

Gas planet seismology and cooling

Steve Markham

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

The logo for the California Institute of Technology (Caltech), featuring the word "Caltech" in a bold, orange, sans-serif font.

California Institute of Technology

Pasadena, California

2022

Defended September 22nd, 2021

Acknowledgement

I would like to thank the California Institute of Technology faculty, especially my advisers Dave Stevenson and Konstantin Batygin, for their wisdom, without which none of this thesis would have been possible. I would like to thank the Geological and Planetary Sciences administrative staff, particularly Margaret Carlos, for making every day easier. I would also like to thank the NASA FINESST program for their funding (No. 80NSSC19K1520). Finally, my wife, for her patience and support.

Abstract

In this thesis I advocate for the enhancement of interdisciplinary expertise between atmospheric and interiors sciences. I illustrate the intimate connection between atmosphere and interior with four projects involving two major topics: giant planet seismology, and convective inhibition by condensation. First I advance a heuristic to evaluate generic localized excitation sources for giant planet seismicity, concluding observed oscillations on Jupiter may be caused by highly energetic rock storms lurking deep beneath the visible clouds. Next I develop a method to use existing spacecraft data to probe for seismic activity on giant planets, applying the method to Cassini data. This method finds possible evidence of p-modes on Saturn, excited to staggering amplitudes warping the surface of Saturn with kilometer scale displacements. Next I explore the impact of convective inhibition on Uranus and Neptune, finding that condensation of methane and water produces non-negligible corrections to these planets' thermal histories. Finally I explore a similar mechanism operating in the limit where condensing species are highly abundant. I find that considering convective inhibition, super-Earths can retain their primordial heat for longer than the age of the universe.

Published Content and Contributions

S. Markham & D. Stevenson, *Icarus* **306** pp. 200-213 (2018)

S.R.M. created the modeling heuristic, computed the stochastic impulse excitation efficiency from first principles order-of-magnitude considerations, drew the essential scaling conclusions, and wrote most the of the text of the manuscript.

<https://doi.org/10.1016/j.icarus.2018.02.015>

S. Markham, D. Durante, L. Iess & D. Stevenson, *PSJ*, **1**, 2 (2020)

S.R.M. derived the mathematical framework, built the software, analyzed the spacecraft residuals data, and wrote the text of the manuscript.

<https://iopscience.iop.org/article/10.3847/PSJ/ac091d>

S. Markham & D. Stevenson, *PSJ*, **2**, 4 (2021)

S.R.M. conceived of the project idea, built the model and software from scratch, created the figures from simulations, and wrote the text of the manuscript.

<https://iopscience.iop.org/article/10.3847/PSJ/ac091d>

Contents

1	Introduction: atmospheres are interiors on gas planets	1
2	Excitation and dissipation of giant planet seismicity	7
2.1	Introduction	8
2.2	Modeling the eigenfunctions of Jupiter’s seismic modes	13
2.2.1	Jupiter interior model	14
2.2.2	Displacement vector eigenfunction generation	15
2.3	Modeling amplitude responses	17
2.3.1	Monopole excitation	18
2.3.2	Dipole excitation	20
2.3.3	Spatial randomness	21
2.3.4	Temporal randomness	21
2.3.5	Excitation duration	24
2.3.6	Spherical harmonic superposition in the power spectrum	26
2.4	Constraining Q	27
2.4.1	Viscous and turbulent damping	27
2.4.2	Radiative damping	29
2.4.3	High-frequency modes: propagation through the stratosphere	33
2.4.4	Ohmic Dissipation by normal modes	34
2.4.5	Normal mode dissipation in the core	35

2.5	Possible physical excitation sources	35
2.5.1	Turbulent convection	36
2.5.2	Meteor strikes	38
2.5.3	Storms	39
2.6	Results and discussion	46
2.6.1	Excitation source parameter constraints	46
2.6.2	Predictions for other frequencies	47
2.6.3	Implications for gravity, Juno, Saturn, and ice giants	48
2.7	Conclusion	55
3	Gravitational seismology	57
3.1	Introduction	58
3.2	The forward model	62
3.2.1	Background	62
3.2.2	Numerical integration and model reproduction	67
3.2.3	Spectral model	68
3.3	Data stacking	70
3.3.1	Error sources	72
3.4	Analysis and results	74
3.5	Discussion	81
3.6	Appendix: intuition, tables, and derivations	86
3.6.1	Potential perturbation	87
3.6.2	Acceleration and velocity	89
3.6.3	Spacecraft sensitivity and model intuition	91
4	The thermal evolution of Uranus and Neptune	95
4.1	Introduction	96
4.2	Intuition and analytic approximations	99

4.3	Atmospheric model	104
4.3.1	Defining the boundaries of the stable layer	105
4.3.2	Radiative transfer across the stable layer	109
4.3.3	Meteorological implications	112
4.3.4	Interior implications	115
4.4	Evolutionary model	116
4.5	Discussion	121
5	The cooling and interiors of super-Earths	129
5.1	Introduction	130
5.2	Convective inhibition by silicate vapor	133
5.3	When does convective inhibition matter?	140
5.3.1	Estimating the thickness of the stable layer	146
5.4	Implications for thermal evolution	148
5.4.1	Formation considerations	155
5.4.2	Hydrogen pollution in the core and commentary on the evolu- tion of intermediate envelope mass planets	157
5.5	Discussion	159

Everything was beautiful, and
nothing hurt.

Kurt Vonnegut

Chapter 1

Introduction: atmospheres are interiors on gas planets

Since things are emerging and dissolving all the time, you cannot specify the point when this division will stop.

Lie Yukou

The field of planetary science originated with the study of our nearest planet: Earth. Earth's surface provides a clear interface between its atmosphere and interior. Earth's interior certainly couples to the atmosphere; much of the atmosphere itself originated from volcanic outgassing [Zahnle et al., 2010], the atmosphere weathers the crust, the ocean formed out of the atmosphere [Elkins-Tanton, 2011] which now likely plays a crucial role in plate tectonics [Mian and Tozer, 1990], and volcanic events today continue to impact atmospheric composition [Robock and Oppenheimer, 2003]. Nevertheless, it is an accurate approximation for many problems to treat the interior and atmosphere of the Earth as totally distinct. For this reason, when Earth scientists began to specialize into sub-disciplines, interiors and atmospheric sciences

likewise emerged as distinct. This distinction has carried over into the modern field of planetary science, with many scientists or research groups focusing primarily on either atmospheres or interiors. On giant planets, however, this distinction becomes meaningless. Without a physical surface to demarcate the interface between atmosphere and interior, any chosen boundary is necessarily arbitrary. Indeed, in a fully convective planet there is absolutely no difference between them at all. A part of the atmosphere today may become a part of the interior tomorrow, and vice versa. This thesis will focus on a few particular problems that illustrate the intimate connection between atmospheric dynamics and planetary interiors on giant planets.

New observations increasingly illustrate the importance of the connection between the interior and atmosphere of giant planets. Gravity experiments aboard the Juno and Cassini missions settled the debate about the depth of Jupiter and Saturn’s zonal winds: the winds howl at great depths, extending far beneath the traditional ‘atmosphere’ [Kaspi et al., 2020]. The ammonia distribution on Jupiter proved to be far more complex than anticipated, with ammonia depletion extending to great depths at most latitudes [Li et al., 2017]. The leading theory to explain this observation require strong updrafts originating from the interior [Guillot et al., 2020b] [Guillot et al., 2020a]. Global oceans in the deep interior of Uranus and Neptune likely set the volatile abundances of these planets’ atmospheres [Bailey and Stevenson, 2021]. Each of these observed phenomena illustrate separate instances of a similar theme: one cannot understand what is happening in a planet’s atmosphere without also considering what happens in the interior, and vice versa. This thesis will focus on two major illustrative examples of this paradigm.

First, a focus on giant planet seismology. Seismology revolutionized our understanding of the sun, moon, and Earth. Currently, seismology is likewise beginning

to revolutionize our understand of new celestial bodies: Mars, with the InSight mission [Knapmeyer-Endrun and Kawamura, 2020], Saturn with Cassini data [Hedman and Nicholson, 2013] [Markham et al., 2020], and soon perhaps Jupiter with ground based [Gaulme et al., 2011] and even Juno data [Durante et al., 2021]. Saturn’s rings provide an exquisitely sensitive and precise natural seismograph, allowing the clearest observations of a giant planet seismology. Seismic modes excite spiral density waves or bending waves in Saturn’s rings at radii resonant with their pattern speed. Dozens of these waves have been detected with optical depth measurements of the rings using stellar occultation measurements. These measurements have already elucidated a variety of unexpected properties of Saturn’s interior, including the existence of an extended dilute core [Fuller, 2014] [Mankovich and Fuller, 2021]. Because Saturn’s rings truncate interior to the D-ring, there exist many theoretical modes that have no resonances with Saturn’s rings, because they are too high in frequency. However, a growing assortment of ground-based and spacecraft measurements hint at truly staggering acoustic oscillations propagating through the interiors of Jupiter and Saturn [Gaulme et al., 2011] [Markham et al., 2020] [Durante et al., 2021]. The excitation source for all these oscillations remain uncertain, but we argue that their excitation is likely to take place near the atmosphere, where the divergence of the displacement eigenfunctions of acoustic oscillations is greatest. Likewise, the most efficient dissipation source we identify resides in the atmosphere. In Chapter 2, we investigate possible excitation and dissipation sources, concluding that no confirmed natural phenomenon is a likely candidate for exciting observed modes on Jupiter. We then speculate about the existence of “rock storms,” moist convective events analogous to thunderstorms powered by chemical reactions of silicates. Such storms, if they exist, could plausibly be sufficiently energetic to excite seismic oscillations to their observed amplitudes. In Chapter 3, we invent a novel technique to use existing spacecraft data from the Cassini mission to probe for seismic activity within Saturn.

We conclude that the observations are compatible with gravity field perturbations associated with huge low-order acoustic oscillations, that may warp the surface of Saturn with kilometer-scale deformation. We comment on whether these observations are consistent with the energetic storm excitation hypothesis, and mention that this same technique can be applied to Juno gravity data. Overall we find acoustic oscillations that resonate throughout the interior of giant planets are likely excited and dissipated by atmospheric phenomena—one cannot consider the interior and atmosphere separately.

Second, a focus on giant planet cooling and interior modeling. Traditional interior models assume adiabatic interiors, using equations of state and an isentropic assumption in order to produce density and temperature profiles of the interior (e.g., [Hubbard and Marley, 1989]). Traditional thermal evolution models neglect the complex dynamics at play in the atmosphere, and set a 1-bar equivalent temperature using simple assumptions about a well-behaved photosphere. It has been demonstrated that these assumptions are inappropriate in the case of condensation [Kurosaki and Ikoma, 2017], and sufficiently abundant condensates can inhibit convection altogether in the atmosphere [Guillot, 1995] [Friedson and Gonzales, 2017] [Leconte et al., 2017]. In this work, we demonstrate that many planets—Uranus, Neptune, and the most common exoplanets in the galaxy—possess sufficiently abundant condensates that this complication cannot be ignored. In Chapter 4, we place constraints on the importance of convective inhibition on the thermal evolution of Uranus and Neptune. We find that convective inhibition, which produces super-adiabatic temperature gradients, tends to influence thermal evolution in the opposite direction as latent heat release, which produces sub-adiabatic temperature gradients. The complex atmospheric dynamics play a crucial role in regulating the cooling rate of these bodies, and should be considered when formulating thermal evolution and interior models. In

Chapter 5, we apply this heuristic to exoplanets, in particular super-Earths. We find that considering silicate vapor profoundly impacts their thermal evolution. Contrary to the canonical picture of super-Earths as magma oceans with an overlying pure gas envelope (e.g., [Ginzburg et al., 2016] [Owen and Wu, 2017] [Vazan et al., 2018]), we find super-Earth cores may instead be supercritical and extremely hot. This high entropy interior state can persist for billions of years or longer if super-Earths retain a hydrogen atmosphere of order 1% its total mass. Based on observational and theoretical predictions, super-Earths with envelopes exceeding this value are likely to be common. If this model is correct, super-Earth internal heat flux can be small (of order Earth’s contemporary heat flux) very early in its evolution, then remains nearly constant for billions of years. These studies demonstrate additional examples in which atmospheric dynamics play a crucial role in the structure and evolution of planetary interiors.

I hope this thesis can serve as a small milestone in our grappling with the fundamental differences between our own planet, and the striking diversity of planets we continue to find populating the cosmos. We must inspect our biases as beings inhabiting one particular planet, and think deeply about the extent to which the assumptions we use here can be exported to other worlds.

Chapter 2

Excitation and dissipation of giant planet seismicity

And don't the stars, each being different, all work toward the same purpose?

Marcus Aurelius

Summary

The objective in this chapter is to formulate a generic heuristic to consider possible excitation and dissipation sources for seismic oscillations within Jupiter, motivated by observations. Recent (2011) results from the Nice Observatory indicate the existence of global seismic modes on Jupiter in the frequency range between 0.7 and 1.5mHz with amplitudes of tens of cm/s. Analysis of Jupiter's gravity field measured by Juno are consistent with ground-based observations. Currently, the driving force behind these modes is a mystery; the measured amplitudes are many orders of magnitude larger than anticipated based on theory analogous to helioseismology (that is, turbulent convection as a source of stochastic excitation). One of the most

promising hypotheses is that these modes are driven by Jovian storms. This work constructs a framework to analytically model the expected equilibrium normal mode amplitudes arising from convective columns in storms. We also place rough constraints on Jupiter’s seismic modal quality factor. Using this model, neither meteor strikes, turbulent convection, nor water storms can feasibly excite the order of magnitude of observed amplitudes. Next we speculate about the potential role of rock storms deeper in Jupiter’s atmosphere, because the rock storms’ expected energy scales make them promising candidates to be the chief source of excitation for Jovian seismic modes, based on simple scaling arguments. We also suggest some general trends in the expected partition of energy between different frequency modes. Finally we supply some commentary on potential applications to gravity, Juno, Cassini and Saturn, and future missions to Uranus and Neptune.

2.1 Introduction

Jupiter is the largest planet in the solar system, and our most accurate nearby representation of thousands of exoplanet analogues which seem to be equally or more massive, and comprised of approximately the same material. Understanding Jupiter’s formation history, then, is of great importance for understanding how planetary systems form in general. Understanding Jupiter’s interior is an essential part of modeling mechanisms for its formation; for example, the most popular explanation for Jupiter’s formation would suggest that the embryo Jupiter was a rocky planet early in its formation history, and we can perhaps expect a many Earth mass core to exist as a relic of that time [Pollack et al., 1996]. Additionally, there is an abundance of information about thermodynamics and materials physics to be learned by probing the detailed structure of Jupiter’s deep interior. Current methods of constraining Jupiter’s interior

(e.g., gravity and magnetic field measurements) are valuable, but cannot uniquely determine the internal structure. Therefore seismology will be an indispensable tool as we continue to try to study Jupiter’s interior [Gaulme et al., 2014]. Techniques applied to Jupiter can also be generalized to other planetary systems, and the scientific community has already expressed interest in applying similar techniques to Uranus, Neptune [Turrini et al., 2014][Elliot et al., 2017], and even Venus [Stevenson et al., 2015][Logonne and Johnson, 2015].

In 2011, a team from the Nice Observatory released a paper which claimed to have detected normal modes from Jupiter using an interferometer called SYMPA to perform Fourier transform spectroscopy [Schmider et al., 2007] [Gaulme et al., 2008] [Gaulme et al., 2011]. SYMPA measures line-of-sight Doppler shifts, so the detected displacements are primarily radial. For modes within the frequency range of sensitivity (high order p-mode overtones with frequencies above about $700\mu\text{Hz}$), SYMPA detected peak oscillation velocities on the order of 50cm/s . As outlined in Section 2.3.6, this value is the result of the superposition of multiple modes, and the velocity amplitudes of individual modes may be lower by a factor of 2 or 3. To put this in perspective, compare this to the maximum velocity amplitude in any single mode found in the sun, around 15cm/s [Christensen-Dalsgaard, 2014]. The total peak velocities measured on the sun can be substantially higher, because the solar observatory’s exquisite spatial resolution allows them to resolve much higher spherical order modes, and therefore more of an effect from superposition. Apparently the surface velocity amplitudes of both bodies are of similar orders of magnitude. It should be noted that since SYMPA’s measurements were limited to eight nights without continuous observations, and because the instrument has low spatial resolution, that these measurements are only relevant to low spherical order, high frequency modes (overtones of global scale modes). The power spectrum for the SYMPA measurements is found

on Figure 2.1.

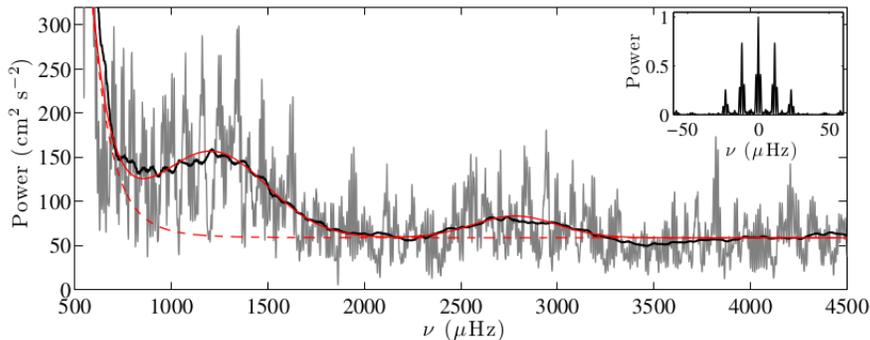


Figure 2.1: The observed power spectrum obtained by Gaulme et. al. [Gaulme et al., 2011]

This result is encouraging because it means the signal is sufficiently strong that meaningful measurements can be taken from Earth. It is puzzling, however, because it requires an excitation mechanism on Jupiter that is fundamentally different from what happens in the sun. We can conduct a simple order-of-magnitude calculation to enumerate the problem here. Since each normal mode behaves as a simple harmonic oscillator, its total energy is equal to its maximum kinetic energy. If its eigenfunction is described by displacement vector eigenfunction ξ (further discussed in Section 2.2 and illustrated in Figure 2.3) normalized to a magnitude of unity at the surface, then integrating over the whole body yields the total energy contained within a given normal mode.

$$E_{\text{mode}} = \frac{1}{2}v^2 \iiint \rho|\xi|^2 dV, \quad (2.1)$$

where v is the velocity amplitude, ρ is the spatially dependent density. $\iiint \rho|\xi|^2 dV$ is called the *modal mass* [Christensen-Dalsgaard, 2014]. The order-of-magnitude behavior of the eigenfunctions in the sun and in Jupiter should be similar, so we can neglect that factor since it is not a significant distinction between Jupiter and the sun. That is, for similar eigenfunction structure ξ , one can approximate the modal mass $\iiint \rho|\xi|^2 dV \sim fM$ to zeroth order—that is, the modal mass scales approximately linearly with the mass of the body [Christensen-Dalsgaard, 2014]. We can therefore

derive a zeroth order scaling relation of the form

$$E_{\text{mode}} \sim Mv^2, \quad (2.2)$$

where M is the mass of the body. Of course, this simplistic analysis ignores relevant details. The density contrast between the shallow and deep parts of the sun is much more extreme than for Jupiter; this affects both the modal mass and the excitation efficiency. Still, as a zeroth order first approximation to introduce the problem, we can place an order of magnitude estimate on the efficiency with which energy is injected into this normal mode by comparing the squared velocity amplitude to the luminosity per unit mass. The luminosity per unit mass in the sun is about $2 \text{ erg g}^{-1}\text{s}^{-1}$, and for Jupiter it's about $2 \times 10^{-6} \text{ erg g}^{-1}\text{s}^{-1}$ [Stevenson, 2021]. The problem then becomes immediately apparent. In order to produce the observed normal modes on Jupiter, the mechanism for injecting energy into the modes and retaining energy within the modes must be millions of times more efficient on Jupiter than on the sun. This excitation is computed in more detail in Section 2.5.1. At the moment, this disparity is not understood. The focus of this paper is to attempt to identify mechanisms that could deposit energy into Jupiter's normal modes orders of magnitude more efficiently than the sun.

Helioseismology revolutionized our understanding of the sun. Studying the sun's seismic modes definitively answered questions ranging from the solar neutrino problem, the sun's convective and radiative zones, the existence of deep jet streams, the age of the sun, and its differential rotation [Deubner and Gough, 1984]. Today, many fundamental questions about Jupiter may be answered with the same treatment. Di-seismology (an alternative word with equivalent meaning to Jovian seismology, first used by Mosser [Mosser, 1994]) could illuminate a condensed or diffuse core. It could

provide more detailed information about the physical properties of liquid metallic hydrogen, and reveal the existence of regions of static stability or exotic chemical cloud decks deep below the visible surface. With so much to gain from dioseismology, it is a worthwhile endeavor to understand.

Unfortunately, the existing data for normal modes has rather low signal-to-noise ratio and is regarded by some as suspect, in part because we lack an understanding of how the modes could be excited. If we can develop a more quantitative understanding of their excitation and dissipation, then we could corroborate the possibility of their existence and motivate future observational programs. Such insights would be useful diagnostic tools to design space-based seismometers for future missions to Jupiter, as well as other planets in the solar system.

The 1994 comet strike of Shoemaker-Levy sparked much interest into the possibility of Jovian seismic mode excitation by the cometary impact. Competing calculations made contradictory predictions at the time. Dombard & Boughn did not predict measurable amplitudes [Dombard and Boughn, 1995], but others such as Lognonne, Mosser and Dahlen predicted measurable amplitudes for a sufficiently energetic impact [Lognonne et al., 1994]. As it turns out, the seismic modes associated with SL9 were never detected [Mosser et al., 1996]. In this work, we generalize the framework constructed by Dombard and Baughn for the expected seismic response to the impact of Shoemaker-Levy with Jupiter [Dombard and Boughn, 1995], as well as the work for the sun and other stars made by Peter Goldreich and others [Goldreich and Keeley, 1977] [Goldreich and Kumar, 1994], to try to propose any plausible candidates for Jovian seismic mode excitations. These mechanisms should be both explanatory and predictive; if a certain model explains the observed results, it can also predict what amplitudes should be expected in frequency ranges which have not yet been detected.

Future measurements, then, can provide support or refutation for different models proposed here.

This paper will begin with an introduction to our model of Jupiter and the treatment of its normal mode displacement eigenfunctions. We will then outline some general mathematical tools to abstractly model and parameterize different types of excitation sources. Next we will investigate a few important dissipation mechanisms to try to place some constraints on Jupiter's modal Q . We will then apply all these tools to some potential physical excitation sources, to try and estimate an order-of-magnitude for what velocity amplitudes these mechanisms might excite. Finally we will discuss our findings, with some brief remarks on potential applications of these findings to Jupiter and other planets.

2.2 Modeling the eigenfunctions of Jupiter's seismic modes

Jupiter, like any other object, can behave as a resonator. The modes of interest for explaining the results from SYMPA are acoustic modes. These modes are trapped in a cavity bounded from below by Snell's law; the ray path enters Jupiter's interior from the surface obliquely. As the ray descends, the sound speed increases, which continuously deflects the ray laterally until it travels tangentially at the minimum radius and begins to return to the surface. Modes below the acoustic cutoff frequency are bounded from above by Jupiter's small scale height (relative to the mode's local wavelength) as it approaches the photosphere. This resonator is rather efficient, since the viscosity in Jupiter is very low. Much work on this basic physics has been done, primarily with applications to helioseismology and asteroseismology in general [Christensen-Dalsgaard, 2014]. There has also been some qualitative work on applying

these ideas to Jupiter [Bercovici and Schubert, 1987]. Some progress can be made by qualitative order of magnitude arguments, but in order to argue for a coherent global picture, a numerical model for the structure of the eigenfunctions, the planetary interior, and the planetary atmosphere must be specified.

2.2.1 Jupiter interior model

The first important step in this modeling process is choosing a suitable Jupiter interior model. This model can in principle be as detailed as desired, but for our purposes we wanted to use the simplest, most generic possible model that can still accurately model Jupiter's behavior because our focus here is on understanding the excitation and dissipation, not the precise evaluation of modal eigenfrequencies. This is desired for simplicity of outcome (no frequency splitting between modes of the same spherical order), as well as simplicity of inputs (homogeneous adiabatic interior), and finally for its ability to easily adapt to explain other planets. We therefore begin with a simple $n = 1$ polytrope equation of state:

$$P = K\rho^2, \tag{2.3}$$

with K chosen to approximate a hydrogen/helium mixture. This model is quite accurate for Jupiter's interior, but does a bad job at accurately describing the behavior near the surface. We therefore adjust the equation of state by adding a $\rho^{1.45}$ term consistent with an adiabatic ideal gas equation of state. The two should connect smoothly in between. The equation of state then takes the form

$$P = K_1\rho^2 + K_2\rho^{1.45}, \tag{2.4}$$

where K_1 and K_2 are chosen to match Galileo measurements for Jupiter's upper troposphere, and to get the right radius and mass.

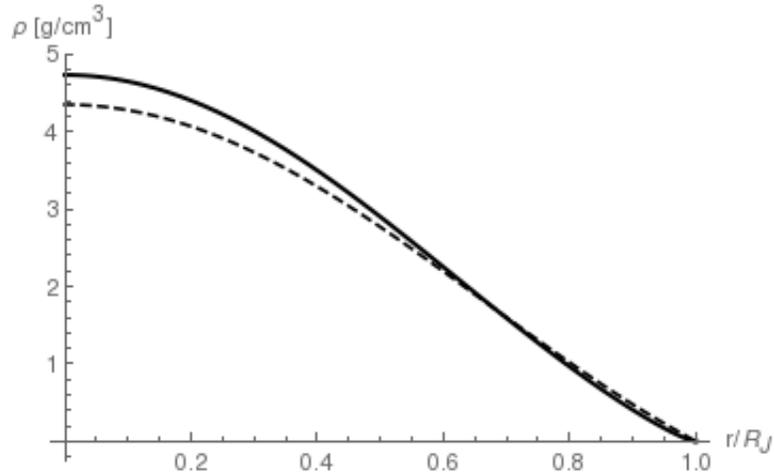


Figure 2.2: Comparison between the hydrostatic interior model using our modified equation of state (solid) and the interior model predicted using an $n=1$ polytropic equation of state (dashed).

Notice that since the ρ^2 term is small near the surface, the ideal gas term will then dominate. Additionally, we investigated the effects to the eigenfunction if we include an isothermal component to the atmosphere above the photosphere. We found that doing so affected observed mode amplitudes by less than 5%. Since the arguments we are making here are generic and correct to no more than an order of magnitude, we elected to neglect the isothermal part of the atmosphere for the purpose of generating the global eigenfunctions. We do, however, discuss the effects of radiative damping in the isothermal part of the atmosphere as it relates to Jupiter’s quality factor in Section 2.4.

2.2.2 Displacement vector eigenfunction generation

After setting upon an interior model which satisfactorily represents the important aspects of Jupiter’s interior, we used the stellar oscillation code GYRE [Townsend et al., 2013] to generate eigenfunctions for Jupiter’s interior. The first four $l=2$ modes are shown on Figure 2.3

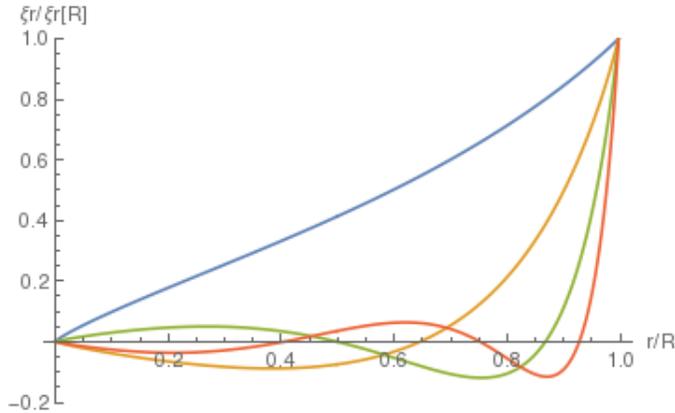


Figure 2.3: An example of the radial eigenfunction produced for our interior model the first four $l = 2$ modes. ξ represents the amplitude of the eigenfunction in the radial direction at that depth, normalized such that $\xi = 1$ at the 1 bar level.

Because we are using a non-rotating, spherically symmetric model for Jupiter, the modes are exactly spherical harmonics. The behavior of the eigenfrequencies is shown on Figure 2.4.

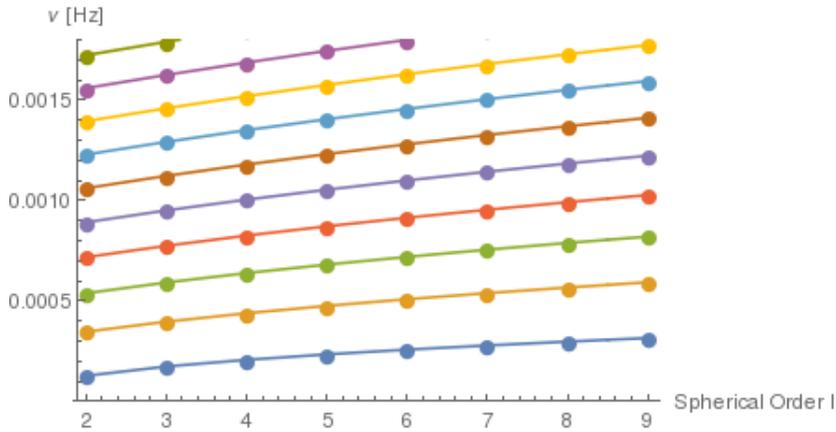


Figure 2.4: Frequencies of low-order modes. Frequency increases gradually with increasing spherical order l and quickly with equal spacing with increasing radial order n , where n defines the number of nodes of the mode as shown in Figure 2.3.

The total observed displacement on the surface of Jupiter is expressed as

$$\mathbf{x}(\mathbf{r}, t) = \sum_{nlm} a_{nlm}(t) \boldsymbol{\xi}_{nlm}(\mathbf{r}), \quad (2.5)$$

where $a_{nlm}(t)$ is a time-dependent amplitude for each normal mode, and $\boldsymbol{\xi}_{nlm}(\mathbf{r})$ is a spatially dependent eigenfunction displacement vector of radial order n , spherical order l and azimuthal order m . Canonically the eigenfunctions are separated into a radial and horizontal part $\xi_r(r)$ and $\xi_h(r)$ so that the full displacement vector eigenfunction takes the form

$$\boldsymbol{\xi}_{nlm}(r, \theta, \phi) = \left[\xi_r(r) \hat{\mathbf{r}} + \xi_h(r) \hat{\boldsymbol{\theta}} \frac{\partial}{\partial \theta} + \frac{\xi_h(r)}{\sin \theta} \hat{\boldsymbol{\phi}} \frac{\partial}{\partial \phi} \right] Y_l^m(\theta, \phi). \quad (2.6)$$

2.3 Modeling amplitude responses

For the purposes of this problem, we will approximate the modes of Jupiter as a set of orthogonal, undamped harmonic oscillators. This is a valid approximation because our assumed timescales for damping are proportional to a very large Q . Specifically, $\tau_{dec} = 2Q/\omega$ for a given mode, and we expect Q to be $\sim 10^6 - 10^8$, which we will justify later in this chapter. Since Q is so large, we will approximate the timescale between excitation events to be much less than the ringdown timescale. As another approximation, we will assume no “leaking” energy between modes, i.e., the modes are linear and non-linear interaction terms are neglected, but this will be discussed in our evaluation of Q . Now we write down the equation of a driven harmonic oscillator

$$\ddot{a} + \omega^2 a = F(t) \quad (2.7)$$

for each mode, where a is the time-dependent coefficient from Equation 2.5, ω is the appropriate eigenfrequency, and $F(t)$ is an effective force. For a mass on a spring, this effective force would simply be the physical force divided by the mass of the object. In this simplified case, the whole driving force acts on the whole mass, but since our excitation sources may be localized, we must define the effective force following Dombard & Boughn [Dombard and Boughn, 1995]. This effective force should account

for the coupling between the eigenfunction ξ and the physical force density vector field $\mathbf{f}(\mathbf{r}, t)$, and scale it by the total modal mass:

$$F(t) = \frac{\iiint \xi \cdot \mathbf{f} dV}{\iiint \rho(\mathbf{r}) |\xi|^2 dV}. \quad (2.8)$$

In the following subsections, a few simple generic models for force density will be examined. Later in the chapter, these generic models can be combined to approximately model physical phenomena to an order of magnitude.

2.3.1 Monopole excitation

An explosion is an example of a monopolar force density field. Following the model of Dombard & Boughn [Dombard and Boughn, 1995] for a comet impact, we can model a spherical explosion centered on a point \mathbf{r}_0 as

$$\mathbf{f}(\mathbf{r}, t) = \delta P \delta(\mathbf{r} - \mathbf{r}_0) \hat{\mathbf{r}}_n \phi(t), \quad (2.9)$$

where δP is the pressure pulse caused by the explosion, $\delta(\mathbf{r} - \mathbf{r}_0)$ is a spherical delta function, $\hat{\mathbf{r}}_n$ is a unit vector pointing away from \mathbf{r}_0 , and $\phi(t)$ is an arbitrary function in time which sets the timescale of the explosion. Substituting this \mathbf{f} into Equation 2.8, using Gauss' theorem, and noting that the energy of the bubble is equal to its pressure perturbation times the volume of the bubble gives

$$F(t) = \frac{E_s/V \phi(t) \iiint_s \nabla \cdot \xi d^3r}{\iiint \rho(r) |\xi|^2 d^3r} \quad (2.10)$$

where V is the volume of the bubble, E_s is its energy, and the integral in the numerator is over the volume of the explosion. Our task is now to compute this expression. Assuming that there is very little non-radial variation in $\nabla \cdot \xi$ (which is a very good approximation near the planetary surface for excitation sources with length scales on

the order of hundreds of kilometers, as long as we are talking about spherical orders less than several thousand), we can simplify

$$\iiint_s \nabla \cdot \xi d^3r \rightarrow \pi \int_{-b}^b \nabla \cdot \xi(x^2 - b^2) dx, \quad (2.11)$$

where b is the radius of the bubble. We can do a Taylor series $\nabla \cdot \xi$ up to a fourth derivative in ξ_r , which is more than a good enough approximation for these length scales with $n < 100$, we can compute this integral directly to be

$$\iiint_s \nabla \cdot \xi d^3r \approx 4/3\pi b^3 \left(\frac{\partial \xi_r}{\partial r} + 1/10\pi b^2 \frac{\partial^3 \xi_r}{\partial r^3} \right), \quad (2.12)$$

since the displacement eigenfunction for low spherical order l modes is primarily radial near the surface. This approximation breaks down for higher spherical order modes, where the tangential component of the eigenfunction is more important. To ensure the accuracy of this method, we compared the exact numerical integration of the divergence of the eigenfunction through the bubble to this approximation, and found excellent agreement for the first 50 modes within less than 1%. In fact, for the first 25 modes (which are the ones in the frequency range of interest), the third order term in also unnecessary. Since $4/3\pi b^3$ is constant, we can take it out of the integral. It's also the volume of the bubble, so we can cancel it with V . Thus if we approximate the spatial and time dependence to be separable quantities, we can write

$$F(t) \simeq \frac{E_s \frac{\partial \xi_r}{\partial r}}{\iiint \rho(r) |\xi|^2 d^3r} \phi(t) \equiv F_0 \phi(t). \quad (2.13)$$

For high radial order modes ($n > 30$), the $\frac{\partial^3 \xi_r}{\partial r^3}$ term from Equation 2.12 should be included for accuracy. Now we can solve the harmonic oscillator equation

$$\ddot{a} + \omega^2 a = F(t) = F_0 \phi(t), \quad (2.14)$$

where F_0 encodes the geometric information, assumed spatially static in space and wrapped in a time-dependent wrapper function $\phi(t)$. Since F_0 is a constant in time, in the one dimensional harmonic oscillator equation it can be considered to be a constant. We now solve this equation by taking its Fourier transform, so that

$$a(t) = \frac{F_0}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{\hat{\Phi}(\nu)}{\omega^2 - \nu^2} e^{i\nu t} d\nu, \quad (2.15)$$

where $\hat{\Phi}(\nu)$ is the Fourier dual of $\phi(t)$. All that is required, then, is to choose a form of $\phi(t)$ and Equation 2.15 is solvable.

2.3.2 Dipole excitation

The simplest way to think of a dipole is two point sources separated by some distance ϵ . This is expressed mathematically as

$$\mathbf{f}(\mathbf{r}, t) = f_0[-\delta(\mathbf{r} - (\mathbf{r}_0 + \epsilon)) + \delta(\mathbf{r} - \mathbf{r}_0)]\hat{\mathbf{r}}, \quad (2.16)$$

where $\hat{\mathbf{r}}$ is the outward pointing radial vector with respect to the center of Jupiter, and f_0 is the normalization coefficient. Provided ϵ is small compared to the wavelength of the mode, a reasonable first order approximation, we can evaluate

$$\begin{aligned} \iiint \mathbf{f}(\mathbf{r}) \cdot \boldsymbol{\xi} dV &\approx -f_0 \frac{\partial \xi_r}{\partial r} \epsilon \\ \implies F_{0,dipole} &\sim \frac{f_0 \frac{\partial \xi_r}{\partial r} \epsilon}{\iiint \rho(r) |\boldsymbol{\xi}|^2 dV} \end{aligned} \quad (2.17)$$

using the fundamental theorem of calculus and the properties of the δ function. For our purposes, this is a sufficient description of a generic dipole excitation. For a specific model, of course, one must evaluate a physically reasonable f_0 in the context of the problem. Note the striking similarity between localized dipole and monopole

excitation sources, which for low spherical and radial order modes are mathematically identical, except with different expressions for F_0 .

