Edited by
Rory Barnes

Formation and Evolution of Exoplanets

Related Titles

R. Dvorak (ed.)

Extrasolar Planets

Formation, Detection and Dynamics

305 pages with 99 figures and 11 tables
2008
Softcover
ISBN: 978-3-527-40671-5

Stahler, S. W., Palla, F.

The Formation of Stars

865 pages with 511 figures and 21 tables
2004
Softcover
ISBN: 978-3-527-40559-6

Shapiro, S. L., Teukolsky, S. A.

Black Holes, White Dwarfs and Neutron Stars

The Physics of Compact Objects

663 pages with 77 figures and 18 tables
1983
Softcover
ISBN: 978-0-471-87316-7

Edited by Rory Barnes

Formation and Evolution of Exoplanets

WILEY-VCH Verlag GmbH & Co. KGaA

The Editor

Dr. Rory Barnes
Lunar and Planetary Lab
University of Arizona
Tucson, USA
rory@astro.washington.edu

Cover Picture
by Lucio Mayer, Zurich

Library of Congress Card No.: applied for

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at http://dnb.d-nb.de.

Cover Adam Design, Weinheim
Typesetting Laserwords Private Limited, Chennai, India
Printing and Binding Bell & Bain Ltd., Glasgow

Printed in Great Britain
Printed on acid-free paper

ISBN: 978-3-527-40896-2

Contents

Formation and Evolution of Exoplanets. Edited by Rory Barnes

ISBN: 978-3-527-40896-2

Preface

Astronomy is a discipline full of breathtaking discoveries that challenge our collective imagination. Yet, in a field of study that literally encompasses the entire universe, few astronomical objects rival the exoplanets' capacity for inspiring awe. These are whole new worlds, which, compared with our familiar solar system, are likely to possess a wide variety of bizarre interiors, fantastic surfaces, and exotic atmospheres. Perhaps the most tantalizing possibility is the remote detection of extraterrestrial life. Such a discovery is years away, but the necessary groundwork is underway toward understanding the formation and evolution of exoplanets.

The first-found exoplanet, 51 Pegasi, was discovered only 15 years ago by Michel Mayor and Didier Queloz. Yet, in the intervening years, over 400 more have been detected by a myriad of techniques. These planets have a breadth of properties that reveals their extreme diversity. Moreover, with sparse data, poor measurement precision, and disparate observational platforms, interpreting data can be daunting to say the least.

On the other hand, the known exoplanets appear so different from the bodies in the Solar System that they pose exciting challenges to long-standing astrophysical concepts. There are giant planets 10 times closer to their star than Mercury is to our Sun. There are exoplanets with elongated orbits, more akin to comets than Solar System planets. There are numerous planets in orbital resonances. There are systems with configurations that seemingly should eject planets within thousands of years. Can theory explain these unexpected discoveries? Perhaps surprisingly, the answer is yes (although often with important caveats). This volume is a guidebook through the twists and turns of exoplanet theory, with frequent confrontations with observations.

In retrospect, the wide diversity should have been expected. The planets (and satellites) in the Solar System alone could never be a representative sample of all the planets in the universe. Nor could the specifics of the formation process of the Solar System be the only route to produce a planetary system. Furthermore, the known exoplanets only represent the tip of the iceberg; new discoveries will undoubtedly overturn emerging paradigms. Nonetheless, this volume provides a critical snapshot of the field of exoplanets as it transitions from case-by-case studies to those which examine and interpret the patterns emerging from the ensemble of discoveries.

Formation and Evolution of Exoplanets. Edited by Rory Barnes

ISBN: 978-3-527-40896-2

Thousands of papers are published annually on exoplanets, an enormous number given the field's relative youth. Not surprisingly, this volume cannot summarize the entire body of work on exoplanets. Nevertheless, the assembled authors have nobly tackled critical aspects of the formation and evolution of planetary systems. From their birth in circumstellar disks, to the evolution of their interiors, to the long-term changes in their orbits, the general trends and remaining puzzles are laid out. This global picture will provide the reader with a broad understanding of both how far the field has come, and how very, very far it has to go.

This book is intended to suit a readership with a wide range of previous knowledge of planetary science, astrophysics, and scientific programming. Expertise in these fields should not be required to grasp the key concepts presented in the forthcoming chapters, although a reasonable grasp of basic physics is probably essential. The topics covered do span the range from gravitational interactions, to gas dynamics, to atmospheric photochemistry, etc. Therefore some chapters may naturally be more accessible to one reader than to others. But taken together, the range of topics provides an introduction to the exciting, interdisciplinary, and fast-moving world of exoplanet science.

Seattle, WA, July 2009 *Rory Barnes*

List of Contributors

Rory Barnes
University of Arizona
Lunar and Planetary Laboratory
Tucson
USA

Jacob L. Bean
Georg-August-Universität
Institut für Astrophysik
Friedrich-Hund-Platz 1
37077 Göttingen
Germany

Richard Greenberg
University of Arizona
Lunar and Planetary Laboratory
1629 E University Blvd
Tucson
AZ 85721-0092
USA

Olenka Hubickyj
NASA
Ames Research Center
Moffett Field
CA 94035
USA

Brian Jackson
University of Arizona
Lunar and Planetary Laboratory
1629 E University Blvd
Tucson
AZ 85721-0092
USA

Wilhelm Kley
University of Tuebingen
Astronomy and Astrophysics
Computational Physics
Auf der Morgenstelle 10
72076 Tuebingen
Germany

Katharina Lodders
Washington University
Department of Earth & Planetary
Sciences and McDonnell Center
for the Space Sciences
Planetary Chemistry Laboratory
One Brookings Drive
Saint Louis
MO 63130
USA

Formation and Evolution of Exoplanets. Edited by Rory Barnes

ISBN: 978-3-527-40896-2

Kevin L. Luhman
Pennsylvania State University
Department of Astronomy and Astrophysics
525 Davey Lab
University Park
PA 16802
USA

F. Marzari
Universita' di Padova
Dipartimento di Fisica
Via Marzolo 8 35131 Padova
Italy

Lucio Mayer
University of Zurich
Institute for Theoretical Physics
CH-8057 Zurich
Switzerland

and

ETH Hoenggerberg Campus
Physics Department
HPT D11
CH-8093 Zurich
Switzerland

Alice C. Quillen
University of Rochester
Department of Physics and Astronomy
500 Wilson Boulevard
Rochester
NY 14627
USA

Sean N. Raymond
Observatoire de Bordeaux and CASA
University of Colorado;
raymond@lasp.colorado.edu
University of Colorado
Center for Astrophysics and Space Astronomy and Center for Astrobiology
Boulder
CO 80309-0389
USA

and

CNRS; Université Bordeaux 1
Laboratoire d'Astrophysique de Bordeaux
33270 Floirac
France

Edward W. Thommes
University of Guelph
Department of Physics
50 Stone Road East
Guelph
Ontario N1G 2W1
Canada

1
Exoplanet Observations

Jacob L. Bean

1.1
Introduction

Observations of exoplanets have revolutionized our view of planetary system formation and evolution. Since the first unambiguous detection of an exoplanet orbiting a normal star in 1995 [1], more than 250 exoplanetary systems have been discovered. The discovery and characterization of these systems revealed that planets have a wider variety of physical and orbital properties than was anticipated. This has stimulated a critical review of every aspect of existing planet formation and evolution paradigms, and necessitated the consideration of new, more general theories.

Despite the diversity of detected exoplanets, some clear trends have emerged. In this chapter, we summarize the main results from the observational study of exoplanetary systems in the context of how they can constrain theories of planet formation and evolution. We focus on the robust findings that have emerged from examining the properties of large samples of detected exoplanets. Despite their success, current observational techniques are still quite limited, relative to what planets might be expected, and it is important to keep this in mind when interpreting results. Therefore, we also describe the limitations and biases of current observational methods and potential overinterpretations of the data that should be avoided. Compelling individual results are mentioned when they offer views beyond the well-explored parameter spaces. The rest of the chapter is separated into sections about observed exoplanet orbital, physical, and host star properties.

1.2
Orbital Properties

The orbital parameters of known exoplanets are perhaps the most well-determined of all their properties. The majority of the data comes from high-precision time-series radial velocity observations, which have also yielded most of the original

Formation and Evolution of Exoplanets. Edited by Rory Barnes

ISBN: 978-3-527-40896-2

discoveries. Radial velocities carry information on a planet's orbital period, eccentricity, and location and time of periastron. Information on a planet's mass is also available in radial velocity data, but this parameter is normally degenerate with the value of the planet's orbital inclination (for a notable exception, see [2]). Barring additional data from transits (see Section 1.2) or astrometry [3], what is then typically determinable is the product of the planet's mass and the sine of the orbital inclination angle ($M \sin i$). This corresponds to the minimum possible value for the planet's mass. When analyzing large samples of radial-velocity-detected exoplanets, the significance of the inclination angle ambiguity is reduced because of the statistical distribution of $\sin i$ [4, 5]. However, when considering a single planet, the uncertainty in the mass should not be ignored because there is the potential that the true mass is significantly different from the minimum or expectation values [6]. With that caveat, we drop the $\sin i$ and use M when speaking about the properties of large samples of planets.

Despite being the best-known properties of exoplanets, the orbital parameters determined from radial velocity data can still potentially suffer from both large statistical and systematic errors. The published statistical uncertainties for orbital parameters are generally underestimated and rarely account for the strongly non-Gaussian nature of the constraints provided by radial velocity data [7]. Systematic errors can arise when there are significant levels of stellar-activity-induced noise in the data (referred to as *jitter*) [8–10]. In some cases, activity signals can be periodic and coherent over typical observational timescales. Care has to be taken in the case of active stars so that not only activity-induced signals are not misinterpreted to be orbiting planets [11] but also real orbital motion signals are not incorrectly attributed to activity [12]. Systematics can also arise when all the planets in a system are not accounted for. In these cases, the determined orbital parameters for the known planets can absorb the signals from the unrecognized planets. The orbital parameters for planets later shown to be in systems harboring additional planets often undergo significant revision when the additional planets are accounted for.

Although high-precision radial velocity observations are possible around stars with spectral types from late F down to the bottom of the main sequence, radial velocity surveys mainly target FGK-type main sequence stars as they are the easiest to observe. High rotational speeds and activity hinder searches around stars with ages below $\sim$1 gigayear [13, 14] and only a few planets are known around such stars. Therefore, the statistics of known exoplanets are dominated by those orbiting roughly solar-type main sequence stars.

1.2.1
Orbital Periods and Eccentricities

For results shown in this section, data on the orbital parameters of the currently known exoplanets (as of September 20, 2008) were taken from the Extrasolar Planets Encyclopaedia.[1)] Such a composite data set is useful because it includes

1) http://exoplanet.eu/

data from recent discoveries and its statistics are consistent with the results from volume-limited [15] and magnitude-limited [16] surveys. We include only objects with $0.3M_{\rm Jup} < M < 20M_{\rm Jup}$ and which were discovered from radial velocity surveys. The considered sample size is 204 exoplanets.

We use an upper mass cutoff for our sample above the accepted boundary between "planets" and "brown dwarfs" (the boundary for which lower mass values imply an object that cannot initiate deuterium fusion, $\sim 13M_{\rm Jup}$) because there is no clear demarcation in the mass distribution of low-mass companions to stars corresponding to this limit. What we are really interested in is in understanding the formation and evolution of objects that ostensibly formed in a circumstellar disk. As planet formation theories suggest that objects with masses greater than $13M_{\rm Jup}$ can form in a disk, it makes sense to consider objects out to the tail of the stellar companion mass distribution rather than an arbitrary cutoff corresponding to a theoretical internal property transition (see Chapter 7).

The lower mass cutoff of our chosen sample corresponds to the mass of Saturn. It was used to select likely gas giants and exclude low-mass planets that are part of the emerging classes of lower mass planets (i.e., Neptune-like ice giants and potentially terrestrial "Super-Earths"). Current surveys are only weakly sensitive to these types of objects compared to their higher mass counterparts, so their statistics are much more incomplete. There is some evidence that these planets' orbital properties might be different from those of gas giants, which would imply they have different formation and evolution histories as well [17–19]. The future results of ongoing planet searches will likely shed more light on this issue.

The motivation for ignoring planets discovered by transit surveys when examining the distribution of planet orbital parameters is that such searches are strongly biased toward short-period planets (i.e., $P < 10$ days) because these planets are the most likely to transit in the first place, and also have multiple such events in the time spans covered by survey campaigns. Radial velocity planet searches are more sensitive to short-period planets as well, but they are still very sensitive to gas giant planets with orbital periods up to ~ 4 years. Therefore, radial-velocity-detected planets are the most uniform sample of planets available, although there are still biases present that should be considered (discussed further below). Disregarding the very low mass planets detected with the radial velocity method also reduces the bias toward shorter periods in the sample because their detectability is strongly dependent on their orbital period. Transiting planets do offer important information on the physical properties of planets, and these planets are discussed in Section 1.2.

The distribution of the orbital periods is shown in Figure 1.1. Inspection of this distribution reveals three potential families. One family, and the most surprising when first discovered, is planets with orbital periods of only a few days. Giant planets with such short orbital periods have come to be known as *Hot Jupiters* because the intensity of the stellar radiation implies that they have high equilibrium temperatures (>1000 K).

The orbital parameters for short-period planets are usually very well determined due to their relatively large induced radial velocity signals and the ease of observing multiple orbits. Current thinking is that these planets did not form

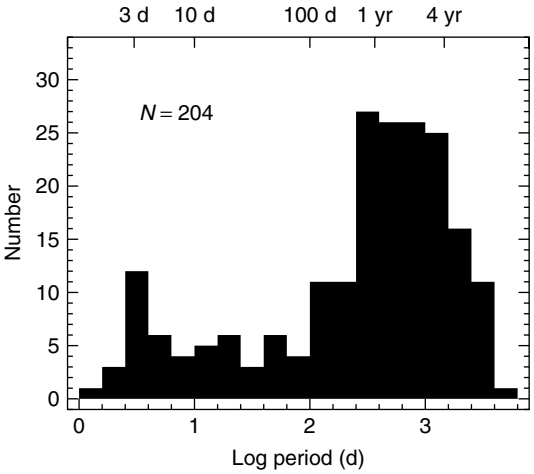


Figure 1.1 Orbital period distribution of gas giant planets identified from radial velocity surveys.

in situ, but rather they formed beyond the snow line as normal gas giants and that some subsequent evolutionary mechanism(s) brought them quite close to their host stars (see Chapter 9).

There is a well-noted "pile-up" of planets near an orbital period of 3 days. Some shorter period planets have been found, mostly from transit surveys, with the shortest known being Wasp-12 b [20] and having $P = 1.09$ days. Interestingly, there appears to be a correlation between high mass planets transiting high metallicity stars with short orbital periods ($P = 1–5$ days) [21, 22]. So it would seem that the mechanisms that gave rise to the short-period planets have a dependence on these properties, or properties that they reflect. The origin of short-period planets is one of the major issues confronting planetary evolution theory, and is discussed at length in Chapters 4–6 and 9–12.

The second observed family in the exoplanet orbital period distribution is the planets with intermediate periods (10 days $< P <$ 100 days). These planets do not display any particularly notable features, but they serve to connect the short-period planets to the longer period planets. The small, flat distribution in this range suggests that, though the orbital evolution mechanisms lead to relatively fewer planets, a continuum of results is still possible.

The third observed family is the planets with longer periods ($P >$ 100 days). There is a steep increase in the number of planets toward longer periods despite the corresponding decrease in sensitivity of the radial velocity method. The large, long-term surveys are only reasonably complete for gas giants with $P < 4$ years, and currently have a time baseline of up to ~10 years for some objects [23]. Many systems exhibit radial velocity trends potentially due to orbiting planets with much longer periods, but the need to observe at least one full orbital cycle precludes robust characterization of most at this time [24].

The data on longer period planets from radial velocity surveys could be interpreted to mean that undetected gas giants with orbits similar to those of Jupiter and Saturn exist. However, precise quantification of the planet formation efficiency and probabilities of evolutionary outcomes should wait until the orbital separation parameter space is well explored. Direct imaging studies have the potential to provide complementary information to radial velocity surveys in the future, and might mitigate the problem of having to obtain decades of observations to probe for Saturn analogs.

Three planetary systems have been identified with direct imaging so far, with planets detected at separations ranging from 8 to 119 AU from their host stars [25–27]. However, these systems all orbit very young A-type stars. Currently, direct imaging results only provide constraints on massive planets having extraordinarily long periods (i.e., $P > 100$ years) around young solar analogs, and even those limits depend on theoretical predictions about the spectral energy distributions of young planets that are of uncertain quality [28, 29]. Improved technology will be needed to close the gap between radial velocity and direct imaging investigations.

In Figure 1.2, the orbital eccentricities of giant planets are plotted as a function of their orbital period, with the objects' minimum masses indicated. Regarding the eccentricity distribution, a segregation into three groups is possible, and in this case, it is guided by theory. The first group is the short-period planets. The orbits of the planets in this region all have measured $e < 0.1$ (see below for some notable exceptions not included in the figure). Some of these planets even have measured $e < 0.01$, which indicates that their orbits are very nearly perfectly circular. However, precise measurements (and estimations of the uncertainties

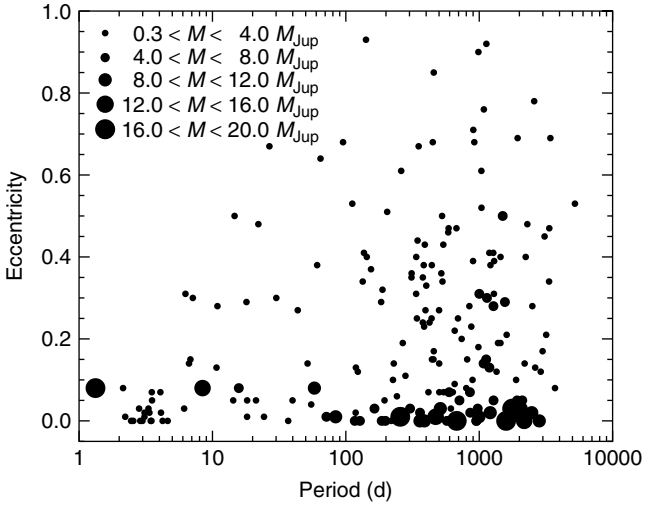


Figure 1.2 Orbital eccentricities versus period for the giant exoplanets. The symbol size indicates the minimum mass of the object.

in these measurements) of low eccentricities are particularly difficult and require a much larger number of radial velocity observations than are usually obtained during initial discovery and characterization of planets [30]. Therefore, while the short-period planets certainly have low eccentricities, it is not known precisely to what level their orbits might be noncircular.

The second obvious eccentricity group is the long-period planets, which have eccentricities over the full range of possible values for bound orbits. The contrast between the eccentricities of the short-period planets and those of the longer period planets suggests that a mechanism operates to affect one distribution and not the other. This mechanism is widely believed to be forces arising from tidal interactions between the planets and their host stars. These forces act on the short-period planets to circularize their orbits on timescales shorter than the age of the system. On the basis of this theory, there is likely a third eccentricity group composed of planets in the transition region between the fully circularized planets and those not noticeably affected by tidal forces. The orbital eccentricities of these planets are damped somewhat, and therefore they have a smaller range of eccentricities than the longer period planets, but not enough time has elapsed for their orbits to become completely circularized. The results of a thorough study of the period–eccentricity relationship for exoplanets using data of fairly uniform quality indicate that the circularization cutoff occurs around $P = 5$ days, the transition region spans 5 days $< P <$ 100 days, and nontidally affected planets have $P >$ 100 days [31].

There are some notable outliers from the nearly circularized orbits of short-period planets trend, all of which were excluded from the data set used for the plots in this section. A number of hypothesis have been proposed to explain the origins of these planets' eccentricities including eccentricity pumping from regular dynamical interactions with another planet or a stellar companion, and longer tidal circularization timescales than the age of the systems. More details on the theory of tides and the interpretations of the observed eccentricities of short planets can be found in Chapter 12.

Another general property of giant planets discernible from Figure 1.2 is the trend for the most massive planets to rarely be in short-period orbits. In the data considered here, only two objects have $M > 5M_{\rm Jup}$ and $P < 50$, and both orbit one component of a multistar system (see Section 1.2.3). A few other similar planets have been found from transiting planet searches, but the trend seen here is still statistically significant when those objects are included. The rarity of massive, short-period planets could be a clue about the orbital evolution mechanism that moves gas giants inward from their birth environment.

1.2.2 Multiplanet Systems

The orbital parameters of planets in multiplanet systems are especially important constraints for theories of planet formation and evolution (see Chapters 4–6). However, detection and characterization of such systems are more challenging. As

mentioned above, signals from additional planets in a system can often be hidden because they are absorbed by the orbital parameters of the known planets when modeling radial velocity data. Generally, for each additional planet in a system, roughly twice as much radial velocity data are needed to constrain the planets' orbits. The amount of needed data increases as the signals from the additional planets decrease, and detection of additional low-mass planets in known planetary systems is particularly challenging.

In some cases, two or more planets in a multiplanet system can be gravitationally interacting. This leads to non-Keplerian motion like periastra precession and eccentricity variations on a wide range of timescales. Most such systems only exhibit interactions on long timescales (i.e., much longer than an observational timescale, see Chapter 3). For these systems, an observationally determined model provides a snapshot of the system and direct integration of the equations of motion is needed to study secular behavior [32]. The implied stability of the determined orbital parameters can also be an issue, with the ''best-fit'' values sometimes describing an unstable system [33–35]. As nearly all exoplanet systems have existed for at least 1 gigayear, such orbital parameters are assumed to be spurious. Stability considerations are sometimes used in these cases to help identify a more likely model. However, such a process involves weighing stability against observational data and there is no straightforward method to incorporate stability in a robust statistical manner (for one such method see [36]). Therefore, modifications to orbital parameters based on stability criteria are sometimes based on judgment and their uncertainties are difficult to quantify.

Multiplanet systems that exhibit planet–planet interactions on short timescales (i.e., similar to or shorter than observational timescales) are rare, but they do exist. For these systems, dynamical models based on the direct integration of the equations of motion must replace the normal Keplerian models when fitting observational data [37, 38]. The proposed existence of an additional very low mass planet ($\sim 5 M_{\oplus}$) in a known planetary system that was later refuted is an extreme example of how erroneous results can be obtained when neglecting the signature of planet–planet interactions [39–41]. A more subtle pitfall is that the orbital parameters determined for these multiplanet systems using a Keplerian model can be inappropriate for studies assessing their long-term dynamics and stability zones.

Despite the challenges of using dynamical models to fit observational data, the systems requiring such analyses present an opportunity to obtain more information from typical observations. If the dynamical perturbations are large enough, and occurring on a short enough timescale, then study of radial velocity data can yield constraints on the interacting planets' true masses and degree of coplanarity [37, 38, 42]. The key to the additional insight is that the planet–planet perturbations discernible in radial velocity data are dependent on the true masses and three-dimensional orbital orientations of the interacting bodies. Therefore, if the perturbations can be characterized well enough, then the planet masses and orbits can be constrained. To date, this kind of analysis has only been applied to one

system (GJ 876, [2, 42]), but this should also be possible for some other known systems when enough data have been collected.

In all the cases of multiplanet systems, the unknown orbital orientations are a problem. The true masses of planets are needed to accurately calculate the dynamical characteristics of multiplanet systems. However, these are rarely known due to the orbital inclination angle degeneracy with planet mass in radial velocity data. Another issue is that the degree of coplanarity of the orbits of planets in all multiplanet systems (except the one mentioned above) is not known. The unknown orbital plane alignments affect the dynamical characteristics of a system. As they are without observational constraint, studies of the dynamics of multiplanet systems have to assume planet masses (usually the minimum or expected values) and the orbital orientations (usually perfectly coplanar).

Currently, there are 33 well-characterized multiplanet systems and ∼30% of systems with only a single identified planet show significant evidence for multiplicity [43]. This percentage must be considered a lower limit to the frequency of multiplanet systems because of the incompleteness of surveys for lower mass and longer period planets. One system is known with five planets (55 Cnc [44]), one is known to harbor four (μ Arae [35]), and the others harbor two or three. The orbital eccentricities in systems known to have three or more planets (median $e = 0.13$) are typically lower than those of planets in other systems (median $e = 0.25$) [43, 45]. This suggests that dynamical instabilities (see Chapter 11) could play an important role in the evolution of planetary systems because the higher order multiples that survive are intrinsically more stable.

A significant number (∼25%) of the multiplanet systems have two planets in or near mean motion resonances [46]. The most common is 2 : 1, with 3 : 1, 4 : 1, 5 : 1, and 3 : 2 also represented. The current best explanation for the observed resonances is differential migration in a disk leading to mutual capture ([47], and see Chapter 9). The prevalence of resonant configurations suggests that this is a common outcome of planetary formation and evolution, and that the responsible mechanism must arise naturally from normal conditions.

1.2.3 Planets in Multiple Star Systems

A significant fraction of stars in the Galaxy are part of a multiple star system [48, 49], and a close stellar companion could disrupt planet formation or influence planets' orbital parameters. Therefore, a complete picture of planet formation and evolution must include consideration of planets around stars that have additional stellar companions, and the details of the effects of stellar multiplicity should be constrained by observations.

Currently, more than 30 exoplanet host stars are known to have stellar companions [50, 51]. Most of the known stellar companions are either widely separated stars of various types, or close-in stars with low masses [52]. The predominance of low-mass stars among close-in companions is likely at least partly an observational bias because radial velocity planet searches mostly exclude near-equal mass binary

stars with separations <5″. The main reason for this is that for all binaries but those with a high mass ratio (large brightness contrast), the obtained spectra contain significant contributions from both objects. These spectra are confusing for the normal algorithms used for high-precision radial velocity measurements. Therefore, the close-in companions to exoplanet host stars have to be sufficiently dim so that they do not affect normal high-precision radial velocity measurements. Many of these close-in stellar companions were not even known a priori, but were discovered with a targeted search after an exoplanet had been found [53].

Some techniques have been developed to overcome the problem of blended spectra from close multiples where the components have similar masses [54, 55], but these have seen limited application. On the basis of these methods [56], a planet was claimed to have been detected around the primary of a close triple star system (HD 188753A), but this is controversial because with the other [57] method the planet was not detected. This further illustrates the difficulty of reliably detecting planets around close multiples.

Among the noncontroversial exoplanet detections, three orbit host stars with a stellar companion separated by ∼20 AU (GJ 86 b [58], γ Cep b [59], and HD 41004 b [60]). These systems pose challenges for planet formation theories because such a close companion could truncate the star's circumstellar disk such that gas giant formation via core accretion becomes unlikely [61, 62].

The orbital properties of exoplanets in multistar systems do exhibit some differences from planets orbiting ostensibly single stars. The main difference is that most of the massive ($M > 2M_{\rm Jup}$) short-period (P < 40 days) planets orbit a star in a multistar system (see Figure 1.2) [52, 63, 64]. Additionally, all short-period planets in multistar systems are in nearly circular orbits ($e < 0.05$) [52, 64]. There is also the suggestion that intermediate-mass planets ($2 < M < 6\,M_{\rm Jup}$) are more common in multistar systems [52]. No correlation has been noticed for planet and companion star orbital properties.

The statistics of planets in multistar systems are still limited owing mainly to the difficulty of detecting planets in close binary systems. Therefore, the observed trends and lack of trends must be treated as preliminary. Continued searches for planets in multiple systems, as well as searches for stellar companions to exoplanet host stars, are needed. In particular, the question of low-mass planets in multistar systems is still very open.

1.2.4
Spin–Orbit Alignment

Radial velocity data for transiting exoplanets offer an additional constraint on their orbital configuration beyond the usual parameters discussed above. The additional information is the angle (λ) between the planets' orbital axes and the rotational axes of their host stars, or their ''spin–orbit'' alignment. This angle is measurable due to the subtle affect of an eclipse of a small portion of a star's surface on its spectrum.

As a planet transits a rotating star, parts of the star with different line-of-sight velocities are obscured. These obscurations lead to changes in spectral line profiles because some component of the Doppler broadened line is missing. Radial velocities measured during a transit then display spurious shifts due to the line profile distortions. If a planet's orbit is prograde, then as it transits its host star, the stellar radial velocities will show a red shift during the first half as some of the blue-shifted light is missing, and then during the second half of the transit the stellar radial velocities will show a blue shift as some of the red-shifted light is missing. The signature of this phenomenon in radial velocities is known as the Rossiter–McLaughlin (RM) effect [65, 66]. The morphology of the RM effect depends on the geometry of the orbit crossing and the rotation of the star [67]. The size of the effect is of the order of tens of meters per second for a Hot Jupiter transiting a main sequence G-type star. Therefore, the effect can be characterized for many transiting planets with current technology, and this characterization yields the spin–orbit angle.

Table 1.1 summarizes the measurements of the RM effect for transiting exoplanets. The main result is that all but one of the exoplanets orbit in the prograde direction and have a spin–orbit angle that is close to or consistent with 0°. These findings are an important constraint for theories of planet formation and evolution because they suggest that stars and their planets originate and inherit their angular momentum from the same collapsing molecular cloud, and that subsequent evolution does not significantly disturb the planets' orbits from their original plane. Of the planets orbiting in the prograde direction with small spin–orbit angles, only HD 209458 b has a more than 3σ detection of a nonzero angle ($\lambda = -4.4^\circ \pm 1.4^\circ$). This value is of the same scale as the spin–orbit angles for the planets of the solar system, which implies some commonalities despite other very different characteristics.

The one possible exception to the spin–orbit alignment rule is XO-3 b ($\lambda = 37.3^\circ \pm 3.7^\circ$). This planet is also unusual for its definite nonzero eccentricity despite its short-period orbit (see Section 1.2.1), and it is also one of the few massive close-in planets ($M = 12.5 \pm 1.9 M_{\rm Jup}$, $P = 3.2$ days [77]). It could be that this object is part

Table 1.1 Measurements of the spin–orbit angle for transiting exoplanets.

Planet	λ (°)	Source
HD 209458 b	-4.4 ± 1.4	[68]
HD 189733 b	-1.4 ± 1.1	[69]
HD 147506 b	0 ± 12	[70, 71]
TrES-1 b	30 ± 21	[72]
HD 149026 b	-12 ± 15	[73]
CoRoT-2 b	7.2 ± 4.5	[74]
HD 17156 b	9.4 ± 9.3	[75, 76]
XO-3 b	37.3 ± 3.7	[77, 78]
TrES-2 b	-9 ± 12	[79]
HAT-P-1 b	3.7 ± 2.1	[80]

of a rare class of planets that had different formation or evolutionary experiences than most of the other detected exoplanets, but this is only speculation. Such counterexamples to otherwise robust statistical trends illustrate the challenges of improving our understanding of the complex interplay of different physical effects leading to planet formation and evolution by studying the inhomogeneous and incomplete sample of known exoplanets.

1.3 Physical Properties

The main observable physical properties of exoplanets are their masses, and, in some cases, their radii. These data constrain planet formation mechanisms, which are generally expected to govern the resulting masses and structures of planets. However, in some individual cases, evolutionary processes could have also played a critical role and must be kept in mind when interpreting results.

1.3.1 Mass Distribution

As described above, the minimum masses of exoplanets can be determined from radial velocity data, and this information is sufficient for statistical studies of large samples of exoplanets. We have already discussed some aspects of how the orbital properties of gas giants depend on planet mass, and now we turn to examining the overall mass distribution.

Figure 1.3 is a plot of the mass function for all known radial-velocity-detected exoplanets, with the data again taken from the the Extrasolar Planets Encyclopaedia. The dominant population is planets with masses on the order of Jupiter's mass. These planets are expected to be composed predominantly of gas, which is confirmed by mass and radius measurements for the subsample of transiting planets in this mass range (Section 1.2.2). The median of the sample is $1.7 M_{\rm Jup}$, and planets with masses up to $10 M_{\rm Jup}$ are not uncommon.

A simple interpretation of the gas giant mass distribution suggests that massive planet formation is generally more efficient than it was in the solar system. However, it is important to consider the biases of radial velocity planet searches to higher mass and shorter period planets when interpreting the statistics. With current radial velocity precisions, the large surveys cannot easily detect Saturn-mass planets ($\approx 0.3 M_{\rm Jup}$) with orbital periods longer than about 300 days [81]. So some fraction of lower mass gas giants might be missing even in the orbital parameter space nominally covered by radial velocity surveys.

Regarding the issue of the short-period bias, nearly all the gas giant planets known likely moved significant distances from their original formation location. If this disturbance was due to migration arising from friction with a disk (see Chapter 9), then the massive gas giant planets might have experienced gas accretion much longer than planets that did not migrate large distances. So the mass distribution

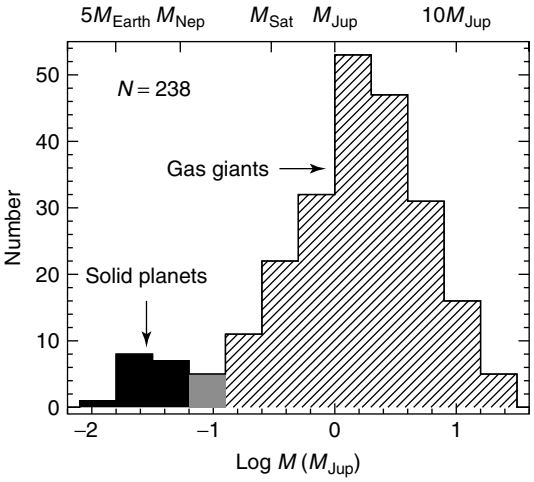


Figure 1.3 Mass distribution of all known exoplanets. There is likely a boundary separating mostly gaseous planets from mostly solid planets around $30M_{\oplus}$ ($\approx 0.1M_{Jup}$), although its exact location is uncertain.

of the known gas giant planets might be different from the mass distribution of the longer period ones yet to be discovered. Despite the biases, the large numbers of detected gas giants imply that their formation is not unusual and must arise from fairly high likelihood events.

Recently, radial velocity surveys have begun to detect lower mass planets that are likely a separate population from the gas giant planets. These planets are expected to be composed mostly of heavy elements (ices and rock). Therefore, they might be referred to as *solid* planets. The likely solid compositions for objects in this mass range are reinforced by the ~90% heavy element abundance inferred for the one known transiting planet in the sample, the $23M_{\oplus}$ GJ 436 b [82].

The exact boundary between the gaseous and solid planets is uncertain, but inspection of the exoplanet mass distribution reveals a potentially significant gap around $30M_{\oplus}$ ($\approx 0.1M_{Jup}$). As these two groups of planets are thought to have very different formation scenarios, it is possible that they differ in other observable properties. Although the sample of solid planets is severely incomplete, even at the shortest periods, there is a hint that the metallicities of the stars hosting these planets are different from those hosting gas giants [83] (see Section 1.3.1 for more discussion). It will be interesting to compare the orbital parameters of the two samples as more solid planets are detected and characterized. In the future, it might become possible to distinguish the mass distributions of rocky "Super-Earth" planets from icy Neptune-like planets.

Interestingly, two out of the seven planet detections from microlensing observations have been of very low mass planets (3 and $5M_{\oplus}$) [84, 85]. These findings suggest that such planets might be quite common because microlensing events are so rare that a single detection points to a large underlying sample. Therefore,

the low-mass planets are similar to the long-period planets in that they represent a potentially large sample of still-to-be-discovered planets. The possible extensive incompleteness of the known planet population should be kept in mind when constraining theory. Simulations of planet system formation and evolution tuned to reproduce the orbital and physical properties of currently known exoplanets might be missing the bigger picture by focusing on what turns out to be the minority of planet populations.

1.3.2
Masses and Radii of Transiting Planets

If a planet is observed to transit its host star, then photometric time-series observations of an event (i.e., a transit light curve) can be used to determine the planet's radius and orbital inclination. This knowledge yields the true mass and density of the planet when combined with the minimum mass obtained from radial velocity data. The mass and radius of a planet is a very interesting data set because it allows inference of the planet's internal structure and composition. Transiting planets also afford a view into their atmospheric properties via time-resolved photometry and spectroscopy as they pass in front of and behind their host stars. These data mainly concern the planets' atmospheric chemistry and are discussed in Chapter 8.

Figure 1.4 is a plot of the masses and radii of all the known transiting objects with $M < 20M_{\rm Jup}$. The data were taken from a compilation [22] and recent individual results [20, 86–103]. The most striking thing about the mass–radii data is that they reveal that Jupiter-mass planets can have a wide range of densities – up to an order of magnitude in dispersion. All of these planets orbit very close to their host stars, so the effect of significant stellar insolation must be taken into account when modeling them. Nevertheless, current models suggest that the range of densities observed is more due to natural variations arising from the planets' different heavy element abundances rather than the different levels of insolation they receive [104]. The heavy elements are thought to be primarily concentrated in a central core, and so planets of the same mass and different densities are expected to have different-sized solid cores (see Chapter 8).

Figure 1.5 shows the estimated core masses for the transiting Jupiter-mass planets from a homogeneous study of transiting planet properties [22]. It should be noted that these values are heavily dependent on theoretical models of uncertain accuracy, and that when different sets of models are used different results may be obtained. However, the diversity of radii seen for these planets definitely implies core masses over a large range. In the case shown, the estimated values range from 0 to $70 + M_{\oplus}$.

The inference of nonzero core masses and heavy element enhancements of many planets has mixed interpretation. Significant cores are a prerequisite of gas giant formation through the core accretion mechanism (see Chapter 5), so the data seem to favor that model. However, simulations of gas giant formation via disk

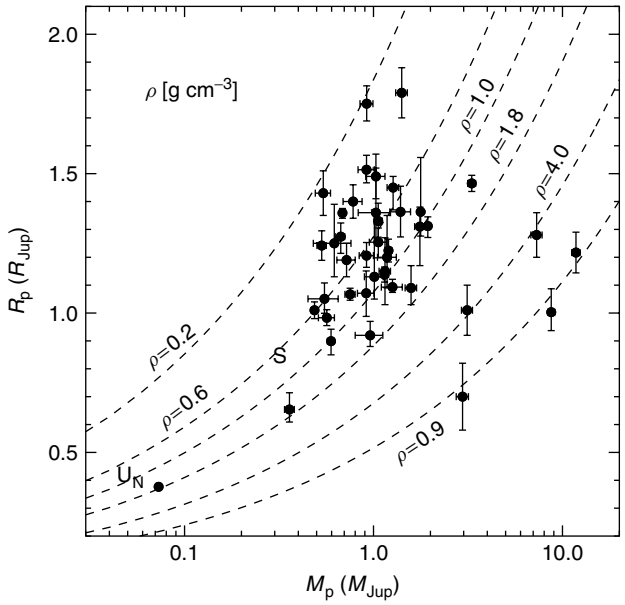


Figure 1.4 Mass–radius diagram for all known transiting exoplanets. The dotted lines are contours of constant density. The locations of Jupiter, Saturn, Neptune, and Uranus are indicated by the labels "J", "S", "N", and "U", respectively.

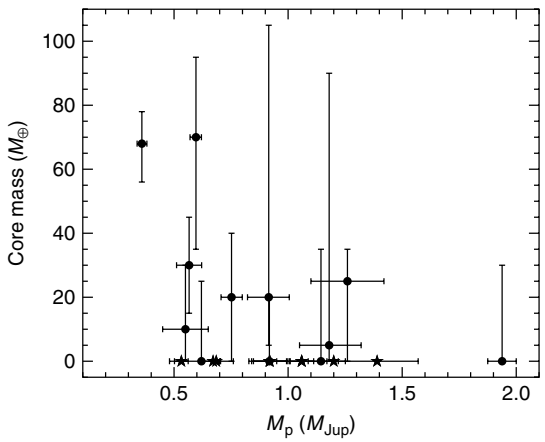


Figure 1.5 Estimated core masses for some transiting exoplanets (data from [22]). The circles represent planets with nonzero or potentially nonzero core masses. The stars represent planets with radii larger than those suggested by coreless models.

instability (see Chapter 4) suggest that sedimentation of solids and planetesimal accretion could potentially lead to the formation of cores in these planets.

A particularly important issue in this area is that the radii of some of the lowest density planets cannot be reproduced even by coreless models – these planets are inexplicably large. The first discovered transiting planet, HD 209458 b [105], has become the prototype of these so-called inflated Hot Jupiters. Its radius is $0.2 R_{\rm Jup}$ larger than nominal coreless models that include insolation [104]. A number of other transiting planets that follow the same trend are now known. So the effect likely arises from a general property of these planets, rather than each of them being unique exceptions. A leading explanation for the larger than expected planets is that they have an unaccounted for source of internal heat [106]. One source could be energy dissipated from ongoing or previous orbital circularization by tidal interactions [107, 108] (see Chapter 12 for more discussion). However, the question of why some Hot Jupiters are inflated is still a major outstanding issue as no consensus on a single mechanism has been reached. A related point is that if coreless models could reproduce some of the planets' sizes, then that would be evidence that they formed via disk instability.

The high end of the distribution of transiting planet densities also offers surprising information. An interesting example is the planet HD 149026 b [109] because heavy elements probably constitute roughly half of its entire mass [110]. Such a high abundance of heavy elements in HD 149026 b and other planets is potentially problematic even for the core accretion mechanism [111, 112], although not all authors agree with this assessment [113]. It has been suggested that the planets with enormous cores could be the products of collisions between planets. This evolutionary scenario suggests the existence of outer planets in the systems. Future observations will be able to test the collision origin theory by looking for these purported companions.

The search for correlations among the properties of transiting planets is still in its preliminary stages. All the investigations to date have utilized small samples of planets, many of which have nonnegligible uncertainties in their measured physical properties, and also theoretical models that are of uncertain quality. As more transiting planets are found and models are improved it will be important to reevaluate the results in this area. The question of the structure of transiting planets is taken up again in Chapters 8 and 12. See also Section 1.3.1 for a discussion of exoplanet host star properties in relation to planet statistics.

1.4 Host Star Properties

Stars and their planets are thought to arise from the same fragment of a collapsed molecular cloud. Consequently, a star's properties likely reflect the physical conditions of its circumstellar disk. The study of correlations between planets' and their host stars' characteristics can therefore give an insight into the

planet formation and evolutionary processes. In this section, we consider results on the correlations between planets and their host stars' compositions and masses.

1.4.1 Planet–Stellar Composition Correlation

The investigation of planet host stars' atmospheric chemical compositions has yielded important constraints on planet formation theories. The most significant finding in this area has been the observed strong correlation between stars' iron abundances[2)] and the frequency of giant planets detected with the radial velocity method [114–116]. Iron is an important element because it traces a star's overall metal content (or metallicity) reasonably well and exhibits numerous lines in the optical spectra of solar-type stars that can be analyzed for determination of abundance. Most studies in this area use a star's iron abundance as a proxy for its metallicity and suggest that the giant planet frequency is correlated with overall metal content rather than iron specifically. This is not strictly correct as the determination of true metallicity requires at least an assessment of the abundances of all the elements that are typically more abundant than iron in stellar photospheres (C, N, O, Ne, Mg, and Si). Such detailed analyses have been carried out for some samples of planet-hosting stars and a few small compositional anomalies have been found (one of which is described below), but mainly the results have confirmed the overall metallicity correlation with iron abundance typically found in field stars [117, 118]. Therefore, the iron abundance trend seen in planet-hosting stars is appropriately interpreted to be broadly a trend with overall metal content.

Figure 1.6 shows data on the giant planet detection frequency as a function of iron abundance for FGK-type dwarf stars from one of the large planet search programs [116]. The sample is limited to FGK-type dwarf stars because these are the stars that make up the majority of planet search targets. Therefore, these are the stars for which the best statistics can be obtained. These stars also have similar masses, so the sought-after correlation between composition and planet frequency is not polluted by another possible correlation with that parameter (see Section 1.3.2). Furthermore, the spectra of FGK-type dwarfs are similar enough so that the same type of abundance analysis method can be used on them. The focus is on giant planets in this correlation study because these are the planets for which robust detection/nondetection statistics are available, and also because they should have a similar formation and evolutionary history.

Numerous studies have reached the conclusion that giant planets are detected more often around metal-rich stars [115, 119, 120]. The origin of the trend has

2) Stellar photospheric chemical abundances are typically given as the logarithm of the fractional abundance relative to the hydrogen abundance and on a scale relative to the solar value. The notation is $[X/H] \equiv \log_{10}(N_X/N_H)_* - \log_{10}(N_X/N_H)_\odot$, where N_X is the number density of element X and N_H is the number density of hydrogen.

