau/4} \quad (4.15)$$

for simplicity, so that T_{dust} equals T_{surface} at the surface of the planet. The dust optical depth to starlight from r to the star is

$$\tau = \frac{3}{4s\rho_{\text{int}}} \int_r^{r_s} \rho x_{\text{dust}} dr + \tau_s. \quad (4.16)$$

The constant τ_s accounts for the optical depth from the outer boundary of our calculation—the sonic point radius r_s —to the star. We arbitrarily set

$$\tau_s = 0.01. \quad (4.17)$$

Our results are insensitive to our assumed value of τ_s as long as it is $\ll 1$. The optical depth at the planet’s surface is $\tau_{\text{surface}} \equiv \tau(r = R)$, where R is the planetary radius, calculated for a given M by assuming a constant bulk density of 5.4 g/cm^3 , the mean density of Mercury.

Latent heat from grain formation is released at a rate given by

$$\Gamma_{\text{lat}} = L_{\text{sub}} \rho v \frac{dx_{\text{dust}}}{dr}, \quad (4.18)$$

where we have used L_{sub} appropriate for pyroxene, as it is more thermally stable than olivine (see Figure 4.1; in any case, L_{sub} for olivine is not that different).

Because we do not model dust condensation from first principles, the dust-to-gas mass ratio $x_{\text{dust}}(r)$ is a free function, which we parametrize as follows:

$$\begin{aligned} \log[x_{\text{dust}}(r)] &= \log(x_{\text{dust,max}}) - \log(x_{\text{dust,amp}}) \cdot \\ &\quad \left\{ 1 - \arctan \left[\frac{8(r - R)}{3(r_s - R)} \right] \left[\arctan \left(\frac{8}{3} \right) \right]^{-1} \right\}. \end{aligned} \quad (4.19)$$

The function x_{dust} increases monotonically with r to a maximum value of $x_{\text{dust,max}}$, starting from a minimum value of $x_{\text{dust,max}}/x_{\text{dust,amp}}$. We have two reasons for choosing x_{dust} to increase rather than decrease with increasing height. The first is physical: as the wind launches into space, it expands and cools, allowing the metal-rich vapor to more easily saturate and condense (see also §4.4.8). The second is motivated by observations: on length scales on the order of the stellar radius, far from the planet, the wind must have a relatively high dust content so as to produce the deep transits observed by *Kepler*. Of course, neither of these reasons constitutes a first-principles proof that x_{dust} actually does increase with r . Indeed the increasingly low density of the wind may frustrate condensation. A physics-based calculation of how grains may (or may not) form in the wind is sorely needed (§4.4.8); our present study merely parameterizes this difficult problem.

As dust condenses from gas and x_{dust} rises, the gas density ρ must fall by mass conservation. For simplicity we neglect the dependence of ρ on x_{dust} and so restrict our analysis to $x_{\text{dust,max}} \ll 1$; in practice, the maximum value of $x_{\text{dust,max}}$ that we consider is $10^{-0.5} \approx 0.32$.

4.2.3 Relaxation Code

The structure of the wind is found by the simultaneous solution of equations (4.10), (4.11), (4.12), & (4.16) with the appropriate boundary conditions. This system constitutes a two-point boundary value problem, as the boundary conditions for ρ and T are defined at the base of the flow, while those for v and τ are defined at the sonic point.

Although generally more complicated, relaxation codes are preferred over shooting methods when solving for the transonic Parker wind. This is because for each transonic solution there are an infinite number of “breeze solutions” where the bulk speed of the wind never reaches supersonic velocities. It is easier to begin with an approximate solution that already crosses the sonic point and to then refine this solution, than to exhaustively shoot in multidimensional space for the sonic point. [Murray-Clay et al. \(2009\)](#) used a relaxation code to find the transonic wind from a hot Jupiter. Here we follow their methodology; we develop a relaxation code based on Section 17.3 of [Press et al. \(1992\)](#).

Finite Difference Equations

We set up a grid of $N = 100$ logarithmically spaced radii that run from the base of the flow located at the surface of the planet ($r = R$) to the sonic point ($r = r_s$). We replace our system of differential equations by a set of finite-difference relations. For ρ and T :

$$\begin{aligned} E_{1,j} &\equiv \Delta_j \rho - \frac{d\rho}{dr} \Delta_j r \\ &= \Delta_j \rho + \rho \left(\frac{2}{r} + \frac{1}{v} \frac{dv}{dr} \right) \Delta_j r = 0, \end{aligned} \quad (4.20)$$

$$\begin{aligned} E_{2,j} &\equiv \Delta_j T - \frac{dT}{dr} \Delta_j r \\ &= \Delta_j T - (\gamma - 1) \left[\frac{T}{\rho} \frac{d\rho}{dr} + \frac{\mu}{k\rho v} (\Gamma_{\text{col}} + \Gamma_{\text{lat}}) \right] \Delta_j r = 0, \end{aligned} \quad (4.21)$$

where finite differences are calculated upwind ($\Delta_j x \equiv x_j - x_{j-1}$), as ρ and T (and by extension E_1 and E_2) have boundary conditions defined at the surface of the planet ($j = 0$). Following [Press et al. \(1992\)](#), we average across adjacent grid points when evaluating variables in (4.20) and (4.21), e.g., $\rho \equiv (\rho_{j-1} + \rho_j)/2$. The finite-difference relations for v and τ are:

$$\begin{aligned} E_{3,j} &\equiv \Delta_j v - \frac{dv}{dr} \Delta_j r \\ &= \Delta_j v - \frac{v}{v^2 - \gamma k T / \mu} \left[\frac{2\gamma k T}{\mu r} - \frac{\gamma - 1}{\rho v} (\Gamma_{\text{col}} + \Gamma_{\text{lat}}) \right. \\ &\quad \left. - GM/r^2 + 3GM_*/a^3 \right] \Delta_j r = 0, \end{aligned} \quad (4.22)$$

$$\begin{aligned} E_{4,j} &\equiv \Delta_j \tau - \frac{d\tau}{dr} \Delta_j r \\ &= \Delta_j \tau + \frac{3}{4s\rho_{\text{int}}} \rho x_{\text{dust}} \Delta_j r = 0. \end{aligned} \quad (4.23)$$

Note that differences are now computed downwind ($\Delta_j x \equiv x_{j+1} - x_j$), as v and τ have boundary conditions defined at the sonic point ($j = N$). When evaluating variables in (4.22) and (4.23), we average across adjacent gridpoints downwind, e.g., $\rho \equiv (\rho_j + \rho_{j+1})/2$.

Boundary Conditions

We need four boundary conditions to solve the system of finite difference equations (4.20)–(4.23). At the base of the atmosphere, we set the boundary conditions for gas density and temperature using relations from §4.2.1:

$$\begin{aligned} E_{1,0} &= \rho_0 - \rho_{\text{vapor}} = 0 \\ E_{2,0} &= T_0 - T_{\text{surface}} = 0. \end{aligned} \quad (4.24)$$

At our outer boundary—the sonic point—we require the bulk velocity of gas to equal the local sound speed, and τ to equal our specified value (§4.2.2):

$$\begin{aligned} E_{3,N} &= v_N - c_s = 0 \\ E_{4,N} &= \tau_N - \tau_s = 0. \end{aligned} \quad (4.25)$$

At every step of the iteration we determine r_s by demanding dv/dr to be finite at the sonic point. As the denominator of equation (4.22) vanishes at $v = c_s$, we require that

$$\left[\frac{2\gamma kT}{\mu r} - \frac{\gamma - 1}{\rho v} (\Gamma_{\text{col}} + \Gamma_{\text{lat}}) - \frac{GM}{r^2} + \frac{3GM_* r}{a^3} \right]_{r_s} = 0. \quad (4.26)$$

Method of Solution

The relaxation method determines the solution by starting with an appropriate guess and iteratively improving it. We use the multidimensional Newton’s method as our iteration scheme, which requires us to evaluate partial derivatives of all $E_{i,j}$ with respect to all $4N$ dependent variables (ρ_j, T_j, v_j, τ_j). We evaluate partial derivatives numerically, by introducing changes of order 10^{-8} in dependent variables and computing the appropriate finite differences. Newton’s method produces a $4N \times 4N$ matrix, which we invert using the `numpy` library in `python`. At each iteration we numerically solve equation (4.26) for r_s and re-map all N gridpoints between R and r_s . We iterate until variables change by less than one part in 10^{10} .

We found our iteration scheme to be rather fragile. The code only converges if initial guesses are already close to the solution. For this reason, we started with a simplified version of the problem with a known analytic solution as the initial guess. We then gradually added the missing physics to the code, obtaining a solution with each new physics input until the full problem was solved. Specifically, we began by solving the isothermal Parker wind problem (e.g., Lamers & Cassinelli 1999). We enforced the isothermal condition by declaring γ to be unity, so that our energy equation read $dT/dr = 0$ (see eq. 4.21). Additionally, we set M_* and

x_{dust} to zero to remove the effects of tidal gravity and dust. We then slowly increased each of the parameters M_* , γ , and x_{dust} (in that order) to their nominal values. Any subsequent parameter change (e.g., M) was also performed in small increments.

In summary, our code contains three main input parameters: $x_{\text{dust, amp}}$ and $x_{\text{dust, max}}$ which prescribe the dust-to-gas profile $x_{\text{dust}}(r)$, and the planet mass M . Our goal is to explore the dependence of the mass loss rate

$$\dot{M} = \Omega \rho v r^2 \quad (4.27)$$

on these three parameters. Here Ω is the solid angle over which the wind is launched, measured from the center of the planet; we set $\Omega = 1$ since the high surface temperatures required to produce a wind are likely to be reached only near the substellar point. We independently vary the parameters $x_{\text{dust, max}}$ and $x_{\text{dust, amp}}$ to find the *maximum* $\dot{M}$ for a given M .

4.3 Results

We begin in §4.3.1 with an isothermal gas model. The isothermal model serves both as a limiting case and as a starting point for developing the full solution which includes a realistic treatment of the energy equation. We provide results for our full model in §4.3.2. In the full solution we focus on three possible planet masses $M = 0.01, 0.03$, and $0.07 M_{\oplus}$, finding that $M \approx 0.01 M_{\oplus}$ yields a *maximum* $\dot{M}$ that is compatible with the observationally inferred $\dot{M}_{1255b}$. We conclude that in the context of our energy equation, the present-day mass of KIC 1255b is *at most* $\sim 0.02 M_{\oplus}$ (since smaller mass planets can generate still higher $\dot{M}$ that are also compatible with $\dot{M}_{1255b}$). In §4.3.3 we integrate back in time to compute the *maximum* formation mass of KIC 1255b.

4.3.1 Isothermal Solution

We begin by solving a steady, dust-free, isothermal wind described by equations (4.10) and (4.11) with M_* set equal to zero. Although this is a highly idealized problem, its solution provides us with a starting point (i.e., an initial model) from which we are able to solve more complicated problems (see §4.2.3). Our code accepts the standard isothermal wind solution (e.g., Chapter 3 of Lamers & Cassinelli 1999) as an initial guess, with minimal relaxation.

Mass loss rates derived for the isothermal model ($T=2145$ K, the temperature given by eq. 4.6 with $\tau_{\text{surface}} = \tau_s = 0.01$) are plotted in Figure 4.2. They depend strongly on planet mass, with large gains in $\dot{M}$ for comparatively small reductions in M , for the basic reason that atmospheric densities are exponentially sensitive to surface gravity. Tidal gravity boosts $\dot{M}$ by reducing the total effective gravity. As M decreases, gravity becomes increasingly irrelevant, the thermal speed of the gas eventually exceeds the surface escape velocity, and the dependence of $\dot{M}$ on M weakens. In the “free-streaming” limit (dotted lines in Figure 4.2), mass loss is no longer influenced by gravity, and occurs at a rate

$$\dot{M}_{\text{free}} \approx \rho_{\text{vapor}} c_s R^2. \quad (4.28)$$

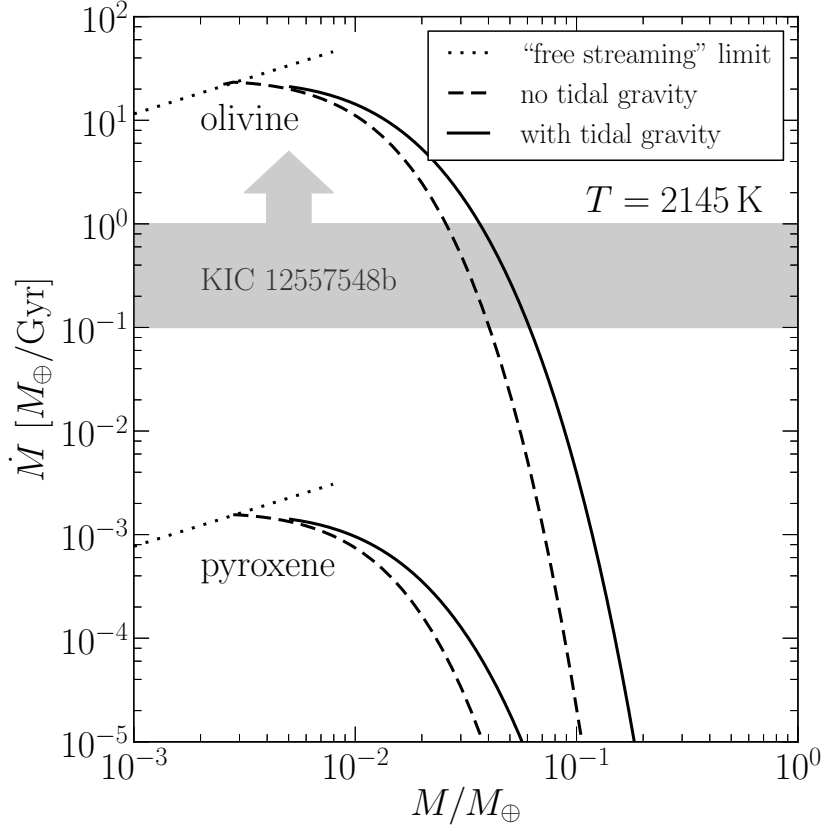


Figure 4.2 Evaporative mass loss rates $\dot{M}$ vs. planet mass M for isothermal dust-free winds. Solid lines are for models that include stellar tidal gravity, while dashed curves are for models that do not. As the planet mass is reduced, all solutions converge toward the free-streaming limit where $\dot{M}$ is not influenced by gravity but instead scales with the surface area of the planet (eq. 4.28). As inferred from observations, possible present-day mass loss rates $\dot{M}_{1255b}$ for KIC 1255b are marked in gray. Technically we have only a lower limit on $\dot{M}_{1255b}$ of $0.1M_\oplus / \text{Gyr}$; for purposes of discussion throughout this Chapter, we adopt $0.1\text{--}1M_\oplus / \text{Gyr}$ as our fiducial range (see discussion surrounding equation 4.5). Clearly KIC 1255b cannot be a pure pyroxene planet. Subsequent figures will refer to planets with pure olivine surfaces (except in §4.4.7 where we consider iron).