2.3.3 Spatial randomness

In all of the above results, the predicted amplitudes implicitly include a spherical harmonic evaluated at a particular point on Jupiter's surface. If at any instant there are N storms within Jupiter's atmosphere then the total displacement would scale as

$$\sum_{i=1}^N |Y_l^m(\theta_i, \phi_i)|^2 = N. \quad (2.18)$$

In the limit of large N and assuming the storms are randomly distributed, the RMS value of this is simply $N^{1/2}$ larger than the amplitude of a single storm, because of the normalization properties of spherical harmonics. Of course this would break down in the limit of small number of storms, or storms with a preferred location, as may be the case. In this case, there may be more complicated dependence of amplitude on the quantum numbers than the results we report below.

2.3.4 Temporal randomness

Having shown that spatial randomness of storm occurrence can be averaged out to be irrelevant, the next logical question is what to do about the issue of the storms being stochastic in time. Because of the findings in the previous section, geometrical effects can be neglected. The amplitude response from a single excitation event j takes the form

$$\mathbf{x}_j(\mathbf{r}, t) = \sum_{nlm} a_{nlm,j} \boldsymbol{\xi}_{nlm} \exp(i\omega_{nlm}(t - t_j)). \quad (2.19)$$

The full expression after N excitations can be written

$$\mathbf{x}(\mathbf{r}, t) = \sum_{nlm} \sum_{j=1}^N [a_{nlm,j} \exp(-i\omega_{nlm}t_j)] \boldsymbol{\xi}_{nlm} \exp(i\omega_{nlm}t). \quad (2.20)$$

The task now is to evaluate

$$\sum_{j=1}^N a_{nlm,j} \exp(-i\omega_{nlm}t_j), \quad (2.21)$$

since t_j is a random variable, and $\exp(-i\omega_{nlm}t_j)$ is a 2π periodic function, the above expression is simply a random walk in the complex plane. The final expression for the amplitude without dissipation after N excitation events then can be written

$$\mathbf{x}(\mathbf{r}, t) \approx \sqrt{N} \sum_{nlm} a_{nlm,j} \boldsymbol{\xi}_{nlm} \cos(\omega_{nlm}t + \phi), \quad (2.22)$$

where ϕ is an arbitrary phase and $a_{nlm,j}$ is now the expected value of amplitude for a given type of excitation. Because the energy of the mode scales as $|x|^2$, energy grows linearly with the number of excitation events, while amplitude grows with its square root.

Now we calculate the equilibrium mode amplitudes including dissipation. If a single excitation imparts energy E_0 , and the expected value for total energy input grows linearly with the number of excitation events, then we can equate average power input to energy dissipation,

$$\frac{E_0}{\tau_s} = \frac{E_{eq}}{\tau_{dec}}, \quad (2.23)$$

where τ_s is the characteristic timescale between excitation events, and τ_{dec} is the decay timescale, related to the quality factor Q according to

$$\tau_{dec} = \frac{2Q}{\omega}. \quad (2.24)$$

Of course this assumes that there is an equilibrium i.e., the time between excitation events is much shorter than the time to decay. If this were not so, it would be evident in continued observations that show a variation of mean amplitude over time. The mean equilibrium energy associated with an excitation source that imparts energy E_0 stochastically in time is

$$E_{eq} = \frac{2E_0Q}{\tau_s\omega}. \quad (2.25)$$

It should be noted that these values are not expected to be constant in time. The arguments here are only statements about the average equilibrium amplitudes; in reality, one observes a specific amplitude at a specific time rather than a long term average. It is therefore perfectly consistent with this framework to have periods of quiescence, and periods of larger amplitudes. The expected value, however, will tend toward the calculations shown here.

As argued in Section 2.1, the energy of a mode described by displacement eigenfunction ξ is

$$E_0 = \frac{1}{2}a^2\omega^2 \iiint \rho|\xi|^2 dV, \quad (2.26)$$

where a is the amplitude response resulting from a single excitation. Ignoring time dependence and focusing on amplitude, we can use $a = F_0/\omega^2$. In reality, the form of a will depend on $\phi(t)$, but that's the focus of the following section. Rewriting E_0 as

$$E_0 = \frac{1}{2}F_0^2\omega^{-2} \iiint \rho|\xi|^2 dV. \quad (2.27)$$

The equilibrium amplitude is

$$a_{eq} = \left(\frac{E_{eq}}{\omega^2 \iiint \rho|\xi|^2 dV} \right)^{1/2}, \quad (2.28)$$

so using Equation 2.25, the equilibrium amplitude can be written

$$a_{eq} = F_0 \left(\frac{Q}{\tau_s \omega^5} \right)^{1/2}. \quad (2.29)$$

This relation is of enormous consequence for Jovian seismic mode excitation. The forcing magnitude of a generic source is proportional to its energy scale. Equation 2.29 implies that the equilibrium amplitude obeys

$$a_{eq} \propto \frac{E_s}{\tau_{ex}^{1/2}}, \quad (2.30)$$

While the power output of these collective excitation sources by definition follows the relationship

$$\dot{E} = \frac{E_s}{\tau_{ex}}. \quad (2.31)$$

Hence, for a fixed power budget, it is more favorable to have less frequent, more energetic excitation events than more frequent, less energetic excitation events.

2.3.5 Excitation duration

The dynamics of storms are immensely complex. Decades of detailed research have gone into modeling storms on Earth for which we have excellent data, and still there is no basic universal picture for their dynamics [Ludlam, 1980]. For the purposes of this paper, the time-dependent aspect of storms as an excitation source will be modeled simplistically. In particular, the δ -function, a Gaussian function, and a single sinusoidal pulse will be considered. Recalling Equation 2.15, we can solve for each of these. For a δ -function,

$$\phi(t) \rightarrow \delta(t) \implies \tilde{\Phi}(\nu) \sim \text{const.} \quad (2.32)$$

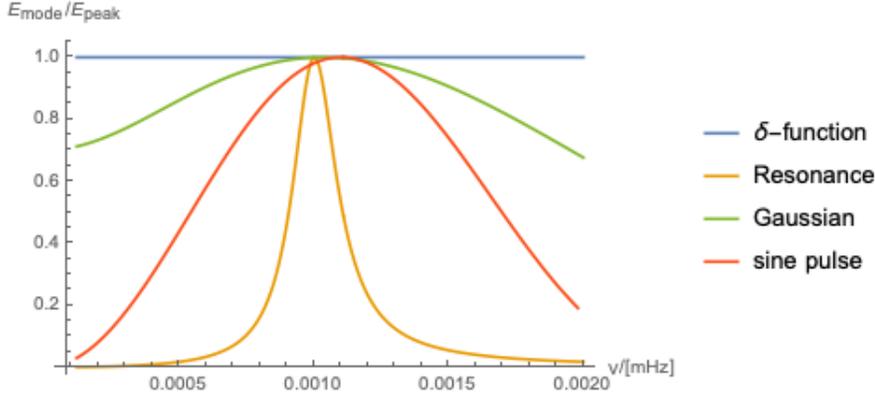


Figure 2.5: Relative excited energy for different forcing mechanisms with a shared characteristic timescale for range of sample eigenfrequencies.

For a Gaussian,

$$\phi(t) \rightarrow \exp(-t^2/\Delta t^2) \implies \tilde{\Phi}(\nu) \sim \exp(\Delta t^2 \nu^2/4) \quad (2.33)$$

where σ sets the width of the Gaussian and has dimensions of time. In this case, the narrower the Gaussian for the input $\phi(t)$, the broader the excitation spectrum in frequency space. For a single sinusoidal pulse,

$$\phi(t) \rightarrow 1 \text{ for } |t| < \Delta t, 0 \text{ elsewhere} \implies \tilde{\Phi}(\nu) \sim \text{sinc}\left(\frac{\nu - 1/\Delta t}{2/\Delta t}\right) - \text{sinc}\left(\frac{\nu + 1/\Delta t}{2/\Delta t}\right). \quad (2.34)$$

Finally, one may consider continuous sinusoidal forcing, for example due to tides, that massively prefers modes with resonant frequencies. The excitation sources considered here are put through Equation 2.15, and the results are plotted in Figure 2.5.

Note the units on the y-axis are arbitrarily normalized to emphasize relative excitation of different frequency modes. For our storm models, we will simply use a δ -function in time so as not to bias our results to match a particular assumption of complex dynamics, but the reader should note that an event with a characteristic

timescale may preferentially excite certain frequencies more efficiently than others. The resonant forcing curve on Figure 2.5 assumes $Q = 10^8$.

2.3.6 Spherical harmonic superposition in the power spectrum

So far these calculations have focused on the excitation of a single mode given some source. This section remarks briefly on the expected power spectrum that would be measured from all visible modes combined. We begin with the general mathematical relationship

$$\int_{\phi=0}^{2\pi} \int_{\theta=0}^{\pi} \left| \sum_{l=0}^N \sum_{m=-l}^l Y_l^m \right|^2 \sin \theta d\theta d\phi = N. \quad (2.35)$$

Recall that the expression for excitation amplitude given the sources investigated here depend on $\frac{\partial \xi_r}{\partial r}$ and ω . The only dependence on Y_l^m is encoded in the denominator, since

$$|\xi|^2 = \boldsymbol{\xi} \cdot \boldsymbol{\xi} = \xi_r^2 |Y_l^m|^2 + \xi_h^2 \left| \frac{\partial Y_l^m}{\partial \theta} \right|^2. \quad (2.36)$$

This expression is integrated over a sphere, so the $|Y_l^m|^2$ averages away. The $\left| \frac{\partial Y_l^m}{\partial \theta} \right|^2$ is retained, but for sufficiently low-order spherical harmonics near the surface, the motions are mostly radial, so the second term can be neglected. Since ξ_r is independent of m and only weakly dependent on l for low spherical order modes, this implies that to a good approximation the excitation amplitude is a function of frequency only. This means that assuming SYMPA is sensitive to spherical orders up to about $l = 3$, the power spectrum calculated for one spherical mode can be approximately doubled to account for the full power spectrum. On the sun, where resolution is greatly enhanced and detection of very high spherical order modes are possible, we expect this principle to have a more substantial effect on peak measured velocity, because the higher resolution implies detection of higher l modes and therefore larger N in

Equation 2.35.

2.4 Constraining Q

As demonstrated in the previous section, our equilibrium mode amplitudes scale as $Q^{1/2}$. Having an idea for the order of magnitude of Jupiter's quality factor, then, is essential to making a predictive theory. One possibility is that the effective Q is actually determined by the interaction of modes with each other rather than intrinsic dissipation. However, these interactions are probably negligible [Luan et al., 2017], so for the moment we will focus on intrinsic processes. Much work has already been done estimating Jupiter's tidal Q [Wu, 2005a]. The primary coupling mechanism between Jupiter and its satellites are inertial modes, which are bounded between $0 < \omega < 2\Omega$, where Ω is Jupiter's spin rate [Wu, 2005b][Wu, 2005a]. The fundamental p-mode of Jupiter has a period on the order of two hours, much shorter than Jupiter's spin rate. Therefore dissipation associated with these inertial modes is irrelevant to the study at hand. Nevertheless, it is possible to place some constraints on our expected value of Q using mechanisms we know must dissipate energy.

2.4.1 Viscous and turbulent damping

The most obvious dissipation mechanism is viscosity. Starting with the standard Stokes-Kirchhof viscous dissipation expression for acoustic waves [Landau and Lifshitz, 1959],

$$\bar{E} = -\frac{1}{2}k^2v_0^2V_0 \left[\left(\frac{4}{3}\eta + \zeta \right) + \kappa \left(\frac{1}{c_V} - \frac{1}{c_p} \right) \right], \quad (2.37)$$

where k is the sound wavenumber, v_0 is the fluid displacement velocity, V_0 is the volume occupied by the sound wave, η is dynamic viscosity, ζ is the second viscosity, κ is the fluid's thermal conductivity, c_V is the specific heat capacity of the fluid at constant volume and c_p is the specific heat capacity at constant pressure. As a

simplifying assumption, assume $\zeta \sim \eta$. Now compare the relative importance of the first and second bracketed terms on the right hand side of Equation 2.37. Noting $\kappa(1/c_V - 1/c_p) = \kappa/c_p(\gamma - 1)$ and plugging in typical values for hydrogen, the second term is $\sim 10^{-12}$ in cgs units, compared to viscosity which is $\sim 10^{-3}$. So the second term can be neglected. Now we write

$$\bar{\dot{E}} \approx -k^2 \omega^2 |\xi|^2 V_0 \eta. \quad (2.38)$$

Integrating over differential volume elements, we get a total average power dissipation of

$$\bar{\dot{E}} = \omega^2 \iiint k^2 \eta |\xi|^2 dV. \quad (2.39)$$

Now to compute Q , note

$$Q \equiv 2\pi \frac{E_{\text{stored}}}{\oint \dot{E} dt} = \omega \frac{\iiint \rho |\xi|^2 dV}{\iiint k^2 \eta |\xi|^2 dV}. \quad (2.40)$$

Now for order-of-magnitude estimates, assume k to be constant to zeroth order in most of the interior. Substitute average, constant values $\bar{\rho}$ and $\bar{\eta}$ and take them out of the integral. The expression for Q then reduces to

$$Q \sim \frac{\omega \bar{\rho}}{k^2 \bar{\eta}}. \quad (2.41)$$

Noting $k \sim \frac{2\pi(n+1)}{R_{\text{J}_4}}$ where n is the radial order, and $\bar{\rho} \approx 1.33$. This gives

$$Q \sim 10^{18} \left(\frac{\omega}{10^{-3} \text{s}^{-1}} \right) \left(\frac{1}{n_r + 1} \right)^2 \left(\frac{10^{-2} \text{cm}^2 \text{s}^{-1}}{\eta} \right), \quad (2.42)$$

where n_r is the radial order of the mode. In reality, turbulence will increase the effective viscosity of the system. Turbulent viscosity should be weak, because Jupiter's convection overturn timescale is much longer than the period of the normal modes,

which means eddies larger than the local scale height do not act viscously [Goldreich and Nicholson, 1977]. Assuming $\eta \sim 10^3$ as is assumed for tides [Goldreich and Nicholson, 1977], the estimate for Q goes to $\sim 10^{13}$. So viscosity and turbulence turn out to be very weak damping mechanisms.

2.4.2 Radiative damping

The most important mandatory loss of energy occurs as a result of radiative damping in Jupiter's stratosphere. Below the tropopause, a displaced parcel of fluid will expand or contract adiabatically, but remain in equilibrium with its convective surroundings, which by definition follow an adiabat. However, the same displacement in the isothermal atmosphere would cause a displaced parcel to warm as it was displaced downward, bringing it out of equilibrium with its surroundings. The warm parcel would then radiate away heat while displaced. Conversely, a parcel displaced upwards will radiate less heat. Importantly, this introduces a phase difference between the oscillations in temperature associated with a wave and the oscillations in pressure or density. The resulting hysteresis is the dissipation arising from radiative damping. We are primarily interested in the case where the tropopause occurs at a location where the waves of interest are no longer propagating (i.e., are evanescent) so that the effect of the wave on the atmosphere is merely the vertical displacement of a column of gas. In the low frequency limit, the fractional density perturbation and the velocity amplitude increases only slightly with height, with a characteristic e-folding distance of $c^2/\omega^2 H \sim R_{\text{t}}$.

First we calculate the radiative damping timescale τ_{rad} . Assuming the atmosphere is optically thin in the stratosphere, and gray opacity such that emission and absorption are described by the same constant, we imagine a parcel in an isothermal environment of temperature T_0 raised to temperature $T_0 + T'$ by being displaced by

seismic modes. It is illuminated from below by the ammonia cloud deck of optical depth unity at Jupiter's effective temperature T_e . The total energy radiated from the plane parcel up and down is

$$2\sigma(T_0 + T')^4 \rho \kappa dz, \quad (2.43)$$

and energy absorbed from below is

$$\sigma T_e^4 \rho \kappa dz. \quad (2.44)$$

In equilibrium with $T' \rightarrow 0$, we obtain the standard result $T_0 = T_e/2^{1/4}$. On the other hand, out of equilibrium with time-dependent T' :

$$\rho c_p dz \frac{dT'}{dt} = -8\sigma T_0^3 T' \rho \kappa dz \quad (2.45)$$

We can write T_0 in terms of T_e from the standard result, so that $8\sigma T_0^3 \rightarrow 4\sigma T_e^3$. Now defining a radiative time constant τ_{rad} according to

$$-\frac{T'}{\tau_{rad}} = \frac{dT'}{dt} \quad (2.46)$$

reveals

$$\tau_{rad} = \frac{c_p}{4\sigma T_e^3 \kappa} \quad (2.47)$$

using values from Galileo, and employing a functional form of pressure dependent opacity for hydrogen as

$$\kappa \sim 10^{-2} \left(\frac{p}{1 \text{ bar}} \right) \text{cm}^2/\text{g} \quad (2.48)$$

yields

$$\tau_{rad} \approx 5 \times 10^7 \left(\frac{1 \text{ bar}}{p} \right) \text{sec}. \quad (2.49)$$

Now to calculate dissipation. Starting with the ideal gas law,

$$p = \frac{k_B}{\mu} \rho T \quad (2.50)$$

$$\implies dp = \frac{k_B}{\mu} (d\rho T + \rho dT). \quad (2.51)$$

We are interested in the part of the pressure perturbation associated with the change in temperature. So

$$\delta p \approx \frac{k_B}{\mu} \rho_0 T' \quad (2.52)$$

to first order. In general for a displaced parcel

$$\frac{\partial T'}{\partial t} = -v \left(\frac{\partial T}{\partial z} - \frac{\partial T}{\partial z} \Big|_{ad} \right) + \frac{T'}{\tau_{rad}}, \quad (2.53)$$

where v is the local velocity of the parcel caused by normal mode oscillations. In the isothermal atmosphere, $\frac{\partial T}{\partial z} \rightarrow 0$. In general for a plane-parallel atmosphere, $\frac{\partial T}{\partial z} \Big|_{ad} = g/c_p$. So assuming v and T' oscillate with the normal mode and are therefore $\propto \exp(i\omega t)$, we can rewrite

$$T' = \frac{vg}{c_p(i\omega + 1/\tau_{rad})}. \quad (2.54)$$

Assuming $\frac{1}{\omega\tau} \ll 1$, true using characteristic values of $\tau \sim 5 \times 10^7$ s and $\omega \sim 10^{-3}$ s $^{-1}$, this can be written as

$$T' \approx \frac{vg}{i\omega c_p} \left(1 - \frac{i}{\omega\tau_{rad}} \right). \quad (2.55)$$

Substituting this into the ideal gas equation yields

$$\delta p \approx \frac{k_B}{\mu} \frac{v}{H c_p i\omega} \left(1 - \frac{i}{\omega\tau_{rad}} \right) p_0, \quad (2.56)$$

by noting $g = \frac{c_s^2}{\gamma H}$ where c_s is the speed of sound and γ is the adiabatic index; and that $p_0 = c_s^2 \rho_0 / \gamma$. The task now is to compute the energy dissipated in one normal

mode period.

$$\oint v \delta p dt = \int_0^{2\pi/\omega} v \delta p dt. \quad (2.57)$$

Now because the quality factor is defined as

$$Q \equiv 2\pi \frac{\text{stored energy}}{\text{energy dissipated in a cycle}} = 2\pi \frac{\omega^2 a^2 \iiint \rho |\xi|^2 dV}{4\pi R_{\gamma_+}^2 \oint v \delta p dt}. \quad (2.58)$$

The complex exponential of the temperature perturbation term is

$$\frac{e^{i\omega t}}{i} - \frac{e^{i\omega t}}{\omega \tau_{rad}} = - \left(\sin(\omega t) + \frac{1}{\omega \tau_{rad}} \cos(\omega t) \right). \quad (2.59)$$

Using the harmonic addition theorem this can be rewritten as a sinusoid with a coefficient and a phase. Again using the fact that $\frac{1}{\omega \tau_{rad}} \ll 1$, we can solve the integral over the period to be

$$\oint v \delta p dt = \frac{k_B}{\mu} \frac{v^2}{H c_p \omega} c \oint \cos(\omega t) \sin(\omega t + \phi) dt \approx \boxed{\frac{\pi k_B v^2 p_0}{\mu H c_p \omega^3 \tau_{rad}}}. \quad (2.60)$$

Now computing Q to an order-of-magnitude, and noting $\iiint \rho dV = M_{\gamma_+}$ and thus taking $\iiint \rho |\xi|^2 dV \sim M_{\gamma_+}/10$ as an order of magnitude approximation based on the behavior of the eigenfunctions, we can write

$$Q \sim \frac{\mu H c_p \omega^3 \tau_{rad} M_{\gamma_+}}{20\pi k_B p_0 R_{\gamma_+}^2} \sim \boxed{10^7 \left(\frac{\omega}{10^{-3} \text{s}^{-1}} \right)^3}. \quad (2.61)$$

This is an upper bound for Q , and only correct to an order of magnitude. Since it's the best to go on, we will use $Q \sim 10^7$ throughout this work. This estimate is an upper bound on dissipation efficiency in the limit of high order p-modes, where the local wave number is not too large compared to the local scale height, and the mode is highly compressible. Dissipation could be considerably less efficient for lower order modes that do not obey these assumptions.

2.4.3 High-frequency modes: propagation through the stratosphere

For modes of frequency above the acoustic cutoff frequency, approximated as

$$\omega_a = \frac{c_s}{2H} \quad (2.62)$$

for an isothermal atmosphere, the modes behave differently. For Jupiter, this corresponds to about 3mHz [Mosser, 1995][Gaulme et al., 2015]. Instead of being trapped in Jupiter's interior, with an evanescent tail in the stratosphere, modes above this cutoff frequency propagate into the atmosphere, and eventually into space, unhindered. In this case, the full power of the waves propagating into the stratosphere is lost, not just the part out of quadrature. The energy density of the waves is given by

$$\frac{dE}{dV} \sim \frac{1}{2}\rho v^2 = \frac{1}{4}\rho\omega^2\xi_r^2, \quad (2.63)$$

where the additional factor of 1/2 comes from averaging square velocity over a period (since ξ_r is an amplitude). These are acoustic modes, so they propagate at the sound speed $c_s = \sqrt{\frac{\gamma k_B T}{\mu}}$. So the energy flux through a unit area is given by

$$\mathcal{F} \sim \frac{1}{4}\rho\omega^2\xi_r^2 c_s. \quad (2.64)$$

The total average power loss then is just $\langle \dot{E} \rangle = 4\pi R_{\frac{1}{4}}^2 \mathcal{F}$. Relating this to Q ,

$$Q \equiv 2\pi \frac{E_{\text{stored}}}{\oint \langle \dot{E} \rangle dt}. \quad (2.65)$$

By definition, $\oint \langle \dot{E} \rangle dt = \frac{2\pi}{\omega} \langle \dot{E} \rangle$ so

$$Q = \omega \frac{\iiint \rho |\xi|^2 dV}{\pi R_{\frac{1}{4}}^2 \rho \xi_r^2 c_s}. \quad (2.66)$$

Substituting approximate values gives

$$Q \sim 6 \times 10^3 \frac{\omega}{10^{-3} \text{s}^{-1}}. \quad (2.67)$$

We will not actually use this value of Q , but we do this calculation to demonstrate that we should not expect any modes with frequencies above the cutoff frequency to have significant amplitudes relative to modes below the cutoff frequency.

2.4.4 Ohmic Dissipation by normal modes

From the induction equation

$$\frac{\partial \mathbf{b}}{\partial t} = -\nabla \times (\lambda \nabla \times \mathbf{b}) + \nabla \times (\mathbf{u} \times \mathbf{B}), \quad (2.68)$$

where \mathbf{b} is the induced field resulting from the action of the normal mode velocity \mathbf{u} acting on the main planetary field \mathbf{B} . The magnetic diffusivity is λ , whose value is small (a metal) deep down but large (a semi-conductor) as one approaches the surface. Evidently

$$|\mathbf{b}| \equiv b \sim \frac{kuB}{i\omega + \lambda k^2}, \quad (2.69)$$

where k is the characteristic wave vector describing the spatial variation of \mathbf{b} . The Ohmic dissipation per unit volume is $\frac{\lambda(\nabla \times \mathbf{b})^2}{\mu_0}$ and scales as $1/\lambda$ at large λ but as λ at small λ . The peak dissipation occurs in the region where $\omega \sim \lambda k^2$. Dividing kinetic energy of the wave by the dissipation per wave period, we see that

$$Q_{\text{Ohmic}} \sim 10 \left(\frac{\omega}{V_A k} \right)^2, \quad (2.70)$$

where V_A is the Alfvén velocity, $\sqrt{B^2/\rho\mu_0}$. The coefficient allows for the fact that the volume of dissipation is much smaller than the entire planet and may be an

underestimate depending on the conductivity profile. This predicts $Q > 10^{10}$ for Jupiter, so we do not expect it to be the dominant dissipation mechanism.

2.4.5 Normal mode dissipation in the core

An alternative tidal dissipation mechanism, suggested long ago [Dermott, 1979] assumes that Q is dominated by the small central core, which dissipates in much the same way as a solid terrestrial planet, but possibly aided by soft rheology [Storch and Lai, 2015] or partial melting. In this picture, the intrinsic Q of the core is low but the Q of the planet as a whole is higher by several orders of magnitude, simply because of the quadratic dependence of tidal potential on the radius and the smallness of the volume involved. For modes of spherical order greater than zero, the core is also expected to be below the lower turning point, where the amplitudes are substantially lower, further reducing its importance. If core dissipation is the correct interpretation of tidal Q for Jupiter then it probably implies a similar, “low” Q (relative to our suggested value) for normal modes, but only for those that have significant amplitude in or near the core. This will not apply to current observations of large n (see Figure 2.3). We cannot exclude this but note that it increases the difficulty of explaining the observed normal mode amplitudes.

2.5 Possible physical excitation sources

This section focuses on possible real excitation sources for Jupiter’s seismic normal modes. Each of these will be modeled crudely. The intent here is not to provide highly accurate detailed descriptions of these excitation mechanisms, but rather to simply test if the general energy scales, timescales, and coupling efficiency expected of them could feasibly be candidates to explain the observed signal.

2.5.1 Turbulent convection

Following the work of [Kumar, 1996], we write the the equation of continuity

$$\rho' + \nabla \cdot (\rho \boldsymbol{\xi}) = 0 \quad (2.71)$$

and the acoustic wave equation with a source term

$$\frac{\partial^2 \rho \xi_i}{\partial t^2} + c^2 \frac{\partial \rho'}{\partial x_i} = \frac{\partial T_{ij}}{\partial x_j}, \quad (2.72)$$

where

$$T_{ij} \equiv \rho v_i v_j + p \delta_{ij} - \rho c^2 \delta_{ij}. \quad (2.73)$$

Combining these equations yields the relationship

$$\frac{\partial^2 \rho \xi_i}{\partial t^2} - c^2 \nabla^2 (\rho \xi_i) = -\frac{\partial T_{ij}}{\partial x_j}. \quad (2.74)$$

Decomposing displacement into eigenfunctions

$$\boldsymbol{\xi} = \sum_{nlm} a_{nlm} \boldsymbol{\xi}_{nlm} \exp(-i\omega t), \quad (2.75)$$

where the amplitudes here are normalized to unit energy according to

$$\omega^2 \int \rho |\boldsymbol{\xi}|^2 dV = 1. \quad (2.76)$$

Solving produces

$$\begin{aligned} \frac{\partial a_{nlm}}{\partial t} &= \frac{-i\omega}{\sqrt{2}} \exp(i\omega t) \int \xi_{q_i} \frac{\partial T_{ij}}{\partial x_j} \\ &= \frac{i\omega}{\sqrt{2}} \exp(i\omega t) \int \frac{\partial \xi_{nlm_i}}{\partial x_j} T_{ij} dV. \end{aligned} \quad (2.77)$$

Following the form of turbulent forcing [Landau and Lifshitz, 1959], $T_{ij} \sim \rho v^2 \delta_{ij}$, we can solve

$$\frac{\partial A_q}{\partial t} \sim \frac{i\omega}{\sqrt{2}} \exp(i\omega t) \int \rho v^2 \frac{\partial \xi_{nlm_r}}{\partial r} dV. \quad (2.78)$$

So the energy input into the mode (n, l, m) follows the time average amplitude squared

$$\boxed{\frac{dE_{nlm}}{dt} \sim 2\pi\omega^2 \int r^2 \rho^2 v_\omega^3 h_\omega^4 \left[\frac{\partial \xi_{nlm_r}}{\partial r} \right]^2 dr}, \quad (2.79)$$

where h_ω and v_ω are the turbulent eddies which are resonant with the mode, i.e., they satisfy

$$\frac{h_\omega}{v_\omega} = \frac{2\pi}{\omega}. \quad (2.80)$$

Assuming a Kolomogorov cascade which obeys

$$v_h = v_H \left(\frac{h}{H} \right)^{1/3}, \quad (2.81)$$

we have everything needed to solve for the energy input once we solve for H and v_H . From mixing length theory, we use the planetary length scale for H , and we know the convective velocity associated with the large scale motion approximately obeys [Stevenson, 2021]

$$v_H \sim 0.1 \left[\frac{LF_{\text{conv}}}{\rho H_T} \right]^{1/3}, \quad (2.82)$$

where H_T is the temperature scale height. Solving this to an order of magnitude assuming Jupiter's entire flux is available for convective flux, using Jupiter's average density and assuming $L/H_T \sim 10$, we obtain $v_H \sim 3 \text{cm s}^{-1}$. Solving for h_ω and v_ω give

$$h_\omega \sim 140 \text{cm} \left(\frac{10^{-3} \text{s}^{-1}}{\omega} \right)^{3/2}, \quad (2.83)$$

$$v_\omega \sim 0.03 \text{cm s}^{-1} \left(\frac{10^{-3} \text{s}^{-1}}{\omega} \right)^{1/2}. \quad (2.84)$$

The Reynold's number for these values is of order $10^2 - 10^3$, so it should still be above the minimum Kolomogorov microscale. Such excitation is tremendously inefficient, creating surface velocity amplitudes $< 10^{-5}$ cm/s, with more power in f-modes and low-order p-modes than in higher frequency modes. It is worth noting that the expected convective velocities increase near the surface, as density rapidly decreases but heat flux remains relatively constant. This can increase convective velocities by two orders of magnitude over a small distance, which can affect the resultant energy input. Indeed, this is the most important contribution to solar p-mode energies. We then repeat the above calculation, but change the scaling of h_ω and v_ω to have coefficients of 20km and 3m/s, respectively. We then calculate the energy input integral only over the top scale height, rather than through the whole interior. We find repeating the calculation in this way does not result in substantially larger mode energies. Therefore we discount stochastic excitation by turbulent convection as a dominant excitation source for the observed modes on Jupiter. These modes may excite the observed f-modes on Saturn if dissipation is significantly less efficient for f-modes than it is for p-modes, as we argued could be the case in Section 2.4.2. We note that the situation is actually considerably worse than our back-of-the-envelope estimate in Section 2.1—not only is Jupiter a million times less luminous than the sun, but an apples-to-apples excitation comparison shows Jupiter is also tremendously less efficient at using its internal energy to excite normal modes with turbulent convection. Therefore, the modes observed on Jupiter must be excited by a fundamentally different mechanism.

2.5.2 Meteor strikes

As much of this paper has, the idea of a meteor strike's excitation will closely follow the work of Dombard and Boughn [Dombard and Boughn, 1995] for the Shoemaker-Levy/9 Jovian cometary impact. Here the primary excitation source is a monopolar

explosion, which occurs after the meteor reaches a certain pressure depth. Since the explosion happens very quickly, we can approximate it as a δ -function so that $\phi(t) \rightarrow \delta(t)$. Assuming the comet explodes at the 50 bar level, and taking the energy of the explosion to be 10^{30} ergs (an optimistic estimate; this corresponds to an upper bound on extremely large impacts like SL9 [Dombard and Boughn, 1995] and should be treated as an upper bound), and assuming an impact of this magnitude happens approximately every 50 years, we get negligible equilibrium amplitudes on the order of microns per second. If we use smaller impact energies, the excitation is correspondingly smaller. We did not bother to include smaller, more frequent impacts in this calculation because as argued above only the most energetic events significantly affect the equilibrium amplitudes. We note that an unlikely extraordinarily energetic impact occurring within the ringdown timescale of the planets could in principle excite larger amplitudes, a hypothesis investigated by [Wu and Lithwick, 2019] for Saturn.

2.5.3 Storms

As all models in this chapter, the formulation for storm models will be greatly simplified. The types of storms we are interested for these purposes form when a parcel of moist air is lifted to the level of free convection (LFC) by some external driving force. Once there, some moisture precipitates out of the parcel, releasing latent heat. This heat causes the parcel to warm and expand, which causes it to become buoyant and rise. As it rises and expands, the parcel cools, allowing more condensation and releasing more latent heat. As this moist parcel rises, it will follow a moist adiabat, causing it to be warmer than the surrounding environment at all levels above the LFC. The parcel will continue to rise until it equilibrates with its surroundings. On Earth, this happens at the inversion layer, or the tropopause. This same basic picture applies to water storms on Jupiter [Stoker, 1986], with the important difference that on Earth water vapor is less dense than the ambient air, while the opposite is true

on Jupiter. To model how such a process would affect the surrounding atmosphere, we consider the relevant forces. As the parcel rises, it pulls air along with it. The characteristic force is the buoyancy of the parcel, so

$$f_0 \sim \Delta\rho gV, \quad (2.85)$$

where V is the volume of the parcel and $\Delta\rho$ is the change in density resulting from the release of latent heat, i.e.,

$$\frac{\Delta\rho}{\rho} = \frac{L_v f}{c_p T}, \quad (2.86)$$

where f is the mass fraction of the condensing constituent and L_v is the latent heat of vaporization. The distance over which this dipole acts would scale with the distance the parcel rises. substituting these values into the equation for dipole forcing, we obtain

$$F_0 \sim \frac{E_s \frac{\partial \xi_r}{\partial r} |_{r=r_0}}{\iiint \rho |\xi|^2 dV}, \quad (2.87)$$

where r_0 is the height of the cloud deck. Now to calculate the appropriate storm energy that couples to the mode. If a rising column of air like this were to originate deep within the atmosphere, it could in principle rise all the way to the stratosphere. However, if it started many order-of-magnitude higher in pressure, the parcel itself would probably break apart and lose its coherence after about a scale height. Alternatively, it could keep rising until it hit a cloud deck above it, providing the lifting needed to lift the parcel in front of it above the LFC, while the droplets that condensed down below have already rained out. The dynamics of how such a situation would proceed are complex and uncertain. We therefore assume that the height the parcel will rise scales with the environmental scale height $\epsilon \propto H$.

The column of rising air will have some characteristic radius r and some height H . A thin parcel of rising air would then have volume $\pi r^2 dz$, implying a buoyant force of

$\pi r^2 \Delta \rho g dz$. Each parcel of rising air starts at the cloud deck, and rises a characteristic distance H . Therefore the work done by each parcel is approximately $\pi r^2 H \Delta \rho g dz$. Now integrating over the height of the column, we find the characteristic storm energy from Equation 2.87 to be about

$$E_s \simeq \pi r^2 H^2 \Delta \rho g. \quad (2.88)$$

The power output by water storms in Jupiter is about 3.3 W m^{-2} [Gierasch, 2000], which is a significant fraction of Jupiter's total heat budget. The characteristic size of convective columns can be large, on the order of 100km or more. If this is the case, the effect of entrainment on column buoyancy is negligible [Stoker, 1986]. When a convective plume rises, it does so by releasing latent heat. The total latent heat released by this process is approximately the total mass of condensate in the column

$$E_L \sim \pi r^2 H \rho f L_v, \quad (2.89)$$

where r is the radius of the convective column. The characteristic timescale between such a column rising, then, is just this energy scale divided by the total power output by storms over the whole of Jupiter's surface. This gives us $E_L \sim 1.3 \times 10^{26} \text{ erg} \implies \tau_s \sim 65 \text{ s}$, and $E_s \sim 3.6 \times 10^{25} \text{ erg}$ if the height of the column is 50km [Stoker, 1986]. This is compatible with our expectations about observed storm activity on Jupiter. We note the duration should scale with the buoyancy timescale (see Section 2.3.5)

$$\Delta t \sim \frac{H}{v}, \quad (2.90)$$

where

$$v^2 \approx \frac{L_v f}{c_p T} r g. \quad (2.91)$$

Following through with the calculation and assuming $Q \sim 10^7$, we obtain the ex-

pected normal mode velocity spectrum in Figure 2.6.

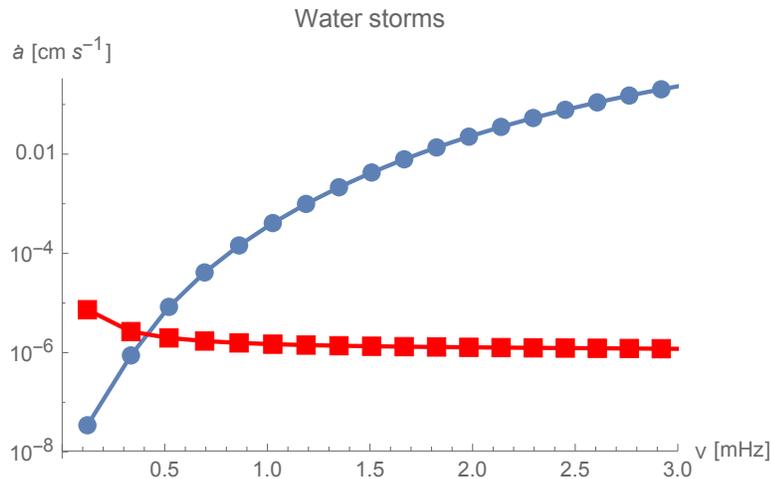


Figure 2.6: Amplitude excitation based on estimates for water storm forcing (blue curve). For comparison, the red curve shows the expected amplitude spectrum from stochastic excitation from turbulent convection.