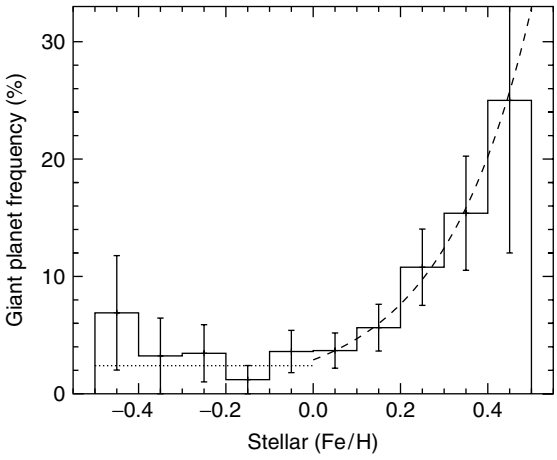


Figure 1.6 The frequency of giant planet detection as a function of stellar iron abundance (data taken from [116]) with errors based on Poisson statistics. The dotted line is a constant fit (2.4) to the data for [Fe/H] < 0.0, and the dashed line is a power law fit ($2.9 \times 10^{2.1 \times [Fe/H]}$) for [Fe/H] > 0.0.

been the subject of much debate, and there have been three main competing explanations. The first idea, which is known as the *primordial* hypothesis, is that the metallicities of the stars reflect the original composition of the protostellar cloud. In this case, giant planets would seem to form preferentially out of circumstellar disks with higher heavy element content. The second idea, which is known as the *pollution* hypothesis, suggests that the formation and/or migration of giant planets leads to postformation accretion of gas-depleted material onto the host star. The third idea suggests that the observed trend is a selection effect arising from planets being easier to detect around metal-rich stars because obtainable radial velocity precisions are better for stars with more and stronger lines in their spectra.

Most of the current data firmly show that the *primordial* hypothesis is the correct one. The supporting results include the lack of a metallicity trend with outer convective zone size among main sequence stars or along the subgiant branch [116, 121]; no observed trend of individual elemental abundances with different condensation temperatures [116, 122]; no degradation of radial velocity precision over the metallicity range covered by radial velocity planet searches [116, 121]; no metallicity trend with planet orbital period [116]; no differences in the abundances between binary stars or stars in a cluster [123–125]; no evidence of ^{6}Li in the photospheres of planet-hosting stars [126]; and asteroseismologic inference that a planet-hosting star's high metallicity is not just concentrated in its atmosphere [127].

It has been pointed out that a single analytic function cannot provide a good fit to the giant planet frequency as a function of host star iron abundance [128]. Instead, a two-piece function with a constant frequency in the subsolar abundance region and a power law for the supersolar region is indicated (see Figure 1.6). This

suggests a dichotomy in giant planet formation. For stars with metal abundances equal to or less than the solar value, the frequency of giant planet formation is ∼2% and is independent of metal abundance. For stars with more metals than the Sun, the giant planet formation rate has a steep correlation with metal content and is ∼25% for stars with three times the solar metal abundance.

The discovered giant planet–metallicity correlation and the evidence for the *primordial* explanation is an important constraint for planet formation theories. However, like the possible presence of heavy element cores in transiting planets, the appropriate interpretation of the data is still uncertain. The data could be interpreted as evidence that core accretion rather than disk instability is the dominant mechanism for the formation of the known gas giants because of the assumption that higher stellar metallicity implies circumstellar disks with a higher surface density of solids. The rocky cores that are the seeds of giant planets in the core accretion model grow faster in such environments. This means the cores can reach the critical size necessary to initiate runaway gas accretion faster. As giant planet formation is a race against the disk gas dissipation, the faster the cores can grow the more giant planets will form [129]. On the other hand, the expected correlation of giant planets with stellar metallicity in the disk instability mechanism is more uncertain, with different authors suggesting that there should [130] or should not [131] be a correlation. See Chapters 4 and 5 for more detailed discussions of the two different gas giant formation mechanisms, and the possible interpretation of the metallicity distribution of exoplanet host stars.

Under the core accretion paradigm, there is most likely a cutoff for the surface density of solids below which giant planets will not form around a solar-type star. This implies the existence of a lower limit for the metallicities of giant planet hosts [132]. Currently, there are three giant planet hosts with $[Fe/H] \sim -0.7$ (HD 114762 [133], HD 47536 [134], and HD 155358 [135]). However, the low-metallicity regime of planet hosting stars has not been well studied, mainly because of selection effects [133, 136]. Therefore, the three lowest metallicity stars might not correspond to the lower limit for giant planet formation.

Another interesting result in this area is that the heavy element abundances of transiting planets and their host stars seem to be correlated [22, 137, 138]. However, what is seen is not a simple linear style relationship, but an increase in the upper limit of the core mass distribution toward higher host star metallicities. This observed trend also suffers from many of the same problems as the other properties of transiting planets mentioned above, and so will need to be reevaluated as more and better data become available.

With regard to individual element enhancement for planet-hosting stars, there has been a finding that giant planet host stars have higher silicon and nickel abundances for a given iron abundance than field stars [139]. Following the *primordial* argument, it would seem that these enhancements further increase the giant planet formation efficiency. However, some caution is warranted in relying too heavily on this result because it is from a single study and its significance, while high, is not overwhelming.

The recent emergence of a sample of potentially solid exoplanets has also motivated study of the compositions of their host stars. There is a preliminary result that such planets are not found preferentially around metal-rich stars, which is in contrast to the trend for giant planet host stars [83]. The current result relies on a small sample of detected planets and therefore suffers from small number statistics. There is also the problem that half of all detected solid planets orbit M dwarfs and this brings up the question of the planet formation correlation with host star mass (see Section 1.3.2). Nevertheless, the composition of solid planet hosts is likely to be a very important area of research as more of these planets are discovered. See Chapter 6 for a detailed discussion of terrestrial planet formation theories.

1.4.2
Planet–Stellar Mass Correlation

It is possible that circumstellar disk mass is proportional to stellar mass. If so, then host star masses are another property that might be correlated with planet properties because total disk mass is likely an important parameter governing planet formation. However, searching for a trend in this area is not straightforward due to the challenges of detecting planets around stars with a wide range of masses. As mentioned above, radial velocity surveys target mostly late F- to mid K-type stars, but these stars only cover a relatively small range of masses. Radial velocity surveys aimed at M dwarfs [140–142], and subgiants and giants [17, 143–145] probe a wider range of masses, but include fewer stars and have different sensitivities and selection biases than the FGK-type star surveys. Nevertheless, there are preliminary data that suggest that the giant planet detection frequency is correlated with stellar mass.

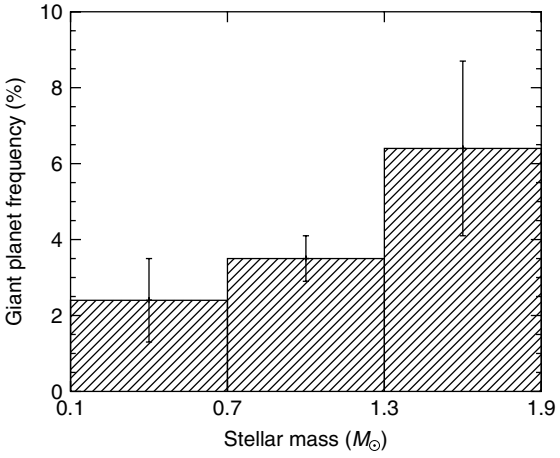


Figure 1.7 The metallicity-corrected frequency of giant planet detection as a function of stellar mass with errors based on Poisson statistics (data taken from [17]).

Figure 1.7 shows the frequency of giant planet detection as a function of stellar mass in one of the large planet search programs (data taken from [17]). There are two caveats to these statistics. First, the data are corrected for stellar metallicity based on the observed trend for solar-type stars. This seems appropriate and necessary, but there is the problem that stellar abundance analyses potentially suffer from systematic errors when applied to different areas of stellar property parameter space [146]. Photospheric abundances for low-mass stars are particularly troublesome to obtain on a scale consistent with solar-type stars because their optical spectra are dominated by molecular lines [147]. Given the steep dependence of planet detection frequency on metallicity in the supersolar regime (see Section 1.3.1), small systematic errors in the metallicities for the mass-correlation sample could be magnified in the applied corrections for these data.

Another caveat is that the detection limits for planets orbiting the high-mass stars are not as stringent as for planets around the solar-type and lower mass stars because the radial velocity signature for a planet of a given mass is inversely proportional to the host star mass. Also, surveys targeting high-mass stars only have about half the time baseline as M dwarf surveys, and about a third of the time baseline as FGK-type star surveys. Therefore, it has been recommended to treat the lower error range in the high-mass bin for the data in Figure 1.7 as a lower limit to the planet detection frequency in this range [17].

From the data in Figure 1.7 it appears that giant planets are more often found around massive stars. Taking the data at face value, giant planets are nearly three times more frequent around stars twice as massive as the Sun than stars with half the mass of the Sun. This is consistent with a number of other findings [140–142, 145], and no evidence to the contrary exists.

Perhaps even more telling than the above data is the fact that there have been no Hot Jupiter detections for M dwarfs despite the ease with which these planets could be detected, given the large radial velocity signal they would have. Assuming 300 M dwarfs have been surveyed for Hot Jupiters [17], no detection implies a 1σ upper limit to their frequency of 0.4%. In contrast, the frequency of Hot Jupiters around FGK-type stars is $1.2\% \pm 0.2\%$ [16].

As it stands, the observed giant planet frequency correlation with stellar mass is potentially in agreement with the core accretion model and inconsistent with the disk instability model. The surface density of solids in disks should scale with the total disk mass [148]. Therefore, if disk mass is correlated with stellar mass, and giant planets form via core accretion, then there is an expected trend for more giant planets around higher mass stars [149, 150]. The frequency of giant planets arising from disk instability is thought to have no relationship with stellar mass [151].

It should again be emphasized that the giant planet–stellar mass correlation is still in its preliminary form. Surveys at both stellar mass ends are continuing and their results will help bring the true underlying relationship into focus. In the future, it will be interesting to take the next step in this area by investigating how the frequency of solid planets correlates with stellar mass.

1.5 Conclusion

We have condensed the main results derived from the study of more than 250 exoplanetary systems over 15 years into a coherent picture that can be used to constrain theories of planet formation and evolution. This amalgamation presumes that a coherent picture exists in the first place, and by necessity washes out the individual characteristics of the systems considered. As discussed extensively above, current observational capabilities are still limited, so the field of observational studies of exoplanets is still progressing and often at a rapid pace. Looking toward the future, the two big issues to tackle observationally as technology improves and observational time baselines are extended are determining the orbital, physical, and host star properties of very low mass exoplanets; and the nature of the exoplanets orbiting beyond 3 AU. However, despite current limitations, the existing data have already proved enlightening. A more in-depth interpretation of the results as they pertain to the various mechanisms and processes of planet formation and evolution are discussed in the following chapters.

References

1 Mayor, M. and Queloz, D. (**1995**) *Nature*, **378**, 355.
2 Rivera, E.J., Lissauer, J.J., Butler, R.P. *et al.* (**2005**) *The Astrophysical Journal*, **634**, 625.
3 Benedict, G.F., McArthur, B.E., Forveille, T. *et al.* (**2002**) *The Astrophysical Journal*, **581**, 115.
4 Zucker, S. and Mazeh, T. (**2001**) *The Astrophysical Journal*, **562**, 1038.
5 Jorissen, A., Mayor, M., and Udry, S. (**2001**) *Astronomy and Astrophysics*, **379**, 992.
6 Bean, J.L., McArthur, B.E., Benedict, G.F. *et al.* (**2007**) *The Astronomical Journal*, **134**, 749.
7 Ford, E.B. (**2005**) *The Astronomical Journal*, **129**, 1706.
8 Saar, S.H., Butler, R.P., and Marcy, G.W. (**1998**) *The Astrophysical Journal*, **498**, 153.
9 Wright, J.T. (**2005**) *The Publications of the Astronomical Society of the Pacific*, **117**, 657.
10 Desort, M., Lagrange, A.-M., Galland, F. *et al.* (**2007**) *Astronomy and Astrophysics*, **473**, 983.
11 Queloz, D., Henry, G.W., Sivan, J.P. *et al.* (**2001**) *Astronomy and Astrophysics*, **379**, 279.
12 Hatzes, A.P., Cochran, W.D., McArthur, B. *et al.* (**2000**) *The Astrophysical Journal*, **544**, 145.
13 Paulson, D.B., Cochran, W.D., and Hatzes, A.P. (**2004**) *The Astronomical Journal*, **127**, 3579.
14 Paulson, D.B. and Yelda, S. (**2006**) *The Publications of the Astronomical Society of the Pacific*, **118**, 706.
15 Udry, S., Mayor, M., Naef, D. *et al.* (**2000**) *Astronomy and Astrophysics*, **356**, 590.
16 Marcy, G., Butler, R.P., Fischer, D. *et al.* (**2005**) *Progress of Theoretical Physics Supplement*, **158**, 24.
17 Johnson, J.A., Butler, R.P., Marcy, G.W. *et al.* (**2007**) *The Astrophysical Journal*, **670**, 833.
18 Mayor, M., Udry, S., Lovis, C. *et al.* (**2009**) *Astronomy and Astrophysics*, **493**, 639.
19 Forveille, T., Bonfils, X., Delfosse, X. *et al.* (**2009**) *Astronomy and Astrophysics*, **493**, 649.

20 Hebb, L., Collier-Cameron, A., Loeillet, B. *et al.* (**2009**) *The Astrophysical Journal*, **693**, 1920.

21 Mazeh, T., Zucker, S., and Pont, F. (**2005**) *Monthly Notices of the Royal Astronomical Society*, **356**, 955.

22 Torres, G., Winn, J.N., and Holman, M.J. (**2008**) *The Astrophysical Journal*, **677**, 1324.

23 Wright, J.T., Marcy, G.W., Butler, R.P. *et al.* (**2008**) *The Astrophysical Journal*, **683**, 63.

24 Wright, J.T., Marcy, G.W., Fischer, D.A. *et al.* (**2007**) *The Astrophysical Journal*, **657**, 533.

25 Kalas, P., Graham, J.R., Chiang, E. *et al.* (**2008**) *Science*, **322**, 1345.

26 Marois, C., Macintosh, B., Barman, T. *et al.* (**2008**) *Science*, **322**, 1348.

27 Lagrange, A.-M., Gratadour, D., Chauvin, G. *et al.* (**2009**) *Astronomy and Astrophysics*, **493**, 21.

28 Lafreniére, D., Doyon, R., Marois, C. *et al.* (**2007**) *The Astrophysical Journal*, **670**, 1367.

29 Nielsen, E.L., Close, L.M., Biller, B.A. *et al.* (**2008**) *The Astrophysical Journal*, **674**, 466.

30 Laughlin, G., Marcy, G.W., Vogt, S.S. *et al.* (**2005**) *The Astrophysical Journal*, **629**, 121.

31 Halbwachs, J.L., Mayor, M., and Udry, S. (**2005**) *Astronomy and Astrophysics*, **431**, 1129.

32 Laughlin, G. and Adams, F.C. (**1999**) *The Astrophysical Journal*, **526**, 88.

33 Barnes, R. and Quinn, T. (**2004**) *The Astrophysical Journal*, **611**, 494.

34 Correia, A.C.M., Udry, S., Mayor, M. *et al.* (**2005**) *Astronomy and Astrophysics*, **440**, 751.

35 Pepe, F., Correia, A.C.M., Mayor, M. *et al.* (**2007**) *Astronomy and Astrophysics*, **462**, 769.

36 Goździewski, K. and Maciejewski, A.J. (**2001**) *The Astrophysical Journal*, **563**, 81.

37 Laughlin, G. and Chambers, J.E. (**2001**) *The Astrophysical Journal*, **551**, 109.

38 Rivera, E.J. and Lissauer, J.J. (**2001**) *The Astrophysical Journal*, **558**, 392.

39 Ribas, I., Font-Ribera, A., and Beaulieu, J.-P. (**2008**) *Astronomy and Astrophysics*, **677**, 59.

40 Alonso, R., Barbieri, M., Rabus, M., Deeg, H.J. *et al.* (**2008**) *Astronomy and Astrophysics*, **487**, 5.

41 Bean, J.L. and Seifahrt, A. (**2008**) *Astronomy and Astrophysics*, **487**, 25.

42 Bean, J.L. and Seifahrt, A. (**2009**) *Astronomy and Astrophysics*, **496**, 249.

43 Wright, J.T., Upadhyay, S., Marcy, G.W. *et al.* (**2009**) *The Astrophysical Journal*, **693**, 1084.

44 Fischer, D.A., Marcy, G.W., Butler, R.P. *et al.* (**2008**) *The Astrophysical Journal*, **675**, 790.

45 Bean, J.L., McArthur, B.E., Benedict, G.F., and Armstrong, A. (**2008**) *The Astrophysical Journal*, **672**, 1202.

46 Butler, R.P., Wright, J.T., Marcy, G.W. *et al.* (**2006**) *The Astrophysical Journal*, **646**, 505.

47 Lee, M.H. (**2004**) *The Astrophysical Journal*, **611**, 517.

48 Duquennoy, A. and Mayor, M. (**1991**) *Astronomy and Astrophysics*, **248**, 485.

49 Lada, C.J. (**2006**) *The Astrophysical Journal*, **640**, 63.

50 Mugrauer, M., Seifahrt, A., and Neuhäuser, R. (**2007**) *Monthly Notices of the Royal Astronomical Society*, **378**, 1328.

51 Eggenberger, A., Udry, S., Chauvin, G. *et al.* (**2007**) *Astronomy and Astrophysics*, **474**, 273.

52 Mugrauer, M., Neuhäuser, R., and Mazeh, T. (**2007**) *Astronomy and Astrophysics*, **469**, 755.

53 Mugrauer, M., Neuhäuser, R., Seifahrt, A. *et al.* (**2005**) *Astronomy and Astrophysics*, **440**, 1051.

54 Zucker, S., Mazeh, T., Santos, N.C. *et al.* (**2003**) *Astronomy and Astrophysics*, **404**, 775.

55 Konacki, M. (**2005**) *The Astrophysical Journal*, **626**, 431.

56 Konacki, M. (**2005**) *Nature*, **436**, 230.

57 Eggenberger, A., Udry, S., Mazeh, T. *et al.* (**2007**) *Astronomy and Astrophysics*, **466**, 1179.

58 Queloz, D., Mayor, M., Weber, L. *et al.* (**2000**) *Astronomy and Astrophysics*, **354**, 99.

59 Hatzes, A.P., Cochran, W.D., Endl, M. *et al.* (**2003**) *The Astrophysical Journal*, **599**, 1383.

60 Zucker, S., Mazeh, T., Santos, N.C. *et al.* (**2004**) *Astronomy and Astrophysics*, **426**, 695.

61 Pichardo, B., Sparke, L.S., and Aguilar, L.A. (**2005**) *Monthly Notices of the Royal Astronomical Society*, **359**, 521.

62 Jang-Condell, H., Mugrauer, M., and Schmidt, T. (**2008**) *The Astrophysical Journal*, **683**, 191.

63 Zucker, S. and Mazeh, T. (**2002**) *The Astrophysical Journal*, **568**, 113.

64 Eggenberger, A., Udry, S., and Mayor, M. (**2004**) *Astronomy and Astrophysics*, **417**, 353.

65 Rossiter, R.A. (**1924**) *The Astrophysical Journal*, **60**, 15.

66 McLaughlin, D.B. (**1924**) *The Astrophysical Journal*, **60**, 22.

67 Gaudi, B.S. and Winn, J.N. (**2007**) *The Astrophysical Journal*, **655**, 550.

68 Winn, J.N., Noyes, R.W., Holman, M.J. *et al.* (**2005**) *The Astrophysical Journal*, **631**, 1215.

69 Winn, J.N., Johnson, J.A., Marcy, G.W. *et al.* (**2006**) *The Astrophysical Journal*, **653**, 69.

70 Winn, J.N., Johnson, J.A., Peek, K.M.G. *et al.* (**2007**) *The Astrophysical Journal*, **665**, 167.

71 Loeillet, B., Shporer, A., Bouchy, F. *et al.* (**2008**) *Astronomy and Astrophysics*, **481**, 529.

72 Narita, N., Enya, K., Sato, B. *et al.* (**2007**) *The Publications of the Astronomical Society of Japan*, **59**, 763.

73 Wolf, A.S., Laughlin, G., Henry, G.W. *et al.* (**2007**) *The Astrophysical Journal*, **667**, 549.

74 Bouchy, F., Queloz, D., Deleuil, M. *et al.* (**2008**) *Astronomy and Astrophysics*, **482**, 25.

75 Narita, N., Sato, B., Ohshima, O., and Winn, J.N. (**2008**) *The Publications of the Astronomical Society of Japan*, **60**, 1.

76 Cochran, W.D., Redfield, S., Endl, M., and Cochran, A.L. (**2008**) *The Astrophysical Journal*, **683**, 59.

77 Hébrard, G., Bouchy, F., Pont, F. *et al.* (**2008**) *Astronomy and Astrophysics*, **488**, 763.

78 Winn, J.N., Johnson, J.A., Fabrycky, D. *et al. The Astrophysical Journal*, **700**, 302, arXiv:0902.3461.

79 Winn, J.N., Johnson, J.A., Narita, N. *et al.* (**2008**) *The Astrophysical Journal*, **682**, 1283.

80 Johnson, J.A., Winn, J.N., Narita, N. *et al.* (**2008**) *The Astrophysical Journal*, **686**, 649.

81 Cumming, A., Butler, R.P., Marcy, G.W. *et al.* (**2008**) *The Publications of the Astronomical Society of the Pacific*, **120**, 531.

82 Bean, J.L., Benedict, G.F., Charbonneau, D. *et al.* (**2008**) *Astronomy and Astrophysics*, **486**, 1039.

83 Sousa, S.G., Santos, N.C., Mayor, M. *et al.* (**2008**) *Astronomy and Astrophysics*, **487**, 373.

84 Beaulieu, J.-P., Bennett, D.P., Fouqué, P. *et al.* (**2006**) *Nature*, **439**, 437.

85 Bennett, D.P., Bond, I.A., Udalski, A. *et al.* (**2008**) *The Astrophysical Journal*, **684**, 663.

86 Winn, J.N., Holman, M.J., Torres, G. *et al.* (**2008**) *The Astrophysical Journal*, **683**, 1076.

87 Irwin, J., Charbonneau, D., Nutzman, P. *et al.* (**2008**) *The Astrophysical Journal*, **681**, 636.

88 Pál, A., Bakos, G.A., Torres, G. *et al.* (**2008**) *The Astrophysical Journal*, **680**, 1450.

89 Pont, F., Tamuz, O., Udalski, A. *et al.* (**2008**) *Astronomy and Astrophysics*, **487**, 749.

90 Udalski, A., Pont, F., Naef, D. *et al.* (**2008**) *Astronomy and Astrophysics*, **482**, 299.

91 Barge, P., Baglin, A., Auvergne, M. *et al.* (**2008**) *Astronomy and Astrophysics*, **482**, 17.

92 Alonso, R., Auvergne, M., Baglin, A. *et al.* (**2008**) *Astronomy and Astrophysics*, **482**, 21.

93 Moutou, C., Bruntt, H., Guillot, T. *et al.* (**2008**) *Astronomy and Astrophysics*, **488**, 47.

94 Anderson, D.R., Gillon, M., Hellier, C. *et al.* (**2008**) *Monthly Notices of the Royal Astronomical Society*, **387**, 4.

95 Wilson, D.M., Gillon, M., Hellier, C. *et al.* (**2008**) *The Astrophysical Journal*, **675**, 113.

96 Pollacco, D., Skillen, I., Collier Cameron, A. *et al.* (**2008**)

Monthly Notices of the Royal Astronomical Society, **385**, 1576.
97 Burke, C.J., McCullough, P.R., Valenti, J.A. *et al.* (**2008**) *The Astrophysical Journal*, **686**, 1331.
98 Shporer, A., Bakos, G.A., Bouchy, F. *et al.* (**2009**) *The Astrophysical Journal*, **690**, 1393.
99 Hellier, C., Anderson, D.R., Gillon, M. *et al.* (**2009**) *The Astrophysical Journal*, **690**, 89.
100 West, R.G., Anderson, D.R., Gillon, M. *et al.* (**2009**) *The Astronomical Journal*, **137**, 483.
101 Joshi, Y.C., Pollacco, D., Cameron, A.C. *et al.* (**2009**) *Monthly Notices of the Royal Astronomical Society*, **392**, 1532.
102 Bakos, G.A., Pál, A., Torres, G. *et al.* (**2009**) *The Astrophysical Journal*, **696**, 1950.
103 Johnson, J.A., Winn, J.N., Cabrera, N.E., and Carter, J.A. (**2009**) *The Astrophysical Journal*, **692**, 100.
104 Fortney, J.J., Marley, M.S., and Barnes, J.W. (**2007**) *The Astrophysical Journal*, **659**, 1661.
105 Charbonneau, D., Brown, T.M., Latham, D.W., and Mayor, M. (**2000**) *The Astrophysical Journal*, **529**, 45.
106 Fabrycky, D.C., Johnson, E.T., and Goodman, J. (**2007**) *The Astrophysical Journal*, **665**, 754.
107 Bodenheimer, P., Laughlin, G., and Lin, D.N.C. (**2003**) *The Astrophysical Journal*, **592**, 555.
108 Jackson, B., Greenberg, R., and Barnes, R. (**2008**) *The Astrophysical Journal*, **681**, 1631.
109 Sato, B., Fischer, D.A., Henry, G.W. *et al.* (**2005**) *The Astrophysical Journal*, **633**, 465.
110 Carter, J.A., Winn, J.N., Gilliland, R., and Holman, M.J. (**2009**) *The Astrophysical Journal*, **696**, 241.
111 Ikoma, M., Guillot, T., Genda, H. *et al.* (**2006**) *Astronomy and Astrophysics*, **650**, 1150.
112 Baraffe, I., Chabrier, G., and Barman, T. (**2008**) *Astronomy and Astrophysics*, **482**, 315.
113 Dodson-Robinson, S.E. and Bodenheimer, P. (**2009**) *The Astrophysical Journal*, **695**, 159.
114 Gonzalez, G. (**1997**) *Monthly Notices of the Royal Astronomical Society*, **285**, 403.
115 Santos, N.C., Israelian, G., Mayor, M. *et al.* (**2005**) *Astronomy and Astrophysics*, **437**, 1127.
116 Fischer, D.A. and Valenti, J. (**2005**) *The Astrophysical Journal*, **622**, 1102.
117 Bodaghee, A., Santos, N.C., Israelian, G., and Mayor, M. (**2003**) *Astronomy and Astrophysics*, **404**, 715.
118 Gilli, G., Israelian, G., Ecuvillon, A. *et al.* (**2006**) *Astronomy and Astrophysics*, **449**, 723.
119 Laws, C., Gonzalez, G., Walker, K.M. *et al.* (**2003**) *The Astronomical Journal*, **125**, 2664.
120 Bond, J.C., Tinney, C.G., Butler, R.P. *et al.* (**2006**) *Monthly Notices of the Royal Astronomical Society*, **370**, 163.
121 Santos, N.C., Israelian, G., Mayor, M. *et al.* (**2003**) *Astronomy and Astrophysics*, **398**, 363.
122 Ecuvillon, A., Israelian, G., Santos, N.C. *et al.* (**2006**) *Astronomy and Astrophysics*, **449**, 809.
123 Paulson, D.B., Sneden, C., and Cochran, W.D. (**2003**) *The Astronomical Journal*, **125**, 3185.
124 Desidera, S., Gratton, R.G., Scuderi, S. *et al.* (**2004**) *Astronomy and Astrophysics*, **420**, 683.
125 Desidera, S., Gratton, R.G., Lucatello, S. *et al.* (**2006**) *Astronomy and Astrophysics*, **454**, 581.
126 Reddy, B.E., Lambert, D.L., Laws, C. *et al.* (**2002**) *Monthly Notices of the Royal Astronomical Society*, **335**, 1005.
127 Vauclair, S., Laymand, M., Bouchy, F. *et al.* (**2008**) *Astronomy and Astrophysics*, **482**, 5.
128 Santos, N.C., Israelian, G., and Mayor, M. (**2004**) *Astronomy and Astrophysics*, **415**, 1153.
129 Pollack, J.B., Hubickyj, O., Bodenheimer, P. *et al.* (**1996**) *Icarus*, **124**, 62.
130 Mayer, L., Lufkin, G., Quinn, T., and Wadsley, J. (**2007**) *The Astrophysical Journal*, **661**, 77.
131 Boss, A. (**2002**) *The Astrophysical Journal*, **567**, 149.
132 Matsuo, T., Shibai, H., Ootsubo, T., and Tamura, M. (**2007**) *The Astrophysical Journal*, **662**, 1282.

133 Fuhrmann, K. (**1998**) *Astronomy and Astrophysics*, **338**, 161.

134 da Silva, L., Girardi, L., Pasquini, L. *et al.* (**2006**) *Astronomy and Astrophysics*, **458**, 609.

135 Cochran, W.D., Endl, M., Wittenmyer, R., and Bean, J.L. (**2007**) *The Astrophysical Journal*, **665**, 1407.

136 Sozzetti, A., Torres, G., Latham, D.W. *et al.* (**2006**) *The Astrophysical Journal*, **649**, 428.

137 Guillot, T., Santos, N.C., Pont, F. *et al.* (**2006**) *Astronomy and Astrophysics*, **453**, 21.

138 Burrows, A., Hubeny, I., Budaj, J., and Hubbard, W.B. (**2007**) *The Astrophysical Journal*, **661**, 502.

139 Robinson, S.E., Laughlin, G., Bodenheimer, P., and Fischer, D. (**2006**) *The Astrophysical Journal*, **643**, 484.

140 Butler, R.P., Johnson, J.A., Marcy, G.W. *et al.* (**2006**) *The Publications of the Astronomical Society of the Pacific*, **118**, 1685.

141 Endl, M., Cochran, W.D., Kürster, M. *et al.* (**2006**) *The Astrophysical Journal*, **649**, 436.

142 Bonfils, X., Mayor, M., Delfosse, X. *et al.* (**2007**) *Astronomy and Astrophysics*, **474**, 293.

143 Sato, B., Ando, H., Kambe, E. *et al.* (**2003**) *The Astrophysical Journal*, **597**, 157.

144 Setiawan, J., Pasquini, L., da Silva, L. *et al.* (**2004**) *Astronomy and Astrophysics*, **421**, 241.

145 Lovis, C. and Mayor, M. (**2007**) *Astronomy and Astrophysics*, **472**, 657.

146 Cayrel de Strobel, G., Soubiran, C., and Ralite, N. (**2001**) *Astronomy and Astrophysics*, **373**, 159.

147 Bean, J.L., Sneden, C., Hauschildt, P.H. *et al.* (**2006**) *The Astrophysical Journal*, **652**, 1604.

148 Ida, S. and Lin, D.N.C. (**2005**) *The Astrophysical Journal*, **626**, 1045.

149 Laughlin, G., Bodenheimer, P., and Adams, F.C. (**2004**) *The Astrophysical Journal*, **612**, 73.

150 Kennedy, G.M. and Kenyon, S.J. (**2008**) *The Astrophysical Journal*, **673**, 502.

151 Boss, A. (**2006**) *The Astrophysical Journal*, **643**, 501.

Further Reading

Johns-Krull, C.M., McCullough, P.R., Burke, C.J. *et al.* (**2008**) *The Astrophysical Journal*, **677**, 657.

2
Pinpointing Planets in Circumstellar Disks

Alice C. Quillen

2.1
Introduction

A crowning achievement in mathematical science was the prediction by Urbain Le Verrier and John Couch Adams in 1845 of a previously unknown planet that was then subsequently discovered. Irregularities in Uranus's orbit were elegantly explained as being due to a single massive object exerting perturbations on Uranus and the position of the object in the sky was predicted. The planet Neptune was subsequently detected by Johann Gottfried Galle at the Berlin Observatory (cf. [1]). Planets can not only perturb each other, as outlined in the original perturbative theory by Laplace, but also perturb the orbits of smaller particles such as dust that can be much more numerous and so can reemit a detectable fraction of the host star luminosity. Disk morphology can provide dynamical evidence for massive bodies that can be subsequently discovered from direct observations. We can refer to historical examples in our solar system. For example, the prediction of a satellite in the Encke gap in Saturn's rings based on wavy edges seen in the gap [2], was followed by the discovery of the satellite, Pan [3].

In this chapter, a broad base of dynamical work relating circumstellar disk morphology to underlying planetary architecture is reviewed. The existence of a planet was predicted on the basis of its interactions with the circumstellar disk in the Fomalhaut system [4] at a semimajor axis of 119 AU. Subsequently, a planet was discovered with an estimated semimajor axis of 115 AU [5], only 4 AU lower than predicted. This is the first planet whose existence was successfully predicted from its interactions with the surrounding circumstellar disk (see Figure 2.1). discovery illustrates that we can, in some cases, predict the presence of planetary bodies from their interactions with disk material. The discrepancies between the predicted planetary orbit and mass from those inferred from observations are and will continue to be interesting. In probing them, we will improve our theoretical understanding and extend our capabilities so that we can accurately infer planetary system architecture at levels below the concurrent detection limits. In this chapter, we discuss some of the broad foundation of work that this prediction and others are built

Formation and Evolution of Exoplanets. Edited by Rory Barnes

ISBN: 978-3-527-40896-2

Figure 2.1 This image, taken with the Advanced Camera for Surveys (ACS) aboard the NASA/ESA Hubble Space Telescope, shows the newly discovered planet, Fomalhaut b, orbiting its parent star, Fomalhaut [5]. A planet at a semimajor axis that is only 4 AU lower than that estimated from the observations was predicted prior to discovery from the 2004 ACS image [4]. The prediction was based on the steeply sloped edge of the dusty disk, the infrared dust opacity, the eccentricity of the ring, and the assumption that a single planet is responsible both for truncating the disk and maintaining the eccentricity of the clearing edge.

upon. We focus more on scenarios and explanations for phenomena than on the extensive accompanying body of observations that has inspired theoretical efforts.

2.2 Signatures of Extrasolar Planets Imprinted on Dusty Disks

The Infrared Astronomical Satellite (IRAS) discovered a number of nearby stars with bright infrared excesses. The emission was subsequently interpreted in terms of dusty circumstellar disks originating from collisions between planetesimal, asteroidal, and/or cometary debris [6]. Because of the short lifetime of dust particles in these disks, and the scarcity of gas, the dust must continually be replenished from collisions of larger orbiting bodies, possibly analogous to Kuiper belt objects. Hence the disks in the older systems are commonly called *debris disks* [7].

At nearly the same time, following the launch of IRAS, bands discovered in the spatial distribution of the zodiacal cloud dust [8, 9] were attributed to dust originating in distinct asteroidal families [10, 11]. Dust particles in our solar system spiral toward the Sun due to radiation drag forces, known as *Poynting–Robertson* ($P-R$) drag, and interactions with the solar wind [12]. Gravitational perturbations from planets can scatter the particles or capture them into orbital resonances with

planets. Since the lifetime of captured particles can greatly exceed the lifetime of particles that are not captured, resonant trapping can sculpt a low-density dusty disk. The existence of a dust-filled resonant circumsolar ring comprised of particles locked in resonances with the Earth was predicted by Jackson and Zook [13]. Subsequently, IRAS observations of dust near the Earth discovered evidence for this ring [14].

More recent studies proposed that extra-solar planets could leave signatures on the dust distribution of their circumstellar disks. Liou and Zook [15] first showed that dust originating from the Kuiper belt was likely to contain both a resonant structure associated with capture of drifting dust particles into mean motion resonances with Neptune and an evacuated region interior to Neptune due to scattering with the giant planets. Planets with masses as low as that of Mars could produce detectable circumstellar rings in exozodiacal disks [16]. Ozernoy *et al.* [17] proposed that the structure of Fomalhaut's and Epsilon Eridani's circumstellar disks, as seen in submillimeter observations [18], could be explained in terms of dust particles captured into resonances with giant planets as they drifted inward due to P–R drag. Models with planets on circular orbits usually do not display strongly peaked or clumped features [19]. Subsequent work also explored models with eccentric planets [20–22] proposing that Vega's and Epsilon Eridani's disk hosted eccentric giant planets at large semimajor axes and predicting the planet's eccentric, semimajor axis, mass, and angle of periastron. The location of peaks in the dust distribution can be predicted using angles associated with different resonances for different masses and eccentricities [23].

These initial works were influenced by the structure seen in the solar system dust distribution and explored models where the dust distribution was dominated by particles trapped in resonance with a planet. Subsequent models have begun to expand the range of physical regimes considered by including collisions and gas/planet coupling. Before exploring physical processes important to interpreting morphology of circumstellar disks in terms of underlying unseen bodies, the morphological features inferred from observations of circumstellar disks are reviewed.

2.3
Morphological Features of Circumstellar Disks

In this section, observed circumstellar disk morphology is first described and some of the associated proposed explanations are listed. A variety of features can be used to place constraints on unseen embedded objects. The slope of a gaseous disk edge can constrain the planet mass if a planet truncates the disk [24]. A change in the surface brightness profile could be due to high eccentricity or unbound small grains produced in a planetesimal disk edge [25, 26]. The thickness of a disk may suggest that massive objects reside in the disk, which have thickened the disk due to gravitational scattering or stirring [27–29]. Spiral features and warps might be explained by variations in illumination [30, 31], instabilities [32], perturbation

by a massive body [33], or a collisional dust avalanche [34]. Clumps in a disk are more challenging to explain, but could be due to resonant trapping of particles by a planet, or particles generated in resonance by objects that have been captured into resonance [35–37]. Clumps could also be transient features due to recent catastrophic events such as collisions between planetesimals [34, 38, 39].

2.3.1 Clearings

Spitzer Space Telescope surveys of young clusters have found that about 5% of young, million-year old stars host circumstellar disks with inner clearings [40]. The clearing edges are inferred from the temperature of the rim as determined from a fit to the spectral energy distribution [41]. Many of the stars with clearings continue to accrete at low rates and contain fainter hot inner disks residing in the clearing [42]. CoKuTau4 has little hot dust and although its clearing was originally interpreted to host a planet just within its clearing edge [43], the central star is now known to be a close stellar binary [44] and so a planetary explanation for the clearing is ruled out.

A gap in a disk is difficult to detect from a spectral energy distribution unless it is very wide, as is believed to be the case for GM Aurigae [41]. Clearing edges could be maintained by a planet [43, 45] or could be due to photoevaporation from the central star [46]. However, disk masses are sufficiently high that photoevaporative models fail to explain all clearings [42]. If planets are proved responsible for clearings seen in gaseous disks hosted by young stars then rough constraints on the masses of the planets can be made from the properties of the disk [43, 47].

Spitzer Space Telescope surveys also find older and lower optical depth systems with dusty disks [48]. In many cases, the temperature of the disks is consistent with a single temperature model [7] suggesting that there is a large clearing. Clearing of dust probably requires multiple planets [49]. Some objects host dust at multiple temperatures suggesting multiple dust rings and so multiple confining planets [50, 51]. The star HD8799 that hosts three giant planets [52] seems to have a wide gap in its dust distribution covering the region containing the three planets (Chen, personal communication). Other than crude estimates based on clearing timescales, little work has been done on what types of planetary systems are likely to reside in systems with multiple dust rings.

While a few hundred objects are known with infrared excesses, only for a handful has it been possible to image the disk. Fomalhaut has an eccentric clearing [53] with a fairly sharp edge [54]. Secular perturbations by an eccentric planet could account for the eccentricity of a ring [55]. If only secular perturbations are considered, the ring eccentricity is insensitive to the planet mass and the planet location is not constrained [56]. If the same planet is also responsible for truncating the clearing edge then the degeneracies of the secular perturbation model can be broken [4]. The ring eccentricity, disk edge slope, and disk opacity can be used to construct a model with a planet just interior to the ring edge that specifies this planet's mass, eccentricity, and semimajor axis [4, 57]. For both the gaseous and particle

disks, these models assume that a disk edge is maintained and stays steep because material interior to the clearing is scattered away by the planet. For the gaseous disk edge, the planet must be massive enough that torques due to spiral density waves driven by the planet overcome viscous diffusion [24]. For the particle disk edge, particles that are scattered into the clearing by collisions are ejected by close approaches to the planet [4, 58].

A number of objects host wide gaps or multiple dust rings [41, 59]. Only the approximate radial distribution can be estimated from a spectral energy distribution alone. While wide gaps (with width to semimajor axis ratio $\delta a/a$ of order 1) have been found, narrow ones such as are common in Saturn's rings (with $\delta a/a \ll 1$) have not yet been clearly seen in circumstellar disks. Multiple planet and planetesimal belt dynamical models are exciting settings for future exploration.

2.3.2 Spiral Arms

A few circumstellar disks, seen in scattered light have spiral-like structure. Circumstellar disks seen in scattered light include HD100546 [60, 61], HD141569A [62], and AB Aurigae [32]. Many models have been proposed to account for this structure but most have been ruled out. A chance encounter with another star could perturb the disk but is unlikely to truncate the disk [30]. Moreover, spiral structure is only excited for a short period of time after a close stellar encounter making such an event extremely improbable. For HD 100546, this scenario was ruled out based on estimates of the mass required for the encountering object to excite two arms and the timescale since encounter required to match winding of the spiral structure. HD 141569A, on the other hand, is a binary system and so could suffer repeated encounters that would both truncate the outer disk and pull out faint spiral arms. For HD 141569A's inner disk, a transient response to secular perturbations have been proposed as a model for its asymmetries [55]. Self-shadowing models [63] for HD100546's spiral arms were also ruled out [61] based on a lack of color variation in the arms. Models for the spectral energy distribution suggest that the disk is quite thick. If gravitational instabilities cause the spiral structure then an extremely cold component must reside in the midplane. Color variations are used to support collision avalanche models [39, 64]; however, these models might not be excluded by the low color gradient.

2.3.3 Clumps

A number of debris disks exhibit clumps or peaks in the dust distribution. AU Mic and Beta Pic are extended systems with long-lasting clumps that have been detected by multiple observations [39, 65, 66]. Models based on an embedded migrating planet tend to produce weaker features [37]; however, recent catastrophic collisions may provide a better explanation for these features [34, 39, 64]. For systems with inner clearings and clumps, such as Vega and Epsilon Eridani, most

models involve resonant trapping of inwardly drifting dust particles or particles generated in resonance [16, 20, 23, 35, 36]. These models tend to produce density contrasts lower than observed; however, collisions could affect the dust distribution and simulations including collisions have only just begun to be explored. Diffusive collision models suggest that resonances only cause structure in a ring when the planet mass is high and can cause a deficit of particles at certain angles and radii rather than stable long-lived clumps [67]. On the other hand, simulations of small particles more strongly affected by radiative forces may not show resonant structure even when the planet mass is high [58].

2.3.4 Important Timescales

Currently imaged circumstellar disks are primarily young objects, which is perhaps not surprising since circumstellar material is eventually blown away by winds and radiation from the star or becomes bound into planets and planetesimals. The fraction of objects with infrared excesses drops steeply with increased age and is interpreted in terms of planet formation, rather than more passive disk evolution [68].

The orbital period for a particle with semimajor axis a is

$$P_{\rm orb} = \sqrt{a^3/GM_*} = 1\ {\rm yr}\left(\frac{a}{\rm AU}\right)^{3/2}\left(\frac{M}{M_*}\right)^{-1/2} \tag{2.1}$$

This timescale can be long, for example, $P \sim 1000$ years for the 130 AU disk hosted by Fomalhaut. It is convenient to work with a timescale $t_{\rm dyn} = n^{-1}$ that is related to the mean motion n, which corresponds to the angular rotation rate for a particle in a circular orbit.

The timescale for dust particles to collide, inferred from infrared excesses, depends on their area filling factor or optical depth, τ, normal to the disk plane. The timescale for collisions is

$$t_{\rm col} \sim t_{\rm dyn}\,\tau^{-1} \tag{2.2}$$

In the zodiacal cloud this timescale is long enough that collisions can be neglected [69]; however, for most observed debris disks τ is in the range of 10^{-3} to 10^{-4} [7], leading to collision timescales only 1000 to 10 000 times longer than the dynamical timescale. The collision timescale estimated from the smallest particles (the dust) is substantially shorter than for planetesimals. Nevertheless, the planetesimal reservoir is suspected to be feeding a collisional cascade that is responsible for dust production. Approximate estimates for the number and masses of planetesimals can be made using power laws [29], though care should be taken when extrapolating over 10 orders of magnitude and when planetesimal growth is not taken into account.