Note that $\dot{M}_{\text{free}}$ decreases when M is reduced, as less surface area ($R^2 \propto M^{2/3}$) is available for evaporation. For $M \sim 0.03 M_{\oplus}$, a pure olivine surface can produce a (dust-free, isothermal) wind for which $\dot{M}$ becomes compatible with the observationally inferred value for KIC 1255b, on the order of $\dot{M}_{1255b} \sim 1 M_{\oplus}/\text{Gyr}$ (this is the observed mass loss rate for both dust and gas combined; see our discussion surrounding equation 4.5 and Rappaport et al. 2012). By contrast, for a pure pyroxene surface, $\dot{M} \ll \dot{M}_{1255b}$ always. Henceforth we will calculate the density ρ_{vapor} of gas at the surface using parameters appropriate for olivine.

4.3.2 Full Solution

We relax the isothermal condition in our code by slowly increasing γ from 1 to 1.3. With this parameter change, we are solving the full energy equation (eq. 4.12). Grains are gradually added by modifying $x_{\text{dust, amp}}$ and $x_{\text{dust, max}}$, necessitating the calculation of optical depth (eq. 4.16).

In Figure 4.3 we show the solution for which $\dot{M}$ was maximized (over the space of possible values of x_{dust}) for $M = 0.03 M_{\oplus}$, or about half the mass of Mercury. Recall from §4.3.1 that when the wind was assumed to be isothermal, a planet of this mass was able to reproduce the inferred mass loss rate of KIC 1255b. Now, with our treatment of the full energy equation and the inclusion of dust, $\max \dot{M}$ plummets by a factor of 40. Relative to the isothermal solution, the reduction in mass loss rate in the full model is mainly caused by the gas temperature dropping from 2095 K at the planet surface to below 500 K at the sonic point, and the consequent reduction in gas pressure. The temperature drops because gas expands in the wind and does PdV work. Gas heating by dust-gas collisions or grain formation could, in principle, offset some of the temperature reduction, but having too many grains also obscures the surface from incident starlight, decreasing T_{surface} and ρ_{vapor} (the latter quantity depends exponentially on the former). The particular dust-to-gas profile x_{dust} that is used in Figure 4.3 is such that the ability of dust to heat gas is balanced against the attenuation of stellar flux by dust, so that $\dot{M}$ is maximized (for this M).

Note that the flow at the base begins with velocities of $\sim 10^{-3}$ the sound speed. At these subsonic velocities, the atmosphere is practically in hydrostatic equilibrium, with gas pressure and gravity nearly balancing. Only when velocities are nearly sonic does advection become a significant term in the momentum equation (bottom right panel of Figure 4.3). Flow speeds are still high enough at the base of the atmosphere to lift dust grains against gravity for all but the highest planet masses considered (§4.4.2; gas speeds as low as ~ 1 m/s in a $\sim 1 \mu\text{bar}$ atmosphere can enable micron-sized grains to escape sub-Mercury-sized planets, as can be verified by equation 4.30; furthermore, the condition that grains not slip relative to gas is conveniently independent of gas velocity in the Epstein free-molecular drag regime, as can be seen in equation 4.29. Our flows are entirely in the Epstein drag regime, as explained below equation 4.29.)

In Figure 4.4, we show how $\dot{M}$ and τ_{surface} vary with the function x_{dust} . As just discussed, $\max \dot{M}$ is reached for a specific choice of parameters which balance gas heating by dust and surface obscuration by dust. (We emphasize, here and elsewhere, that we have not identified

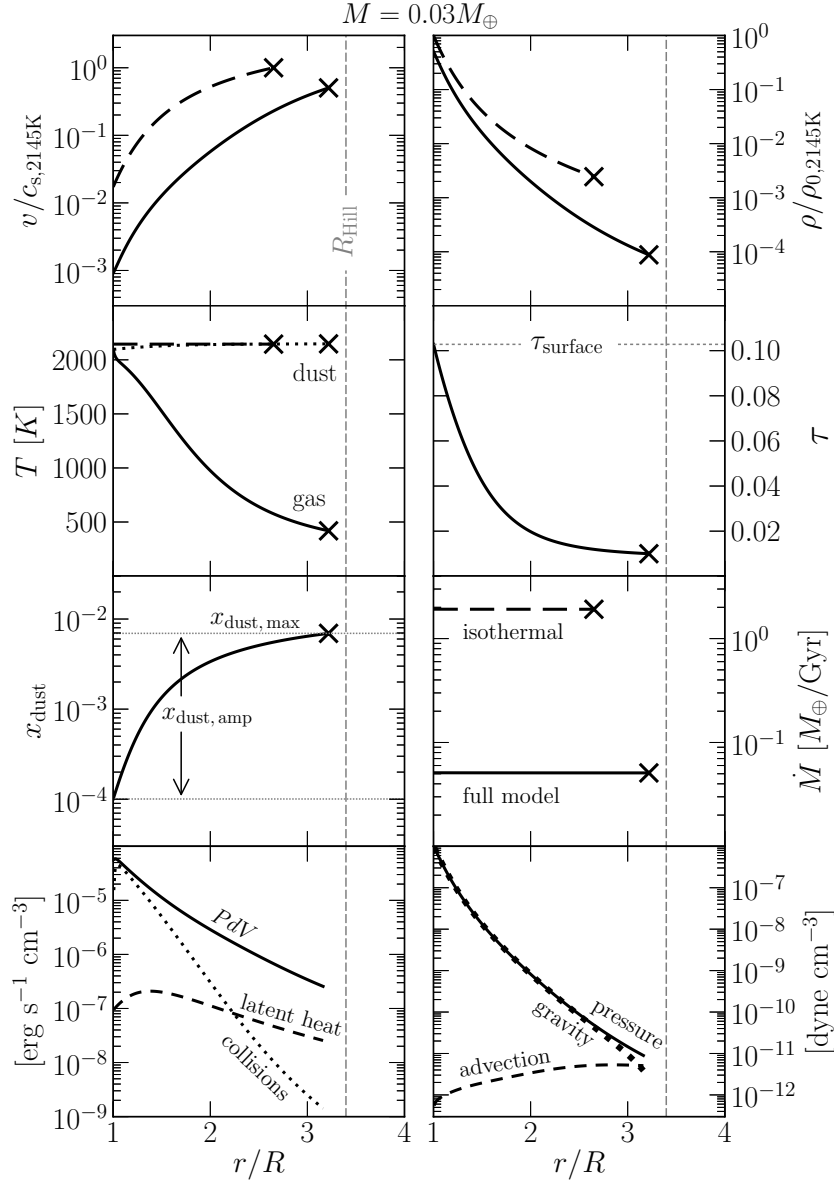


Figure 4.3 Radial dependence of wind properties in the full solution for a planet of mass $M = 0.03M_{\oplus}$. Parameters of the x_{dust} function are chosen to maximize $\dot{M}$. In the upper six panels, the full solution is shown by solid curves, while the dust-free isothermal solution is marked by dashed curves. Values for v have been normalized to the sound speed at $T = 2145$ K, and values for ρ have been normalized to the density $\rho_{0,2145K} = \mu P_{\text{vapor}}/(kT)$ evaluated also at $T = 2145$ K. The $\times$ -symbol marks the location of the sonic point (the outer boundary of our solution), which occurs close to the Hill radius R_{Hill} (marked by a vertical line). The two lower panels show the contributions of the individual terms in the energy and momentum equations (left and right panels, respectively).

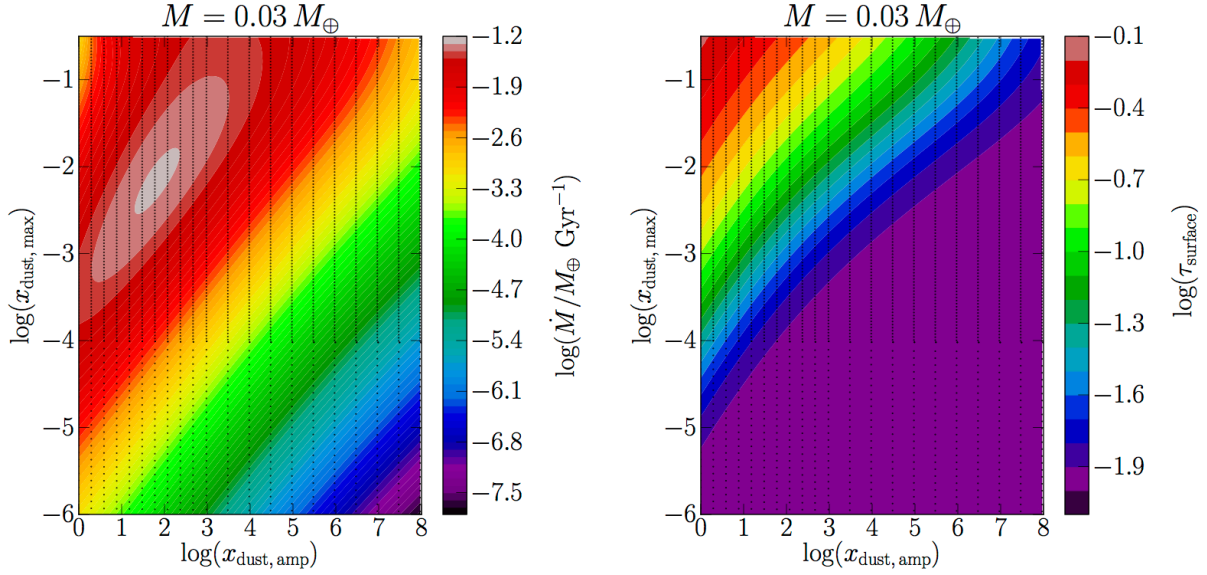


Figure 4.4 *Left panel:* Dependence of $\dot{M}$ on dust abundance for a planet of mass $M = 0.03 M_{\oplus}$. The dust abundance parameters $x_{\text{dust, max}}$ and $x_{\text{dust, amp}}$ are defined in equation (4.19); see also Figure 4.3. Both parameters are varied independently to find the maximum $\dot{M}$, which occurs at $\log(x_{\text{dust, amp}}, x_{\text{dust, max}}) \sim (1.7, -2.2)$. The full model corresponding to this maximum $\dot{M}$ is detailed in Figure 4.3. Each black dot corresponds to a full model solution. Colour contours for $\dot{M}$ are interpolated using a cubic polynomial. *Right panel:* Optical depth τ_{surface} between the star and the planetary surface as a function of dust abundance for a planet of mass $M = 0.03 M_{\oplus}$. Apart from the region near the peak in $\dot{M}$, contours of constant τ_{surface} are roughly parallel to contours of constant $\dot{M}$, suggesting that the value of τ_{surface} is more important for determining $\dot{M}$ than the specific functional form of x_{dust} .

a physical reason why actual systems like KIC 1255b should have mass loss rates $\dot{M}$ that equal their theoretically allowed maximum values.) Apart from the region near the peak in $\dot{M}$, contours of constant τ_{surface} are roughly parallel to contours of constant $\dot{M}$, suggesting that the value of τ_{surface} is more important for determining $\dot{M}$ than the specific functional form of x_{dust} . This finding increases the confidence we have in the robustness of our solution.

It is clear from Figures 4.3 and 4.4 that a planet of mass $M = 0.03 M_{\oplus}$ can only emit winds for which $\max \dot{M} < \min \dot{M}_{1255b} \sim 0.1 M_{\oplus} / \text{Gyr}$ (at least within the context of our energy equation). In Figure 4.5 we show results for $M = 0.01 M_{\oplus}$, for which $\max \dot{M} > \min \dot{M}_{1255b}$. For this M , the mass loss rate $\dot{M}$ is maximized when the planet surface is essentially unobscured by dust. At our arbitrarily chosen value for $x_{\text{dust}} \sim 3 \times 10^{-7}$, the wind is not significantly heated by dust and expands practically adiabatically. The maximum $\dot{M}$ is reached for an essentially dust-free solution because the wind is blowing at too high a speed for dust-gas collisions to be important. That is, the timescale over which gas travels from the planet’s surface to the sonic point is shorter than the time it takes a gas particle to collide with a dust grain: gas is thermally decoupled from dust. Were we to increase x_{dust} to make dust-gas heating important, the flow would become optically thick and the wind would shut down. If the model shown in Figure 4.5 does represent KIC 1255b, then the dust grains that occult the star must condense outside the sonic point, beyond the Hill sphere.

Figure 4.6 shows the solution for $M = 0.07 M_{\oplus}$. In this case $\dot{M}$ is maximized when enough dust is present to heat the gas significantly above the adiabat. At this comparatively large planet mass, initial wind speeds are low enough that a gas particle collides many times with dust over the gas travel time, so that dust-gas collisions heat the gas effectively near the surface of the planet. Further downwind, near the sonic point, latent heating overtakes PdV cooling and actually increases the temperature of the gas with increasing altitude. The wind is much more sensitive to heating near the base of the flow than near the sonic point: if we omit latent heating from our model—so that gas cools to ~ 500 K at the sonic point—then $\dot{M}$ is reduced by only a factor of two. In comparison, $\dot{M}$ drops by several orders of magnitude if dust-gas collisions are omitted.

The dependence of $\dot{M}$ on x_{dust} in both the “dust-free low-mass” and the “dusty high-mass” limits is illustrated in Figure 4.7. In the low-mass limit, $\dot{M} = \max \dot{M}$ when no dust is present. In this regime, gas moves too quickly for heating by dust to be significant; dust influences the flow only by attenuating starlight, and as long as $\tau_{\text{surface}} \ll 1$, dust hardly affects $\dot{M}$. By contrast, in the high-mass limit, $\dot{M} = \max \dot{M}$ for the dustiest flows we consider (i.e., the highest values of $x_{\text{dust, max}}$). In this regime, $\dot{M}$ is very sensitive to dust abundance, with values spanning six orders of magnitude over the explored parameter space.

In Figure 4.8 we show mass loss rates as a function of planet mass for the full model. We emphasize that these are maximum mass loss rates, found by varying x_{dust} . At low planet masses, $\max \dot{M}$ for the full model converges with $\dot{M}$ for the isothermal model because both approach the free-streaming limit, where gravity becomes irrelevant and $\dot{M}$ is set entirely by conditions at the surface of the planet (eq. 4.28). To reach $\dot{M}_{1255b} > 0.1 M_{\oplus} / \text{Gyr}$, the present-day mass of KIC 1255b must be $\lesssim 0.02 M_{\oplus}$, or less than about twice a lunar mass.

We have verified *a posteriori* for the models shown in Figures 4.3, 4.5, and 4.6 that the

sonic point is attained at an altitude where the collisional mean free path of gas molecules is smaller than r_s , so that the hydrodynamic approximation embodied in equations (4.10)–(4.12) is valid. In other words, the exobase lies outside the sonic point in these $\dot{M} = \max \dot{M}$ models. The margin of safety is largest for the lowest mass models.

4.3.3 Mass-Loss Histories

We calculate mass-loss histories $M(t)$ by time-integrating our solution, shown in Figure 4.8, for $\max \dot{M}(M)$. Before we integrate, however, we introduce $0 < f_{\text{duty}} < 1$ to account for the duty cycle of the wind: we define $f_{\text{duty}} \cdot \max \dot{M}$ as the time-averaged mass-loss rate. We estimate that $f_{\text{duty}} \sim 0.5$ based on the statistics of transit depths compiled by Brogi et al. (2012, see their figure 2). We time-integrate $f_{\text{duty}} \cdot \max \dot{M}$ to obtain the $M(t)$ curves shown in Figure 4.9. (This factor of 2 correction for the duty cycle should not obscure the fact that the actual mass loss rate, time-averaged or otherwise, could still be much lower than the theoretically allowed $\max \dot{M}$, a possibility we return to throughout this Chapter.)