Clearly, the amplitudes are orders of magnitude too small to explain the SYMPA data. However, the behavior is qualitatively different from the result of turbulent convection; whereas turbulent convection is expected to deposit most energy in low-order modes, storm excitation expects more energy in higher order modes. This is an important distinction, and these two broad classes of excitation sources can be compared as data at lower frequencies becomes available.

However, we have not solved the problem of exciting larger amplitudes than would be expected from turbulent convection. Thermodynamically we expect there to be more cloud levels deeper in Jupiter’s interior. Detailed calculations about the behavior of chemical equilibria and condensation in Jupiter’s shallow interior have been carried out by Fegley and Lodders [Fegley and Lodders, 1994], including the posited existence of rock clouds. Silicate and iron clouds have been observed on brown dwarfs and posited on hot exoplanets [Marley and Ackerman, 1999], and there has even been some modeling of their storm dynamics [Lunine et al., 1989]. Similar dynamics may

well be at play in Jupiter. These comparatively refractory species will have much higher latent heats, and can thus be expected to be more energetic than water storms. If this were the case, we could follow through the same analysis but assume the length scales H and r used to calculate E_s and E_L are proportional to the relative pressure scale heights between the water cloud deck and the rock cloud deck. We also substitute the latent heat of vaporization of water ($2.3 \times 10^{10} \text{erg g}^{-1}$) with the appropriate value for silica ($1.2 \times 10^{11} \text{erg g}^{-1}$). Rock storms must occur deeper in the atmosphere, where pressure, temperature, and density are higher. We will use parameters at 10kbar in pressure at around 2000K, roughly where we expect silane gas to start producing silica droplets. A visualization of this difference is illustrated in Figure 2.7.

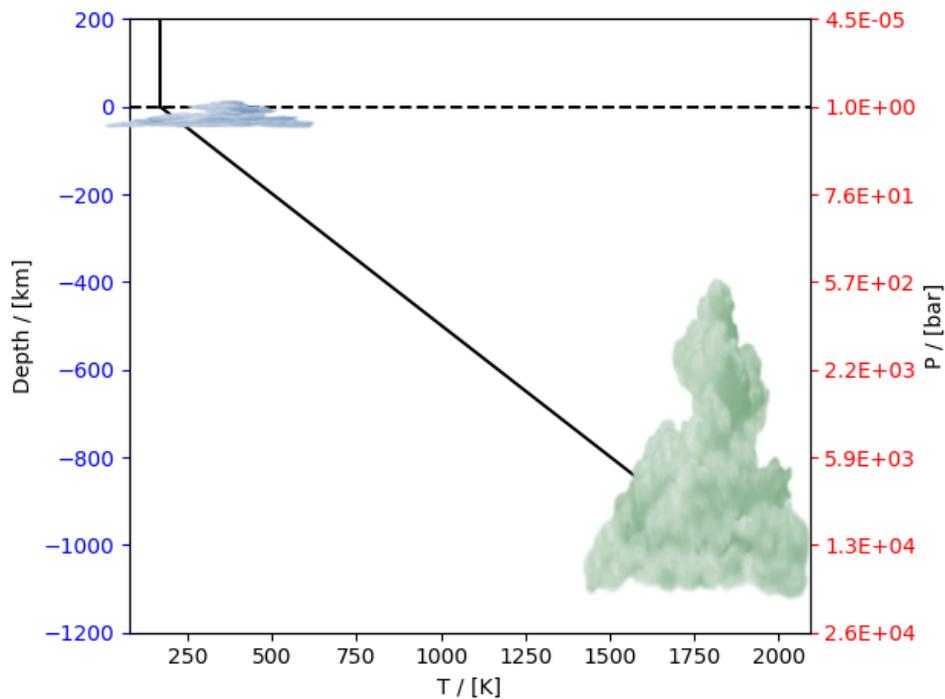


Figure 2.7: A cartoon depicting the relative dimensions of water and rock storms. As one dives into the interior, the scale height increases rapidly, which is important for our estimates of storm length scales at these depths. The left y-axis shows depth while the right y-axis shows corresponding pressure. The blue cloud represents the height and location of water storms, while the green cloud represents these same parameters for rock storms.

This different depth affects the coupling efficiency for higher frequency modes. This is one of several factors which are ignored in Figure 2.9. The justification for using the latent heat of a silica phase transition as a stand-in for silicate droplet condensation is not immediately obvious, since based on thermodynamic equilibrium chemistry we expect this transition to be a complicated multi-component chemical reaction of silane, iron-carrying vapor, magnesium-carrying vapor, and water vapor to form silicate droplets. The dynamics of how such reactions would unfold need future inspection to complete a detailed picture, but for our purposes we are not overly concerned with the details, only the order-of-magnitude energy scales. If we assume the dominant reaction is e.g., silane to silica instead of a silica vapor to liquid phase transition, it affects the outcome by less than 30%, which is negligible in the context of our order-of-magnitude consideration. Therefore we take a silica phase transition to be a proxy for potentially complicated chemical reactions, noting that the important aspect is the release of heat, not the specific mechanism which causes it. As such, we combine the total abundances of silicon, magnesium, and iron and take this to be the concentration of silica vapor, in order to simplify the model. Finally, we assume that the available energy budget for rock storms is the same as for water storms relative to Jupiter’s luminosity. Using these parameters and allowing the storm column radius to grow, one can justify using parameters like $E_s \sim 5 \times 10^{31} \implies \tau_s \sim 1.5 \times 10^7$. Using these parameters, coupling to five kilobar level (the midpoint of the storm on Figure 2.7), the same model produces Figure 2.8.

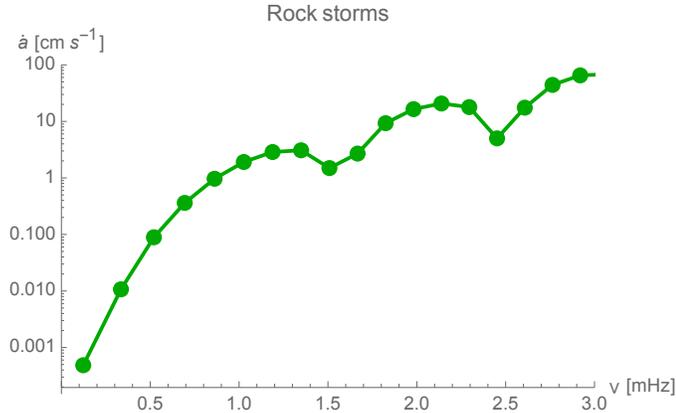


Figure 2.8: Amplitude excitation based on preliminary estimates for rock cloud forcing. The non-smooth structure results from proximity to nodes in the relevant eigenfunctions at the silicate cloud level, relevant for high frequency overtones at these non-negligible depths. The specific structure of the curve shouldn't be taken too seriously; the point is the order of magnitude of the velocities which begin to approach the observed values on order of tens of cm/s.

One could easily argue that these parameters are all highly uncertain, and that this is an issue of fine tuning. After all, we can adjust the storm parameters to yield any order of magnitude equilibrium mode amplitude we like, in principle. But the important point here is not to make an accurate prediction of the behavior of these hypothetical rock storms, whose existence and behavior is largely unconstrained. Instead, since we know nothing about rock storms, this analysis is intended to place constraints on the necessary parameters of storm-like activity which could produce the observed equilibrium amplitudes. The details of the dynamics of a hypothetical rock storm are highly speculative. In this chapter we assumed the dynamics were identical to water storms, and just scaled the parameters to their appropriate values accordingly. This exercise serves simply to demonstrate an example of a physically plausible mechanism which could excite the observed amplitudes.

2.6 Results and discussion

No excitation mechanism investigated here seems to be a clear candidate for producing the observed amplitudes of Jovian seismic modes. However, if we are to believe the results, we can place meaningful constraints on the type of source that may cause these observations, and make some predictions about other frequencies based on this.

2.6.1 Excitation source parameter constraints

The expected turbulent convection is insufficient to explain the observed amplitudes of normal modes. Point source excitations, either storms, meteor strikes, or something else, may be able to explain these amplitudes if analyzed more carefully. Both monopole and dipole excitation types are of the same form, to first order.

$$a_{eq} \sim \frac{E_s \frac{\partial \xi_r}{\partial r} |_{r=r_0}}{\iiint \rho |\xi|^2 dV} \left(\frac{2Q}{\tau_s \omega^3} \right)^{1/2} \xi_r(R). \quad (2.92)$$

Using this general form, one can place order-of-magnitude constraints on the necessary bulk parameters needed to excite the amplitudes observed by SYMPA.

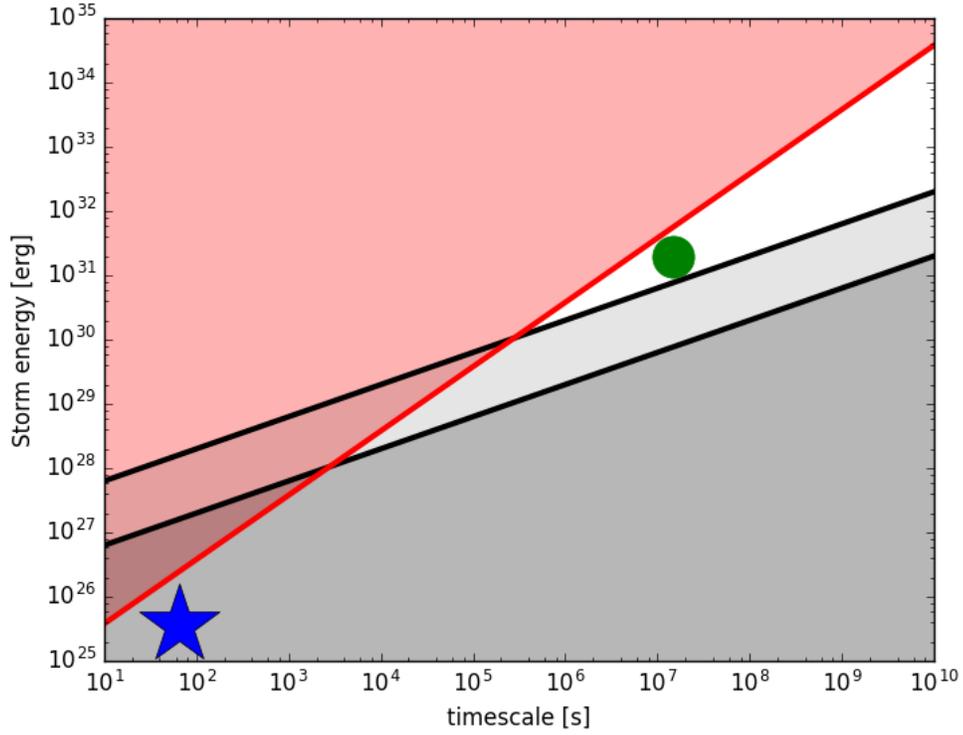


Figure 2.9: Assuming a storm-like excitation and holding all other parameters constant, any viable candidate must lie above the black curve in order to explain the Gaulme et. al. results [Gaulme et al., 2011], and below the red curve to satisfy Jupiter’s luminosity constraint. The two black curves represent different values of Q . The lowest line represents an idealistic $Q = 10^8$, above that a more pessimistic $Q = 10^6$. The blue star represents the excitation from water storms in this parameter space. The green point represents the same model scaled to rock clouds.

Any such mechanism must not violate Jupiter’s total energy budget, but must be energetic and frequent enough to excite modes of the observed amplitude in the steady state. There is a sliver of parameter space as shown in Figure 2.9 that could theoretically satisfy these constraints.

2.6.2 Predictions for other frequencies

Using the storm or meteor strike model, or any generic short-lived, localized, stochastic excitation source, we obtain some general features of the power spectrum. In

particular, low frequencies generated in this way are orders of magnitude smaller than their overtones, since the local gradient of the radial eigenfunction near the surface is much smaller for lower frequencies, and the coupling is therefore weaker. In contrast, the red curve on Figure 2.6 shows more power in lower frequency modes compared to overtones. Future observations which show the power spectrum with better resolution, and in lower frequencies could distinguish between these two basic classes of excitation: global or point source.

2.6.3 Implications for gravity, Juno, Saturn, and ice giants

Because no unique candidate for excitation has been determined, it's difficult to make predictions for how this may affect Juno's results. If the excitation sources are point sources of the sort described in this work, the amplitudes for f-modes, which would most significantly perturb Jupiter's gravity field, would be orders of magnitude smaller than the overtones detected by SYMPA. This means that even though the displacement amplitude of normal mode overtones may be on the order of fifty meters, the fundamental modes could self-consistently have displacement amplitudes of mere centimeters. The gravity field perturbation caused by the normal modes is still strongest for the lowest frequency modes, since the global coherence of zeroth radial order modes as shown in Figure 2.3 makes them perturb the gravity field much more strongly than oscillatory, higher order modes.

We can decompose the gravity field into a sum of gravity harmonics

$$\Phi(r, \theta, \phi) = \frac{1}{R} \sum_{l=0}^{\infty} \sum_{m=0}^l \left(\frac{R}{r}\right)^{l+1} (C_{lm} \cos(m\phi) + S_{lm} \sin(m\phi) P_l^m(\cos \theta)). \quad (2.93)$$

Because both gravity harmonics and normal modes are defined by spherical harmonics, a given normal mode's gravity perturbation can be completely described by a

single gravity harmonic term. If we wish to ask whether a given normal mode will be detectable, we can compute an illustrative example by considering how J_2 is affected by ξ_{n20} . To calculate this change, we must compute the density perturbation $\delta\rho_{nlm}$ from a displacement eigenfunction ξ_{nlm} . We can do this simply by using the continuity equation

$$\delta\rho = \nabla \cdot (\rho\xi). \quad (2.94)$$

The shape of these density eigenfunctions are shown on Figure 2.10.

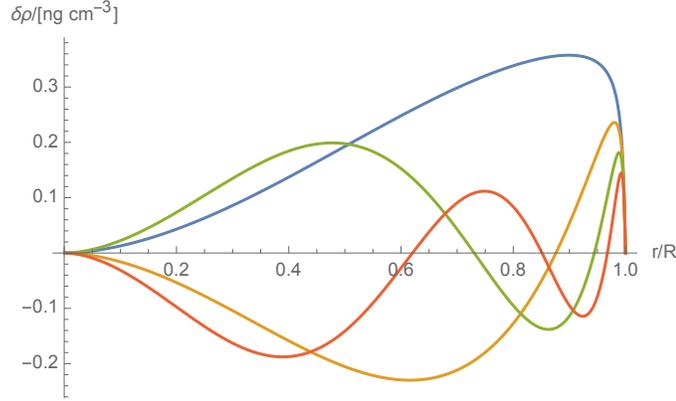


Figure 2.10: Normalized density eigenfunctions for the first few $l=2$ modes. Notice that the $n=1$ density eigenfunction has no nodes, even though its corresponding displacement eigenfunction has one. This is a simple consequence of Equation 2.94, since the density is the divergence of the displacement.

To calculate the change in J_l associated with mode ξ_{nl0} , we use

$$J_l = -\frac{1}{MR^l} \int r'^l P_l(\cos \theta') \delta\rho(\mathbf{r}') d^3\mathbf{r}'. \quad (2.95)$$

Juno's ΔJ_2 3σ uncertainty for gravity perturbations is about $10^{-8}\Phi_{\mathcal{J}}$ [Bolton et al., 2017], so we can compute the required amplitudes for gravitational detection of normal modes by Juno. This is shown in Figure 2.11.

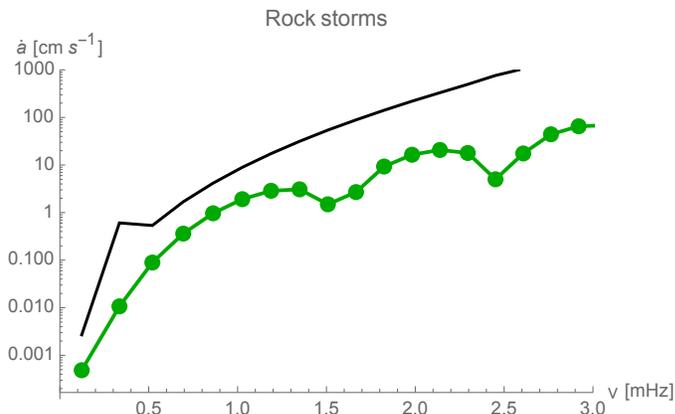


Figure 2.11: The black curve represents the 3 sigma sensitivity limit for Juno detecting a variation in J_2 , and the green curve is identical to Figure 2.8.

Evidently under the assumptions of our model, detection of some normal modes from Juno gravity is plausible. However it's right on the edge, and since our results are very imprecise, detection or lack of detection are both plausible outcomes.

Identical calculations to the ones carried out for Jupiter can be replicated for any planetary model by simply changing input parameters. In addition to Jupiter, we have carried out these calculations for Saturn. Kronoseismology has developed in a different trajectory from dioseismology, since the seismometers employed for Saturn are the rings themselves. Kronoseismology is therefore most sensitive to modes which can resonate with the orbits of ring particles. Because there is a gap between the surface of Saturn and the C-ring, only the lowest frequency modes can be detected this way. In contrast, dioseismology is performed using time series Doppler imaging, which is most sensitive to the largest velocities and shorter periods, i.e., overtones. Jupiter and Saturn are very similar planets, with similar compositions, radii, and heat budgets. It is therefore probable that they each behave much more like each other than like stars. Turbulent convection as a source of normal mode excitation suffers the same deficiency on Saturn as it does on Jupiter; small convective velocities. Convective velocities are on the order of 3cm s^{-1} for both, much smaller than the sound speed in both cases. This would indicate a power spectrum comparable to the

red curve on Figure 2.6. We must ensure that our storm excitation mechanism, which was used to explain large mode amplitudes on Jupiter, does not produce excessively large amplitudes on Saturn compared to estimates for Saturn’s mixed f and g-modes [Fuller, 2014]. In particular, we can compute the mode excitation by observed storms expected based on our model. Based on the arguments leading up to Equation 2.29, the water storms on Saturn may be much more important for mode excitation than the water storms of Jupiter. While Jupiter has continuous thunderstorms happening all over its surface, Saturn has just one hugely energetic storm every few decades [Li and Ingersoll, 2015]. The most recent Great Storm on Saturn occurred in 2011, and was observed by Cassini, ground based telescopes, and amateur astronomers. Similar Great Storms have been seen throughout Saturn’s history, occurring on a characteristic timescale of roughly 30 years. As demonstrated, this type of excitation (infrequent, large energy) is the most favorable situation to produce high amplitude normal modes. The great storm on Saturn releases as much energy as the whole of Saturn does in a year [Fischer et al., 2011]. Assuming $E_s/E_L \sim 10\%$, as is the case for water storms on Jupiter, this provides an approximation for $E_s \sim 4 \times 10^{30}$ ergs. We know events like these occur roughly every 30 years, which directly provides the relevant τ_s . We can do a similar analysis to the one applied to Jupiter, but apply parameters relevant to the Saturnian Great Storms and scale our calculated dissipation due to radiative damping to Saturn. This produces a value of $Q \sim 5 \times 10^6$ which is consistent with (although much larger than) the observational lower bound of $Q > 10^4$ [Hedman and Nicholson, 2013]. Using these inputs we obtain Figure 2.12.

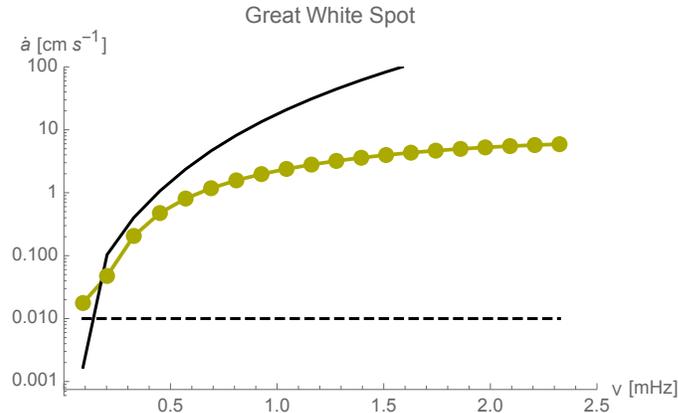


Figure 2.12: Saturn velocity amplitudes based on estimates for the Great White Spot 30 year quasi-periodic super storm. The yellow curve represents the expected amplitudes, while the black curve represents the detection limit for Cassini gravity, and the dashed gray line corresponds roughly to Jim Fuller’s prediction for f-mode amplitudes on Saturn based on inspection of optical depth variations in the spiral density waves in Saturn’s rings raised by its normal modes [Fuller, 2014].

We can use Figure 2.12 to compare our predictions to expected measurements. This calculation did not include dissipation from the core, which could be more important on Saturn than on Jupiter since Saturn’s core is known to be relatively large. This indicates that for the lowest order modes, storm activity may be comparable in importance to turbulent convection, and that for higher frequency overtones Saturn may have comparable normal mode amplitudes to Jupiter. Importantly, the storm excites relatively small amplitudes for Saturn’s low-order modes. If those excitation predictions were too large, it would be evidence against our storm excitation model, since it would be inconsistent with observations. Additionally, rock clouds may also play a role in Saturn as they do in Jupiter. However, our analysis suggests the Great White Spot alone could theoretically produce p-mode amplitudes on Saturn of the same order as have been observed on Jupiter, an interesting result on its own. For this reason we will refrain from further speculation about additional excitation sources. Doppler imaging of Saturn may take additional technical advances or dedicated time on larger telescopes, because the light from Saturn that reaches Earth is significantly fainter than that of Jupiter. As with Jupiter, it is unclear whether a gravity sig-

nal from the normal modes can be expected. Certainly additional excitation from rock storms on Saturn could put it over the edge. However, stochastic excitation from turbulent convection as we have calculated it certainly cannot produce normal mode amplitudes large enough to produce a gravity signal [Marley, 1991] [Marley and Porco, 1993]. Therefore if one wishes to invoke normal modes as the explanation for the unexplained component of Saturn’s gravity field measured by Cassini [Iess et al., 2019], one must consider storms or some other excitation source.

In addition to the gas giants, ice giants may prove to be of similar interest for performing planetary seismology from orbit [Elliot et al., 2017]. Three of four multi-billion dollar proposals for missions to either Uranus or Neptune in the coming decades include a doppler imager, which would ideally be capable of detecting seismic normal modes. Attempts have been made to measure poseidoseismology (seismology on Neptune) using Kepler K2, although only the reflection of solar oscillations were detected [Gaulme et al., 2016]. Unfortunately, it is difficult to put constraints on what amplitudes to expect without a coherent understanding of the excitation source or an a priori knowledge of the planetary interior. Indeed, complicated interactions between the atmosphere and the mantle of the ice giants, immense uncertainty about interior dynamics, general ignorance of the ice giants’ bulk interior structure including possible dissipation mechanisms, and universal uncertainty about normal mode excitation theory in giant planets makes constraining the expected normal mode amplitudes exceedingly difficult. Rather than attempting a naive quantitative analysis here, we will simply provide some remarks for future work. Using an approximation of the equation of state from previous studies of the ice giants’ interiors [Helled et al., 2002], we constructed hypothetical eigenfunctions for Uranus and Neptune which, although highly uncertain, provide an order-of-magnitude estimate for the general scale of the inertia of these modes and gradients near the surface. Uranus and Nep-

tune have much smaller energy fluxes than Saturn and Jupiter, even relative to their total masses. Convective velocities should be on the order of 1cm s^{-1} , insufficient to excite amplitudes larger than microns per second. However, methane storms have been observed from Earth on Uranus [Gibbard et al., 2002], so it is possible that this activity could excite higher amplitude normal mode responses. Storm systems observable by telescope are methane storms, but just as rock storms could be at play deeper in Jupiter, water storms could behave similarly deeper in the ice giants. Of course, the eventual amplitude depends strongly on the energy and timescales of the storm, as shown in Figure 2.9. Neptune has a larger luminosity than Uranus, and could therefore in principle produce higher amplitude modes. It is possible of course that solid-phase seismic activity in the mantle could couple very efficiently to the atmosphere to provide higher amplitude responses in the upper atmosphere. A Uranus quake occurring in a solid phase mantle, for example, could couple efficiently to the dense overlying atmosphere and produce a high amplitude signal in the stratosphere. Such a mechanism, however, is beyond the scope of this paper. Indeed, it is very difficult to place theoretical constraints on ice giant seismic mode amplitudes without making an enormous array of assumptions. Since we don't even understand the very basics of ice giant interiors, such assumptions are difficult to defend. A more focused effort to characterize normal mode couplings in the ice giants, as well as an elementary understanding of deep moist convection in gaseous interiors, could provide some basic theoretical predictions for normal mode amplitudes for the ice giants, which would be necessary for calibrating a Doppler imager on board a future mission. Before such a method could be reliably employed, much further study of giant planet seismology must be carried out, both on the observational and theoretical fronts, as well as further study of ice giant and gas giant interior dynamics.

2.7 Conclusion

The observed amplitudes of normal modes on Jupiter are in great excess of what would be expected based on turbulent convective theory. Meteor strikes do not occur frequently enough or with sufficient energy to excite the observed amplitudes either. Water storms are extremely frequent, but relatively low energy and with very weak coupling to the normal modes. Therefore they cannot come anywhere close to explaining the observed modes. The only viable candidate examined in this paper is rock storms. It should be mentioned that there are other possible excitation mechanisms not examined in this work that may warrant further study. For example, baroclinic instabilities may play a role in seismic mode excitation. Additionally, dynamics in the helium rain layer or in a region of deep static stability are potentially worth consideration. If the primary excitation source is rock storms, as suggested here, the specific dynamics of the rock storms could significantly affect the outcome. In particular, the timescale associated with a rock storm's duration, and the length scales associated with such a storm, might differ significantly from the basic simplifying assumptions presented here. However, rock clouds are a promising candidate given the large latent heat of silicates compared to water, as well as the large length scales expected at such a depth with an atmospheric scale height much larger than the upper troposphere. Preliminary crude calculations indicate that any storm mechanism invoked to explain the observed amplitudes must occur below the red curve and at least above the lowest black curve on Figure 2.9. Jupiter may have a rich abundance of storm activity below the visible surface. This work suggests this storm activity could feasibly be responsible for the much larger normal mode amplitudes seen on Jupiter compared to predictions. More sophisticated models of storm activity may show better coupling between storms and normal modes than we estimated here, which could make these storms a candidate to explain Jupiter's normal modes. Similar storms and large scale convection may excite normal modes on the ice giants in a similar fashion, and this

topic warrants further study.

The following link directs to the published version of this chapter:

<https://www.sciencedirect.com/science/article/abs/pii/S0019103517307558>

Chapter 3

Gravitational seismology

First you will come to the Sirens
who enchant all who come near
them. If anyone unwarily draws
in too close and hears the
singing of the Sirens...

Homer

Summary

The goal of this chapter is to build a heuristic for gravitational seismology, wherein seismic signals can be searched for within existing gravity data sets of giant planets. We analyze the residual range rate metadata from Cassini's gravity experiment that cannot be explained with a static, zonally symmetric gravity field. In this chapter we reproduce the data using a simple forward model of gravity perturbations from normal modes. To do this, we stack data from multiple flybys to improve sensitivity. We find a partially degenerate set of normal mode energy spectra that successfully reproduce the unknown gravity signal from Cassini's flybys. The most likely models are dominated by gravitational contributions from p-modes between $500\text{-}700\mu\text{Hz}$, despite the fact that f-modes have a stronger gravity signal for a given amplitude.

This suggests strong frequency dependence in normal mode excitation on Saturn. Favorable models have a narrow distribution around the peak frequency. We predict peak amplitudes for p-modes on the order of several kilometers, at least an order of magnitude larger than the peak amplitudes inferred by Earth-based observations of Jupiter. The large p-mode amplitudes we predict on Saturn, if they are steady state, would imply weak damping with a lower bound of $Q > 10^7$ for these modes, consistent with theoretical predictions.

3.1 Introduction

Gravity field measurements allow us to probe the interior structure of a planet by measuring its deviation from spherical symmetry. For giant planets, the planet's response to its own rotation breaks its spherical symmetry. The deviation away from spherical symmetry depends on the planet's internal density distribution [Stevenson, 2021]. Therefore a detailed mapping of a planet's gravity field can corroborate or refute interior models. Saturn's non-spherical gravity field was first inferred from spacecraft tracking data of Pioneer 11 [Null et al., 1981] [Hubbard et al., 1980], and were later improved using Voyager data [Campbell and Anderson, 1989]. The arrival of the Cassini spacecraft in the Saturnian system yielded more accurate determination of the gravity field of the gas giant by first looking at the orbits of its satellites. Now the Grand Finale of the Cassini mission has produced exquisite gravity field data for Saturn, providing the first concrete constraints for Saturn's ring mass, zonal wind depths, and evidence for internal differential rotation by offering gravity field measurements up to J_{12} [Iess et al., 2019] [Galanti et al., 2019]. But behind these spectacular new findings lurks a dark side: a small component of Saturn's gravity field which cannot be explained with the canonical static, zonally symmetric gravity

field expected of gas giants.

Cassini’s radioscience experiment is carried out by measuring the Doppler shift of a microwave signal in a two-way configuration: the signal is sent from a ground station to the spacecraft, which retransmits it back to the station preserving phase coherency. The Doppler shift is, to first order, proportional to the relative velocity of the spacecraft with respect to the station. These measurements are compared with predictions based on dynamical and observation models to obtain data residuals. The data we used in this study are two-way Doppler residuals, converted in a radial velocity time series, obtained by removing the effect of empirical acceleration from the reference solution given by Iess, et al., 2019.

This additional and unknown source of gravity can be fit with a variety of models. A static tesseral gravity field is possible, but there is no convincing low-order fit [Iess et al., 2019]. A low-order tesseral field does not provide a predictive solution with the available data, and also depends on the assumed rotation rate of Saturn. That is, a given gravity harmonic solution for a subset of flybys will not accurately predict the next flyby and requires additional harmonic terms. The nominal method which was employed for the published gravity harmonic results was an agnostic “empirical acceleration” model which, due to the unknown origin of the source of the additional gravity, included random acceleration vectors which changed on a ten minute timescale. In this context, “random” means that each acceleration vector is allowed to have any direction with an a priori amplitude of $\pm 4 \times 10^{-10}$ km/s [Iess et al., 2019]. They could be correlated (non-random) even when the process used to create them allows for randomness. This timescale between changing acceleration vectors was determined empirically as the longest timescale which can successfully reduce range rate residuals to the noise level.

A time-dependent signal does not necessarily require normal modes. For example, there may be a time-dependent or non-symmetric signature from large scale convection [Kong et al., 2016]. Additionally, Saturn’s envelope is differentially rotating [Galanti et al., 2019] [Chachan and Stevenson, 2019]. If mass anomalies were embedded at different depths or latitudes, then a spacecraft could encounter measurably different quasi-static tesseral gravity fields during each flyby [Iess et al., 2019]. However, differentially rotating tesseral structure in Saturn’s gravity field has been shown to produce structures in the rings [El Moutamid et al., 2017], and the magnitude of the potential perturbation inferred from observation is orders of magnitude too small to explain the anomalous signal. Because of Saturn’s expected internal differential rotation rate (about 5%) [Galanti et al., 2019] [Chachan and Stevenson, 2019], it is unlikely that such structure could measurably affect the spacecraft trajectory without showing clear structure from resonances in the rings.

This work will specifically explore the hypothesis that Saturn’s residual gravity is a consequence of normal mode oscillations. It has already been demonstrated that normal modes are capable of eliminating the range-rate residuals to the noise level [Iess et al., 2019]. This has been done by computationally optimizing for individual mode amplitudes using a large number of free parameters. One possible solution involves only zonal f-modes. This solution, however, is affected by model assumptions such as maximum modeled spherical degree, whether to permit p-modes or g-modes, whether to permit non-zonal normal modes, etc. These uncertainties occur because, when optimizing with a large number of free parameters, there is a risk of over-fitting the data using too complex of a model. These issues are not important in the context of constraining Saturn’s zonal gravity harmonics and ring mass because the uncertainty can simply be absorbed in the error ellipses for these values. However in this

work we revisit the residuals data with a different purpose: to try to extract a preferred normal mode spectrum which is predictive for further flybys, robust to changes in model assumptions, and as simple as possible to capture the qualitative behavior of the spectrum without over-fitting the data. Bearing this in mind, although we find a statistical preference using our simple model for signals dominated by low-order p-modes, readers should remember that our findings are not conclusive proof of such a spectrum on Saturn.

Our investigation has at least two important applications: first, normal modes are themselves a promising method by which to probe the interior structure of giant planets, and this analysis provides some evidence of their power spectrum. Second, any gravitational signal from normal modes above the noise level contaminates spacecraft tracking data and may be aliased into the static model. As we will see, the behavior of the modes is partially degenerate and the solution is non-unique. However, the solutions are clustered in parameter space and predict a high probability of reproducing the observed unexplained gravity signal. The most successful models indicate the signal is likely to be dominated by p-modes between 500 and 700 μ Hz (see Figure 3.4).

In the second section, we outline some fundamentals of giant planet seismology, spacecraft tracking, and our forward model. In the third section we discuss our data reduction method including a novel data stacking technique, as well as error sources, and a fitting procedure. In the fourth section we present the results of our analysis, finding a simple two parameter model that has a high probability of producing a good fit to the spacecraft signal. In the final section we discuss the implications of our findings.

3.2 The forward model

In order to accurately model seismic effects on Cassini’s gravity signal, we must determine the mode’s eigenfrequencies, and the scaling relationship between a mode’s displacement amplitude and its effect on Saturn’s gravity field. These issues are addressed in Section 3.2.1. Next we must model how given gravity potential perturbations affect the spacecraft tracking signal, which is done in Section 3.2.2. Next we need an agnostic parametric model for modal energy spectrum, discussed in Section 3.2.3. Finally we account for the intrinsic stochasticity of the problem; a given mode cannot be modeled deterministically, because we have no way of knowing what the temporal phase of each mode was when Cassini was at periapse. This is partially circumvented with our stacking technique, discussed in Section 3.3 with further technical information in the appendix.

3.2.1 Background

In this paper we approximate Saturn as an adiabatic, spherical, uniformly rotating planet. We neglect rotation to compute the eigenfunctions and potential perturbations, but account for rotation when considering Coriolis force frequency splitting and the rotating gravity potential encountered in an inertial frame. In this case, giant planet oscillations can be decomposed into a discrete set of orthogonal normal modes with quantum numbers (n, l, m) . n corresponds to the number of radial nodes in the displacement eigenfunction, l to the spherical harmonic degree, and $m = [-l..l]$ to the azimuthal degree. Each mode has a unique displacement eigenfunction

$$\xi_{nlm}(\mathbf{r}) = \left(\xi_{r,nlm}(r)\hat{\mathbf{r}} + \xi_{h,nlm}(r)\hat{\theta}\frac{\partial}{\partial\theta} + \frac{\xi_{h,nlm}(r)}{\sin\theta}\hat{\phi}\frac{\partial}{\partial\phi} \right) Y_{lm}(\theta, \phi) \quad (3.1)$$

and a characteristic eigenfrequency ω_{nlm} so that the total displacement as a function of time is $\xi_{nlm} \cos(\omega_{nlm}t + \alpha_{nlm})$. Because ω_{nlm} is not precisely determined, the phase α_{nlm}

cannot be coherently specified between flybys and is assumed random for each mode for each flyby. ξ_r and ξ_h correspond to the radial and horizontal eigenfunctions respectively, which together specify the fluid displacement at any point within the planetary sphere. For our purposes the eigenfunctions were obtained using GYRE stellar oscillation code [Townsend et al., 2013], with an $n=1$ nonrotating polytrope model for Saturn’s interior. Our goal here is independent of accurate interior modeling; we are interested in the relative gravity signal between modes and their order-of-magnitude, which is not strongly sensitive to small changes in the interior model. However, using an adiabatic interior model precludes g-modes, so we account for contributions from g-modes separately.

Because Saturn’s interior structure is not precisely determined, we performed our full analysis on a variety of interior model assumptions to demonstrate that the results are not sensitive to small errors in modal eigenfrequencies. We tested eigenfrequencies produced by this same polytrope model generated with GYRE [Townsend et al., 2013], as well as a sampling computed using a more sophisticated Saturn interior model [Gudkova and Zharkov, 2006]. The nominal model uses the eigenfrequencies from Sa8. In addition, we accounted for mode-splitting due to Coriolis forces [Christensen-Dalsgaard, 2014]. These split according to

$$\delta\omega_{nlm} = m\beta_{nl}\Omega, \quad (3.2)$$

where Ω is Saturn’s spin rate and

$$\beta_{nl} \equiv \frac{\int_0^R (\xi_r^2 + l(l+1)\xi_h^2 - 2\xi_r\xi_h - \xi_h^2)r^2\rho dr}{\int_0^R (\xi_r^2 + l(l+1)\xi_h^2)r^2\rho dr}. \quad (3.3)$$

The nominal frequencies for this paper are plotted in Figure 3.1.