Dust particles in a circular orbit decay due to P–R drag on a timescale proportional β_r^{-1}, where β_r is the ratio of radiation to gravitational (from the star) forces. It is

convenient to write

$$\beta_r \sim \frac{0.2}{s_{\mu m}} \left(\frac{L_*}{L_\odot}\right) \left(\frac{M_*}{M_\odot}\right)^{-1} \tag{2.3}$$

where $s_{\mu m}$ is the radius of the particle in μm and L_* is the luminosity of the star [70]. For extremely small particles, β_r approaches a constant and the above approximation is no longer valid [69].

The timescale for orbital decay due to P–R drag for $\beta_r < 1$ is

$$t_{pr} \sim 400 t_{dyn} \beta_r^{-1} \left(\frac{r}{AU}\right)^{1/2} \tag{2.4}$$

where the extra factor of r arises because the decay timescale depends on the ratio of the particle's velocity to the speed of light [69].

Most known debris disk systems have $t_{pr} > t_{col}$ [71] implying that the currently observed circumstellar disks should not be modeled without collisions. It is likely that the resonant population of dust particles is much reduced in simulations that include collisions.

The younger systems contain gas and so a viscous or accretion timescale is relevant. The accretion timescale for a gas disk is

$$t_\nu = t_{dyn} \alpha^{-1} \left(\frac{r}{h}\right)^2 \tag{2.5}$$

where h/r is the ratio of the disk height to radius and α is the parameter setting the disk viscosity. For a disk aspect ratio of 0.05, based on thermal models of the disk, we estimate that t_ν is 10^4 to 10^5 times longer than the dynamical timescale. For planets that open gaps, planet migration is also expected to take place on this timescale [72]. In some cases, a disk of particles with frequent inelastic collisions can be considered an accretion disk. If the collisions are infrequent or destructive, rather than nearly elastic, then the disk will not act like a viscous accretion disk.

When a planet is present in a disk, timescales associated with the planet are important. We first consider secular perturbations or those present when we average over the orbital motion. The orbit of a particle exterior to a planet precesses at a frequency

$$\dot{\varpi} \sim \frac{n\mu\alpha}{4} b^1_{3/2}(\alpha) \tag{2.6}$$

where $b^1_{3/2}(a/a_p)$ is a Laplace coefficient, μ is the mass ratio of planet to host star, $\alpha = a/a_p$ and a and a_p are the particle's and planet's semimajor axis. The precession rate is faster nearer the planet, diverging as da^{-1}, where $da = a - a_p$ is the difference in semimajor axis between particle and planet. When multiple planets are present the secular dynamics can be much richer [73] .

Also relevant is a timescale for a particle initially in a circular orbit to be first scattered by a planet. For multiple planet systems, Chambers *et al.* [74] measured a relationship between planet spacing and timescale to the first close planet/planet encounter. This timescale depends on the spacing and planet mass, and the number of planets, though it approaches a limiting value as the number of planets

increases. For systems with equally spaced (logarithmically), equal mass planets the time to first close encounter is

$$\log_{10}(t_e/P_{\text{inner}}) \approx -0.1 - \log_{10}(\mu/10^{-7}) + 3.2(\delta/\mu^{1/4}) \tag{2.7}$$

where $1+\delta$ is the ratio of the distance between two neighboring planets and t_e is given in terms of the orbital period of the innermost planet, P_{inner} [49]. It is not known why this relationship scale with the planet-mass ratio to the power of 1/4. This timescale was used to roughly estimate the total of number planets likely to reside inside a clearing in a dust disk assuming that the dust and planetesimals are removed from the clearing on a timescale shorter than the age of the system [49]. Quillen [4] numerically estimated the timescale for particles scattered via collisions into the corotation region near a planet to be scattered away by the planet and predicted the mass of a planet likely to reside just within a dust ring. Better estimates for clearing would help in better understanding how the timescales and regions of stability depend on planetary number, masses, and configuration.

The last timescale that is now introduced is the period of libration in a mean motion resonance. This timescale is important as spiral density waves can only be driven into a particulate disk when the libration period exceeds the collision timescale. The timescale also determines whether a drifting system can be captured into resonance. Hence the mechanism for angular momentum transfer between planet and disk depends on this last timescale. To estimate this timescale in terms of planet mass and resonance, we must understand mean motion resonances.

2.4 The Role of Mean Motion Resonances

The mass of even giant planets is very low compared to that of a host star and as such gravitational forces from planets are always perturbative unless an object approaches within the planet's Hill sphere. For example, the mass of Jupiter is only a thousandth of that of the Sun. Nevertheless, Jupiter has a strong influence on other bodies in the solar system. It is responsible for most bodies that have been ejected from the solar system into the Oort cloud. Resonances with Jupiter have sculpted the asteroid belt. In the inner asteroid belt, empty regions in the semimajor axis distribution, known as *Kirkwood gaps*, are regions where initially low eccentricity objects have their eccentricities increased and are removed by planets when they cross planetary orbits. In the outer asteroid belt and Kuiper belt, resonances can be regions of stability where objects are prevented from close approaches to planets by their strong resonant interactions. The gravitational force from a planet is primarily important in resonances or within the planet's Hill sphere.

Resonant interactions are thought to be important in a number of different settings. They are the locations where spiral density waves can be driven into a

cold collisional or gaseous disk [75]. Thus angular momentum transfer between a planet and a gaseous or collision-dominated disk depends on resonant coupling.

Resonances can transport particles to regions of high eccentricity where a particle can be scattered by a planet. A sharp dynamical boundary exists where mean motion resonances overlap near the planet's corotation radius. Inside this region, known as the *chaotic zone*, orbits are strongly chaotic [76]. Because the mean motion resonances overlap [76], a particle in this region will cross the planet's orbit and so eventually be scattered by the planet [4]. Thus resonances are important for models of disk clearing and interpretation of disks with inner clearings as inferred from their images or spectral energy distributions.

When a system is drifting, particles can be captured into resonance. Planets migrating outward or inward can capture particles into resonance. Dust particles spiraling inward can also be captured into resonance with a planet.

2.4.1 General Theory for First-Order Mean Motion Resonances

Mean motion resonances occur when the particle mean motion, n, can be related to the mean motion of the planet, n_p:

$$in \approx jn_p \tag{2.8}$$

with i, j integers.

We can integrate Eq. (2.8) with respect to time, finding

$$\phi = i\lambda - j\lambda_p \approx \text{constant} \tag{2.9}$$

where λ_p is the planet's mean longitude. We refer to the slowly varying angle ϕ as a resonant angle.

This is analogous to a pendulum with Hamiltonian

$$H(p, \phi) = \frac{1}{2}p^2 + A\cos\phi \tag{2.10}$$

which exhibits two types of motion, each separated by a limiting curve called a *separatrix*. Motion about the fixed point, $p = 0, \phi = \pi$ is similar to that of a harmonic oscillator with period $2\pi/\sqrt{A}$. Here the angle ϕ librates about its equilibrium value. However, for large p, all motion involves continuously increasing values of ϕ. Physically, this corresponds to the pendulum swinging in a large circle rather than gently oscillating about its equilibrium state.

Near mean motion resonances, a resonant angle given as ϕ in Eq. (2.9) varies slowly. Hence, these are regions where small perturbations can become important and cause the resonant angle to librate about a fixed value rather than oscillate. We consider a particle trapped in resonance when the resonant angle librates about a particular value. A particle is not in resonance when the resonant angle continuously increases or decreases.

For the Keplerian system restricted to the plane, we employ the Poincaré coordinates $\lambda = M + \overline{\omega}$, $\gamma = -\overline{\omega}$ and their associated momenta

$$L = \sqrt{GM_* a}, \qquad \Gamma = \sqrt{GM_* a}(1 - \sqrt{1 - e^2}) \tag{2.11}$$

where M_* is the mass of the star, λ is the mean longitude, M is the mean anomaly, $\overline{\omega}$ is the longitude of pericenter, a is the semimajor axis, and e is the eccentricity. These variables describe the orbit of a particle or planetesimal in a plane. For the Keplerian system the unperturbed Hamiltonian, or that lacking the disturbing function describing perturbations from a planet, is $H_0(L, \lambda; \Gamma, \gamma) = -1/2L^2$. When expanded near a mean motion resonance, the Hamiltonian is

$$K_0(\Lambda, \psi; \Gamma, \gamma) = a'\Lambda^2 + b'\Lambda + \text{constant} \tag{2.12}$$

with coefficients given, for example, by Quillen and Faber [57]. Here the angles for a $j : j - k$ exterior (to the planet) resonance are $\psi = j\lambda - (j - k)\lambda_p$ and $\gamma = -\overline{\omega}$. Subtracting the factor depending on the mean longitude of the planet, λ_p, (as seen in the ψ angle), is akin to working in the frame rotating with the planet. We note that Eq. (2.12) has been expanded to second order in the momenta. Consider a dynamical system like the pendulum expanded containing only a first-order term in p,

$$H(p, \phi) = \kappa p + A\cos\phi \tag{2.13}$$

This only has the solution $\dot{\phi} = \kappa$. Hence ϕ increases forever; solutions only oscillate and never librate about a fixed value. To exhibit both types of motion, libration, and oscillation, an expansion to second order in the momenta variables is required. Another problem with lower order expansions is that solutions can incorrectly blow up near resonances. To exhibit the richer dynamics present in resonances, at least a second-order (in momenta) expansion is required.

For the Keplerian system, perturbations are described by the disturbing function. Taking into account only resonant first-order terms (and neglecting secular terms), this can be written for a first-order resonance

$$R = g_0\Gamma^{1/2}\cos(\psi - \overline{\omega}) + g_1\Gamma_p^{1/2}\cos(\psi - \overline{\omega}_p) \tag{2.14}$$

where we have used the approximation $e^2 \sim 2\Gamma/L \sim 2\Gamma\alpha^{1/2}$ where α is the ratio of the particle and planet semimajor axes. Here the perturbation strengths, g_0 and g_1, are functions of semimajor axis, j and Laplace coefficients, which are proportional to the planet mass. The longitudes of pericenter, $\overline{\omega}, \overline{\omega}_p$ are those for the particle and planet, respectively. The second term is sometimes called the *corotation resonance* and is often ignored. The parameter g_1 depends on the planet's eccentricity. The resonant angle for the dominant resonance term is $\phi = \psi - \overline{\omega}$, where ψ depends on the mean longitude of the planet and particle.

The Hamiltonian near a first-order mean motion resonance is the sum of Eq. (2.12) (representing the unperturbed system) and (2.14) (the dominant terms in the disturbing function). Following canonical transformation to new coordinates, specifying one to be the resonant angle, ϕ, and neglecting the corotation resonance term,

$$H(J, \phi) = a'J^2 + J\delta - g_0J^{1/2}\cos\phi \tag{2.15}$$

Here δ describes the distance from resonance and is dependent on the semimajor axis. The parameter g_0 is proportional to the planet mass.

It is useful to estimate dimensions from the parameters describing the Hamiltonian. The parameter δ only sets the distance to resonance, so we can neglect it during rescaling. Dimensionless coordinates can be chosen as j, τ with

$$j = |g_0|^{2/3}|a'|^{-2/3}J \tag{2.16}$$

$$\tau = |g_0|^{2/3}|a'|^{1/3}t \tag{2.17}$$

Equation 2.16 gives a timescale

$$t_{\rm res} = |g_0|^{-2/3}|a'|^{-1/3} \tag{2.18}$$

which can be computed for any resonance (see [77]) and approximately corresponds to the period of libration or eccentricity variation in the resonance. A comparison between this timescale and the collision timescale determines if spiral density waves can be driven into the disk. For example, using the numbers listed in Table 2.1 by Quillen [77], we estimate that the libration timescale for the 4 : 3 exterior mean motion resonance is

$$t_{\rm res} \sim 20\left(\frac{q}{10^{-3}}\right)^{2/3} n^{-1} \tag{2.19}$$

where n is the mean motion of the planet and q is the ratio of planetary to stellar mass. The external 3 : 2 resonance has a timescale that is about 1.7 times longer and the 2 : 1 resonance is about 10 times longer. The 2 : 1 resonance is weak because the indirect terms partly cancel the direct terms in the disturbing function.

The Hamiltonian lacking the corotation term can be written in dimensionless form as

$$H(j,\phi) = j^2 + j\overline{\delta} - \overline{\beta}j^{1/2}\cos\phi \tag{2.20}$$

where $\overline{\delta}$ and $\overline{\beta}$ depend on the scaling factors given in Eqs. (2.16) and (2.17) and the coefficients in Eq. (2.15).

A particle that is migrating or drifting has a mean motion that is varying. The square of $t_{\rm res}$ is relevant in this situation. A comparison between the square of $t_{\rm res}$ and a drifting or migration timescale determines when drifting is sufficiently fast so that resonances fail to capture particles because they are drifting in the nonadiabatic limit [77]. Because small particles drift more swiftly due to radiation forces, they are less likely to be captured into resonance. For example, Table 2.2 by Quillen [77] estimates that the planetary drift rate allowing objects to be captured into the external 3 : 2 resonance must be lower than

$$\dot{n} < 0.02n^{-2}\left(\frac{q}{10^{-3}}\right)^{4/3} \tag{2.21}$$

The critical drift rate is about 1/4 of this for the 3 : 2 resonance and about 400 times slower for the external 2 : 1 resonance. A similar expression can be used to determine the minimum-sized dust particle that can be captured into resonance as it drifts due to P–R drag. Particles that drift too quickly cannot be captured into weak resonances with long libration times.

The role of eccentricity can be estimated by considering the role of the corotation resonance term in Eq. (2.14). If orbits are nearly closed (or have zero-free

eccentricity) then the dynamics are identical to that at low eccentricity near a zero eccentricity planet [57]. Otherwise, the corotation resonance can prevent resonance capture. If particles are scattered by planetesimals then the eccentricity distribution may become sufficiently wide that resonance capture is reduced [78]. The eccentricity required to reduce the probability of capture can be computed using the rescaling coefficients for momentum shown in Eq. (2.16).

Mininum eccentricities ensuring capture in the adiabatic limit, $e_{\rm lim}$, for first-order resonances are approximately proportional to the cube root of the planet-mass ratio [77] or about

$$e_{\rm lim} \sim 0.01 \left(\frac{q}{10^{-3}}\right)^{1/3} \tag{2.22}$$

2.5 When are Spiral Density Waves Important?

The separation between collisional and collisionless disks is important as the opacity of the disk (setting the collision timescale) is an observable. Spiral density waves are efficiently driven at a Lindblad resonance by a satellite when the collision timescale is above a critical one, $t_{\rm crit}$, where $t_{\rm crit} \propto \mu^{-2/3}$, and $\mu \equiv \frac{m_p}{M_*}$ is the ratio of the planet mass divided by that of the star [75, 79]. This scaling is predicted by comparing the period of excited epicyclic oscillations at a Lindblad resonance with the collision timescale [79]. The period of excitations at a $j:j-1$ mean motion resonance is approximately equal to the renormalization factor in Eq. (2.17) or

$$p_e \sim n^{-1}|\beta a'|^{-2/3} = n^{-1}\left|\frac{3j^2\mu\sqrt{2}}{2\pi\, da}\right|^{-2/3} \tag{2.23}$$

with coefficients evaluated in the limit of large j (see [77]). In the limit of small da, and setting the critical timescale to this period, $t_{\rm crit} = p_e$,

$$t_{\rm crit} n \sim \mu^{-2/3} j^{-2} \sim \mu^{-2/3} da^2 \tag{2.24}$$

We have recovered the scaling with the planet predicted by previous works. Resonant libration periods are longest for the mean motion resonances most distant from the planet. As the opacity drops, the last mean resonance capable of driving spiral density waves would be the 2 : 1 mean motion resonance.

2.5.1 The Chaotic Zone Boundary

Quillen [4] proposed that a natural boundary for a collisional disk in the proximity of a planet would be the chaotic zone associated with the region of corotation. Nearing the planet, mean motion resonances become closer together and stronger. Their widths can be estimated using the scaling in Eq. (2.16), as the momentum in this equation sets the particle eccentricity and the semimajor axis and eccentricity are

related in the resonance. When the resonance widths exceed the distance between them, motion is chaotic [76]. The zone boundary is at

$$da_z \sim 1.3\mu^{2/7} \tag{2.25}$$

where da_z is the difference between the zone edge semimajor axis and that of the planet divided by the semimajor axis of the planet. For planet-mass objects, the width of this chaotic zone exceeds that set by the Hill radius because 2/7 is smaller than 1/3. The Hill radius is more relevant for setting the distance to the planet if the disk is gaseous [24, 80]. Simulations incorporating collisions in a diffusive limit suggest that the chaotic zone boundary does set the distance between the disk and the planet [67] but may be larger than that shown in Eq. (2.25). Instead of 1.3 the coefficient might be as large as 1.5 [57] or 2.0 [58]. If the disk is composed of particles in nearly circular orbits and truncated by a low eccentricity planet, then the chaotic zone boundary is the same as given by Eq. (2.25). This is true because dynamics at low free eccentricity near a low eccentricity planet is similar to that at zero eccentricity near a planet in a circular orbit [57].

2.6
Minimum Gap Opening Planet Masses

Instead of asking the question, ''when can we predict a planet?'', it is interesting instead to ask, ''when can we rule out the presence of a planet?'' The most natural setting for this is a gapless disk.

For a viscous disk, the planet is coupled to the disk via driving spiral density waves. A planet will open a gap in a viscous disk if the sum of torques due to all-driven spiral density waves exceeds the viscous torque [72, 81, 82]. This condition is approximately equivalent to

$$\mu > 50\mathcal{R}^{-1} \tag{2.26}$$

where μ is the planet-mass ratio, $\mathcal{R} = \frac{r_p^2 n_p}{\nu}$ is the Reynolds number, r_p and n_p are the radius and mean motion of the planet, and ν is the disk viscosity [80]. The viscosity depends on the sound speed in the disk and so depends on the thermal structure of the disk. On the basis of a simple temperature model for the disk, the minimum gap-opening planet mass is estimated to be

$$\mu_{\min} \sim 4 \times 10^{-4} \left(\frac{\alpha}{0.01}\right)^{0.8} \left(\frac{L_*}{L_\odot}\right)^{-0.08} \times \left(\frac{\dot{M}}{10^{-8} M_\odot/\mathrm{yr}}\right)^{0.48} \left(\frac{M_*}{M_\odot}\right)^{-0.58} \tag{2.27}$$

where $\dot{M}$ is the disk accretion rate [47]. Similar estimates were made befor by Menou and Goodman [83]. The actual value of the minimum gap-opening planet mass as a function of radius depends on the temperature structure of the disk and so depends on whether the disk is primarily heated by accretion or radiation

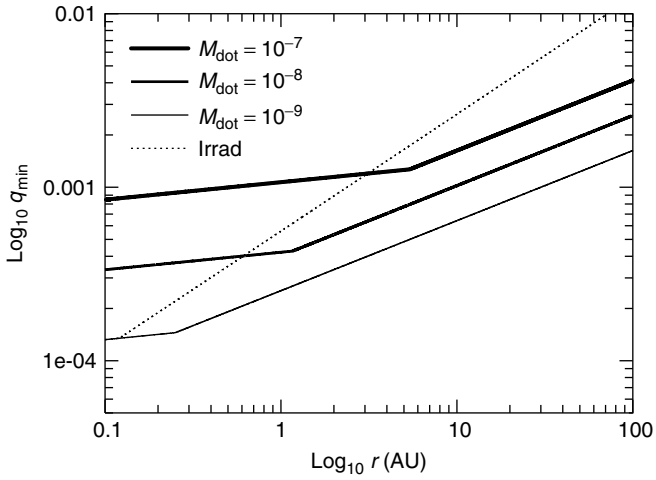


Figure 2.2 The minimum gap-opening planet-mass ratio as a function of radius for gaseous disks with accretion rates $\dot{M} = 10^{-7}, 10^{-8}$, and $10^{-9} M_{\odot}$/year (from [47]). The transition radius is evident where the dotted line intersects the solid lines. This radius is where midplane temperature from viscous dissipation is similar to that from stellar radiation. Inside this radius the minimum gap-opening planet mass is not strongly sensitive to radius. Outside this radius a larger planet mass is required to open a gap, as long as the disk is sufficiently flared to be heated by starlight. Lower mass planets can open a gap at the larger transition radius (slope change of the solid lines) in self-shadowed disks.

from the central star. Figure 2.2 shows minimum gap-opening planet masses as a function of radius for a simple accretion disk model, showing that only gas giants are likely to open gaps and only when the accretion rate drops at the late stages of evolution.

When the disk is not viscous, the minimum gap-opening mass is more difficult to estimate. A natural boundary for collisionless particles is the edge of the chaotic zone associated with the planet's corotation radius [4]. However, unless there is a diffusive mechanism operating in the disk edge, planet of any mass could open a gap. A diffusive model for collisions was investigated to determine the minimum gap-opening planet mass in intermediate opacity disks [67]. In this regime, less massive planets can open gaps than in the viscous setting (see Figure 2.3).

2.7 Fomalhaut

Prior to the detection of a planetary object in the Fomalhaut system [5], there was speculation that the disk clearing hosted one or more planets. An asymmetry in the disk was detected in submillimeter and infrared wavelengths [50, 53, 56] with clearing edge eccentricity, $e \sim 0.1$, most precisely measured at visible wavelengths [54]. A single planet on an eccentric orbit was proposed as a possible explanation for

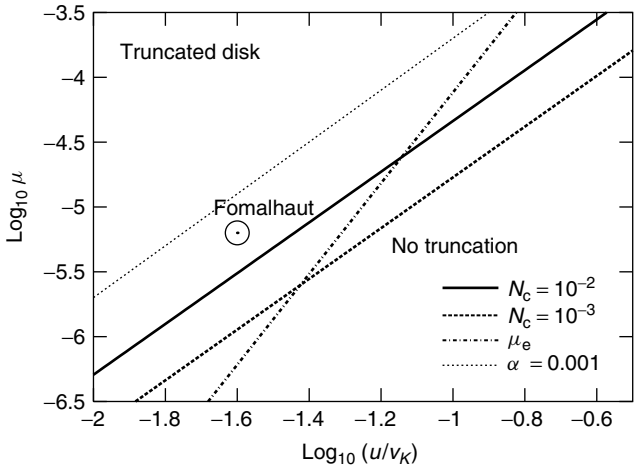


Figure 2.3 Particle disks can be truncated by planets with a mass ratio that lies above the lines drawn (from [77]). The *x*-axis is the disk velocity dispersion in units of the circular velocity and can be estimated from the disk thickness. The *y*-axis is the minimum gap-opening planet-mass ratio. The solid and dashed line refer to particle disks with $N_c = 10^{-2}$ and 10^{-3} collisions per particle per orbit, respectively. The thin dotted line is the gap-opening criteria expected for an α accretion disk with $\alpha = 0.001$. A lower mass planet can open a gap in a particle disk than an accretion disk. The dot dashed line is a limit set by requiring that the disk thickness be lower than chaotic zone width. A circle has been placed at the estimated planet mass lower limit for a planet inside Fomalhaut's disk edge using $N_c = 3 \times 10^{-2}$ expected from the normal optical depth, $\tau_n = 1.6 \times 10^{-3}$ and velocity dispersion estimated from the disk edge slope $u/v_K \sim 0.02$.

the ring's eccentricity in the context of the pericenter glow model [53]. This model uses first-order secular perturbation theory to account for the ring eccentricity and was originally proposed for the HR 4796A system [55]. As Fomalhaut's ring is eccentric, axisymmetric scenarios that do not involve planets and involving radiative forces that could account for the steepness of the ring edge [84] are less attractive explanations for the eccentricity of the ring. Now that a planetary object has been discovered just interior to the ring edge in the Fomalhaut system, explanations that do not involve planets would only be considered for other disks with clearings, such as HR 4796A, if planets are ruled out. However, eccentric modes could account for lopsidedness and radiative effects may truncate a disk in the absence of planets.

Within the context of the pericenter glow model, the ring eccentricity is insensitive to planet mass. Only a function of both the planet eccentricity and semimajor axis is constrained, though March *et al.* [56] used the secular perturbation timescale to place limits on the planet's mass and location. Quillen [4] proposed that a single planet could account for both the ring eccentricity and edge sharpness. This assumption was used to break the degeneracies of the pericenter glow model. The Fomalhaut system was used to break these degeneracies because its sharp disk edge is better resolved than other disk edges. When the planet and ring are in proximity, the eccentricity and orbital alignment of the planet and ring are approximately

equal. The planet mass was constrained to be sufficiently massive that it could keep the clearing empty but would not cause the disk edge to be thicker than observed. Thus the predicted planet mass is sensitive to the collision timescale estimated from the dust opacity. The chaotic zone was proposed as a possible boundary separating planet and disk. The prediction was a planet with orbital eccentricity of that of the ring, located just interior to the edge and apsidally aligned with the ring edge. The predicted planet has a mass between that of Neptune and Saturn. The semimajor axis of the planet was estimated to be 119 AU based on this mass range and setting the disk edge at the chaotic zone boundary.

Remarkably, the discovered planet is located near the disk edge confirming Quillen's hypothesis that a planet is truncating the disk. As only two epochs of observation with the planetary object exist, the constraints on the orbit are not currently tight, though the object is clearly in orbit about the star Fomalhaut. If the object has an eccentricity similar to the ring and aligned with the ring, as proposed by Quillen [4], then the planet's semimajor axis can be estimated. The difference between the estimated and predicted semimajor axis is only 4 AU, with that estimated from the observations at 115 AU. Thus, the prediction [4] can be considered accurate.

The discovered planet does reside near the disk edge, but is brighter than the expected luminosity of a Saturn mass object. From the two optical wavelength points, and their color, it is likely that emission from the planet itself has not been detected [5]. Kalas *et al.* propose that reflected light from a ring system has been detected, thus they only give an upper limit of $3M_J$ for the mass of the planet.

Chiang *et al.* [58] reduced the 4 AU discrepancy in predicted versus estimated semimajor axis with a somewhat larger planet mass (increasing the chaotic zone width) and larger chaotic zone than that adopted by Quillen [4]. Chiang *et al.* considered small dust particles affected by radiation from the star and released by collisions in the ring. They found that a larger planet than estimated by Quillen [4] could truncate a featureless ring but with a larger distance between planet and disk edge. Their estimated planet mass of $\sim 0.5-1\ M_J$ conflicts with diffusional collision simulations that suggest that a planet of this mass would induce nonaxisymmetric structure in the disk [67]. However, it is nontrivial to accurately simulate dust dynamics in proximity to planetary embryos and a planet; thus future work can improve upon these early models and simulations and so improve on the estimated planet mass.

The proposed clearing mechanism requires that the planet scatter dust particles and planetesimals via close approaches. If there is a plentiful supply of planetesimals making close approaches to the planet, then a ring system could still be accreting mass. In hindsight, perhaps one could have predicted bright ring systems for planetary systems undergoing bombardment. The nearby Fomalhaut system is likely to continue to yield surprises as this system continues to be studied.

2.8 Summary

The morphology of circumstellar disks is diverse. The more detailed and deeper the observations, the more the structure that has been revealed. Exoplanetary systems could have the complexity of Saturn's rings and we have only seen their largest, brightest and crudest features. The simple dynamical models and scenarios we have so far explored to explain their morphology have primarily been inspired by solar system models. However the detected exoplanetary disks differ in age, scale, opacity, thickness, and morphology from disks present in our solar system even though some of them may be analogs of a younger version of our solar system. Because they differ, exoplanetary systems are challenging to understand dynamically and motivate a better understanding of the processes likely to have taken place in our solar system when it was younger. For example, the role of collisions among dust particles in the opacity regime appropriate for debris disks has only partly been explored; however, it is possible that all exosolar systems harbor planetesimal debris and exhibit a bright debris disk phase. Most studies have necessarily focused on the simplest models or those harboring a single planet. However, many systems have disks with clearings and some are known to harbor multiple dust rings, suggesting that systems with a single planet could be rare. Future study can explore interactions between multiple planets and disk systems and aim to better understand evolving dusty planetesimal disks in the presence of planetary bodies.

Some of the pioneering studies that predicted planetary bodies embedded in disks have been ruled out [20, 43], other models are now obsolete [21], and yet others are only weakly constraining [55]. Previous studies have explored ''signatures'' of planets in circumstellar disks [15–17, 19], however, future work may instead neglect models that lack planets or planetary embryos, finding that they are ubiquitous and that circumstellar disks cannot be interpreted without them. A remarkably accurate prediction for the Fomalhaut system, building on these pioneering works, motivates further study relating the dynamics of disks to underlying planetary architecture and evolution.

References

1 Danjon, A. (**1946**) Le centenaire de la d'écouverte de Neptune. *Ciel et Terre*, **62**, 369–383.

2 Cuzzi, J.N. and Scargle, J.D. (**1985**) Wavy edges suggest moonlet in Encke's gap. *Astrophysical Journal*, **292**, 276–290.

3 Showalter, M.R. (**1991**) Visual detection of 1981 S 13, Saturn's eighteenth satellite, and its role in the Encke gap. *Nature*, **351**, 709–713, DOI:10.1038/351709a0.

4 Quillen, A.C. (**2006**) Predictions for a planet just inside Fomalhaut's eccentric ring. *Monthly Notices of the Royal Astronomical Society*, **372**, L14–L18.

5 Kalas, P., Graham, J.R., Chiang, E. *et al.* (**2008**), Optical images of an exosolar planet 25 light years from earth. *Science*, DOI:10.1126/science.1166609.

6 Aumann, H.H. (**1985**) IRAS observations of matter around nearby stars. *Publications of the Astronomical Society of the Pacific*, **97**, 885–891.

7 Chen, C.H., Sargent, B.A., Bohac, C. *et al.* (**2006**) Spitzer IRS spectroscopy of IRAS-discovered Debris Disks. *Astrophysical Journal Supplement Series*, **166**, 351.

8 Sykes, M.V. (**1988**) IRAS observations of extended zodiacal structures. *Astrophysical Journal*, **334**, L55–L58.

9 Sykes, M.V. (**1990**) Zodiacal dust bands-their relation to asteroid families. *Icarus*, **85**, 267–289.

10 Dermott, S.F., Nicholson, P.D., Burns, J.A., and Houck, J.R. (**1984**) Origin of the solar system dust bands discovered by IRAS. *Nature*, **312**, 505–509.

11 Nesvorny, D., Bottke, W.F., Vokrouhlicky, D. *et al.* (**2008**) Origin of the near-ecliptic circumsolar dust band. *Astrophysical Journal*, **679**, L143–L146.

12 Burns, J.A., Lamy, O.L., and Soter, S. (**1979**) Radiation forces on small particles in the solar system. *Icarus*, **40**, 1–48.

13 Jackson, A.A. and Zook, H.A. (**1989**) A solar system dust ring with earth as its shepherd. *Nature*, **337**, 629–631.

14 Dermott, S.F., Jayaraman, S., Xu, Y.L. *et al.* (**1994**) A circumsolar ring of asteroidal dust in resonant lock with the Earth. *Nature*, **369**, 719–723.

15 Liou, J.-C. and Zook, H.A. (**1999**) Signatures of the giant planets imprinted on the edgeworth-kuiper belt dust disk. *Astronomical Journal*, **118**, 580–590.

16 Stark, C.C. and Kuchner, M.J. (**2008**) The detectability of exo earths and super earths via resonant signatures in exozodiacal clouds. *The Astrophysical Journal*, **686**, 637–648.

17 Ozernoy, L.M., Gorkavyi, N.N., Mather, J.C., and Taidakova, T.A. (**2000**) Signatures of exosolar planets in dust debris disks. *Astrophysical Journal*, **537**, L147–L151.

18 Greaves, J.S., Holland, W.S., Moriarty-Schieven, G. *et al.* (**1998**) A dust ring around epsilon eridani: analog to the young solar system. *Astrophysical Journal Letters*, **506**, L133–L137.

19 Moro-Martin, A., Wolf, S., and Malhotra, R. (**2005**) Signatures of planets in spatially unresolved debris disks. *Astrophysical Journal*, **621**, 1079–1097.

20 Wilner, D.J., Holman, M.J., Kuchner, M.J., and Ho, P.T.P. (**2002**) Structure in the dusty debris around Vega. *Astrophysical Journal*, **569**, L115–L119.

21 Quillen, A.C. and Thorndike, S. (**2002**) Structure in the epsilon eridani dusty disk caused by mean motion resonances with a 0.3 eccentricity planet at periastron. *Astrophysical Journal*, **578**, L149–L152.

22 Deller, A.T. and Maddison, S.T. (**2005**) Numerical modeling of dusty debris disks. *Astrophysical Journal*, **625**, 398–413.

23 Kuchner, M.J. and Holman, M.J. (**2003**) The geometry of resonant signatures in debris disks with planets. *Astrophysical Journal*, **588**, 1110.

24 Varnière, P., Quillen, A.C., and Frank, A. (**2004**) The evolution of protoplanetary disk edges. *Astrophysical Journal*, **612**, 1152–1162.

25 Strubbe, L.E. and Chiang, E.I. (**2006**) Dust dynamics, surface brightness profiles, and thermal spectra of debris disks: the case of AU microscopii. *Astrophysical Journal*, **648**, 652.

26 Su, K.Y.L., Rieke, G.H., Misselt, K.A. *et al.* (**2005**) The vega debris disk: a surprise from spitzer. *Astrophysical Journal*, **628**, 487.

27 Artymowicz, P. (**1997**) Beta pictoris: an early solar system? *Annual Review of Earth and Planetary Sciences*, **25**, 175–219.

28 Edgar, R. and Artymowicz, P. (**2004**) Pumping of a planetesimal disc by a rapidly migrating planet. *Monthly Notices of the Royal Astronomical Society*, **354**, 769.

29 Quillen, A.C., Morbidelli, A., and Moore, A. (**2007**) Planetary embryos and planetesimals residing in thin debris discs. *Monthly Notices of the Royal Astronomical Society*, **380**, 1642–1648.

30 Quillen, A.C., Varnière, P., Minchev, I., and Frank, A. (**2005**) Driving spiral arms in the circumstellar disks of HD 100546 and HD 141569A. *Astronomical Journal*, **129**, 2481.

31 Wisniewski, J.P., Clampin, M., Grady, C.A. *et al.* (**2008**) The HD 163296 circumstellar disk in scattered light: evidence of time-variable self-shadowing. *Astrophysical Journal*, **682**, 548.

32 Fukagawa, M., Hayashi, M., Tamura, M. *et al.* (**2004**) Spiral structure in the circumstellar disk around AB aurigae. *Astrophysical Journal*, **605**, L53.

33 Kalas, P., Deltorn, J.-M., and Larwood, J. (**2001**) Stellar encounters with the beta pictoris planetesimal system. *Astrophysical Journal*, **553**, 410.

34 Grigorieva, A., Artymowicz, P., and Thebault, Ph. (**2007**) Collisional dust avalanches in debris discs. *Astronomy and Astrophysics*, **461**, 537.

35 Wyatt, M.C. (**2003**) Resonant trapping of planetesimals by planet migration: debris disk clumps and Vega's similarity to the solar system. *Astrophysical Journal*, **598**, 1321.

36 Krivov, A.V., Queck, M., Löhne, T., and Sremcevic, M. (**2007**) On the nature of clumps in debris disks. *Astronomy and Astrophysics*, **462**, 199.

37 Reche, R., Beust, H., Augereau, J.-C., and Absil, O. (**2008**) On the observability of resonant structures in planetesimal disks due to planetary migration. *Astronomy and Astrophysics*, **480**, 551.

38 Kenyon, S.J. and Bromley, B.C. (**2004**) Collisional cascades in planetesimal disks. II. Embedded planets. *Astronomical Journal*, **127**, 513.

39 Telesco, C.M. *et al.*. (**2005**) Mid-infrared images of Pictoris and the possible role of planetesimal collisions in the central disk. *Nature*, **433**, 133.

40 Muzerolle, J., Adame, L., D'Alessio, P. *et al.* (**2006**) 24 μm detections of Circum(sub)stellar disks in IC 348: grain growth and inner holes? *Astrophysical Journal*, **643**, 1003–1010.

41 Calvet, N., D'Alessio, P., Watson, D.M. *et al.* (**2005**) Disks in Transition in the Taurus Population: Spitzer IRS Spectra of GM Aurigae and DM Tauri . *Astrophysical Journal*, **630**, L185–L188.

42 Najita, J.R., Strom, S.E., and Muzerolle, J. (**2007**) Demographics of transition objects. *Monthly Notices of the Royal Astronomical Society*, **378**, 369–378.

43 Quillen, A.C., Blackman, E.G., Frank, A., and Varnière, P. (**2004**) On the planet and the disk of COKU TAURI/4. *Astrophysical Journal*, **612**, L137–L140.

44 Ireland, M.J. and Kraus, A.L. (**2008**) The disk around CoKu tauri/4: circumbinary, not transitional. *Astrophysical Journal*, **678**, L59–L62.

45 Edgar, R.G. and Quillen, A.C. (**2008**) The vertical structure of planet-induced gaps in protoplanetary discs. *Monthly Notices of the Royal Astronomical Society*, **387**, 387.

46 Alexander, R.D., Clarke, C.J., and Pringle, J.E. (**2006**) Photoevaporation of protoplanetary discs - II. Evolutionary models and observable properties. *Monthly Notices of the Royal Astronomical Society*, **369**, 229.

47 Edgar, R.G., Quillen, A.C., and Park, J. (**2007**) The minimum gap-opening planet mass in an irradiated circumstellar accretion disc. *Monthly Notices of the Royal Astronomical Society*, **381**, 1280.

48 Trilling, D.E., Bryden, G., Beichman, C.A. *et al.* Debris disks around Sun-like stars. *Astrophysical Journal*, **674**, 1086.

49 Faber, P. and Quillen, A.C. (**2007**) The total number of giant planets in debris discs with central clearings. *Monthly Notices of the Royal Astronomical Society*, **382**, 1823.

50 Backman, D., Marengo, M., Stapelfeldt, K. *et al.* (**2008**) Epsilon Eridani's planetary debris disk: structure and dynamics based on spitzer and CSO observations. *Astrophysical Journal*, **690**, 1522 arXiv:0810.4564.

51 Stapelfeldt, K.R. *et al.* (**2004**) First look at the fomalhaut debris disk with the spitzer space telescope. *Astrophysical Journal Supplement Series*, **154**, 458.

52 Marois, C., Macintosh, B., Barman, T. *et al.* (**2008**) Direct imaging of multiple planets orbiting the star HR 8799. *Science*, DOI:10.1126/science.1166585.

53 Holland, W.S. *et al.* (**2003**) Submillimeter observations of an asymmetric dust disk around fomalhaut. *Astrophysical Journal*, **582**, 1141.

54 Kalas, P., Graham, J.R., and Clampin, M. (**2005**) A planetary system as the origin of structure in Fomalhaut's dust belt. *Nature*, **435**, 1067.

55 Wyatt, M.C., Dermott, S.F., Telesco, C.M. *et al.* (**1999**) How observations of circumstellar disk asymmetries can reveal hidden planets: pericenter glow and its application to the HR 4796 disk. *The Astrophysical Journal*, **527**, 918–944.

56 Marsh, K.A., Velusamy, T., Dowell, C.D. *et al.* (**2005**) Image of fomalhaut dust ring at 350 microns: the relative column density map shows pericenter-apocenter asymmetry. *Astrophysical Journal*, **620**, L47.

57 Quillen, A.C. and Faber, P. (**2006**) Chaotic zone boundary for low free eccentricity particles near an eccentric planet. *Monthly Notices of the Royal Astronomical Society*, **373**, 1245–1250.

58 Chiang, E., Kite, E., Kalas, P. *et al.* (**2008**) Fomalhaut's debris disk and planet: constraining the mass of fomalhaut b from disk morphology. *Astrophysical Journal*, **693**, 734 arXiv0811.1985.

59 Lisse, C.M., Chen, C.H., Wyatt, M.C., and Morlok, A. (**2008**) Circumstellar dust created by terrestrial planet formation in HD 113766. *Astrophysical Journal*, **673**, 1106–1122.

60 Grady, C. *et al.* (**2001**) The disk and environment of the Herbig Be Star HGD 100646. *Astronomical Journal*, **122**, 3396–3406.

61 Ardila, D.R., Golimowski, D.A., Krist, J.E. *et al.* (**2007**) Hubble space telescope advanced camera for surveys coronagraphic observations of the dust surrounding HD 100546. *Astrophysical Journal*, **665**, 512.

62 Clampin, M. *et al.* (**2003**) HST/ACS coronagraphic imaging of the circumstellar disk around HD1415659A. *Astronomical Journal*, **126**, 385.

63 Quillen, A.C. (**2006**) The warped circumstellar disk of HD 100546. *Astrophysical Journal*, **640**, 1078.

64 Fitzgerald, M.P., Kalas, P.G., Duchene, G. *et al.* (**2007**) The AU microscopii debris disk multiwavelength imaging and modeling. *Astrophysical Journal*, **670**, 536.

65 Liu, M.C. (**2004**) Substructure in the circumstellar disk around the young star AU microscopii. *Science*, **305**, 1442.

66 Metchev, S., Eisner, J.A., Hillenbrand, L.A., and Wolf, S. (**2005**) AO imaging of AU MIC. *Astrophysical Journal*, **622**, 451.

67 Quillen, A.C. (**2007**) Diffusive low optical depth particle discs truncated by planets. *Monthly Notices of the Royal Astronomical Society*, **377**, 1287.

68 Hillenbrand, L.A. (**2008**) Disk-dispersal and planet-formation timescales. *Physica Scripta*, **130**, 014024.

69 Gustafson, B.A.S. (**1994**) Physics of zodiacal dust. *Annual Reviews of Earth and Planet Sciences*, **22**, 553.

70 Sicardy, B., Beauge, C., Ferraz-Mello, S. *et al.* (**1993**) Capture of grains into resonances through Poynting-Robertson drag. *Celestial Mechanics and Dynamical Astronomy*, **57**, 373.

71 Wyatt, M.C. (**2005**) The insignificance of P-R drag in detectable extrasolar planetesimal belts. *Astronomy and Astrophysics*, **433**, 1007.

72 Ward, W.R. (**1997**) Protoplanet Migration by Nebula Tides . *Icarus*, **126**, 261.

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

74 Chambers, J.E., Wetherill, G.W., and Boss, A.P. (**1996**) The stability of multiplanet systems. *Icarus*, **119**, 261.

75 Goldreich, P. and Tremaine, S. (**1980**) Disk-satellite interactions. *Astrophysical Journal*, **241**, 425.

76 Wisdom, J. (**1980**) The resonance overlap criterion and the onset of stochastic behavior in the restricted three-body problem. *Astronomical Journal*, **85**, 1122.

77 Quillen, A.C. (**2006**) Reducing the probability of capture into resonance. *Monthly Notices of the Royal Astronomical Society*, **365**, 1367.

78 Moore, A.J., Quillen, A.C., and Edgar, R.G. (**2008**), Planet migration through a self-gravitating planetesimal disk. *Astrophysics*, arXiv0809.2855

79 Lissauer, J.J. and Espresate, J. (**1998**) Resonant satellite torques on low optical depth particulate disks. *Icarus*, **134**, 155–162.

80 Crida, A., Morbidelli, A., and Masset, F. (**2006**) On the width and shape of gaps in protoplanetary disks. *Icarus*, **181**, 587–604.

81 Lin, D.N.C. and Papaloizou, J.C.B. (**1979**) Tidal torques on accretion discs in binary systems with extreme mass ratios. *Monthly Notices of the Royal Astronomical Society*, **186**, 799–812.

82 Bryden, G., Chen, X., Lin, D.N.C. *et al.* (**1999**) Tidally induced gap formation in protostellar disks: gap clearing and

suppression of protoplanetary growth. *Astrophysical Journal*, **514**, 344–367.

83 Menou, K. and Goodman, J. (**2004**) Low mass protoplanet migration in T tauri α-disks. *Astrophysical Journal*, **606**, 520.

84 Klahr, H. and Lin, D.N.C. (**2005**) Dust distribution in gas disks. II. Self-induced ring formation through a clumping instability. *Astrophysical Journal*, **632**, 1113.