We highlight the case of a planet with a lifetime of $t_{\text{life}} = 5$ Gyr. Under our full (non-isothermal) model (right panel of Figure 4.9), such a planet has a mass at formation of $0.06 M_{\oplus}$ and gradually erodes over several Gyr until it reaches $M \sim 0.03 M_{\oplus}$ —whereupon the remaining mass is lost catastrophically over a short time (just how short is estimated for the specific case of KIC 1255b below). Contrast this example with planets having initial masses $\gtrsim 0.07 M_{\oplus}$ —these have such low $\dot{M}$ that they barely lose any mass over Gyr timescales.

For comparison, we also present mass-loss histories using a dust-free isothermal model with stellar tidal gravity (left panel of Figure 4.9). Compared to our non-isothermal solutions—for which gas temperatures fall immediately as the wind lifts off the planet surface—the isothermal solution corresponds to a flow which stays relatively pressurized and which therefore enjoys the largest $\dot{M}$ for a given M . The isothermal wind thus represents an endmember case. However, the mass-loss histories under the isothermal approximation do not differ qualitatively from those using the full energy equation. For isothermal winds, the dividing mass between planet survival and destruction within 10 Gyr is about $0.11 M_{\oplus}$, only $\sim 40\%$ larger than the value cited above for our non-isothermal solutions.

4.3.4 Possible Mass Loss Histories for KIC 1255b

Figures 4.8 and 4.9 can be used to sketch possible mass-loss histories for KIC 1255b. For our first scenario, we employ the full model that solves the full energy equation. Today, to satisfy the observational constraint that $\dot{M}_{1255b} > 0.1 M_{\oplus}/\text{Gyr}$ (see the derivation of our equation 4.5), the planet must have a present-day mass of *at most* $M \approx 0.02 M_{\oplus}$, or roughly twice the mass of the Moon (Figure 4.8). Such a low mass implies that currently KIC 1255b is in a “catastrophic evaporation” phase (vertical straight lines in Figure 4.9). Depending on the age of the planet (i.e., the age of the star: 1–10 Gyr), KIC 1255b originally had a mass of *at most* 0.04 – $0.07 M_{\oplus}$: 2–4 times its maximum current mass, or about the mass of Mercury. Starting from today, the time the planet has before it disintegrates completely is on the order

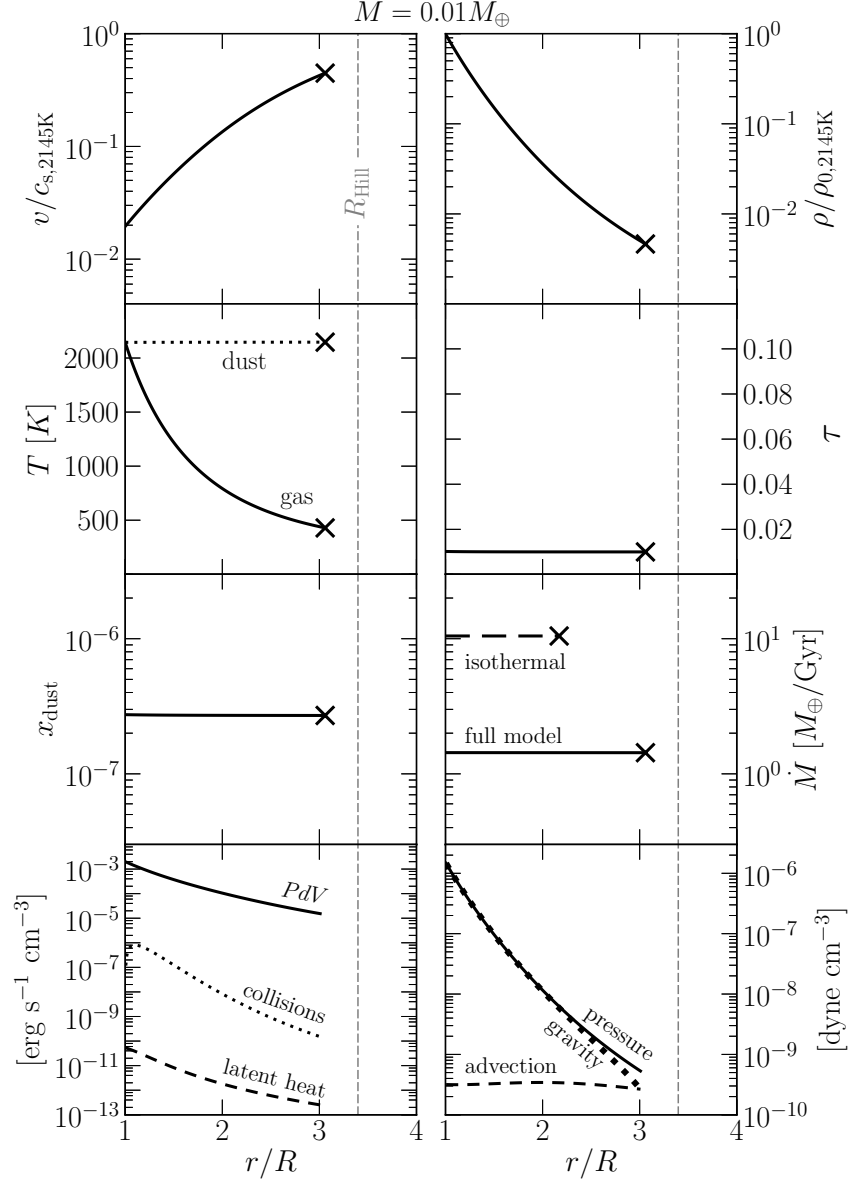


Figure 4.5 Same as Figure 4.3, but for a planet of mass $M = 0.01 M_{\oplus}$. For this planet mass, $\dot{M}$ is maximized when essentially no dust is present (i.e., the atmosphere is essentially transparent) as gas moves too quickly for dust-gas collisions to heat the gas. As such, all heating terms are negligible and the gas expands adiabatically.

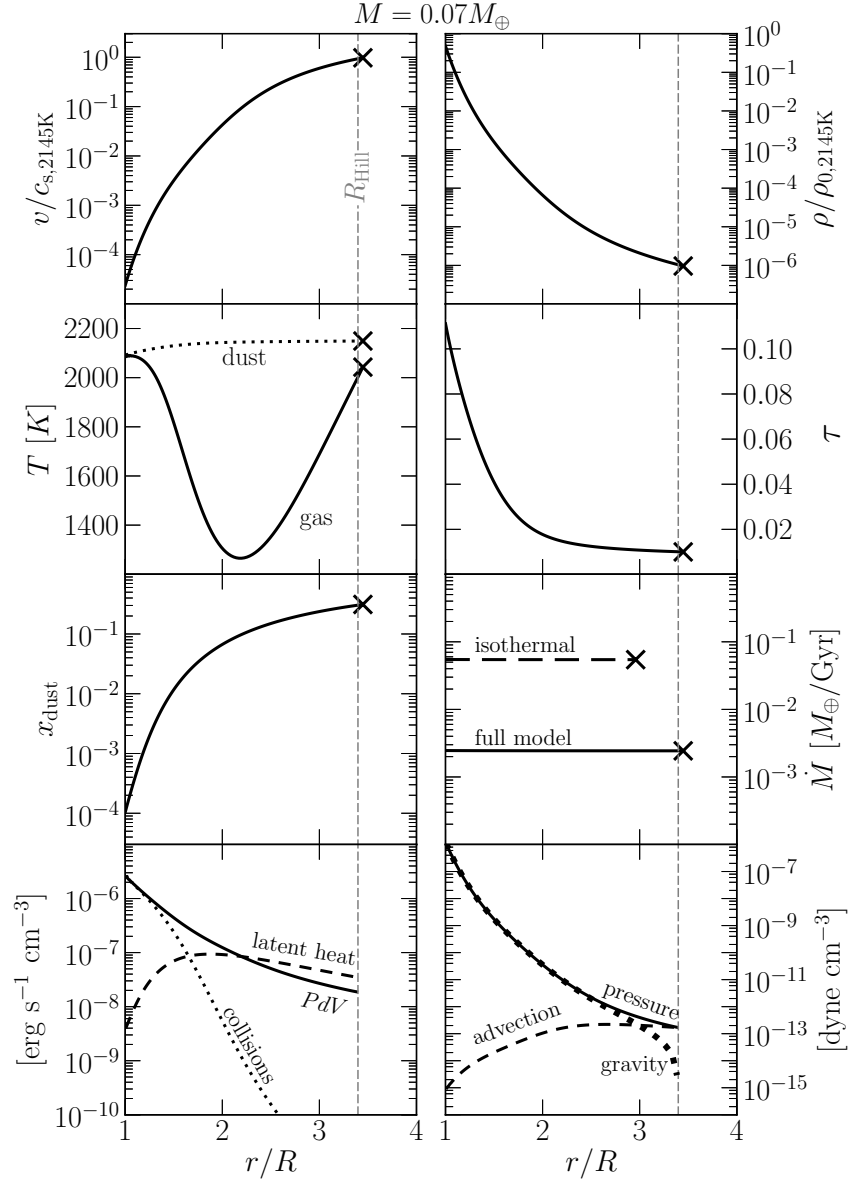


Figure 4.6 Same as Figure 4.3, but for a planet of mass $M = 0.07 M_{\oplus}$. For this mass, $\dot{M}$ is maximized for the dustiest flows considered in this work. The outflow is launched at such low velocities that collisional dust-gas heating is important. Near the sonic point, latent heating overtakes PdV cooling and the temperature of the gas rises.

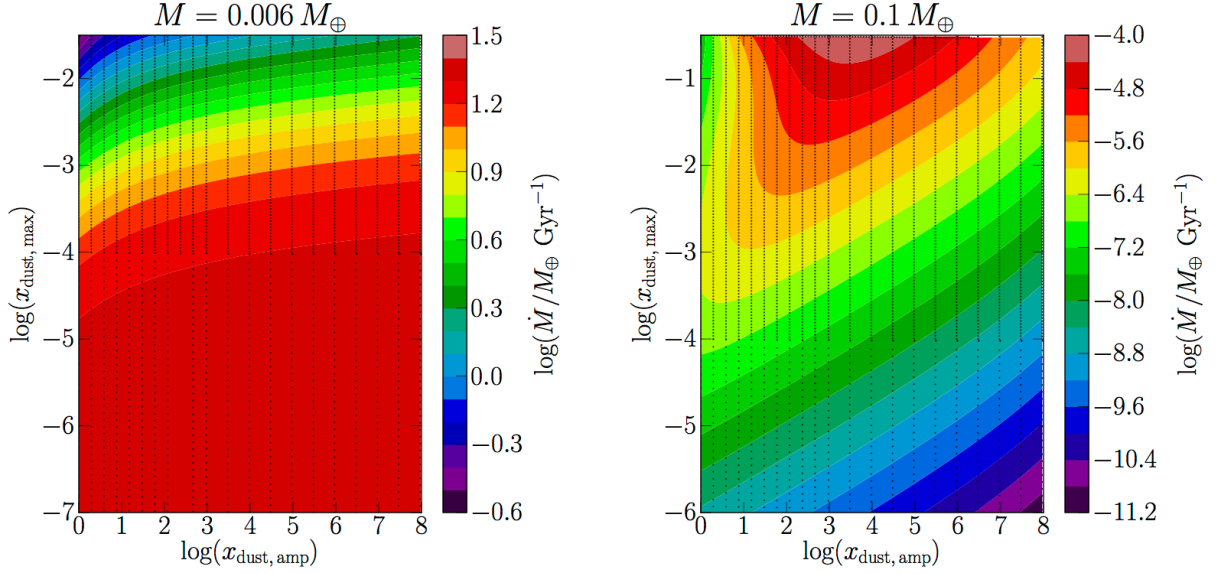


Figure 4.7 *Left panel:* Same as the left panel of Figure 4.4, but for a planet of mass $M = 0.006 M_{\oplus}$. In this low-mass limit, $\dot{M}$ is maximized when no dust is present (i.e., when the atmosphere is transparent) because gas is moving too quickly for heating by dust to be significant. *Right panel:* Same as left panel, but for a planet of mass $M = 0.1 M_{\oplus}$. In this high-mass limit, $\dot{M}$ is maximized for the dustiest flows we consider (i.e., the highest values of $x_{\text{dust, max}}$) because flow speeds near the surface are slow enough for dust-gas energy exchange to be significant.

of $t_{\text{evap}} \sim M / (f_{\text{duty}} \cdot \max \dot{M})$, which is as long as 400 (40) Myr for $M = 0.02 M_{\oplus}$, $f_{\text{duty}} = 0.5$, and $\max \dot{M} = 0.1$ (1) $M_{\oplus} / \text{Gyr}$.

If our energy equation were somehow in error and the wind better described as isothermal, then the numbers cited above would change somewhat. The present-day mass of KIC 1255b would be at most $M \sim 0.07 M_{\oplus}$; the maximum formation mass would be between 0.08 – $0.11 M_{\oplus}$; and the evaporation timescale starting from today could be as long as $t_{\text{evap}} \sim 400$ Myr if the present-day mass $M \sim 0.07 M_{\oplus}$.

As we have stressed throughout, a critical assumption we have made is that the time-averaged mass loss rate is $f_{\text{duty}} \cdot \max \dot{M}$ with f_{duty} as high as 0.5. We have not identified a physical reason why the actual mass loss rate should be comparable to the theoretically allowed $\max \dot{M}$ for a given M . If our assumption were in error and actual mass loss rates $\dot{M} \ll \max \dot{M}$, then the present-day mass of KIC 1255b, the corresponding evaporation time, and the formation mass of KIC 1255 would all decrease from the upper bounds cited above.

The degeneracy of possible masses and evaporative histories outlined in this section could be broken with observational searches for the progenitors of KIC 1255b-like objects (§4.4.6).