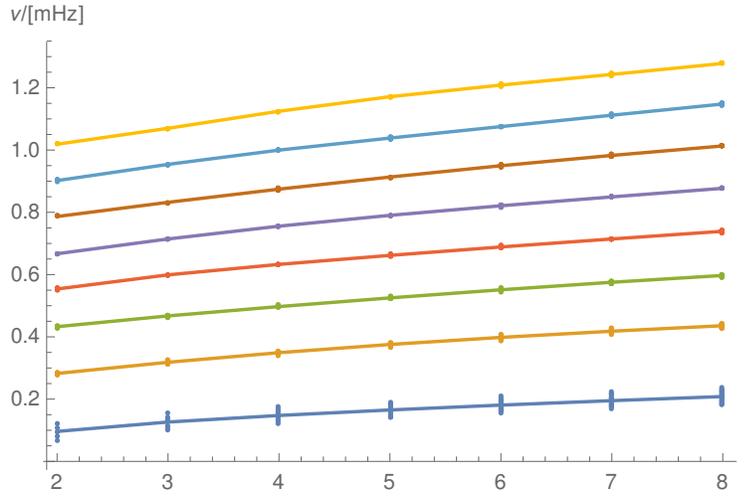


Figure 3.1: Eigenfrequencies including splitting due to the Coriolis force in the rotating frame. Each curve corresponds to rising radial order n for $m = 0$ modes, with l rising along the x-axis. Each eigenfrequency ω_{nlm} is shown as a point, with $m \neq 0$ modes deviating from the $m = 0$ curve. This frequency splitting effect is most important for f-modes.

For computational reasons, we consider a finite subset of modes in our model. In the nominal model we consider f-modes and p-modes up to $l = 8$, $n = 7$. We found equivalent results when using instead $l = 10$, $n = 5$ as bounds on parameter space. We also specially tested f-modes only up to $l = 20$. We do not expect g-modes to dominate the signal for at least two reasons: first, because the stable layer where they resonate is so deep, its effect on the gravity field would be very weak unless its amplitude were extremely large. Second, we do not expect its amplitude to be extremely large, because its eigenfunction is evanescent near the surface where mode excitation is expected to be most efficient. Nevertheless for the sake of completeness we tested g-modes using published eigenfrequencies [Gudkova and Zharkov, 2006].

After choosing eigenfrequencies, we compute the scaling between displacement eigenfunctions and gravity potential perturbations. The gravity field perturbation associated with displacement eigenfunction ξ can be obtained by integrating over the

material sphere and accounting for fluid point displacements according to

$$\delta\Phi = G \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - (\mathbf{r}' + \xi(\mathbf{r}'))|} d^3r' - \frac{GM}{r}. \quad (3.4)$$

For linear perturbations, this is equivalent to the Eulerian density perturbation from continuity $\delta\rho = \nabla \cdot (\rho\xi)$ so that

$$\delta\Phi = G \int \frac{\delta\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3r'. \quad (3.5)$$

Decomposing this potential perturbation into

$$\delta\Phi = \frac{GM}{r} \sum_{n=0}^{\infty} \left(- \sum_{l=2}^{\infty} \sum_{m=-l}^l \left(\frac{R}{r} \right)^l P_l^m(\cos\theta) [\delta C_{nlm} \cos m\phi + \delta S_{nlm} \sin m\phi] \right), \quad (3.6)$$

one can show that the gravity harmonic coefficient perturbation associated with the normal mode is [Stevenson, 2021] [Marley and Porco, 1993]

$$\delta C_{nlm}(t) = \frac{1}{MR^l} \frac{4\pi}{2l+1} \cos(\omega_{nlm}t) \int_0^R r^{l+2} \delta\rho_{nlm,r}(r) dr, \quad (3.7)$$

where $\delta\rho_{nlm}(\mathbf{r}) \equiv \delta\rho_{nlm,r}(r) P_l^m(\theta) \cos(m\phi)$ normalized such that the mode surface displacement is 1cm at the planet surface. With appropriate choice of coordinates, $\delta S_{nlm} \rightarrow 0$. This leads to Figure 3.2 which illustrates why f-modes are a priori favored as sources of gravity perturbations.

Higher order p-modes have nodes in their eigenfunction, leading to destructive interference of the gravitational signature. Therefore in order for p-modes to dominate the signal, they must have more than an order-of-magnitude larger energy than f-modes (two orders of magnitude larger amplitude).

A test particle outside of the planet on a prescribed trajectory $\mathbf{r}(t)$ encounters the

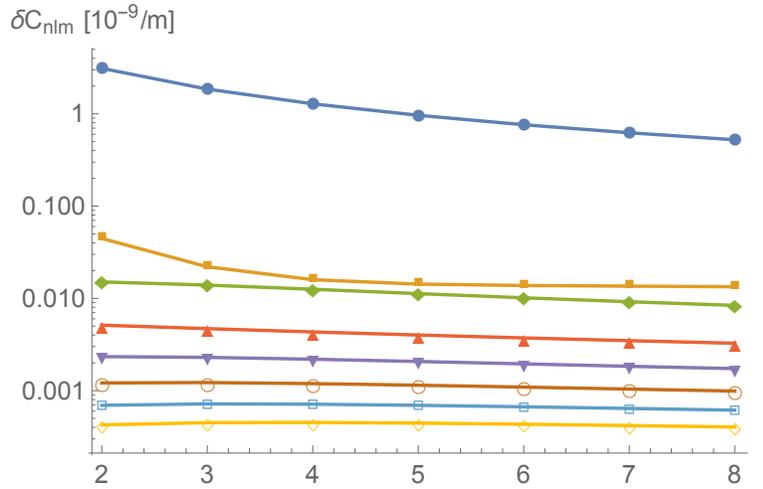


Figure 3.2: Gravity harmonic coefficient perturbations for various modes per meter of surface displacement amplitude. The blue curve (top) represents f-modes, which have the most prominent gravitational signature for a given surface amplitude, while p-modes (below) need larger amplitudes to be detected.

potential perturbation

$$\delta\Phi(t) = \frac{GM}{r} \sum_{nlm} a_{nlm} \delta C_{nlm} \left(\frac{R}{r(t)} \right)^{l+1} P_{lm}(\cos\theta(t)) \cos[m(\phi(t) - \Omega t - \phi_0) - \omega_{nlm}t - \alpha_{nlm}], \quad (3.8)$$

where a_{nlm} is the maximum surface displacement of mode (n, l, m) in cm, and Ω is Saturn's spin rate. Notice that this equation includes two random variables: ϕ_0 the initial longitudinal orientation of the modes with respect to our coordinates, and α_{nlm} is the initial temporal phase of the mode. ϕ_0 is shared between all modes, but is random for each flyby. α_{nlm} is a random variable for each mode and for each flyby. Although the phase difference between flybys can in principle be determined from the mode's eigenfrequency, in practice this is impossible. Eigenfrequencies depend on Saturn's interior structure and cannot be predicted with perfect precision. Because the time between encounters is much longer than the period of a mode, in practice the phase of each mode must be regarded as randomized for each flyby. This stochasticity introduces a complication for modeling the flyby—we do not know the initial phase

and orientation of the modes when each flyby occurred. This issue will be addressed in Section 3.3.

3.2.2 Numerical integration and model reproduction

The gravity experiments were conducted with an edge-on geometry from Earth's perspective for maximum signal to noise. The orbits were highly inclined and highly eccentric. The closest approach (C/A) of the spacecraft is about 5% of Saturn's radius from its cloud tops, approaching and receding from the planet very quickly during ingress and egress. Because of this orbital geometry, we only expect a significant signal from the planet within about an hour of C/A. Therefore we use the spacecraft orbital elements two hours before C/A to compute the initial conditions. We then numerically integrate the equation of motion for Saturn plus the potential perturbation associated with normal modes, neglecting oblateness. Including the measured static zonal gravity [Iess et al., 2019] affects the simulated normal mode range rate signal by less than 1%. We integrate the equation of motion using Mathematica's built-in integrator to generate a three-dimensional velocity time series. We then subtract the Keplerian solution from the numerical solution to isolate the signal from the spectrum of modes we are modeling. Finally we project the three-dimensional velocity vector onto the line-of-sight vector connecting Saturn to the Earth. We verify the accuracy of this method by reproducing the signature from static zonal J 's [Iess et al., 2019] using the method from this chapter, finding good agreement. This method is fully general for any potential perturbation, and we will use it to inspect the behavior of normal modes.

We verified empirically the approximate linearity of combining the velocity per-

turbation from various sources. That is,

$$\delta\mathbf{v}(t) = \sum_q \mathbf{v}_q(t) \quad (3.9)$$

within $<0.1\%$ for the perturbation magnitudes in question. Strictly speaking this linearity does not hold absolutely; although gravity potential perturbations are exactly linear, a test particle encountering these perturbations may be perturbed from its trajectory—if this perturbation is sufficiently large the linearity breaks down. But for the small perturbation of interest, this non-linearity is not important.

It is important to discuss at this point a fundamental ambiguity in probing for normal modes from the spacecraft’s perspective. The spacecraft is observing two sources of variation of the gravity signal: the intrinsic geometric variation, and the temporal variation. The geometric variation is the physical shape of the mode, which attenuates with distance and varies with the spacecraft’s latitude and longitude relative to Saturn. As the spacecraft approaches and recedes from the planet, traveling from the north to south and west to east, even a static gravity perturbation would have a time-dependent signal from the spacecraft’s frame of reference. On the other hand, the potential perturbation itself varies with time. The convolution of these effects makes it difficult to have a simple intuition for Cassini’s response to each mode. See the Appendix for further discussion.

3.2.3 Spectral model

For the spectral model, we aim to be as agnostic as possible. We do not know with certainty by what mechanism seismic activity is excited on Saturn, although meteor impacts [Wu and Lithwick, 2019] or exotic meteorological phenomena [Markham and Stevenson, 2018] have been suggested. Therefore we only use simple parametric

models that scale the energy of each mode as a function of parameters. We tested a variety of scaling relationships, including power law dependence in frequency, as well as power law dependence on quantum numbers n , l , and $\frac{l-|m|}{l}$. We also tested equipartition. None of these models provided convincing fits to the data.

The nominal model used is a Gaussian frequency dependent model, although here too we aimed to be as agnostic as possible. The Gaussian is convenient because it has the power to probe for a diverse variety of frequency dependencies by only varying two parameters. Using this assumption, the mode energy is a function of its eigenfrequency according to

$$E(\omega_{nlm}) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(\omega_{nlm} - \omega_0)^2}{2\sigma^2}\right]. \quad (3.10)$$

We tested extreme parameters, varying the peak of the Gaussian between 0 and 5mHz, well above the acoustic cutoff frequency. We also tested widths between an extremely narrow distribution of $10\mu\text{Hz}$ and an extremely wide distribution of 5mHz (see Figure 3.3). By varying the parameters so widely, we can capture a wide variety of possible frequency dependent behavior. A Gaussian with a far away peak behaves approximately like an exponential relationship. A Gaussian with an extremely narrow peak behaves approximately like a δ -function, and one with an extremely wide peak approximates equipartition. As discussed in Section 3.4, we find narrowly peaked distributions centered on low-order p-modes to be the most likely to fit the data, although there is considerable degeneracy within that region of parameter space. We settled on the Gaussian dependence on frequency after trying a variety of parametric models because it provided the best fit to the data, and is flexible in qualitatively approximating many diverse behaviors. We do not claim that the real power spectrum behaves in exactly this way.

This frequency dependence on energy has a straightforward connection to excitation and dissipation efficiency, if the energy spectrum is in steady state. In this case,

$$E(\omega) = \frac{\dot{E}_{\text{in}}(\omega)Q(\omega)}{\omega}, \quad (3.11)$$

where $Q(\omega)$ is the frequency-dependent quality factor and $\dot{E}_{\text{in}}(\omega)$ is the frequency-dependent excitation rate. The mode energy is $E_{\text{nlm}} = a_{\text{nlm}}^2 \omega_{\text{nlm}}^2 M_{\text{nlm}}$ where M_{nlm} is the modal inertia uniquely defined for each normal mode according to $M_{\text{nlm}} = \int \rho |\xi_{\text{nlm}}|^2 d^3r$, where ξ_{nlm} is the mode eigenfunction normalized such that the surface displacement is 1cm. Therefore the energy scales as the square of the amplitude.

3.3 Data stacking

There is a fundamental ambiguity when modeling normal modes that does not exist for a static gravity field: the phase of the mode in question. If we only had one flyby, breaking this ambiguity would be hopeless; since we have multiple, we can do better. By combining multiple flybys, we can average out the effect of initial phases. This can be done perfectly if there is a large number of identical flybys. Indeed, one can show (see Appendix) for a particle on a prescribed trajectory $\mathbf{r}(t)$ encountering a potential perturbation of the form of Equation 3.8 where ϕ_0 and α_{nlm} are random variables for each flyby, that the summed squared potential obeys the asymptotic relationship

$$\sum_i^N \delta\Phi_i^2(t) \sim \frac{N}{2} \sum_{\text{nlm}} A_{\text{nlm}}^2 f_{\text{nlm}}^2(t), \quad (3.12)$$

where $A_{\text{nlm}} \equiv a_{\text{nlm}} \delta C_{\text{nlm}}$, and $f_{\text{nlm}}(t)$ is a deterministic function of time independent of ϕ_0 and α_{nlm} . This approximation is valid at large N . It is possible to derive a similar expression for acceleration perturbations, simply the gradient of the potential perturbations, and for velocity perturbations (see Appendix). In fact these derivations

depend on assuming a prescribed trajectory, but in reality the potential perturbations perturb the trajectory itself. We verify empirically that for a sufficiently large number of flybys

$$\sum_i^N \delta v_i^2 = \frac{N}{2} \sum_{\text{nlm}} a_{\text{nlm}}^2 \sum_i^N \delta v_{\text{nlm},i}^2, \quad (3.13)$$

where $\delta v_{\text{nlm},i}^2$ is a randomly generated squared time series of velocity perturbation associated with the spacecraft encountering the potential perturbation due to a 1cm displacement amplitude mode with quantum numbers (n, l, m) .

We note that the stochastic behavior of the modes approaches deterministic behavior when summing over a large number of flybys to demonstrate why such an exercise is useful: it reduces the stochastic component of the signal and amplifies the deterministic component. In the real experiment, however, there were only five flybys, which is not large enough to simply stack the data and compare it against the asymptotic average. Therefore, we ran a Monte-Carlo simulation, leaving the initial phase as a free random variable, and combined the signal from five randomly selected flybys with an input spectrum scaled according to Section 3.2.3. In this case, we use the relationship from Equation 3.9 to obtain

$$\sum_i^N \delta v_i^2 = \sum_i^N \left(\sum_{\text{nlm}} a_{\text{nlm}} \delta v_{\text{nlm}} \right)^2. \quad (3.14)$$

This expression is equivalent to Equation 3.13 in the limit of large N , but for finite N has stochastic components which should be accounted.

With $N = 5$ we can eliminate significant ambiguity. The raw data is shown in the red scattered points in Figure 3.6 for each flyby (the black curves are model fits to the data). The raw data were obtained by subtracting the observed spacecraft signal from a model excluding stochastic acceleration [Iess et al., 2019]. We take

these points and bin them into 150 second windows so that most points in time will have contributions from all flybys (see Section 3.3.1 for why this is important). We then average the square value of the corresponding data point across the five flybys to obtain an average value. After accounting for various quantifiable sources of error, we produce Figure 3.5 that shows the stacked data with error bars in red, with a black curve as a good fit forward model.

3.3.1 Error sources

To average this data, we must propagate the errors from the input data, and account for additional errors from the stacking process. We have identified three quantifiable sources of error in the tracking system, which we use for the error bars. The first source of error is the intrinsic noise in the system [Iess et al., 2019]. This source of error affects all data points.

The second source of error is the fact that part of the “real” non-static, non-zonal gravitational signal may have been aliased into the uncertainty about the static zonal gravity harmonic coefficient J ’s [Iess et al., 2019]. To understand how this impacts the data, we ran a Monte Carlo simulation systematically adding the gravitational signal δJ_l for each zonal gravity harmonic l , and running that modified data through our stacking pipeline. We modeled each δJ_l as a statistically independent normally distributed random variable using the published 1σ formal uncertainty [Iess et al., 2019] (although the total values for different J_l ’s are correlated, small deviations δJ_l can be approximately independent). We found the impact of this effect by taking the standard deviation of the stacked data for 1000 such simulations and used those values as an additional independent source of error to add in quadrature with instrumental noise. This source of error is most important near closest approach.

The third source of error is only applicable to a subset of points, but is the most important source of error for those points. Because we will be comparing this data to simulations without gaps or sampling issues, we need to account for the fact that some data points do not average all five flybys. This occurs because the time window in which the spacecraft is blocked by Saturn's rings is slightly different for each flyby, and because some of the data sets end before others. When we only average a subset of data points together, there will be a systematic offset from the otherwise smooth behavior of the average. We quantify this offset by taking samples of points that have data from all five flybys, then calculate the average systematic offset caused by using only a subset of those data points. We use this average value as an additional source of error, which is simply a function of the number of data points averaged. If there are five data points, this source of error is zero. Note that this error is systematic, so a series of points all missing one data set will not be randomly scattered around the main curve but will be systematically offset from it. Accounting for these three quantified sources of error produces the red points and error bars in Figure 3.5, which is the time series data set we will attempt to reproduce (with an example black curve model fit).

There are additional sources of error that are likely to prevent us from getting a perfect fit to the data. First, we do not know the actual eigenfrequencies of the modes. We attempted multiple assumptions for the frequencies to verify that our conclusion is not affected by different choices. In fact, the signature from a flyby is a slowly varying function of mode frequency, so expected errors (less than 10% in frequency) should not affect the general, qualitative behavior of the flyby signature. Nevertheless, errors in frequency yield systematic modeling errors, small temporal offsets for small errors and slowly varying qualitative behavior for larger errors, such

that the fit will not be perfect. This impacts the goodness of fit. An additional source of possible modeling error is the simplicity of our assumptions (a smoothly varying amplitude spectrum). For example, the excitation mechanism may be partly stochastic in nature [Goldreich et al., 1994] [Markham and Stevenson, 2018], and the real frequency dependence may be more complicated or jittery than a simple Gaussian. Accounting for this possibility, however, would violate the purpose of this investigation: to keep the number of parameters small, and the spectral model simple. Another cause of error we have not formally accounted is the difference in geometry between the flybys. To first order, the orbit is similar and the Saturn-Earth orientation is nearby during each flyby. But the subtle differences in geometry means we should not expect the assumption of fixed geometry and identical orbit initial conditions to reproduce the data exactly. Nevertheless, this is a necessary assumption in order to use the stacking method to amplify the deterministic component of the signal. We note these sources of error not to rigorously quantify their effect, but to justify our relatively lax error tolerance for goodness of fit; the upshot here is that we are trying to evaluate the probability of reproducing the general qualitative behavior of the signature for a given power spectrum, not to provide a single exact reconstruction of the gravity field Cassini encountered (doing so would be impossible with the available data anyway).

3.4 Analysis and results

Now that we have added error bars to account for the straightforwardly quantifiable sources of error, we can attempt to fit them. We do this with a reduced χ^2 test according to

$$\chi^2 = \frac{1}{D-f} \sum_i^D \frac{(x_i - m_i)^2}{e_i^2}, \quad (3.15)$$

where D is the number of data points, f is the number of degrees of freedom in the model (three in our case: the two Gaussian parameters and the scaling coefficient), x_i is data point i , m_i is its corresponding modeled value, and e_i is the error. Choosing an appropriate model is subtle. One choice is to use the asymptotic average of an infinite number of flybys for a given spectral model. As demonstrated in Section 3.3, we expect the data to converge toward this average. But given the finite number of flybys, there will be variation from this asymptotic mean. Therefore in order to evaluate the likelihood of a given model, we conducted 2,000 tests of five simulated flybys for each modeled spectrum.

To produce our forward model, we ran 10^4 simulated signals from individual modes, for the subset of considered modes (recall our results are not sensitive to the specific choices of considered modes or computed eigenfrequencies. For more discussion see Section 3.2). After simulating a large number of range rate signals for each individual mode, we chose five to combine their squared signal using Equation 3.14, where a_{nlm} for a given model is computed according to Equation 3.10 with $E_{\text{nlm}} = a_{\text{nlm}}^2 \omega_{\text{nlm}}^2 M_{\text{nlm}}$. This is the forward model we use to try to fit Figure 3.5. Our tolerance threshold for goodness of fit is $\chi^2 = 50$. This is a large value, but we consider it sufficient to qualitatively reproduce the essential shape of the data (for further discussion as to why this is appropriate, see Section 3.3.1). Choosing a different threshold does not significantly affect the results, but reduces probability of fitting within the tolerance threshold for all models. We then use the large number of experiments to assign a probability of reproducing the data within tolerance for a given input spectrum. The results show a degenerate set of distributions which can reproduce the signal. We found a strongly favored region of parameter space after coarse sampling using extreme parameters, then followed up with a finer sampling in that region.

The probability plots are shown in Figure 3.3, which illustrates a clearly preferred region of parameter space, but degeneracy within that region. Each grid cell of the Figure 3.3 represents a particular spectral model. The shading indicates the probability of reproducing the data within our tolerance threshold, if that spectrum were Saturn's normal mode spectrum. The degeneracy can be partly understood by considering the contributions to the gravity signal a moving spacecraft encounters. The degeneracy can be understood in two ways. First, as shown in Figure 3.4, there is a great deal of overlap between favored models. Second, because the spacecraft is moving through space, a static field would have a time-dependent signature. Because normal modes oscillate in time, there is another source of time dependence which would be experienced even by a stationary test particle. The synthesis of these two contributions allows gravity perturbations with different properties to produce a similar signal along the Saturn-Earth line-of-sight axis from the spacecraft's frame of reference. We elaborate on this second degeneracy source in Section 3.2.2 and in the Appendix.

Although we cannot identify a single conclusive power spectrum, we can exclude a wide variety of simple spectra, and find the highest probability models favor a relatively narrowly peaked distribution. The location of the peak is also constrained, with the most likely models having a peak between 500 and 700 μHz .

Although the exact width and frequency peak cannot be precisely determined, we can exclude a wide variety of models as implausible, and note a clear clustering of models which have a high probability of reproducing the observed signal. low-order f-modes lie generally below 200 μHz (see Figure 3.1). No good-fit models favor significant contributions from f-modes. We demonstrate an example of what we consider

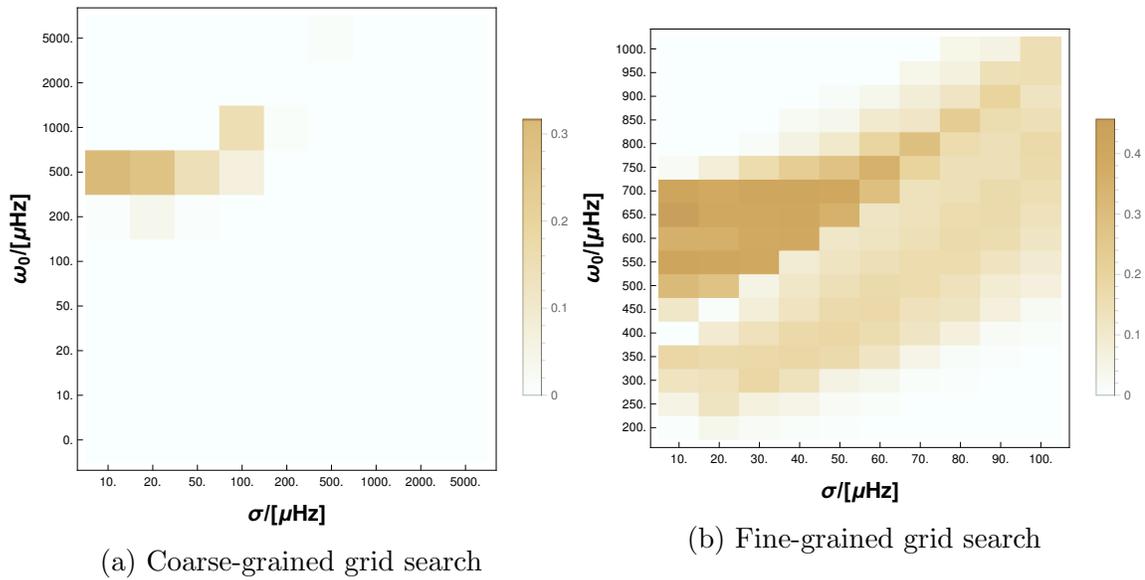


Figure 3.3: Probability maps for different input parameters. Darker colors correspond to models which have a higher probability of satisfactorily fitting the data. (a) Left shows a coarse grained plot, which searches a wide range of parameter space including models which approximate exponential behavior, delta-functions, or white noise, indicating a preferred region of parameter space. (b) Right shows a finer grained sampling in this region, illustrating the degeneracy within that region.

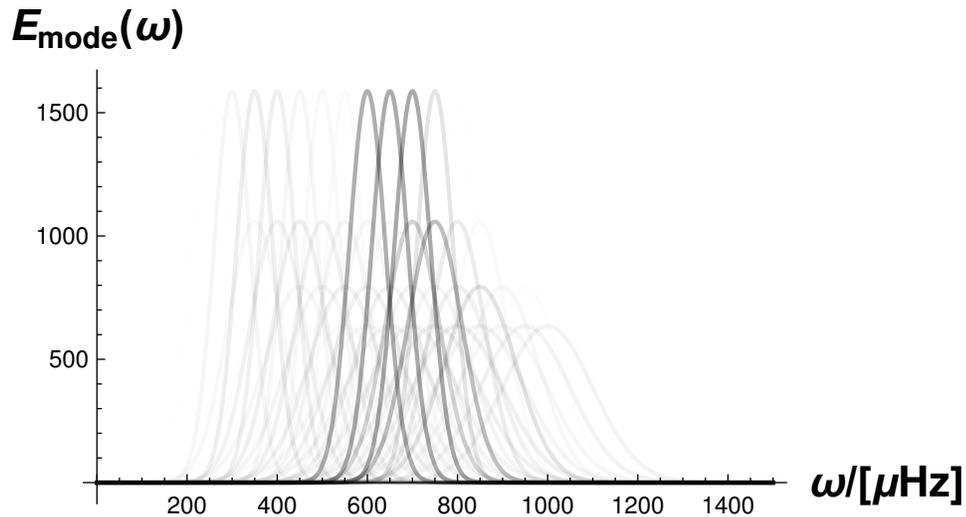


Figure 3.4: Plotting different allowable solutions, with the darkness of the curve corresponding to the probability that, if that spectrum is correct, we would observe the data within our tolerance level. The curves are normalized in the plot such that their integrated value is unity (using Hz rather than μHz as the ordinate).

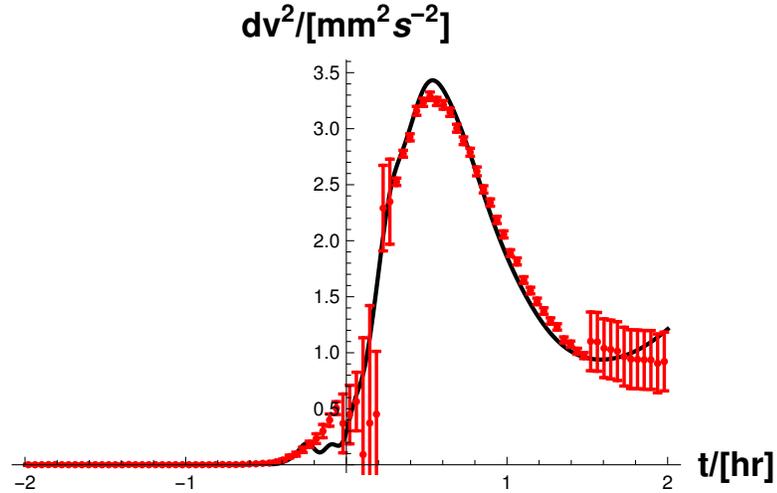


Figure 3.5: An example fit to the stacked data. The data is represented with error bars according to Section 3.3.1, with the black curve corresponding to an example energy spectrum with $\omega_0 = 600\mu\text{Hz}$ and $\sigma = 40\mu\text{Hz}$.

to be a “plausible” fit near the cutoff threshold in Figure 3.5. This particular run has $\chi^2 = 4$, among the better fits we were able to obtain.

We also assessed the frequency content of the residual data. This proved to be less diagnostic than fitting the time series. This is perhaps unsurprising, because in order to fit the frequency content, you only need a model that varies on the correct timescales, driven largely by the geometric $(r/R)^l$ effect. By contrast, in order to fit the time series data, you have to match much more specific behavior. Looking at Figure 3.5, the time series model must fit several specific phenomena. From left to right on the figure, the best models the steepness of the “ramp up” before C/A, the timing of the peak after C/A, the width of the main curve, the timing of the plateau/turnover, and other features. We find that the plausible time series fits are also compatible with the data’s Fourier transform. However, the Fourier transform is much more degenerate and possible to fit with a wide variety of models, and it is

difficult to obtain any new information.

To verify that our results were not excessively biased by our assumptions, we ran a variety of tests and alternatives, in addition to trying various parametric models as described in Section 3.2.3. We also explored the possibility that the signal may be dominated by a single mode, by testing that hypothesis against each mode in our sample. This possibility seems plausible based on our results given the narrowness of the peak in many best fit cases. We found some modes within the preferred region of parameter space had a finite probability of reproducing the data, but the probability was lower than our preferred spectral models. Consistent with our spectral method, f-modes were not favored. No f-mode had a probability higher than 2% of producing the observed signal.

We separately tested all f-modes up to spherical order 20 for completeness, because of the a priori expectation that they should be the most gravitationally important modes, and because some of their frequencies overlap with the degenerate region of parameter space that can have some probability of providing a tolerable fit to the data. High degree f-modes are discussed in more detail in the appendix. Even allowing for higher order f-modes, we did not find any simple combinations that satisfactorily reproduce the data. Perhaps f-modes are inefficiently excited for reasons beyond their frequency; for example, some excitation models depend on compressibility [Goldreich et al., 1994], and in the sun mode power declines for increasing l even at fixed eigenfrequency.

We also tested g-modes. Although our interior model assumptions did not produce g-modes, we tested them specially using published eigenfrequencies [Gudkova and Zharkov, 2006], and separately tested asymptotic approximations for their eigen-

frequencies [Tassoul, 1980]. Without the eigenfunctions, we tested the spectral model by varying the gravity harmonic coefficient perturbations directly with frequency dependence. We did not find any solution which could satisfactorily reproduce the data with g-modes. We also explored the possibility of using a given mode's eigenfrequency as a tunable parameter, varying our expectation for each f-mode's eigenfrequency between half and thrice its theoretically predicted value. Although some frequencies fit better than others, none came close to the goodness of fit we obtain with our spectral model. We also tested the full pipeline omitting one flyby, testing each subset of four flybys to ensure the results were consistent, and not a spurious peculiarity of these five particular flybys. That is, we wanted to ensure that if one of the flybys had had a problem such that it did not successfully transmit data, that it would not have altered our conclusion. As expected, with fewer flybys the preferred region of parameter space could not be as tightly constrained, but the results were consistent and favored the same region shown in Figure 3.3. If future missions can perform the same experiment with a larger number of flybys, we may be able to make stronger conclusions.

All plausible spectral models predict large peak mode amplitudes on the order of several kilometers for a small number of modes (of order 5-10) near the peak frequency. Mode amplitudes inferred from the velocity map time series of Jupiter are of order 100m [Gaulme et al., 2011], so in order to explain our findings we require the peak amplitudes to be at least an order-of-magnitude larger on Saturn than have been observed on Jupiter.

This method also allows us to fit the range rate residuals from each individual flyby. We begin with a sample amplitude spectrum with a high probability of reproducing the data (see Figure 3.3). We then run a suite of simulations with random initial phases of each mode and show the best fit results for each flyby in Figure 3.6. Each

fit optimizes for the best fit scaling coefficient, and all are in agreement within a factor of two.

3.5 Discussion

We can reproduce the behavior of the non-zonal and/or non-static component of Saturn’s gravity field using a simple three parameter forward model for mode amplitudes (Gaussian peak, width, and scaling). Our model uses a simple interior model and is not sensitive to detailed assumptions about Saturn’s interior structure, spin rate, or rotation profile. For the amplitude spectrum of the modes, we rely only on a general understanding of Saturn’s eigenfunctions and the equation of continuity in order to compute its gravitational effect for a given amplitude. We can also compute the mode inertia to scale the gravity signal for a given mode energy.

The best way to think of these results is in the Bayesian sense; we do not claim incontrovertible proof of p-modes on Saturn. Rather, given this particular data set, we present the probability that different models will reproduce this data in Figure 3.3. These probabilities should be used to update prior assumptions about Saturn’s normal mode spectrum, bearing in mind that more complex models with more degrees of freedom (e.g., from Iess et al 2019) are not captured by our analysis.

We find a moderate preference for models which have a frequency peak between about 500 and 700 μHz with a narrow width, although models with peaks as low as 250 μHz or as high as 1000 μHz also have nonzero probability. Intriguingly, the inferred narrowness of the peak is analogous to the narrowly peaked five minute modes observed on the sun. Observing the power spectrum of the solar modes, one finds a peak frequency near 3000 μHz with a full width half maximum (FWHM) of order

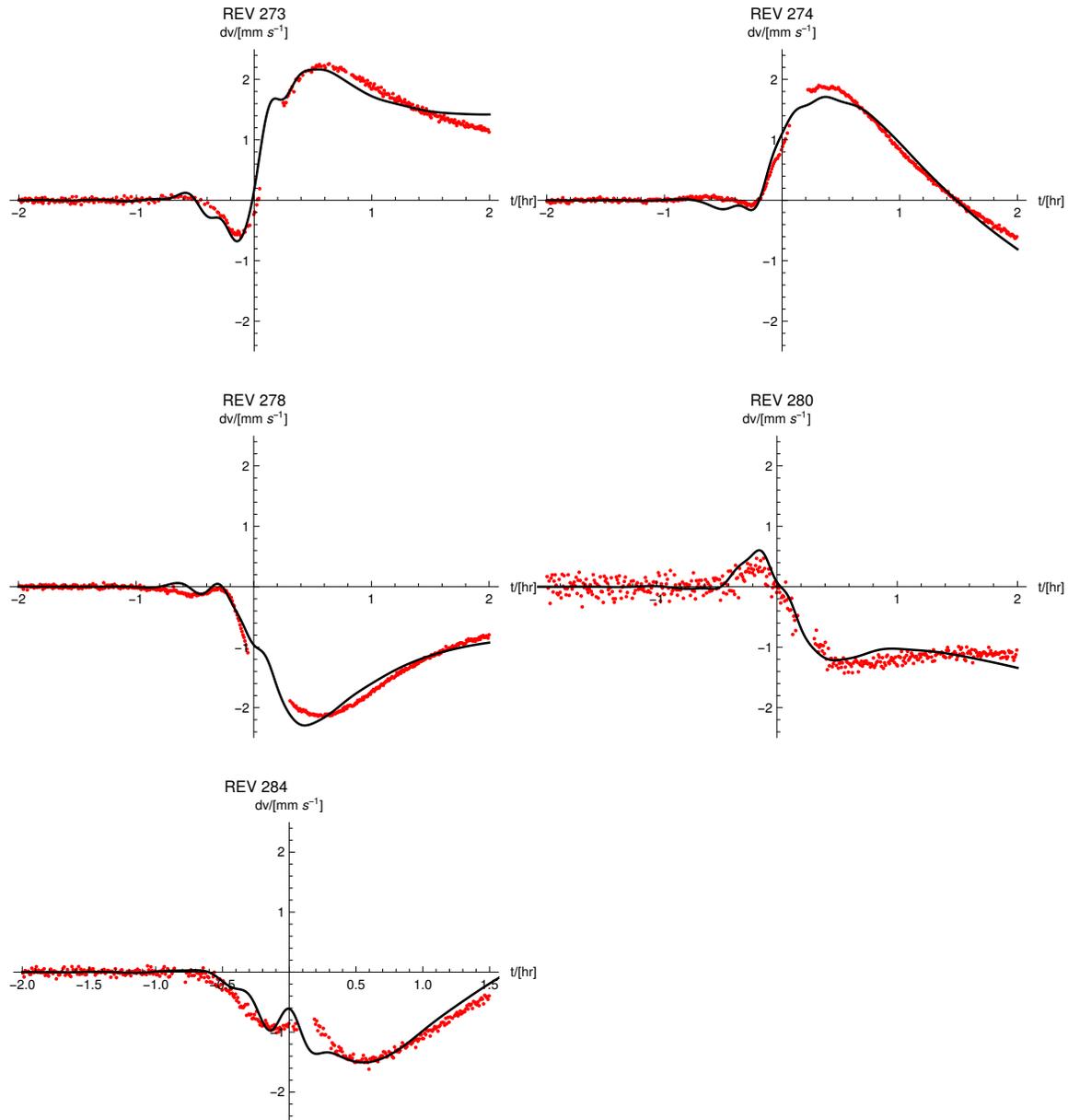


Figure 3.6: Forward calculation fits for individual flybys using our simple spectral model. The parameters used here are $\omega_0 = 600\mu\text{Hz}$ and $\sigma = 40\mu\text{Hz}$. The red scattered points are the data, and the black curves are the best fits models.