3
Planet–Planet Interactions

Rory Barnes

3.1
Introduction

One of the most striking differences between known exoplanets and the giant planets of our solar system involves the observed orbits. Exoplanets tend to have large eccentricities e, and small semimajor axes a, whereas the gas giants of the solar system have small e and large a. Recently, however, it has been shown that the dynamical interactions in many multiple-planet systems (including the solar system) show certain features in common [1–3]. These shared traits suggest that the character of dynamical interactions (over 10^3–10^4 years), rather than the observed orbits, may be a more meaningful constraint on the origins of planetary systems [4]. Considerations of shared dynamical properties have even resulted in the first successful prediction of the mass and orbit of an extrasolar planet, HD 74156 d [5–8].

Now, with over 300 extrasolar planets known, including >30 multiplanet systems, we can look at this population as a whole with increasing confidence regarding which common characteristics may be more than statistical flukes. Some commonalities among planetary interactions may be key constraints on the origins of planetary systems. Moreover, parameters that quantify interactions have the power to quantitatively constrain planet formation models.

This chapter describes how planets in a system can interact gravitationally with each other in general, how known multiplanet systems interact, and what the distributions of those interactions are. For a more complete review of planetary system dynamics, see [9]. Some basics of planet–planet interactions are reviewed in Section 3.2. The interactions of known exoplanets are presented in Section 3.3. Finally, general conclusions are drawn in Section 3.4.

3.2
Review of Orbital Theory

For convenience, dynamical analyses may be divided into two categories: analytic and N-body. Analytic methods make assumptions that produce analytical equations,

Formation and Evolution of Exoplanets. Edited by Rory Barnes

ISBN: 978-3-527-40896-2

but the equations can often only be solved by numerical methods (i.e., with computers). *N*-body methods use computers to directly calculate the gravitational forces among all bodies in a system in order to determine their motions, that is, how the positions and velocities change with time.

Analytical methods generally ignore short-period changes (usually variations over an orbital period), which are assumed to average to zero over long timescales. These methods model the evolution of orbits, not the actual positions of planets. Since these methods have individual terms whose numerical values can be calculated, the relative importance of each effect can be quantified. The disadvantage of these methods is that they are often accurate only in certain regimes, like low eccentricity. More terms can be added to a semianalytic description to improve accuracy, but eventually the complexity outweighs the advantage of averaging short-period effects. In these situations, it is best to turn to numerical simulations. These *N*-body integrations are grounded in first principles as they solve the fundamental laws of gravity and motion, that is, they are "self-consistent." Modern computing power limits the simulation time and/or the number of bodies that can be considered, but for many systems the changes in the shapes of the orbits are periodic, and the equations that describe the motion need only be integrated for one period ($\lesssim 10^5$ years) to reveal the dynamics. In summary, analytic methods approximate the long-term motion and *N*-body simulations show the true motion. When used appropriately, these approaches complement each other and provide powerful insight into the dynamics and origins of planetary systems.

3.2.1 Analytic Methods

The basis for analytical methods lies in the consideration of the "disturbing function," the difference between the gravitational potential of a planet due to a star and that due to a star and one or more additional planets. Here, we do not review the disturbing function, which is an infinite series Fourier expansion, but focus on two commonly used methods that utilize only a few of these terms: resonant and secular theory. These approaches assume certain terms in the disturbing function average to zero, which can therefore be ignored when modeling orbits. Secular theory ignores all terms that depend on the mean motion, n (the orbital frequency if the planet were on a circular orbit), and, in many cases, all terms that are of order 3 or higher, that is, $e^3 \approx 0$. Resonant theory adds terms that do depend on mean motions, but only those related to the resonance in question. For a detailed description of the disturbing function, secular theory, and resonance theory, refer to [9].

3.2.1.1 Secular Theory

To visualize secular theory, imagine the distribution of the masses of the planets over a long timescale: the distribution would be a ring of matter whose mass equals the planet's mass. (In other words, this assumption implies that perturbations that occur at conjunctions are negligible.) Secular theory predicts how the shapes of these rings change with time, and hence how the orbits change with time. In this

second-order approximation, motion out of the orbital plane is decoupled from motion in the plane.

In nearly all planetary system cases, secular theory predicts the e's and $\Delta\varpi$'s (the difference between two longitudes of periastron, ϖ) oscillate. Secular theory assumes that a remains constant, and therefore conservation of angular momentum requires that as one planet's eccentricity drops, the other's rises (orbital angular momentum is proportional to e). If $\Delta\varpi$ oscillates about 0, the system is experiencing aligned libration. If $\Delta\varpi$ librates about π, then the system is undergoing antialigned libration. If $\Delta\varpi$ oscillates through 2π then the "apsides" (the points of closest and furthest approach to the origin) are circulating. The type of oscillation depends on initial conditions. Exoplanet examples of these types of behavior are shown in Figure 3.1. The motion of $\Delta\varpi$ is analogous to that of a swinging pendulum. When the pendulum swings back and forth, the oscillation is libration. If the pendulum swings all the way around, the oscillation is circulation. Note that there is a clear boundary between these types of motion: the swing that brings the pendulum up to a perfectly vertical position. This boundary between qualitatively different types of oscillation is known as a "separatrix." In apsidal behavior, the analogous separatrix is the boundary between librating and circulating apsides. (Note that general relativity can also affect apsidal behavior in exoplanet systems, see, e.g., [11].)

Recently, it has been noted that many systems lie near an "apsidal separatrix" [2, 13, 15]. The simplest apsidal separatrix is the boundary between libration (either aligned or antialigned) and circulation. For systems of just two planets, the apsidal

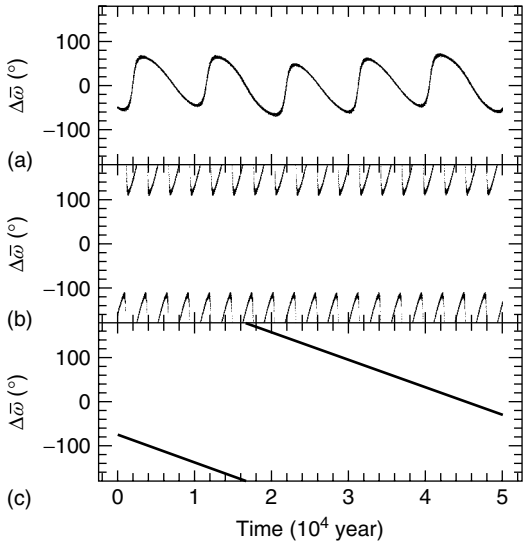


Figure 3.1 Examples of apsidal behavior in known exoplanet systems. (a) The planets HD 37124 c and d undergo aligned libration. (b) The two planets HIP 14810 b and c (from [10]) undergo antialigned libration. (c) The planets in HD 38529 undergo apsidal circulation.

separatrix can only separate circulation and libration. This type of separatrix is a *libration–circulation separatrix*. An example is shown in (a) and (b) of Figure 3.2.

In systems of more than two planets, things get more complicated. In addition to the *libration-circulation separatrix*, the system may interact with different numbers of rotations of $\Delta\varpi$ through 360° during one eccentricity oscillation. The boundary between interactions with different numbers of circulations in one eccentricity cycle is called a *circulation-mode separatrix*, and an example is shown in (c) and (d) of Figure 3.2.

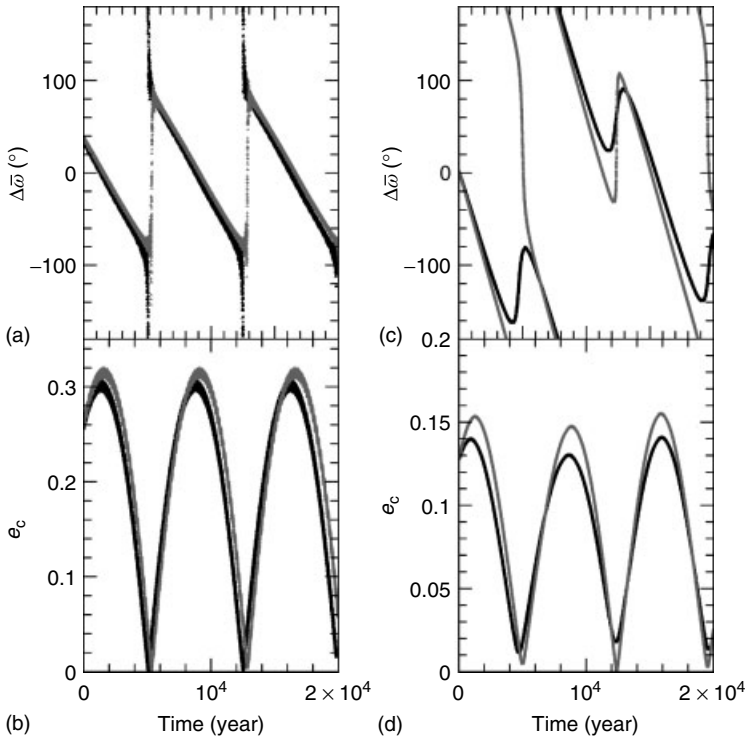


Figure 3.2 Examples of near-separatrix motion in planetary systems. (a, b) A *libration–circulation separatrix*. Two possible evolutions of υ And c and d (the middle and outer planet of the system) assuming different estimates of the current orbits. The black points are the system from [12], the red from [13]. Although the best-fit orbits in these two cases are very similar, they result in qualitatively different types of evolution of $\Delta\varpi$. The older data predict aligned libration, whereas the updated data predict circulation (a). Note that the evolution of e_c is similar in both cases and periodically reach near-zero values (b). (c, d) A *circulation-mode separatrix*. HD 69830 c and d evolve near the *circulation-mode separatrix*. The black data are from [14], and the red data are for a fictitious system in which the inner planet's (b's) eccentricity was changed from 0.1 to 0.15. In the first 10^4 years, the actual $\Delta\varpi$ undergoes one complete rotation through 360^o, but in the fictitious system, $\Delta\varpi$ undergoes two complete circulations (c). We again see that the middle planet's eccentricity periodically drops to near-zero values in both cases (d).

For an interaction to lie near a separatrix, the amplitude of eccentricity oscillations is generally of two orders of magnitude or more. Since $0 \leq e < 1$ for bound planets, this means that at least one planet in near-separatrix interactions (both libration–circulation and circulation-mode) periodically is on a nearly circular orbit. The proximity to the separatrix can be parameterized by ϵ, which is approximately equal to the minimum e divided by the average e over a secular cycle [2]. When ϵ is small ($\lesssim 0.1$), the system is near-separatrix. If $\epsilon = 0$ the pair is on an apsidal separatrix, and one eccentricity periodically reaches zero.

Although secular theory provides a method for identifying components of the dynamics of planetary systems, it must be used with caution on extrasolar planetary systems. Their proximity to the apsidal separatrix can obscure the true motion of the system [15]. The inclusion of additional terms may be more useful in these cases [16–19]; however, *N*-body methods may be the most practical method for determining the secular behavior of exoplanets.

At this point, we must digress and discuss the term *secular resonance*. Currently two definitions exist in the literature for this term, which often leads to confusion. While one definition states that $\Delta\varpi$ is librating, the other is that the ratio of two (or more) precessional frequencies is close to a ratio of two small integers. The latter definition is preferable as it is closer to the true meaning of a resonance, a commensurability of frequencies; so we use this definition. The former should be referred to as *apsidal libration*. For more on libration and secular resonances, consult Barnes and Greenberg [15], who also show that a secular resonance is impossible in a two-planet system.

The above discussion has focused on interactions in which in-plane and out-of-plane motion are decoupled. Such a situation will generally be true when inclinations and eccentricities are small. In other configurations, e and i oscillations can be coupled, which is the phenomenon usually referred to as *Kozai oscillations* or a *Kozai resonance* [20]. For large mutual inclinations ($\gtrsim 40^\circ$), angular momentum can be transferred between planets, resulting in large variations of e and i. This chapter does not discuss this type of interaction in detail; instead, the reader is referred to [21–26].

3.2.1.2 Resonant Interactions

Two bodies may be in a mean motion resonance (MMR) when the ratio of their periods is close to a ratio of small integers. When this occurs, the planets periodically line up at the same points in their orbits, which introduces a repetitive force that cannot be assumed to average to zero over long timescales. Resonant effects can be comparable to secular effects, depending on masses and orbits. Ten exoplanet systems have two planets that are probably in a resonance.

Resonances can stabilize a system by preventing close approaches that might eject a planet. Stable resonances tend to prevent conjunction from occurring near the minimum distance between two orbits. Consider the 3 : 2 case of Neptune and Pluto–although the orbits cross, the resonance prevents conjunctions from occurring at this danger zone.

Resonances are often described in terms of the mean longitudes λ of the planets. The mean longitude measures the position of a planet assuming its angular velocity is constant (only true for a circular orbit). When resonances occur, the mean longitudes and angles of periastron evolve in certain, regular ways. Suppose conjunction occurs at both periastra, the gravitational force will often be strong enough to alter the planetary orbits. Then, at the following conjunction, the apsides will not be perfectly aligned because the change in e will change the apsidal frequency. This nonalignment will introduce a net torque on the orbit that tends to pull the orbits back toward alignment. In this way, resonances maintain themselves, but the alignment will oscillate about an equilibrium position.

From the qualitative description above, we see that a resonance occurs if certain combinations of angles librate about fixed values. If we denote the outer planet with a prime, then the resonant dynamics are important if the ''resonant argument,''

$$\phi = j_1\lambda' + j_2\lambda + j_3\varpi' + j_4\varpi, \tag{3.1}$$

varies slowly relative to the orbital motion. Note that the integers j_k obey the relation

$$j_1 + j_2 + j_3 + j_4 = 0 \tag{3.2}$$

in all terms of the disturbing function. For any pair of planets, integers can be identified that solve Equations (3.1 and 3.2), but the resonance will only be important if its ''order'' is low enough. The order of a resonance is defined as the difference between $|j_1|$ and $|j_2|$. If the order is $\sim$4 or less and the larger number (j_1) is small ($\lesssim$5), then the resonance is at least as important as secular effects in an exoplanet system. High-order resonances are present and important in the solar system, including resonances between three planets [27], but their role is unknown in exoplanets because the observational errors are too large for the effects of these interactions to be unambiguously determined. In exoplanet systems, some resonances show simple behavior, such as one or more resonant arguments always being librating. However, some peculiar examples of resonances have been uncovered. The planets around HD 108874 reported by Butler *et al.* [12], for example, evolve with one resonant argument always librating, but the other arguments alternate between libration and circulation [2]. Note that this system has an ϵ value of 0.2, suggesting that it lies far from the apsidal separatrix. However, the resonance alters the apsidal motion, and we conclude that ϵ is not always a valid description of near-separatrix motion.

If ϕ librates for multiple combinations of j's, then the system is in an ''apsidal corotation resonance'' [28, 29], and $\Delta\varpi$ will also librate. For more on the physics of resonances, consult [9, 30–32], or Chapters 2, 8, and 11.

3.2.2
N-Body Integrations

The most accurate method to determine the evolution of a system is through an N-body calculation. Although more accurate than analytic methods, it does

not provide the researcher with terms that may be interpreted and may require substantial computational resources.

In general, an *N*-body code solves the second-order differential equations of acceleration due to gravity. The accuracy is contingent on two factors: the size of the timestep, Δt, and the order of the integration, which is a measure of the accuracy of the method itself. *N*-body codes update a coordinate, $r(t)$, to $r_0 + v_r \Delta t$, where r_0 is the position at the start of the timestep, and v_r is the velocity at the start of the timestep, which is determined in an analogous manner from the acceleration. This example is of a first-order scheme. Higher order schemes involve calculating positions, velocities and accelerations more frequently through the timestep. They have higher accuracy, but also have more terms and more calculations. As the order increases, the fractional gain in accuracy decreases, which creates an optimal order for algorithms: the order that maximizes speed, but minimizes truncation errors. Most algorithms use second- to fourth-order schemes.

Modern integration methods are ''symplectic,'' which means that truncation errors grow linearly with time, and are therefore the preferred method for *N*-body integrations (errors in nonsymplectic methods, such as the Runge–Kutta methods, grow faster). For exoplanet systems, symplectic codes need only conserve energy to 1 part in 10^4 to produce reliable results [1]. For more on symplectic integrators, refer to [33] or [34]. Several symplectic *N*-body codes are publicly available and are widely used throughout the planetary dynamics community. These codes are well-tested and reliable. The three most prevalent are SWIFT [35][1)], HNBODY[2)], and MERCURY [36][3)].

A code like MERCURY can integrate a few bodies for upward of 1 Gyr in $<$1 month on a 3-GHz processor and therefore permits integrations for the lifetime of a planetary system. Alternatively, these codes may be used to run numerous shorter simulations to explore parameter space of known planetary systems [1, 37] or to model late-stage planet formation [36, 38, 39]. These codes therefore provide tools to understand both long-term behavior and general characteristics of planetary systems.

3.2.3
Dynamical Stability and Chaos

Chaos is a general term that describes a system with nonrepeating motion over a given timescale, that is, the motion appears random. *Stability* describes the ''boundedness'' of a system; a system is stable if changes in its evolution are confined to a certain range. Therefore, one of the most fundamental features of a chaotic system is stability (for a more complete review of chaos theory, consult [40]). For example, the solar system is a chaotic system, but is stable in the sense that the orbits of the planets do not interchange or become unbound, and the oscillations of orbital elements, like eccentricity, occur over a finite range.

1) http://www.boulder.swri.edu/~hal/swift.html
2) http://janus.astro.umd.edu/HNBody
3) http://star.arm.ac.uk/~jec/mercury/mercury6.tar

Alternatively, the solar system is unstable in the sense that the minor planets' orbits can evolve in a nonrepeating manner, as was spectacularly displayed when comet Shoemaker–Levy 9 impacted Jupiter.

So is the solar system stable? It depends on the bodies in question and the timescale. The orbits of many comets are not stable on timescales comparable to the age of the solar system, but the orbits of the planets clearly are (they are still here undergoing periodic evolution). However, on longer timescales, the planets' orbits are not stable; the most unstable planet in the solar system, Mercury, may be lost to the solar system in another 10^{12} years or less [41, 42]. The example of our solar system elucidates an important dichotomy in chaotic systems: a system may be formally unstable, but, for all practical purposes, it is stable. It is irrelevant if Mercury could collide with Venus or the Sun in 10^{12} years because it will be engulfed by the Sun when it enters its red giant phase in 5×10^9 years. However, from a rigorous definition from chaos theory, the planets cannot be said to be stable; the solar system's lifetime is just less than the timescale for instabilities to arise.

For a system to be chaotic, its motion must be (i) governed by nonlinear equations and (ii) sensitive to initial conditions. These requirements are met for systems with two or more planets that are close enough to each other, or near resonance. How close is ''close enough'' is a subject of intense research. In linear, nonchaotic motion, two nearby trajectories diverge at a constant rate, like two balls thrown together; their random motions increase their separation at a constant rate (their relative velocity times the time). In chaotic systems, two nearby trajectories diverge at an exponential rate. Take, for example, two water molecules in a stream that begin right next to each other. Although, in general, the water flows downhill, the paths of the molecules will eventually become divergent due to rocks, vortices, tributaries, and so on. Once one molecule reaches the ocean, the other may be stuck kilometers upstream. A planetary example of chaotic motion is represented in the Kirkwood gaps in the asteroid belt [43–45]. These gaps result from the ejection of asteroids in resonances with Jupiter. Asteroids next to the gaps have evolved regularly (the motion is repeating) for billions of years, but those in the gap were ejected in just millions of years [41].

Most exoplanet systems of two or more planets are chaotic; and we would like to know if they are dynamically stable. Several meanings of stability with regard to planetary systems have arisen that complicate discussions. A system in which no planet is ejected and the semimajor axes remain bounded for all time is known as *Lagrange stability*. This definition is the preferable definition of stability, as it implies that a system will continue to behave the way it does now. Unfortunately, there is no known way to prove Lagrange stability at this time (although numerical simulations may disprove it). A more subtle form of stability exists when the ordering of the planets remains constant. This type is known as *Hill stability* or *hierarchical stability* and it can be proven analytically for nonresonant, two-planet systems [46–50]. In this type of stability the outermost planet may escape, but not the inner one; the ordering of the bodies continues to remain constant.

The term *hierarchical* has two meanings that must be explained here. In stability analyses, a system is hierarchical if it satisfies a simple equation. However, the term

hierarchical is now also employed to describe exoplanet systems for which the ratio of the semimajor axes (a/a') is low (≤ 0.3) [51, 52]. These conflicting definitions naturally lead to the problem that a ''hierarchical planetary system'' may not be ''hierarchically stable'' if the eccentricities are large enough.

Recently, it has been shown that the Hill and Lagrange boundaries may be quite close to each other (see 3.2.3 or [3, 50]). The proximity of a system to the Hill boundary may be parameterized by β. If $\beta = 1$, the system is on the Hill boundary, and if $\beta > 1$, the system is Hill stable. Barnes and Greenberg [50] found that for two systems (47 UMa and HD 12661), the Lagrange stability boundary (estimated via N-body integrations) corresponded to β values of about 1.02 and 1.1, respectively. Although the expression for Hill stability is only valid for systems of two planets out of resonance, many observed systems have only two known companions. Therefore, β can be calculated for the majority of observed multiple-planet systems to determine their proximity to instability [3].

A system's sensitivity to initial conditions is often measured by the Lyapunov time. This time is a measure of the divergence between two initially nearby trajectories. The Lyapunov time does not necessarily predict the onset of irregular (nonrepeating) motion. The Earth has a Lyapunov time of about 5 million years [53–55] . This value does not mean that in 5 million years the Earth's orbit will begin to change wildly; it just means that 5 million years from now the Earth's position cannot be known with arbitrarily high precision. Furthermore, one must not think of the Lyapunov time as a measure of the ''degree'' or ''amount'' of chaos. A system is either chaotic or it is not. For more on chaos in planetary systems, see [41].

The Lyapunov exponent has been exploited in one code in common use in dynamical analyses of exoplanets: MEGNO [56]. This code determines the Lyapunov time in a grid of parameter space, and stability is inferred from this time. Although evolution from a given set of initial conditions was not proven to be stable or unstable, if the Lyapunov time is long enough, the configuration is assumed stable (again, ''long enough'' is not rigorously defined); the Lyapunov time is assumed to be a proxy for stability.

3.3 Distributions of Dynamical Properties

In this section, some emerging trends in multiple-planet system interactions are highlighted. Since only 33 multiple-planet systems are confirmed (as of Aug. 2009), these distributions may not represent the actual population of multiple-planet systems. Nonetheless, the distribution of planetary orbits can validate, as well as inspire, models of planet formation. A naive distribution would be a tabulation of the frequency of individual orbital elements, that is, the number of planets within a certain range of a or e, and so on. This approach leads to the problem that a two-planet system is described by 12 parameters (5 orbital elements and 1 mass per planet). Therefore, research has been focused instead on describing the dynamical interactions. Often, these interactions can be characterized by a single parameter,

which describes the system as a whole. Moreover, multiple-planet systems are dynamical systems, and research should focus on their dynamical properties, not the orbital elements that the planets happen to have today.

3.3.1 Types of Interactions

Three types of dynamical effects can dominate the interactions between adjacent pairs of planets: secular, resonant, and tidal. All three can affect the interaction, but here we assume that only one tends to be dominant. A pair is tidally dominated if the inner planet's a is less than $\sim$0.1 AU [57], resonantly dominated if one or more resonant argument librates, and secularly dominated if the former two are not important. These ''classes'' provide a quick description of the motion [2]. However, as orbital elements are revised, a system's classification may also change.

Currently the secular class appears to be the most common with 21 of 46 pairs in this type of interaction, including the gas giants in our solar system. Tidal pairs account for 15 interactions and resonant interactions dominate the remaining 10 pairs. The observational uncertainties associated with radial velocity surveys are such that resonant interactions are difficult to identify, but identifying tidal pairs is relatively easy. Therefore, we might expect the actual frequency of resonant interactions to be higher and that of tidal interactions to be lower.

3.3.2 Frequency of Mean Motion Resonances

The most dramatic (dynamically speaking) aspect of a planetary system is the presence of MMRs, and considerable work has investigated its presence in exoplanet systems, for example, [24, 29, 58–76]. Of the 46 known pairs of exoplanets, 10 appear to be in an MMR, and 2 (55 Cnc b-c and HD 108874 b-c) have previously had fits that placed them in an MMR. Furthermore, several researchers have suggested that other systems may be in higher order resonances: HD 12661 in 11 : 2 or 6 : 1 [16, 77], and 47 UMa in 5 : 2 [78] or 7 : 3 [79]. From the 10 systems with best fits that predict libration of resonance angles, it appears that the 2 : 1 resonance is most common. If this trend is real, it may be because this resonance is the strongest, and therefore the most efficient at trapping planets [80–82] .

In the case of GJ 876, the resonant interaction is so strong that the evolution of the orbits may be observed in real time [60]. In this case, the orbital parameters of the system change so quickly that an N-body model is necessary to determine the properties of the system. Although this system is the only one known to evolve on such a short timescale, it demonstrates the importance of considering planet–planet interactions among planetary systems. Similarly, it has been speculated that resonances could change the timing of transits over timescales of years [83, 84], but a definitive detection of such a ''transit timing variation'' has yet to be made.

Also of interest is a system's proximity to an MMR. Even if a system is not in resonance, the resonance can still affect the secular motion. The classic example

of this perturbation is the ''Great Inequality''; the orbits of Jupiter and Saturn are close to, but not in, a 5 : 2 MMR [85, 86]. Similar phenomena may be present in HD 12661 and υ And [24]. The HD 38529, HD 168443 and HD 169830 systems are not affected by MMRs, despite the latter's proximity to the 9 : 1 MMR [24].

3.3.3 Apsidal Motion

Considerable attention has also been directed toward identifying the apsidal motions, for example, [2, 4, 15, 23, 32, 87–95], which should only be determined by *N*-body calculations; see Section 3.2.1 and [15]. About one-third of all planetary interactions, including the gas giants of the solar system, interact with $\epsilon < 0.01$, and their apsidal motion is best described as near-separatrix; see Figure 3.3. When this type of interaction was identified [13], it was suggested that the ejection of an original Jupiter mass planet may have produced this near-separatrix behavior. However, subsequent analysis has demonstrated that this mechanism is unlikely

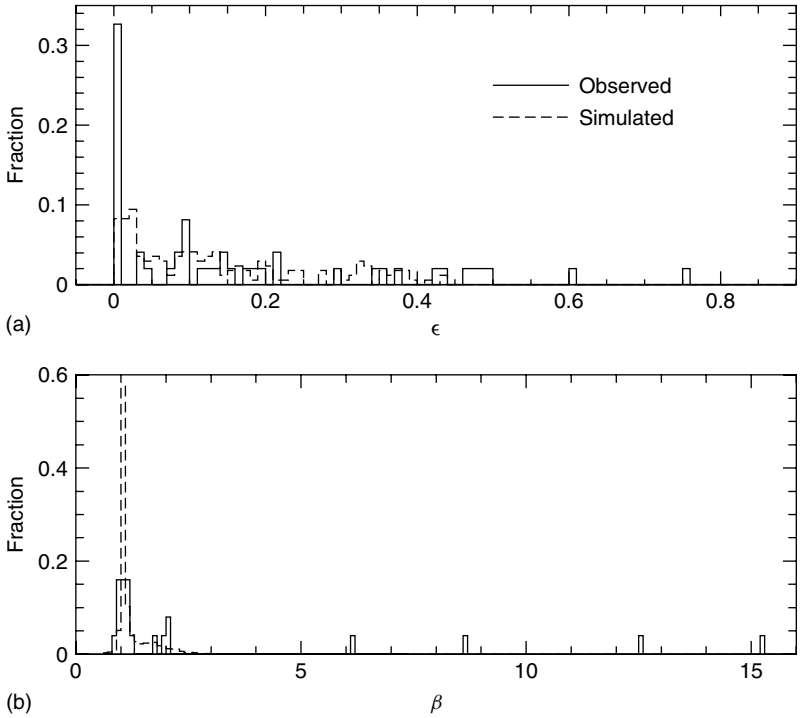


Figure 3.3 (a) Distributions of ϵ. The solid line is the observed distribution, and the dashed line is that predicted by a model involving the ejection of an additional Jupiter mass planet [4, 13]. (b) Distributions of β. The solid line is the observed distribution for known two-planet systems, and the dashed line is that predicted by a model involving the ejection of an additional companion (see [96] for more details). Note that we may only calculate β for two-planet systems.

to result in motion near an apsidal separatrix [4], leaving the origin of this type of interaction unclear.

From the best fits, an overwhelming majority of adjacent pairs appear to undergo apsidal circulation [2], contrary to initial beliefs that libration was the typical state [97]. Of the nine librating systems, seven undergo antialigned libration, and two are aligned. The statistics of these nine systems is too small to draw any reliable conclusions, but at this point it appears that, among librating systems, antialignment is favored.

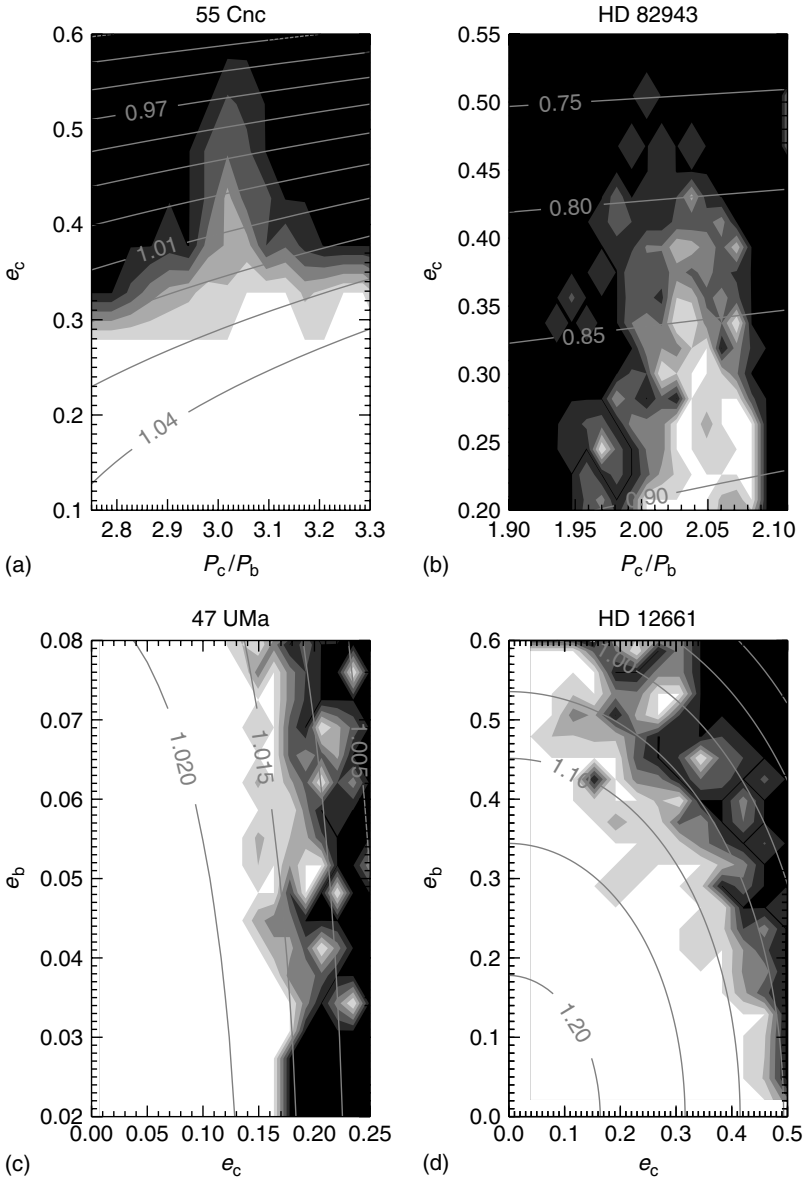
3.3.4
Proximity to Dynamical Instability

The most critical aspect of dynamical systems is their stability, and, therefore, a substantial amount of research has investigated the dynamical stability of planetary systems, for example, [1, 3, 37, 50, 58, 98–115]. Many known systems appear to lie near Lagrange instability (at least one planet is ejected from the system; see Section 3.2.3).

Investigations into stability have found that Lagrange unstable regions exist within the one standard deviation error ellipses for systems with MMRs, such as 55 Cnc and HD 82943, as well as those without, such as HD 12661 and 47 UMa as shown in Figure 3.4. This figure shows the results of numerical simulations within observationally permitted parameter space for these four systems (see [1] for more details). The parameter space was sampled as a Gaussian with a peak at the best-fit values (at the time of the simulations) with a standard deviation equal to the published error. Therefore, the centers of each panel are more highly sampled than the edges. The shading indicates the fraction of initial conditions, in a certain range of orbital element space, that give Lagrange stable behavior (no ejections or exchanges) after $\sim 10^6$ years: white regions contain only stable configurations, black only unstable, and darkening shades of gray correspond to decreasing fractions of simulations, which predict stability. In this figure, P_c/P_b is the ratio of the orbital periods.

The contours represent values of β, the proximity to the Hill stability boundary [3, 50]. For nonresonant systems, the limit of Lagrange stability corresponds to

Figure 3.4 Lagrange stability boundary in relation to the Hill stability boundary for some exoplanetary systems (see text for a discussion of the simulations summarized in these figures). In these plots, white regions represent bins in which all configurations were stable, black contain no stable configurations, darker shades of gray correspond to regions in which the fraction of stable simulations were smaller (see [1] for more details). The curves represent contour lines of β. Contour lines follow the shape of the Lagrange stability boundary, except in resonance, where the Lagrange stability region is larger. (a) Stability of the 55 Cnc system depends on the parameters of the 3 : 1 resonant pair, the eccentricity of the larger planet, and the ratio of the periods. When $P_c/P_b \neq 3$, the Lagrange stability boundary is located at $\beta \approx 1.03$. (b) HD 82943's stability depends on the eccentricity of the larger planet and the ratio of the planets' periods. The Lagrange stable region shown exists wholly in a region that would be considered unstable from Hill stability theory. (c) The stability of 47 UMa depends on the eccentricities of the two planets. The Lagrange stability boundary corresponds to $\beta \approx 1.015$. (d) The stability of the HD 12661 system depends on the eccentricities of the two outer planets. The Lagrange stability boundary lies near the $\beta = 1.1$ contour.


55 Cnc
0.97
1.01
1.04
0.6
0.5
0.4
0.3
0.2
0.1
e_c
2.8 2.9 3.0 3.1 3.2 3.3
P_c/P_b
(a)
HD 82943
0.75
0.80
0.85
0.90
0.55
0.50
0.45
0.40
0.35
0.30
0.25
0.20
e_c
1.90 1.95 2.00 2.05 2.10
P_c/P_b
(b)
47 UMa
1.020
1.015
1.005
0.08
0.07
0.06
0.05
0.04
0.03
0.02
e_b
0.00 0.05 0.10 0.15 0.20 0.25
e_c
(c)
HD 12661
1.00
1.10
1.20
0.6
0.5
0.4
0.3
0.2
0.1
0.0
e_b
0.0 0.1 0.2 0.3 0.4 0.5
e_c
(d)

values of β slightly greater than 1. However, in the presence of a resonance, Lagrange stable orbits can be found at β values as small as 0.75. Lagrange stability boundaries are qualitatively different for resonant and nonresonant pairs. The former have a ''stability peninsula'' located at the resonance, while the latter only have a large contiguous ''stability plateau.''

The apparent correspondence between Lagrange stability and values of β is evidence that the expression for β (Equation [1] in [50]) is a valid representation of the limits of dynamical stability in systems of two planets. In Figure 3.3(b), we plot the current distribution of β values for two-planet systems. Most systems lie near $\beta = 1$, and resonant systems tend to have $\beta < 1$ (the resonance protects the system from instability) [3].

The observed distribution may be explained by the planet–planet scattering model ([116–118] ; Chapter 11). Raymond *et al.* [96] integrated several thousand hypothetical three-planet systems, which are initially unstable, to determine the likelihood that this model could reproduce the observed β distribution. They found that about 25% of the cases ejected two planets, but for the cases that only ejected one planet, 95% had $\beta \leq 2$ and 60% had $\beta \leq 1.1$. The simulated distribution is shown as the dashed line in Figure 3.3(b). Therefore, it appears that scattering is even more likely to produce planets near the stability limit, although we caution that [96] only considers systems with initially three planets, which could be an important simplification. Note that the simulated distribution in Figure 3.3 has had the mass-inclination degeneracy removed.

Figures 3.3 and 3.4 suggest that many planetary systems are dynamically full (no additional companions can survive in between the observed planets) and leads to the packed planetary systems (PPS) hypothesis [1, 5, 6, 119]: all pairs of planets formed close to dynamical instability. The observation that so many systems (using minimum masses) lie near Lagrange instability, despite incompleteness issues (see Chapter 1), has led to investigations to identify regions of Lagrange stability between the few pairs that are more separated [5, 6, 108, 119, 120–130]. This research has shown that the gaps between HD 74156 b-c, HD 38529 b-c, HD 47186 b-c, and 55 Cnc c-d are large enough to support additional planets. The detection of planets in these locations would support the packed nature of planetary systems.

Since its inception in 2004 [5], two planets have been discovered in stable gaps: HD 74156 d [7] and 55 Cnc f [131]. The former now ''fills'' the gap as no additional companions can exist in between planets b and c. The latter, however, does not, and additional planets may still lie in between planets f and d [132]. Although these discoveries provide important corroborations of the PPS model [8], recent work has also shown that Saturn-mass companions do not appear to be present in 20 single-planet systems [133], leaving the veracity of this hypothesis unclear.

Moreover, in order to verify or disprove the PPS hypothesis, a quantitative definition of ''close'' is required. As stated in Section 3.2.3 there exists no quantitative definition of Lagrange stability for any number of planets, and no definition for Hill stability for systems with more than two planets. With limited data available, it appears that when $\beta \lesssim 1.5$ to 2 a system is packed [3]. Future work should reveal how robust this limit is and identify the packing limit for systems of more than two planets.

Table 3.1 Summary of dynamical properties of multiple-planet systems.

System	Pair	MMR	AM	ϵ	β	Class
47 UMa	b-c	–	C[a]	0	1.025	S
55 Cnc	e-b	–	C	3.2×10^{-4}	–	T
	b-c	–	C	8.9×10^{-4}	–	S
	c-f	–	C	0.16	–	S
	f-d	–	C	0.14	–	S
GJ 581	e-b	C[a]	0	0	–	T
	b-c	–	C[a]	0	–	T
	c-d	–	C	0.18	–	T
GJ 876	d-c	–	C[a]	0	–	T
	c-b	2 : 1	A	0.34	–	R
HD 11506	c-b	–	C	0.35	1.17	S
HD 11964	b-c	–	C	0.47	2.04	T
HD 12661	b-c	–	C	0.003	1.2	S
HD 37124	b-c	–	C	0.009	–	S
	c-d	–	A	0.096	–	S
HD 38529	b-c	–	C	0.44	2.07	S
HD 40307	b-c	–	C[a]	?[b]	–	T
	c-d	–	C[a]	?[b]	–	T
HD 45364	b-c	3 : 2	AA	0.76	0.989	R
HD 47186	b-c	–	C	0.49	6.1	S
HD 60532	b-c	3 : 1	AA	0.08	1.05	R
HD 69830	b-c	–	C	0.095	–	T
	c-d	–	C	0.04	–	S
HD 73526	b-c	2 : 1	AA	0.006	0.982	R
HD 74156	b-d	–	C	3.7×10^{-4}	–	S
	d-c	–	C	3.7×10^{-4}	–	S
HD 82943	b-c	2 : 1	C	0.004	0.946	R
HD 128311	b-c	2 : 1	C	0.091	0.968	R
HD 108874	b-c	–	AA	0.6	1.107	S
HD 128311	b-c	2 : 1	C	0.09	0.968	R
HD 155358	b-c	–	AA	0.21	1.043	S
HD 168443	b-c	–	C	0.22	1.939	S
HD 169830	b-c	–	C	0.33	1.280	S
HD 177830	c-b	–	C	0.295	1.05	S
HD 183263	b-c	–	AA	0.12	1.07	S
HD 187123	b-c	–	C	0.43	15.3	T
HD 190360	c-b	–	C	0.38	1.701	T
HD 202206	b-c	5 : 1	C	0.096	0.883	R
HD 208487	b-c	7 : 1	C	0.46	1.2	R
HD 217107	b-c	–	C	0.46	7.191	T
HIP 14810	b-c	–	AA	0.38	–	T
	c-d	–	C	0.14	–	S
μ Ara	c-d	–	C	0.002	–	T
	d-b	2 : 1	C	0.003	–	R

(continued overleaf)

Table 3.1 (*Continued*)

System	Pair	MMR	AM	ϵ	β	Class
	b-e	–	C	0.13	–	S
υ And	b-c	–	C	1.8×10^{-4}	–	T
	c-d	–	C	2.8×10^{-4}	–	S
SS	J-S	–	C	0.19	–	S
	S-U	–	C	0.006	–	S
	U-N	–	C	0.004	–	S

a) The current eccentricity of one planet is 0, placing the pair on an apsidal separatrix.
b) All orbital eccentricities are reported to be zero.

3.4 Conclusions

This chapter has laid out analyzation methods and preliminary trends in exoplanetary dynamics, using best-determined orbits. These results are summarized in Table 3.1[4]. In this table, AM stands for ''apsidal motion'' and the possibilities are circulation (C), aligned libration (A), or antialigned libration (AA). The MMR column lists the resonance, if applicable. The current proximities to the apsidal separatrix, ϵ, and proximities to the Hill stability boundary, β, are also shown. The ''class'' distinguishes orbits whose evolution are dominated by tidal (T), resonant (R), or secular (S) interactions. Table 3.1 includes the dynamical properties of the giant planets in our solar system for comparison.

About half of the planetary systems are probably multiple [10], predictions of additional companions are being borne out [6, 7], and the current distribution of planet masses suggest that there will be many planets with a mass equal to that of Saturn or less [134]. These three observations imply that many multiple planet systems will be detected in the future. Hence, the characterization of planet–planet interactions is a critical aspect of the study of exoplanets.

References

1 Barnes, R. and Quinn, T.R. (**2004**) The (In)stability of planetary systems. *The Astrophysical Journal*, **611**, 494–516.

2 Barnes, R. and Greenberg, R. (**2006**) Behavior of apsidal orientations in planetary systems. *The Astrophysical Journal*, **652**, L53–L56.

3 Barnes, R. and Greenberg, R. (**2007**) Stability limits in resonant planetary systems. *The Astrophysical Journal*, **665**, L67–L71.

4 Barnes, R. and Greenberg, R. (**2007**) Apsidal behavior among planetary orbits: testing the planet-planet scattering

4) see http://xsp.astro.washington.edu for an up-to-date list of these properties

model. *The Astrophysical Journal*, **659**, L53–L56.

5 Barnes, R. and Raymond, S.N. (**2004**) Predicting planets in known extrasolar planetary systems. I. Test particle simulations. *The Astrophysical Journal*, **617**, 569–574.

6 Raymond, S.N. and Barnes, R. (**2005**) Predicting planets in known extrasolar planetary systems. II. Testing for saturn mass planets. *The Astrophysical Journal*, **619**, 549–557.

7 Bean, J.L. *et al.* (**2008**) Detection of a third planet in the HD 74156 system using the Hobby-Eberly telescope. *The Astrophysical Journal*, **672**, 1202–1208.

8 Barnes, R., Goździewski, K., and Raymond, S.N. (**2008**) The successful prediction of the extrasolar planet HD 74156d. *The Astrophysical Journal*, **680**, L57–L60.

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

10 Wright, J.T., Marcy, G.W., Fischer, D.A., *et al.* (**2007**) Four new exoplanets and hints of additional substellar companions to exoplanet host stars. *The Astrophysical Journal*, **657**, 533–545.

11 Adams, F.C. and Laughlin, G. (**2006**) Relativistic effects in extrasolar planetary systems. *International Journal of Modern Physics D*, **15**, 2133–2140.

12 Butler, R.P., Wright, J.T., Marcy, G.W., *et al.* (**2006**) Catalog of nearby exoplanets. *The Astrophysical Journal*, **646**, 505–522.

13 Ford, E.B., Lystad, V., and Rasio, F.A. (**2005**) Planet-planet scattering in the upsilon Andromedae system. *Nature*, **434**, 873–876.