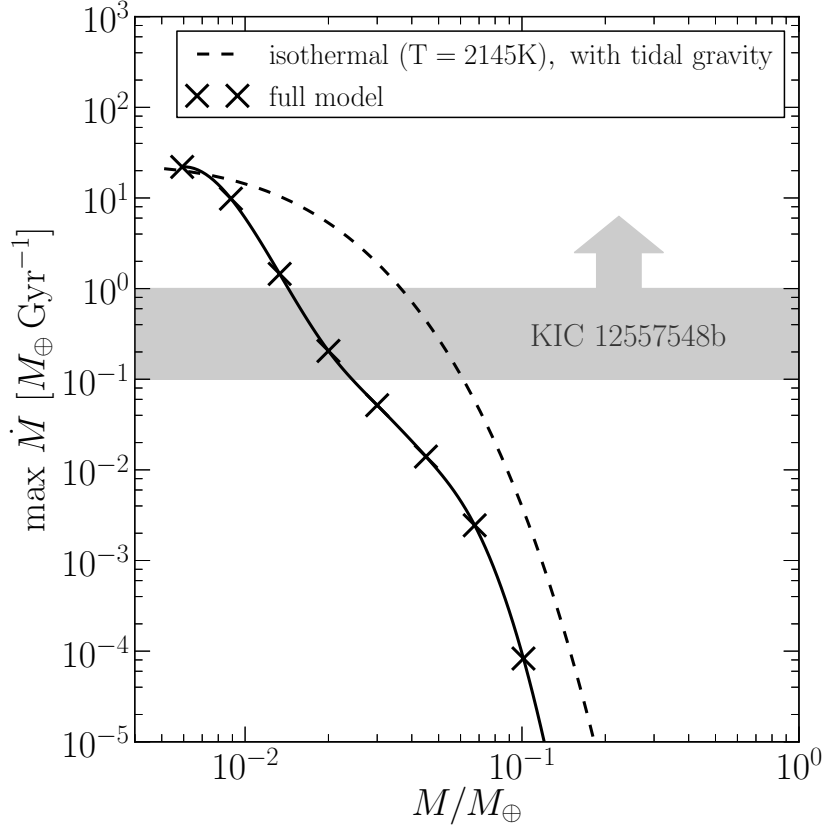


Figure 4.8 Maximum mass loss rates $\dot{M}$ vs. planet mass M for the full model. At each $\times$ -marked mass, $\max \dot{M}$ was found by varying x_{dust} as described, e.g., in Figure 4.4. The solid curve is a cubic spline interpolation. Mass loss rates for the full model are generally lower than for the isothermal model (dashed curve) because the dust-gas heating terms we have included in our full model turn out to be inefficient. At low planet masses, the isothermal and full models converge because both approach the free-streaming limit, where gravity becomes irrelevant and $\dot{M}$ is set entirely by conditions at the surface (i.e., surface area, equilibrium vapor pressure, and sound speed; eq. 4.28). According to the full model, the present-day mass of KIC 1255b is $\lesssim 0.02 M_{\oplus}$, or less than twice the mass of the Moon; for such masses, $\max \dot{M} > \min \dot{M}_{1255b} \sim 0.1 M_{\oplus} / \text{Gyr}$.

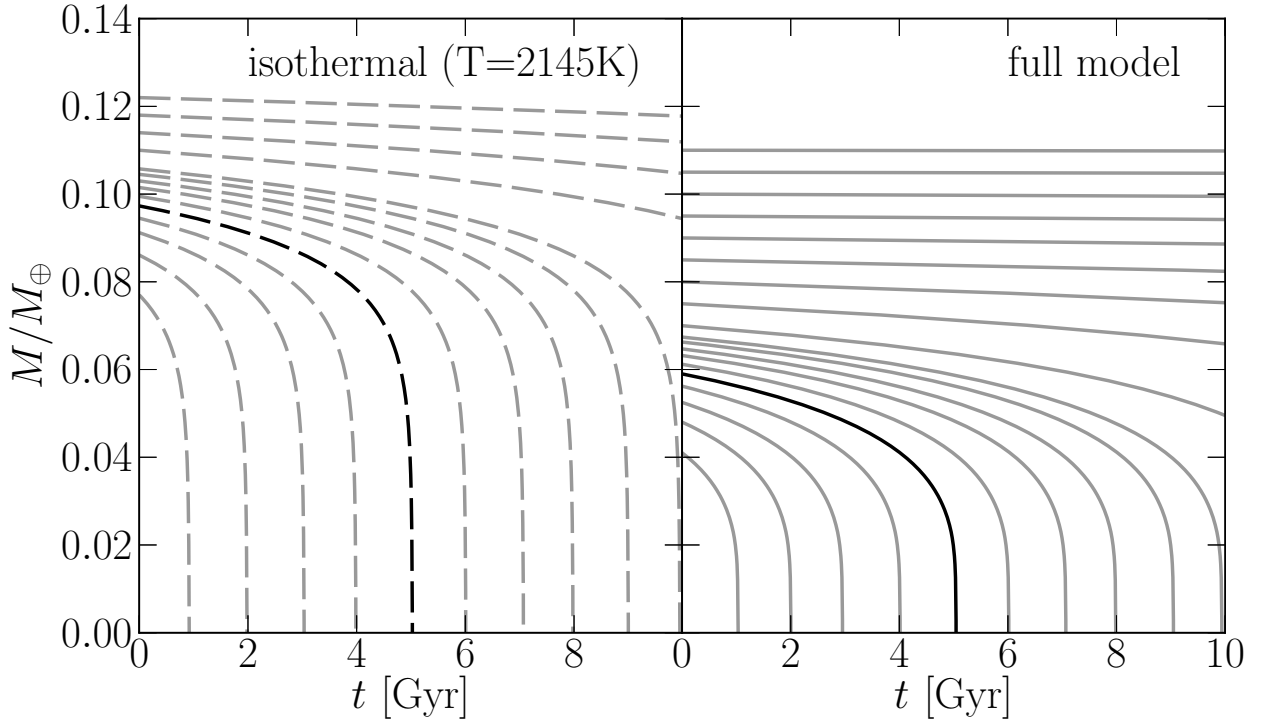


Figure 4.9 Mass-loss histories $M(t)$ obtained by time-integrating $f_{\text{duty}} \cdot \max \dot{M}(M)$ with $f_{\text{duty}} = 0.5$ for the isothermal and full models. We highlight the case of a planet with a 5 Gyr lifetime. For the full model (right panel), this corresponds to an initial mass of $\sim 0.06 M_{\oplus}$, slightly larger than the mass of Mercury. Such a planet could have formed *in situ* and slowly eroded over several Gyr until reaching $\sim 0.03 M_{\oplus}$, whereupon the planet evaporates completely in a few hundred Myr. By contrast, planets with formation masses $\gtrsim 0.07 M_{\oplus}$ survive (in the full model) for tens of Gyr without significant mass loss.

4.4 Discussion

We discuss how mass loss is not “energy-limited” in §4.4.1; how gas entrains dust in §4.4.2; how the evaporative wind can be time-variable in §4.4.3; the effects of winds on orbital evolution in §4.4.4; the interaction of the planetary wind with the stellar wind in §4.4.5; the occurrence rate of quiescent progenitors of catastrophically evaporating planets in §4.4.6; the possibility that evaporating planets can be naked iron cores in §4.4.7; and the ability of dust to condense out of the wind in §4.4.8.

4.4.1 Mass Loss is Not Energy-Limited

Atmospheric mass loss rates are commonly assumed to be “UV-energy-limited”: it is assumed that a fixed, order-unity fraction ϵ of the incident UV flux F_{UV} does PdV work and lifts material out of the planet’s gravitational well. Under the energy-limited assumption, the mass loss rate is $\dot{M} \sim \epsilon \pi R^2 F_{\text{UV}} / (GM/R)$. The validity of this formula has been tested by [Watson et al. \(1981\)](#) for terrestrial atmospheres and by [Murray-Clay et al. \(2009\)](#) for hot Jupiters.

The energy-limited formula does not apply at all to the evaporating rocky planets considered in this Chapter. The stellar UV flux is not essential for low-mass planets because their escape velocities are so small that energy deposition by photoionization is not necessary for driving a wind. Even optical photons—essentially the stellar bolometric spectrum—can vaporize close-in planets. Merely replacing F_{UV} with $F_{\text{bolometric}}$ in the energy-limited formula would still be misleading, however, because most of the incident energy is radiated away and does no mechanical work. More to the point, ϵ is not constant; as discussed in §4.3.2, the energetics of the flow changes qualitatively with planet mass. If we were to insist on using the energy-limited formula, we would find $\epsilon \sim 10^{-8}$ for $M \sim 0.1M_{\oplus}$, and $\epsilon \sim 10^{-4}$ for $M \sim 0.01M_{\oplus}$ (see Figure 4.8).

4.4.2 Dust-Gas Dynamics

In our model we have assumed that dust grains condensing within the wind are carried along without any relative motion between dust and gas (1-fluid approximation). Grains are well entrained if their momentum stopping times

$$t_{\text{stop}} \sim \frac{m_{\text{grain}} v_{\text{rel}}}{\rho v_{\text{rel}} c_s s^2} \sim \frac{s \rho_{\text{int}}}{\rho c_s} \quad (4.29)$$

are shorter than the grain advection time $t_{\text{adv}} \sim R/v$. In equation (4.29), m_{grain} is the mass of an individual grain and v_{rel} is the relative gas-grain velocity. The aerodynamic drag force in the denominator is given by the Epstein law, which is appropriate for $v_{\text{rel}} \lesssim c_s$ and grain sizes s smaller than the gas collisional mean free path (for our flows, $\lambda_{\text{mfp}} \sim \mu/(\rho\sigma) \gtrsim 2$ cm, where $\sigma \sim 3 \times 10^{-15}$ cm²).

We find using our full model (for $\dot{M} = \max \dot{M}$) that $t_{\text{stop}}/t_{\text{adv}} < 1$ everywhere for $M < 0.03 M_{\oplus}$, confirming our assumption that micron-sized (and smaller) grains are well-entrained

in the winds emanating from such low-mass planets. When $M = 0.03 M_{\oplus}$, $t_{\text{stop}} < t_{\text{adv}}$ close to the planet’s surface, but as gas nears the sonic point, t_{stop} becomes comparable to t_{adv} . For $M > 0.03 M_{\oplus}$, the 1-fluid approximation breaks down near the sonic point. For such large planet masses, future models should account for gas-grain relative motion—in addition to other effects that become increasingly important near the sonic point / Hill sphere boundary (e.g., Coriolis forces and stellar radiation pressure). Note that the 1-fluid approximation should be valid for the present-day dynamics of KIC 1255b, since according to the full model its current mass is at most $M \approx 0.02 M_{\oplus} < 0.03 M_{\oplus}$.

Although dust can slip relative to gas, dust can still be transported outward and escape the planet if the aerodynamic drag force exceeds the force of gravity:

$$\rho v_{\text{rel}} c_s s^2 \gtrsim \rho_{\text{int}} s^3 g, \quad (4.30)$$

where $g = G(M/r^2 - 3M_*r/a^3)$ is the total gravitational acceleration. Grains that manage to be lifted beyond the Hill sphere (which is situated close to the sonic point in all our models) are no longer bound to the planet. Outside the Hill sphere, dust decouples from gas and is swept into a comet-like tail by stellar radiation pressure and Coriolis forces.

For $M \approx 0.03 M_{\oplus}$ and our adopted grain size $s = 1 \mu\text{m}$, the forces in equation (4.30) are comparable near the sonic point when $v_{\text{rel}} \sim c_s$ (i.e., when grains are barely lifted by drag). For $M > 0.03 M_{\oplus}$, micron-sized and larger grains will not be dragged past the sonic point according to (4.30)—thus our assumption that they do fails for such high-mass planets, and will need to be rectified in future models. Smaller, sub-micron-sized grains can, however, be lifted outward. Moreover, sub-micron sized grains may heat the gas more effectively because they have greater geometric surface area for dust-gas collisions (at fixed x_{dust}) and because their temperatures exceed those of blackbodies (their efficiencies for emission at infrared wavelengths are much less than their absorption efficiencies at optical wavelengths). The superior momentum and thermal coupling enjoyed by grains having sizes $s \ll 1 \mu\text{m}$, plus their relative transparency at optical wavelengths—which helps them avoid shadowing the planet surface from stellar radiation—motivate their inclusion in the next generation of models.

Our model could be improved still further by self-consistently accounting for how the gas density ρ must decrease as dust grains condense out of gas. Moreover, the restriction that $x_{\text{dust}} < 1$ could be lifted.

4.4.3 Time Variability

Occultations of KIC 12557548 vary in depth from a maximum of 1.3% to a minimum of $\lesssim 0.2\%$ on orbital timescales. [Rappaport et al. \(2012\)](#) discussed qualitatively the possibility that such transit depth variations arise from a limit cycle that alternates between high- $\dot{M}$ and low- $\dot{M}$ phases. A high- $\dot{M}$ phase that produces a deep eclipse would also shadow the planet surface from starlight. The resultant cooling would lower the surface vapor pressure and lead to a low- $\dot{M}$ phase—after which the atmosphere would clear, the surface would re-heat, and the cycle would begin anew. Such limit cycle behavior could be punctuated by random

explosive events that release dust, similar to those observed on Io (Geissler 2003; see also the references in §4.2 of Rappaport et al. 2012).

Order-of-magnitude variations in $\dot{M}$ arise from only small fractional changes in surface temperature because the vapor pressure of gas over rock depends exponentially on temperature. For example, from equations (4.6) and (4.7), we see that increasing the surface optical depth τ_{surface} from 0.1 to 0.4 reduces the surface temperature T_{surface} by 150 K and the base density ρ_{vapor} by a factor of 10. Such changes would cause $\dot{M}$ to drop by more than a factor of 10, because $\dot{M}$ scales super-linearly with ρ_{vapor} : a linear dependence results simply because $\dot{M}$ scales with gas density, while an additional dependence arises because the wind speed increases with the gas pressure gradient, which in turn scales as $\rho_{\text{vapor}} T_{\text{surface}}$.

To reproduce the orbit-to-orbit variations in transit depth observed for KIC 1255b, the dynamical time t_{dyn} of the wind cannot be much longer than the planet’s orbital period of $P_{\text{orb}} = 15.7$ hr. The dynamical time is that required for dust to be advected from the planet surface to the end of the comet-like tail that occults the star: it is the minimum timescale over which the planet’s transit signature “refreshes”. If $t_{\text{dyn}} \gg P_{\text{orb}}$, then we would expect transit depths to be correlated from one orbit to the next—in violation of the observations.

Referring to the full model for a KIC 1255b-like mass of $M \approx 0.01 M_{\oplus}$ (Figure 4.5), we estimate that

$$t_{\text{dyn}} = \int_R^{r_s} \frac{dr}{v} + \int_{r_s}^{0.1R_*} \frac{dr}{v} \sim 13 \text{ hr} + \frac{0.1R_*}{v_2} \sim 14 \text{ hr}, \quad (4.31)$$

where we have split t_{dyn} into two parts: the first integral is the time for dust to reach the sonic point (the outer boundary of our calculation), while the second integral is the time for dust to travel out to $0.1R_*$ (a circular, optically disk of this radius would generate a 1% transit depth). The first integral is performed numerically using our full model (upper left panel of Figure 4.5), while the second integral is estimated to order-of-magnitude using a characteristic grain velocity (well outside the planet’s Hill sphere) of $v_2 \sim 10 \text{ km s}^{-1}$ (Rappaport et al. 2012; see their equation 6 and related commentary). Since $t_{\text{dyn}} \lesssim P_{\text{orb}}$, we conclude that the wind/cometary tail can refresh itself quickly enough to change its appearance from orbit to orbit.

We can go one step further. The timescale over which $\dot{M}$ changes should be the timescale over which the stellar insolation at the planetary surface changes—in other words, the timescale over which the “weather” at the substellar point, where the wind is launched, changes from, e.g., “overcast” to “clear”. Ignoring the possibility of volcanic eruptions, we estimate this variability timescale as the time for the wind to reach the Hill sphere boundary, at which point the Coriolis force has turned the wind by an order-unity angle away from the substellar ray joining the planet to the star (i.e., beyond the Hill sphere, the dust-laden wind no longer blocks stellar radiation from hitting the substellar region where the wind is launched). Because the Hill radius is situated near the sonic point, this variability timescale is given approximately by the first integral in equation (4.31): $t_{\text{dyn}}^a \approx 13 \text{ hr}$.