300 μHz [Frohlich et al., 1997], indicating a ratio of the FWHM to peak frequency of order 1/10. Converting the standard deviation of our Gaussian frequency-dependence functions to the corresponding FWHM, we likewise obtain a solution of order 1/10 (for example in the case of the fit shown in Figure 3.5). This may indicate some similarity between the two systems, for example that the peak frequency is set by some dynamical process with a characteristic timescale. On the sun this timescale is the eddy turnover time in the top scale height of the convective zone. Although the same mechanism cannot excite the observed amplitudes on Saturn, less frequent moist convective events with a characteristic turnover timescale (see e.g., [Markham and Stevenson, 2018]) could produce similar strong frequency dependence. On Jupiter, new theories to explain the ammonia distribution require updrafts which traverse 100km in 1000s [Guillot et al., 2020b] [Guillot et al., 2020a]. If similar dynamics occur on Saturn, the timescale is roughly consistent with the peak frequencies inferred by this work. Others have suggested a large impact as a source of Saturn’s oscillations [Wu and Lithwick, 2019]. Although a 150km impactor could in principle excite km-scale oscillations in p-modes, the scaling suggests the gravity signal from f-modes should always dominate. Therefore if this is indeed the dominant excitation mechanism on Saturn, there must be some other reason to preferentially dissipate f-modes or preferentially amplify p-modes.

Most notably, we cannot reproduce the time series data with f-modes; neither with a single f-mode dominating the signal nor with a straightforward superposition of f-modes. This finding is consistent with inferred amplitudes of f-modes which have been measured using Saturn’s rings (kronoseismology) [Hedman and Nicholson, 2013] [Wu and Lithwick, 2019] [Fuller, 2014], which are determined to be on the order of a meter in amplitude and should not produce this large of a signal in Cassini’s gravity experiment. Assuming the amplitudes inferred from ring data, the detectability

should have been marginal ($\delta v \sim 0.05 \text{mm s}^{-1}$ instead of the observed $\sim 2 \text{mm s}^{-1}$). The p-modes that are required to produce the observed signal would need surface amplitude on the order of kilometers, implying radial velocities of meters per second. These p-modes despite their large amplitudes, are not expected to show structure in the rings, because the relevant resonant radius for these frequencies is well inside Saturn's C ring.

The required p-mode amplitudes are at least an order-of-magnitude larger than were observed on Jupiter [Gaulme et al., 2011]. Interestingly, early analysis of the Juno mission indicates a similar unexplained gravity signal on Jupiter that is approximately 20 times weaker than the signal observed on Saturn [Durante et al., 2020]. This is interesting, because this analysis indicates that if the relevant amplitudes were those observed by Gaulme et al 2011, then we should expect a similar time-dependent signal diminished in scale by about an order-of-magnitude. Replicating our analysis of Cassini's gravity data for Juno, which has many more planned gravity orbits than Cassini, may be a promising future application of the method outlined in this paper. We can test to see if the inferred normal mode spectrum from gravity measurements on Jupiter is consistent with the corresponding power spectrum obtained with Earth-based observations.

Because we predict large peak amplitudes, we must consider if these are plausible and consistent with existing data. Voyager radio occultation measurements of Saturn has error estimates between 6 and 10km, and the measurements found incompatible radii between the northern and Southern hemisphere on the order of 10km [Lindal et al., 1985]. These uncertainties are compatible with the time-dependent shape variations our analysis predicts. Our analysis here would predict future measurements of Saturn's shape cannot obtain better accuracy than around a few kilometers. A series

of highly accurate measurements of Saturn's shape should have a time-dependent component on the order kilometers.

We must also consider how our findings can constrain Q . Using a set of N modes excited to 10km amplitude a (a high estimate, see Table 3.1), powered by the full luminosity of Saturn \mathcal{L} as a lower bound on Q , we compute

$$Q > \frac{a^2 \omega^3 M N}{\mathcal{L}} \sim 10^7 \quad (3.16)$$

for the relevant peak modes we have identified, where M is the modal mass and $N \sim 5 - 10$ corresponds to the number of modes with significant amplitude (for example, greater than a kilometer). This Q is compatible with estimates so far based on theory [Markham and Stevenson, 2018] [Wu and Lithwick, 2019].

We must also check that the modes can still be approximated as linear perturbations, i.e., $\mathbf{u} \cdot \nabla \mathbf{u} \ll \frac{d\mathbf{u}}{dt}$ where \mathbf{u} is the velocity vector. For the peak modes we identified, the frequency is sufficiently high that the motion is almost purely vertical. $\frac{du}{dz}$ is most significant near the surface when the atmospheric properties vary quickly. We know in this region $\frac{du}{dz} \sim \left(1 - \sqrt{1 - \left(\frac{\omega}{\omega_c}\right)^2}\right) u/H$ [Christensen-Dalsgaard, 2014], where H is the scale height and ω_c is the acoustic cutoff frequency. Therefore the condition that the system can be treated linearly is

$$\left(1 - \sqrt{1 - \left(\frac{\omega}{\omega_c}\right)^2}\right) \ll \omega H/u, \quad (3.17)$$

the ratio of the left hand side to the right hand side using $\omega \sim 5 \times 10^{-3}$ for 1km amplitude modes is about 10^{-3} . This is the maximum value near the surface; the value is much smaller in the interior where most of the mode inertia is. If the inferred

amplitudes are correct, nonlinear effects are important for p-modes on both Saturn and Jupiter.

Many questions remain, and it is clear that the field of giant planet seismology—both observational and theoretical—is in its infancy. Here we demonstrate that the unexpected and unexplained components of Saturn’s “dark” gravity field can be straightforwardly modeled as simple frequency dependent seismic activity. This provides one more piece of plausible evidence that the giant planets are seismically active, and should motivate further observations and theoretical study.

3.6 Appendix: intuition, tables, and derivations

Table 3.1: Sample spectrum using $\omega_0 = 600\text{m}\mu\text{Hz}$ and $\sigma = 40\mu\text{Hz}$, listing modes with amplitudes larger than 100m. This spectrum should not be taken too seriously as good fit solutions are degenerate and non-unique—the general orders of magnitude should be paid attention to more than the specific modes.

(\mathbf{n}, \mathbf{l})	$\omega_{\mathbf{n}\mathbf{l}\mathbf{m}}/[\mu\text{Hz}]$	$a_{\mathbf{n}\mathbf{l}\mathbf{m}}/[\text{m}]$
(2,4)	497	107
(2,5)	525	461
(2,6)	551	1220
(2,7)	575	2140
(2,8)	597	2700
(3,2)	553	1190
(3,3)	599	2350
(3,4)	633	1900
(3,5)	662	965
(3,6)	689	334
(4,2)	667	794

3.6.1 Potential perturbation

In the following derivation, we will assume the spacecraft is on a prescribed Keplerian orbit, and assess the gravity potential field it encounters as a function of time. In reality, the data we have is a velocity time series; we cannot measure the gravity potential directly. However, we demonstrate how a gravity potential with stochastic elements can be averaged toward deterministic behavior that isolates information about the amplitude spectrum. This same basic procedure will be employed to forward model the velocity time series, although the details in that case are considerably more complicated.

We begin with Equation 3.8. The full potential of all modes for a given flyby i is

$$\Phi_i(t) = \sum_{\text{nlm}} A_{\text{nlm}} \left(\frac{R}{r(t)} \right)^{l+1} P_{lm}(\cos(\theta(t))) \cos[m(\phi(t) - \Omega t - \phi_i) - \omega_{\text{nlm}} t - \alpha_{\text{nlm},i}], \quad (3.18)$$

with $A_{\text{nlm}} \equiv a_{\text{nlm}} \delta C_{\text{nlm}}$. ϕ_i is the random initial longitudinal orientation of Saturn (the same for each mode, but random for each orbit because Saturn's spin rate is not precisely known), and $\alpha_{\text{nlm},i}$ is the temporal phase of each mode—random for each mode and for each flyby because we do not know the eigenfrequency with sufficient precision to impose phase coherency between subsequent close encounters. The uncertainty in ϕ_0 can be absorbed into α ; for zonal modes, ϕ_0 does not matter, and for tesseral/sectoral modes it can be added into α so that there is only one relevant random variable. We can rewrite this expression using the harmonic addition theorem:

$$\Phi_i(t) = \sum_{\text{nlm}} A_{\text{nlm}} f_{lm}(t) (\cos(g_m(t)) \cos \alpha_{\text{nlm},i} - \sin(g_m(t)) \sin \alpha_{\text{nlm},i}), \quad (3.19)$$

where $f_{lm}(t) = \left(\frac{R}{r(t)} \right)^{l+1} P_{lm}(\cos(\theta(t)))$ and $g_m(t) = m(\phi(t) - \Omega t) - \omega t$ is the same for every orbit and does not depend on the random variable $\alpha_{\text{nlm},i}$. Now we square this

expression

$$\begin{aligned}
\Phi_i(t)^2 = & \sum_q \sum_{q'} A_q A_{q'} f_q(t) f_{q'}(t) (\cos(g_q(t)) \cos(g_{q'}(t)) \cos \alpha_{q,i} \cos \alpha_{q',i} \\
& - \cos(g_q(t)) \sin(g_{q'}(t)) \cos \alpha_{q,i} \sin \alpha_{q',i} - \sin(g_q(t)) \cos(g_{q'}(t)) \sin \alpha_{q,i} \cos \alpha_{q',i} \\
& + \sin(g_q(t)) \sin(g_{q'}(t)) \sin \alpha_{q,i} \sin \alpha_{q',i}),
\end{aligned} \tag{3.20}$$

where we substitute a single index q to refer to a given mode (n, l, m) for notation convenience.

From here, we perform the crucial step of summing over many such flybys. This total can be expressed

$$\begin{aligned}
\sum_{i=1}^N \Phi_i(t)^2 = & \sum_{i=1}^N \sum_q \sum_{q'} A_q A_{q'} f_q(t) f_{q'}(t) (\cos(g_q(t)) \cos(g_{q'}(t)) \cos \alpha_{q,i} \cos \alpha_{q',i} \\
& - \cos(g_q(t)) \sin(g_{q'}(t)) \cos \alpha_{q,i} \sin \alpha_{q',i} \\
& - \sin(g_q(t)) \cos(g_{q'}(t)) \sin \alpha_{q,i} \cos \alpha_{q',i} \\
& + \sin(g_q(t)) \sin(g_{q'}(t)) \sin \alpha_{q,i} \sin \alpha_{q',i}).
\end{aligned} \tag{3.21}$$

In this case, we can assume random variables behave as true statistical averages. We make use of the fact that $\sum_{q,q',i,j} \cos \alpha_{q,i} \cos \alpha_{q',j} \rightarrow \delta_{qq'} \delta_{ij}$. Similarly $\sum_{q,q',i,j} \cos \alpha_{q,i} \sin \alpha_{q',j} \rightarrow 0$. Using this asymptotic behavior, we can write for a sufficiently large number of N flybys,

$$\sum_i^N \Phi_i^2(t) \sim \frac{N}{2} \sum_{nlm} A_{nlm}^2 f_{nlm}^2(t). \tag{3.22}$$

Thus a particle on a prescribed trajectory encountering potential perturbations due to normal modes will, when averaging over many such encounters, approach a deterministic curve that does not depend on the initial phase of each mode.

3.6.2 Acceleration and velocity

The gravity experiment does not directly measure the gravity potential. The data we have is the velocity perturbation along a single axis. Our data stacking method is most straightforwardly derived for gravity potential perturbations. To calculate gravitational acceleration we apply $\delta g_{nlm} = -\nabla \delta \Phi_{nlm}$ to obtain

$$\begin{aligned} \frac{\delta \mathbf{g}_{nlm}}{A_{nlm}} = \frac{GM}{r^2} \left(\frac{R}{r} \right)^l & \left[-(l+1) \cos[m(\phi - \Omega t) - \omega t - \alpha] P_l^m(\cos \theta) \hat{\mathbf{r}} \right. \\ & + \cos[m(\phi - \Omega t) - \omega t - \alpha] \frac{\partial P_l^m(\cos \theta)}{\partial \theta} \hat{\boldsymbol{\theta}} \\ & \left. - \frac{m}{\sin \theta} \sin[m(\phi - \Omega t) - \omega t - \alpha] P_l^m(\cos \theta) \hat{\boldsymbol{\phi}} \right]. \end{aligned} \quad (3.23)$$

This gives us the components we need to project onto $\hat{\oplus}$, the unit vector pointing toward Earth. The total gravitational acceleration perturbation from normal modes can be written as

$$\delta g_{\oplus}(t) = \sum_{nlm} A_{nlm} g_{\oplus, nlm}(t), \quad (3.24)$$

where we can express

$$g_{\oplus, nlm}(t) = A_{nlm} (f_1(t) + f_2(t)) (f_3(t) \cos \alpha + f_4(t) \sin \alpha), \quad (3.25)$$

where $f_i(t)$ are tedious but nevertheless well-defined functions of time, independent of α . Just as above, if we have an expression in this form, we can express the asymptotic average of the sum of a large number of flybys as

$$\sum_{i=1}^N g_{\oplus, i}^2(t) = \frac{N}{2} \sum_{nlm} A_{nlm}^2 f_{nlm}(t)^2. \quad (3.26)$$

where $f(t)^2$ is a function of $f_1(t)$ and $f_2(t)$.

A similar procedure can be followed for velocity perturbations, and this is the

source of frequency dependence (absent in averaged potential perturbations) ω_{nlm} by temporally integrating the gravitational perturbations. In fact it is not this simple; a perturbed spacecraft will deviate from its Keplerian trajectory, which leads to errors from the exact solution of order $(\Delta r/r)\Delta t$. A better approximation is to dynamically solve the equation of motion, accounting for displacement from the initial Keplerian trajectory to linear order, although this method produces higher order errors from the exact solution. The method we eventually used, solving the exact equation of motion explicitly then subtracting the Keplerian solution, is not conducive to an analytic expression. Because of this, we verified these results by testing them numerically against simulations to be sure they did approach a deterministic curve when stacked.

To do this, we performed 10^4 Monte Carlo simulations of random flybys for each mode. We then averaged random subsets of these samples to verify that they asymptotically approach an asymptotic curve. We also directly simulated various examples of a spectral superposition of different modes, running 10^4 flybys. We then compared these stacked results against the superposition of a spectrum of averaged squared modes, finding excellent agreement with Equation 3.13. We therefore have an analytic approximation which motivates the stacking procedure, as well as more exact numerical tests to make sure the results from the analytic approximation are robust to this application. These tests verify the validity of our data stacking procedure. In practice, we have a finite number $N = 5$ flybys. A random set of 5 flybys will deviate somewhat from the asymptotic behavior of 10^4 flybys, so for the actual probabilistic fitting routine we used Equation 3.14. We used the statistical method outlined in the Analysis and Results section to account for this, by directly solving for the probability that a given spectrum will produce a good fit to the data.

3.6.3 Spacecraft sensitivity and model intuition

We must consider whether the particular orbit of Cassini biases its ability to detect certain modes. If the convolution of a particular mode's space dependent eigenfunction and time-dependent eigenfrequency appears stationary in Cassini's frame during close approach, then Cassini will preferentially detect signals from this mode. For example, because Cassini's orbit is nearly polar, there may be an intrinsic preference for sectoral ($m = \pm l$) modes, as seems to be true for f-modes (see Figure 3.7). As a different example, there may be a preference for modes whose half period is near the timescale of Cassini's motion from the northern to the southern hemisphere. This appears to be the case for tesseral modes with $m = l - 1$ in Figure 3.7. Moreover, because Cassini's orbit moves from west to east, it is plausible that it would preferentially detect $m > 0$ modes whose pattern rotates in the prograde direction, an intuition also supported by Figure 3.7.

According to Figure 3.7, this quasi-resonant effect does indeed make Cassini more sensitive to certain modes. All these effects are implicitly accounted for in the forward model in the main text.

Another important consideration is high-degree f-modes. If the strong frequency dependence we predict applies equally to f-modes, then we must consider if Cassini would detect them. There are two sources of attenuation: the intrinsic geometric attenuation with distance of high degree gravity harmonics $(R/r)^l$, and the monotonically decreasing gravity potential coefficient response to each mode which obeys $(l(l+1))^{-1/2}$ [Wu and Lithwick, 2019]. We use a characteristic radius for Cassini's close encounter $r_c \sim \frac{v}{R} \int_{-R/2v}^{R/2v} r(t) dt \sim 1.15R$ where v is the periape velocity. High degree f-mode eigenfrequencies obey $\omega^2 \sim \frac{GM}{R^3} (l(l+1))^{1/2}$, so the modes in the most likely peak frequency region have $l \sim 50 - 100$. Using our attenuation estimates,

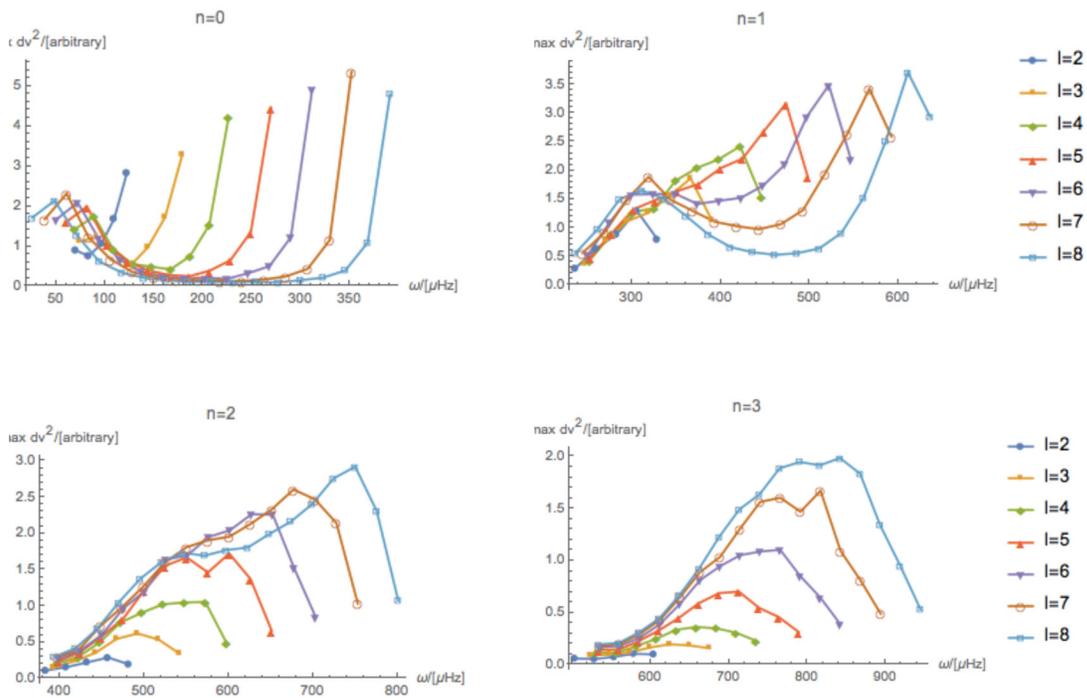


Figure 3.7: In the above plots the x-axis is the frequency in the inertial frame in which Cassini's orbit is defined. This is why different m -values are so spread out in frequency when in Saturn's rotating frame they are very close. The y-axis is the average maximum squared velocity response to a particular mode, using fixed gravity potential perturbation coefficients for each mode.

these modes would be reduced by at least 5 orders of magnitude compared to low-order f-modes, significantly more severe than the 2-3 orders of magnitude reduction in low-order p-modes. Therefore, we do not expect for Cassini to detect high order f-modes, even if they likewise had km-scale amplitudes.

The following link directs to the published version of this chapter:

<https://iopscience.iop.org/article/10.3847/PSJ/ab9f21/meta>

Chapter 4

The thermal evolution of Uranus and Neptune

In the midst of winter, I found
there was, within me, an
invincible summer.

Albert Camus

Summary

The internal heat flows of both Uranus and Neptune remain major outstanding problems in planetary science. Uranus' surprisingly cold effective temperature is inconsistent with adiabatic thermal evolution models, while Neptune's substantial internal heat flow is twice its received insolation. In this work we constrain the magnitude of influence condensation, including latent heat and inhibition of convection, can have on the thermal evolution of these bodies. We find that while the effect can be significant, it is insufficient to solve the Uranus faintness problem on its own. Self-consistently considering the effects of both latent heat release and stable stratification, methane condensation can speed up the cool-down time of Uranus and Neptune by

no more than 15%, assuming 5% molar methane abundance. Water condensation works in the opposite direction; water condensation can slow down the cool-down timescale of Uranus and Neptune by no more than 15% assuming 12% molar water abundance. We also constrain the meteorological implications of convective inhibition. We demonstrate that sufficiently abundant condensates will relax to a state of radiative-convective equilibrium requiring finite activation energy to disrupt. We also comment on the importance of considering convective inhibition when modeling planetary interiors.

4.1 Introduction

Giant planet atmospheres are primarily heated by a combination of sunlight and internal heat leftover from formation. All giant planets except Uranus are observed to emit more infrared radiation into space than the absorbed sunlight, by approximately a factor of two. Jupiter's present day luminosity can be approximately explained by convective cooling from an initially hot state over the age of the solar system [Hubbard, 1977][Hubbard et al., 1999]. To accurately reproduce Saturn's present day state, one may need to account for additional heating by the settling of helium rain from the envelope into the interior [Hubbard et al., 1999][Stevenson, 1983]. However, luminosity is a crude indicator of thermal evolution since planets can store heat internally and may have internal heat sources (e.g., differentiation). The present luminosities of Uranus and Neptune are not well understood because even their basic structures, including composition, internal structure, and thermal transport properties, are not well understood.

Measurements of the ice giants' electromagnetic emission to space began in the

1960s [Kellermann and Pauliny-Toth, 1966], with high-quality far infrared measurements constraining the effective temperatures beginning in the 1970s [Fazio et al., 1976][Loewenstein et al., 1977][Stier et al., 1978]. These early observations concluded that Uranus appeared approximately in equilibrium with its received sunlight, while Neptune emitted more than twice the radiation it received. These observations were corroborated by higher quality analysis after the Voyager 2 flybys [Pearl et al., 1990][Pearl and Conrath, 1991]. The 1σ upper limit for Uranus’ energy balance (the ratio between its emitted and absorbed thermal flux) is 1.14. The lower limit is below unity, indicating the results are consistent with zero internal heat flow. However, we know the heat flow cannot be zero because Uranus has a magnetic field. Moreover, the higher microwave temperatures at long wavelengths (e.g., [Gulkis et al., 1983]) are compatible with heat flow from depth. Uranus *must* be convective at depth.

Theoretical attempts to explain these observations began promptly. It was immediately clear that the ice giants could not have the same thermal histories as the gas giants. Early studies concluded that, if these planets cool convectively like the gas giants, they must have formed at a temperature not much warmer than their current states [Hubbard, 1978][Hubbard and MacFarlane, 1980], a highly unlikely interpretation because the energy of accretion $\sim GM^2/R$ far exceeds their current heat content for any plausible assumption of structure. Alternative theories suggested a large fraction of gravitational heat of formation remains trapped in the interior, but by some mechanism cannot escape to space [Podolak et al., 1991]. More recent studies suggest that there is no problem for Neptune [Fortney et al., 2011][Linder et al., 2019], or even that Neptune’s present luminosity is higher than expected [Nettelmann et al., 2016][Scheibe et al., 2019]. Uranus’ very low internal heat flux, sometimes known as the faintness problem [Helled et al., 2020], remains largely unsolved, although it has been suggested that the problem can be solved by modeling thin layers of static

stability near phase boundaries [Nettelmann et al., 2016]. Today it is largely accepted that the adiabatic assumption for the interior is probably inappropriate for Uranus and Neptune [Helled et al., 2020].

In this work we present a mechanism that inhibits convection near the methane cloud level, thereby trapping internal heat beneath the clouds. This mechanism has already been theorized and discussed e.g., [Leconte et al., 2017] [Friedson and Gonzales, 2017] [Guillot, 2005], but the effect of methane on the ice giants has not yet been explicitly quantified and worked into a thermal evolutionary model. In hydrogen atmospheres, sufficiently abundant condensible species can shut off convection near the cloud level [Guillot, 1995][Guillot, 2005]. By “sufficiently abundant” we mean greater than an analytically calculable critical mole fraction q_{crit} . This value is about 1.4% for methane and 1.2% for water under the relevant conditions in Uranus and Neptune. Recent theoretical study confirms this effect is also stable against double diffusive convection in a saturated medium in the fast precipitation limit, indicating radiation would be the only remaining efficient thermal transport mechanism [Leconte et al., 2017] [Friedson and Gonzales, 2017]. Methane is certainly sufficiently abundant for convective inhibition to occur [Helled et al., 2020]. The long-term survival of the configuration against entrainment is still a subject of research; the configuration may be intermittently eroded, destroyed, and reformed [Friedson and Gonzales, 2017].

In Section 4.2 we begin by laying the heuristic groundwork and providing analytic order-of-magnitude estimates of the effect of convective inhibition. Then in Section 4.3 we outline a more detailed atmospheric model. We then use this model to quantify the effects of condensation on the difference between the planet’s observed effective temperature and its internal entropy. We also comment on the meteorological implications arising from convective inhibition on Uranus, Neptune, and Saturn,

as well as its importance for interior modeling. In Section 4.4 we constrain the importance of both methane and water condensation on the planets' thermal evolution. Finally in Section 4.5 we make recommendations for future missions to the ice giants, and comment on additional applications of this mechanism to the thermal histories of exoplanets, especially super-Earths.

4.2 Intuition and analytic approximations

A hydrogen atmosphere becomes stable against convection at a critical value of the condensate mole fraction q_{crit} that depends on temperature and the properties of the condensate and the gas mixture [Guillot, 1995] [Guillot, 2005] [Li and Ingersoll, 2015] [Leconte et al., 2017] [Friedson and Gonzales, 2017]. The mechanism is as follows: consider an isobaric open system hydrogen gas parcel saturated with a vapor species of higher molecular weight. Assume there exists a finite reservoir of liquid condensate in equilibrium with the saturated parcel, outside but in contact with the system. If the parcel is relatively cool, the effect of the condensate will be a small correction, and the parcel will approximately behave like an ideal gas such that density decreases as temperature increases. However, as temperature increases at fixed pressure, the mixing ratio of the condensate likewise increases. Because the condensate vapor is heavier than the dry air, there comes a crossover where the Arrhenius relationship in temperature governing vapor pressure saturation overcomes the linear relationship in temperature governing mean spacing between molecules in a gas. After this crossover point for the system outlined above, increasing temperature actually increases the density of the parcel. For this reason, a hydrogen atmosphere with sufficiently abundant condensate ($q_{\text{max}} > q_{\text{crit}}$) with an internal heat source will not convect. It does not convect because the warmer underlying gas is Ledoux stable

to convection for any temperature gradient if saturated at all levels. In our notation,

$$q_{\text{crit}}^{-1} = \left(\frac{\mu_c L}{RT} - 1 \right) (\epsilon - 1), \quad (4.1)$$

where R is the ideal gas constant, L is the latent heat of vaporization, μ_c is the molecular weight of the condensate, and $\epsilon \equiv \mu_c/\mu_d$ is the ratio of the condensate molecular mass to that of dry air. Note that our definition of q is the molar mixing ratio, and is different from the quantity defined as q in [Leconte et al., 2017].

If we neglect radiative transfer of heat, we can analytically approximate the temperature difference between the top and bottom of the stable layer. The bottom of the upper saturated level satisfies $q = q_{\text{crit}}$, while the deep well-mixed convective atmosphere has uniform composition of the vapor satisfying $q = q_{\text{max}}$. In the limit of no thermal transport, the stable layer will reduce to a stable interface, with an unsaturated convective level beneath a saturated convective level. Thus, we can approximate the top of the deep well mixed convective layer and the bottom of the upper saturated convective layer to be at the same pressure level p . We define the temperature at the bottom of the saturated convective layer to be T_1 , and the temperature at the top of the well-mixed convective layer to be T_2 . In this case, if the saturation vapor pressure p_s satisfies the Clausius-Clapeyron relationship where $q = p_s/p$, we can immediately write down

$$\frac{1}{T_2} - \frac{1}{T_1} = \frac{R}{\mu_c L} \log \left(\frac{q_{\text{crit}}}{q_{\text{max}}} \right). \quad (4.2)$$

Equation 4.2 turns out to be a good approximation for very deep clouds (pressure of order a hundred bars, e.g., the water cloud level in the contemporary ice giants), where the atmosphere is relatively opaque. This approximation is less accurate when the atmosphere is less opaque and radiative transfer is more efficient, such as the

methane cloud deck of the contemporary ice giants, or the water cloud deck earlier in their thermal histories. Equation 4.2 will also always be an upper bound to the difference between T_1 and T_2 because it neglects thermal transport. In practice the difference between the pressure levels at the top and bottom of the stable layer play an important role. In Section 4.3, we solve the problem of radiative convective equilibrium explicitly.

The goal of this analysis is to quantify the effect of condensation on a planet's effective temperature. For a given internal entropy, there will be three important theoretical effective temperatures we discuss throughout this paper: T_e , T_{int} , and T_{ab} . We define these quantities here, referring to Figure 4.1 to illustrate a sample atmospheric profile that corresponds to its respective temperature. The effective temperature T_e (corresponding to the solid red curve on Figure 4.1) accounts for convective inhibition, and should correspond to the observed effective temperature of the planet from the outside. The internal effective temperature T_{int} (faded solid blue curve on Figure 4.1) accounts for latent heat but not convective inhibition. Finally the adiabatic effective temperature T_{ab} (black curve on Figure 4.1) is the effective temperature the planet would have if no condensation occurred at all, i.e., if the whole troposphere were dry adiabatic.

We now approximate how this temperature difference between T_1 and T_2 changes the effective temperature of the planet. Consider an initially convective atmosphere (for example, mixed by a cosmic ladle) that has effective temperature T_{int} . If $q_{max} > q_{crit}$, then part of this atmosphere will be stable to convection. The atmosphere will cool from the top, but cannot carry that heat out convectively, causing the upper layer to relax onto a cooler adiabat until radiative-convective equilibrium is reached. We wish to estimate the difference between the initial effective temperature

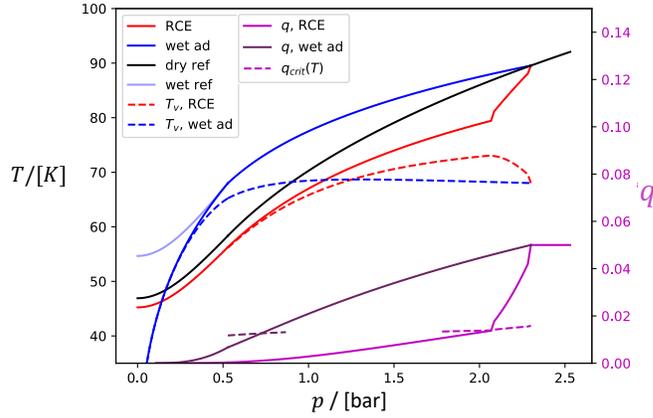


Figure 4.1: Sample atmospheric profile useful for intuition. Blue curves show wet adiabatic (wet ad) temperature (solid) and virtual temperature (dashed) profiles. Red curves show our radiative-convective equilibrium (RCE) solutions, and the black curve shows the dry adiabatic reference solution, all corresponding to the same interior entropy. Magenta and purple curves correspond to saturated abundance of methane for RCE and wet ad respectively, while the corresponding dashed curves show the local critical mixing ratio q_{crit} .

T_{int} and the final effective temperature T_e in equilibrium. Entropy increment scales as $dS \propto dT/T$ in an isobaric environment, according to the first law of thermodynamics, and adiabatic processes are isentropic. Therefore if the difference between T_1 and T_2 is not too large, we can approximate

$$\frac{T_e - T_{int}}{T_{int}} \sim \frac{T_1 - T_2}{T_2} \sim \frac{RT_2}{\mu_c L} \log \left(\frac{q_{crit}}{q_{max}} \right), \quad (4.3)$$

where T_e is the observed effective temperature of the planet after accounting for convective inhibition. This approximation neglects the non-adiabaticity due to latent heat release, which we address in the following paragraph.

In reality we expect two effects arising from condensation: the tendency toward sub-adiabatic wet pseudo-adiabaticity arising from latent heat, and the tendency toward super-adiabatic stable stratification arising from convective inhibition. These

two effects will be in opposite directions; while latent heat will tend to produce a warmer effective temperature, convective inhibition will tend to produce a cooler effective temperature with the same interior entropy. The former effect has been studied in detail [Kurosaki and Ikoma, 2017], while the latter effect is the subject of this work. We can estimate the magnitude of the latent heat effect using the definition of equivalent potential temperature that is conserved along a moist adiabat.

$$\theta_e(p, T) = \theta(p, T) \exp \left[\frac{\epsilon p_s(T)}{c_p T (p - p_s(T))} \right], \quad (4.4)$$

where $\theta \equiv T(p_0/p)^{\nabla_{ab}}$ is the potential temperature, $\nabla_{ab} \equiv \frac{\gamma-1}{\gamma}$ is the adiabatic gradient, and γ is the Grüneisen parameter. Using this, we can estimate the temperature change due to moist adiabaticity as $\frac{T_e - T_{int}}{T_{int}} \sim \exp \left[\frac{\nabla_{ab} \mu_c q_{max}}{RT_2} \right]$. Comparing this to Equation 4.3 demonstrates these two quantities should not scale in the same way. However, they are comparable in order-of-magnitude under the conditions of interest. Therefore both effects must be accounted for explicitly in order to fully understand the effect of condensation on thermal evolution of planets with polluted hydrogen atmospheres. We perform this calculation in Section 4.3.

The importance of these effects on thermal evolution are as follows. Because potential temperature relates linearly to a reference temperature, the temperature at all pressures will scale linearly with the temperature at some reference pressure. Likewise, the effective temperature of a planet scales linearly with a reference temperature in the adiabatic region, assuming constant opacity (this is actually a poor assumption, and a fully complete model must include opacity variations due to condensation explicitly. See Section 4.5 for further details). Therefore, it is possible to model thermal evolution by assuming a planet's effective temperature is linearly related to its internal heat content. The purpose of calculating the difference between the ob-

served effective temperature T_e and the adiabatic equivalent effective temperature T_{ab} is to explicitly quantify the non-linearity arising from condensation, so that thermal evolution can be modeled self-consistently. We carry this out in Section 4.4.

4.3 Atmospheric model

We model a radiative-convective equilibrium atmosphere using a two stream gray opacity approximation for thermal radiative transfer. We seek to uniquely define the apparent effective temperature T_e as a function of a planet’s internal equivalent effective temperature T_{int} and condensate abundance q_{max} . Planetary and physical properties, such as surface gravity and the physical properties of the gas mixture, are considered to be fixed.

We also assume the planet is subject to intermittent moist convective events that overcome the potential barrier of the stable layer. These could occur due to instabilities caused by entrainment over long timescales [Friedson and Gonzales, 2017], rare impact events, or strong updrafts from the interior. The equilibrium configuration then is reached by gradual cooling, with the upper layer relaxing onto a moist pseudo-adiabat set by a different potential temperature than the adiabat that sets the interior. The stable layer meanwhile will have a super-adiabatic temperature gradient set by thermal radiative equilibrium, see Figure 4.2.

In this section we explore the effect of methane abundances varying between 2–5%. The most commonly cited number for the deep mixing ratio of methane in Uranus and Neptune are 2.3% and 2% respectively, because these are the nominal values in the first published work on the atmospheric structure of these planets derived from radio refractivity data from Voyager [Lindal et al., 1987] [Lindal, 1992]. However, these

early works provide solutions to the refractivity data using assumptions of methane abundance between 1% and 4%, finding all these solutions to be theoretically compatible with the observations. Subsequent analysis from ground based and Hubble observations have likewise found a range of acceptable values for both planets ranging between roughly 2%-4% for both planets, as well as latitudinal variation in methane mixing ratio [Baines et al., 1995][Rages et al., 1991] [Baines and Hayden Smith, 1990] [Karkoschka and Tomasko, 2011]. In this work we are interested in understanding how thermal evolution is affected by methane condensation. Since the exact mixing ratio is not precisely constrained and we are interested in this question broadly, we take sample values for methane concentration between 2-5% to understand how the effect changes with methane abundance. Unsurprisingly, the effect becomes monotonically more important as concentration increases within this range, as shown in Figures 4.3 and 4.6.

4.3.1 Defining the boundaries of the stable layer

In this subsection we quantify important pressure boundaries we need to define the radiative transfer model in the following subsection. We are interested in the case where condensate is sufficiently abundant to inhibit convection, $q_{max} > q_{crit}$. We consider an atmosphere where optical depth unity in the IR is at lower pressure than the level at which convection is expected. At deeper levels (higher pressure) we assume there is a region of rapidly varying condensate mixing ratio in the vapor phase; this region can be convectively stable as discussed in Section 4.2. Deeper still, below the conventionally defined cloud deck, the condensate mixing ratio is a constant because the vapor pressure is always less than the saturated vapor pressure at that temperature. We refer to this below as the “bulk mixing ratio” though it is strictly only applicable to whatever deep, well mixed layer lies beneath the clouds and says

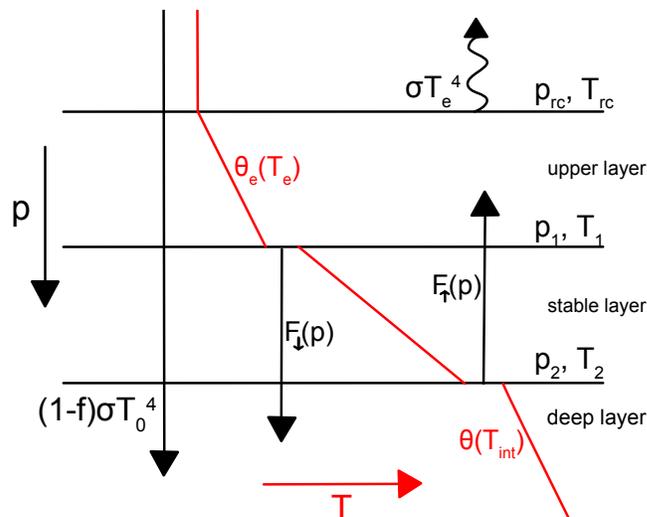


Figure 4.2: A schematic sketch of the model (not to scale). The red curves represent temperature increasing to the right, while black horizontal lines demarcate important pressure levels. The convective layers are labeled $\theta_e(T_e)$ and $\theta(T_{int})$ respectively, to indicate that their temperature structure is dictated by setting θ_e and θ to a constant value uniquely determined by T_e and T_{int} .

nothing about the actual methane abundance at far deeper levels (i.e., the methane abundance of the planet as a whole). Accordingly, our atmosphere has (from the top downward) a radiative layer (the stratosphere) a convective layer, another radiative layer (called the “stable layer” below), and a deep convective layer.