14 Lovis, C., Mayor, M., Pepe, F., *et al.* (**2006**) An extrasolar planetary system with three Neptune-mass planets. *Nature*, **441**, 305–309.

15 Barnes, R. and Greenberg, R. (**2006**) Extrasolar planetary systems near a secular separatrix. *The Astrophysical Journal*, **638**, 478–487.

16 Lee, M.H. and Peale, S. (**2003**) Secular evolution of hierarchical planetary systems. *The Astrophysical Journal*, **592**, 1201–1216.

17 Michtchenko, T.A. and Malhotra, R. (**2004**) Secular dynamics of the three-body problem: application to the υ Andromedae planetary system. *Icarus*, **168**, 237–248.

18 Libert, A.-S. and Henrard, J. (**2005**) Analytical approach to the secular behaviour of exoplanetary systems. *Celestial Mechanics and Dynamical Astronomy*, **93**, 187–200.

19 Veras, D. and Armitage, P. (**2007**) Extrasolar planetary dynamics with a generalized planar Laplace-Lagrange secular theory. *The Astrophysical Journal*, **661**, 1311–1322.

20 Kozai, Y. (**1962**) Secular perturbations of asteroids with high inclination and eccentricity. *Astronomical Journal*, **67**, 591–598.

21 Wu, Y. and Murray, N. (**2003**) Planet migration and binary companions: the case of HD 80606b. *The Astrophysical Journal*, **589**, 605–614.

22 Takeda, G. and Rasio, F.R. (**2005**) High orbital eccentricities of extrasolar planets induced by the Kozai mechanism. *The Astrophysical Journal*, **627**, 1001–1010.

23 Michtchenko, T.A., Ferraz-Mello, S., and Beaugé, S. (**2006**) Modeling the 3-D secular planetary three-body problem. *Icarus*, **181**, 555–571.

24 Libert, A.-S. and Henrard, J. (**2007**) Analytical study of the proximity of exoplanetary systems to mean-motion resonances. *Astronomy and Astrophysics*, **461**, 759–763.

25 Fabrycky, D. and Tremaine, S. (**2007**) Shrinking binary and planetary orbits by Kozai cycles with tidal friction. *The Astrophysical Journal*, **669**, 1298–1315.

26 Libert, A.-S. and Tsiganis, K. (**2009**) Kozai resonance in extrasolar systems. *Astronomy and Astrophysics*, **493**, 677–686.

27 Murray, N. and Holman, M. (**1999**) The origin of Chaos in the outer solar system. *Science*, **283**, 1877–1881.

28 Ferraz-Mello, S., Michtchenko, T.A., and Beaugé, C. (**2005**) The orbits of the extrasolar planets HD 82943c and b. *The Astrophysical Journal*, **621**, 473–481.

29 Michtchenko, T.A., Beaugé, C., and Ferraz-Mello, S. (**2006**) Stationary orbits

in resonant extrasolar planetary systems. *Celestial Mechanics and Dynamical Astronomy*, **94**, 411–432.

30 Peale, S. (**1976**) Orbital resonances in the solar system. *Annual Review of Astronomy and Astrophysics*, **14**, 215–246.

31 Greenberg, R. (**1977**) Orbit-orbit resonances in the solar system - Varieties and similarities. *Vistas in Astronomy*, **21**, 209–239.

32 Beaugé, C., Ferraz-Mello, S., and Michtchenko, T.A. (**2003**) Extrasolar planets in mean-motion resonance: apses alignment and asymmetric stationary solutions. *The Astrophysical Journal*, **593**, 1124–1133.

33 Gladman, B., Duncan, M., and Candy, J. (**1991**) Symplectic integrators for long-term integrations in celestial mechanics. *Celestial Mechanics and Dynamical Astronomy*, **52**, 221–240.

34 Yoshida, H. (**1993**) Recent progress in the theory and application of symplectic integrators. *Celestial Mechanics and Dynamical Astronomy*, **56**, 27–43.

35 Levison, H. and Duncan, M. (**1994**) The long-term dynamical behavior of short-period comets. *Icarus*, **108**, 18–36.

36 Chambers, J.E. (**1999**) A hybrid symplectic integrator that permits close encounters between massive bodies. *Monthly Notices of the Royal Astronomical Society*, **304**, 793–799.

37 Barnes, R. and Quinn, T.R. (**2001**) A statistical examination of the short-term stability of the υ andromedae planetary system. *The Astrophysical Journal*, **550**, 884–889.

38 Raymond, S.N., Quinn, T.R., and Lunine, J.I. (**2004**) Making other earths: dynamical simulations of terrestrial planet formation and water delivery. *Icarus*, **168**, 1–17.

39 Lissauer, J.J. (**2007**) Planets formed in habitable zones of M dwarf stars probably are deficient in volatiles. *The Astrophysical Journal*, **660**, L149–L152.

40 Chirikov, B.V. (**1979**) A universal instability of many-dimensional oscillator systems. *Physics Reports*, **52**, 263–379.

41 Lecar, M. *et al.* (**2001**) Chaos in the solar system. *Annual Review of Astronomy and Astrophysics*, **39**, 581–631.

42 Laskar, J. and Gastineau, M. (**2009**) Existence of collisional trajectories of Mercury, Mars and Venus with the Earth. *Nature*, **459**, 817–819.

43 Kirkwood, D. (**1888**) *The Asteroids, or Minor Planets between Mars and Jupiter*, J.B. Lippencott, Philadelphia.

44 Moons, M. (**1997**) Review of the dynamics in the Kirkwood gaps. *Celestial Mechanics and Dynamical Astronomy*, **65**, 175–204.

45 Tsiganis, K., Varvoglis, H., and Hadjidemetriou, J.D. (**2002**) Stable chaos versus kirkwood gaps in the asteroid belt: a comparative study of mean motion resonances. *Icarus*, **159**, 284–299.

46 Marchal, C. and Bozis, G. (**1982**) Hill stability and distance curves for the general three-body problem. *Celestial Mechanics and Dynamical Astronomy*, **26**, 311–333.

47 Milani, A. and Nobili, A.M. (**1983**) On topological stability in the general three-body problem. *Celestial Mechanics and Dynamical Astronomy*, **31**, 213–240.

48 Gladman, B. (**1993**) Dynamics of systems of two close planets. *Icarus*, **106**, 247–263.

49 Chambers, J.E., Wetherill, G.W., and Boss, A. (**1996**) The stability of multi-planet systems. *Icarus*, **119**, 261–268.

50 Barnes, R. and Greenberg, R. (**2006**) Stability limits in extrasolar planetary systems. *The Astrophysical Journal*, **647**, L153–L156.

51 Lee, M.H. and Peale, S. (**2002**) Dynamics and origin of the 2:1 orbital resonances of the GJ 876 planets. *The Astrophysical Journal*, **567**, 596–609.

52 Goździewski, K. and Konacki, M. (**2004**) Dynamical properties of the multi-planet system around HD 169830. *The Astrophysical Journal*, **610**, 1093–1106.

53 Laskar, J. (**1989**) A numerical experiment on the chaotic behaviour of the solar system. *Nature*, **338**, 237–238.

54 Sussman, G.J. and Wisdom, J. (**1988**) Numerical evidence that the motion of Pluto is chaotic. *Science*, **241**, 433–437.

55 Sussman, G.J. and Wisdom, J. (**1992**) Chaotic evolution of the solar system. *Science*, **257**, 56–62.

56 Cincotta, P. and Simó, C. (**2000**) Simple tools to study global dynamics in non-axisymmetric galactic potentials – I. *Astronomy and Astrophysics Supplement*, **147**, 205–228.

57 Rasio, F.A., Tout, C.A., Lubow, S.H., and Livio, M. (**1996**) Tidal decay of close planetary orbits. *The Astrophysical Journal*, **470**, 1187–1191.

58 Goździewski, K. and Maciejewski, A. (**2001**) Dynamical analysis of the orbital parameters of the HD 82943 planetary system. *The Astrophysical Journal*, **563**, L81–L85.

59 Rivera, E.J. and Lissauer, J.J. (**2001**) Dynamical models of the resonant pair of planets orbiting the star GJ 876. *The Astrophysical Journal*, **558**, 392–402.

60 Laughlin, G. and Chambers, J.E. (**2001**) Short-term dynamical interactions among extrasolar planets. *The Astrophysical Journal*, **551**, L109–L113.

61 Kinoshita, H. and Nakai, H. (**2001**) Stability of the GJ 876 planetary system. *Publications of the Astronomical Society of Japan*, **53**, L25–L26.

62 Laughlin, G. and Chambers, J.E. (**2002**) Extrasolar trojans: the viability and detectability of planets in the 1:1 resonance. *Astronomical Journal*, **124**, 592–600.

63 Ji, J., Liu, L., and Li, G.-Y. (**2002**) The dynamical simulations of the planets orbiting GJ 876. *The Astrophysical Journal*, **572**, 1041–1047.

64 Ji, J., Liu, L., Kinoshita, H., and Nikai, H. (**2002**) The stabilising mechanism of the HD 82943 planetary system. *Chinese Astronomy and Astrophysics*, **26**, 379–385.

65 Hadjidemetriou, J.D. (**2002**) Resonant periodic motion and the stability of extrasolar planetary systems. *Celestial Mechanics and Dynamical Astronomy*, **83**, 141–154.

66 Ji, J., Kinoshita, H., Liu, L., and Li, G. (**2003**) Could the 55 cancri planetary system really be in the 3:1 mean motion resonance? *The Astrophysical Journal*, **585**, L139–L142.

67 Lee, M.H. (**2004**) Diversity and origin of 2:1 orbital resonances in extrasolar planetary systems. *The Astrophysical Journal*, **611**, 517–527.

68 Beaugé, C., Michtchenko, T.A., and Ferraz-Mello, S. (**2005**) Planetary migration and extrasolar planets in the 2/1 mean-motion resonance. *Monthly Notices of the Royal Astronomical Society*, **365**, 1160–1170.

69 Marzari, F., Scholl, H., and Tricarico, P. (**2005**) Frequency map analysis of the 3/1 resonance between planets b and c in the 55 Cancri system. *Astronomy and Astrophysics*, **442**, 359–364.

70 Laughlin, G., Butler, R.P., Fischer, D.A., *et al.* (**2005**) The GJ 876 planetary system: a progress report. *The Astrophysical Journal*, **622**, 1182–1190.

71 Psychoyos, D. and Hadjidemtriou, J.D. (**2005**) Dynamics of 2/1 resonant extrasolar systems application to HD82943 and GLIESE876. *Celestial Mechanics and Dynamical Astronomy*, **92**, 135–156.

72 Lee, M.H., Butler, R.P., Fischer, D.A., *et al.* (**2006**) On the 2:1 orbital resonance in the HD 82943 planetary system. *The Astrophysical Journal*, **641**, 1178–1187.

73 Marzari, F., Scholl, H., and Tricarico, P. (**2006**) A numerical study of the 2:1 planetary resonance. *Astronomy and Astrophysics*, **453**, 341–348.

74 Voyatzis, G. and Hadjidemetriou, J.D. (**2006**) Symmetric and asymmetric 3:1 resonant periodic orbits with an application to the 55Cnc extra-solar system. *Celestial Mechanics and Dynamical Astronomy*, **95**, 259–271.

75 Hadjidemetriou, J.D. (**2006**) Symmetric and asymmetric librations in extrasolar planetary systems: a global view. *Celestial Mechanics and Dynamical Astronomy*, **95**, 225–244.

76 Ji, J.-H. and Liu, L. (**2006**) Stability and 2:1 resonance in the planetary system HD 829431. *Chinese Astronomy and Astrophysics*, **30**, 75–86.

77 Goździewski, K. (**2003**) Stability of the HD 12661 planetary system. *Astronomy and Astrophysics*, **398**, 1151–1161.

78 Psychoyos, D. and Hadjidemetriou, J.D. (**2005**) Dynamics of populations of planetary systems, *Proceedings of IAU Colloquium, 197* (eds Z. Knezevic and A. Milani), Cambridge University Press, Cambridge, pp. 55–62.

79 Laughlin, G., Chambers, J.E., and Fischer, D.A. (**2002**) A dynamical analysis of the 47 ursae majoris planetary system. *The Astrophysical Journal*, **579**, 455–467.

80 Kley, W., Peitz, J., and Bryden, G. (**2004**) Evolution of planetary systems in resonance. *Astronomy and Astrophysics*, **414**, 735–747.

81 Kley, W., Lee, M.H., Murray, N., and Peale, S.J. (**2005**) Modeling the resonant planetary system GJ 876. *Astronomy and Astrophysics*, **437**, 727–742.

82 Sándor, Zs. and Kley, W. (**2006**) On the evolution of the resonant planetary system HD 128311. *Astronomy and Astrophysics*, **451**, L31–L34.

83 Agol, E., Steffen, J., Sari, R., and Clarkson, W. (**2005**) On detecting terrestrial planets with timing of giant planet transits. *Monthly Notices of the Royal Astronomical Society*, **359**, 567–579.

84 Holman, M.J. and Murray, N.W. (**2005**) The use of transit timing to detect terrestrial-mass extrasolar planets. *Science*, **307**, 1288–1291.

85 Varadi, F., Ghil, M., and Kaula, W.M. (**1999**) Jupiter, Saturn, and the edge of chaos. *Icarus*, **139**, 286–294.

86 Michtchenko, T.A. and Ferraz-Mello, S. (**2001**) Modeling the 5 : 2 mean-motion resonance in the Jupiter-Saturn planetary system. *Icarus*, **149**, 357–374.

87 Chiang, E.I., Tabachnik, S., and Tremaine, S. (**2001**) Apsidal alignment in υ andromedae. *Astronomical Journal*, **122**, 1607–1615.

88 Chiang, E.I. and Murray, N. (**2002**) Eccentricity excitation and apsidal resonance capture in the planetary system υ andromedae. *The Astrophysical Journal*, **576**, 473–477.

89 Malhotra, R. (**2002**) A dynamical mechanism for establishing apsidal resonance. *The Astrophysical Journal*, **575**, L33–L36.

90 Ji, J., Liu, L., Zhou, J.L., and Kinoshita, K. (**2003**) The apsidal motion in multiple planetary systems. *Chinese Astronomy and Astrophysics*, **27**, 127–132.

91 Ji, J., Liu, L., and Kinoshita, H. (**2003**) The librating companions in HD 37124, HD 12661, HD 82943, 47 ursa majoris, and GJ 876: alignment or antialignment? *The Astrophysical Journal*, **591**, L57–L60.

92 Zhou, L.Y., Lehto, H.J., Sun, Y.S., and Zheng, J.Q. (**2004**) Apsidal corotation in mean motion resonance: the 55 Cancri system as an example. *Monthly Notices of the Royal Astronomical Society*, **350**, 1495–1502.

93 Libert, A.-S., and Henrard, J. (**2006**) Secular apsidal configuration of non-resonant exoplanetary systems. *Icarus*, **183**, 186–192.

94 Rodríguez, A., and Gallardo, T. (**2005**) The dynamics of the HD 12661 extrasolar planetary system. *The Astrophysical Journal*, **628**, 1006–1013.

95 Ji, J., Kinoshita, H., Liu, L., and Li, G. (**2007**) The secular evolution and dynamical architecture of the neptunian triplet planetary system HD 69830. *The Astrophysical Journal*, **657**, 1092–1097.

96 Raymond, S.N., Barnes, R., Veras, D., *et al.* (**2009**) Planet-planet scattering leads to tightly packed planetary systems. *The Astrophysical Journal*, **696**, L98–L101.

97 Zhou, J.-L. and Sun, Y.-S. (**2003**) Occurrence and stability of apsidal resonance in multiple planetary systems. *The Astrophysical Journal*, **598**, 1290–1300.

98 Laughlin, G. and Adams, F.C. (**1999**) Stability and chaos in the υ andromedae planetary system. *The Astrophysical Journal*, **526**, 881–889.

99 Laskar, J. (**2000**) On the spacing of planetary systems. *Physics Review Letters*, **84**, 3240–3243.

100 Stepinski, T.F., Malhotra, R., and Black, D.C. (**2000**) The υ andromedae system: models and stability. *The Astrophysical Journal*, **545**, 1044–1057.

101 Rivera, E.J. and Lissauer, J.J. (**2000**) Stability analysis of the planetary system orbiting υ andromedae. *The Astrophysical Journal*, **530**, 454–463.

102 Lissauer, J.J. and Rivera, E.J. (**2001**) Stability analysis of the planetary system orbiting υ andromedae. II. Simulations using new lick observatory fits. *The Astrophysical Journal*, **554**, 1141–1150.

103 Gozdziewski, K., Bois, E., Maciejewski, A.J., and Kiseleva-Eggleton, L. (**2001**) Global dynamics of planetary systems

with the MEGNO criterion. *Astronomy and Astrophysics*, **378**, 569–586.

104 Ito, T. and Miyama, S.M. (**2001**) An estimation of upper limit masses of ŏ3c5 andromedae planets. *The Astrophysical Journal*, **552**, 372–379.

105 Kiseleva-Eggleton, L., Bois, E., Rambaux, N., and Dvorak, R. (**2002**) Global dynamics and stability limits for planetary systems around HD 12661, HD 38529, HD 37124, and HD 160691. *The Astrophysical Journal*, **578**, L145–L148.

106 Goździewski, K. (**2002**) Stability of the 47 UMa planetary system. *Astronomy and Astrophysics*, **393**, 997–1013.

107 Bois, E., Kiseleva-Eggleton, L., Rambaux, N., and Pilat-Lohinger, E. (**2003**) Conditions of dynamical stability for the HD 160691 planetary system. *The Astrophysical Journal*, **598**, 1312–1320.

108 Dvorak, R., Pilat-Lohinger, E., Funk, B., and Freistetter, F. (**2003**) A study of the stable regions in the planetary system HD 74156 - Can it host earthlike planets in habitable zones? *Astronomy and Astrophysics*, **410**, L13–L16.

109 Goździewski, K. and Maciejewski, A. (**2003**) The janus head of the HD 12661 planetary system. *The Astrophysical Journal*, **586**, L153–L156.

110 Érdi, B., Dvorak, R., Sándor, Zs., *et al.* (**2004**) The dynamical structure of the habitable zone in the HD 38529, HD 168443 and HD 169830 systems. *Monthly Notices of the Royal Astronomical Society*, **351**, 1043–1048.

111 Dvorak, R., Pilat-Lohinger, E., Schwarz, R., and Freistetter, F. (**2004**) Extrasolar Trojan planets close to habitable zones. *Astronomy and Astrophysics*, **426**, L37–L40.

112 Veras, D. and Armitage, P. (**2004**) The dynamics of two massive planets on inclined orbits. *Icarus*, **172**, 349–371.

113 Érdi, B. and Sándor, Zs. (**2005**) Stability of co-orbital motion in exoplanetary systems. *Celestial Mechanics and Dynamical Astronomy*, **92**, 113–121.

114 Goździewski, K., Konacki, M., and Maciejewski, A. (**2006**) Orbital configurations and dynamical stability of multiplanet systems around sun-like stars HD 202206, 14 herculis, HD 37124, and HD 108874. *The Astrophysical Journal*, **645**, 688–703.

115 Goździewski, K., Maciejewski, A., and Migaszewski, C. (**2007**) On the extrasolar multiplanet system around HD 160691. *The Astrophysical Journal*, **657**, 546–558.

116 Weidenschilling, S. and Marzari, F. (**1996**) Gravitational scattering as a possible origin for giant planets at small stellar distances. *Nature*, **384**, 619–621.

117 Rasio, F.A. and Ford, E.B. (**1996**) Dynamical instabilities and the formation of extrasolar planetary systems. *Science*, **274**, 954–956.

118 Marzari, F. and Weidenschilling, S. (**2002**) Eccentric extrasolar planets: the jumping Jupiter model. *Icarus*, **156**, 570–579.

119 Raymond, S.N., Barnes, R., and Kaib, N.A. (**2006**) Predicting planets in known extrasolar planetary systems. III. Forming terrestrial planets. *The Astrophysical Journal*, **644**, 1223–1231.

120 Jones, B.W., Sleep, P.N., and Chambers, J.E. (**2001**) The stability of the orbits of terrestrial planets in the habitable zones of known exoplanetary systems. *Astronomy and Astrophysics*, **366**, 254–262.

121 Jones, B.W. and Sleep, P.N. (**2002**) The stability of the orbits of Earth-mass planets in the habitable zone of 47 Ursae Majoris. *Astronomy and Astrophysics*, **393**, 1015–1026.

122 Noble, M., Musielak, Z.E., and Cuntz, M. (**2002**) Orbital stability of terrestrial planets inside the habitable zones of extrasolar planetary systems. *The Astrophysical Journal*, **572**, 1024–1030.

123 Cuntz, M., von Bloh, W., Bounama, C., and Franck, S. (**2003**) On the possibility of earth-type habitable planets around 47 UMa. *Icarus*, **162**, 214–221.

124 Menou, K. and Tabachnik, S. (**2003**) Dynamical habitability of known extrasolar planetary systems. *The Astrophysical Journal*, **583**, 473–488.

125 von Bloh, W., Cuntz, M., Franck, S., and Bounama, C. (**2003**) On the possibility of earth-type habitable planets in the 55 Cancri system. *Astrobiology*, **3**, 681–688.

126 Asghari, N., Broeg, C., Carone, L., *et al.* (**2004**) Stability of terrestrial planets in the habitable zone of Gl 777 A, HD 72659, Gl 614, 47 Uma and HD 4208. *Astronomy and Astrophysics*, **426**, 353–365.

127 Funk, B., Pilat-Lohinger, E., Dvorak, R., *et al.* (**2004**) Resonances in multiple planetary systems. *Celestial Mechanics and Dynamical Astronomy*, **90**, 43–50.

128 Ji, J., Liu, L., Kinoshita, H., and Li, G. (**2005**) Could the 47 Ursae majoris planetary system be a second solar system? Predicting the earth-like planets. *The Astrophysical Journal*, **631**, 1191–1197.

129 Rivera, E.J. and Haghighipour, N. (**2007**) On the stability of test particles in extrasolar multiple planet systems. *Monthly Notices of the Royal Astronomical Society*, **374**, 599–613.

130 Kopparapu, R., Raymond, S.N., and Barnes, R. (**2009**) Stability of additional planets in and around the habitable zone of the HD 47186 planetary system. *The Astrophysical Journal*, **695**, L181–L184.

131 Fischer, D.A., Marcy, G.W., Butler, R.P., *et al.* (**2008**) Five planets orbiting 55 Cancri. *The Astrophysical Journal*, **675**, 790–801.

132 Raymond, S.N., Barnes, R., and Gorelick, N. (**2008**) A dynamical perspective on additional planets in 55 Cancri. *The Astrophysical Journal*, **689**, L478–L481.

133 Wittenmyer, R.A., Endl, M., Cochran, W.D. *et al.* (**2009**) A search for multi-planet systems using the hobby-eberly telescope. *The Astrophysical Journal*, **182**, 97–119.

134 Marcy, G.W. *et al.* (**2005**) Five new extrasolar planets. *The Astrophysical Journal*, **619**, 570–584.

Bodenheimer, P., Laughlin, G., and Lin, D.N.C. (**2003**) On the radii of extrasolar giant planets. *The Astrophysical Journal*, **592**, 555–563.

Butler, R.P. *et al.* (**1999**) Evidence for multiple companions to υ andromedae. *The Astrophysical Journal*, **526**, 916–927.

Butler, R.P. *et al.* (**2002**) On the double-planet system around HD 83443. *The Astrophysical Journal*, **578**, 565–572.

Cochran, W.D. *et al.* (**2007**) A planetary system around HD 155358: the lowest metallicity planet host star. *The Astrophysical Journal*, **665**, 1407–1412.

Correia, A.C.M. *et al.* (**2005**) The CORALIE survey for southern extra-solar planets. XIII. A pair of planets around HD 202206 or a circumbinary planet? *Astronomy and Astrophysics*, **440**, 751–758.

Goździewski, K. and Konacki, M. (**2006**) Trojan pairs in the HD 128311 and HD 82943 planetary systems? *The Astrophysical Journal*, **647**, 573–586.

Papaloizou, J.C.B. and Terquem, C. (**2006**) Planet formation and migration. *Reports on Progress Physics*, **69**, 119–180.

Pepe, F. *et al.* (**2007**) The HARPS search for southern extra-solar planets. VIII. μ Arae, a system with four planets. *Astronomy and Astrophysics*, **462**, 769–776.

Tinney, C.G. *et al.* (**2006**) The 2:1 resonant exoplanetary system orbiting HD 73526. *The Astrophysical Journal*, **647**, 594–599.

Veras, D. and Armitage, P. (**2006**) Extrasolar planetary dynamics with a generalized planar laplace-lagrange secular theory. *The Astrophysical Journal*, **661**, 1311–1322.

Veras, D. and Ford, E.B. (**2009**) Secular evolution of HD 12661: a system caught at an unlikely time. *The Astrophysical Journal*, **690**, L1–L4.

Further Reading

Beaugé, C. *et al.* (**2007**) Co-orbital terrestrial planets in exoplanetary systems: a formation scenario. *Astronomy and Astrophysics*, **463**, 359–367.

4
Formation via Disk Instability

Lucio Mayer

4.1
Introduction—Basic Notions of Disk Instability

In the disk instability model, giant planets form from direct collapse of the gas in the protoplanetary disk. The disk fragments into gravitationally bound clumps with sizes comparable to giant planets as a result of gravitational instability [1]. The idea that planets in the solar system could condense directly out of the nebula is one of the oldest in astrophysics since it dates back to Laplace in the eighteenth century. The concept was revived by Kuiper [2] and Cameron [3] in the context of the solar system. Since the discovery of extrasolar planets, the interest in disk instability has grown because fragmentation, when feasible, occurs quickly, naturally producing multiple giant planets [1, 4, 5] on a timescale comparable to the orbital time in the protoplanetary disk, that is, $<$100 years. A short formation timescale is attractive because the presence of protoplanetary disks becomes increasingly rare around stars older than a few million years [6], suggesting fairly short disk dissipation timescales.

Gravitational instability requires high mass densities and low temperatures as it demands that the self-gravity of the gas, which drives the collapse, prevails over the stabilizing effect of thermal pressure. This translates into the requirement that, at some stage, the protoplanetary disk had to be very cold and massive. Furthermore, a protoplanetary disk revolving around its host star is in differential rotation. This causes shear, which also opposes the tendency toward collapse promoted by gravity. The competing effects of self-gravity, thermal pressure, and shear in the disk are nicely encapsulated in the Toomre Q parameter [7]:

$$Q = \frac{\kappa c_s^2}{\pi G \Sigma} \tag{4.1}$$

In Eq. (4.1), the gas surface density Σ measures the importance of self-gravity, the sound speed c_s is related to thermal pressure, and κ, the epicyclic frequency, determines the strength of the shear. The Toomre parameter can be derived via a *local* perturbative stability analysis of infinitesimally thin differentially rotating disks [8]. If $Q < 1$ then the gas is unstable to gravitational collapse, while if $Q > 1$

Formation and Evolution of Exoplanets. Edited by Rory Barnes

ISBN: 978-3-527-40896-2

stability is achieved. If the disk is essentially Keplerian, as expected if its mass is substantially lower than that of the star, the epicyclic frequency is $\kappa \sim \Omega$, where Ω is the Keplerian angular velocity. The Toomre parameter indicates that, given equivalent temperatures and surface densities, a disk rotating around a lower mass star will be more prone to gravitational instability compared to one revolving around a more massive star (because $\kappa^2 \sim \Omega^2 \sim M_*$). Moreover, closer to the star, disk instability is reduced because the shear induced by differential rotation is stronger (κ is an increasing function of the radial distance r in a nearly Keplerian disk, and so is $d\kappa/dr$). Provided that the mass of the star is fixed, whether Q can drop below unity anywhere in a protoplanetary disk depends on the temperature distribution and the surface density profile of the disk.

The standard assumption in disk instability is that the surface mass density in the disk follows a power law of the type $r^{-\alpha}$, with $\alpha \sim 1$–1.5 and that the total disk mass within 20–30 AU is in the range 0.05–$0.2 M_\odot$ [1, 4, 5, 9–11] corresponding to the high end of T Tauri disk masses [12, 13]. The initial temperature profile is also assumed to be a power law with values going from a few hundred kelvin at ~1 AU to a few tens of kelvin at greater than 10 AU. Such temperature profiles are suggested by radiative transfer calculations of protoplanetary disks heated by the star and by accretion of mass from the molecular envelope [4, 14–16]. Unfortunately, little information can be drawn from observations of protoplanetary disks that are sensitive to the temperature of the surface of the disk, rather than its midplane. Moreover, gas and dust observables trace the conditions of the disk within the inner 1 AU or outside 20 AU, but not in the 1–20 AU region, which is the most relevant for giant planet formation [17]. Constraints on the outer disk temperature come mostly from the cosmochemistry of the protosolar nebula, in particular, from the composition of comets in the solar system [18], since little is known about the gas component at several AUs in other protoplanetary disks (see review by Durisen *et al.* [19]. Temperatures of order 50–100 K at ~10 AU from the star would be consistent with radiative transfer models of the dust emission in protoplanetary disks [20]. With the profiles just outlined, a disk should have $Q \sim 1$–1.5 at distances between 5 and 50 AU [4, 9–11, 19, 21, 22, 23]. At very large radii, the lower surface density tends to stabilize a disk with power-law density profile, but Q may still approach 1 if the temperature decreases to 10–20 K. In the early stages of protostellar collapse, the disk might be even more massive than $0.2 M_\odot$, and have enough surface density to support disk instability even at radii of 100 AU or larger [24, 25].

The requirement that $Q \geq 1$ for stability implies the existence of a large wavelength that can become unstable, the so-called Toomre or critical wavelength, $\lambda_{\rm crit} = 4\pi^2 G\Sigma/\kappa^2$. Perturbations with larger scales are stabilized by shear. Only perturbations with scales smaller than the Toomre wavelength and larger than a minimum length equal to the Jeans length, $\lambda_j = c_s^2/\pi\Sigma$ can grow (assuming an infinitesimally thin slab). Scales smaller than the Jeans length are stabilized by pressure. Finally, perturbation analysis also yields the most unstable wavelength, $\lambda_{\rm mu} \sim 0.5\lambda_{\rm crit}$, namely, the wavelength for which the growth of the perturbation is fastest [8]. A gas parcel with a size comparable to $\lambda_{\rm mu}$ would then collapse

first. Moreover, the mass associated with a spherical region with diameter of order λ_{mu} defines the typical mass scale of the collapsing clump, while that associated with λ_j and λ_{crit} define, respectively, the minimum and maximum mass of clumps [26, 27]. The maximum mass of collapse, for example, will be given by $M_{max} = \rho(\lambda_{crit}/2)^3$.

For the typical temperatures and densities of disks reported above, the maximum mass varies from slightly less than a Jupiter mass to a few times the mass for radii between a few and a few tens of astronomical units [26]. Since both shear and gas temperature are expected to increase toward the star, a disk with mass $\sim 0.1–0.2 M_\odot$ cannot have $Q < 1$ at small radii, $R < 1–3$ AU. The limiting radius for fragmentation translates into a minimum clump mass of roughly a Saturn mass [9, 26]. More massive disks, or disks around lower mass stars, will tend to produce larger clumps (λ_{crit} is larger). While these scalings are informative, as soon as the system ventures in the nonlinear regime the perturbative approach, and thus any considerations drawn from it, are not expected to remain valid. This means that the actual masses and sizes of the objects formed by direct collapse need to be assessed via a fully self-consistent, nonlinear three-dimensional (3D) calculation. We discuss this point later in Section 4.2, where we also discuss the masses of extrasolar planets.

Another important limitation to be kept in mind regarding analytic theory is that the Toomre parameter is derived as a *local* stability criterion and specifically addresses the response of the gas to axisymmetric perturbations only. For the study of global stability of a disk with finite thickness undergoing an arbitrary perturbation that, in general, will be nonaxisymmetric [28], one has to resort to a numerical simulation. Since the pioneering works in the early 1990s [29], nearly two decades of numerical simulations have established that the stability threshold given by the Toomre parameter is approximately correct even in a global sense, and for nonaxisymmetric perturbations [19]; in particular, vigorous nonaxisymmetric perturbations, possibly resulting in fragmentation, require that $Q < 1.4$ [9, 19, 30, 31], as shown in Figure 4.1. Spiral modes are the result of density waves sheared by rotation as they propagate through a differentially rotating disk. Several spiral modes can be simultaneously present, and can exchange power in a complex nonlinear interaction [32]. What happens after a vigorous spiral pattern has developed depends on the dissipation mechanisms in the disk. Gas is compressed and eventually shocked as it encounters a spiral wave, increasing its temperature (Figure 4.2). This tends to stabilize the gas as the Toomre parameter is increased. Hence the instability itself acts as a source of heating for the disk, perhaps resulting in self-regulation. Other sources of heating, such as kinematical or magnetic viscosity, accretion from the molecular cloud envelope, and/or irradiation from the parent star or neighboring stars, may add further to the heating.

If the gas cools efficiently enough to dissipate the thermal energy generated by shocks and other heating sources, the spiral arms become increasingly more prominent until fragmentation eventually begins along them [1, 5]. Both 2D and 3D simulations concur that the cooling timescale t_{cool} must be of the order of the local orbital time T_{orb}, for dissipation of energy to be effective and to allow fragmentation

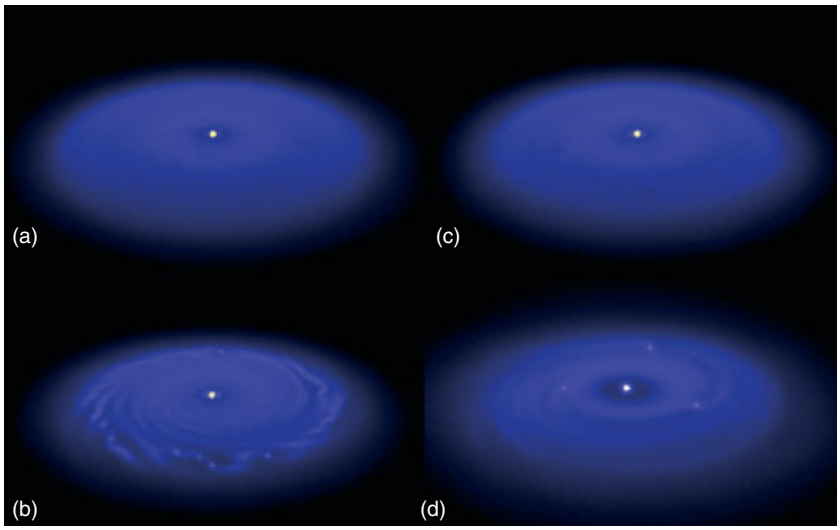


Figure 4.1 A disk instability calculation that starts with a locally isothermal equation of state and switches to an adiabatic equation of state after fragmentation [5, 9]. The disk extends from 4 to 25 AU around a solar mass star and grows in mass over time to mimic accretion from a molecular cloud envelope. From (a)–(d), an evolutionary sequence covering about 1000 years is shown. In (a), the disk has a minimum Toomre parameter of $Q \sim 1.8$, thus displaying only mild spiral structure (the disk mass is $\sim 0.06 M_{\odot}$). At the time represented in (b) the minimum Toomre parameter has dropped to $Q \sim 1.4$ as the disk mass has grown to $\sim 0.085 M_{\odot}$, so that the system approaches fragmentation. About 300 years later, the disk has already fragmented (d). A small number of clumps survives after merging with other clumps.

[10, 11, 30, 33, 34]. In particular, varying thermodynamical parameters such as the adiabatic index of the gas indicates that fragmentation occurs for $t_{\rm cool} = A T_{\rm orb}$ with $1 < A < 2$ [35]; see Figure 4.2. If the cooling time condition is satisfied, fragmentation happens as soon as Q drops below unity.

If fragmentation occurs and produces clumps, will these clumps become self-gravitating and survive long enough to be considered plausible precursors of giant planets? The fate of clumps is discussed in Section 4.3 onward. The current section provides an overview on simulations of disk instability, while Section 4.2 briefly discusses disk thermodynamics.

4.2 Simulations of Disk Instability

Since gravitational instability is a highly nonlinear process, it needs to be studied with numerical simulations that solve for the coupled hydrodynamical, gravitational, and radiative partial differential equations in three dimensions. Instead of directly solving the radiative transfer equation (including cooling and heating

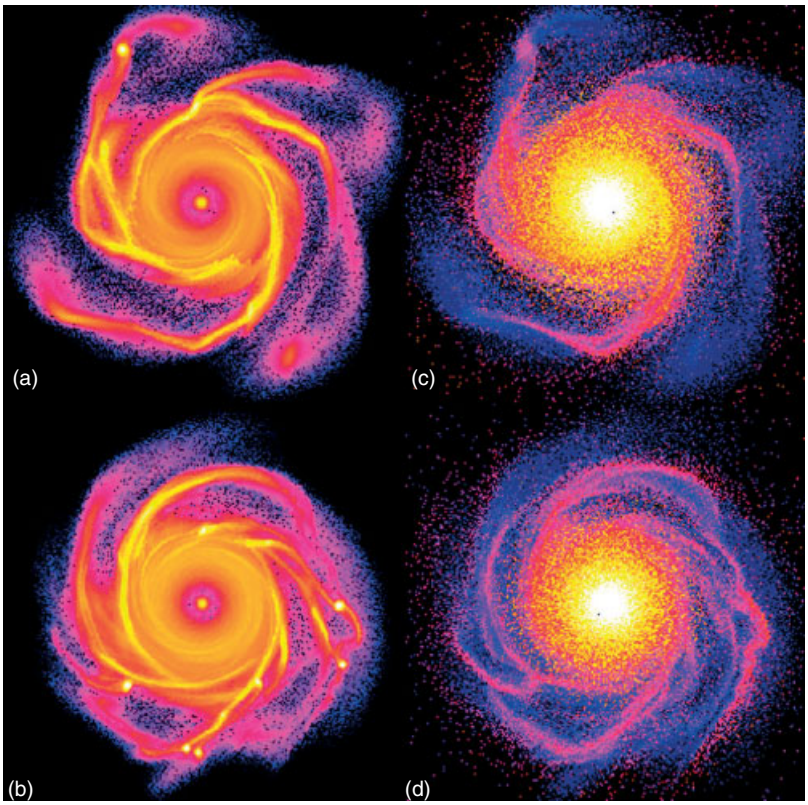


Figure 4.2 Color-coded logarithmic density (a, b) and temperature (c, d) maps for two simulations evolving identical disks but with different cooling times, $t_{\rm cool} = 1.5T_{\rm orb}$ (a, c) and $t_{\rm cool} = 1T_{\rm orb}$ (b, d). The brighter the colors the higher the density (temperature). Clearly, fragmentation is reduced with increasing cooling time [33]. The two simulations adopt an adiabatic index $\gamma = 7/5$ and a molecular weight $\mu = 2.4$ in the ideal gas equation of state.

terms), a simpler approach, when possible, would be to describe gas thermodynamics using an equation of state. All methods to solve this system of differential equations rely on discretization of the fluid continuum. The two main classes of methods, Eulerian grid-based techniques and Lagrangian particle-based techniques, discretize the system using, respectively, cells at a fixed spatial location or particles as Lagrangian fluid tracers.

Grid codes come in different varieties depending on how exactly they solve the fluid equations, using finite differences or finite volume. For self-gravitating protoplanetary disks, finite difference spherical grid codes [36], cylindrical grid codes [37–41, 65], and smoothed particle hydrodynamics (SPH) codes [5, 9–11, 24, 25, 33, 35, 37, 42–48] have been used. We refer the reader to Durisen *et al.* [19] and Mayer *et al.* [49] for a detailed description of how these codes work. Briefly, in Eulerian grid codes one assigns physical variables to node points inside cells

organized in a grid and solves the relevant equations on the grid, typically in conservation form.

In Lagrangian SPH codes, one solves the fluid equations using particles as Lagrangian tracers of the fluid properties. The method is based on interpolation. Basic quantities such as density, pressure, and velocity are calculated at the location of a given particle by summing over a set of neighboring particles and smoothing by means of a kernel function (the spline kernel being the most popular choice, see [50]). Among these codes are the 3D SPH GASOLINE code [51] used in the calculations of Mayer and his collaborators, and the 2D and 3D SPH codes used by Nelson *et al.* [37, 52], Nelson [42, 53], Rice *et al.* [10, 11, 44–46], and Stamatellos *et al.* [24, 25, 47, 48], all based on a code originally developed by Benz [54].

Both grid and SPH codes have advantages and disadvantages. SPH codes, owing to their Lagrangian nature, are naturally adaptive in space and time; particles simply accumulate where the density of the flow increases, allowing a variable spatial resolution. This is a welcome feature in the case of a high dynamic range problem such as disk instability, in which densities and dynamical times can span several orders of magnitude. In a fixed grid code, the spatial resolution is set at the beginning of the simulation. SPH codes couple naturally with gravity solvers based on a tree algorithm [55]. Treecodes directly calculate the gravitational forces between neighboring particles, while they approximate the forces from more distant masses using a multipole expansion. They are faster and more accurate compared to Poisson solvers employed in grid codes. The accuracy and speed in computing gravity is, of course, crucial in modeling gravitational instability. On the other end, grid codes have better shock-capturing properties and better handling of discontinuities and turbulence in the flow [56]. Since grid codes solve differential equations on a finite grid, they need to specify boundary conditions that can also alter the character of the global flow if not chosen in a sensible way. SPH codes do not need boundary conditions since particles trace the fluid with no limitations on the computational volume and/or the values of the physical variables.

The discrete representation of the fluid continuum adopted by SPH forces the introduction of a diffusion term in the momentum and energy equation, called *artificial viscosity*, to stabilize the fluid flow against spurious fluctuations at the scale of the interparticle separation and to guarantee stable shock profiles [57]. It has been firmly established that artificial viscosity tends to suppress clump formation [9, 19, 49]. Artificial viscosity is not needed in grid codes using the most accurate, higher order methods to solve the fluid equations (although some numerical diffusion is always present). In some grid codes artificial viscosity also needs to be implemented to stabilize the flow, and can have both positive and negative effects on fragmentation, depending on the specific formulation used (e.g., [58, 59]). Nevertheless, at high resolution, SPH artificial viscosity produces flow characteristics in the disk comparable to those of high-order grid methods (e.g., the piecewise parabolic method (PPM)–[60]) that need no explicit artificial viscosity [61].

In both grid and particle codes, high resolution is crucial in order to achieve the correct result [19], but is more of an issue for grid codes since they are not adaptive

and are thus constrained by the initial resolution. In particular, not resolving the local Jeans or Toomre length can produce artificial fragmentation [53, 62] and miss physical fragmentation if the fastest growing spiral modes have characteristic wavelengths below the resolution [61].

Recently, simulations of disk instability with adaptive mesh refinement (AMR) codes have appeared in the literature [61]. These are grid codes in which grid spacing, and thus resolution, can be increased in a region of interest (e.g., in a collapsing region) [19, 61]. Adaptive meshes combine the advantages of both SPH and static grids, although they still need to specify boundary conditions.

A good resolution in the gravitational force is crucial in self-gravitating systems. Boss [58] has shown that excluding high order terms in the spherical harmonics expansion can lead to premature dissolution of clumps, and a similar effect of limited resolution in the potential solver was found by Pickett *et al.* [63]. In SPH codes, the resolution in the gravitational force is set by the so-called gravitational softening, which can be fixed (e.g., [5, 9]) or be varied according to the smoothing length used to compute hydrodynamical forces (e.g., [10, 11, 53]). In essence, the gravitational softening smoothens the gravitational force at distances comparable to the interparticle separation in order to avoid large accelerations at the resolution limit. If the softening scale is comparable to or is larger than the most unstable Toomre wavelength, spiral modes can be artificially damped, which can stifle fragmentation [9].

Since analytical solutions do not exist, code comparisons between radically different codes, such as SPH, cartesian AMR, and cylindrical fixed grids, are the only means to assess convergence of different methods. Currently, fragmentation in disks with $Q \sim 1.4$ and a locally isothermal equation of state is confirmed by a variety of codes [19, 61].

Existing codes also differ in the way they treat gas thermodynamics, in particular whether they use a fixed equation of state or solve the internal energy equation coupled with some approximate treatment of the radiative transfer equation. The role of thermodynamics in disk instability is of paramount importance and is discussed in the next section.