The fact that t_{dyn}^a is neither much shorter than nor much longer than P_{orb} supports our proposal that the observed time variability of KIC 1255b is driven by a kind of limit cycle involving stellar insolation and mass loss. Clearly if $t_{\text{dyn}}^a \gg P_{\text{orb}}$, orbit-to-orbit variations

in $\dot{M}$ would be impossible. Conversely, if $t_{\text{dyn}}^a \ll P_{\text{orb}}$ —or more precisely if t_{dyn}^a were much shorter than the transit duration of 1.5 hr—then the wind would vary so rapidly that each transit observation would time-integrate over many cycles, yielding a smeared-out average transit depth that would not vary from orbit to orbit as is observed.

4.4.4 Long-Term Orbital Evolution

When computing planet lifetimes in §4.3.3, we have assumed that the planet does not undergo any orbital evolution while losing mass. This approximation is valid because gas leaves the planet’s surface at velocities $v_{\text{launch}} \lesssim c_s \sim 1 \text{ km s}^{-1}$. Launch velocities are so low compared to the orbital velocity of the planet $v_{\text{orb}} \sim 200 \text{ km s}^{-1}$ that the total momentum imparted by the wind is a tiny fraction of the planet’s orbital momentum.

We quantify this as follows. We estimate the total change in semimajor axis a and eccentricity e from Gauss’ equations (see, e.g., [Murray & Dermott 1999](#)). Consider the case in which the gas is launched at a velocity v_{launch} at an angle α from the substellar ray and in the plane of the planet’s orbit. The angle α could be non-zero because of surface inhomogeneities or asynchronous rotation of the planet. Neglecting terms of order e^2 and αe , we find that a and e evolve according to

$$\begin{aligned} \frac{da}{dt} &\sim a \frac{v_{\text{launch}}}{v_{\text{orb}}} \frac{\dot{M}}{M} (e + \alpha), \\ \frac{de}{dt} &\sim \frac{v_{\text{launch}}}{v_{\text{orb}}} \frac{\dot{M}}{M} (1 + \alpha), \end{aligned} \quad (4.32)$$

where the first term in parentheses on the right-hand-side of each equation is due to the radial component of the perturbation and the second term is due to the azimuthal component (where radius and azimuth are measured in a cylindrical coordinate system centered at the star and in the plane of the planet’s orbit). Integrating equations (4.32) yields the changes in a and e accumulated over the planet’s age:

$$\begin{aligned} \frac{\Delta a}{a} &\sim \ln \left(\frac{M_{\text{initial}}}{M_{\text{final}}} \right) \frac{v_{\text{launch}}}{v_{\text{orb}}} (e + \alpha) \lesssim 0.01 (e + \alpha), \\ \Delta e &\sim \ln \left(\frac{M_{\text{initial}}}{M_{\text{final}}} \right) \frac{v_{\text{launch}}}{v_{\text{orb}}} \lesssim 0.01. \end{aligned} \quad (4.33)$$

When evaluating equation (4.33), we have used the maximum possible launch velocity $v_{\text{launch}} \sim c_s$ (valid in the “free-streaming” limit) and $M_{\text{initial}}/M_{\text{final}} \sim 5$ (see §4.3.3).

The fact that evaporating planets undergo negligible orbital evolution suggests they could have resided on their current orbits for Gyrs—indeed they might even have formed *in situ*. [Swift et al. \(2013\)](#) would disfavor *in-situ* formation of KIC 1225b because at the planet’s orbital distance, dust grains readily sublimate, and ostensibly there would have been no solid material in the primordial disk out of which rocky planets could have formed. However,

a less strict *in-situ* formation scenario is still viable. Solid bodies larger than dust grains obviously have longer evaporation times. Such large planetesimals could have drifted inward through the primordial gas disk and then assembled into the progenitors of objects like KIC 1255b at their current close-in distances (e.g., Youdin & Shu 2002; Hansen & Murray 2012; Chiang & Laughlin 2013). Solid particles can avoid vaporization if they merge faster than they evaporate.

4.4.5 Flow Confinement by Stellar Wind

When computing the transonic solution for the planetary wind, we implicitly assumed that the outflow was expanding into a vacuum. In reality, the outflow from the planet will collide with the stellar wind and form a “bubble” around the planet.³ For our solution to be valid, the radius of the bubble—i.e., the surface of pressure balance between the winds—has to be downstream of the sonic point where the flow is supersonic. This way, any pressure disturbance created at the interface cannot propagate upstream and influence our solution.

At the wind-wind interface, the normal components of the pressures of the planetary and stellar winds balance. The total pressure includes the ram, thermal, and magnetic contributions. We compare the pressure of the stellar wind at $a \sim 0.01 \text{ AU} \sim 4R_*$ with that of the planetary wind computed at r_s to gauge whether the interface occurs downstream of r_s . For the pressure P_* of a main-sequence, solar-type star we are guided by the Solar wind. Flow speeds of the Solar wind are measured by tracking the trajectories of coronal features (Sheeley et al. 1997; Quémerais et al. 2007). At $a = 0.01 \text{ AU}$ the Solar wind⁴ is still accelerating with typical flow speeds of $v_* \sim 100 \text{ km s}^{-1}$. For this v_* , and a canonical Solar mass-loss rate of $2 \times 10^{-14} \text{ M}_\odot \text{ yr}^{-1}$, the proton number density is $n_* = 2 \times 10^5 \text{ cm}^{-3}$. We take the proton temperature to be $T_* \sim 10^6 \text{ K}$ (Sheeley et al. 1997) and the heliospheric magnetic field strength to be $B_* \approx 0.1 \text{ G}$ at 0.01 AU (Kim et al. 2012). The total stellar pressure at 0.01 AU is then $P_* \sim n_* m_H v_*^2 + n_* k T_* + B_*^2 / (8\pi) \sim 4 \times 10^{-4} \text{ dyne cm}^{-2}$, with the magnetic pressure dominating other terms by an order of magnitude.

The ram and thermal pressures of the planet’s wind are equal at the sonic point and add up to (for $\dot{M} = \max \dot{M}$) $P_s \sim 2 \times 10^{-3} (1 \times 10^{-4}) \text{ dyne cm}^{-2}$ when $M = 0.03 (0.07) M_\oplus$. Thus for $M \lesssim 0.03 M_\oplus$ (a range that includes possible present-day masses of KIC 1255b), $P_s > P_*$ so that the planetary wind blows a bubble that extends beyond the sonic point and the transonic solutions we have computed are self-consistent. By contrast for $M \sim 0.07 M_\oplus$, $P_s \sim P_*/4 < P_*$ so that the stellar wind pressure will balance the planetary wind pressure inside of r_s . The stellar wind will prevent the planetary wind from reaching supersonic

³The considerations in this section parallel those for hot Jupiter winds and their bow shocks; see Tremblin & Chiang (2013), Vidotto et al. (2010, 2011a,b), and Llama et al. (2011).

⁴These stellar wind parameters are appropriate for the “slow” Solar wind blowing primarily near the equatorial plane of the Sun. In addition, the Solar wind also contains a “fast” component which emerges from coronal holes. During Solar minimum, the fast Solar wind is confined mainly to large heliographic latitudes, but may reach the equator plane during periods of increased solar activity (Kohl et al. 1998; McComas et al. 2003). Because most planets are expected to orbit near their stellar equatorial planes, we take the slow component of the Solar wind to be a better guide.

velocities; the planetary outflow will conform to a “breeze” solution with a reduced $\dot{M}$. If the colliding winds in our problem behave like those of hot Jupiters and their host stars, then $P_s \sim P_*/4$ will reduce $\dot{M}$ by $\sim 30\%$ (see figure 12 of [Murray-Clay et al. 2009](#)).

We have found P_* to be dominated by magnetic pressure. Since magnetic fields only interact with the ionized component of the planetary wind, and since the planetary wind may not be fully ionized, we might be over-estimating the effect of P_* . Dust grains in the planetary wind may absorb free charges so that the wind may be too weakly ionized to couple with the stellar magnetic field. If we ignore the magnetic contribution to the stellar wind, then $P_s > P_*$ for all planet masses that we have considered, and all our transonic solutions would be self-consistent.

What about the planet’s magnetic field? A sufficiently ionized outflow could be confined by the planet’s own magnetic field. To assess the plausibility of this scenario, we use Mercury’s field as a guide for KIC 1255b. The flyby of *Mariner 10* and recent measurements by *MESSENGER* show that Mercury possesses a weak magnetic field with a surface strength of $\sim 3 \times 10^{-3}$ Gauss ([Ness et al. 1974](#); [Anderson et al. 2008](#)). The associated magnetic pressure $\sim 4 \times 10^{-7} (R/r)^6$ dyne cm^{-2} is small compared to both the thermal pressures that we have computed at the planet surface and the hydrodynamic pressures at the sonic point. For planetary magnetic fields to confine the wind, they are required to have surface strengths in excess of ~ 30 Gauss.

4.4.6 Occurrence Rates of Close-In Progenitors

According to our analysis in §4.3.4, currently KIC 1255b has a mass of at most $0.02\text{--}0.07 M_\oplus$, and is in a final, short-lived, catastrophically evaporating phase possibly lasting another $t_{\text{evap}} \sim 40\text{--}400$ Myr during which dust is ejected at a large enough rate to produce eclipse depths of order 1%. But for most of its presumably Gyrs-long life, KIC 1255b was a more quiescent planet—larger but still less than roughly Mercury in size (§4.3.4), with a stronger gravity and emitting a much more tenuous wind. We therefore expect that for every KIC 1255b-like object discovered, there are many more progenitors in the quiescent phase—i.e., planets having sizes up to about that of Mercury orbiting main sequence K-type stars at ~ 0.01 AU.

Out of the $\sim 45,000$ main sequence K-type stars in the Kepler Input Catalogue ([Batalha et al. 2010](#)), there is apparently only 1 object like KIC 1255b. Then

$$f_{\text{observed}} \sim f_{\text{progenitor}} f_{\text{transit}} f_{\text{evap}} \sim \frac{1}{45000}, \quad (4.34)$$

where $f_{\text{progenitor}}$ is the intrinsic occurrence rate of progenitors around K stars with orbital periods shorter than a day, $f_{\text{transit}} \sim R_*/a \sim 1/4$ is the geometric probability of transit, and $f_{\text{evap}} = t_{\text{evap}}/t_{\text{life}} \sim \{40, 400\} \text{ Myr} / 5 \text{ Gyr} \sim \{0.8, 8\}\%$ is the fraction of the planet’s lifetime spent in the catastrophic mass-loss stage. Inverting equation (4.34), we estimate that $f_{\text{progenitor}} \sim \{1, 0.1\}\%$ of K stars harbor a close-in planet having less than the mass of Mercury.

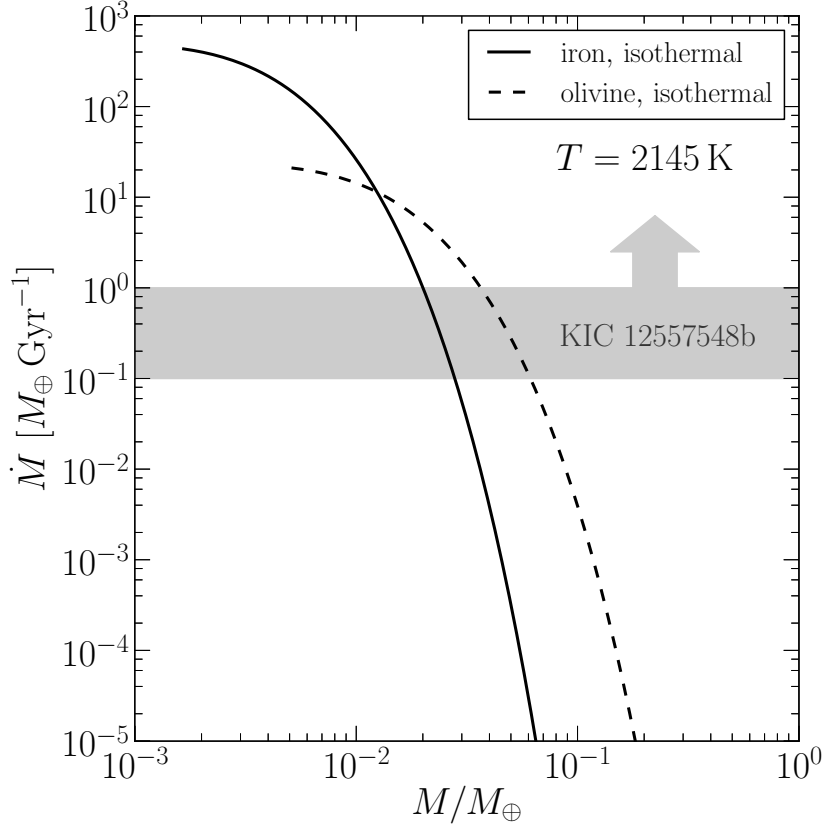


Figure 4.10 Mass loss rates $\dot{M}$, computed using the isothermal model at $T = 2145 \text{ K}$, for iron planets and olivine planets of varying M . In the low- M , free-streaming limit, the higher vapor pressure of iron allows for a higher $\dot{M}$ than for olivine. At larger M , mass loss rates are lower for iron planets than for olivine planets because of the higher molecular weight of iron. Our estimate for the maximum present-day mass of KIC 1255b varies by a factor of two between the iron and olivine scenarios.

For each KIC 1255b-like object there should be $f_{\text{evap}}^{-1} = t_{\text{life}}/t_{\text{evap}} \sim \{130, 13\}$ planets with radii no larger than about Mercury's ($\lesssim 0.4 R_\oplus$), transiting K-type stars with sub-day periods. These planets have transit depths of $(R/R_*)^2 \lesssim 10^{-5}$ —small but possibly detectable by *Kepler* if light curves are folded over enough periods (S. Rappaport 2012, personal communication) and if $\dot{M} \sim \max \dot{M}$ so that the progenitor masses attain their maximum, Mercury-like values. If $\dot{M}$ were actually $\ll \max \dot{M}$, the progenitor sizes would be less than that of Mercury. The smaller sizes would decrease their lifetimes t_{evap} and thus increase their expected number f_{evap}^{-1} , but would also render them undetectable even with *Kepler*.

4.4.7 Iron Planets

In our model, we found the (maximum) present-day mass of KIC 1255b to be about 1/3 of its (maximum) mass at formation. If KIC 1255b began its life similar in composition to the rocky planets in the Solar System, evaporation may have stripped the planet of its silicate mantle, so that only its iron core remains: KIC 1255b could be an evaporating iron planet today.

We estimate mass loss rates of a pure iron planet using our isothermal model, including tidal gravity. We set the mean molecular weight of the gas to $\mu_{\text{Fe}} \approx 56m_{\text{H}}$ and use a bulk density for the planet of 8.0 g cm^{-3} , appropriate for a pure iron planet of mass $0.01M_{\oplus}$ (Fortney et al. 2007). We fit laboratory measurements of the iron vapor pressure⁵ at $T = 2200 \text{ K}$ (Desai 1986) to equation (4.7), obtaining $e^b = 7.8 \times 10^{11} \text{ dyne cm}^{-2}$ for a latent heat of sublimation $L_{\text{sub}} = 6.3 \times 10^{10} \text{ erg g}^{-1}$ (Desai 1986) and $m = \mu_{\text{Fe}}$.