Assuming a two-stream approximation with collimated light beams, the thermal structure of an atmosphere in radiative equilibrium as a function of optical depth is $T(\tau) = T_e(\tau + 1/2)^{1/4}$. We assume the IR opacity to be dominated by pressure-induced opacity of hydrogen collisions that approximately obeys $\kappa \sim \kappa_0(p/p_0)$, where $\kappa_0 = 10^{-2} \text{g}^{-1} \text{cm}^2$ and $p_0 = 1 \text{ bar}$. Assuming a different κ_0 does not significantly affect our findings. An atmosphere in radiative equilibrium becomes unstable to convection at the point where its lapse rate becomes superadiabatic. Under our assumptions,

this places the pressure level of the radiative-convective boundary at

$$p_{\text{rc}}^2 = \frac{4gp_0R_d}{(c_p - 2R_d)\kappa_0}, \quad (4.5)$$

where R_d is the specific gas constant of dry air, and c_p is its constant pressure heat capacity. Note for ideal gasses $R_d/c_p = \nabla_{ab}$. We use a dry adiabatic lapse rate to set the radiative-convective boundary because we assume the effect of moist adiabaticity is small in this relatively cold part of the atmosphere. Beneath the boundary, we assume the atmosphere to be moist adiabatic. A moist adiabatic atmosphere conserves the equivalent potential temperature, Equation 4.4. We set the moist adiabatic equivalent potential temperature using the temperature and pressure at the radiative-convective boundary. By doing this, we define a unique moist adiabat for a given effective temperatures.

In equilibrium, the mole mixing ratio is set by the condensate's saturated vapor pressure $q(p, T) = p_s(T)/p(T)$. When we define a moist adiabat, the temperature is uniquely defined at every pressure level. Therefore we can solve for the level p_1 where the atmosphere becomes stable to convection by solving $q(p_1) = q_{\text{crit}}(T(p_1))$, where q_{crit} is defined in Equation 4.1.

Similarly, we can solve for the bottom of the stable layer by solving $q(p_2) = q_{\text{max}}$, where q_{max} is the bulk abundance of the condensate species. This is set using the pseudoadiabat corresponding to a planet with effective temperature that neglects convective inhibition. By fixing the uninhibited effective temperature T_{int} that defines p_2 , we can solve for the corresponding effective temperature T_e that satisfied radiative convective equilibrium, i.e., $F_{\uparrow}(p_2) - F_{\downarrow}(p_2) = F_{\uparrow}(p_1) - F_{\downarrow}(p_1) = \sigma(T_e^4 - fT_0^4)$ where f is the fraction of sunlight absorbed above the stable layer, and T_0 is the equilibrium

effective temperature with the sun in the absence of internal heat (see Figure 4.2).

Figure 4.2 contains apparent temperature discontinuities, which exist in the model. Of course, temperature discontinuities are not stable in natural media, as conductive heat transport will be infinite. Additionally, a temperature discontinuity—even if stabilized by a compositional difference—will lead to negligible temperature differences due to thin thermal boundary layer convection in an inviscid fluid. The temperature discontinuities in Figure 4.2 therefore do not actually represent discontinuities in nature, but steep temperature gradients. We provide the following order-of-magnitude analysis to determine how important these steep quasi-discontinuities are in the context of the model. The relaxation timescale for thermal diffusion in a medium with thermal conductivity k_t is $\tau_c \sim \frac{\rho c_p D^2}{k_t}$, where D is the vertical length scale of relevance. The timescale for radiative relaxation is the radiative time constant $\tau_{\text{rad}} \sim \frac{c_p}{8\sigma T^3 \kappa}$. We solve for the thickness of the conductive layer by equating the two timescales solving for D . Using appropriate parameters for hydrogen around 1 bar ($k_t \sim 10^4 \text{g cm s}^{-3} \text{K}^{-1}$, $\rho\kappa \sim 10^{-6}$) we find the length scale to be of order 10-100 meters, small compared to atmospheric length scales (i.e., the scale height). Then using Fourier's Law, we find heat conduction to be of order $10^{-2} \text{erg cm}^{-2} \text{s}^{-1}$ for discontinuities of order 1K (scaling linearly with the size of the temperature discontinuity), about four orders of magnitude smaller than σT_e^4 and therefore not included in the model. Thus the discontinuities in Figure 4.2 are really there in the model, but are physically understood to be steep temperature gradients nevertheless unimportant for the purposes of calculating total heat flow.

We must also comment on entrainment by convection outside the stable layer. Entrainment will tend to erode and thin the stable layer over time [Friedson and Gonzales, 2017]. In general for water, the erosion timescale is greater than the cooling

timescale, indicating that the equilibrium configuration should exist at some times. However as the stable layer is eroded and becomes thinner, heat transport across the stable layer will be enhanced, reducing the difference between T_e and T_{int} . Eventually the thinning stable region will reduce to an interface that may be stable, unstable, or conditionally stable. We acknowledge that these complications are confounding factors for our model, and therefore our results that neglect entrainment erosion of the stable layer should be thought of as an upper bound on the magnitude of the effect on $\Delta T \equiv T_e - T_{int}$ and on evolution.

4.3.2 Radiative transfer across the stable layer

Figure 4.2 is a useful visual reference for this section. In order to compute the radiative-convective equilibrium solution, we first solve for the equilibrium heat flow for a system specifying the boundaries of the stable layer p_1 and p_2 , along with their corresponding temperatures T_1 and T_2 . The temperature structure above p_1 is moist adiabatic, while the temperature structure below p_2 is dry adiabatic. In equilibrium $\frac{dT}{dt} \propto \frac{dF}{dp} = 0$. Then using the two stream approximation of radiative transfer,

$$\frac{dF}{dp_{\uparrow}} = \frac{dF}{dp_{\downarrow}} = \frac{\kappa}{2g}(F_{\uparrow} - F_{\downarrow}) \quad (4.6)$$

with appropriate boundary conditions, we can analytically solve for the upward and downward heat flux at every level in the stable layer. Of interest for our problem is the net heat flow from the deep/stable layers to the shallow layer that is convectively coupled to the photosphere. This is

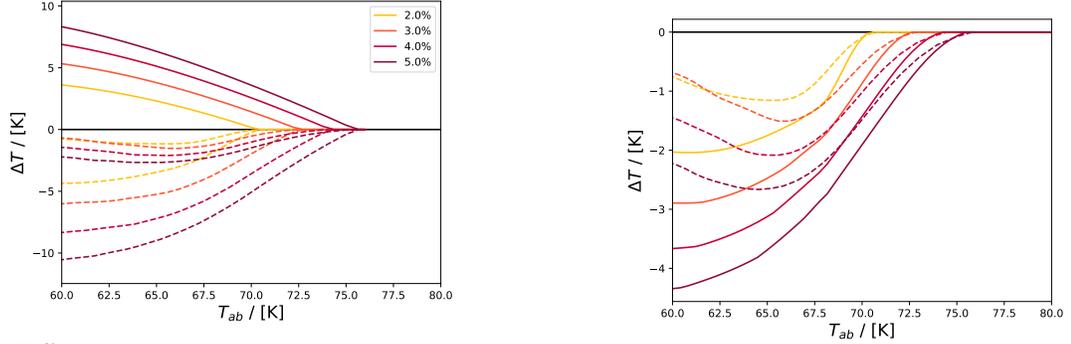
$$F_{\uparrow}(p_1) - F_{\downarrow}(p_1) = \frac{4F_2gp_0 + F_1(p_2^2 - p_1^2)\kappa_0}{4gp_0 + (p_2^2 - p_1^2)\kappa_0} - F_1, \quad (4.7)$$

where $F_1 = F_{\downarrow}(p_1)$ and $F_2 = F_{\uparrow}(p_2)$ are the boundary conditions. We can solve for these boundary conditions using the uniquely determined temperature structures above and below the stable layer, i.e.,

$$F_1 = \frac{\kappa_0 \sigma}{g} \int_0^{p_1} \exp \left[\frac{\kappa_0 (p^2 - p_1^2)}{2gp_0} \right] T(p)^4 (p/p_0) dp, \quad (4.8)$$

$$F_2 = \frac{\kappa_0 \sigma T_2^4}{g} \int_{p_2}^{\infty} \exp \left[\frac{\kappa_0 (p_2^2 - p^2)}{2gp_0} \right] (p/p_2)^{4R/c_p} (p/p_0) dp. \quad (4.9)$$

Using this process, for a given $(T_{\text{int}}, q_{\text{max}}) \implies (p_2, T_2)$ we solve the above nonlinear equation to obtain $T_e \implies (p_1, T_1)$ using the condition $F_{\uparrow}(p_1) - F_{\downarrow}(p_1) = \sigma(T_e^4 - fT_0^4)$. This defines the observed effective temperature as a function of the internal effective temperature $T_e(T_{\text{int}})$, shown as the dashed-dotted curves in Figure 4.4a. In order to relate the observed effective temperature T_e to the adiabatic effective temperature T_{ab} , one must additionally consider the effect of latent heat, shown as solid curves in Figure 4.4a. Then the net effect, $T_e(T_{ab})$ is shown as dashed curves in Figure 4.3, where $\Delta T = T_e - T_{ab}$. Figure 4.1 shows a sample temperature/pressure profile to illustrate the contributions from each source. Similar results are shown for water in Figure 4.4. These figures show ΔT under two assumptions: an initially dry adiabatic atmospheric profile (solid curves) where moist adiabaticity is not considered, and an initially moist adiabatic upper layer (dashed curves), as described in detail in the text. In the dry adiabatic case, $T_{\text{int}} = T_{ab}$. One may notice that $\Delta T = T_e - T_{ab}$ arising from the dry adiabatic case is substantially smaller than $T_e - T_{\text{int}}$ in the wet adiabatic case (dashed-dotted curves in Figure 4.4a). There are two reasons for this. First, the dry adiabatic lapse rate is steeper than the wet pseudo-adiabatic lapse rate, which improves the efficiency of radiative transfer. Second, the super-adiabatic stable region makes less of a difference when the initial profile was already comparatively steeper than the moist adiabatic case. For this reason, the results are less sensitive to assumptions about the initial temperature profile than one might initially expect.

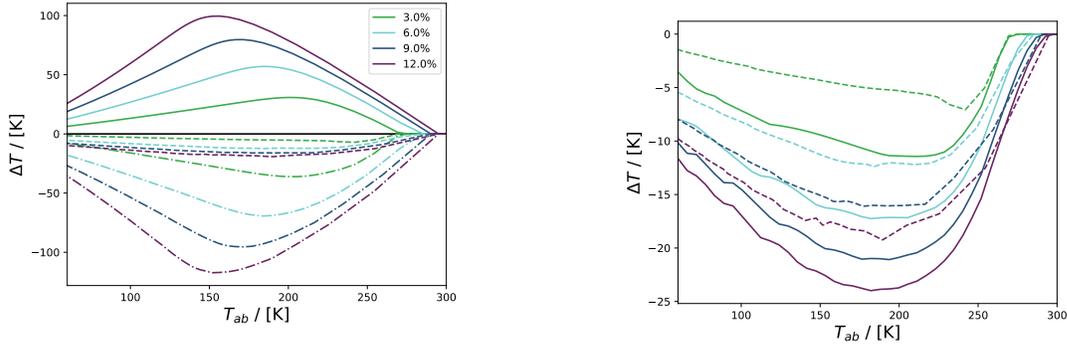


(a) Effective temperature changes arising from different sources. Latent heat (solid), convective inhibition (dashed-dotted), and the net effect (dashed).

(b) Net temperature differences, for an initially dry adiabetic atmosphere (solid) and an initially moist adiabetic upper layer (dashed)

Figure 4.3: $\Delta T(T_{ab})$ for different envelope abundances of methane q_{max} between 2-5%, where $\Delta T \equiv T_e - T_{ab}$. The dashed (net ΔT) curves are the same in (a) and (b).

Finally we must consider the deposition of sunlight. The primary absorber of sunlight in Uranus' troposphere is methane vapor [Marley and McKay, 1999], while the primary absorber of infrared light is hydrogen collisions. At 1 bar with 2% methane for example, accounting for methane absorption plus Rayleigh scattering vs. thermal absorption by hydrogen, the ratio between visible to thermal infrared opacity is approximately $\kappa_{\odot}/\kappa_t < 10^{-2}$ averaging over a broad band in wavelength. This value of course is not unique; there are windows at certain wavelengths, the methane abundance changes rapidly with depth, and this simple calculation neglects absorption by haze and cloud particles. Nevertheless κ_{\odot}/κ_t is sufficiently small that we can plausibly argue that sunlight penetrates significantly beyond the 1 bar level, and most of the sunlight is absorbed deeper in the atmosphere. This can be parameterized by simply arguing some fraction f of sunlight is absorbed above the cloud level, and $1 - f$ is absorbed below. In principle this procedure accommodates any value of f , but for our purposes for simplicity we approximate using the limiting cases $f \rightarrow 0$ for



(a) Effective temperature changes arising from different sources. Latent heat (solid), convective inhibition (dashed-dotted), and the net effect (dashed).

(b) Net temperature differences, for an initially dry adiabatic atmosphere (solid) and an initially moist adiabatic upper layer (dashed)

Figure 4.4: $\Delta T(T_{ab})$ for different envelope abundances of water q_{max} between 3-12%, where $\Delta T \equiv T_e - T_{ab}$. The dashed (net ΔT) curves are the same in (a) and (b).

the shallow (~ 1 bar) methane condensation level, while $f \rightarrow 1$ for deep (~ 100 bar) water clouds. The difference for the stable layer is that in radiative equilibrium, the flux through the stable layer must balance with $\sigma(T_e^4 - fT_0^4)$ as explained above.

4.3.3 Meteorological implications

The equilibrium solution will be stable if the virtual potential temperature θ_v is monotonically decreasing with increasing pressure between p_1 and p_2 :

$$\theta_v(p, T) = T_v(p, T) \left(\frac{p_0}{p} \right)^{\nabla_{ab}}, \quad (4.10)$$

$$T_v(p, T) = T (1 - q(1 - \epsilon))^{-1}. \quad (4.11)$$

In the equilibrium cases discussed here, this condition is always satisfied (see dashed red curve in Figure 4.1). This result should be fully general; it does not depend on, for example, the choice of atmospheric opacity. In order to trigger a convective instability, the upper layer would need to become more dense than the deep layer

(note in this section we use the word “dense” in the virtual potential temperature sense, i.e., accounting implicitly for adiabatic expansion/compression). In order to accomplish this by cooling, the upper layer would need to cool such that the bottom of the upper layer satisfies $q < q_{crit}$, so that further cooling makes the gas mixture more dense rather than less dense. However, as soon as q becomes infinitesimally less than q_{crit} , it will be more dense than the material in the stable layer directly beneath it, even while remaining less dense than the well mixed gas in the deep layer. This will cause a small convective instability wherein a portion of the stable layer is eroded into the upper layer, causing the stable layer to thin and restoring the upper layer to satisfy $q = q_{crit}$ at a new pressure level. This will happen in all cases where the stable layer is of finite thickness. Therefore, in order to trigger a convective instability with the deep layer, the stable layer must vanish completely, reducing to a compositional discontinuity between the upper and deep layers. Such a scenario in the limit of heat transport by thermal conduction and an inviscid fluid results in an infinitesimal thermal boundary layer with an infinitesimal temperature discontinuity. In the optically thin radiative transfer case, the situation is somewhat more subtle, because rather than heat flux diverging to infinity, it converges on a finite value (in the two stream approximation: the upward heat flux from below minus the downward heat flux from above). Nevertheless, if there is a significant temperature difference between the upper and deep layer (as would be required in order to trigger a convective instability), this finite value is orders of magnitude larger than the luminosity of the planets in question, for any appreciable temperature discontinuity. This result contrasts with previous findings [Li and Ingersoll, 2015], which posited the system may behave as a relaxation oscillator when the upper layer cools sufficiently to become over-dense and trigger a convective instability. This previous study did not explicitly account for radiative transfer across the stable layer, instead dynamically cooling from above and treating the stable layer as a perfect insulator. In our case, we find instead that the

profile relaxes into a state of global radiative-convective equilibrium; eventually the upper layer stops cooling as the heat it loses to space balances with the heat radiated across the stable layer.

After solving for the equilibrium atmospheric profile, it is possible to compute the convective available potential energy (CAPE), as well as the activation energy required to disrupt the equilibrium stable layer. Both quantities can be computed in the same way, using

$$E = \int_{p_a}^{p_b} R_d(T_{v,f} - T_{v,i}) \frac{dp}{p}, \quad (4.12)$$

where $T_{v,f}$ is the final virtual temperature profile (for example, a moist adiabat), $T_{v,i}$ is the initial virtual temperature profile (e.g., in radiative-convective equilibrium with a stable layer), and p_a and p_b are bounding pressures that satisfy $T_{v,i} = T_{v,f}$. The blue and red dashed curves in Figure 4.1 provide a visual example for how this calculation can be done. This computed activation energy in units of energy per mass, can be converted into updraft velocities required to disrupt the stable layer. In the case of Saturn, if $q_{crit} \sim q_{max}$ as suggested in [Li and Ingersoll, 2015], then the activation energy required to disrupt this stable equilibrium is quite small, requiring updraft velocities of only a few meters per second. Therefore the basic premise of [Li and Ingersoll, 2015] can still be valid, although it may require some additional mechanism to jump-start the process, for example a strong updraft or entrainment erosion as described in [Friedson and Gonzales, 2017]. In the case of Uranus and Neptune with methane mixing ratios further from the critical value, the activation energy necessary to initiate a convective instability is correspondingly larger, requiring updraft velocities of order tens of meters per second, far larger than expected convective velocities. However, if such a disruption were able to occur by some anomalous updraft, the resulting storm would be extremely energetic, with CAPE exceeding 10 J/g, substantially larger than even the most extreme weather

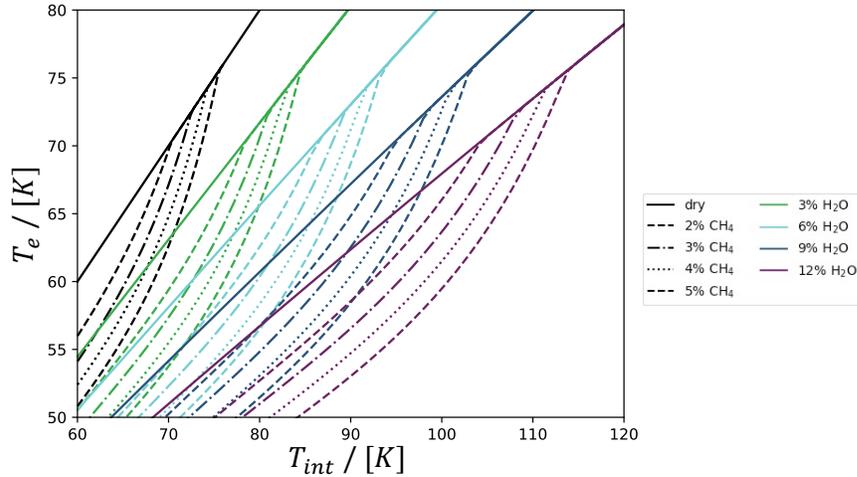


Figure 4.5: Apparent effective temperature T_e as a function of internal equivalent effective temperature T_{int} considering both methane and water. Colors correspond to water abundance, increasing downward from 0% to 12% water. The curve style corresponds to methane abundance, increasing downward from 0% to 5% abundance.

events on Earth.

4.3.4 Interior implications

Interior models of giant planets generally assume a 1-bar equivalent temperature that sets the internal entropy of the planet. This is usually done using the atmospheric temperature, and corrected for any non-adiabatic behavior in the atmosphere. Non-adiabatic models for Uranus' and Neptune's interiors and thermal evolution have recently been carried out (e.g., [Nettelmann et al., 2016], [Vazan and Helled, 2020], [Scheibe et al., 2021]), finding self-consistent solutions to Uranus' contemporary heat flow. Our results are still relevant to many of these models. The so-called thermal boundary layers at depth from [Nettelmann et al., 2016] and [Vazan and Helled, 2020] retain an adiabatic convective envelope set by a 1-bar equivalent temperature. Furthermore, the U-1 and U-2 models from [Vazan and Helled, 2020] involve a convective envelope of homogeneous composition, whose temperature profile is also set using a

1-bar equivalent temperature. Although our results make no direct statement about the behavior of the deep interior of the planet that may include extended regions of static stability, it still applies to the thermal evolution of these upper envelopes. Because Uranus and Neptune possess magnetic fields, and the latter possesses a substantial internal heat flow, we deduce that both must be convective at depth, and that this adiabatic description is probably relevant for at least some fraction of the interior of Uranus and Neptune. The primary source of non-adiabatic behavior in the atmosphere is condensation, usually accounted for by the sub-adiabatic gradients caused by latent heat of condensation. In Figure 4.5, we show the apparent effective temperature T_e against the internal equivalent effective temperature T_{int} for various envelope condensate abundances. If convective inhibition does occur, it will largely cancel out the effect of moist adiabaticity. Therefore, using a dry adiabat to guess the 1-bar equivalent temperature is a better approximation than accounting for latent heat alone but neglecting convective inhibition. A better approach could be to use the analytic scaling relationships in Section 2 to estimate the magnitude of these effects. The best approach would be to explicitly model the effect of convective inhibition, using methods from this work or [Leconte et al., 2017].

4.4 Evolutionary model

We present an adiabatic thermal evolution model of Uranus and Neptune. As discussed, treating the interior as adiabatic is probably inappropriate for the ice giants [Helled et al., 2020]. Nevertheless it provides a convenient framework to understand the effect of methane condensation on these planets' thermal histories in the absence of an accepted interior model. For an adiabatic model, we assume the total heat content of the planet's interior to be a linear function of its adiabatic equivalent effective

temperature T_{ab} :

$$\int_{M_{\min}}^M c_p T dm = A \bar{c}_p M T_{ab}. \quad (4.13)$$

One way of thinking about this equation is to imagine a small set of layers, or possibly even one layer, in the form of concentric shells, each of which is isentropic and homogeneous but of different composition to neighboring layers, with negligible thermal boundary layers between them as would be fluid dynamically expected for a low-viscosity system. Beneath this set of shells there could be a region, possibly a substantial fraction of the planet, where there is a compositional gradient and therefore inefficient convective transport. This deeper region would not contribute to A or to the resulting thermal evolution of the planet because it stores primordial heat. In reality, there would be non-zero thermal diffusion from a stably stratified interior portion of the planet and its convective envelope, slowing down planetary cooling. However, if the diffusion timescale for the planet is longer than the age of the solar system, then this contribution would be small. It is not guaranteed that the diffusion timescale is in fact longer than the age of the solar system. It depends on the (unknown) thermal transport properties of the (unknown) compositional constituents of the ice giants' interiors, and is further complicated by the possibility of thermal transport by double diffusive convection. Therefore we acknowledge this description of the interior evolution is imperfect, but it does at least approximately describe a wide variety of possible interior behaviors, and provides a convenient framework to self-consistently assess the relative influence of convective inhibition on thermal evolution while remaining agnostic about the details of the ice giants' interior structures. A full description of Uranus' and Neptune's thermal evolutions would require a detailed interior model. Nevertheless under our assumptions, the parameter A is approximately constant through time because the Gruneisen parameter is rather insensitive to temperature, and its value is set by the fraction of the total mass that is fully convective. Some fraction of the planets must be convective in order to generate

their observed magnetic fields.

The rate of cooling depends on the apparent effective temperature T_e , while the heat content of the bulk of the interior is linearly related to the adiabatic equivalent effective temperature T_{ab} . The equation governing thermal evolution is then

$$4\pi R^2 \sigma (T_e^4 - T_0^4) = -A\bar{c}_p M \frac{dT_{ab}}{dt}, \quad (4.14)$$

where T_0 is the equilibrium effective temperature of the planet with sunlight if there were no internal heat source. The radius R is treated as constant because we are concerned with most of the evolution where the body is degenerate, not any early very hot phase. The steady increase in solar luminosity (i.e., time variation of T_0) is ignored. In order to solve the thermal evolution equation, we need an explicit relationship between T_e and T_{ab} . This is done using the method from Section 4.3, with the results for methane in Figure 4.3 and for water in Figure 4.4. In general the relationship between T_e and T_{ab} depends on the condensate bulk interior abundance q_{max} . Without condensation, the relationship is $T_e = T_{ab}$. In this case, Equation 4.14 can be non-dimensionalized into the canonical form

$$\frac{x^4 - x_0^4}{1 - x_0^4} = -\tau_K \frac{dx}{dt}, \quad (4.15)$$

where $x \equiv T_e/T_e^{(\theta)}$, $T_e^{(\theta)}$ is the apparent effective temperature today, and $x_0 = T_0/T_e^{(\theta)}$. The Kelvin timescale $\tau_K = \frac{A\bar{c}_p M T_e^{(\theta)}}{4\pi R^2 \sigma} (T_e^{(\theta)4} - T_0^4)^{-1}$ scales how long it takes to cool to $T_e^{(\theta)}$ from an initial arbitrarily hot state. In the asymptotic case $T_e^{(\theta)} \gg T_0$ (not true for Uranus!), this is about $\tau_K/4$, but can be a different fraction of τ_K in general.

Accounting for condensation, Equation 4.14 becomes

$$\frac{dx}{dx_{ab}} \frac{x^4 - x_0^4}{1 - x_0^4} = -\tau_K \frac{dx}{dt}, \quad (4.16)$$

where $x_{ab} = T_{ab}/T_e^{(\theta)}$. The difference between Equations 4.15 and 4.16 then straightforwardly demonstrates the effect of convective inhibition by condensation on the planets' thermal evolution: it alters the rate of cooling by a factor of $\frac{dx}{dx_{ab}}$. As we will see, this factor can be greater than or less than unity. This means the effect can either speed up or slow down the rate of change of the planets' apparent effective temperature at different points in its thermal history. This is especially important for understanding the results for water.

The fact that Equation 4.16 retains the Kelvin timescale τ_K makes this formulation especially convenient. This allows us to directly compare the fraction of that timescale that a given evolutionary model takes to cool from arbitrarily hot bodies to their current temperatures for different assumptions of the condensate abundance q_{max} . Leaving the the effect in terms of the Kelvin timescale allows our results to be roughly independent of accurate interior models, because τ_K implicitly encodes an arbitrary interior model. The results for methane are shown in Figure 4.6, and for water in Figure 4.7. The results for methane are relatively straightforward; for the early stages of Uranus and Neptune's thermal histories, the effect of methane is unimportant, because the atmosphere is warm enough that methane does not condense anywhere. As the atmosphere cools, methane begins to condense, at first in the stratosphere above the radiative convective boundary. As cooling continues, convective inhibition begins to extend the radiative-convective boundary downward, as superadiabatic gradients can be stable. At this point, the measured effective temperature T_e departs from its adiabatic equivalent T_{ab} , causing the effective temperature to drop faster than the interior is cooling. As cooling continues, the layered system described in Section 4.3 emerges, and perhaps persists today [Guillot, 1995].

The case of water, shown in Figure 4.7, the behavior is more subtle. In this case,

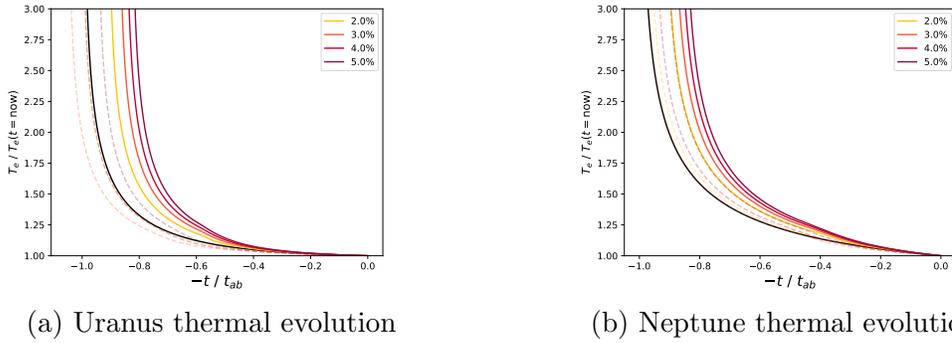


Figure 4.6: Thermal evolution model for Uranus and Neptune, with different colored curves representing different methane abundances. The x-axis is the time before the present day, scaled to the cool-down time in the dry adiabatic (no condensation) case. The y-axis is x from Equation 4.16. Line styles and colors are identical to Figure 4.3.

thermal evolution is actually slowed down compared to the adiabatic case. This contrasts with previous findings [Kurosaki and Ikoma, 2017], which considered the effect of moist adiabaticity (i.e., latent heat) but did not quantify the effects of convective inhibition. If we consider latent heat only, we obtain results in good agreement with this previous study. In their case, thermal evolution is sped up, because the atmosphere initially remains warm while the interior cools. This allows the planet to lose heat efficiently when condensation first occurs, speeding up evolution. Our findings demonstrate that the effect of convective inhibition overwhelms the effect of moist adiabaticity, so that our story is the opposite. Early on, as condensation occurs, we find the atmosphere cools faster than the interior, reducing cooling efficiency. Later on, $\frac{dx}{dx_{ab}}$ from Equation 4.16 becomes less than unity, as demonstrated by the negative slope at low temperatures in Figure 4.4. Therefore in the case of water, condensation early in the ice giants' thermal histories caused the effective temperature to drop faster than the internal temperature, analogous to what happened with methane condensation more recently. However, this temporary speedup of $\frac{dT_e}{dt}$ coincides with a loss of luminosity, slowing down the rate at which the interior loses heat. Then, over subsequent evolution, the interior cools inefficiently, and in recent history the effec-

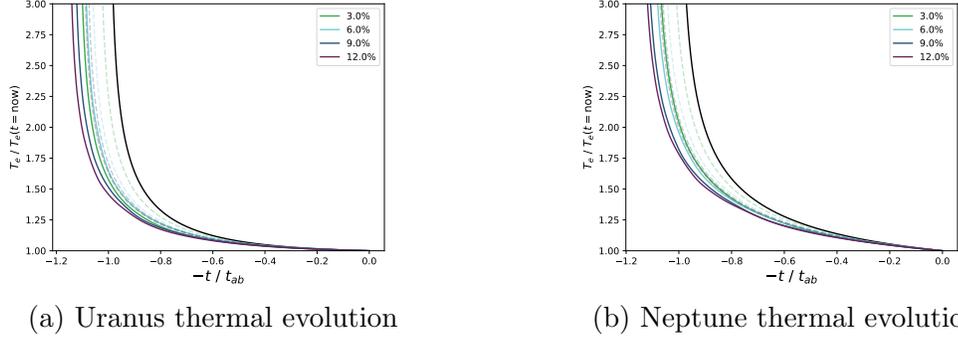


Figure 4.7: Thermal evolution model for Uranus and Neptune, with different colored curves representing different water abundances. Axes are identical to Figure 4.6. Line styles and colors are identical to Figure 4.4.

tive temperature changes *slower* than the internal temperature. The net effect is a cool-down time that is longer than the dry adiabatic case.

Because the water and methane cloud decks are well separated, the superposition of the two effects is straightforward. Immediately beneath the methane cloud deck, the behavior can be accurately modeled as a dry adiabat, because the water mixing ratio is so small at these relatively low temperatures. Therefore the results of modeling the whole atmosphere with both cloud decks explicitly is virtually identical to using $\Delta T_{tot} = \Delta T_{CH_4} + \Delta T_{H_2O}$ from Figures 4.3 and 4.4.

4.5 Discussion

Provided $q_{max} > q_{crit}$, when the planet cools to a temperature low enough for condensation, convection can be interrupted. At this point the apparent effective temperature departs from the internal equivalent effective temperature, by the mechanism described in Section 4.3. We can solve for the equilibrium configuration to derive the apparent effective temperature T_e as a function of q_{max} and T_{int} . These results are shown as the dashed-dotted lines in Figure 4.3 for methane and Figure 4.4 for water.

Interior modelers should bear in mind that the internal 1-bar equivalent temperature may depart from simple adiabatic extrapolation of the troposphere by nearly a significant factor (see Section 4.3.4 and Figure 4.5).

This behavior leads to the evolutionary behavior observed in Figure 4.6. Before condensation occurs, the planet cools normally. Upon the onset cloud formation, the apparent effective temperature drops rapidly. However, upon reaching the minimum, the apparent effective temperature actually begins to decrease more slowly than the fully adiabatic case. This effect is present in Figure 4.3 but is more apparent in Figure 4.4. The net effect for the ice giants is a net speedup of thermal evolution for methane, and a net slowdown for water. The magnitude of this speedup or slow down can be no more than 15% in either case, assuming 5% methane molar abundance or 12% water molar abundance. Both effects can occur simultaneously, and superimpose straightforwardly because their cloud decks are well separated.

As Neptune continues to cool, methane will begin to behave similarly to water, exhibiting a local minimum in $\Delta T(T_{ab})$. This local minimum is seen most clearly for water in Figure 4.4 around $T_{ab} = 180\text{K}$, but can also be seen for the dashed curves in Figure 4.3 around 75K, and would be present in the solid curves if the x-axis extended to lower temperatures. As the planet continues to cool below this local minimum, the slope of $\Delta T(T_{ab})$ becomes negative, and the rate of change of the thermal state of the atmosphere slows. Consider the implications of this for methane clouds near the 1-bar level. This state persists for longer from a thermal evolution perspective than an arbitrary/random thermal state. That is, this state is a local minimum in $\Delta T(T_{ab})$, meaning the planet reaches this state faster than it would if it were cooling adiabatically, and leaves this state more slowly than it would if it were cooling adiabatically. Therefore these planets will spend a longer portion of their thermal histories in the

state where the cloud level is $\sim 1 - 10$ bars than they would in a thermal evolution model that does not consider convective inhibition by condensation. Perhaps this consideration renders the surprising similarity of Uranus and Neptunes' atmospheres' shallow temperature structures despite their vast difference in insolation somewhat less improbable than it first appears.

We must consider whether this atmospheric structure is compatible with existing data, especially Voyager radio refractivity data. The current data has been shown to be consistent with many different models, including subadiabatic, adiabatic, moist adiabatic, and superadiabatic temperature gradients [Helled et al., 2020]. The data has also been shown to be compatible with a wide range of temperature structures and methane abundances [Lindal et al., 1987] [Lindal, 1992]. The data itself shows a layer of rapidly varying refractivity near the condensation level, generally interpreted to be methane clouds [Lindal et al., 1987] [Lindal, 1992][Marley and McKay, 1999]. Another interpretation of the same data supports a layer of superadiabatic temperature lapse rate in the cloud-forming regions of these planets [Guillot, 1995]. In general, our understanding of the thermal structure of the ice giant atmospheres is incomplete, as the results from Voyager 2 refractivity data are model dependent, with a particular degeneracy between assumed methane enrichment and temperature structure. In order to disentangle these variables and have a more confident understanding of these planets' atmospheres' thermal structures, we must return with a mission. It should be a priority for a future mission to independently measure methane abundance and temperature, perhaps with entry probes or a well designed microwave radiometer experiment.

These general findings do not consider the long term stability of stable layers in the atmosphere. As long as the stability timescale is greater than the relaxation

timescale for a stable layer, the results should approximately reflect reality. However, the stability timescale is poorly constrained [Friedson and Gonzales, 2017]. If it is sufficiently short, this could further complicate the dynamics. If that condition is satisfied, then even if stable layers are intermittently interrupted by massive internal plumes, large meteor impacts, or instability due to long term erosion by entrainment, they will reform again on geologically short timescales ($\sim 100\text{yr}$). Therefore the thermal evolution will be governed primarily by the equilibrium state, and not possible intermittent periods of enhanced activity. Intermediate states where the equilibrium configuration is thinned over time but not totally destroyed by entrainment erosion would in general reduce the magnitude of $\Delta T(T_{int})$, so the findings in this paper should be considered an upper bound. Furthermore, we use a highly simplified thermal evolution model, not considering changes in planetary radius or explicitly accounting for the effects of non-adiabaticity at depth. Seasonal variations in insolation were not included in the model, as these variations average out over geologic time. However, seasonal variations have been shown to create local temperature variations of order 10K [Orton et al., 2007], comparable to the magnitude of the effect of convective inhibition by methane. The possible dynamical and evolutionary consequences could be the subject of future work. Our atmospheric model also did not explicitly include the condensate opacities, and may therefore not capture possible feedback mechanisms. We discuss further the possible effects of opacity variation due to condensation in the following paragraphs. For these reasons, this work should be considered exploratory, and further work is needed in order to more confidently establish the thermal histories of the ice giants while accounting for convective inhibition.