4.3
Disk Thermodynamics and Fragmentation

The ability of the gas disk to cool is crucial for instability to lead to fragmentation. After fragmentation, a clump will be able to collapse only if it can efficiently radiate away the energy generated by compression during the collapse. For quite some time work done in the context of disk instability has relied on using locally isothermal equations of state that assume perfect thermal balance at all times and everywhere, so that the disk is forced to maintain the initially assigned temperature profile [1, 29]. However, since spiral shocks can generate sudden, localized and copious heating, it is not obvious that thermal balance will be maintained in the disk. Various numerical experiments of the late 1990s showed that fragmentation is

strongly dependent on the assumed equation of state. Pickett *et al.* [63–65] showed that a fragmenting isothermal disk would self-regulate if evolved adiabatically. This happens because the Toomre Q rises well above the stability threshold as a result of compressional and shock heating. If compressional heating was included but irreversible shock heating was neglected, for example, when using an isentropic equation of state such as a polytrope, the disk was still able to achieve stability but could develop stronger overdensities, differing less from the isothermal case [65]. The question is, then, what is the most appropriate model for gas thermodynamics in the disk.

Efficient cooling requires the gas to be optically thin, that is, $\tau < 1$. The massive disks required for fragmentation are instead mostly optically thick in the disk midplane because of their high densities. In order to achieve $Q < 1.4$ at a radius of 5–10 AU, a disk around a solar mass star, with a surface density profile of the type $\sim r^{-3/2}$ requires a mass of $\sim 0.1 M_{\odot}$ within 20 AU for outer disk temperatures of order 40–60 K [1, 4, 5, 9, 21, 66]. The surface density for such a disk at 5 AU is close to 2000 g cm^{-3}, which translates into an optical depth $\tau > 50$ in the midplane for standard opacities. The high optical depths suggest that radiative losses may not be important. Such high optical depths permit the transport of radiation to be treated with the diffusion approximation [21, 22, 38, 40, 43, 67, 68, 69]. Boss [21, 67] found that in the clump-forming regions along the spiral arms the gas is so optically thick ($\tau > 100$) that it behaves nearly adiabatically. Indeed Mayer *et al.* [5, 9] used a switch between an isothermal and an adiabatic equation of state above a critical gas density threshold, finding more moderate fragmentation in adiabatic simulations than in the case of isothermal simulations.

However, simple assumption of two extreme regimes, optically thick or optically thin, ceases to be valid in the diffuse disk atmosphere, where τ may hover around unity. Indeed, the treatment of the cooling near the disk boundary is still a major issue. Different groups are using different schemes to model cooling in this intermediate regime, which could be the most important cause of the discrepancies among different disk instability calculations. Some authors find that the disk is able to cool globally on a timescale comparable to the orbital time and subsequently fragment by a combination of efficient midplane convection and efficient radiative cooling at the optically thin disk edge [43, 58, 67, 69, 70]. On the other end, others find that convection does not occur and the disk midplane is not able to lose the thermal energy generated in spiral shocks fast enough to allow for fragmentation (e.g., [25, 40, 48]. A point of agreement is that radiative diffusion alone cannot cool the disk midplane fast enough. The timescale $\tau_{\rm tr}$ for transporting thermal energy away from the midplane is indeed larger than 10^4 years at 10 AU in a typical disk, and is hence much longer than the orbital timescale, which is about 30 years at such distances. The radiative diffusion timescale $\tau_{\rm rad} = \kappa \rho h^2 / c$ would be very short ($\sim$0.03 years) if we assume typical values such as $\kappa = 1\,{\rm cm}^2\ {\rm g}^{-1}$ for the opacity, $\rho = 10^{-10}$ g cm^{-2}, and $h \sim 1$ AU for the disk scale height (c is the speed of light). However, thermal compression across the midplane in a massive disk is 10^6 times larger than radiation pressure [67], so that $\tau_{\rm tr} = 10^6 \tau_{\rm rad}$ is the actual time needed for radiation to carry the midplane thermal energy to the edge of the disk.

As a consequence, a nonradiative mechanism for transporting thermal energy away from the midplane is needed for the disk to fragment. Is convection such a missing mechanism? Boley *et al.* [38] find no convection and no fragmentation. They notice upwellings of gas from the midplane, but only in association with the dissipation of spiral shocks away from the midplane, ''shock bores'' [71]. Conversely, Boss [67–70] and Mayer *et al.* [43] find evidence for intermittent convection in the midplane. This appears to maintain a temperature of ~100 K even in the presence of strong spiral shocks, allowing fragmentation for at least some of their disk models. These authors observe steep vertical temperature profiles that are superadiabatic at the location of overdense regions, so that the Schwarzschild criterion for convection seems locally satisfied. All these works adopt the diffusion approximation in the disk midplane. In their most recent works, Boss and Mayer *et al.* employ a flux limiter to model the transition to the optically thin regime, while Boley *et al.* model directly the transport of radiation in the intermediate optical depth regime near the disk boundary. Mayer *et al.* [43] cool the disk boundary using a simple blackbody approximation with an efficiency parameter that depends on the effective radiative surface area. They find that fragmentation can be achieved or suppressed while varying the size of the radiating surface area by less than a factor of 2, suggesting that their disk models are really on the verge of fragmentation. Cai *et al.* [72] also include irradiation from the central star, which, as expected, stifles instability by reducing the net cooling at the disk boundary. Recently, Stamatellos *et al.* [47] have considered an approximation to radiative transfer in which gas parcels can cool according to a local opacity but cannot exchange energy. They found that fragmentation can only occur at large radii, $R > 40$ AU [25, 48].

Rafikov [73], using an analytical approach for a static disk, finds that even with convective cooling disks should not fragment at distances of less than about 100 AU because the photosphere of the disk cools too slowly for realistic disk parameters. He finds that the required masses, $M_{\rm disk} > 0.5 M_{\odot}$, and photospheric temperatures, several hundred kelvin, would be too high if compared with those of T Tauri disks [12]. On the other end, Mayer and Gawryszczak [61] suggest that the conditions set by Rafikov [73] can be considerably relaxed once a realistic dynamic disk is considered. In particular, the natural mass redistribution in a self-gravitating disk leads to local surface densities along spiral arms that are comparable to those in Rafikov's unstable analytic models (>2000 g cm^{-3}) despite the fact that the total disk mass is small enough to be in agreement with T Tauri disks observations, $M_{\rm disk} < 0.2 M_{\odot}$. Fragmenting disks appear to have photospheric temperatures lower than those found by Rafikov [73], although discrepancies are within a factor of 2. This important issue certainly requires further investigation [58].

In the outermost part of disks, at $R > 50$–100 AU, optical depths are much lower and the radiative cooling becomes considerably more efficient from the midplane to the surface. However, Boss [68] finds that fragmentation becomes unlikely at very large radii, $R > 100$ AU, because the already low surface densities are reduced further by the mass transport induced by spiral arms. An opposite result, in agreement with analytical predictions, is found by Stamatellos and Whitworth [48] and Boley [41]. Boley [41], however, highlights how mass loading onto the disk,

for example, by the surrounding envelope, is crucial at such distances in order to trigger a low Toomre Q. Boss [22] finds, instead, fragmentation into longlasting clumps at intermediate distances of 30–50 AU, suggesting that this might be how ice giants in the solar system formed (see Section 4.5.1). Fragmentation at such intermediate distances is marginally consistent with Rafikov's [73] model but has not been studied in simulations other than in [22]. Clearly, the location of fragmentation regions will also depend on the detailed shape of the disk mass and temperature profiles.

4.4 Beyond the Fragmentation Stage

In this section, we discuss the basic properties of the clumps formed via disk instability and how these depend on the methods used to evolve the disk beyond the fragmentation stage. We caution that fully hydrodynamical simulations of disk instability, due to computational limits, have only been able to follow clumps for a few thousand years [19]. The latter is comparable to only a few hundred orbits of the clumps, and it is much shorter than the 1–10 megayear disk lifetimes inferred from observations [6, 13]. As described below, techniques have been developed, for instance, the introduction of sink particles, to extend the time span of the simulations while not giving up on a fully hydrodynamical description of the system.

4.4.1 Masses, Numbers, and Evolution of Clumps Produced by Fragmentation

The masses of the clumps depend on both the local disk properties prior to fragmentation, which set the local Jeans and Toomre length, and on the thermodynamics of gas during and after fragmentation, which essentially determine how the collapse of the clumps proceeds.

The clumps are rotating oblate structures with a range of obliquities, but typically their rotation axis is perpendicular to the disk plane [9]. In simulations, clumps are identified as gravitationally bound objects. Mayer *et al.* [8] use the "virial" criterion to define boundaries, namely, where $2E_{\rm th} + E_{\rm kin} + W = 0$, where $E_{\rm th}$ is the thermal energy of a clump, W is the gravitational binding energy, and $E_{\rm kin}$ is its kinetic energy.

Matsuo *et al.* [26], assuming simply that the scale of clumps is imprinted by the conditions in the disk prior to fragmentation, find that disk instability produces planets with masses between 0.2 and 15 Jupiter masses. Mayer *et al.* [8] showed directly from the simulations that their lowest mass, fragmenting disks with a minimum temperature of 30 K at a few tens of astronomical units of a $0.085 M_{\odot}$ star, could produce Saturn-size clumps. Clumps do not necessarily contain all the mass within a sphere of radius $\sim\lambda_{\rm mu}$. It is only until the disk is barely nonlinear, that is, when the maximum densities along the spiral arms are only a few orders

of magnitude higher than the average local density, that the typical scale size of overdensities is still comparable to λ_{mu} [9].

Once formed, clumps can collide and merge with other clumps (Figure 4.1) as their orbits cross (due to eccentricity; see Figure 4.2). This is a nonlinear effect that perturbation theory cannot account for. In addition, as soon as a clump forms it begins to collapse further. At this point, thermodynamical effects become very important because the gas becomes completely opaque to radiation as its density continues to increase [67, 70]. This is another aspect of the physics that is missed by perturbation theory (which implicitly assumes isothermal conditions). The gas behaves nearly adiabatically in this regime [67]. The heat generated by compression generates a strong pressure gradient, which acts against the collapse. As a result, the clump density grows less than it does in isothermal conditions so that the object becomes more susceptible to shear [65, 9]. In fact, simulations switching from a locally isothermal to an adiabatic equation of state past fragmentation produce fewer and less-massive clumps, which grow less via accretion compared to those in locally isothermal simulations. A few tens of orbits after fragmentation, 2–3 clumps per disk survive and have masses in the range of 1–5 Jupiter masses when switching to adiabatic simulation as opposed to >10 per disk with masses in the range of 5–15 Jupiter masses when the calculation is continued as locally isothermal. ([5, 9]; see Figure 4.3). In either case, clumps form at distances between 5 and 15 AU, as this is the region where $Q < 1.4$. Boss [4, 22, 67, 69], using similar disk models, finds fragmentation at similar radii, 5–12 AU from the star. These results are valid for solar mass stars, while fragmentation can occur at smaller distances, 2–5 AU, for lower mass stars [9, 74]. Accretion rates onto clumps in locally isothermal calculations are unrealistically high, $\sim 10^{-4} M_{\odot}$ per year, so that most of the disk is consumed by the clumps in a few hundred years. Instead, in adiabatic simulations, accretion rates are $<10^{-5} M_{\odot}$ per year and the disk survives intact for several thousand years [19].

Instead of a fixed equation of state, more realistic models use a cooling time proportional to the local orbital time [10, 11, 44] and switch-off cooling at densities for which the optical depth τ is well above unity [33]. In this case, the mass accretion rates are in the range 10^{-6} to $10^{-5} M_{\odot}$ per year and the number of clumps is comparable to that of simulations with an adiabatic switch (Figure 4.2). A similar result is seen in even more realistic runs with flux-limited diffusion, in which cooling occurs via a combination of radiation and convection [43]. fragmentation at radii The mass function of clumps in flux-limited diffusion runs is presented in Figure 4.3, where it is also compared with that of locally isothermal runs with an adiabatic switch. Boss [67–70, 74] also finds a few clumps with masses of up to a few Jupiter masses in his runs adopting the diffusion approximation with a flux limiter or a constant envelope temperature as boundary condition. However, the clumps in the latter simulations are only marginally bound and it is unclear whether they would survive over long timescales.

In order to extend computations beyond ~100 clump orbits (roughly a few thousand years), some authors turn clumps into "sink particles" above a specified density threshold (e.g., [10, 11, 23, 44]). This way the later stage of the computation

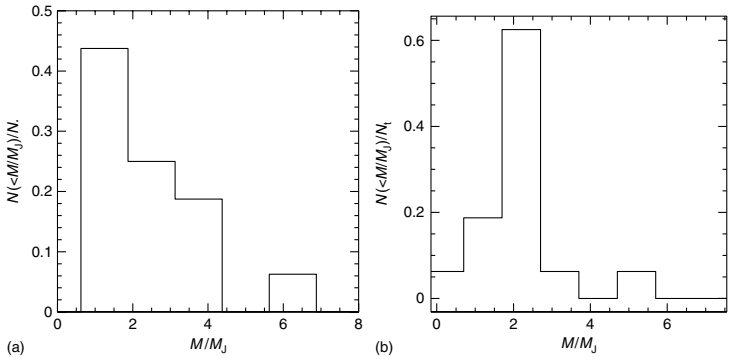


Figure 4.3 Mass function of clumps in the disk instability calculations presented in Mayer *et al*. [5, 9], that adopted a locally isothermal equation of state before fragmentation and an adiabatic one after (a), and in the combined set of three of the recent simulations including flux-limited diffusion presented in Mayer *et al*. [43] (b).

is not constrained by the short timesteps occurring inside the dense clumps. With this method, every clump survives the tidal stresses and the shear since the gravitational collapse of the clump is not modeled further. In addition, clumps become point masses before the replacement with sinks, thus offering reduced cross sections to interactions and mergers with other clumps. These last two features of sinks favor clump survival, and possibly explain why they produce many more clumps, ∼100, than the simulations that do not employ sinks [11, 19]. With sinks, the mass function of clumps is skewed toward lower values compared to fully hydrodynamical simulations, being roughly flat from 0.1 to 1 Jupiter mass, and then rapidly declining afterward as a power law. Owing to the sinks, Rice *et al*. [11] were able to evolve their simulations up to 10^4 years, at which point most of the gas in the disk was accreted onto the clumps or onto the central star. They then mapped the velocities and positions of the sinks and the central star into a gas-free, pure *N*-body calculation to follow their dynamics further. After about 10^5 years only one ∼7.8 Jupiter mass planet remained, since all the other smaller bodies had been scattered out of the system because of two-body encounters.

The mass distribution of observed extrasolar planets peaks about 1 Jupiter mass [75]. While the initial mass distribution of clumps right after fragmentation would be consistent with this observational result, the final masses of planets in disk instability simulations are typically higher, around a few Jupiter masses [9]. There are some important remarks to make regarding this discrepancy. First, existing simulations lack the resolution necessary to model the collapse of clumps realistically, which robustly determine their final masses. A resimulation of some of the radiative transfer models presented in [43] (and whose mass function is shown in Figure 4.3) with a resolution improved by more than an order of magnitude, now approaching ∼5 Jupiter radii, shows that a core-envelope structure not present in low-resolution simulations rapidly develops in the collapsing clump (see

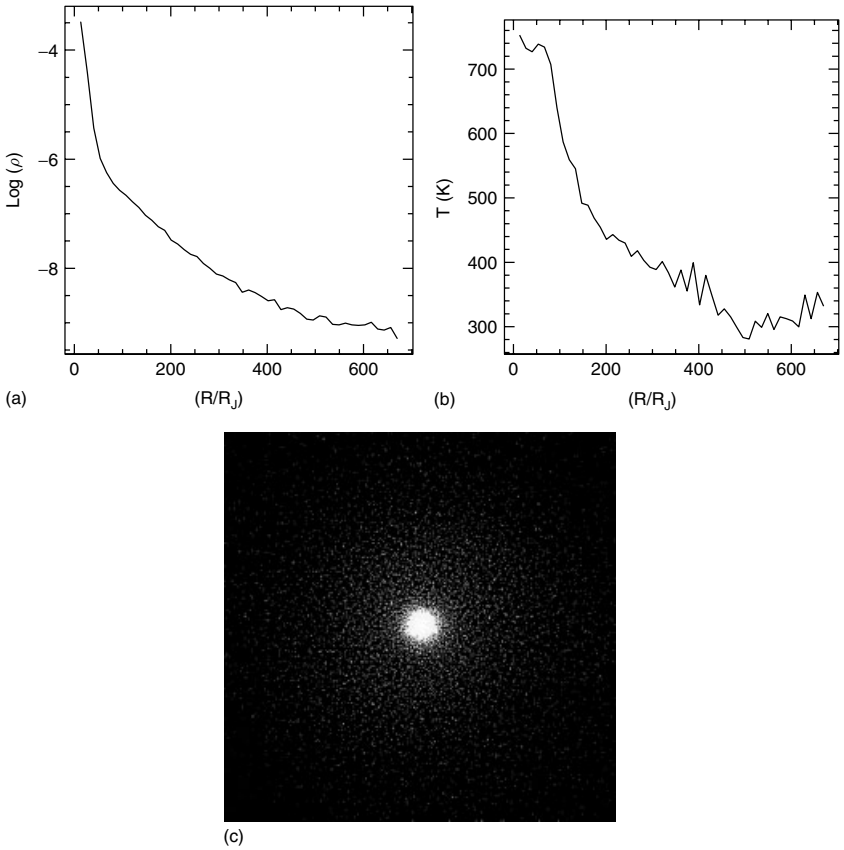


Figure 4.4 Density profile(density is measured in grams per cubic centimeter) (a), temperature profile (b), and color-coded logarithmic density map (c) of a clump formed in a higher resolution version of one of the fragmenting calculations presented in [43], adopting a molecular weight $\mu = 2.7$. The spatial resolution is ∼5 Jupiter radii. The clump has a core-halo structure. The gravitationally bound core has a mass of ∼1.6 Jupiter masses and is shown about 50 years after formation in (c), which is ∼0.1 AU on one side.

Figure 4.4). The core collapses to very high densities ($>10^{-4}$ g cm^{-3}) in a fraction of a disk orbital time and reaches a central temperature of ∼800 K (Figure 4.4). Interestingly, the mass of the core, the tightly bound component that can be identified as the protoplanet, is a factor of 2 lower than the mass of the bound clump found in a simulation with an order of magnitude less resolution. This suggests that the actual mass of the planetary-sized object at the end of the collapse may indeed be quite close to a Jupiter mass.

Similar cautionary remarks have to be made regarding the maximum mass that a planet can reach in disk instability. Nelson and Benz [76, 77] noted that accretion rates onto planetary-sized object computed in planet–disk interaction simulations ([78] for a review) are feeding rates rather than actual accretion rates.

Feeding rates are the rates at which the disk can supply the mass rather than the actual rate at which the planet can accept it. Moreover, they showed that the Kelvin–Helmoltz timescale imposes a stricter limit on accretion relative to that provided by feeding rates measured in the simulations. The Kelvin–Helmoltz timescale simply accounts for the fact that as the protoplanet contracts, it has to overcome the increased pressure gradient (see above). The associated limiting accretion rate corresponds to about 10^{-3} Jupiter masses per year for a Jupiter mass planet as opposed to up to 10^{-2} Jupiter masses per year for an accretion rate equal to the feeding rate. This Kelvin–Helmoltz limited accretion rate yields an upper limit of 10 Jupiter masses, assuming a constant rate over 10^4 years. However, in principle, a planet can grow well beyond 10 Jupiter masses if the gas disk survives for a million years. Therefore the real question is, what stops gas feeding for a newly born giant planet and keeps the mass within the range observed for extrasolar planets? Note that this question has nothing to do with the nature of disk instability but applies to a generic massive planet present in a gas disk with mass comparable to or larger than the minimum mass solar nebula, as shown by Nelson and Benz [76, 77].

An important role in determining the gas feeding rate is played by gap opening. Massive planets open a deep gap, at least in non-self-gravitating disks [78]. While the inflow of gas toward the planet does not stop in a realistic 3D disk (e.g., [79, 80]), the feeding rate is reduced by orders of magnitude. Simulations find that the maximum mass that a planet can achieve in a minimum solar nebula disk is around six Jupiter masses after gap formation [81]. Whether a planet is able to open a gap depends on its mass, the thickness of the disk (which depends on the local temperature and on the mass of the star), and viscosity.

Lodato and Rice [35, 82], following up on pioneering work [83], have shown that a gravitationally unstable disk can be described as a viscous α disk [84] in which gravitational torques provide the source of viscosity. They find that a disk that fragments reaches a critical maximum $\alpha \sim 0.06$. These high viscosities would make it difficult to open a clean, deep gap. Simulations of disk–planet interaction in a massive, locally isothermal self-gravitating disk indeed show that Jupiter-sized planets are fed substantially, even beyond gap opening, at a nearly steady rate of 10^{-3} to 10^{-2} Jupiter masses per year [85] for disk masses ranging between 0.01 and $0.1 M_{\odot}$. These high rates suggest that the Kelvin–Helmoltz timescale (see above) would still be the main limiting factor to the growth of the planet's mass. However, the feeding rates in [85] are likely exaggerated by the assumption of a locally isothermal equation of state.

In conclusion, while it is still unclear which mechanism can reduce the feeding rate substantially and limit the growth of the planet, we want to emphasize that disk instability simulations do not preferentially produce planets with masses comparable to the most massive among extrasolar giant planets (>5 Jupiter masses). The mass range yielded by the simulations after several thousand years of evolution, 1–10 Jupiter masses, is in reasonable agreement with the mass distribution of extrasolar planets. Containing further growth is a generic problem of planet-formation models, not just of disk instability. Disk dissipation mechanisms

can be invoked to reduce substantially the surrounding disk mass, thus lowering gas feeding rates. One such mechanism is gravitational torques arising during a gravitationally unstable phase of the disk that rapidly transport mass inward and angular momentum outward. In disk instability, accretion rates of the disks onto the central star are of order 10^{-2} Jupiter masses per year [9]; hence, they exceed the accretion rate limit onto the planet set by the Kelvin–Helmoltz timescale by one order of magnitude. Future simulations capable of probing disk evolution on much longer timescales than currently feasible will be needed to test if gravitational instability itself can lead to disk dissipation.

On the other hand, disk instability can easily produce planets with masses larger than the limiting mass of ~6 Jupiter masses estimated in disk–planet interaction simulations that assume the conditions typical of the core-accretion scenario [81]. Disk instability thus provides a better match to the high mass tail of extrasolar planets. Concerning the low-mass end of extrasolar planets, from Neptune-sized objects to super-Earths, additional mechanisms have been proposed to produce them in the context of disk instability, and are discussed in Section 4.5.

4.4.2 Orbits and Orbital Evolution

In disk instability, clumps form along spiral arms. Since spiral arms are non-axisymmetric features in the disk, it follows automatically that clumps form on noncircular orbits traced by the path of the arms. The average eccentricity is in the range of $e = 0.1–0.3$ for clumps formed in a variety of simulations [5, 9, 33], thus being consistent with the $e = 0.29$ mean value found for the orbital eccentricities of extrasolar planets [75]. In the longest calculation carried out so far, however, the orbital eccentricity of the two surviving clumps drops below $e = 0.1$. Yet, the gas thermodynamics in the latter simulation is so simplified that this result has to be taken with caution. Yet it appears that in disk instability, as in core accretion, one needs additional mechanisms to explain very high eccentricities ($e > 0.5$). Planet–planet gravitational interactions after the gas disk has been mostly dissipated might be playing a role (see Chapter 11). Studies of the dynamical interactions in a system comprising many massive planets formed via disk instability suggest that high eccentricity ($e > 0.5$) may naturally be obtained for the planets remaining in the system following a few ejections [86]. In the same models, high eccentricities arise also for lower mass giant planets eventually formed on a longer timescale by, for example, core accretion [87]. In their extended simulation with almost 100 gravitationally interacting clumps replaced with sink particles, Rice *et al.* [44] find that several planets are ejected, a few remain in the system but have large orbital distances well exceeding those of observed extrasolar planets ($R > 500$ AU), but two remain in the inner region of the disk ($R >\sim 5$ AU), having acquired an eccentricity as high as $e = 0.67$.

Because of the difficulty of running simulations over long timescales, migration and gap opening in the context of disk instability still have to be understood. Yet the few works that have explored migration seem to agree that there is no fast inward

migration for massive protoplanets formed by disk instability [19, 23, 44, 85]. This removes one of the major problems of the core-accretion scenario (see Chapter 9). However, this also means that disk instability does not provide a natural explanation for the existence of hot Jupiters, since *in situ* formation via fragmentation at less than 1 AU is extremely unlikely (see Section 4.1). Nevertheless, hot Jupiters may form as a result of dynamical relaxation of a system of many massive gas giants formed rapidly via disk instability and originally located at large (>10 AU) distances from the star [44, 86].

4.5 Comparison with Solar and Extrasolar Giant Planets

The impressive wealth of data on extrasolar planets gathered in the last decade or so shows that planets encompass a wide parameter space in terms of masses, orbits, metallicities of the host star, spectral type of the host star [75] and, perhaps, their own composition [88]. In this section, we discuss more in detail the comparison between the predictions of disk instability calculations and the various types of massive planets detected so far inside and outside the solar system, namely gas giants, ice giants, and super-Earths. Subsequently, we explore the predictions of disk instability concerning the correlation between the frequency of planets and the metallicity of the host star, the presence of planets around stars later than G, and the interior composition of the planets.

4.5.1 From Gas Giants to Ice Giants and Super-Earths

As discussed above, the disk instability model per se produces masses typically between 1 and 10 Jupiter mass. It is at present unclear if the model can explain why the typical mass of extrasolar gas giants is close to roughly one Jupiter mass [89], a question that only future simulations with enough resolution to follow the collapse of individual clumps will be able to answer (Section 4.4.1). Fragmentation cannot directly produce planets with masses lower than about Saturn (see Section 4.4). This would seem to exclude automatically disk instability as a viable formation model for a fairly large fraction of the planets. The tendency for more massive planets to be more frequent among those with larger orbital periods [90] would be instead natural in disk instability since fragmentation is more effective at large radii, although several possible explanations based on migration theory within standard core accretion [75] or multiple chaotic planet–planet interactions [91] have also been proposed.

Additional mechanisms have been proposed to explain the lower mass planets in the context of disk instability. Boss [22, 58, 92] has noted that if fragmentation occurs in disks inhabiting high-density star-forming regions, the fate of the protoplanets might be heavily affected by the local ultraviolet (EUV and FUV) radiation fields produced by nearby massive O and B stars. The UV radiation can heat and then evaporate the disk down to a radius determined by the balance between increased

hydrostatic pressure and the gravitational restoring force of the star. For stars with the mass of the Sun, Boss [22] finds that evaporation might occur down to a radius of 15–20 AU. Massive gas clumps there would also be partially evaporated. As we discuss later, a core of up to 10 Earth masses may form in gas giants formed via disk instability as a result of planetesimal accretion and dust grain sedimentation (see Section 4.5). If such relatively massive solid core is present, photoevaporation of a several-Jupiter-masses clump would leave behind the solid core and blow off most of the gaseous envelope, resulting in a planet with composition and mass similar to that of ice giants in the solar system [22].

If the gas clump is relatively small to begin with, say not exceeding a Jupiter mass, a proportionally smaller solid core would be left behind by photoevaporation, resulting in a rocky planet whose mass is comparable to that of the discovered Super-Earths. According to Boss [92], such a mechanism can work down to a few astronomical units in the case of M dwarfs, because for such lower mass stars photoevaporation of the gas disk is more effective. Super-Earths around M dwarfs are known at distances of <1 AU. Some migration seems to be required in order to bring the planet inward from a more plausible formation site. As described above (Section 4.3) whether significant migration happens for planets formed via disk instability is not clear yet. However, the migration of super-Earths may occur in conditions more similar to standard type-I migration because of their lower masses and because the surrounding disk would have become much lighter as a result of photoevaporation, approaching the conditions usually assumed in standard disk–planet interaction simulations. Simulations tailored to this specific scenario including photoevaporation will have to be designed.

4.5.2
Fragmentation around Different Stellar Types

The stellar mass range probed by most existing radial velocity surveys is quite small (FGK dwarfs with masses ranging from 0.7 to $1.3 M_{\odot}$). The increasing data on the frequency of planets around M dwarfs suggests that giant planets may be sparser around such stars, although this applies mostly to short-period planets [93]. Boss [74] has run disk instability simulations including radiative transfer with the diffusion approximation finding that fragmentation into gas giants is as likely as it is around normal stars. Instead, core accretion would be slower and less efficient around an M dwarf because of the longer orbital times around lower mass stars [94]. A more complete census of the frequency of planets around M dwarfs may thus provide valuable constraints on formation models (see also Section 4.9 on similar constraints coming from surveys of planets in binary systems).

4.5.3
The Correlation between Metallicity of Host Star and Planet Frequency

The correlation between the frequency of planets and the metallicity of the host star is a remarkable property of the extrasolar planets population and one that can

potentially constrain their formation mechanism [75, 89]. The observed correlation mostly reflects the conditions of the nebular gas from which both the star and the planets arise, rather than some event following planet formation itself. In core accretion, the observed correlation should arise naturally since the growth of a solid core should be more efficient in a nebula with a higher metallicity, at least assuming, as usual, that the metallicity of the star correlates with the dust-to-gas ratio in the nebula ([95–97]; Chapter 4). Nevertheless, the correlation almost vanishes at low metallicities, $Fe/H \sim -0.2.$, where the frequency of planets becomes nearly constant over a range of metallicities. This plateau at low metallicities suggests that planets can form even around metal-poor stars.

The shape of the mass function of planets as a function of metallicity of the host star predicted by the core-accretion model has a steep drop toward low metallicities rather than a plateau [98]. The predictions for the disk instability model concerning whether one expects a correlation with stellar metallicity are quite controversial at the moment. Naively, one would expect no correlation with metallicity since gravitational instability involves the direct formation of a giant planet from the gas phase rather than a combined action of the gas and solid phase. However, radiative cooling is affected by metallicity, and cooling dictates whether fragmentation happens or not (see Section 4.1). The effect of metallicity on cooling is twofold. A higher metallicity enhances cooling in the optically thin regime, namely, near the disk surface, while it suppresses cooling in the optically thick regime, namely, in the disk midplane [19]. Boss [67] explored the effect of changing metallicity by only modifying the opacity in the optically thick regime. He found no effect on the behavior of gravitational instability even when the opacity was changed by over two orders of magnitude. This is understandable since the effective radiative cooling time of the midplane of a massive disk is not just two, but several orders of magnitude longer than the orbital time in the disk midplane (see Section 4.1).

Cai *et al.* [99, 100] examined the effect in the optically thin disk boundary, finding that cooling becomes more or less efficient by increasing or decreasing the metallicity within the range of observed host star metallicities. Note, however, that all the models of [99] do not fragment, even for supersolar metallicities comparable to those of the most metal-rich stars in observed samples. Therefore, both Cai's and Boss' works, despite their different conclusions regarding fragmentation, suggest that giant planet formation via disk instability is rather insensitive to metallicity. Boss [58] argues that this would be consistent with the plateau of the observed metallicity correlation. On the other hand, it would be unclear as to how to explain the remarkable correlation seen at high metallicities. Since most of the planets discovered so far are found around stars with metallicities comparable to, or larger than, solar, Matsuo *et al.* [26] concluded that core accretion is likely responsible for the formation of about 90% of the population of extrasolar planets while disk instability could be the correct formation mechanism for planets around low-metallicity stars. Rice *et al.* [11] also noticed that planets around metal-poor stars are preferentially more massive than Jupiter, which would also be consistent with disk instability being the main formation mechanism in this case.

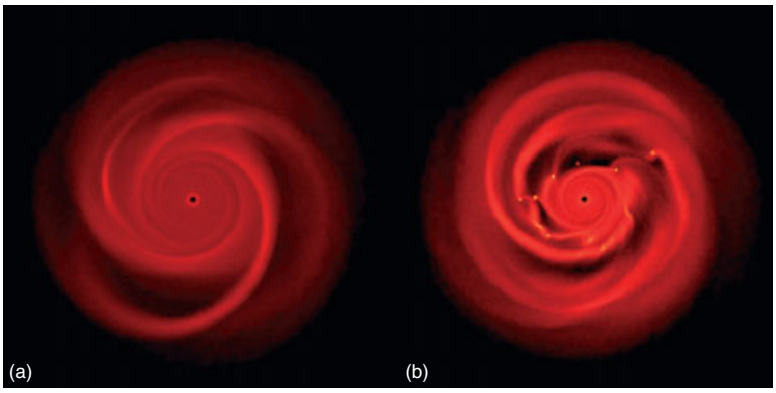


Figure 4.5 Color-coded logarithmic density maps of two flux-limited diffusion simulations presented in [43] (see Figure 4.2) that adopt the same efficiency of cooling at the disk boundary but different molecular weights ($\mu = 2.4$ (a) and $\mu = 2.7$ (b)). Note how fragmentation happens only with the higher molecular weight (see text).

Mayer *et al.* [43] proposed a new scenario in which a higher metallicity can indeed promote disk instability. They found that fragmentation is enhanced if the gas has a higher molecular weight μ. Since fragmentation occurs along spiral arms for the remainder of this section, we assume that the relevant molecular weight will be the local one along the spiral arms. Fragmentation can be enhanced because with higher molecular weight the gas undergoes a lower compression ($P \sim \frac{1}{\mu}$); hence heating along the spiral arms, which tends to stabilize the gas against fragmentation, is reduced. In Figure 4.5, we show that changing the molecular weight from solar ($\mu = 2.4$) to just 10% above solar ($\mu = 2.7$) brings the disk from a marginally unstable state to fragmentation. It is important to recall that the dependence of molecular weight on metallicity is very weak. Significant changes in the gas molecular weight can only occur when (i) very large variations of the dust-to-gas ratio occur and (ii) the metal content of the dust grains is returned to the gas phase.

4.6
The Solid Content of Planets Formed by Disk Instability

Recently, a number of works have attempted to address the solid content of planet formed by disk instability. Such works employ semianalytical models or a mixture of N-body and semianalytical models to describe the interaction between the gas clump and the solid component in the disk.

Helled *et al.* [101, 102] and Helled and Schubert [103] start from the density and temperature structure of a Jupiter-sized clump formed in one Boss [4] simulation, then calculate the collapse of such clumps assuming spherical symmetry,

computing the sedimentation of dust and the accretion of planetesimals. They consider a range of compositions for the solids, such as rocks (i.e., silicate materials) and organic material (CHON). Planetesimals can be destroyed and returned to the dust phase as a result of the thermal and mechanical ablation while they move through the clump atmosphere. Dust growth is also taken into account via a standard coagulation equation, and convection in the gas envelope is also accounted for. Solids that sediment down to the center of the clump contribute to the core. The authors find that rocky cores of 2–12 Earth masses can be formed in Jupiter-mass clumps and that their envelopes can be enriched beyond solar up to the levels observed in Jupiter and Saturn (40 Earth masses, see also [104]). The sedimentation of grains into the core occurs in a few thousand years during the contraction phase of the clump. How much material ends up in the core depends on the temperature and density evolutions of the clumps. If the center of the clump becomes too hot too rapidly during contraction, then all types of dust grains are vaporized before they can be deposited in the center, and no core would form. On the other hand, feeding of solid material via planetesimal accretion, especially for large planetesimals with sizes of order 100 km, is more efficient during later stages of collapse when the clump has become denser.

Helled and Schubert [103] extended their model to clumps with masses up to 10 Jupiter masses, but considered only the sedimentation of dust grains. They found that protoplanets with several Jupiter masses should not develop a core, mainly because temperatures in these more massive clumps reach values as high as 1000 K much too rapidly during contraction, sublimating both ices and silicates on a short timescale. This yields the interesting prediction that larger planets should have proportionally smaller cores if they formed via disk instability. This could be a problem for a scenario in which ice-giant planets form via photoevaporation of gas giants with masses much larger than Jupiter and proportionally more massive cores [22, 92].

A new calculation of core formation and envelope enrichment in clumps formed via disk instability has been performed [105] using the structure of clumps formed in a higher resolution version of one of the simulations presented in [43]. The impact model adopted in this case is that used by Alibert *et al.* [96, 97] in semianalytical models of core accretion. Only accretion/sedimentation of planetesimals present in the Hill radius of a Jupiter-sized clump was included (no dust grains were assumed to be already present within the clump since the beginning). The impact parameters and the fraction of accreting planetesimals with a range of sizes (1–1000 km) and composition (ices and silicates) were determined via an *N*-body simulation of a minimum mass solar nebula planetesimal disk with an already formed Jupiter-sized gravitating body, a few Jupiter radii in size. The contraction of the clump to such a small radius occurs in only about 100 years instead of the $>10^5$ years predicted by the quasi-static spherical collapse model used in Helled *et al.* [101, 102]. The contracting clump develops a clear core-envelope structure at such high resolution, with the core reaching very high densities (Figure 4.4). This allows the protoplanetary core to accrete planetesimals more efficiently because its Roche lobe is much larger due to its deeper potential well. On the other hand,

temperatures inside the rapidly contracting protoplanet rise faster, increasing the effectiveness of thermal ablation of the accreted planetesimals and the vaporization of the accreted grains relative to a colder, lower density clump. The final mass of the core achieved ranges from a few to 10 Earth masses, while the maximum envelope enrichment is at the most comparable with Jupiter. We caution that these results, while interesting, stem from a model that is not really self-consistent. In principle, one should take into account the fact that the solid content of a massive disk, arranged in both planetesimals and dust grains, may be much higher than that of a minimum solar nebula [106]. Planetesimal formation in such a disk may also take a completely different path compared to standard collisional accumulation as it may be triggered by gravitational instability in the dust layer itself [45]. The main conclusion is that there is no fundamental obstacle to forming a core in a giant planet formed via disk instability or to enriching its envelope significantly beyond solar.

First attempts toward self-consistent disk simulations that incorporate both the gas and the solid component have been made in the last few years. Mejia *et al.* [107] have followed a global 3D simulation of one of the gravitationally unstable gas disk models presented in [9] including dust grains with a range of sizes and composition but only for a few orbits. They showed that dust concentrates in the spiral arms owing to gas drag in less than a disk orbital time. The point to note is that it concentrates in the same sites where clumps would form if fragmentation occurs, as argued by previous 2D calculations [108, 109]. Rice *et al.* [45, 46] have run simulations with 50-cm planetesimals and included their self-gravity, showing that once the particles concentrate in the spiral arms, the local surface density of the layer of solids can increase significantly and render it gravitationally unstable. Irrespective of whether or not the solid component becomes locally gravitationally unstable, an important result of these works is that the dust-to-gas ratio can increase by an order of magnitude or more at the density maxima along the spiral arms. This suggests that a clump formed by fragmentation of the gas disk, which would typically emerge from a spiral arm, may also have a dust-to-gas ratio higher than the mean. This would promote core formation and envelope enrichment.

4.7
Disk Formation and Disk Masses – a Key Issue

Fragmentation via disk instability requires a disk that is massive and can cool fast. As we have already discussed, the second requirement involves the detailed interplay between the various cooling and heating processes. The first requirement is fundamentally determined by the balance between the mass accretion from the molecular envelope onto the disk and the mass transport from the disk to the central star. In general, the relative importance of the two mass transport rates can be a function of time and distance from the star, and will determine the mass and surface density of the disk at any time and location [110].

Computationally, the challenge is daunting since one should simultaneously resolve large scales (>0.1 pc, the typical size of a molecular core) and small scales (<0.1 AU, the size of disk inner boundary) in order to compute the differential mass transport in the disk. Such calculations should also be carried out in three dimensions and with realistic radiation physics. In addition, the initial state of the collapse should reflect the large-scale properties of the turbulent interstellar medium. For massive stars, substantial progress in this direction has been achieved with high-resolution AMR calculations [111, 112] that properly include self-gravity and radiation physics, and start from a realistic turbulent cloud. In the latter, the early accreting protostellar disk can become rapidly gravitationally unstable and form secondary stellar companions, although accounting for radiative transfer effects appears to reduce the chances of fragmentation.

For low-mass stars, fully hydrodynamical collapse calculations begin with idealized models of rotating nonturbulent gas clouds and so far can cover only the first phase of protostellar disk formation ($<10^5$ years after the collapse of the core begins) during which the disk has a mass still comparable with that of the central star and is still undergoing substantial mass redistribution [14, 49, 113–115]. In this stage, the disk is seen to accrete vigorously from the molecular envelope, at rates exceeding $10^{-4} M_\odot$ per year. Some of these calculations predict that the disk should reach $Q \sim 1$ in the very early stages, but others suggest that the disk may self-regulate to $Q \sim 1.5$–2 and avoid fragmentation [116].

Recent simulations that include magnetohydrodynamical effects show that disks are more stable in magnetized collapses [117], although processes that can diminish the coupling between magnetic fields and matter, for instance, ambipolar diffusion, are not accounted for yet. If the disk fragments early, when its mass is still as large as that of the star, it will likely form large objects such as brown dwarfs, which may consume a large fraction of its gas and prevent fragmentation into planetary-sized clumps at a later stage ([48] –note, however, that a few clumps in their simulations have masses comparable to the largest among extrasolar giant planets).

As elucidated in the previous sections, fragmentation of disks into Jupiter-sized objects can be achieved for disk masses in the range 0.1–$0.2 M_\odot$, and hence smaller than those found in the early stages of protostellar disk formation. It is unclear if such masses characterize a later stage of disk evolution. They would correspond to the most massive among T Tauri disks, which are typically in the range 0.01–$0.05 M_\odot$ [12, 13]. Yet, when comparing with observed disk masses, a major caveat is that gas masses at >5 AU are not directly determined from the molecular hydrogen content of the disks but rather from dust emission or emission by secondary molecular gas tracers such as CO [118]. Models of self-regulated accretion rates in T Tauri disks that match the observationally estimated accretion rates onto the star in such systems suggest that T Tauri disk masses could be underestimated [110]. Alternatively, it is very likely that disks with masses of $0.1 M_\odot$ characterize a phase later than class 0 but earlier than T Tauri. Interestingly, while traditionally planet-formation models for the solar system have assumed the $\sim 0.01 M_\odot$ minimum mass solar nebula (MMSN) model of Weidenschilling [119] as an initial condition for the disk mass recently it has been argued that a much

more massive protosolar nebula, with mass $\sim 0.1 M_{\odot}$ might be more consistent with the so-called Nice model [106], which successfully explains several aspects of the architecture of our solar system [120]. Over the next decade, upcoming instruments such as the Atacama Large Millimeter Array (ALMA) and the James Webb Space Telescope (JWST) will shed light on the evolution of disk mass during the phases preceding the T Tauri stage, when the disk is still embedded in the molecular cloud envelope and still vigorously accreting from it. This will certainly boost our knowledge of disk formation and evolution, clarifying the conditions of protostellar disks at various stages.

4.8 The Role of Stellar Companions and Interactions

About two-thirds of stars in the disk of our Galaxy are binaries [121]. In one of the two main models of star formation, competitive accretion, even relatively isolated stars form in high-density regions where gravitational interactions with other stars and protostellar disks should be frequent [122, 123]. In either case, the early massive protoplanetary disk might be affected by frequent gravitational interactions with perturbers. The main question is whether such interactions will enhance or reduce the chances of fragmentation. The interaction with a companion in a binary system was studied in the pioneering work of Nelson [42], and later by Boss [124] and Mayer *et al.* [33]. Lodato *et al.* [125] studied the case of fly-byes, namely, interactions with perturbers on unbound orbits. These various works are summarized and compared in the recent review paper by Mayer *et al.* [49]. Most experiments considered two nearly equal mass star+disk systems and orbits with moderate eccentricity ($e < 0.5$). Mayer *et al.* [33] and Nelson [42] found that fragmentation by disk instability is unlikely in relatively tight binaries with separations less than 60 AU. Mayer *et al.* [33] carried out a systematic comparison between isolated and binary protoplanetary disks showing that the tidal perturbation of the companion heats the disk by compressing it near pericenter passages and by exciting stronger spiral shocks compared to those occurring in isolation. As a result, unrealistically short cooling times, less than half of an orbital time, are needed for fragmentation to occur in a relatively massive disk ($M \sim 0.05$–$0.1 M_{\odot}$). One of these simulations was repeated with flux-limited diffusion finding that fragmentation does not occur because such short cooling times cannot be achieved [49]. Boss [124], instead, found that binarity can promote fragmentation because a disk that would be Toomre stable in isolation can become unstable if perturbed. In his simulations, the heating associated with the emergence of the strong tidally triggered spiral arms is not effective at stifling the development of overdensities.