In Figure 4.10 we compare mass loss rates derived for the isothermal model with $T = 2145 \text{ K}$ for both an iron planet and a pure olivine planet. On the one hand, at this temperature, the vapor pressure for iron $P_{\text{vapor, Fe}} = 1.8 \times 10^3 \text{ dyne cm}^{-2}$ is about 50 times higher than for olivine, which raises the mass loss rate by a similar factor in the free-streaming limit appropriate for low masses (see equation 4.28). On the other hand, at higher masses, $\dot{M}$ for iron drops below that for olivine because the iron atmosphere, with its higher molecular weight, is harder to blow off. These two effects counteract each other so that an iron planet and a silicate planet both reach $\dot{M}_{1255b} \sim 1 M_{\oplus} / \text{Gyr}$ at a similar M . Thus at least within the context of isothermal winds, our estimate for the present-day mass of KIC 1255b is insensitive to whether the evaporating surface of the planet is composed of iron or silicates. Figure 4.11 displays mass loss histories for an iron planet using our isothermal model at $T = 2145 \text{ K}$. Initial planet masses are up to a factor of two lower than those of our full model using silicates.

According to Figure 4.11, iron planets with $M \gtrsim 0.05M_{\oplus}$ survive for tens of Gyrs. Compare this result to its counterpart in Figure 4.9 (left panel), which shows that olivine planets with $M \lesssim 0.1M_{\oplus}$ evaporate within $\sim 10 \text{ Gyr}$. This comparison suggests that for planets with the right proportion of silicates in the mantle to iron in the core, mantles may be completely vaporized, leaving behind essentially non-evaporating iron cores. Such massive, quiescent iron cores might be detectable by *Kepler* via direct transits (see §4.4.6).

4.4.8 Dust Condensation

Recall that we have not modelled the microphysics of dust formation, but treated the dust-to-gas ratio as a free function. Are the conditions of our flow actually favorable for the condensation of dust? Any gas whose partial pressure is greater than its vapor pressure (at that T) might condense and form droplets (modulo the many complications discussed below). Condensation might proceed until all vapor in excess of saturation is in cloud particles. If

⁵We have found measurements of the iron vapor pressure at $T \sim 2000 \text{ K}$ to differ by a factor of two in the literature (see, e.g., Nuth et al. 2003).

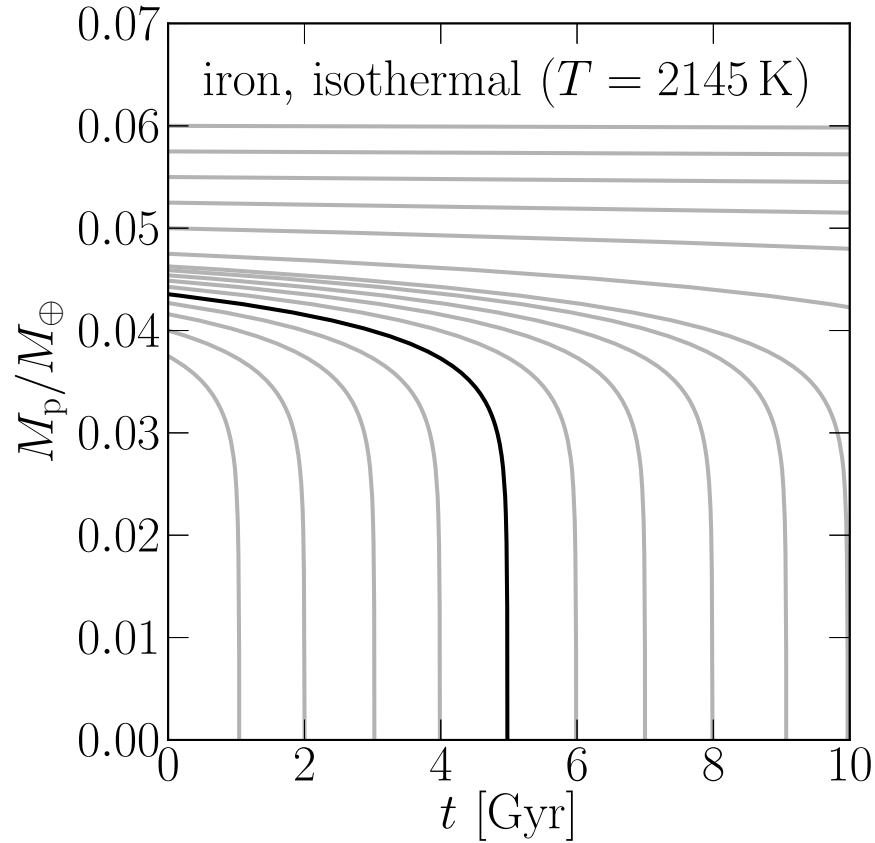


Figure 4.11 Mass-loss histories $M(t)$ for an iron planet, obtained by time-integrating $f_{\text{duty}} \cdot \max \dot{M}$ with $f_{\text{duty}} = 0.5$ for an isothermal wind. An iron planet with a 5 Gyr lifetime will have an initial mass of $\sim 0.044 M_{\oplus}$. As was the case for olivine planets (see Figure 4.9), the catastrophic evaporation stage lasts only for $\sim 100 \text{ Myr}$. Iron planets (or planetary iron cores) with masses greater than $0.05 M_{\oplus}$ survive for over 10 Gyr.

the condensates have sedimentation velocities larger than updraft speeds they will rain out of the atmosphere; otherwise they remain aloft (see, e.g., [Lewis 1969](#); [Marley et al. 1999](#); [Ackerman & Marley 2001](#); [Helling et al. 2001](#)).

In Figure 4.12, we compare gas pressures of our solutions (for $\dot{M} = \max \dot{M}$) at $M = 0.01$, 0.03 , and $0.07 M_{\oplus}$ with the equilibrium vapor pressures of olivine and pyroxene. At the base of the flow, the gas pressure equals the saturation vapor pressure of olivine by construction. As gas expands and cools, its pressure generally remains above the saturation pressure of pyroxene, so conditions are favorable for the condensation of pyroxene grains. Unfortunately we have an embarrassment of riches: at $T \lesssim 1700$ K, our gas pressures are many orders of magnitude higher than silicate vapor pressures, and the concern is that gas condensation will lower pressures precipitously to the point where winds shut down. This fate could be avoided if grains take too long to condense out of the outflowing and rapidly rarefying gas—i.e., the finite timescale of grain condensation might permit gas to remain supersaturated. Furthermore, even before bulk condensation can begin, seed particles must first nucleate from the gas phase. These issues call for a time-dependent analysis of grain formation, perhaps along the lines made for brown dwarf and exoplanet atmospheres ([Helling et al. 2001, 2008b,a](#); [Helling & Rietmeijer 2009](#)). Accounting for the extra heating from small grains with sizes $\ll 1 \mu\text{m}$ as they first condense out of the wind (§4.4.2) would keep the gas closer to isothermal and help to prevent catastrophic condensation.

Of course, by the wind leaves the sonic point and is diverted into the trailing comet-like tail, it must be full of grains to explain the observed occultations of KIC 1255b. In this Chapter, we have not modeled the dusty tail at all. Future work should address the dynamics of the tail and compute its extinction profile, with the aim of reproducing the observed transit lightcurves.

4.5 Conclusions

Our work supports the hypothesis that the observed occultations of the K-star KIC 12557548 originate from a dusty wind emitted by an evaporating planetary companion (“KIC 1255b”). We reach the following conclusions.

1. *Maximum present-day mass*

Our estimates for the maximum present-day mass of KIC 1255b range from $\max M \approx 0.02 M_{\oplus}$ (roughly twice the mass of the Moon) to $\max M \approx 0.07 M_{\oplus}$ (slightly larger than the mass of Mercury), depending on how strongly the wind is heated by dust and can remain isothermal (Figure 4.8). For these maximum masses, calculated mass loss rates peak at $\max \dot{M} \sim 0.1 M_{\oplus}/\text{Gyr}$, which is just large enough to produce the observed transit depths of order 1%. Smaller planets with weaker gravities yield larger mass loss rates and are also compatible with the observations.

2. *Maximum formation mass*

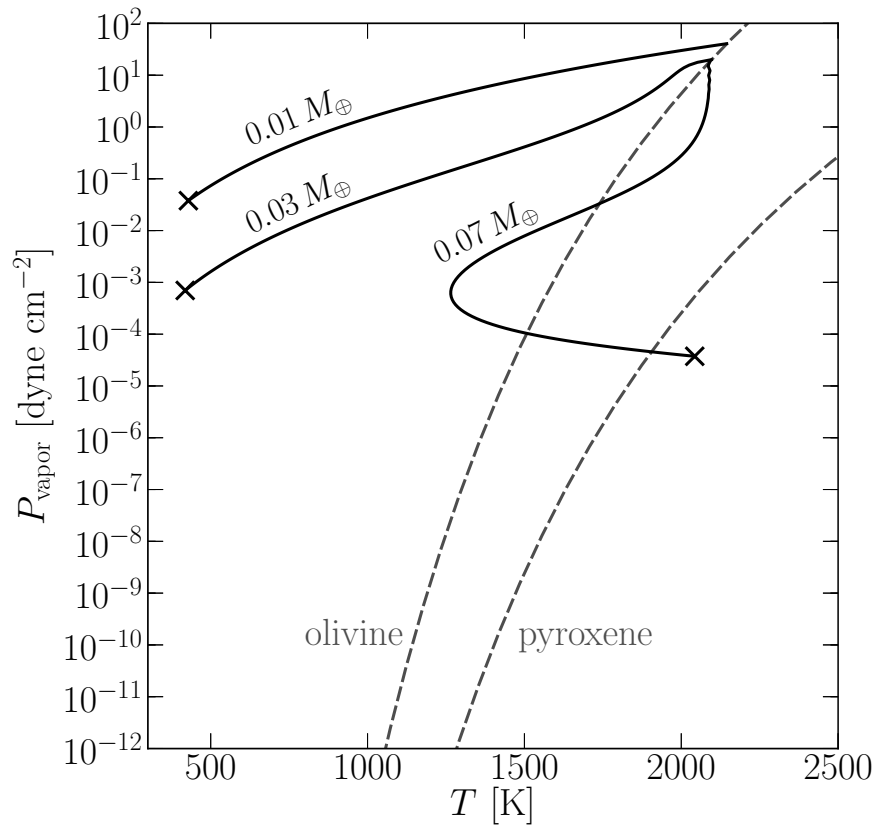


Figure 4.12 Trajectories of the wind in pressure-temperature space, computed using the full model for $M = 0.01, 0.03$, and $0.07M_{\oplus}$. Gas pressures generally remain above the equilibrium vapor pressures of pyroxene, permitting the condensation of pyroxene grains.

For an assumed planet age of ~ 5 Gyr, the maximum mass of KIC 1255b at formation ranges from $0.06 M_{\oplus}$ to $0.1 M_{\oplus}$, again depending on the energetics of the wind (Figure 4.9).

3. *Mass threshold for catastrophic evaporation*

A pure rock planet of mass $\gtrsim 0.1 M_{\oplus}$ and surface temperature $\lesssim 2200$ K will survive with negligible mass loss for tens of Gyrs.

4. *Time variability*

The observed occultations of KIC 12557548 vary by up to an order of magnitude in depth without any apparent correlation between orbits. The implied order-of-magnitude variations in $\dot{M}$ can be explained in principle by the exponential sensitivity of $\dot{M}$ to conditions at the planet surface. The dynamical “refresh” timescale of the wind $t_{\text{dyn}} \sim 14$ h is similar to the orbital period $P_{\text{orb}} = 15.7$ h. This supports our dusty wind model because were $t_{\text{dyn}} \gg P_{\text{orb}}$ or $t_{\text{dyn}} \ll P_{\text{orb}}$, then eclipse depths would correlate from orbit to orbit, in violation of the observations.

5. *Progenitors and occurrence rates*

KIC 1255b’s current catastrophic mass loss phase may represent only the final few percent of the planet’s life. As such, for every KIC 1255b-like object, there could be anywhere from 10 to 100 larger planets in earlier stages of mass loss. These close-in, relatively quiescent progenitors may be detectable by *Kepler* through conventional “hard-sphere” transits if they are as large as Mercury. We cannot, however, rule out the possibility that the progenitors are lunar-sized or even smaller. If KIC 1255b remains the only catastrophically evaporating planet in the *Kepler* database, then the occurrence rate of close-in progenitors orbiting K-stars with sub-day periods is $> 0.1\%$, with larger occurrence rates for progenitors increasingly smaller than Mercury.

6. *Iron planet*

KIC 1255b may have lost $\sim 70\%$ of its formation mass to its thermal wind. It seems possible that today only the iron core of KIC 1255b remains and is evaporating. Transmission spectra of the occulting dust cloud might reveal whether the planet’s surface is composed primarily of iron or silicates.

7. *Future modelling and observations*

Keeping the wind hot as it lifts off the planet surface significantly enhances mass loss rates. In our model we have included heating of the gas by micron-sized grains embedded in the flow. But these grains can also shadow the surface from starlight and reduce $\dot{M}$ if they are too abundant. Additional heating may be provided by super-blackbody grains with sizes $\ll 1 \mu\text{m}$ which do not significantly attenuate starlight. Future models should incorporate such tiny condensates—and treat the dusty comet-like tail that our work has completely ignored. More information about grain size distributions and

compositions might be revealed by *Hubble* observations of KIC 12557548 scheduled for early 2013.