Here we must include a discussion about the effects of opacity variation due to condensation, which are not considered in this model but which are certainly impor-

tant for a fully complete understanding of Uranus and Neptune’s thermal states and thermal histories, and has been considered explicitly by prior works, e.g., [Kurosaki and Ikoma, 2017]. The variation of opacity affects our results in two important ways: first, by changing the radiative-convective boundary as vapor condenses out of the stratosphere; and second, by affecting radiative transfer within and across the layer of stable stratification caused by convective inhibition.

We begin by discussing the stratospheric effect of opacity variations due to condensation. Water and methane are both more opaque than hydrogen in the thermal infrared, therefore as the planet cools and these volatiles begin to condense and rain out of the stratosphere, the stratosphere becomes more transparent and the radiative-convective boundary deepens. At fixed effective temperature, the temperature at the radiative-convective boundary is relatively unchanged, therefore decreasing the opacity of the stratosphere has the net effect of decreasing the entropy of the troposphere at fixed effective temperature. Therefore during this stage, the temperature of the troposphere is cooling faster than the effective temperature of the planet as the stratosphere extends downward. Convective inhibition only begins to become relevant when the stratosphere has cooled sufficiently such that the radiative-convective boundary has a lower vapor mixing ratio than the bulk abundance. By the time this occurs, the bulk of the stratosphere is cooler than the radiative convective boundary by approximately a factor of $2^{1/4}$, and therefore relatively dry due to the highly sensitive dependence of saturation vapor pressure on temperature. So the important stratospheric effect due to opacity variation we have just described qualitatively, essentially predates the onset of convective inhibition. This allows us to neglect these dynamics in our context, although we caution the reader that a fully realistic consideration of the effects of condensation must also include the effects of opacity variations, which are important.

The second effect of opacity variation is on the radiative transfer across the stable layer. To estimate the importance of this effect, we modified our method so that we increased the opacity of the deep layer and stable layer (see Figure 4.2) by an order-of-magnitude. This changes our results for contemporary methane clouds by no more than 2%, and for contemporary deep water clouds by $<0.01\%$. The effect for water clouds is larger earlier in its evolution, but is always $<2\%$. In either case, the direction of this consideration is to increase the magnitude of ΔT . The reason this matters less for deep clouds is because the opacity is already very large, and the effect due to deep clouds is accurately approximated by Equation 4.2, which assumes a high opacity limit. For shallower clouds where the details of thermal transport are more relevant, it affects the results, but only as a relatively small correction even assuming a very large order-of-magnitude change in opacity.

If there are indeed layers of static stability in the troposphere or deep atmosphere of Uranus and/or Neptune, then they should support gravity waves. Whether we expect gravity waves to be excited, what their general characteristic would be, and whether they could be detected from space (for example using an Doppler imager) is a subject worthy of future theoretical consideration.

Whatever the uncertainties about the specifics, the basic physical mechanism is likely to be important in the ice giants because of their highly enriched atmospheres. There may be additional stable layers, for example a silicate cloud level beneath the water cloud level, or a sulfide/ammonia cloud level. We focus on only two in this work to demonstrate the general principle without getting bogged down by largely unconstrained assumptions about the envelope enrichment in each species. However, the intuition we build here for methane and water can be straightforwardly applied

to other cloud levels using exactly the same method. This method is also likely to be applicable to the majority of exoplanets, ranging from super-Earths and water worlds with hydrogen envelopes, to metal-enriched gas giants. It is clear from this work that thermal evolution and internal thermal structure may be profoundly influenced by convective inhibition by condensation. Any complete model of thermal evolution or internal structure is advised to consider convective inhibition.

The following link directs to the published version of this chapter:

<https://iopscience.iop.org/article/10.3847/PSJ/ac091d>

Chapter 5

The cooling and interiors of super-Earths

I have come to lead you to the
other shore; into eternal
darkness; into fire and into ice.

Dante Alighieri

Summary

In this section we will propose a new model for super-Earth internal structure and evolution, based on thermodynamic arguments about the coexistence of hydrogen and silicate vapor. We argue that for hydrogen envelope masses exceeding roughly $10^{-3} - 10^{-2} M_{\oplus}$, the convective contact between the envelope and core may shut down. The core then cools inefficiently, potentially remaining in a high-entropy supercritical state for long timescales. The core cooling time scales with envelope mass, cooling over (Gyr) geologic time for lower mass ($\sim 10^{-3} - 10^{-2} M_{\oplus}$) envelopes, and longer than the age of the universe for higher mass ($> 10^{-2} M_{\oplus}$) envelopes. We predict that super-Earth internal luminosity decreases with increasing envelope mass, and should

be roughly independent of planetary system age for higher mass envelopes. This model predicts small internal heat flow (of order Earth’s internal heat flux today) even at early times. If core-powered mass loss is indeed a dominant mechanism for atmospheric escape, our model predicts it cannot operate efficiently for envelopes more massive than $\sim 10^{-2}M_{\oplus}$. Overall we find that super-Earths, rather than mere larger cousins to our home world, may be truly and profoundly alien with no concrete “surface” to speak of.

5.1 Introduction

Super-Earths were among the very first observed exoplanets, with Poltergeist and Phobos found orbiting the pulsar Lich in 1992 [Wolszczan and Frail, 1992]. The structure and composition of these bodies, intermediate in mass between Earth and Neptune with no analogues in the solar system, has only gradually come into focus since. The deluge of data from the Kepler mission represents the greatest advancement in our understanding of these bodies, making clear that planets do not have to become very much larger than Earth before they begin to rapidly increase in size [Borucki et al., 2011]. These bodies’ mean densities can be described by a composition dominated by water [Seager et al., 2007], although the dearth of observed high mass, high density planets suggests many are likely to be a mixture of primarily silicates and hydrogen. Canonical models usually involve an extended gas envelope in thermal contact with a magma ocean (e.g., [Ginzburg et al., 2016] [Vazan et al., 2018]). This model, involving a silicate core with an overlying gas envelope, will be the focus of this chapter. In particular, we will discuss the implications of convective inhibition (see e.g., [Guillot, 1995] [Friedson and Gonzales, 2017] [Leconte et al., 2017] [Markham and Stevenson, 2021]) as a result of the condensation of silicate vapor.

Recent observations of Jupiter and Saturn indicate both planets possess an extended diffuse core [Stevenson, 2020] [Mankovich and Fuller, 2021], indicating that significant metal enrichment may go alongside gas accretion during planetary formation. In this chapter we argue that if the initial envelope of super-Earths is substantially enriched in silicates, as recent formation models contend [Bodenheimer et al., 2018] [Brouwers and Ormel, 2020], this can meaningfully complicate our understanding of super-Earth structure and evolution. We argue that such enrichment leads naturally to a layer of static stability at depth such that the envelope convectively decouples from the core, if the envelope mass exceeds $\sim 10^{-3} - 10^{-2}M_{\oplus}$. This mechanism is similar to convective inhibition by condensation [Guillot, 1995] [Friedson and Gonzales, 2017] [Leconte et al., 2017] [Markham and Stevenson, 2021], and we extend these arguments to apply more generally to phase separation of coexisting mixtures. Planets that satisfy these criteria are likely to be among the most common in the universe [Borucki et al., 2011]. A layer of static stability at depth dramatically reduces super-Earth luminosity at early times, so that the core loses heat extremely inefficiently. Because of the expected low thermal conductivity under the relevant conditions (see Section 5.3.1), the layer of static stability is likely to be thin compared to the full extent of the atmosphere, and the temperature change across the stable layer can be thousands of Kelvin. In the thin stable layer limit, we can precisely predict a super-Earth’s steady state internal heat flow based on the mass of its envelope. For massive envelopes with very low luminosity, there can be a non-negligible correction to the thin stable layer approximation.

We further argue that the core of these planets can be a supercritical mixture of gases and silicates if the mass of its atmosphere exceed about $10^{-3}M_{\oplus}$, and that this high entropy state can persist on geologically long timescales. This model predicts that the luminosity of planets drops rapidly after isolation, but then remains

nearly constant for geologically long timescales. This is an exploratory work focused primarily on constraining the relevant orders of magnitude, but we comment on the potential import of these findings.

A supercritical core could mix with hydrogen in all proportions, potentially leading to less dense regions of the core, or extended stable compositional gradients depending on its formation conditions. Such a scenario has been explored in prior works [Bodenheimer et al., 2018] [Brouwers and Ormel, 2020], although these works did not account for the possibility of convective inhibition due to the coexistence of hydrogen with silicate vapor. The existence of a polluted “outer core” (following the verbiage of [Bodenheimer et al., 2018]) or “inner envelope” (following the verbiage of [Brouwers and Ormel, 2020]) is not the subject of this work, although we do find that such an interior configuration can persist for longer than the age of the universe.

Exoplanet observations demonstrate a bimodal distribution of planetary radii, with a relative absence of planets in the size range between $1.5 - 2R_{\oplus}$. The prevailing theories to explain this observation posit rapid atmospheric mass loss early in a planet’s evolution, either by intense stellar radiation [Owen and Wu, 2017] or by the heat flow originating from a hot core [Gupta and Schlichting, 2019], or a combination of both effects. This work does not question either result, but adds a caveat to the core powered mass loss mechanism. If the atmospheric mass loss rate is set by heat flow from the core, we argue the magnitude of this heat flow may be limited by convective inhibition. As we will argue, the luminosity of planets with envelopes less massive than $\sim 10^{-3}M_{\oplus}$ is not significantly affected by the inhibition of convection. We note that if the luminosity of the core is indeed a major source of atmospheric mass loss as has been suggested [Gupta and Schlichting, 2019], that we expect this process to be considerably less efficient for higher envelope masses. We note this as

a possible future application of the principles we will outline in this chapter, but do not explore this particular possibility in detail, leaving it instead as a potentially ripe subject for related future work.

This work more generally predicts that luminosities for sub-Neptune/super-Earth class planets with envelope masses exceeding about 2% M_{\oplus} will be roughly independent of the planetary system's age. For planets with lighter envelopes, between 0.1-1% M_{\oplus} , this state can persist for billions of years, but may evolve over geologic time. Subsequent evolution would involve gradual hydrogen degassing from a polluted core.

5.2 Convective inhibition by silicate vapor

This work largely follows the logic of convective inhibition by condensation in hydrogen atmospheres [Guillot, 1995]. In the previous chapter, we investigated a situation wherein the atmosphere is dominated by hydrogen, with some pollution by volatile condensible species. In this chapter, we investigate the same mechanism operating in the opposite extreme—a relatively low-mass hydrogen atmosphere with an infinite reservoir of volatile condensibles. We note that the difference between a super-Earth planet and a Neptune-like planet is subtle. In this case, rather than the mixing ratio trending toward some value $q_{max} \ll 1$, the mixing ratio approaches unity. If the relevant condensing species are silicates, this situation roughly describes super-Earths, as a deep magma ocean or silicate vapor core provides an infinite reservoir of silicate condensates available to dissolve in the atmosphere.

First we will present a generalized argument for convective inhibition involving a mixture of phases. In general, convection will be shut off for a negative temperature

gradient (that is, temperature increasing with depth) if the coefficient of thermal expansion becomes negative. This is a familiar enough phenomenon on Earth in the case of water, where the local minimum of density is above the freezing point. In the case of water, as you cool from the top convection is shut off as cooler, lower density water is stable against convection by warmer, higher density water—hence why lakes freeze from the top down. We therefore begin by considering two coexisting species that we call “dry” (subscript d) and “condensate” (subscript c). At relatively cool temperatures (“cool” refers to the relative magnitude of $k_B T$ and inter-molecular bond energies), the substances are nearly unmixed and the coexisting phases are nearly pure. Above the critical temperature, both phases become indistinguishable and can mix in all proportions. In the limits of low pressure and domination of the dry species by molar abundance, the partial pressure of the condensate in the gas phase is described by an Arrhenius relationship $\propto \exp\left(-\frac{L}{k_B T}\right)$ where L is the molar latent heat of vaporization. Meanwhile, the abundance of the dry species in the liquid phase is described by Henry’s Law, negligible to first order. In this case, the critical mixing ratio is the ratio described in Chapter 3 and prior works.

This approximation breaks down at higher pressures and temperatures likely to be relevant on super-Earths. At higher temperatures, one cannot extrapolate the saturation vapor pressure to pressures approaching the hydrostatic pressure in a coexisting multi-component system—one must use an empirical phase diagram to determine the coexistence under these conditions. At higher pressures, the relative abundances of both species become more symmetrical, and the approximations of both saturation vapor pressure and Henry’s Law (or negligible pollution) break down. We therefore present the following derivation that makes no explicit assumption about the nature of the relative abundances within the coexisting phases, and seek the point at which the coefficient of thermal expansion α becomes negative. At this point, convection

is shut off. This derivation still makes the assumption that the gas phase behaves like an ideal gas, and that the system is in thermodynamic equilibrium (saturation). The latter especially is a nontrivial assumption as it relies on transport (e.g., diffusive, turbulent) of silicates into the gas phase, but prior works demonstrate that even imperfect transport can nevertheless exhibit convective inhibition [Friedson and Gonzales, 2017] [Leconte et al., 2017].

We begin with the simple ideal gas case. If we increase the temperature of a parcel of gas at temperature T by an increment δT at constant pressure, its specific volume per molecule changes by an amount independent of composition or molecular mass, $\delta v/v = \delta T/T$. If the composition is fixed this yields the familiar coefficient of thermal expansion $\alpha = 1/T$. We now inspect the same temperature increment δT , but do so while the gas remains saturated with some condensing vapor. For simplicity we consider a fixed mass of gas whose total number of molecules changes from N to $N + \delta N$ as additional condensate vapor molecules enter the system. The corresponding change in the molar ratio x is δx . By conservation of mass, the total quantity $N[x\mu_c + (1-x)\mu_d]$ must remain invariant before and after the temperature increment. Therefore $\delta x(\mu_c - \mu_d) = -\delta N/N[x\mu_c + (1-x)\mu_d]$. For convective inhibition, we seek the point at which $\alpha \rightarrow 0$, i.e., $\delta(Nv) = v\delta N + N\delta v = 0$. Under the ideal gas approximation, we obtain the criterion for convective inhibition,

$$\frac{\delta T}{T} = \delta x \frac{\mu_c - \mu_d}{x\mu_c + (1-x)\mu_d}. \quad (5.1)$$

If we assume saturation vapor pressure obeys an Arrhenius relationship (as it will for low pressure and temperature), $p_s/p \equiv x \propto \exp\left(-\frac{L}{k_B T}\right)$, then at constant pressure $\frac{dx}{dT} = -\frac{L}{k_B T^2} x$. In general if we assume $\frac{dx}{dT}$ is linear in x , Equation 5.1 is equivalent to

the statement

$$x_{\text{crit}}^{-1} = \left(\frac{d \ln x}{d \ln T} - 1 \right) (\epsilon - 1), \quad (5.2)$$

where x_{crit} is the so called “critical” concentration where the coefficient of thermal expansion is zero. We use the word “critical” here in the sense of the minimum mixing ratio to initiate convective inhibition following the language of prior works (e.g., [Guillot, 1995] [Leconte et al., 2017] [Friedson and Gonzales, 2017]), but note that this critical mixing ratio is distinct from the critical temperature and critical pressure at which the fluid becomes a supercritical state of matter. For all saturated concentrations above this limit, negative temperature gradients (temperature increasing with depth) will be stable against convection. If $x(T)$ obeys an Arrhenius relationship, then Equation 5.1 yields an expression identical to Equation 4.1 in Chapter 3 for x_{crit} (q_{crit} in the notation of that chapter).

Now that we have identified the mixing ratio at which convection is shut off, what then will be the mixing ratio at which convection reasserts itself? In the previous chapter, this occurs after the mixing ratio reaches its “bulk” value called q_{max} . However, this bulk value as such only applies in the limit where the system is dominated by gases. In the limit we consider now, where there exists an infinite reservoir of condensible species (for example in an ocean), what then will be the mixing ratio at depth? The naive answer is unity, if one is considering simply the saturation vapor pressure. Under this line of reasoning, the mixing ratio of condensate will continue to increase as temperature increases, until the saturation vapor pressure of the condensate approaches the hydrostatic pressure. Then a stable gas atmosphere would possess a stable layer that terminates at an ocean with condensate mixing ratio unity. In reality, it depends on the local pressure.

Figure 5.1 shows empirical coexistence curves of a gas/liquid mixture of water

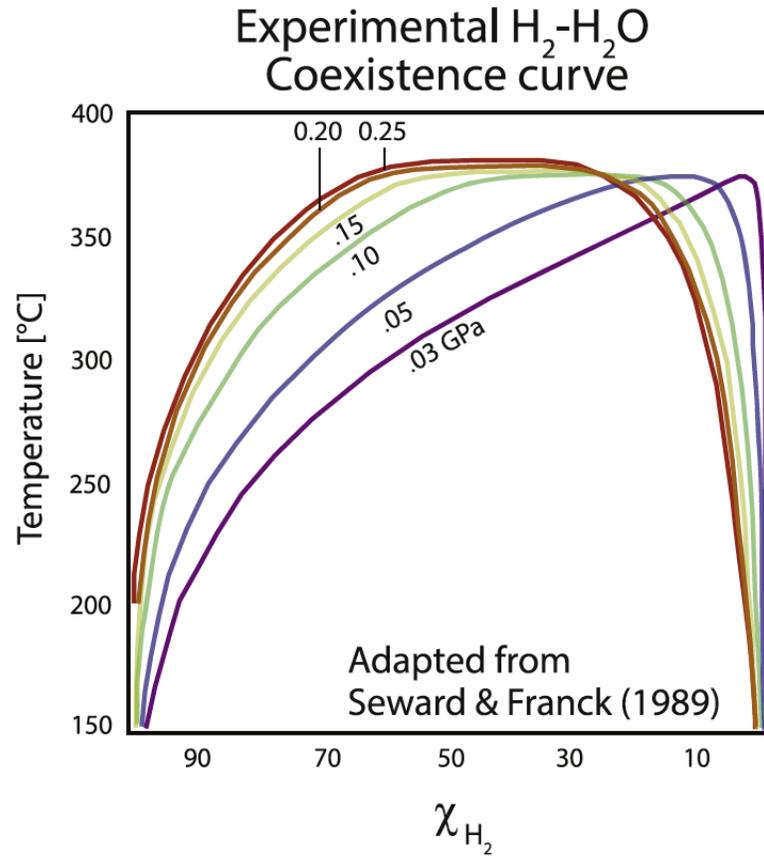
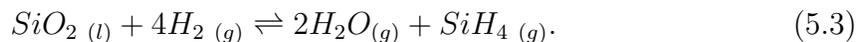


Figure 5.1: Empirical coexistence curves between water and hydrogen. At low pressures, saturation vapor pressure is a reasonably good approximation, but at higher pressures the coexistence is more symmetrical. Data from [Seward and Franck, 1981], Figure from [Bailey and Stevenson, 2021].

and hydrogen. At low pressures (e.g., .03GPa purple curve), the liquid phase remains nearly pure, while more and more water enters the gas phase as temperature increases. Above the critical temperature of water (about 650K, note the y-axis in Figure 5.1 is in Centigrade), the phases no longer separate and both species can coexist in a well-mixed supercritical fluid. We expect the basic thermodynamics to play out in much the same way for a mixture of silicates and hydrogen. Likely the coexistence curves will look qualitatively similar, albeit increasing both pressure and temperatures by about an order-of-magnitude owing to the order-of-magnitude larger bond energies within silicate atoms compared to weaker hydrogen bonding between water molecules. Therefore we expect the mixing ratio beneath the stable layer to depend on the pressure formation conditions. In the high pressure limit $x \rightarrow 1/2$, and in the low pressure limit $x \rightarrow 1$. This complication may be important from the perspective of understanding planetary densities and internal composition, and may impact evolution in the sense that hydrogen may degas from the core as it cools. We will discuss hydrogen pollution of the core in the conceptual sense in more detail in this thesis, but will leave the more involved topic of a numerical analysis as the subject of a future work. From the perspective of understanding the heat content of the core, this detail is of secondary importance. Even in the limit where the mole ratio is 1:1 between silicates and hydrogen, the mass ratio will still be dominated by silicates. Therefore in this work, we operate under the assumption that the core remains pure in order to constrain the circumstances in which this analysis will be relevant. If convection is inhibited at or above the critical pressure of silicates, then regardless of the mixing ratio at depth, the temperature must be the critical temperature at the top of the core. Below the critical temperature, the composition of the vapor phase continues to become increasingly enriched in silicates as temperature increases and the coefficient of thermal expansion remains negative, therefore the critical temperature uniquely determines the temperature beneath the stable layer and a new convective

layer begins. If convective inhibition initiates below the critical pressure, then the temperature beneath the stable layer cannot be uniquely determined, and depends on the formation and evolutionary state. In this case, the core will be a magma ocean as previous studies predicted, although the thermal contact between the core and envelope will be less efficient due to the existence of a superadiabatic stable layer. In this work, we will focus on the cases where the pressure at which convection is shut off is greater than the critical pressure, and the core is a supercritical fluid.

We must add the following caveat particular to silicates that does not apply to ices that condense at temperatures cool enough that individual molecules do not dissociate. For silicates, unlike water, the atomic composition of the condensate in the condensed phase will not in general be identical to the atomic composition of the vapor phase [Xiao and Stixrude, 2018]. Furthermore, unlike a mixture of hydrogen and water, hydrogen can react chemically with silicates, complicating the approximation of pure substances in coexisting phases. The chemistry is varied and complex, but likely to be of greatest importance especially under extreme partial pressures of hydrogen is the equilibrium between silicate vapor and silane gas, e.g.,



Le Chatelier’s principle states that the equilibrium concentration scales with chemical activity raised to the power of its stoichiometric coefficient. In high pressure hydrogen dominated environments, the above chemical equilibrium reaction will be driven strongly to the right, indicating that our approximated treatment of silicate vapor obeying results from ab initio simulations that did not include hydrogen is incomplete. Detailed modeling of the complicated series of hundreds of simultaneous equilibrium chemical reactions is beyond the scope of this work (but has been modeled in prior

works, see e.g., [Schaefer and Fegley, 2009]), though we note its effect with respect to convective inhibition can still be accurately modeled using the generalized approach from the preceding paragraphs. The derivation will be different, because the number of gaseous molecules is actually greater on the left hand side of the equation, such that δN will be negative when δT is positive. Our more generalized framework is more flexible and intuitive than previous formulations, allowing for this general principle to be re-derived straightforwardly in more exotic thermodynamic environments.

In the following sections, we will neglect these caveats, treating the coexistence between liquid and vapor phases as an Arrhenius relationship while the liquid phase remains pure. For the behavior of silicate vapor, we use the results from ab initio quantum mechanical simulations [Xiao and Stixrude, 2018].

5.3 When does convective inhibition matter?

Having established the theoretical foundations for convective inhibition by silicate vapor, we now seek to delineate under what circumstances these considerations are relevant. We will do this under the following approximations: neglecting self-gravity of the atmosphere, an ideal gas equation of state, an adiabatic temperature gradient below the radiative-convective boundary, and a plane-parallel geometry. We further use the approximations for the coexistence of silicate vapor with hydrogen gas outlined in the previous section. Under the cold Murnahan equation of state, the radius of a silicate planet scales as the quarter power of its mass, $R_c = R_{\oplus}(M_c/M_{\oplus})^{1/4}$.

We model the saturation vapor pressure as

$$p_s = \exp(A - B/T), \quad (5.4)$$

where p_s is the saturation vapor pressure in GPa, and the parameters A and B are 11.8 and 45,000 respectively, fitting ab initio simulations from [Xiao and Stixrude, 2018]. Convective inhibition is initiated at pressure p_1 and temperature T_1 satisfying

$$\frac{p_s(T_1)}{p_1} = x_{crit}(T_1). \quad (5.5)$$

Under our approximations, the total pressure at the bottom of the envelope (or equivalently the top of the core) is $p_c = \frac{M_e M_e g}{4\pi R_c^2}$ where M_e is the total mass in the envelope. Conveniently, because $R_c \propto M_c^{1/4}$, this can be rewritten to be independent of core mass and radius:

$$p_c = \frac{GM_e M_\oplus}{4\pi R_\oplus^4}. \quad (5.6)$$

The total pressure at the core includes both gas mass M_{XY} and some mass from silicate vapor, which can be non-negligible. The total mass of the envelope, then, is

$$M_e = M_{XY} + 4\pi R_\oplus^4 \int_{p_{rcb}}^{p_1} \frac{\epsilon p_s(T(p))}{pGM_\oplus} dp, \quad (5.7)$$

where p_{rcb} is the pressure at the radiative-convective boundary. This quantity is likewise independent of core mass. Therefore for a given envelope mass, the mass at the top of the core is uniquely determined.

We now make the assumption that the stable layer is thin compared to the full extent of the envelope, and possesses negligible gas mass. This is equivalent to assuming a high opacity, low thermal conductivity limit. We assess the validity of these assumptions in Section 5.3.1. In this case, we can insist $p_1 = p_c$. Because T_1 is uniquely determined for a given p_1 according to Equation 5.5, this sets the adiabat for the convective part of the envelope. Highly irradiated super-Earths are expected to have effective temperatures very near their equilibrium temperatures (e.g., [Ginzburg

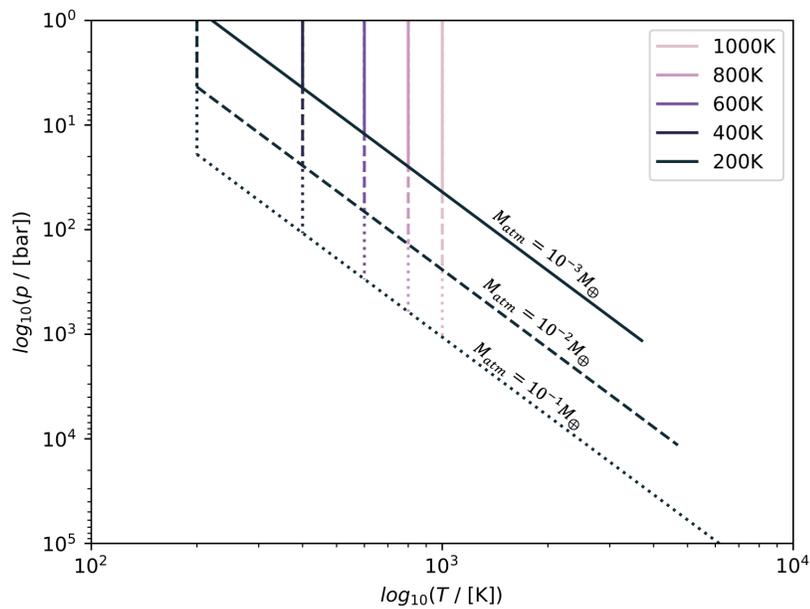


Figure 5.2: Pressure temperature profiles for a sample of envelope properties. Note the thermodynamic conditions at the base of the envelope, (p_1, T_1) , depend on envelope mass only. Higher equilibrium temperature planets have deeper RCBs.

et al., 2016]), and the temperature at the radiative-convective boundary will be near the effective temperature. Using the adiabatic relationship

$$T_1 = T_{eq} \left(\frac{p_1}{p_{rcb}} \right)^{\nabla_{ab}}, \quad (5.8)$$

where $\nabla_{ab} \equiv \frac{d \ln T}{d \ln p}|_{ab}$ is the adiabatic temperature gradient, invariant if the Gruneisen parameter is invariant. By combining these considerations, we can uniquely determine the radiative-convective boundary pressure by specifying the total mass M_{XY} of gas in the envelope. The luminosity scales with the depth of the radiative-convective boundary according to [Ginzburg et al., 2016]

$$L = \frac{64\pi\sigma T_{rcb}^4 R'_B}{3(\rho\kappa)_{rcb}}, \quad (5.9)$$

where R'_B is the so-called “modified Bondi radius,”

$$R'_B \equiv \frac{\gamma - 1}{\gamma} \frac{GM_c \mu d}{k_B T_{rcb}}. \quad (5.10)$$

Equation 5.9 intuitively states that the planetary flux is inversely proportional to the optical depth of the radiative convective boundary, i.e., $F \sim \sigma T_e^4 / \tau$, and is valid for large τ . Following [Freedman et al., 2008], we assume the opacity scales linearly with pressure, so that the planetary flux diminishes rapidly with increasing depth of the radiative-convective boundary. We plot the planetary flux as a function of envelope mass in Figure 5.3.

Figure 5.3 is centered around an envelope mass of $10^{-2}M_\oplus$. We note for clarity that this value is in Earth masses, not as a fraction of the core mass. So, for example, a $5M_\oplus$ sub-Neptune/super-Earth that is 1% gas by mass should be read as 5×10^{-2} on Figures 5.3 and 5.4. We further note that an atmospheric mass of

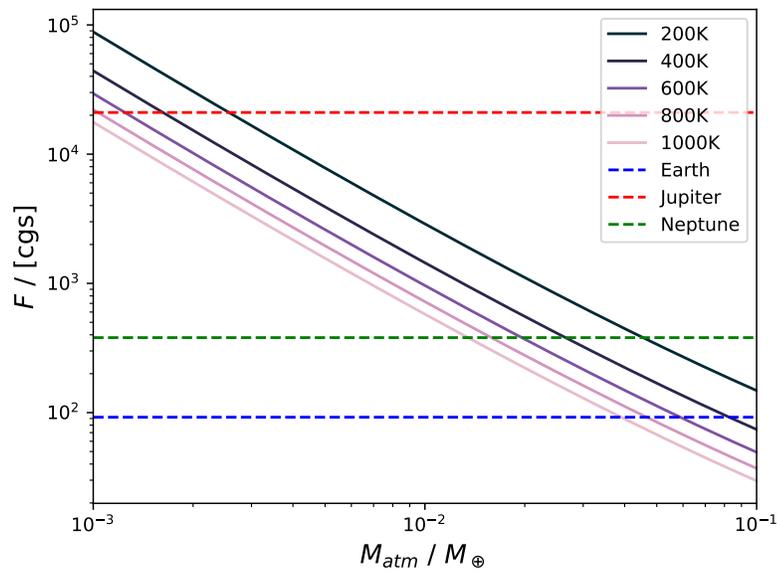


Figure 5.3: Core-envelope heat flux in cgs units as a function of envelope mass. This value should be roughly constant until the core temperature cools to T_1 . Dashed horizontal lines show the contemporary flux for three sample planets in our solar system for comparison.

$\sim 10^{-2}M_c$ is a realistic estimate for super-Earths based on hydrostatic equilibrium between the core and the gas disk during formation (see e.g., [Ginzburg et al., 2016]).

The behavior of Figure 5.3 shows a monotonically decreasing relationship for planetary flux as a function of envelope mass. Intuitively, one can understand the behavior this way: the pressure at the bottom of the envelope scales linearly with envelope mass. Under the thin stable layer approximation, we can then use the relationship from 5.1 to uniquely determine the temperature T_1 at the bottom of envelope such that $x \rightarrow x_{crit}$. Given the effective temperature, we can use the adiabatic relationship to place the radiative-convective boundary. Assuming fixed T_1 , we can then use Equation 5.9 to estimate that the luminosity (or flux, for fixed radius) should scale approximately as the inverse square of the envelope mass. In reality this is an overestimate; accounting for the fact that T_1 will increase as p_1 increases accounts for the somewhat shallower dependence of flux on envelope mass.

Figure 5.3 also shows that the flux of the planet is not too sensitive to the planet's effective temperature, and that the flux decreases with increasing effective temperature. From Equation 5.9, we see that $F \sim T_e^4/p_{rcb}^2$, where the flux is the core-envelope internal heat flux. For fixed $M_{atm} \implies (p_1, T_1)$ will likewise be fixed. Therefore we can infer the expected scaling for p_{rcb} from Equation 5.8. We then find the scaling relationship for flux as a function to be $F \sim T_e^{-1}$, consistent with Figure 5.3. We note that this result is sensitive to the adiabatic lapse rate, i.e., the Grüneisen parameter.

We note that the critical pressure of silicates, about 1.4kbar, corresponds to an envelope mass of about $10^{-3}M_\oplus$. We see from Figure 5.3 that the luminosity of planets near this limit is relatively large, and the core can cool efficiently. Therefore we consider this the lower bound for which the consideration of convective inhibition is

interesting. For much lower mass, the flux is sufficiently high that the core can cool quickly, and prior models subsequently work to a good approximation. Therefore for all cases of interest to us, the core will initially be in a supercritical state, rather than a magma ocean.

5.3.1 Estimating the thickness of the stable layer

We must challenge our assumption that the stable layer can be treated as infinitesimally thin. Furthermore, we will challenge our assumption that the layers of the planet behave distinctly—either fully convective or fully stable—given the intrusion of eddy diffusivity. The thermal transport properties of the materials of interest under the relevant thermodynamic conditions are poorly constrained, but we can set reasonable estimates on their order-of-magnitude. We consider three thermal transport mechanisms of interest: conductive, radiative, and advective.

In the absence of thermal transport data at high temperatures and pressures, we resort to the ideal gas estimate $k \sim \rho \lambda c_v \sqrt{\frac{2k_B T}{\pi m}}$. Thermal conductivity should be roughly independent of density, and the mean free path λ is inversely proportional to density. Using characteristic numbers, and using the critical temperature of silicates as an upper bound, we find the thermal conductivity to be $k \sim 10^5$.

Radiative heat transport in the limit where the mean free path is small compared the length scales of the problem (valid for high pressures), we can estimate the equivalent thermal conductivity due to radiative transfer as $k \sim \frac{4\sigma T^3}{\rho\kappa}$. The relevant parameters are reasonably well-constrained in this context, except the opacity. It is unlikely that the opacity is extremely small. According to hydrogen's absorption spectrum, hydrogen can be nearly transparent for blackbody temperatures of order

1400K. However, in our case the stable layer is a mixture of hydrogen and silicates. In this case, the stable layer should be a cosmic mixture of many elements, not dominated by hydrogen or any other single element. The diversity and abundance of composition will make atmospheric windows unlikely. Following [Freedman et al., 2008], the opacity for kbar to 100s of kbar (the pressure range of interest) should be significant, plausibly of order $1\text{-}100\text{cm}^2\text{g}^{-1}$. To be cautious we assume a low opacity of order $1\text{cm}^2\text{g}^{-1}$. In this case we obtain a characteristic thermal conductivity of order 10^7 . If the material is more opaque, the corresponding equivalent thermal conductivity will be smaller, plausibly comparable to the microscopic thermal conductivity $\sim 10^5$. Our upper estimate is two orders of magnitude larger than the estimate for conduction, and would therefore dominate thermal transport.

Finally we consider advective heat transport. Although the stable layer is called such because it is stable against ordinary large scale overturning convection, we cannot assume this layer is completely stagnant. We know statically stable layers of the Earth's atmosphere can nevertheless involve considerable eddy diffusivity due to, for example, breaking gravity waves [Dornbrack, 1998] [Garcia and Solomon, 1985]. We can attempt to constrain the order-of-magnitude of this process, although it is highly uncertain. In order to move material upward, one must do work against gravity. We can assume as a first approximation that the eddy diffusivity behaves like thermal diffusivity, but acts on deviations away from adiabaticity. For heat flux of order 10^3cgs (see Figure 5.3), this predicts an eddy diffusivity of comparable order-of-magnitude to ordinary thermal diffusivity. Accounting for all these considerations, it appears radiative heat transport is most likely to dominate.

Using our rough upper bound for equivalent thermal conductivity of the system, we can estimate the thickness of the stable layer. From Fourier's Law, $F = -k\nabla T$,

or $L \sim k\Delta T/F$. For a $1\%M_{\oplus}$ envelope, we estimate that F should be of order 10^3 cgs, and in general $\Delta T \sim 10^3$ K. Therefore the thickness of the stable layer should be of order 100km, or about $<1\%$ the expected radius of a super-Earth. The corresponding pressure drop would be of order a kbar, about 10% the pressure overlying the stable layer for $1M_{\oplus}$. For smaller atmospheric masses, the thermal flux will be greater and the density/opacity smaller, while for larger atmospheric masses the thermal flux can be smaller but the density/opacity larger. Therefore for our conservative estimates for relatively efficient radiative heat transport, the stable layer can be non-negligible in thickness and mass content, and this will add a small but non-negligible correction to Figures 5.3 and 5.4.

5.4 Implications for thermal evolution

Previous works (e.g., [Vazan et al., 2018]) have identified that the existence of a gas envelope can considerably slow the thermal evolution process, such that a magma ocean evolutionary phase may persist for billions of years. Our findings here indicate that thermal evolution of the core could take place even more slowly, owing to the inefficient thermal transport through a stable layer. We begin by computing the cooling timescale of a body beginning in the state described in the previous section, with the top of core at the critical temperature and the bottom of the envelope at T_1 such that convection is inhibited. We assume in the radiative-convective quasi-equilibrium state that the envelope temperature profile is frozen, that is does not change in time. This occurs because, if the luminosity of the planet exceeds the heat transport from the core the envelope, that the convective envelope will cool, further thinning the stable layer, until the stable layer is sufficiently thin that the heat flow through the stable layer balances the heat flow out of the planet. As we will argue later in this section, even if the planetary luminosity initially exceeds the heat flow

between the core and envelope, the envelope will quickly relax into the equilibrium state. In radiative-convective quasi-equilibrium,

$$L_{planet} = L_{core} = L_{radiogenic} - M_{core}c_v \frac{dT}{dt}. \quad (5.11)$$

The cooling timescale, then, is

$$t_{cool} = \frac{L_{planet} - L_{radiogenic}}{M_{core}c_v}. \quad (5.12)$$

For simplicity, we assume the radiogenic heat production of silicates matches the radiogenic heat production in the Earth today. Following [Guillot et al., 1995] we use $L_{radiogenic} \sim 2 \times 10^{20} M_c / M_{\oplus} \text{erg s}^{-1}$, although of course this quantity is time-dependent. Furthermore we assume the heat capacity of the core agrees with the ideal gas limit [Bolmatov et al., 2013]. We use the Earth’s contemporary value 5Gyr after formation as a lower bound; radiogenic heating is likely substantially larger at early times, plausibly by more than an order-of-magnitude [Nettelmann et al., 2011]. Assuming the luminosities computed in Figure 5.3, we compute cooling timescales shown in Figure 5.4 assuming a $2M_{\oplus}$ core.