The disagreement between these works occurs for separations <60 AU, while for separations larger than 100 AU, both Mayer *et al.* [33] and Boss [124] agree that fragmentation can take place in a manner quite similar to the isolated disk case, requiring cooling times commensurate with the orbital time. A likely reason of the discrepant results at separations of 60 AU is that Nelson [42] and Mayer *et al.* [33]

include artificial viscosity in the thermal energy equation, while Boss [124] does not (although some intrinsic numerical viscosity is present to guarantee code stability). Artificial viscosity tends to suppress clump formation (see Section 4.1) via increased heating and augmented random kinetic energy [9,49]. Repeating these experiments with good shock-capturing codes that do not need artificial viscosity, for example, PPM grid codes, is the only way to assess which of these works provides the correct answer.

If fragmentation is really suppressed by binarity as suggested by Mayer *et al.* [33], then we have a new way to constrain planet formation mechanisms. Currently, there are hints of a deficiency of giant planets around binary stars with separations $<$100 AU [126], although samples are still small ([127] for a larger sample that is currently under study). The perturbation of the companion has effects on the solid constituents of the disk as well. In massive disks ($M \sim 0.05$–$0.1 M_{\odot}$), the resulting intense heating creates temperatures of $\sim$200–300 K at $R \sim 5$–15 AU, which would vaporize ice grains. Since ice grains should comprise about 30% of the solid material in a standard protosolar nebula model, one would expect that solid core formation may be hampered as well, so that even core accretion would be unlikely in such systems [42].

Mayer *et al.* [33] showed that for light disks with masses comparable to the minimum solar nebula, $M \sim 0.01 M_{\odot}$, the companion triggers only transient heating at the pericenter, so that the disk temperature stays unchanged on average. Hence, in such disks, survival and coagulation of grains into planetesimals should occur much in the same way as in isolation. Light disks respond differently from massive disks to the companion's perturbation because their self-gravity is lower. In fact, when spiral arms are excited by the perturbation of the companion, the strong amplification of spiral modes, and thus the ramping up of the associated shock heating, requires that the disk is strongly self-gravitating, and hence quite massive [49]. In conclusion, in the light-disks, core-accretion models proceed unaffected by binarity. Therefore, one would expect no trend between binarity and the frequency of giant planets if core accretion is the main giant planet-formation mechanism.

4.9 Looking into the Future

We live in an exciting time for planet formation. In addition to the over 300 planets discovered by radial velocity surveys, in the last year the first planetary-sized bodies with masses in the range of gas giants have been discovered by direct imaging at distances larger than 50 AU from the central star [128–130]. Core accretion *in situ* formation of gas giants beyond 10–20 AU is hardly achievable [98] while fragmentation at such large radii should be likely [41, 48, 73]. However, one cannot exclude that such planets formed closer in by core accretion and then migrated outward. Assessing whether the observed configurations are the result of the dynamical evolution of an initially more compact system or whether the system

had to start with the planets on very wide orbits will thus constrain formation models.

For fragmentation at smaller radii, 5–20 AU, the jury is still out. Current disagreement likely stems from different treatments of radiation physics. This will be clarified in the next few years since detailed code comparisons are already under way. If fragmentation at $R < 20$ AU is unlikely, giant planets located at such distances could still have formed via disk instability; they could have formed much further out and migrated inward. Such a scenario cannot be excluded even for the solar system because disk instability is compatible with the presence of solid cores and the envelope enrichment of its gas giants. Making further progress in understanding the role of disk instability will inevitably demand calculations capable of following the different stages of disk formation and evolution over sufficiently long timescales, thereby connecting star and planet formation.

We face a challenging computational endeavor calling for new, highly scalable hydrodynamical codes capable of taking advantage of the new generation of petascale parallel supercomputers. At the same time, the next generation of space-born and ground-based astronomical observatories, such as ALMA, JWST, and HERSCHEL, will shed light on the various phases of disk formation and evolution. The synergy between the new groundbreaking observations and the new computer models will be key to confirming the true nature of the planet-formation process.

References

1 Boss, A.P. (**1997**) *Science*, **276**, 1836.
2 Kuiper, G.P. (**1951**) *Proceedings of the National Academy of Sciences of the United States of America*, **37**, 1.
3 Cameron, A.G.W. (**1978**) *Moon Planets*, **18**, 5.
4 Boss, A.P. (**1998**) *The Astrophysical Journal*, **503**, 923.
5 Mayer, L., Quinn, T., Wadsley, J., and Stadel, J. (**2002**) *Science*, **298**, 1756.
6 Haisch, K.E. Jr., Lada, E.A., and Lada, C.J. (**2001**) *The Astrophysical Journal*, **533**, L153.
7 Toomre, A. (**1964**) *The Astrophysical Journal*, **139**, 1217.
8 Binney, J., and Tremaine, S. (**1987**), Princeton University Press.
9 Mayer, L., Quinn, T., Wadsley, J., and Stadel, J. (**2004**) *The Astrophysical Journal*, **609**, 1045.
10 Rice, W.K.M., Armitage, P.J., Bate, M.R., and Bonnell, I.A. (**2003**) *Monthly Notices of the Royal Astronomical Society*, **339**, 1025.
11 Rice, W.K.M. *et al.* (**2003**) *Monthly Notices of the Royal Astronomical Society*, **346**, L36.
12 Beckwith, S.V.W., Sargent, A., Chini, R.S., and Guesten, R. (**1990**) *The Astrophysical Journal*, **99**, 924.
13 Andrews, A. and Williams, J.P. (**2005**) *The Astrophysical Journal*, **631**, 1134.
14 Yorke, H.W. and Bodenheimer, P. (**1999**) *The Astrophysical Journal*, **525**, 330.
15 Boss, A.P. (**1996**) *The Astrophysical Journal*, **469**, 906.
16 Boss, A.P. (**1998**) *Annual Review of Earth and Planetary Sciences*, **26**, 53.
17 Carmona, A. (**2009**), *Proceedings of the Conference "Formation and Evolution of Planets"*, Ascona, Switzerland, June 29–July 2, 2008, in press on Earth Moon and Planets.
18 Kawakita, H. *et al.* (**2001**) *Science*, **294**, 1089.
19 Durisen, R.H. *et al.* (**2007**) *Protostars and Planets V* (eds. B. Reipurth, D.

Jewitt, and K. Keil), University of Arizona Press, Tucson, p. 607.
20 D'Alessio, P., Calvet, N., and Hartmann, L. (**2001**) *The Astrophysical Journal*, **553**, 321.
21 Boss, A.P. (**2001**) *The Astrophysical Journal*, **563**, 367.
22 Boss, A.P. (**2003**) *The Astrophysical Journal*, **599**, 577.
23 Boss, A.P. (**2005**) *The Astrophysical Journal*, **629**, 535.
24 Stamatellos, D., Hubber, D., and Whitworth, A.P. (**2007**) *Monthly Notices of the Royal Astronomical Society*, **382**, L30.
25 Stamatellos, D. and Whitworth, A.P. (**2009**) *Monthly Notices of the Royal Astronomical Society*, **392**, 413.
26 Matsuo, T., Shibai, H., Ootsubo, T., and Tamura, M. (**2007**) *The Astrophysical Journal*, **622**, 128.
27 Escala, A. and Larson, R. (**2008**) *The Astrophysical Journal*, **685**, L31.
28 Papaloizou, J.C. and Savonije, G.J. (**1991**) *Monthly Notices of the Royal Astronomical Society*, **248**, 353.
29 Laughlin, G. and Bodenheimer, P. (**1994**) *The Astrophysical Journal*, **436**, 335.
30 Gammie, C.F. (**2001**) *The Astrophysical Journal*, **553**, 174.
31 Johnson, B. and Gammie, C.F. (**2003**) *The Astrophysical Journal*, **597**, 131.
32 Laughlin, G., Korchagin, V., and Adams, F.C. (**1997**) *The Astrophysical Journal*, **477**, 410.
33 Mayer, L., Wadsley, J., Quinn, T., and Stadel, J. (**2005**) *Monthly Notices of the Royal Astronomical Society*, **363**, 641.
34 Mejia, A.C., Durisen, R.H., Pickett, M.K., and Cai, K. (**2005**) *The Astrophysical Journal*, **619**, 1098.
35 Lodato, G. and Rice, W.K.M. (**2004**) *Monthly Notices of the Royal Astronomical Society*, **351**, 630.
36 Boss, A.P. and Myhill, E.A. (**1992**) *The Astrophysical Journal Supplement Series*, **83**, 311.
37 Nelson, A.F., Benz, W., Adams, F.C., and Arnett, W.D. (**1998**) *The Astrophysical Journal*, **502**, 342.
38 Boley, A.C. *et al.* (**2006**) *The Astrophysical Journal*, **651**, 517.
39 Boley, A.C., Hartquist, T.W., Durisen, R.H., and Scott, M. (**2007**) *The Astrophysical Journal*, **656**, L89.
40 Boley, A.C., Durisen, R.H., Nordlund, A., and Lord, J., *et al.* (**2007**) *The Astrophysical Journal*, **655**, 1254.
41 Boley, A.C. (**2009**) *The Astrophysical Journal*, **696**, L53.
42 Nelson, A.F. (**2000**) *The Astrophysical Journal*, **537**, L65.
43 Mayer, L., Lufkin, G., Quinn, T., and Wadsley, J. (**2007**) *The Astrophysical Journal*, **661**, L77.
44 Rice, W.K.M., Lodato, G.,Pringle, J.E. *et al.* (**2004**) *Monthly Notices of the Royal Astronomical Society*, **355**, 543.
45 Rice, W.K.M., Lodato, G., and Armitage, P.J. (**2005**) *Monthly Notices of the Royal Astronomical Society*, **364**, L56.
46 Rice, W.K.M., Lodato, G., Pringle, J.E. *et al.* (**2006**) *Monthly Notices of the Royal Astronomical Society*, **327**, L9.
47 Stamatellos, D., Whitworth, A.P., Bisbas, T., and Goodwin, S. (**2007**) *Astronomy and Astrophysics*, **475**, 37S.
48 Stamatellos, D. and Whitworth, A.P. (**2008**) *Astronomy and Astrophysics*, **480**, 879.
49 Mayer, L., Boss, A.P., and Nelson, A.F. (**2008**) *Review Chapter in Volume Planets in Binary Star Systems*, Springer, Nader Haghighipour.
50 Hernquist, L. and Katz, N. (**1989**) *The Astrophysical Journal Supplement Series*, **70**, 419.
51 Wadsley, J., Stadel, J., and Quinn, T.R. (**2004**) *New Astronomy*, **9**, 137.
52 Nelson, A.F., Benz, W., and Ruzmaikina, T.V. (**2000**) *The Astrophysical Journal*, **529**, 357.
53 Nelson, A.F. (**2006**) *Monthly Notices of the Royal Astronomical Society*, **373**, 1039–1073.
54 Benz, W. (**1990**) *The Numerical Modeling of Stellar Pulsation* (ed. J.R. Buchler), Kluwer, Dordrecht, p. 269.
55 Barnes, J. and Hut, P. (**1986**) *Nature*, **324**, 446.
56 Agertz, O. *et al.* (**2007**) *Monthly Notices of the Royal Astronomical Society*, **380**, 963.
57 Monaghan, J.J. (**1992**) *Annual Review of Astronomy and Astrophysics*, **30**, 543.

58 Boss, A.P. (**2007**) *The Astrophysical Journal*, **661**, L73.

59 Pickett, M.K. and Durisen, R.H. (**2007**) *The Astrophysical Journal Letters*, **654**, L155.

60 Colella, P. and Woodward, P.R. (**1984**) *Journal of Computational Physics*, **54**, 174.

61 Mayer, L. and Gawryszczak, A. (**2008**) *Extreme Solar Systems*, Vol. 398, ASP Conference Series, proceedings of the conference held 25–29 June, 2007 (eds. D. Fischer, F. Rasio, S. E. Thorsett and A. Wolszczan), Santorini Island, ASP, Greece, p. 2.

62 Bate, M.R. and Burkert, A. (**1997**) *Monthly Notices of the Royal Astronomical Society*, **228**, 1060.

63 Pickett, B.K. *et al.* (**2000**) *The Astrophysical Journal*, **529**, 1034.

64 Pickett, B.K., Durisen, R.H., and Davis, G.A. (**1996**) *The Astrophysical Journal*, **458**, 714.

65 Pickett, B.K., Cassen, P., Durisen, R.H., and Link, R. (**1998**) *The Astrophysical Journal*, **504**, 468–491.

66 Boss, A.P. (**2000**) *The Astrophysical Journal*, **545**, L61.

67 Boss, A.P. (**2002**) *The Astrophysical Journal*, **576**, 462.

68 Boss, A.P. (**2006**) *The Astrophysical Journal*, **637**, L137.

69 Boss, A.P. (**2008**) *The Astrophysical Journal*, **677**, 607.

70 Boss, A.P. (**2004**) *The Astrophysical Journal*, **610**, 456.

71 Boley, A.C. and Durisen, R.H. (**2006**) *The Astrophysical Journal*, **641**, 534.

72 Cai, K. *et al.* (**2008**) *The Astrophysical Journal*, **673**, 1138.

73 Rafikov, R.R. (**2007**) *The Astrophysical Journal*, **662**, 642.

74 Boss, A.P. (**2006**) *The Astrophysical Journal*, **643**, 501.

75 Udry, S. and Santos, N.C. (**2007**) *Annual Review of Astronomy and Astrophysics*, **45**, 397.

76 Nelson, A.F. and Benz, W. (**2003**) *The Astrophysical Journal*, **589**, 556.

77 Nelson, A.F. and Benz, W. (**2003**) *The Astrophysical Journal*, **589**, 578.

78 Papaloizou, J.C.B., Nelson, R.P., Kley, W. *et al.* (**2007**) *Protostars and Planets V*, vol. 951 (eds. B. Reipurth, D. Jewitt, and K. Keil), University of Arizona Press, Tucson, pp. 655–668.

79 Klahr, H. and Kley, W. (**2006**) *Astronomy and Astrophysics*, **445**, 747.

80 Fouchet, L. and Mayer, L. (**2008**) *Icarus*, 2008arXiv0806.3975F, in press.

81 Bate, M.R., Lubow, S.H., Ogilvie, G.I., and Miller, K.A. (**2003**) *Monthly Notices of the Royal Astronomical Society*, **341**, 213.

82 Lodato, G. and Rice, W.K.M. (**2005**) *Monthly Notices of the Royal Astronomical Society*, **358**, 1489.

83 Lin, D.N.C. and Pringle, J.E. (**1987**) *Monthly Notices of the Royal Astronomical Society*, **225**, 607.

84 Shakura, N.I. and Sunyaev, R.A. (**1976**) *Monthly Notices of the Royal Astronomical Society*, **175**, 613.

85 Lufkin, G., Quinn, T.,Wadsley, J. *et al.* (**2004**) *Monthly Notices of the Royal Astronomical Society*, **347**, 421.

86 Papaloizou, J.C.B. and Terquem, C. (**2001**) *Monthly Notices of the Royal Astronomical Society*, **325**, 221.

87 Terquem, C. and Papaloizou, J.C.B. (**2002**) *Monthly Notices of the Royal Astronomical Society*, **332**, L39.

88 Guillot, T. (**2008**) *Physica Scripta*, **130**, 014023.

89 Marcy, G., Butler, R.P., Fischer, D. *et al.* (**2005**) *Progress of Theoretical Physics Supplement*, **158**, 24–42.

90 Udry, S., Mayor, M., and Santos, N.C. (**2003**) *Astronomy and Astrophysics*, **407**, 369.

91 Rasio, F.A. and Ford, E.B. (**1996**) *Science*, **274**, 954.

92 Boss, A.P. (**2006**) *The Astrophysical Journal*, **644**, L79.

93 Bonfils, X., Delfosse, X., Udry, S. *et al.* (**2006**) *Proceedings of the Tenth Anniversary of 51 Peg-b: Status of and Prospects for hot Jupiter Studies* (eds. L. Arnold, F. Bouchy, and C. Moutou), published by Frontier Group, Paris, p. 111.

94 Laughlin, G., Bodenheimer, P., and Adams, F.C. (**2004**) *The Astrophysical Journal*, **612**, L73.

95 Ida, S. and Lin, D.N.C. (**2004**) *The Astrophysical Journal*, **616**, 567.

96 Alibert, Y., Mordasini, C., Benz, W., and Winisdoerffer, C. (**2005**) *Astronomy and Astrophysics*, **434**, 343.

97 Alibert, Y., Mousis, O., Mordasini, C., and Benz, W. (**2005**) *The Astrophysical Journal*, **626**, L57.

98 Mordasini, C., Alibert, Y., Benz, W., and Naef, D. (**2009**) *Astronomy and Astrophysics*, **501**, 1162.

99 Cai, K. *et al.* (**2006**) *The Astrophysical Journal*, **636**, L149–L159.

100 Cai, K. *et al.* (**2006**) *The Astrophysical Journal*, **624**, L173, (Erratum).

101 Helled, R., Podolak, M., and Kovetz, A. (**2006**) *Icarus*, **185**, 64.

102 Helled, R., Podolak, M., and Kovetz, A. (**2008**) *Icarus*, **195**, 863.

103 Helled, R. and Schubert, G. (**2008**) *Icarus*, **198**, 156.

104 Guillot, T. (**2007**) *Annual Review of Earth and Planetary Sciences*, **33**, 493.

105 Carron, G. (**2008**) Master thesis University of Bern.

106 Desch, S.J. (**2007**) *The Astrophysical Journal*, **671**, 878.

107 Mejia, A.C., Quinn, T.R., and Mayer, L. (**2005**) Protostars and planets V, Proceedings of the Conference, October 24–28, 2005, LPI Contribution 1286, 8210.

108 Haghighipour, N. and Boss, A.P. (**2003**) *The Astrophysical Journal*, **583**, 996.

109 Haghighipour, N. and Boss, A.P. (**2003**) *The Astrophysical Journal*, **598**, 1301.

110 Vorobyov, E.I. and Basu, S. (**2009**) *The Astrophysical Journal*, **692**, 1609.

111 Banerjee, R., Pudritz, R.E., and Anderson, D.W. (**2006**) *Monthly Notices of the Royal Astronomical Society*, **373**, 1091.

112 Krumholz, M.R., Klein, R.I., and McKee, C.F. (**2007**) *The Astrophysical Journal*, **656**, 959.

113 Hueso, R. and Guillot, T. (**2005**) *Astronomy and Astrophysics*, **442**, 703.

114 Banerjee, R., Pudritz, R.E., and Holmes, L. (**2004**) *Monthly Notices of the Royal Astronomical Society*, **355**, 248.

115 Kratter, K.M., Matzner, C.D., and Krumholz, M.R. (**2008**) *The Astrophysical Journal*, **681**, 375.

116 Vorobyov, E.I. and Basu, S. (**2007**) *Monthly Notices of the Royal Astronomical Society*, **381**, 1009.

117 Price, D.J. and Bate, M.R. (**2007**) *Astrophysics and Space Science*, **311**, 75.

118 Meyer, M.R., Backman, D.E., Weinberger, A.J., and Wyatt, M.C. (**2007**) *Protostars and Planets V* (eds. B. Reipurth, D. Jewitt, and K. Keil), University Arizona Press, Tucson, p. 537.

119 Weidenschilling, S.J. (**1977**) *Astrophysics and Space Science*, **51**, 153.

120 Tsiganis, K., Gomes, R., Morbidelli, A., and Levison, H.F. (**2007**) *Nature*, **435**, 459.

121 Duquennoy, A. and Mayor, M. (**1991**) *Astronomy and Astrophysics*, **248**, 485.

122 Bonnell, I.A., Bate, M.R., and Vine, S.G. (**2003**) *Monthly Notices of the Royal Astronomical Society*, **343**, 413.

123 Bonnell, I.A., Clark, P., and Bate, M.R. (**2008**) *Monthly Notices of the Royal Astronomical Society*, **389**, 1556.

124 Boss, A.P. (**2006**) *The Astrophysical Journal*, **641**, 1148.

125 Lodato, G., Meru, F., Clarke, C., and Rice, W.K.M. (**2007**) *Monthly Notices of the Royal Astronomical Society*, **374**, 590.

126 Eggenberger, A., Udry, S., and Mayor, M. (**2004**) *Astronomy and Astrophysics*, **417**, 353.

127 Eggenberger, A., Udry, S., Chauvin, G. *et al.* (**2007**) *Astronomy and Astrophysics*, **474**, 273.

128 Kalas, P. *et al.* (**2008**) *Science*, **322**, 1345.

129 Marois, C. *et al.* (**2008**) *Science*, **322**, 1348.

130 Lafreniere, D., Marois, C., Doyon, R., and Barman, T. (**2009**) *The Astrophysical Journal*, **694**, L148.

Further Reading

Aikawa, Y. (**2003**) *Proceedings of the Conference Star Formation at High Angular Resolution*, vol. **67**, (eds. M. Burton, R. Jayawardhana, and T. Bourke), International Astronomical Union Symposium 221, XXV General Assembly of the IAU, IAU, Sydney, 22–25 July 2003, p. 2.

Boley, A.C. and Durisen, R.H. (**2008**) *The Astrophysical Journal*, **685**, 1193.

Boss, A.P. (**1993**) *The Astrophysical Journal*, **417**, 351.

Boss, A.P. and Yorke, H.W. (**1993**) *The Astrophysical Journal*, **411**, L99.

Boss, A.P. and Yorke, H.W. (**1996**) *The Astrophysical Journal*, **469**, 366.

Chauvin, G. *et al.* (**2006**) *Astronomy and Astrophysics*, **456**, 1165.

Durisen, R.H., Cai, K., Mejia, A.C., and Pickett, M.K. (**2005**) *Icarus*, **173**, 417.

Evrard, A.E. (**1990**) *The Astrophysical Journal*, **363**, 349.

Gingold, R.A. and Monaghan, J.J. (**1977**) *Monthly Notices of the Royal Astronomical Society*, **181**, 375.

Lucy, L.B. (**1977**) *The Astronomical Journal*, **82**, 1013.

Patience, J. *et al.* (**2002**) *The Astrophysical Journal*, **581**, 654.

Pickett, B.K., Durisen, R.H., and Link, R. (**1997**) *Icarus*, **126**, 243.

Pickett, B.K. *et al.* (**2003**) *The Astrophysical Journal*, **590**, 1060.

Raghavan, D. *et al.* (**2006**) *The Astrophysical Journal*, **646**, 523.

Thébault, P. *et al.* (**2004**) *Astronomy and Astrophysics*, **427**, 1097.

Tomley, L., Cassen, P. and Steiman-Cameron, T. (**1991**) *The Astrophysical Journal*, **382**, 530.

Tomley, L., Steiman-Cameron, T.Y., and Cassen, P. (**1994**) *The Astrophysical Journal*, **422**, 850.

Udry, S., Eggenberger, A., Beuzit, J.L., Lagrange, A.M., Mayor, M., and Chauvin, G. (**2004**) *The Environment and Evolution of Double and Multiple Stars* (eds. C. Allen and C. Scarfe), Proceedings of IAU Colloquium 191,. Revista Mexicana de Astronomica, 3–7 February, 2002, IAU, Merida.

5
Core-accretion Model

Olenka Hubickyj

5.1
Introduction

Fifteen years ago planetary formation models were constrained by the observational data of the planets in the solar system. Since the discovery of the companion orbiting 51 Peg [1, 2], planetary formation models need to explain the planets in the solar system *and* the over 300 confirmed–observed exoplanets *and* over 30 candidates of multiple planet systems around stars other than the Sun. We have found that these exoplanets are diverse in their characteristics. The mission of searching for Earth-sized planets, Kepler, will increase the number and diversity of planets and pose new challenges to planetary formation models.

The goal of this chapter is to review the core-accretion gas capture ((CAGC) or core accretion) model for the formation of the gas giant planets. In this model, planetesimals accrete into a solid core that eventually gravitationally accretes gas from the solar nebula. This chapter provides the basic information that an interested reader would need to acquire a robust understanding of the model in its theoretical and computational form.

Section 5.2 contains a brief historical review of the progenitors of this model and the results of and the progress made by these early investigators. Section 5.3 describes the observational constraints on this proposed model. Section 5.4 provides an overview of the core-accretion model describing the protoplanet's growth from an embryo to a gas giant planet. Section 5.5 introduces the mathematical structure of the model, the boundary conditions, and the assumptions applied in the simulation. Section 5.6 discusses the results of the simulations, focusing on the description of the work by the NASA/ARC and UCSC group. Section 5.7 presents a discussion of topics related to the core-accretion model and final remarks.

5.2
Historical Background of the Development of the Core Accretion Model

Since the time planets were distinguished from stars, people have thought about their origin. Descartes [3], Kant [4], Laplace [5], and Herschel [6] formalized the

Formation and Evolution of Exoplanets. Edited by Rory Barnes

ISBN: 978-3-527-40896-2

earliest models proposing that the planets condensed from material left over from the Sun's formation. This theory was called the *nebular theory*. An alternate model, a very early precursor to the core-accretion model, was presented by Buffon *et al.* [7]. He considered a ''building up'' formation process from solar material pulled off by a comet passing close to the Sun and then accreted into planets. However, the nebular theory remained the dominant model for planetary formation until the very early twentieth century. A rudimentary form of the core-accretion model reemerged when Chamberlin [8] wrote that Earth might have formed by accretion of cold particles that he called *planetesimals*. Moulton [9], in collaboration with Chamberlin, proposed that tidal action of a near encounter with another star released solar material that condensed into planetesimals, which then accreted by mutual gravitational attraction to form planets.

The modern core-accretion model dates from Safronov's work in the 1960s. He created an analytical formulation for the accumulation of solid particles from the protoplanetary cloud into planets. The Safronov [10] accretion model and the burgeoning capabilities of computers prompted extensive work on planet formation. Wetherill [11] was one of the first to apply Safronov's model by developing a computer code to simulate the formation of Earth. Perri and Cameron [12] examined the formation of the gas giants by considering the quantitative approach of hydrodynamic instabilities in the solar nebula in the vicinity of an already-condensed planetary core. The studies by Mizuno [13] and Mizuno *et al.* [14] incorporated the important effect of the gaseous part of the solar nebula in the formation of Jupiter and Saturn and initiated core-accretion computer simulations. This early work consisted of ''time snapshot'' calculations: a series of hydrostatic equilibrium protoplanetary models with discrete and increasing core masses surrounded by a gaseous envelope that extended to the protoplanet's *tidal* radius. This radius is the distance from the center of the planet to the point at which the protoplanet's gravitational pull equals that of the Sun. Mizuno and collaborators described their models with an important parameter derived from the giant planet simulation: the ''critical'' mass, M_{crit}. This is the planetary mass for which a static solution with a core mass greater than M_{crit} was not possible, implying that further gas accretion was a hydrodynamical process/calculation. Typically, for Jupiter and Saturn values, M_{crit} was determined to be $\approx 10 M_{\oplus}$.

The first *evolutionary* (i.e., continuous time dependent) calculation of gas giant planets based on the core accretion model was done by Bodenheimer and Pollack [15] (hereafter referred to as BP86). This calculation consisted of a stellar evolution code that was adapted to planetary conditions and a constant planetesimal accretion rate was applied. With the success of these computer simulations, Pollack *et al.* ([16], hereafter referred to as Paper 1) extended the investigation by computing the formation and evolution of the gas giant planets in which *both* the planetesimal and the gas accretion rates were calculated in a self-consistent and interactive fashion. The characterizing parameter of these simulations is M_{cross}, the mass for which the solid and gas mass is equal in the forming protoplanet. M_{cross} (replacing M_{crit}) is important because it marks the evolution time when gas runaway is initialized and large amounts of gas are accreted from the solar nebula in a very short time.

By 1995 when the companion planet to 51 Peg was reported [1, 2], the core-accretion model had become a sophisticated model with computer simulations that explained many features of the gas giant planets in our solar system. Calculations of the gas giant planets' interiors (e.g., [17, 18]) are an important component in the general investigation of gas giant planet formation. Actual observations of the giant planets in the solar system (e.g., gravitational moments) provide substantial information about the presence of a solid core and the size and the composition of the core can be extracted when structural model parameters are matched with observed values [19]. All of these different approaches to studying gas giant planets could begin to provide an integrated view of the formation process. The original sample of the four gas giants in our solar system has increased significantly and the core-accretion model is challenged to explain this varied sample of gas giant planets. Brush [20] has written a very interesting historical and comprehensive account of the evolution of planetary ideas and models which is well worth reading.

5.3 Observational Constraints

Observations are ''answers in the back of the book'' for any theoretical model. In the case of the gas giant planets, a successful theoretical model needs to explain these observational characteristics: (i) the observed bulk composition characteristics of Jupiter, Saturn, Uranus and Neptune; (ii) the quick formation of giant planets; (iii) the wide range of eccentricities and semimajor axes exhibited by the extrasolar planets.

Observations of the gas giants in the solar system indicate (i) the similarity of the total heavy element contents of the four giants (about $25M_{\oplus}$), (ii) the very massive H_2 and He envelopes of Jupiter and Saturn (about $300M_{\oplus}$ and $70M_{\oplus}$, respectively) and much less massive gaseous envelopes of Uranus and Neptune (about $2–4M_{\oplus}$); and (iii) the higher heavy element content in the atmospheres of the four giant planets compared to that in our solar system [21–24]. Interior models of Jupiter [18] indicate that the total solid mass ranges from 8 to $39M_{\oplus}$, of which $0–11M_{\oplus}$ is concentrated in the core. Saturn models indicate a total heavy element mass of $13–28M_{\oplus}$, with a core mass between $9–22M_{\oplus}$. Uranus and Neptune models indicate heavy element masses ranging from 10 to $15M_{\oplus}$ and a gaseous mass between 2 and $4M_{\oplus}$ [21].

The giant planets need to form quickly, because observations of dust disks around young stellar objects indicate disk ages of <10 megayears [25–29]. Hillenbrand [30] summarized current observations of protoplanetary dust and gas disks. Observations of disks believed to be <3 megayears show that there is a wide dispersion in spectral energy distributions, indicating a variety of sizes for the dust particles in the disk. This dispersion is significantly reduced for disks $\sim$5 megayears old, indicating that most of the dust mass is agglomerated into large objects. For $\sim$10-megayear-old disks, the spectral energy distribution shows no infrared excess

implying that the dust is gone. The consensus of these observations indicates that the gas giant planets should form within 5 megayears.

The exoplanets exhibit a range of diverse characteristics. (The values quoted are taken from the compiled data reported in Jean Schneider's website: The Extrasolar Planet Encyclopedia: exoplanet.eu.) Observed masses range from 0.02 to $\sim 10 M_{\rm Jup}$. Orbital distances from parent stars span from $\sim$0.02 to 5.8 AU, but most are in the 3–4 AU. Eccentricities vary widely: 0.008–0.95. The range of periods of the orbits around the parent star is from 0.85 days to 12 years. A few cases are long-period, low-eccentricity planets whose orbits are comparable to that of Jupiter. However, for the most part, the exoplanets do not resemble Jupiter.

5.4 General Description of the Core-accretion Model

The core-accretion model proposes that gas giants are formed in two stages: a coagulation of solid planetesimals into a massive solid core followed by the gravitational capture of a gaseous envelope. The finer details of the accretion scenario occurs in the following sequence [31]:

1) Dust particles in the solar nebula form planetesimals [10] that accrete into a solid core surrounded by a low-mass gaseous envelope. Initially, solid runaway accretion occurs, and the gas accretion rate is much slower than that of solids. As the solid material in the feeding zone is depleted, the solid accretion rate is reduced.
2) The gas accretion rate steadily increases and eventually exceeds the solid accretion rate. The protoplanet continues to grow as the gas accretes at a relatively constant rate. The mass of the solid core also increases, but at a slower rate and, eventually, the core and envelope masses become equal ([12, 13], Paper 1).
3) Runaway gas accretion occurs and the protoplanet grows rapidly. The evolution (1) through (3) is referred to as the *nebular* stage, because the outer boundary of the protoplanetary envelope is in contact with the solar nebula and the density and temperature at this interface are the nebular values.
4) The gas accretion rate reaches a limiting value defined by the rate at which the nebula can transport gas to the vicinity of the planet. After this point, the equilibrium region of the protoplanet contracts inside the effective accretion radius, and gas accretes hydrodynamically onto this equilibrium region. This part of the evolution is considered to be the *transition* stage.
5) Accretion is stopped by either the opening of a gap in the disk as a consequence of accretion and the tidal effect of the planet [32–34], or by dissipation of the nebula. Once accretion stops, the planet enters the *isolation* stage.
6) The planet contracts and cools to the present state at constant mass (BP86, [35, 36]).

The core-accretion model is successful in explaining many of the observations in our solar system and the extrasolar planets. The bulk composition and the higher heavy element content in the envelopes of gas giant planets in the solar system are a direct consequence of the CAGC process. Early accretion creates a solid core that can acquire a gaseous envelope which is eventually massive enough that the incoming planetesimals either break up or ablate in the envelope, thereby depositing solids away from the core (Paper 1). The model is also successful in explaining the possibility of *in situ* formation of extrasolar planets [31], and, forming Jupiter with a low-mass core within reasonable formation times ([37], hereafter referred to as HBL05). The low-mass core requirement is called for by the interior models of Jupiter and Saturn [18, 38]. In the case of the extrasolar planets, observations have determined that there is a positive correlation between planet frequency and the metallicity of the host star [39–44], which is consistent with the CAGC model [45, 46].

5.5 The Model and the Computer Code

An important aspect of modern theoretical research is the computer code. Since Safronov's introduction of his planetesimal accretion model for terrestrial planets and Mizuno's extension of this to the formation of Jupiter and Saturn, theoreticians have made substantial progress in understanding planet formation in the last few decades. There are three different computer simulation methods. The first is exemplified by the research of Ikoma *et al.* [47], Inaba *et al.* [48], Inaba and Ikoma [49], and Ida and Lin [45]. The core-accretion calculations are accomplished by using semianalytical expressions for the rate of accretion and for protoplanetary models with discrete core masses. The second approach is a fully computational code as in BP86 and Paper 1. The third approach is a fully time-dependent radiation-hydrodynamical computer code. Wuchterl [50, 51] was the first to apply this approach to the core-accretion model. Since then other investigators [52, 53] have used a hydrodynamical approach.

The computer code structure described in this chapter is based on the that used by the group of collaborators from NASA Ames Research Center and Lick Observatory at the University of California at Santa Cruz and is referred to as the ARC/UCSC code. The governing equations, the boundary conditions, and the assumptions of the ARC/UCSC code are shared by most of the (fully hydrodynamic and hydrostatic) codes investigating the core-accretion model. The overall planetary formation model lends itself to a scaled down stellar formation approach, replacing the stellar core nuclear energy source with the mass accretion as the main energy source of the planet. The ARC/UCSC code consists of three components: the calculation of the rate of solid planetesimal accretion onto the planet, a computation of the planetesimal interaction with the evolving envelope, and the calculation of the gas accretion rate and the thermal evolution of the protoplanet. Detailed descriptions of these components are provided in the following subsections.

5.5.1 The Code Components

The first code component computes the rate of solids accreted onto the protoplanet. The rate is an updated version of the classical theory of planetary growth [10] and is given as

$$\frac{dM_p}{dt} = \pi R_c^2 \sigma \Omega F_g \tag{5.1}$$

where M_p is the mass of the protoplanet, R_c is the capture radius, defined as the radius in the envelope of the protoplanet where the planetesimals are gravitationally caught from the feeding zone, σ is the surface mass density of planetesimals, Ω is the orbital frequency of the protoplanet, and F_G is the ''gravitational enhancement factor''. The gravitational enhancement factor is the ratio of the total effective accretion cross section to the geometric cross section. F_G is an analytical expression that was derived to fit the data from the numerical calculations of Greenzweig and Lissauer [54], consisting of a large number of three-body (Sun, protoplanet, and planetesimal) orbital interaction simulations. The planet's accretion (feeding) zone is assumed to be an annulus centered at the protoplanet's orbital distance, a, from the Sun and it extends to a radial distance, a_f, on either side of its orbit. According to the simulations of Greenzweig and Lissauer [55] and Kary and Lissauer [56], a_f is well approximated by the following:

$$a_f = \sqrt{12 + e_H^2}\ R_H, \quad R_H = a\left(\frac{M_p}{3M_\odot}\right)^{1/3}, \quad e_H = \frac{a}{R_H}\ e \tag{5.2}$$

where R_H is the tidal radius, also called the *Hill-sphere radius*, e_H is the rms value of the planetesimals' orbital eccentricities in units of Hill sphere units, and $M_\odot$ is the mass of the Sun.

The accretion zone grows as the planet gains mass, independent of whether the accreted mass is a solid or gas. The mass of planetesimals in the accretion zone is assumed to equal the initial mass of planetesimals in the current accretion zone minus the amount that has already been accreted by the protoplanet. Radial migration of planetesimals into and out of the accretion zone is neglected. Random scatterings are assumed to spread the unaccreted planetesimals uniformly over its accretion zone.

The second component of the evolution code is the calculation of the interaction of the accreted planetesimals with the gas in the envelope. This is a two-body trajectory calculation (the planetesimal and the protoplanet) and it determines whether the planetesimals reach the core, are dissolved in the envelope, or it is a combination of the two [57]. The trajectories of the incoming planetesimals are computed by integrating the equations of motion for particles under the influence of gravitational acceleration due to the protoplanet (core and envelope masses), the gas-drag force, the effects of radiative and ablative heating (as the planetesimal traverses the envelope), and fragmentation by gas dynamical pressure. These equations are integrated using a fourth order Runge–Kutta scheme. In addition, these trajectory calculations provide the radius in the envelope at which the planetesimal is

captured, R_c, (required to compute the accretion rate of the planetesimals), and the energy deposited by the planetesimal in the envelope (required for the structure computation, the third component). The formulae and specific constants of this calculation cannot be easily summarized in this chapter. The important physical concepts are noted here so that the reader can have a better understanding of the code. However, for a clear and thorough discussion, the reader is referred to [58] and [57].

The final component is the calculation of the gas accretion rate and thermal evolution of the protoplanet under the assumption that the planet is spherical and that the standard equations of stellar structure apply. The conventional stellar structure equations of conservation of mass and energy, hydrostatic equilibrium, and radiative or convective energy transport are used.[1] The energy generation rate is the result of the accretion of planetesimals and the quasi-static contraction of the envelope. The envelope is evolved with the standard stellar structure equations in spherical symmetry (BP86): the equation of mass conservation:

$$\frac{\partial M_r}{\partial t} = 4\pi r^2 \rho \tag{5.3}$$

the equation of hydrostatic equilibrium:

$$\frac{1}{\rho}\frac{\partial P}{\partial r} = -\frac{GM_r}{r^2} \tag{5.4}$$

the equation of energy conservation:

$$\frac{\partial L_r}{\partial t} = 4\pi r^2 \rho \left(\epsilon - P\frac{\partial V}{\partial t} - \frac{\partial E}{\partial r} \right) \tag{5.5}$$

the equation for radiative transfer:

$$\frac{\partial T}{\partial r} = \frac{3}{4ac}\frac{\kappa\rho}{T^3}\frac{L_r}{4\pi r^2} \tag{5.6}$$

where M_r is the mass interior to radius r, P is the pressure, L_r is the energy crossing a sphere at radius r, T is the temperature, ρ is the density, and $V = 1/\rho$. The energy equation includes three sources: ϵ, the heat generated by the captured planetesimals, $P\partial V$, work from compression by gravitational forces, and, E, cooling from the release of internal heat per mass. In regions of convective instability, where the radiative temperature gradient exceeds the adiabatic gradient, the temperature gradient is taken to be the adiabatic gradient.

1) Wuchterl's [50, 51]) hydrodynamical models showed that once the envelope mass became comparable to the core mass, a dynamical instability develops that results in the ejection of much of the envelope. This did not occur for models computed in disk with a high solid surface density. However, the assumption of hydrostatic equilibrium was numerically substantiated by Tajima and Nakagawa [59] who performed an instability analysis on the models reported in BP86.

5.5.2 The Code Boundary Conditions

The inner and outer boundary conditions are set at the bottom and top of the envelope. The core is assumed to have a uniform density and to be composed of a combination of ice, CHON (Carbon, Hydrogen, Oxygen, Nitrogen), and rock. At the inner edge of the envelope, the luminosity is zero, $L_r = 0$, the mass equals the core mass, $M_r = M_{\text{core}}$, and the radius equals the core radius, $r = R_{\text{core}}$. The outer boundary condition of the protoplanet is applied differently depending on the stage of evolution (as described in Section 5.4). For the first stage of evolution, the nebular stage, the envelope is in direct contact with the protoplanetary nebula. The density and temperature of the outer edge of the envelope are ρ_{neb} and T_{neb}, respectively, are constant, and are prescribed by nebula evolution models (e.g., Bell *et al.* [60], D'Angelo *et al.* [61]). In the nebular stage, the outer radius of the planet is assumed to fall at a modified accretion radius, R_a, which is given by

$$R_a = \frac{GM_p}{c^2 + \frac{GM_p}{R_H}} \tag{5.7}$$

where c is the sound speed in the nebula. In the limits of large and small R_H, this expression reduces to the accretion radius, (GM_p/c^2), and the tidal radius, (R_H), respectively. The gas accretion rate is determined by the requirement that the outer radius of the protoplanet be close to R_a, within a small tolerance (set to $\leq$10% in the computer code calculations). At every timestep, mass is added at the outer edge so that this requirement is satisfied.

As gas runaway occurs, the gas accretion rate eventually reaches a limiting value determined by the ability of the nebula to supply gas at the required rate. Once the limiting rate is reached, the planet contracts inside R_a, and the evolution enters the transition stage. The planet is still assumed to be in hydrostatic equilibrium, but gas is accreting hydrodynamically onto it at near free-fall velocities. The luminosity of the protoplanet consists of two contributions, the normal luminosity (L_{contr}) caused by contraction and accretion of solids and a gas accretion luminosity given by

$$L_{\text{acc}} = GM_p\dot{M}_{XY,\max}(1/R_p - 1/R_a) \tag{5.8}$$

where R_p is the planetary radius and $\dot{M}_{XY,\max}$ is the limiting accretion rate.

The density and temperature at R_p are also modified by considering accretion shock conditions. The density at the inner edge of the infalling flow, ρ_0, is obtained from

$$\rho_0 = \frac{\dot{M}_{XY,\max}}{4\pi R_p^2 v_{\text{ff}}} \tag{5.9}$$

v_{ff} is the free-fall velocity from R_a to R_p. The density, ρ_0, is multiplied by the square of the infall Mach number to get the planetary boundary density ρ_s, under the assumption that the boundary shock is isothermal. To get the temperature, the radiative diffusion equation

$$\frac{dT^4}{dr} = -\frac{3}{4\sigma}\frac{\kappa\rho_0 R_p^{1.5}}{r^{1.5}}\frac{L_r}{4\pi r^2} \tag{5.10}$$

is integrated and solved approximately under the assumptions that the Rosseland mean opacity κ and the total luminosity $L_r = L_{\rm acc} + L_{\rm contr}$ are constants as a function of distance r. In the limit that the envelope is optically thin the result is simply that the boundary shock temperature is $T_s = T_{\rm neb}$. These approximations have been checked against more detailed solutions of the radiative transfer equation and found to be quite adequate.

The supply of gas to the planet is eventually assumed to be exhausted (e.g., as a result of tidal truncation of the nebula, removal of the gas by effects of the star, and/or the accretion of all nearby gas by the planet). The planet's mass levels off and the planet enters the final stage of its formation, the isolation stage. The boundary conditions are the standard photospheric conditions

$$L = 4\pi\sigma R_p^2 T_s^4, \kappa P = \frac{2}{3}g \tag{5.11}$$

where P is the photospheric pressure and g is the acceleration of gravity at R_p. If the planet is close to the central star, the solar radiation energy received on the surface of the planet, (insolation), is taken into account according to the procedure described by Stringfellow [62].

It should be noted here that until recently, the gas accretion rate was modeled to reduce smoothly to zero as the predetermined limiting planetary mass value (defined by the planetary system that is being modeled) was approached. Previously, the limiting mass transfer rate in the calculations was set to $3 \times 10^{-8} M_\odot$ per year or $\approx 10^{-2} M_\oplus$ per year, which is due to viscous effects in a typical nebula. Lissauer [63] has replaced this approach with a more realistic gap clearing calculation, which took into account the planet's interactions with the protoplanetary disk which were calculated using 3D hydrodynamic simulations [61].