Bibliography

- Ackerman, A. S., & Marley, M. S. 2001, [Astrophys. J.](#), 556, 872
- Aikawa, Y., Miyama, S. M., Nakano, T., & Umebayashi, T. 1996, [Astrophys. J.](#), 467, 684
- Aldrovandi, S. M. V., & Pequignot, D. 1973, *Astron. Astrophys.*, 25, 137
- Alexander, R. D., Clarke, C. J., & Pringle, J. E. 2005, [Mon. Not. Roy. Astron. Soc.](#), 358, 283
- Allamandola, L. J., Hudgins, D. M., & Sandford, S. A. 1999, [ApJ](#), 511, L115
- Allamandola, L. J., Tielens, A. G. G. M., & Barker, J. R. 1989, [ApJS](#), 71, 733
- Anderson, B. J., Acuña, M. H., Korth, H., et al. 2008, [Science](#), 321, 82
- Andrews, D., Holton, J., & Leovy, C. 1987, *Middle Atmosphere Dynamics* (Academic Press, New York)
- Andrews, S. M., & Williams, J. P. 2005, [Astrophys. J.](#), 631, 1134
- Artymowicz, P., & Lubow, S. H. 1994, [Astrophys. J.](#), 421, 651
- . 1996, [ApJ](#), 467, L77
- Bai, X.-N. 2011, [Astrophys. J.](#), 739, 51
- Bai, X.-N., & Goodman, J. 2009, [Astrophys. J.](#), 701, 737
- Bai, X.-N., & Stone, J. M. 2011, [Astrophys. J.](#), 736, 144
- Balbus, S. A. 2009, ArXiv e-prints, [arXiv:0906.0854 \[astro-ph.SR\]](#)
- Balbus, S. A., & Terquem, C. 2001, [Astrophys. J.](#), 552, 235
- Batalha, N. M., Borucki, W. J., Koch, D. G., et al. 2010, [ApJ](#), 713, L109
- Ben-Jaffel, L. 2007, [ApJ](#), 671, L61
- . 2008, [Astrophys. J.](#), 688, 1352
- Bergin, E., Calvet, N., D'Alessio, P., & Herczeg, G. J. 2003, [ApJ](#), 591, L159
- Bergin, E. A., Aikawa, Y., Blake, G. A., & van Dishoeck, E. F. 2007, *Protostars and Planets V*, 751
- Bida, T. A., Killen, R. M., & Morgan, T. H. 2000, *Nature*, 404, 159
- Blaes, O. M., & Balbus, S. A. 1994, [Astrophys. J.](#), 421, 163
- Blum, J., & Wurm, G. 2008, [ARA&A](#), 46, 21
- Borucki, W. J., & Whitten, R. C. 2008, [Planet. Space Sci.](#), 56, 19
- Borucki, W. J., Koch, D., Jenkins, J., et al. 2009, [Science](#), 325, 709
- Brandenburg, A., & Zweibel, E. G. 1994, [ApJ](#), 427, L91
- Bretherton, C. S., & Smolarkiewicz, P. K. 1989, *Journal of Atmospheric Sciences*, 46, 740
- Bretherton, C. S., & Sobel, A. H. 2003, *Journal of Atmospheric Sciences*, 60, 451
- Brogi, M., Keller, C. U., de Juan Ovelar, M., et al. 2012, *Astron. Astrophys.*, 545, L5

- Brown, J. M., Blake, G. A., Qi, C., et al. 2009, [Astrophys. J.](#), 704, 496
- Browning, M. K., Tumlinson, J., & Shull, J. M. 2003, [Astrophys. J.](#), 582, 810
- Burns, J. A., Lamy, P. L., & Soter, S. 1979, [Icarus](#), 40, 1
- Calvet, N., D'Alessio, P., Hartmann, L., et al. 2002, [Astrophys. J.](#), 568, 1008
- Calvet, N., & Gullbring, E. 1998, [Astrophys. J.](#), 509, 802
- Calvet, N., Muzerolle, J., Briceño, C., et al. 2004, [AJ](#), 128, 1294
- Calvet, N., D'Alessio, P., Watson, D. M., et al. 2005, [ApJ](#), 630, L185
- Caselli, P., Walmsley, C. M., Terzieva, R., & Herbst, E. 1998, [Astrophys. J.](#), 499, 234
- Catling, D. C., & Zahnle, K. J. 2009, [Scientific American](#), 300, 050000
- Cazaux, S., & Tielens, A. G. G. M. 2010, [Astrophys. J.](#), 715, 698
- Chiang, E., & Laughlin, G. 2013, [Mon. Not. Roy. Astron. Soc.](#), 431, 3444
- Chiang, E., & Murray-Clay, R. 2007, [Nature Physics](#), 3, 604
- Chiang, E., & Youdin, A. N. 2010a, [Annual Review of Earth and Planetary Sciences](#), 38, 493
- . 2010b, [Annual Review of Earth and Planetary Sciences](#), 38, 493
- Chiang, E. I., & Goldreich, P. 1997, [Astrophys. J.](#), 490, 368
- Chiang, E. I., Joun, M. K., Creech-Eakman, M. J., et al. 2001, [Astrophys. J.](#), 547, 1077
- Cho, J. Y.-K., Menou, K., Hansen, B. M. S., & Seager, S. 2003, [ApJ](#), 587, L117
- . 2008, [Astrophys. J.](#), 675, 817
- Cho, J. Y.-K., & Polvani, L. M. 1996, [Physics of Fluids](#), 8, 1531
- Cooper, C. S., & Showman, A. P. 2005, [ApJ](#), 629, L45
- Cowan, N. B., & Agol, E. 2011, [Astrophys. J.](#), 729, 54
- Cowan, N. B., Agol, E., & Charbonneau, D. 2007, [Mon. Not. Roy. Astron. Soc.](#), 379, 641
- Cowan, N. B., Machalek, P., Croll, B., et al. 2012, [Astrophys. J.](#), 747, 82
- Crida, A., & Morbidelli, A. 2007, [Mon. Not. Roy. Astron. Soc.](#), 377, 1324
- Crossfield, I. J. M., Hansen, B. M. S., Harrington, J., et al. 2010, [Astrophys. J.](#), 723, 1436
- Crudace, R., Paresce, F., Bowyer, S., & Lampton, M. 1974, [Astrophys. J.](#), 187, 497
- D'Alessio, P., Calvet, N., Hartmann, L., Franco-Hernández, R., & Servín, H. 2006, [Astrophys. J.](#), 638, 314
- D'Alessio, P., Calvet, N., Hartmann, L., Lizano, S., & Cantó, J. 1999, [Astrophys. J.](#), 527, 893
- D'Alessio, P., Hartmann, L., Calvet, N., et al. 2005, [Astrophys. J.](#), 621, 461
- Dalgarno, A., & Stephens, T. L. 1970, [ApJ](#), 160, L107
- de Jong, T., Boland, W., & Dalgarno, A. 1980, [Astron. Astrophys.](#), 91, 68
- Deming, D., & Seager, S. 2009, [Nature](#), 462, 301
- Desai, P. D. 1986, [Journal of Physical and Chemical Reference Data](#), 15, 967
- Desch, S. J. 2004, [Astrophys. J.](#), 608, 509
- Dowling, T. E., & Ingersoll, A. P. 1989, [Journal of Atmospheric Sciences](#), 46, 3256
- Draine, B. T., Roberge, W. G., & Dalgarno, A. 1983, [Astrophys. J.](#), 264, 485
- Draine, B. T., & Sutin, B. 1987, [Astrophys. J.](#), 320, 803
- Dullemond, C. P., & Dominik, C. 2004, [Astron. Astrophys.](#), 421, 1075
- Dullemond, C. P., Dominik, C., & Natta, A. 2001, [Astrophys. J.](#), 560, 957
- Dullemond, C. P., Henning, T., Visser, R., et al. 2007, [Astron. Astrophys.](#), 473, 457
- Eisner, J. A., Doppmann, G. W., Najita, J. R., et al. 2010, [ApJ](#), 722, L28

- Eisner, J. A., Monnier, J. D., Tuthill, P., & Lacour, S. 2009, [ApJ](#), 698, L169
- Ekenbäck, A., Holmström, M., Wurz, P., et al. 2010, [Astrophys. J.](#), 709, 670
- Espaillet, C., Calvet, N., D'Alessio, P., et al. 2007a, [ApJ](#), 670, L135
- . 2007b, [ApJ](#), 664, L111
- Espaillet, C., Muzerolle, J., Hernández, J., et al. 2008, [ApJ](#), 689, L145
- Espaillet, C., D'Alessio, P., Hernández, J., et al. 2010, [Astrophys. J.](#), 717, 441
- Fleming, T. P., Stone, J. M., & Hawley, J. F. 2000, [Astrophys. J.](#), 530, 464
- Fortney, J. J., Lodders, K., Marley, M. S., & Freedman, R. S. 2008, [Astrophys. J.](#), 678, 1419
- Fortney, J. J., Marley, M. S., & Barnes, J. W. 2007, [Astrophys. J.](#), 659, 1661
- Fossati, L., Haswell, C. A., Froning, C. S., et al. 2010, [ApJ](#), 714, L222
- Fromang, S., Terquem, C., & Balbus, S. A. 2002, [Mon. Not. Roy. Astron. Soc.](#), 329, 18
- Gammie, C. F. 1996, [Astrophys. J.](#), 457, 355
- García Muñoz, A. 2007, [Planet. Space Sci.](#), 55, 1426
- Geers, V. C., van Dishoeck, E. F., Visser, R., et al. 2007, [Astron. Astrophys.](#), 476, 279
- Geers, V. C., Augereau, J.-C., Pontoppidan, K. M., et al. 2006, [Astron. Astrophys.](#), 459, 545
- Geissler, P. E. 2003, [Annual Review of Earth and Planetary Sciences](#), 31, 175
- Getman, K. V., Feigelson, E. D., Broos, P. S., Micela, G., & Garmire, G. P. 2008a, [Astrophys. J.](#), 688, 418
- Getman, K. V., Feigelson, E. D., Micela, G., et al. 2008b, [Astrophys. J.](#), 688, 437
- Gill, A. E. 1980, [Quarterly Journal of the Royal Meteorological Society](#), 106, 447
- Gill, A. E. 1982, *Atmosphere-Ocean Dynamics* (Academic Press, New York)
- Glassgold, A. E., Najita, J., & Igea, J. 1997, [Astrophys. J.](#), 480, 344
- . 2004, [Astrophys. J.](#), 615, 972
- Glassgold, A. E., Najita, J. R., & Igea, J. 2007, [Astrophys. J.](#), 656, 515
- Goldreich, P., & Tremaine, S. 1980, [Astrophys. J.](#), 241, 425
- Gorti, U., Dullemond, C. P., & Hollenbach, D. 2009, [Astrophys. J.](#), 705, 1237
- Gorti, U., & Hollenbach, D. 2004, [Astrophys. J.](#), 613, 424
- . 2008, [Astrophys. J.](#), 683, 287
- Goto, M., Henning, T., Kouchi, A., et al. 2009, [Astrophys. J.](#), 693, 610
- Gould, R. J. 1978, [Astrophys. J.](#), 219, 250
- Gould, R. J., & Salpeter, E. E. 1963, [Astrophys. J.](#), 138, 393
- Guillot, T., Burrows, A., Hubbard, W. B., Lunine, J. I., & Saumon, D. 1996, [ApJ](#), 459, L35
- Habart, E., Natta, A., & Krügel, E. 2004, [Astron. Astrophys.](#), 427, 179
- Hack, J. J., & Jakob, R. 1992, Tech. rep., National Center for Atmospheric Research Technical note NCAR/TN-343+STR, Boulder, CO
- Hansen, B. M. S., & Murray, N. 2012, [Astrophys. J.](#), 751, 158
- Harrington, J., Hansen, B. M., Luszcz, S. H., et al. 2006, [Science](#), 314, 623
- Hartmann, L., D'Alessio, P., Calvet, N., & Muzerolle, J. 2006, [Astrophys. J.](#), 648, 484
- Hartmann, L., Hewett, R., & Calvet, N. 1994, [Astrophys. J.](#), 426, 669
- Hashimoto, A. 1990, [Nature](#), 347, 53
- Hawley, J. F., & Stone, J. M. 1998, [Astrophys. J.](#), 501, 758
- Helling, C., Oevermann, M., Lüttke, M. J. H., Klein, R., & Sedlmayr, E. 2001, [Astron.](#)

[Astrophys.](#), 376, 194

- Helling, C., & Rietmeijer, F. J. M. 2009, [International Journal of Astrobiology](#), 8, 3
- Helling, C., Woitke, P., & Thi, W.-F. 2008a, [Astron. Astrophys.](#), 485, 547
- Helling, C., Ackerman, A., Allard, F., et al. 2008b, [Mon. Not. Roy. Astron. Soc.](#), 391, 1854
- Heng, K., Menou, K., & Phillipps, P. J. 2011, [Mon. Not. Roy. Astron. Soc.](#), 413, 2380
- Herczeg, G. J., & Hillenbrand, L. A. 2008, [Astrophys. J.](#), 681, 594
- Holmström, M., Ekenbäck, A., Selsis, F., et al. 2008, [Nature](#), 451, 970
- Holton, J. 2004, *An Introduction to Dynamic Meteorology*, 4th Ed. (Academic Press, San Diego)
- Hughes, A. M., Wilner, D. J., Andrews, S. M., Qi, C., & Hogerheijde, M. R. 2011, [Astrophys. J.](#), 727, 85
- Hughes, A. M., Wilner, D. J., Calvet, N., et al. 2007, [Astrophys. J.](#), 664, 536
- Igea, J., & Glassgold, A. E. 1999, [Astrophys. J.](#), 518, 848
- Ilgner, M., & Nelson, R. P. 2006a, [Astron. Astrophys.](#), 445, 205
- . 2006b, [Astron. Astrophys.](#), 445, 223
- Inutsuka, S.-i., & Sano, T. 2005, [ApJ](#), 628, L155
- Ireland, M. J., & Kraus, A. L. 2008, [ApJ](#), 678, L59
- Jackson, J. D. 1975, *Classical electrodynamics*
- Jakosky, B. M., & Phillips, R. J. 2001, [Nature](#), 412, 237
- Jenkins, E. B. 2009, [Astrophys. J.](#), 700, 1299
- Jensen, E. J., & Thomas, G. E. 1991, [J. Geophys. Res.](#), 96, 18603
- Johns-Krull, C. M., Valenti, J. A., & Linsky, J. L. 2000, [Astrophys. J.](#), 539, 815
- Johnson, R. S. 1997, *A Modern Introduction to the Mathematical Theory of Water Waves* (Cambridge University Press, Cambridge, UK)
- Kasting, J. F., & Pollack, J. B. 1983, [Icarus](#), 53, 479
- Kaufman, M. J., Wolfire, M. G., Hollenbach, D. J., & Luhman, M. L. 1999, [Astrophys. J.](#), 527, 795
- Keller, L. D., Sloan, G. C., Forrest, W. J., et al. 2008, [Astrophys. J.](#), 684, 411
- Killen, R. M., Sarantos, M., Potter, A. E., & Reiff, P. 2004, [Icarus](#), 171, 1
- Kim, K. H., Watson, D. M., Manoj, P., et al. 2009, [Astrophys. J.](#), 700, 1017
- Kim, R.-S., Gopalswamy, N., Moon, Y.-J., Cho, K.-S., & Yashiro, S. 2012, [Astrophys. J.](#), 746, 118
- Kimura, H., Mann, I., Biesecker, D. A., & Jessberger, E. K. 2002, [Icarus](#), 159, 529
- Knutson, H. 2011, AGU Fall Meeting Abstracts, C4
- Knutson, H. A., Charbonneau, D., Cowan, N. B., et al. 2009a, [Astrophys. J.](#), 703, 769
- Knutson, H. A., Charbonneau, D., Allen, L. E., et al. 2007, [Nature](#), 447, 183
- Knutson, H. A., Charbonneau, D., Cowan, N. B., et al. 2009b, [Astrophys. J.](#), 690, 822
- Knutson, H. A., Lewis, N., Fortney, J. J., et al. 2012, [Astrophys. J.](#), 754, 22
- Kohl, J. L., Noci, G., Antonucci, E., et al. 1998, [ApJ](#), 501, L127
- Kunz, M. W., & Balbus, S. A. 2004, [Mon. Not. Roy. Astron. Soc.](#), 348, 355
- Lahuis, F., van Dishoeck, E. F., Blake, G. A., et al. 2007, [Astrophys. J.](#), 665, 492
- Lai, D., Helling, C., & van den Heuvel, E. P. J. 2010, [Astrophys. J.](#), 721, 923