As can be inferred from Figure 5.4, the cooling timescale of the core can exceed the age of the universe for a $2M_{\oplus}$ planet with an envelope 1% of its mass. For smaller envelope masses, its cooling timescale can still be geologically long (billions of years). From these considerations, we infer it is likely that the cores of most observed sub-Neptune/super-Earths may in fact currently be a high-entropy supercritical fluid, rather than a magma ocean or a solid core.

We must now inspect the assumption that we can treat the envelope as initialized in the equilibrium state. An envelope 1% the mass of the core contains about 10% of

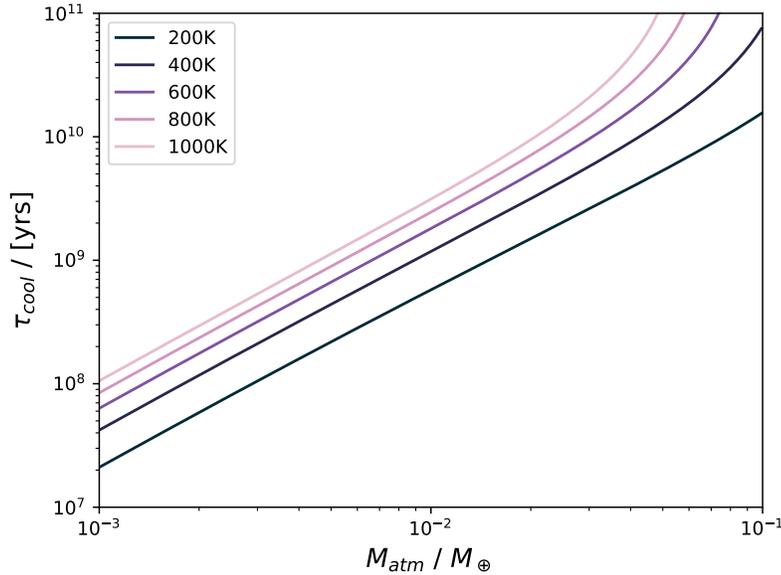


Figure 5.4: Cooling timescales for a $2M_{\oplus}$ super-Earth.

the molecules. Therefore if the envelope were warm at early time, then the envelope could cool to a lower luminosity state on timescales short compared to the time it would take to cool the core and envelope simultaneously. We demonstrate this with a toy model simulation of the early atmospheric evolution, assuming initially the core is in a marginally supercritical state and the envelope is adiabatic, with the radiative-convective boundary at 10 bars. In this sample simulation, the core mass is $2M_{\oplus}$ and the envelope is 0.1% of the core mass. We compare two evolutionary pathways in Figure 5.5.

Over the first million years of evolution, we see that the radiative-convective boundary (RCB) and the luminosity of the inhibited planet (solid) plunges rapidly, while these quantities change much more slowly in the uninhibited (dashed) case. We truncate the simulation after the RCB extends to a depth such that the top of the stable layer approaches the top of the core, at which point the atmosphere will freeze

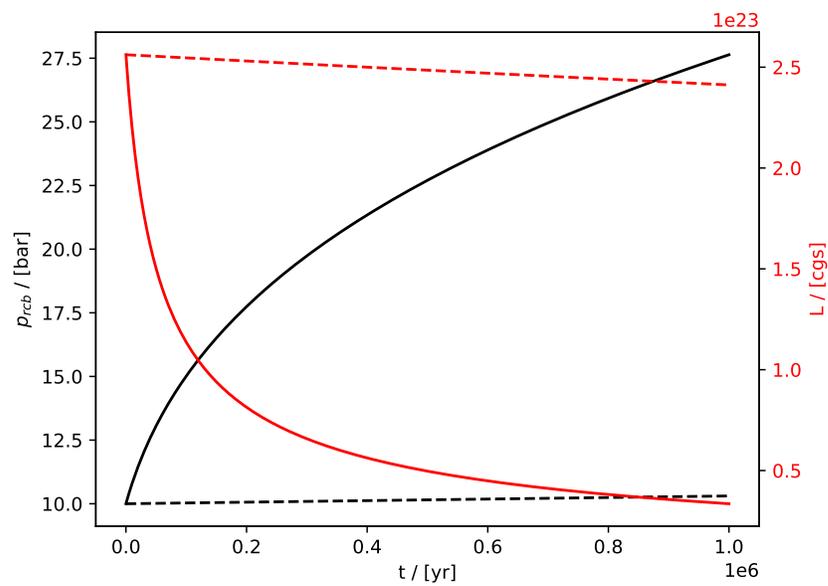


Figure 5.5: Early evolution from an initially adiabatic envelope assuming convective inhibition as described in the preceding section occurs (solid), and no convective inhibition (dashed). The radiative-convective boundary pressure level is shown in black and the left vertical axis, while the luminosity is shown in red and the right vertical axis.

and the core begins to cool. If the envelope were initially more massive, the timescale for it to cool off will be correspondingly larger. This demonstrative simulation is not intended to represent realistic initial conditions (see Section 5.4.1 for further commentary), but merely intended to demonstrate that even an initially hot state can relax onto the equilibrium state discussed in the previous section on a short timescale.

In order to compute the thermal evolution of a planet, we assume radiative-convective quasi-equilibrium at all times. We compute the total energy contained within the planet, both internal heat content and gravitational potential. The rate of change of planetary energy is equal to the magnitude of its luminosity. Assuming the planet's equilibrium temperature is invariant, and the effective temperature is near the equilibrium temperature, then according to Equation 5.9 luminosity is set by the depth of the RCB. Therefore, we are interested in the quantity

$$\frac{dp_{rcb}}{dt} = -\frac{dp_{rcb}}{dE} L(p_{rcb}), \quad (5.13)$$

where L is the luminosity and E is the total energy content of the planet. $E = Q + U$, where Q is the internal heat and U is gravitational energy. Therefore we are interested in the total derivative

$$\frac{dE}{dp_{rcb}} = \frac{dQ}{dp_{rcb}} + \frac{dU}{dp_{rcb}}. \quad (5.14)$$

Focusing first on the internal heat, we begin by splitting up the layers of the planet (stratosphere, troposphere, stable layer, core):

$$Q = Q_s + Q_t + Q_{stab} + Q_c. \quad (5.15)$$

We now compute these quantities analytically, under our assumptions for the envelope

outlined in the previous section.

$$Q_s = 2\pi R_B c_p T_e \rho_{rcb} H^2, \quad (5.16)$$

where R_B is the Bondi radius and H is the scale height.

$$Q_t = \frac{4\pi R_{ic}^2 \gamma k_B}{\mu(\gamma - 1)(2\gamma - 1)g} (p_1 T_1 - p_{rcb} T_e), \quad (5.17)$$

where T_1 is the temperature at the bottom of the troposphere.

$$Q_{stab} = 2\pi R_{ic}^2 c_p / g (T_{oc} - T_1)(p_c - p_1), \quad (5.18)$$

assuming temperature varies linearly with temperature in the stable layer (this is a contestable assumption, but it matters very little because the stable layer is so short lived and contains a very small fraction of the total planetary heat anyway).

Next we compute the gravitational energy contributions.

$$U = U_s + U_t + U_{stab}, \quad (5.19)$$

where we neglect thermal inflation of the core consistent with our assumption of the cold Murnahan equation of state. The stratosphere is straightforward:

$$\frac{dU_s}{dp_{rcb}} = \frac{4\pi R_{ic}^2 k_B T_e}{g\mu_D}. \quad (5.20)$$

The troposphere's gravitational energy is

$$U_t \sim 4\pi R_{ic}^2 g \int_{r_1}^{r_1 + \Delta r} \rho dr + M_{Z,t} g r_1, \quad (5.21)$$

where r_1 is the radius corresponding to pressure level p_1 , Δr is the troposphere's thickness (recall this is under the plane-parallel approximation). $p_{cloud} = \frac{M_{Z,t}g}{4\pi R_{ic}^2}$ is the additional pressure at the bottom of the troposphere caused by the integrated metal mass in the troposphere. Here we model the metal mass of all the vapor in the troposphere as a shell at a fixed radius, and count the additional density of gas above that point. It makes the calculation much easier to assume all this mass is concentrated at the cloud level, rather than allowing the mean molecular weight to vary in the stratosphere. This should be a good approximation, because most of the metal mass is indeed near the cloud level. Using the adiabatic temperature gradient $\frac{dT}{dr} = -g/c_p$ and hydrostatic equilibrium we can derive the density structure

$$\rho(r) = \frac{(p_1 - p_{cloud})\mu_{XY}(c_p T_{oc} + g(r_1 - r))^{\nabla_{ab}}}{k_B(T_{oc} + (r_1 - r)g/c_p)}, \quad (5.22)$$

$$r_1 = \frac{k_B T_{oc}}{\mu_{oc}g(\nabla_{ab} - 1)} \left[\left(\frac{p_{ic}}{p_{oc}} \right)^{\nabla_{ab}} - 1 \right] + R_{ic}, \quad (5.23)$$

$$\Delta r = \frac{k_B T_e}{\mu_{XY}g(\nabla_{ab} - 1)} \left[\left(\frac{p_1 - p_{cloud}}{p_{rcb}} \right)^{\nabla_{ab}} - 1 \right], \quad (5.24)$$

$$\frac{dU_t}{dp_{rcb}} = \frac{\partial U_t}{\partial p_1} \frac{dp_1}{dp_{rcb}} + \frac{\partial U_t}{\partial T_1} \frac{dT_1}{dp_{rcb}} + \frac{\partial U_t}{\partial(\Delta r)} \frac{d(\Delta r)}{dp_{rcb}} + \frac{\partial U_t}{\partial r_1} \frac{dr_1}{dp_{rcb}} + \frac{\partial U_t}{\partial M_{Z,t}} \frac{dM_{Z,t}}{dp_{rcb}}. \quad (5.25)$$

The integral in Equation 5.21 can be solved analytically, and its derivatives with respect to r_1 and Δr are straightforward using the fundamental theorem of calculus.

This is the method used in order to compute the early envelope evolution using Equation 5.5. The difference between the dashed and solid curves are whether we include the contribution of the heat content of the core in the cooling $Q_c = T_c c_v M_c$, or if we assume while the stable layer is highly extended that negligible heat escapes the core. Modeling subsequent evolution is more complicated and will be discussed further in Section 5.4.2, laying the groundwork for future work.

5.4.1 Formation considerations

Our analysis so far has focused exclusively on the planet after formation, and assumes conservation of mass. We must include a brief discussion of the dependence of these results on formation conditions. The consequential elements to consider are the initial temperature of the core, the initial placement of the RCB, atmospheric loss at early times, the distribution of silicate vapor in the envelope, and possible pollution of hydrogen in the core. We will discuss each of these points individually.

The upper bound for the initial temperature of the core is the accretion energy. The gravitational binding energy is $\frac{3GM_c^2}{5R_c^2}$. Under our assumptions for the relationship between mass and radius for silicate cores, this corresponds to a temperature of approximately

$$T_{accretion} \sim 2.9 \times 10^4 \left(\frac{M_c}{M_\oplus} \right)^{7/4} \text{ K}, \quad (5.26)$$

well above the critical temperature for silicate super-Earths. However, such extreme temperatures imply extremely efficient cooling according to the Stefan-Boltzmann Law. Therefore a detailed formation model is needed in order to determine the actual core temperature at early times. Prior modeling of super-Earth planets shows the core temperature can be maintained at high $\sim 10^4\text{K}$ temperatures at early times after the emplacement of the hydrogen envelope [Bodenheimer et al., 2018] [Brouwers and Ormel, 2020]. We point to these results from detailed formation studies to justify our assumption that the core can exist at or above the critical temperature of $\sim 6.6 \times 10^3\text{K}$ [Xiao and Stixrude, 2018] at early times.

The radiative-convective boundary at early time will be determined by the disk conditions during formation. Formation models find this value to initialize near roughly 10 bars [Bodenheimer et al., 2018] [Brouwers and Ormel, 2020], varying

somewhat depending on specific formation conditions. This motivates our choice of an initial RCB at 10 bars in our illustrative simulation from Figure 5.5, that simply demonstrates how rapidly a low mass, convectively decoupled envelope can cool off at early times.

We further did not consider atmospheric loss at early times, a highly complex and poorly understood phenomenon that is likely a diagnostic phase of planetary formation and early evolution. This phenomenon has been investigated in detail in prior works, both atmospheric loss by photoevaporation [Owen and Wu, 2017] or by core luminosity [Gupta and Schlichting, 2019]. With the exception of the toy demonstration model for early evolution and cooling of the envelope, this should not be relevant to our analysis. We are not in this work performing detailed evolution models, but rather considering a planet in equilibrium in a snapshot in time. We do comment, however, that the core powered mass loss mechanism may be interrupted when considering convective inhibition, as it effectively shuts off significant heat flow from the core to the envelope until the luminosity has been substantially reduced.

Our model assumes the envelope is saturated in silicate vapor, and that the core is pure silicates. The assumption of a saturated envelope is supported by formation models that explicitly consider silicate vapor [Brouwers and Ormel, 2020] [Bodenheimer et al., 2018]. The mechanism for this is the simultaneous accretion of hydrogen and silicates, where at early times accretion is dominated by silicates with some gas pollution, and at later times dominated by gas accretion with silicate pollution ranging from pebbles to planetesimals vaporizing in the atmosphere before cooling and raining out, ensuring a saturated envelope. These formation models also find a so called “outer core” polluted with hydrogen as a consequence of the stage of formation when some hydrogen is accreted alongside the continued accretion of silicates. At

later evolutionary phases, this outer core may condense and rain out magma, effectively growing the inner, embryonic core [Brouwers and Ormel, 2020]. According to our model, we expect a polluted outer core to gradually undergo phase separation, as we will describe in more detail in Section 5.4.2, although a full evolutionary model is beyond the scope of this exploratory work. We do however lay the theoretical groundwork to motivate future work on this topic. We further note that according to Figure 5.4, for sufficiently massive envelopes the core can stay hot for longer than the age of the universe, and evolutionary considerations are of secondary importance. However, we consider the topic of interest because many super-Earth planets with less massive envelopes may undergo evolution over geologic time.

5.4.2 Hydrogen pollution in the core and commentary on the evolution of intermediate envelope mass planets

Recent formation models of super-Earths indicate a high metallicity region outside of the embryonic core polluted with hydrogen. Above the critical temperature and pressure, silicates and hydrogen can mix in all proportions. The exact nature of the mixture depends on thermodynamics, and the formation scenario. In the thermodynamic sense, the mixing ratio depends on the ambient pressure (Figure 5.1), ranging from a 1:1 mixture in the high pressure limit, and nearly pure silicate in the lower-pressure limit. As gas accretes on top of the core, the pressure at the top of the core will increase, potentially allowing more hydrogen to dissolve into the core. Here we will outline a general heuristic for the subsequent evolution of a hydrogen polluted outer core underlying an envelope, although a detailed evolutionary model will be reserved as the subject of future work.

We begin in the configuration outlined in the previous sections: the envelope consists of a deep RCB, a convective troposphere, a thin stable layer truncated above

by the critical mixing ratio of a saturated envelope, and below by a supercritical core. The top of the core will be at the critical temperature. The difference for the current consideration involves an “outer core” consisting of a non-negligible mixing ratio of hydrogen. How does the core subsequently cool? Initially, the stable layer is sufficiently thin that the luminosity of the planet is balanced by the luminosity of the core, and the envelope temperature structure remains frozen. Allow for an infinitesimal time step in which the core cools by a small quantity $\delta\theta$. As this happens, a thin shell at the top of the core will cool to a marginally subcritical temperature, and the mixture will phase separate into two coexisting phases of somewhat higher and somewhat lower metallicity. Each phase will have different density depending on their composition. The higher density phase will rain down into the outer core and dissolve in the supercritical mixture, net enriching the outer core in silicates. The temperature gradient within the subcritical top layer of the outer core will then cause the layer to become stable against convection due to the arguments from Section 5.2, joining the overlying thin stable layer. In a timescale short compared to geologic time, then, an equilibrium temperature gradient will emerge, and the stable layer will become thicker than it was before. This thicker stable layer will no longer accommodate the full luminosity of the planet with heat flow between the core and the envelope, meaning the envelope can now continue to cool. The RCB will increment deeper, until the stable layer has thinned sufficiently that the full heat flow of the planet can be accommodated by the core. At this point the configuration has reset, matching the initial conditions but with a somewhat enriched core, a somewhat more massive envelope, and a deeper RCB. By this process, the outer core can gradually degas, separating into an increasingly enriched outer core and an increasingly massive envelope, until the outer core approaches a silicate mixing ratio of unity and the configuration considered in the previous section emerges.

The timescale for this cooling can be estimated to first order from Figure 5.4, although a detailed model is necessary for more precise results. The cooling of a hydrogen-polluted core will proceed slower than a simple cooling of a pure core, because the luminosity of the planet will continue to appreciably decrease as the RCB increases in depth, and the phase separation and settling of higher density of material at depth will provide an additional source of heat through gravitational energy. A detailed model of this generalized evolution will be the subject of a future work, following this heuristic outlined above and the equations from the main Section 5.4. In this more generalized case, the derivatives $\frac{dQ}{dp_{rcb}}$ and $\frac{dU}{dp_{rcb}}$ will be considerably more complicated when the relevant pressure levels and compositions are time variable. Additionally, the ideal gas approximation is unlikely to be appropriate for the extreme pressures present in the outer core, and a more sophisticated equation of state must be adopted.

5.5 Discussion

We have presented a new model for the interior and evolution of sub-Neptune/super-Earth class planets, likely to be the most common planets in the universe [Silburt et al., 2015]. Similar to the concept of convective inhibition that has been in the literature for decades [Guillot, 1995] and has generated renewed interest in recent years [Leconte et al., 2017] [Friedson and Gonzales, 2017] [Markham and Stevenson, 2021], we generalize these arguments to apply in the limit where the condensing species rather than the dry gas dominates in abundance. We find an extreme case of convective inhibition, wherein a hydrogen atmosphere with a layer of static stability can effectively insulate a core at very high temperatures ($10^3 - 10^4\text{K}$) for geologic time, or potentially longer than the age of the universe.

This model should be thought of as exploratory, and further work is necessary. To begin with, more work needs to be done demonstrating the general mechanism of convective inhibition in either a laboratory setting or a high-quality physics-based simulation as an emergent phenomenon. Details about our assumptions, including the assumption of saturation and thermodynamic equilibrium, must be assessed in a dynamical context. Furthermore, this model must be coupled with a realistic formation scenario, as the subsequent evolution may be sensitive to the planet's initial conditions. Additionally our simplified assumptions of a plane-parallel atmosphere and an ideal gas equation of state are not likely to be appropriate, as the radii observed on super-Earths from transit depths indicate nearly half the total radius may be the envelope, implying the necessity of a spherical geometric treatment. The stable layer, which for low heat flows may not be extremely thin, must also be considered explicitly along with relevant thermal transport properties. Finally a serious investigation must explicitly account for the possibility of hydrogen pollution in the core, likely to be substantial based on both formation and thermodynamics arguments.

Our model is of interest for the following reasons. First, we predict the luminosity of these planets will be very low, roughly between the luminosity of Earth and Neptune or less, and will also be rather insensitive to the planetary system's age. This can potentially be tested by direct imaging of further out exoplanets in the mass range of interest. Second, if the luminosity of the planet does not significantly exceed the production of heat by the core, e.g., by radiogenic heating, convection within the core may shut off, thereby shutting down any dynamo producing a magnetic field. Third, the temperatures within the core may be so extreme that the contribution to the planet's density due to heat cannot be neglected, and the required quantity of hydrogen to match observed exoplanet radii may be smaller than usually thought (see e.g., [Bodenheimer et al., 2018]). Finally, this model for super-Earth interiors

further demonstrates the importance of considering exoplanetary systems holistically, thinking about fundamental thermodynamics, and remembering that our own planet may be atypical. Indeed, the term “super-Earth” itself may be misleading, as bodies significantly larger than Earth likely do not resemble our planet at all.

Bibliography

- [Bailey and Stevenson, 2021] Bailey, E. and Stevenson, D. (2021). *PSJ*, 2(2).
- [Baines and Hayden Smith, 1990] Baines, K. and Hayden Smith, W. (1990). *Icarus*, 85:65–108.
- [Baines et al., 1995] Baines, K., Mickelson, M., Larson, L., and Ferguson, D. (1995). *Icarus*, 114:328–340.
- [Bercovici and Schubert, 1987] Bercovici, D. and Schubert, G. (1987). *Icarus*, 69:557–565.
- [Bodenheimer et al., 2018] Bodenheimer, P., Stevenson, D., Lissauer, J., and D’Angelo, G. (2018). *ApJ*, 868(138).
- [Bolmatov et al., 2013] Bolmatov, D., Brazhkin, V., and Trachenko, K. (2013). *Nature Comm.*, 4(2331).
- [Bolton et al., 2017] Bolton, S. et al. (2017). *Science*, 356(821).
- [Borucki et al., 2011] Borucki, W., Koch, D., Basri, G., Batalha, N., Boss, A., et al. (2011). *ApJ*, 728(2).
- [Brouwers and Ormel, 2020] Brouwers, M. and Ormel, C. (2020). *A&A*, 634(A15).
- [Campbell and Anderson, 1989] Campbell, J. and Anderson, J. (1989). *ApJ*, 97(5).

- [Chachan and Stevenson, 2019] Chachan, Y. and Stevenson, D. (2019). *Icarus*, 323:87–98.
- [Christensen-Dalsgaard, 2014] Christensen-Dalsgaard, J. (2014). *Lecture Notes on Stellar Oscillations*. Aarhus University.
- [Dermott, 1979] Dermott, S. (1979). *Icarus*, 37:310–321.
- [Deubner and Gough, 1984] Deubner, F. and Gough, D. (1984). *Ann. Rev. A&S*, 22:593–619.
- [Dombard and Boughn, 1995] Dombard, A. and Boughn, S. (1995). *ApJ*, 443:L89–92.
- [Dornbrack, 1998] Dornbrack, A. (1998). *J. Fluid Mechanics*, 375:113–141.
- [Durante et al., 2021] Durante, D., Bolton, S., Mankovich, C., Markham, S., Stevenson, D., et al. (2021). *In prep*.
- [Durante et al., 2020] Durante, D., Parisi, M., Serra, D., et al. (2020). *GRL*, 47(4).
- [El Moutamid et al., 2017] El Moutamid, M., Hedman, M., and Nicholson, P. (2017). *2017dps*.
- [Elkins-Tanton, 2011] Elkins-Tanton, L. (2011). *ApSS*, 332:359–364.
- [Elliot et al., 2017] Elliot, J. et al. (2017). *NASA JPL*, D-100520.
- [Fazio et al., 1976] Fazio, G., Traub, W., Wright, E., Low, F., and Trafton, L. (1976). *ApJ*, 209:633–657.
- [Fegley and Lodders, 1994] Fegley, B. and Lodders, K. (1994). *Icarus*, 110:117–154.
- [Fischer et al., 2011] Fischer, G. et al. (2011). *Nature Let.*, 475:75–77.
- [Fortney et al., 2011] Fortney, J., Ikoma, M., Nettelmann, N., Guillot, T., and Marley, M. (2011). *ApJ*, 729(32).

- [Freedman et al., 2008] Freedman, R., Marley, M., and Lodders, K. (2008). *ApJS*, 174(2).
- [Friedson and Gonzales, 2017] Friedson, A. and Gonzales, E. (2017). *Icarus*, 297:160–178.
- [Frohlich et al., 1997] Frohlich, C., Andersen, B., Appourchaux, T., et al. (1997). *SoPh*, 170:1–25.
- [Fuller, 2014] Fuller, J. (2014). *Icarus*, 242:283–296.
- [Galanti et al., 2019] Galanti, E., Kaspi, Y., Miguel, Y., et al. (2019). *GeoRL*, 46(2).
- [Garcia and Solomon, 1985] Garcia, R. and Solomon, S. (1985). *JGR Atmospheres*, 90(D2):3850–3868.
- [Gaulme et al., 2008] Gaulme, P. et al. (2008). *A&A*, 490(2):859–871.
- [Gaulme et al., 2014] Gaulme, P. et al. (2014). *ArXiv [astro=ph.EP]*, 1411(1740v2).
- [Gaulme et al., 2015] Gaulme, P. et al. (2015). Cambridge University Press.
- [Gaulme et al., 2016] Gaulme, P., Rowe, J., Bedding, T., Benomar, O., et al. (2016). *ApJ Let.*, 833(1).
- [Gaulme et al., 2011] Gaulme, P., Schmider, F., Gay, J., Guillot, T., and Jacob, C. (2011). *A&A*, 531(A104).
- [Gibbard et al., 2002] Gibbard, S. et al. (2002). *Icarus*, 156:1–15.
- [Gierasch, 2000] Gierasch, P. (2000). *Nature Let.*, 403:628–629.
- [Ginzburg et al., 2016] Ginzburg, S., Schlichting, H., and Sari, R. (2016). *ApJ*, 825(29).
- [Goldreich and Keeley, 1977] Goldreich, P. and Keeley, D. (1977). *ApJ*, 212:243–251.

- [Goldreich and Kumar, 1994] Goldreich, P. and Kumar, P. (1994). *ApJ*, 424:466–479.
- [Goldreich et al., 1994] Goldreich, P., Murray, N., and Kumar, P. (1994). *ApJ*, 424(1):466–479.
- [Goldreich and Nicholson, 1977] Goldreich, P. and Nicholson, P. (1977). *Icarus*, 30:301–304.
- [Gudkova and Zharkov, 2006] Gudkova, T. and Zharkov, V. (2006). *AdSpR*, 38(4):764–769.
- [Guillot, 1995] Guillot, T. (1995). *Science*, 269(5231):1697–1699.
- [Guillot, 2005] Guillot, T. (2005). *AREPS*, 33:493–530.
- [Guillot et al., 1995] Guillot, T., Chabrier, G., Gautier, D., and Morel, P. (1995). *ApJ*, 450:463–472.
- [Guillot et al., 2020a] Guillot, T., Li, C., Bolton, S., Brown, S., Ingersoll, A., et al. (2020a). *JGR Planets*, 125(8).
- [Guillot et al., 2020b] Guillot, T., Stevenson, D., Atreya, S., Bolton, S., and Becker, H. (2020b). *JGR Planets*, 125(8).
- [Gulkis et al., 1983] Gulkis, S., Olsen, E., Klein, M., and Thompson, T. (1983). *Science*, 221(4609):453–455.
- [Gupta and Schlichting, 2019] Gupta, A. and Schlichting, H. (2019). *MNRAS*, 487:24–33.
- [Hedman and Nicholson, 2013] Hedman, M. and Nicholson, P. (2013). *ApJ*, 146(1).
- [Helled et al., 2002] Helled, R. et al. (2002). *arXiv:1010.5546v1*.

- [Helled et al., 2020] Helled, R., Nettelmann, N., and Guillot, T. (2020). *SSRv*, 216(38).
- [Hubbard, 1977] Hubbard, W. (1977). *Icarus*, 30(2):305–310.
- [Hubbard, 1978] Hubbard, W. (1978). *Icarus*, 35:177–181.
- [Hubbard et al., 1999] Hubbard, W., Guillot, T., Marley, M., Burrows, A., Lunine, J., and Saumon, D. (1999). *P&SS*, 47:1175–1182.
- [Hubbard and MacFarlane, 1980] Hubbard, W. and MacFarlane, J. (1980). *JGR*, 85(B1).
- [Hubbard et al., 1980] Hubbard, W., MacFarlane, J., Anderson, J., Null, G., and Biller, E. (1980). *JGR*, 85(A11):5909–5916.
- [Hubbard and Marley, 1989] Hubbard, W. and Marley, M. (1989). *Icarus*, 78:102–118.
- [Iess et al., 2019] Iess, L., Militzer, B., Kaspi, Y., Nicholson, P., Durante, D., et al. (2019). *Science*, 364(6445).
- [Karkoschka and Tomasko, 2011] Karkoschka, E. and Tomasko, M. (2011). *Icarus*, 211:780–797.
- [Kaspi et al., 2020] Kaspi, Y., Galanti, E., Showman, A., Stevenson, D., Guillot, T., et al. (2020). *SSR*, 216(84).
- [Kellermann and Pauliny-Toth, 1966] Kellermann, K. and Pauliny-Toth, I. (1966). *ApJ*, 145(3):954–956.
- [Knapmeyer-Endrun and Kawamura, 2020] Knapmeyer-Endrun, B. and Kawamura, T. (2020). *Nature Comm.*, 11(1451).
- [Kong et al., 2016] Kong, D., K., Z., and Schuber, G. (2016). *NatSR*, 6(23497).

- [Kumar, 1996] Kumar, P. (1996). *arXiv: [astro-ph]*, 961210v1.
- [Kurosaki and Ikoma, 2017] Kurosaki, K. and Ikoma, M. (2017). *AJ*, 153(260).
- [Landau and Lifshitz, 1959] Landau, L. and Lifshitz, E. (1959). *Fluid Mechanics*. Pergamon Press.
- [Leconte et al., 2017] Leconte, J., Selsis, F., Hersant, F., and Guillot, T. (2017). *A&A*, 598(A98).
- [Li and Ingersoll, 2015] Li, C. and Ingersoll, A. (2015). *Nature Geoscience*, 8:398–403.
- [Li et al., 2017] Li, C., Ingersoll, A., Janssen, M., Levin, S., Bolton, S., et al. (2017). *GRL*, 44(11):5317–5325.
- [Lindal, 1992] Lindal, G. (1992). *ApJ*, 103(3).
- [Lindal et al., 1987] Lindal, G., Lyon, J., Sweetnam, D., Eshleman, V., et al. (1987). *JGRA*, 92(A13):14987–15001.
- [Lindal et al., 1985] Lindal, G., Sweetnam, D., and Eshleman, V. (1985). *ApJ*, 90(6).
- [Linder et al., 2019] Linder, E., Mordasimi, C., Molliere, P., Dominik-Marleau, G., Malik, M., et al. (2019). *A&A*, 623(A85).
- [Loewenstein et al., 1977] Loewenstein, R., Harper, D., and Moseley, H. (1977). *ApJ*, 218:L145–L146.
- [Lognonne et al., 1994] Lognonne, P., Mosser, B., and Dahlen, F. (1994). *Icarus*, 110(2):180–195.
- [Lognonne and Johnson, 2015] Lognonne, P. and Johnson, C. (2015). *Treatise on Geophysics*.

- [Luan et al., 2017] Luan, J., Fuller, J., and Quataert, E. (2017). *arXiv:1707.02510v1 [astro-ph.EP]*.
- [Ludlam, 1980] Ludlam, F. (1980). *Clouds and Storms*. Penn State Press, London.
- [Lunine et al., 1989] Lunine, J. et al. (1989). *ApJ*, 338:314–337.
- [Mankovich and Fuller, 2021] Mankovich, C. and Fuller, J. (2021). *arXiv:2104.13385*.
- [Markham et al., 2020] Markham, S., Durante, D., Iess, L., and Stevenson, D. (2020). *PSJ*, 1(2).
- [Markham and Stevenson, 2018] Markham, S. and Stevenson, D. (2018). *Icarus*, 306:200–213.
- [Markham and Stevenson, 2021] Markham, S. and Stevenson, D. (2021). *PSJ*, 2(4).
- [Marley, 1991] Marley, M. (1991). *Icarus*, 94(2):420–435.
- [Marley and Ackerman, 1999] Marley, M. and Ackerman, A. (1999). *ASP Conf. Ser.*, 3e8.
- [Marley and McKay, 1999] Marley, M. and McKay, C. (1999). *Icarus*, 138(2):268–286.
- [Marley and Porco, 1993] Marley, M. and Porco, C. (1993). *Icarus*, 106(2):508–524.
- [Mian and Tozer, 1990] Mian, Z. and Tozer, D. (1990). *Terra Nova*, 2:455–459.
- [Mosser, 1994] Mosser, B. (1994). *IAU Colloq.*, 147:481.
- [Mosser, 1995] Mosser, B. (1995). *A&A*, 293:586–593.
- [Mosser et al., 1996] Mosser, B., Galdemard, P., Lagage, P., Pantin, E., et al. (1996). *Icarus*, 121(2):331–340.

- [Nettelmann et al., 2011] Nettelmann, N., Fortney, J., Kramm, U., and Redmer, R. (2011). *ApJ*, 733(2).
- [Nettelmann et al., 2016] Nettelmann, N., Wang, K., Fortney, J., Hamel, S., Yellamilli, S., et al. (2016). *Icarus*, 275:107–116.
- [Null et al., 1981] Null, G., Lau, E., Biller, E., and Anderson, J. (1981). *ApJ*, 86(3).
- [Orton et al., 2007] Orton, G., Encrenaz, T., Leyrat, C., Puetter, R., and Friedson, A. (2007). *A&A*, 473(L5-L8).
- [Owen and Wu, 2017] Owen, J. and Wu, Y. (2017). *ApJ*, 847(29).
- [Pearl and Conrath, 1991] Pearl, J. and Conrath, B. (1991). *JGR*, 96:18921–18930.
- [Pearl et al., 1990] Pearl, J., Conrath, B., Hanel, R., Pirraglia, J., and Coustenis, A. (1990). *Icarus*, 84:12–28.
- [Podolak et al., 1991] Podolak, M., Hubbard, W., and Stevenson, D. (1991). Models of uranus’ interior and magnetic field. In *Uranus*, pages 29–61. University of Arizona Press, Tucson, AZ.
- [Pollack et al., 1996] Pollack, J. et al. (1996). *Icarus*, 124:62–85.
- [Rages et al., 1991] Rages, K., Pollack, J., Tomasko, M., and Doose, L. (1991). *Icarus*, 89:359–376.
- [Robock and Oppenheimer, 2003] Robock, A. and Oppenheimer, C. (2003). *Geophys. Monogr. Ser.*, 139:360.
- [Schaefer and Fegley, 2009] Schaefer, L. and Fegley, B. (2009). *ApJ*, 703:L113–L117.
- [Scheibe et al., 2019] Scheibe, L., Nettelmann, N., and Redmer, R. (2019). *A&A*, 632(A70).

- [Scheibe et al., 2021] Scheibe, L., Nettelmann, N., and Redmer, R. (2021). *A&A*, in review.
- [Schmider et al., 2007] Schmider, F. et al. (2007). *A&A*, 473(3):1073–1080.
- [Seager et al., 2007] Seager, S., Kuchner, M., Hier-Majumder, C., and Militzer, B. (2007). *ApJ*, 669(2).
- [Seward and Franck, 1981] Seward, T. and Franck, E. (1981). *Berichte der Bunsengesellschaft für physikalische Chemie*, 85(2).
- [Silburt et al., 2015] Silburt, A., Gaidos, E., and Wu, Y. (2015). *ApJ*, 799(180).
- [Stevenson, 1983] Stevenson, D. (1983). *JGR*, 88(B3):2445–2455.
- [Stevenson, 2020] Stevenson, D. (2020). *Ann. Rev. EPS*, 48:465–489.
- [Stevenson, 2021] Stevenson, D. (2021). *Lectures on Planetary Interiors*. California Institute of Technology.
- [Stevenson et al., 2015] Stevenson, J. et al. (2015). *KISS Venus Seismology Study Team*.
- [Stier et al., 1978] Stier, M., Traub, W., Fazio, G., Wright, E., and Low, F. (1978). *ApJ*, 226:347–349.
- [Stoker, 1986] Stoker, C. (1986). *Icarus*, 67:106–125.
- [Storch and Lai, 2015] Storch, N. and Lai, D. (2015). *MNRAS*, 450(4):3952–3957.
- [Tassoul, 1980] Tassoul, M. (1980). *ApJ*, 43:469–490.
- [Townsend et al., 2013] Townsend, R., Teitler, S., and Paxton, B. (2013). *arXiv*, 1309(4455).
- [Turrini et al., 2014] Turrini, D. et al. (2014). *ArXiv*, 1402(2474).

- [Vazan and Helled, 2020] Vazan, A. and Helled, R. (2020). *A&A*, 633(A50).
- [Vazan et al., 2018] Vazan, A., Ormel, C., Noack, L., and Dominik, C. (2018). *ApJ*, 869(163).
- [Wolszczan and Frail, 1992] Wolszczan, A. and Frail, D. (1992). *Nature*, 355(6356).
- [Wu, 2005a] Wu, Y. (2005a). *ApJ*, 635(1).
- [Wu, 2005b] Wu, Y. (2005b). *ApJ*, 635:647–687.
- [Wu and Lithwick, 2019] Wu, Y. and Lithwick, Y. (2019). *ApJ*, 881(2).
- [Xiao and Stixrude, 2018] Xiao, B. and Stixrude, L. (2018). *PNAS*, 115(21):5371–5376.
- [Zahnle et al., 2010] Zahnle, K., Schaefer, L., and Fegley, B. (2010). *Cold Spring Harb. Perspect. Biol.*, 2(10):a004895.