- Lamers, H. J. G. L. M., & Cassinelli, J. P. 1999, Introduction to Stellar Winds
- Lamers, H. J. G. L. M., Lecavelier Des Etangs, A., & Vidal-Madjar, A. 1997, *Astron. Astrophys.*, 328, 321
- Lammer, H., & Bauer, S. J. 1997, *Planet. Space Sci.*, 45, 73
- Lange, N. A., & Forker, G. M. 1967, Handbook of Chemistry
- Laughlin, G., Deming, D., Langton, J., et al. 2009, *Nature*, 457, 562
- Lecavelier Des Etangs, A., Vidal-Madjar, A., Burki, G., et al. 1997, *Astron. Astrophys.*, 328, 311
- Lecavelier Des Etangs, A., Ehrenreich, D., Vidal-Madjar, A., et al. 2010, *Astron. Astrophys.*, 514, A72
- Lecavelier des Etangs, A., Bourrier, V., Wheatley, P. J., et al. 2012, *Astron. Astrophys.*, 543, L4
- Lewis, J. S. 1969, *Icarus*, 10, 365
- Lewis, N. K., Knutson, H. A., Showman, A. P., et al. 2013, *Astrophys. J.*, 766, 95
- Li, A., & Draine, B. T. 2001, *Astrophys. J.*, 554, 778
- Li, J., & Goodman, J. 2010, *Astrophys. J.*, 725, 1146
- Linsky, J. L., Yang, H., France, K., et al. 2010, *Astrophys. J.*, 717, 1291
- Liu, B., & Showman, A. P. 2013, *Astrophys. J.*, 770, 42
- Llama, J., Wood, K., Jardine, M., et al. 2011, *Mon. Not. Roy. Astron. Soc.*, 416, L41
- Lodders, K. 2003, *Astrophys. J.*, 591, 1220
- Lopez, E. D., Fortney, J. J., & Miller, N. 2012, *Astrophys. J.*, 761, 59
- Lovis, C., & Fischer, D. 2010, Exoplanets (ed. S. Seager, University of Arizona Press, Tucson)
- Lubow, S. H., & D'Angelo, G. 2006, *Astrophys. J.*, 641, 526
- Lubow, S. H., Seibert, M., & Artymowicz, P. 1999, *Astrophys. J.*, 526, 1001
- Mac Low, M.-M., Norman, M. L., Konigl, A., & Wardle, M. 1995, *Astrophys. J.*, 442, 726
- Marley, M. S., Gelino, C., Stephens, D., Lunine, J. I., & Freedman, R. 1999, *Astrophys. J.*, 513, 879
- Marsh, N., & Svensmark, H. 2003, *Journal of Geophysical Research (Atmospheres)*, 108, 4195
- Martin-Zaïdi, C., Deleuil, M., Le Bourlot, J., et al. 2008, *Astron. Astrophys.*, 484, 225
- Masset, F., & Snellgrove, M. 2001, *Mon. Not. Roy. Astron. Soc.*, 320, L55
- Matsuno, T. 1966, *J. Meteorol. Soc. Japan*, 44, 25
- Maxted, P. F. L., Anderson, D. R., Doyle, A. P., et al. 2013, *Mon. Not. Roy. Astron. Soc.*, 428, 2645
- McCabe, C., Duchêne, G., & Ghez, A. M. 2003, *ApJ*, 588, L113
- McComas, D. J., Elliott, H. A., Schwadron, N. A., et al. 2003, *Geophys. Res. Lett.*, 30, 1517
- Meijerink, R., Glassgold, A. E., & Najita, J. R. 2008, *Astrophys. J.*, 676, 518
- Melosh, H. J., & Vickery, A. M. 1989, *Nature*, 338, 487
- Menou, K., Cho, J. Y.-K., Seager, S., & Hansen, B. M. S. 2003, *ApJ*, 587, L113
- Merín, B., Brown, J. M., Oliveira, I., et al. 2010, *Astrophys. J.*, 718, 1200
- Millar, T. J., Farquhar, P. R. A., & Willacy, K. 1997, *A&AS*, 121, 139
- Mohanty, S., Ercolano, B., & Turner, N. J. 2013, *Astrophys. J.*, 764, 65
- Morbidelli, A., & Crida, A. 2007, *Icarus*, 191, 158

- Mura, A., Wurz, P., Schneider, J., et al. 2011, [Icarus](#), **211**, 1
- Murray, C. D., & Dermott, S. F. 1999, Solar system dynamics
- Murray-Clay, R. A., Chiang, E. I., & Murray, N. 2009, [Astrophys. J.](#), **693**, 23
- Muzerolle, J., Luhman, K. L., Briceño, C., Hartmann, L., & Calvet, N. 2005, [Astrophys. J.](#), **625**, 906
- Nagahara, H., Mysen, B. O., & Kushiro, I. 1994, [Geochim. Cosmochim. Acta](#), **58**, 1951
- Najita, J. R., Strom, S. E., & Muzerolle, J. 2007, [Mon. Not. Roy. Astron. Soc.](#), **378**, 369
- Natanson, G. L. 1960, *Sov. Phys. Tech. Phys.*, **5**, 538
- Natta, A., Testi, L., Calvet, N., et al. 2007, Protostars and Planets V, 767
- Ness, N. F., Behannon, K. W., Lepping, R. P., Whang, Y. C., & Schatten, K. H. 1974, [Science](#), **185**, 151
- Nuth, III, J. A., Ferguson, F. T., Johnson, N., & Martinez, D. 2003, in Lunar and Planetary Institute Science Conference Abstracts, Vol. 34, Lunar and Planetary Institute Science Conference Abstracts, ed. S. Mackwell & E. Stansbery, 1598
- Ochi, Y., Sugimoto, K., & Hanawa, T. 2005, [Astrophys. J.](#), **623**, 922
- Oppenheimer, M., & Dalgarno, A. 1974, [Astrophys. J.](#), **192**, 29
- Paardekooper, S.-J., & Mellema, G. 2006, [Astron. Astrophys.](#), **453**, 1129
- Parthasarathy, R. 1976, [J. Geophys. Res.](#), **81**, 2392
- Pendleton, Y. J., & Allamandola, L. J. 2002, [ApJS](#), **138**, 75
- Pequignot, D., Petitjean, P., & Boisson, C. 1991, *Astron. Astrophys.*, **251**, 680
- Perez-Becker, D., & Chiang, E. 2011, [Astrophys. J.](#), **727**, 2
- Perna, R., Heng, K., & Pont, F. 2012, [Astrophys. J.](#), **751**, 59
- Perna, R., Menou, K., & Rauscher, E. 2010, [Astrophys. J.](#), **719**, 1421
- Pessah, M. E., Chan, C.-k., & Psaltis, D. 2007, [ApJ](#), **668**, L51
- Pinte, C., Woitke, P., Ménard, F., et al. 2010, [Astron. Astrophys.](#), **518**, L126
- Potter, A. E., & Morgan, T. H. 1990, [Science](#), **248**, 835
- . 1997, [Planet. Space Sci.](#), **45**, 95
- Preibisch, T., Kim, Y.-C., Favata, F., et al. 2005, [ApJS](#), **160**, 401
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 1992, Numerical recipes in C. The art of scientific computing
- Quémerais, E., Lallement, R., Koutroumpa, D., & Lamy, P. 2007, [Astrophys. J.](#), **667**, 1229
- Rappaport, S., Levine, A., Chiang, E., et al. 2012, [Astrophys. J.](#), **752**, 1
- Ratzka, T., Leinert, C., Henning, T., et al. 2007, [Astron. Astrophys.](#), **471**, 173
- Rauscher, E., & Menou, K. 2010, [Astrophys. J.](#), **714**, 1334
- . 2013, [Astrophys. J.](#), **764**, 103
- Reedy, R. C. 1987, [J. Geophys. Res.](#), **92**, 697
- Rice, W. K. M., Armitage, P. J., Wood, K., & Lodato, G. 2006, [Mon. Not. Roy. Astron. Soc.](#), **373**, 1619
- Robertson, S., & Sternovsky, Z. 2008, [Physics of Plasmas](#), **15**, 040702
- Salyk, C., Blake, G. A., Boogert, A. C. A., & Brown, J. M. 2007, [ApJ](#), **655**, L105
- Sano, T., & Inutsuka, S.-i. 2001, [ApJ](#), **561**, L179
- Sano, T., Miyama, S. M., Umebayashi, T., & Nakano, T. 2000, [Astrophys. J.](#), **543**, 486

- Sano, T., & Stone, J. M. 2002, [Astrophys. J.](#), 570, 314
- Savage, B. D., & Sembach, K. R. 1996, [ARA&A](#), 34, 279
- Schaefer, L., & Fegley, B. 2009, [ApJ](#), 703, L113
- Schunk, R. W., & Nagy, A. F. 2004, Ionospheres
- Scott, R. K., & Polvani, L. M. 2007, [Journal of Atmospheric Sciences](#), 64, 3158
- Semenov, D., Wiebe, D., & Henning, T. 2004, [Astron. Astrophys.](#), 417, 93
- Sheeley, Jr., N. R., Wang, Y.-M., Hawley, S. H., et al. 1997, [Astrophys. J.](#), 484, 472
- Shell, K. M., & Held, I. M. 2004, [Journal of Atmospheric Sciences](#), 61, 2928
- Showman, A. P. 2007, [Journal of Atmospheric Sciences](#), 64, 3132
- Showman, A. P., Cho, J. Y.-K., & Menou, K. 2010, Exoplanets (ed. S. Seager, University of Arizona Press, Tucson)
- Showman, A. P., Cooper, C. S., Fortney, J. J., & Marley, M. S. 2008, [Astrophys. J.](#), 682, 559
- Showman, A. P., Fortney, J. J., Lewis, N. K., & Shabram, M. 2013a, [Astrophys. J.](#), 762, 24
- Showman, A. P., & Guillot, T. 2002, [Astron. Astrophys.](#), 385, 166
- Showman, A. P., & Polvani, L. M. 2010, [Geophys. Res. Lett.](#), 37, 18811
- . 2011, [Astrophys. J.](#), 738, 71
- Showman, A. P., Wordsworth, R. D., Merlis, T. M., & Kaspi, Y. 2013b, ArXiv e-prints, [arXiv:1306.2418 \[astro-ph.EP\]](#)
- Sloan, G. C., Keller, L. D., Forrest, W. J., et al. 2005, [Astrophys. J.](#), 632, 956
- Sobel, A. H. 2002, [Chaos](#), 12, 451
- Sobel, A. H., Nilsson, J., & Polvani, L. M. 2001, [Journal of Atmospheric Sciences](#), 58, 3650
- Stelzer, B., & Schmitt, J. H. M. M. 2004, [Astron. Astrophys.](#), 418, 687
- Strom, S. E., Edwards, S., & Skrutskie, M. F. 1990, in Astronomical Society of the Pacific Conference Series, Vol. 9, Cool Stars, Stellar Systems, and the Sun, ed. G. Wallerstein, 275
- Strom, S. E., Edwards, S., & Skrutskie, M. F. 1993, in Protostars and Planets III, ed. E. H. Levy & J. I. Lunine, 837
- Sturm, B., Bouwman, J., Henning, T., et al. 2010, [Astron. Astrophys.](#), 518, L129
- Svensmark, H. 1998, [Physical Review Letters](#), 81, 5027
- Swift, J. J., Johnson, J. A., Morton, T. D., et al. 2013, [Astrophys. J.](#), 764, 105
- Tachibana, S., Tsuchiyama, A., & Nagahara, H. 2002, [Geochim. Cosmochim. Acta](#), 66, 713
- Telleschi, A., Güdel, M., Briggs, K. R., Audard, M., & Palla, F. 2007, [Astron. Astrophys.](#), 468, 425
- Tian, F., Toon, O. B., Pavlov, A. A., & De Sterck, H. 2005, [Astrophys. J.](#), 621, 1049
- Tielens, A. G. G. M., & Hollenbach, D. 1985a, [Astrophys. J.](#), 291, 747
- . 1985b, [Astrophys. J.](#), 291, 722
- Tremblin, P., & Chiang, E. 2013, [Mon. Not. Roy. Astron. Soc.](#), 428, 2565
- Turner, N. J., Carballido, A., & Sano, T. 2010, [Astrophys. J.](#), 708, 188
- Turner, N. J., & Drake, J. F. 2009, [Astrophys. J.](#), 703, 2152
- Turner, N. J., Sano, T., & Dziourkevitch, N. 2007, [Astrophys. J.](#), 659, 729
- Umebayashi, T., & Nakano, T. 1981, [PASJ](#), 33, 617
- Vallis, G. K. 2006, Atmospheric and Oceanic Fluid Dynamics (Cambridge University Press, Cambridge)

- Vasyunin, A. I., Semenov, D., Henning, T., et al. 2008, [Astrophys. J.](#), **672**, 629
- Vidal-Madjar, A., Lecavelier des Etangs, A., Désert, J.-M., et al. 2003, [Nature](#), **422**, 143
- . 2008, [ApJ](#), **676**, L57
- Vidal-Madjar, A., Désert, J.-M., Lecavelier des Etangs, A., et al. 2004, [ApJ](#), **604**, L69
- Vidotto, A. A., Jardine, M., & Helling, C. 2010, [ApJ](#), **722**, L168
- . 2011a, [Mon. Not. Roy. Astron. Soc.](#), **411**, L46
- . 2011b, [Mon. Not. Roy. Astron. Soc.](#), **414**, 1573
- Visser, R., Geers, V. C., Dullemond, C. P., et al. 2007, [Astron. Astrophys.](#), **466**, 229
- Ward, W. R. 2009, in Lunar and Planetary Institute Science Conference Abstracts, Vol. 40, Lunar and Planetary Institute Science Conference Abstracts, 1477
- Wardle, M. 1999, [Mon. Not. Roy. Astron. Soc.](#), **307**, 849
- . 2007, [Ap&SS](#), **311**, 35
- Wardle, M., & Salmeron, R. 2012, [Mon. Not. Roy. Astron. Soc.](#), **422**, 2737
- Watson, A. J., Donahue, T. M., & Walker, J. C. G. 1981, [Icarus](#), **48**, 150
- Watson, A. M., Stapelfeldt, K. R., Wood, K., & Ménard, F. 2007a, Protostars and Planets V, 523
- Watson, D. M., Bohac, C. J., Hull, C., et al. 2007b, [Nature](#), **448**, 1026
- Whitten, R. C., Borucki, W. J., & Tripathi, S. 2007, [Journal of Geophysical Research \(Planets\)](#), **112**, 4001
- Woitke, P., Pinte, C., Tilling, I., et al. 2010, [Mon. Not. Roy. Astron. Soc.](#), **405**, L26
- Wolk, S. J., Harnden, Jr., F. R., Flaccomio, E., et al. 2005, [ApJS](#), **160**, 423
- Woodall, J., Agúndez, M., Markwick-Kemper, A. J., & Millar, T. J. 2007, [Astron. Astrophys.](#), **466**, 1197
- Wyatt, M. C. 2008, [ARA&A](#), **46**, 339
- Yelle, R. V. 2004, [Icarus](#), **170**, 167
- . 2006, [Icarus](#), **183**, 508
- Youdin, A. N., & Shu, F. H. 2002, [Astrophys. J.](#), **580**, 494
- Zahnle, K., Pollack, J. B., Grinspoon, D., & Dones, L. 1992, [Icarus](#), **95**, 1
- Zahnle, K. J., & Kasting, J. F. 1986, [Icarus](#), **68**, 462
- Zhu, Z., Nelson, R. P., Hartmann, L., Espaillat, C., & Calvet, N. 2011, [Astrophys. J.](#), **729**, 47