5.5.3 The Code Assumptions

The following assumptions were applied to the calculation of a forming planet consisting of a rock and silicate solid core of constant density, 3.2 g cm^{-3}, and an envelope with a protosolar composition of $X = 0.74$ (hydrogen), $Y = 0.243$ (helium), and $Z = 0.017$ (everything else). The core is an inert component in the calculation and as long as the core is at least $0.1 M_\oplus$, it is massive enough to gravitationally bind the accreted nebula gas and this envelope is optically thick enough for it to be in hydrostatic equilibrium.

The assumptions applied and implied in the code structure are as follows:

1) The growing protoplanet is a lone embryo surrounded by a disk with an initially uniform surface density in the region of the protoplanet consisting of planetesimals with uniform mass and radius.
2) The protoplanet's feeding zone is assumed to be an annulus extending to a radial distance of about 4 Hill-sphere radii on either side of its orbit [56], which grows as the planet gains mass. Planetesimals are spread uniformly over the zone and do not migrate into or out of the feeding zone.

3) The equation of state is nonideal and the tables used are based on the calculations of Saumon *et al.* [64], interpolated to a near-protosolar composition of $X = 0.74$, $Y = 0.243$, $Z = 0.017$. The opacity tables are derived from the calculations of Pollack *et al.* [65] and Alexander and Ferguson [66].
4) The planetesimal is considered to be captured by the protoplanet if the planetesimal deposits 50% or more of its mass into the envelope during its trajectory.
5) Once the mass and energy profiles in the envelope have been determined, the planetesimals are assumed to sink to the core, liberating additional energy in the process.
6) The rate of planetesimal accretion near gas runaway is limited to its value at crossover in order to account for the depletion of the planetesimal disk by accretion onto neighboring embryos.

5.6 Results

It is fitting that as Jupiter was the first planet that was observed, which ushered in the observational era in astronomy, it was a Jupiter-like planet that has heralded the age of planet discovery. By observationally and theoretically studying the gas giants in the solar system, planetary scientists were provided with a basis from which they could model the extrasolar planets. An overview of the general results from the ARC/UCSC simulations is presented here.

The core-accretion simulations are characterized by three major phases illustrated in Figure 5.1. The interfaces of the phases are denoted by dashed, vertical lines. In the first phase, the planet's mass consists primarily of solid material. The planetesimal accretion rate, which dominates that of gas, rapidly increases owing to runaway accretion, and then decreases as the planet's feeding zone is depleted of solids. During the second phase, both solid and gas accretion rates are small and nearly independent of time. The overall evolutionary timescale is generally determined by the length of the second phase. The nature of phase 2 is such that its length depends on either the accretion of the planetesimals timescale (the dominant energy source at the start of phase 2) or the gravitational contraction (Kelvin–Helmholtz) timescale of the envelope (dominant at the end of phase 2 of the evolution, just before the start of phase 3). The third phase, marked by runaway gas accretion, starts when the solid and gas masses are equal, $M_{\rm cross}$. The high rate of gas accretion is a result of a strong positive feedback driven by the rapid contraction of the gaseous envelope. The contraction triggers the gas inflow and, in turn, the outer boundary expands because the planetary mass is increasing. Runaway occurs very quickly and most of the gaseous mass of the planet is accreted during this time (e.g., see Figure 5.1a).

A short formation time is a crucial hallmark of a successful planetary formation model. Although the timescale problem related to core-accretion model is not as dire as it once was ($>$30 megayears [13]), $<$10 megayears (Paper 1), $\sim$3 megayears

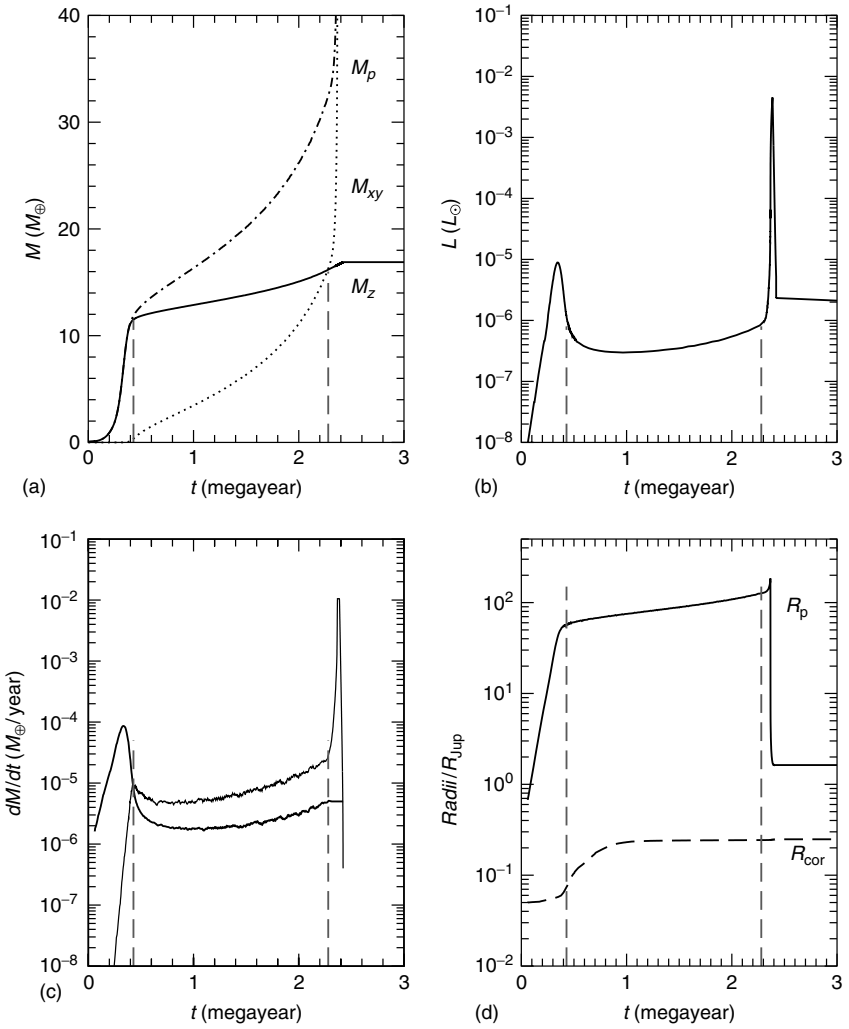


Figure 5.1 Evolution of the baseline case (HBL05) with $\sigma_{\mathrm{init},Z} = 10$ g cm^{-2} and grain opacity at 2% of the interstellar medium value (ISM). The gray dashed, vertical lines mark the interfaces of the phases of evolution. (a) The mass is plotted as a function of time. The solid line denotes the solid core mass, M_Z, the dotted line denotes the envelope mass, M_{XY}, and the dash-dotted line denotes the total mass of the protoplanet, M_p. The value of M_p levels off at 1 M_J after 2.3 megayears, but the plot only displays masses up to the limit of $M_p = 40 M_\oplus$ in order to better display the crucial stages. (b) The luminosity (in units of $L_\odot$) is plotted on a logarithmic scale as a function of time. (c) The solid line denotes the core-accretion rate, $\dot{M}_Z$. The dotted line denotes the envelope accretion rate, $\dot{M}_{XY}$. (d) The radii plotted on a logarithmic scale as a function of time.

(HBL05), other problems related to the CAGC model remain. These concerns can be discussed within the context of the following parameters that affect the timescales: (i) the opacity due to grains in the envelope; (ii) the mass of the core; (iii) the initial solid surface density of the solar nebula (i.e., minimum mass solar nebula (MMSN) or not MMSN); (iv) the migration of the protoplanet.

5.6.1 Opacity

The opacity due to grains (hereafter referred to as "grain opacity") in the protoplanetary envelope has a major effect on the formation timescale but has no effect on the core mass. Opacity controls the rate at which the planet can radiate energy, which, in turn, dictates the rate at which the protoplanet's envelope shrinks and allows more gas to enter its gravitational domain. The models of Jupiter in Paper 1 were computed with the grain opacity of the planet's upper atmosphere assumed to be that of the interstellar medium (ISM) and then randomly reduced to 2% of ISM, resulting in formation times of ∼8 and ∼3 megayears, respectively. It was noted in Paper 1 and [47] and demonstrated with calculations in HBL05 that a substantial reduction in the grain opacity to 2% ISM significantly reduces the formation timescale. The effect of the change in grain opacity is illustrated in Figure 5.2 with the mass and luminosity plotted as a function of time.

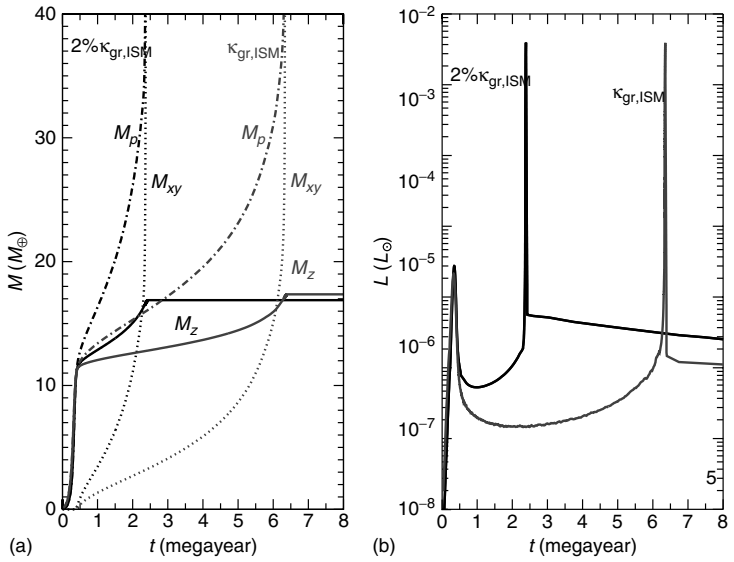


Figure 5.2 The effect of grain opacity on the baseline Jupiter model (HBL05) with $\sigma_{\mathrm{init},Z} = 10$ g cm^{-2}. (a) The mass is plotted as a function of time. The line style denotation is the same as in Figure 5.1. (b) The luminosity is plotted as a function of time. The black curves show the mass and luminosity for models with 2% ISM grain opacity; the gray curves show the mass and luminosity for models with full ISM grain opacity.

In all these investigations, the opacity has been treated as a free parameter. Recently, Podolak [67] and Movshovitz and Podolak [68] have developed a detailed numerical model for the growth and sedimentation of grains in a protoplanetary atmosphere, coupled with a procedure for calculating the opacity at each depth. Their simulations indicate grain opacity values lower than ISM values throughout the atmosphere. They report that the probable reduction of grain opacities is even greater, as compared to interstellar values, than the factor 50 assumed in Paper 1, [47], and HBL05. Their most recent results show that opacities [68] do not scale linearly with ISM values, with some regions of the atmosphere having opacity well above 2% ISM and others well below this value. For models computed with the grain opacity below interstellar values, formation times are short and the final core mass is unaffected. The baseline case computed in HBL05 shows that Jupiter can form at 5 AU in just over 2 megayears, and the core mass is $16M_{\oplus}$. The interior models for Jupiter and Saturn calculated by Guillot [69] indicate that the gas giants have small solid cores (Section 5.3). Ikoma *et al.* [47] and then Hubickyj *et al.* (HBL05) showed that it is not possible for a gas giant to form with a small solid core on a short timescale for models computed with the grain opacity equal to that for typical interstellar material. Decreasing the grain opacity affects the formation timescale but the final core mass is unaffected.

5.6.2 Core Mass

The effect on the formation timescale due to the core mass is considered here from two points of view: (i) the mass of the core as considered in the interior models of Jupiter and Saturn (Guillot [69]; see also Section 5.3) and (ii) the model for the rate of accretion of the solids. With respect to the first point, the wide range of possible core masses for Jupiter and Saturn is a consequence of uncertainties in the equation of state of a mixture of hydrogen and helium at megabar pressures. At present, the preferred core mass for Jupiter is less than $3M_{\oplus}$. Recently, Militzer *et al.* [70] presented results of computer simulations of hydrogen–helium mixtures at the conditions in Jupiter's interior. The result implies that Jupiter has a $14–18M_{\oplus}$ central mass. Even with these ranges of core mass, it is clear that there is a substantial increase in overall heavy element content in all four planets; what is not clear is the division between core and envelope, especially in Jupiter.

How and under what conditions will the core-accretion model produce a gas giant planet (e.g., Jupiter) within 5 megayears, maybe 3 megayears, with a core of $3M_{\oplus}$? In HBL05 this question is investigated by examining the various mass values to determine whether or not gas runaway can occur for small mass cores on a reasonable timescale. Figure 5.3 shows the mass and luminosity as a function of time for models with varying core masses. Halting planetesimal accretion at low core mass simulates the presence of a competing embryo. The size of the core mass at solid accretion cutoff will change the formation timescale. For the Jupiter simulation with $\sigma_{\mathrm{init},Z} = 10\ \mathrm{g\ cm^{-2}}$ and the grain opacity uniformly reduced to 2% ISM value, the formation time is 2.3 megayears, about a third that is required for

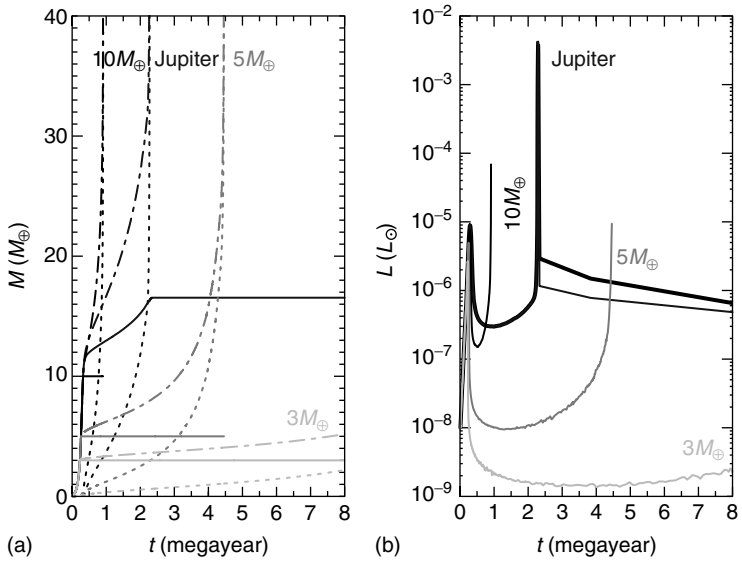


Figure 5.3 The effect of the core mass on the baseline model (HBL05) labeled "Jupiter", with $\sigma_{\mathrm{init},Z} = 10$ g cm^{-2} and grain opacity reduced to 2% ISM. (a) The mass is plotted as a function of time. The line style denotation is the same as in Figure 5.1. (b) The luminosity is plotted as a function of time. The curves are labeled with the mass of the core mass, the curves labeled "Jupiter" have a mass of $16M_\oplus$.

the case with 100% ISM grain opacity (see Figure 5.2). The resulting core mass is $16M_\oplus$. If the solid accretion rate is stopped at a protoplanetary core mass of $10M_\oplus$ or $5M_\oplus$, the formation time becomes 1 megayear and 4.5 megayears, respectively. These timescales fit well with observations of dust disks around young stellar objects (see Section 5.3). The low core masses ($10M_\oplus$ and $5M_\oplus$) agree with models for present day Jupiter [71, 69]. The core mass of $16M_\oplus$ fits with the more recent models of Militzer *et al.* [70].

The second point of view in considering the core mass was noted in the list of assumptions: the core-accretion code uses an oversimplified model to determine rates of solids accretion and the calculations start with a Mars-sized core, $0.1M_\oplus$. The overall results of the core-accretion code indicate that it is phase 2 that determines the final properties of the planet. However, results from recent work indicate that for low atmospheric opacity and/or a cutoff in accretion of solids, phase 2 can be relatively short. Thus, the time for phase 1, the solid core-accretion phase, may be the determining factor for the timescale for the formation of a giant planet.

In the work of the ARC/UCSC group, and in the code used by Alibert *et al.* [72], it is assumed that the protoplanet is an isolated embryo within a "sea" of much smaller planetesimals. The accretion rate of these planetesimals onto the protoplanet is estimated using the formulae given by Greenzweig and Lissauer [54]. This method represents an advance over the constant solids accretion rate employed in BP86.

Simulations of the core-accretion process that take into account multiple embryos and a number of other physical processes (e.g., the interaction of the accreted planetesimal and atmosphere of the protoplanet) [48, 73, 74], indicate that the core formation times are longer than those computed by the ARC/UCSC studies. Fortier *et al.* [75] motivated by Thommes [73], Inaba and Ikoma [49], and Ida and Makino [76], incorporated solid core accretion into their accretion code. They adopted Thommes' oligarchic growth model but modified it to include gas drag of the planetesimal's encounter with the envelope. Their results require a solar nebula density many times larger than MMSN (6–10 MMSN) Thus, the question remains as to how large an enhancement of solid surface density, as compared with that in the minimum mass solar nebula, is needed to form a giant planet in a few megayears.

5.6.3
Should MMSN be Used

What is the surface density value of the MMSN? Why should we use MMSN? What is the typical surface density of observed protoplanetary disks? These questions are always present. The earlier calculated timescales of the core-accretion model for the formation time for Jupiter models with an MMSN were >10 megayears ([13], Paper 1). Lissauer [77] showed that in order to get reasonable core formation times this quantity must be enhanced over that in the MMSN by a factor of several. The numerical simulations (Paper 1, [47]) indicate that the factor is probably in the range 3–4, although more recent detailed calculations of the core-accretion process [48] indicate that the required factor is larger, in the range 4–8. The standard value assumed here, for a proto-Jupiter at 5 AU, is 10 g cm^{-2}, which is about three times that in the MMSN. It should be noted that according to the minimum mass solar nebula model of Weidenschilling [78], the surface density for solids in the solar nebula can vary between 1.7 and 34 g cm^{-2} at Jupiter's position. Thus the value of 10 g cm^{-2} used in the simulations is reasonable. Figure 5.4 is a plot of the mass and luminosity of models for two different values of the surface density for solids. By increasing the initial solid surface density in the disk to about two times that of the MMSN, it is still possible to form Jupiter in less than 5 megayears if the core accretion is cutoff at $5M_{\oplus}$.

5.7
Discussion

The observations of exoplanets and the theoretical sophistication of the computer simulations have provided a wealth of knowledge to enable researchers to move ahead. With this knowledge and ability to theoretically probe aspects other than timescale issues, we can consider other outstanding topics: migration, the metallicity correlation, and applying these models to predict observational characteristics.

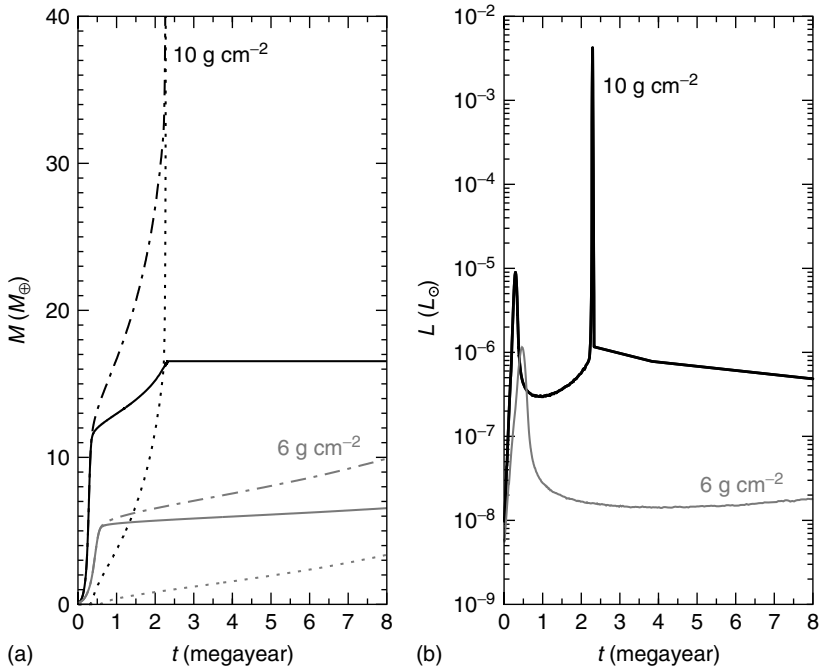


Figure 5.4 The effect of the surface density of solids for the baseline Jupiter model with the grain opacity of 2% ISM. (a) The mass is plotted as a function of time. The line style denotation is the same as in Figure 5.1. (b) The luminosity is plotted as a function of time. The curves are labeled with the solid surface density. The black curves denote $\sigma_{\mathrm{init},Z} = 10$ g cm^{-2} and the gray curves denote $\sigma_{\mathrm{init},Z} = 6$ g cm^{-2}.

5.7.1 Migration

A persistent topic in giant planet formation is the effect of migration induced by the protoplanetary disk. Migration occurs when there is a dynamical tidal interaction of the growing protoplanet with the disk, which leads to two phenomena: migration and gap formation ([32, 79, 80]; see also Chapter 9). For low-mass planets, the tidal interaction is a linear function of mass and the migration is type I: inward migration with no gap opening. Higher mass planets open a gap, leading to a reduction of the inward migration, and this is referred to as *type II migration*. Observations of extrasolar giant planets have reinforced the importance of considering migration during planet formation, but migration rates of planets with a mass smaller than a few tens of Earth masses are rather uncertain. Since Chapter 9 is devoted to this topic, migration of forming protoplanets, as it pertains to the CAGC model, is briefly mentioned here.

Migration is more than likely a viable aspect of gas giant planet formation in explaining the wide range of eccentricities and semimajor axes deduced from the observations of extrasolar planets. Alibert *et al.* [72] examine the effect of migration

in the core accretion scenario. Their code follows the structure outlined in this chapter. The migration part of their code sets the rate for type I migration as a free parameter, whereby the starting location of the embryo is adjusted for each choice of the migration rate, in order for the protoplanet to reach the crossover mass at 5.5 AU. The profile of the initial disk surface density, Σ, is the same for each of the migration rates, namely a power law $\Sigma \propto r^{-2}$. This profile corresponds to the case in Paper 1 (case J2) that was computed with $\sigma_{\mathrm{init},Z} = 7.5$ g cm^{-2}, the solid surface density that is about twice the MMSN. The disk structure and its evolution are based on the method of Papaloizou and Terquem [81] which is in the framework of the α-disk formulation of Shakura and Sunyaev [82]. The models yield satisfactory results but more importantly, their work showed the importance of migration to the planet's accretion of the planet's solids. However, the specific results of these simulations are a result of an arbitrary choice of the migration rate and should not be interpreted as a real result.

The good news is that migration prevents the depletion of the feeding zone that occurs in *in situ* formation. With an everpresent supply of matter to accrete, the protoplanet's formation timescale is reduced from that of *in situ* formation and it is not necessary to consider massive disks as proposed by Lissauer [77]. The bad news is that the migration results in a timescale for the traveling cores to be shorter than the time necessary for the protoplanet to accrete to higher masses. Further research is necessary to resolve these timescale problems related to planetary migration models.

5.7.2 Metallicity

The correlation of stars with high metallic content and the presence of planets was first reported by Gonzalez [83] and further explored by Fischer and Valenti [42, 43].

While the metallicity correlation has been found to be consistent with simplified CAGC models [45, 46], it has also been suggested [84] that the correlation is a result of preferential migration of planets in high-metallicity disks into the period range where they are observed. Sozzetti [85] suggests that there is a correlation between observed orbital period and the host star metallicity in the sense that the higher metallicity stars are more likely to have short-period planets. This tendency would be consistent with the migration scenario; however, the correlation is weak, and Santos *et al*. [86] do not find any correlation. On the theoretical side, the simple model of Livio and Pringle [87] results in only a small difference in migration rates in metal-rich and metal-poor disks, which is not sufficient to explain the trend reported in [42]. Thus, this correlation is more likely to be a result of the formation mechanism itself. Although a higher metallicity planet has a higher opacity in the envelope, which results in longer formation times, increases in opacity due to grains by only a factor of 2 (the typical range of stellar metallicities compared with solar) have a very small effect on the time; factors of 50 or more in grain opacity changes are required to make significant differences in formation times.

The core-accretion model does provide results that are in good agreement with the observed correlation between planet frequency and stellar metallicity [39, 41, 43, 45, 46, 88], and it explains many aspects of the giant planets in the solar system.

5.7.3 Observational Predictions for New Planets

At this stage of the development of the CAGC model and with the extensive inventory of extrasolar planets, the core-accretion model may be able to provide observational directives for planet hunters. Adaptive optics, coronagraphs, and other methods (e.g., [89]) hold the promise of directly detecting Jupiter mass companions to young, relatively nearby stars (e.g., Formalhaut b). Until recently, the theoretical basis for understanding substellar or planetary companions at very young ages was lacking. Two groups [36, 88] have done evolution calculations without modeling the actual accretion of planetary mass objects and their calculations are based on somewhat arbitrary initial conditions. The luminosities of young objects are very sensitive to initial accretion rates and the arbitrary initial structure and radius are thus highly suspect [91, 92].

Two important uncertainties limit the ability to predict the luminosity as a function of time for young planets. First, the formation time (particularly, $t = 0$) is highly uncertain, because of uncertainties in the accretion timescales arising from the treatment of dust opacity in the atmosphere. This affects measurements of the age of a planet and its initial effective temperature. For example, if a $T_{eff} = 500$ K companion is observed orbiting a 10-megayear star, the question is how long has it been since it stopped accreting? The answer will depend on whether the companion is a 1, 3, or 9 Jupiter mass object. Secondly, the initial temperature of the atmosphere and internal entropy of the young planet are highly model dependent. The arbitrary initial conditions of planets used in ''standard" evolution models look nothing like planets produced by the CAGC calculations. *Combined, these two issues translate into highly uncertain luminosities for model planets at young ages.* Predicting luminosities of young planets would be more reliable from a self-consistent evolution model such as the CAGC model.

The core-accretion model is successful in explaining many features of the planets in this solar system and in those around other stars. As new planets are discovered the core-accretion model will be challenged to explain the diversity of the observations.

References

1 Mayor, M. and Queloz, D. (**1995**) A Jupiter-mass companion to a solar-type star. *Nature*, **378**, 355–359.

2 Marcy, G.W. and Butler, R.P. (**1995**) 51 Pegasi. *IAU Circulars*, **6251**, 1.

3 Descartes, R. (**1644**) Principia Philosophiae (in Latin), Elzevir, Amsterdam.

4 Kant, I. (**1755**) Allgemeine Naturgeschichte und Theorie des

Himmels, Johann Friederich Petersen, Koenigsberg and Leipzig.
5 Laplace, P.S. (**1796**) Exposition de Système du Monde, Circle-Sociale, Paris, English translation: Harte, H.H. (1830) The System of the World, University Press, Dublin.
6 Herschel, W. (**1811**) Astronomical observations relating to the construction of the heavens, arranged for the purpose of a critical examination, the result of which appears to throw some new light upon the organization of the celestial bodies. *Philosophical Transactions of the Royal Society of London*, **101**, 269.
7 Buffon, G.-L. and Leclerc, Comte de (**1749**) Histoire Naturelle, Générale et Particulière, avec la Description du Cabinet du Roi, Imprimerie Royal, Paris.
8 Chamberlin, T.C. (**1899**) On Lord Kelvin's address on the age of the Earth as an abode to life. *Science*, **9–10**, 11,889.
9 Moulton, F.R. (**1905**) On the evolution of the solar system. *The Astrophysical Journal*, **22**, 165.
10 Safronov, V.S. (**1969**) Evolution of the Protoplanetary Cloud and Formation of the Earth and Planets In Russian, Nauka Press, Moscow, English translation: NASA–TTF–677, 1972.
11 Wetherill, G.W. (**1977**) Accretion of the terrestrial planets. *Meteoritics*, **12**, 387.
12 Perri, F. and Cameron, A.G.W. (**1974**) Hydrodynamic instability of the solar nebula in the presence of a planetary core. *Icarus*, **22**, 416–425.
13 Mizuno, H. (**1980**) Formation of the giant planets. *Progress in Theoretical Physics*, **64**, 544–557.
14 Mizuno, H., Nakazawa, K., and Hayashi, C. (**1978**) Instability of a gaseous envelope surrounding a planetary core and formation of giant planets. *Progress in Theoretical Physics*, **60**, 699–710.
15 Bodenheimer, P. and Pollack, J.B. (**1986**) Calculations of the accretion and evolution of giant planets: the effects of solid cores. *Icarus*, **67**, 391–408.
16 Pollack, J.B., Hubickyj, O., Bodenheimer, P. *et al.* (**1996**) Formation of the giant planets by concurrent accretion of solids and gas. *Icarus*, **67**, 409–443.
17 Hubbard, W.B., Guillot, T., Marley, T. *et al.* (**1999**) Comparative evolution of Jupiter and Saturn. *Planetary and Space Science*, **47**, 1175–1182.
18 Saumon, D. and Guillot, T. (**2004**) Shock compression of deuterium and the interiors of Jupiter and Saturn. *The Astrophysical Journal*, **609**, 1170–1180.
19 Marley, M.S. (**1999**) Interiors of the giant planets, in Encyclopedia of the Solar System (eds. P.Weissman, L.A.McFadden, and T.V.Johnson), Academic Press, New York, pp. 339–355.
20 Brush, S. (**1990**) Theories of the origin of the solar system 1956-1985. *Reviews of Modern Physics*, **62**, 43–112.
21 Pollack, J.B. and Bodenheimer, P. (**1989**) Theories of the origin and evolution of the giant planets, in Origin and Evolution of Planetary and Satellite Atmospheres, University of Arizona Press, Tucson, pp. 564–602.
22 Owen, T., Mahaffy, P., Nieman, H.B. *et al.* (**1999**) A low-temperature origin for the planetesimals that formed Jupiter. *Nature*, **402**, 269–270.
23 Young, R.E. (**2003**) The Galileo probe: how it has changed our understanding of Jupiter. *New Astronomy Reviews*, **47**, 1–51.
24 Mousis, O., Marboeuf, U., Lunine, J.I., Alibert, Y., Fletcher, L.N., Orton, G.S., Pauzat, F., and Ellinger, Y. (**2009**) Determination of the minimum masses of heavy elements in the envelopes of Jupiter and Saturn. *Astrophys. J.* **696**, 1348–1354.
25 Cassen, P. and Woolum, D. (**1999**) Encyclopedia of the Solar System (eds. P.R.Weissman, L.A.McFadden, and T.V.Johnson), Academic Press, p. 35.
26 Haisch, K.E. Jr, Lada, E.A., and Lada, C.J. (**2001**) Disk frequencies and lifetimes in young clusters. *The Astrophysical Journal*, **553**, L153–L156.
27 Lada, E.A. (**2003**). Evolution of circumstellar disks in young stellar clusters. *Bulletin of the American Astronomical Society*, **35**, 24.06, 730.

28 Chen, C.H. and Kamp, I. (**2004**) Are giant planets forming around HR 4796A? *The Astrophysical Journal*, **602**, 985–992.

29 Metchev, S.A., Hillenbrand, L., and Meyer, M. (**2004**) Ten micron observations of nearby young stars. *The Astrophysical Journal*, **600**, 435–450.

30 Hillenbrand, L.A. (**2008**) Disk-dispersal and planet-formation timescales. *Physica Scripta*, **130**, 14024.

31 Bodenheimer, P., Hubickyj, O., and Lissauer, J.J. (**2000**) Models of the *in situ* formation of detected extrasolar giant planets. *Icarus*, **143**, 2–14.

32 Lin, D.N.C. and Papaloizou, J. (**1979**) On the structure of circumbinary accretion disks and the tidal evolution of commensurable satellites. *Monthly Notices of the Royal Astronomical Society*, **188**, 191–201.

33 Lin, D.N.C. and Papaloizou, J. (**1986**) On the tidal interaction between protoplanets and the protoplanetary disk. III - orbital migration of protoplanets. *The Astrophysical Journal*, **309**, 846–857.

34 Lin, D.N.C., Papaloizou, J.C.B., and Kley, W. (**1993**) On the nonaxisymmetric convective instabilities in accretion disks. *The Astrophysical Journal*, **16**, 689–699.

35 Saumon, D., Chabrier, G., and Van Horn, H.M. (**1996**) An equation of state for low-mass stars and giant planets. *The Astrophysical Journal. Supplement Series*, **99**, 713–741.

36 Burrows, A., Marley, M., Hubbard, W.B. *et al.* (**1997**) A nongray theory of extrasolar giant planets and brown dwarfs. *The Astrophysical Journal*, **491**, 856–875.

37 Hubickyj, O., Bodenheimer, P., and Lissauer, J.J. (**2005**) Accretion of the gaseous envelope of Jupiter around a 5 10 Earth-mass core. *Icarus*, **179**, 415–431.

38 Guillot, T., Gautier, D., and Hubbard, W.B. (**1997**) New constraints on the composition of Jupiter from Galileo measurements and interior models. *Icarus*, **130**, 534–539.

39 Gonzalez, G. (**1998**) Spectroscopic analyses of the parent stars of extrasolar planetary system candidates. *Astronomy and Astrophysics*, **334**, 221–238.

40 Santos, N.C., Israelian, G., and Mayor, M. (**2001**) The metal-rich nature of stars with planets. *Astronomy and Astrophysics*, **373**, 1019–1031.

41 Santos, N.C., Israelian, G., and Mayor, M. (**2004**) Spectroscopic Fe/H for 98 extra–solar planet-host stars. Exploring the probability of planet formation. *Astronomy and Astrophysics*, **415**, 1153–1166.

42 Fischer, D.A. and Valenti, J.A. (**2003**) Metallicities of stars with extrasolar planets, in *Scientific Frontiers in Research on Extrasolar Planets*, ASP Conference Series, vol. **294** (eds. D. Deming and S. Seager), Astronomical Society of the Pacific, San Francisco, pp. 117–128.

43 Fischer, D.A. and Valenti, J.A. (**2005**) The planet-metallicity correlation. *The Astrophysical Journal*, **622**, 1102–1117.

44 Valenti, J.A. and Fischer, D.A. (**2008**) Relationship between giant planet frequency and stellar metallicity. *Physica Scripta*, **130**, 14003.

45 Ida, S. and Lin, D.N.C. (**2004**) Toward a deterministic model of planetary formation. II. The formation and retention of gas giant planets around stars with a range of metallicities. *The Astrophysical Journal*, **616**, 567–572.

46 Kornet, K., Bodenheimer, P., Rózyczka, M., and Stepinski, T.F. (**2005**) Formation of giant planets in disks with different metallicities. *Astronomy and Astrophysics*, **430**, 1133–1138.

47 Ikoma, M., Nakazawa, K., and Emori, H. (**2000**) Formation of giant planets: dependences on core accretion rate and grain opacity. *The Astrophysical Journal*, **537**, 1013–1025.

48 Inaba, S., Wetherill, G.W., and Ikoma, M. (**2003**) Formation of gas giant planets: core accretion models with fragmentation and planetary envelope. *Icarus*, **166**, 46–62.

49 Inaba, S. and Ikoma, M. (**2003**) Enhanced collisional growth of a protoplanet that has an atmosphere. *Astronomy and Astrophysics*, **410**, 711–723.

50 Wuchterl, G. (**1991**) Hydrodynamics of giant planet formation II: model equations and critical mass. *Icarus*, **91**, 39–52.

51 Wuchterl, G. (**1991**) Hydrodynamics of giant planet formation III: Jupiter's nucleated instability. *Icarus*, **91**, 53–64.

52 Benvenuto, O.G. and Brunini, A. (**2005**) Methods for computing giant planet formation and evolution. *Monthly Notices of the Royal Astronomical Society*, **356**, 1383–1395.

53 Morbidelli, A. and Crida, A. (**2007**) The dynamics of Jupiter and Saturn in the gaseous protoplanetary disk. *Icarus*, **191**, 158–171.

54 Greenzweig, Y. and Lissauer, J.J. (**1992**) Accretion rates of protoplanets. II. Gaussian distributions of planetesimal velocities. *Icarus*, **100**, 440–463.

55 Greenzweig, Y. and Lissauer, J.J. (**1990**) Accretion rates of protoplanets. *Icarus*, **87**, 40–77.

56 Kary, D.M. and Lissauer, J.J. (**1994**) Numerical simulations of planetary growth, in Numerical Simulations in Astrophysics (eds. J.Franco, S.Lizano, L.Aguilar, and E.Daltabuit), Cambridge University Press, Cambridge, pp. 364–373.

57 Podolak, M., Pollack, J.B., and Reynolds, R.T. (**1988**) Interactions of planetesimals with protoplanetary atmospheres. *Icarus*, **73**, 163–179.

58 Pollack, J.B., Podolak, M., Bodenheimer, P., and Christofferson, B. (**1986**) Planetesimal dissolution in the envelopes of the forming, giant planets. *Icarus*, **73**, 163–179.

59 Tajima, N. and Nakagawa, Y. (**1997**) Evolution and dynamical stability of the proto-giant-planet envelopes. *Icarus*, **126**, 282–292.

60 Bell, K.R., Cassen, P.M., Klahr, H.H., and Henning, Th. (**1997**) The structure and appearance of protostellar accretion disks: limits on disk flaring. *The Astrophysical Journal*, **486**, 372–387.

61 D'Angelo, G., Henning, T., and Kley, W. (**2003**) Thermodynamics of circumstellar disks with high-mass planets. *The Astrophysical Journal*, **599**, 548–576.

62 Stringfellow, G. S., Black, D. C., and Bodenheimer, P. (**1990**) Brown dwarfs as close companions to white dwarfs. *Astrophys. J.* **349**, L59–L62.

63 Lissauer, J. J, Hubickyj, O., D'Angelo, G., and Bodenheimer, P. (**2009**) Models of Jupiter's growth incorporating thermal and hydrodynamic constraints. *Icarus* **199**, 338–350.

64 Saumon, D., Chabrier, G., and Van Horn, H.M. (**1995**) An equation of state for low-mass stars and giant planets. *The Astrophysical Journal. Supplement Series*, **99**, 713–741.

65 Pollack, J.B., McKay, C.P., and Christofferson, B. (**1985**) A calculation of the Rosseland mean opacity of dust grains in primordial Solar System nebulae. *Icarus*, **64**, 471–492.

66 Alexander, D.R. and Ferguson, J.W. (**1994**) Low-temperature Rosseland opacities. *The Astrophysical Journal*, **437**, 879–891.

67 Podolak, M. (**2003**) The contribution of small grains to the opacity of protoplanetary atmospheres. *Icarus*, **165**, 428–437.

68 Movshovitz, N. and Podolak, M. (**2008**) The opacity of grains in protoplanetary atmospheres. *Icarus*, **194**, 368–378.

69 Guillot, T. (**2005**) The interiors of giant planets. Models and outstanding questions. *Annual Review of Earth and Planetary Sciences*, **33**, 493–530.

70 Militzer, B., Hubbard, W.B., Vorberger, J. *et al.* (**2008**) A massive core in jupiter predicted from first-principles simulations. *The Astrophysical Journal*, **688**, L45–L48.

71 Hubbard, W.B., Burrows, A., and Lunine, J.I. (**2002**) Theory of giant planets. *Annual Review of Astronomy and Astrophysics*, **40**, 103–136.

72 Alibert, Y., Mordasini, C., Benz, W., and Winisdoerffer, C. (**2005**) Models of giant planet formation with migration and disk evolution. *Astronomy and Astrophysics*, **434**, 343–353.

73 Thommes, E.W., Duncan, M.J., and Levison, H.F. (**2003**) Oligarchic growth of giant planets. *Icarus*, **161**, 431–455.

74 Kokubo, E. and Ida, S. (**2002**) Formation of protoplanet systems and diversity of planetary systems. *The Astrophysical Journal*, **581**, 666–680.

75 Fortier, A., Benvenuto, O.G., and Brunini, A. (**2007**) Oligarchic planetesimal accretion and giant planet

formation. *Astronomy and Astrophysics*, **473**, 311–322.

76 Ida, S. and Makino (**1993**) Toward a deterministic model of planetary formation. II. The formation and retention of gas giant planets around stars with a range of metallicities. *The Astrophysical Journal*, **616**, 567–572.

77 Lissauer, J.J. (**1987**) Timescales for planetary accretion and the structure of the protoplanetary disk. *Icarus*, **69**, 249–265.

78 Weidenschilling, S.J. (**1977**) The distribution of mass in the planetary system and solar nebula. *Astrophysics and Space Science*, **51**, 153–158.

79 Ward, W.R. (**1997**) Survival of planetary systems. *The Astrophysical Journal*, **482**, L211–L214.

80 Tanaka, H., Takeuchi, T., and Ward, W.R. (**2002**) Three-dimensional interaction between a planet and an isothermal gaseous disk. I. Corotation and Lindblad torques and planet migration. *The Astrophysical Journal*, **565**, 1257–1274.

81 Papaloizou, J.C.B. and Terquem, C. (**1999**) Critical planetary core masses in protoplanetary disks and the formation of short-period giant planets. *The Astrophysical Journal*, **521**, 823–838.

82 Shakura, N.I. and Sunyaev, R.A. (**1973**) Black holes in binary systems. Observational appearance. *Astronomy and Astrophysics*, **24**, 337–355.

83 Gonzalez, G. (**1997**) The stellar metallicity-giant planet connection. *Monthly Notices of the Royal Astronomical Society*, **285**, 403–412.

84 Sigurdsson, S., Richter, H.B., Hansen, B.M. *et al.* (**2003**) A young white dwarf companion to pulsar B1620-26: evidence for early planet formation. *Science*, **301**, 193–196.

85 Sozzetti, A. (**2004**) On the correlation between the orbital periods of extrasolar planets and the metallicity of the host stars. *Monthly Notices of the Royal Astronomical Society*, **354**, 1194–1200.

86 Santos, N.C., Israelian, G., Mayor, M. *et al.* (**2003**) Statistical properties of exoplanets. II. Metallicity, orbital parameters, and space velocities. *Astronomy and Astrophysics*, **398**, 363–376.

87 Livio, M. and Pringle, J.E. (**2003**) Metallicity, planetary formation, and migration. *Monthly Notices of the Royal Astronomical Society*, **346**, L42–L44.

88 Udry, S. and Santos, N.C. (**2007**) Statistical properties of exoplanets. *Annual Review of Astronomy and Astrophysics*, **45**, 397–439.

89 Macintosh, B., Max, C., Zuckerman, B. *et al.* (**2001**) Keck adaptive optics observations of TW hydrae association members, in Young Stars Near Earth: Progress and Prospects, ASP Conference Series, vol. **244** (eds. R.Jayawardhana and T.Greene), Astronomical Society of the Pacific, San Francisco, pp. 309.

90 Chabrier, G., Baraffe, I., Allard, F., and Hauschildt, P.H. (**2000**) Evolutionary models for very low-mass stars and brown dwarfs with dusty atmospheres. *The Astrophysical Journal*, **542**, 464–472.

91 Marley, M.S. (**2001**) Models of young brown dwarfs and giant planets, in Young Stars Near Earth: Progress and Prospects, ASP Conference Series, vol. **244** (eds. R. Jayawardhana and T. Greene), Astronomical Society of the Pacific, San Francisco, p. 316.

92 Baraffe, I., Chabrier, G., Barman, T.S. *et al.* (**2003**) Evolutionary models for cool brown dwarfs and extrasolar giant planets. The case of HD 209458. *Astronomy and Astrophysics*, **402**, 701–712.

Further Reading

Weidenschilling, S.J., Spaute, D., Davis, D.R., Marzari, F., and Ohtsuki, K. (**1997**) Accretional evolution of a planetesimal swarm: 2. The terrestrial zone. *Icarus*, **128**, 429–455.

Wuchterl, G., Guillot, T. and Lissauer, J.J. (**2000**) Giant planet formation, in *Protostars and Planets IV* (eds. V. Mannings, A.P. Boss, and S. Russell), University of Arizona Press, Tucson, pp. 1081–1109.

6
Formation of Terrestrial Planets

Sean N. Raymond

6.1
Introduction

The current paradigm for the growth of rocky planets in protoplanetary disks includes several distinct dynamical steps. These stages of growth are illustrated in Figure 6.1 and described in detail in Section 6.3. First, micron-sized dust grains settle to a thin disk midplane on a 10^4-year timescale. Next, grains grow quickly via sticky collisions to roughly 1 m in size (the steep slope starting at $\sim 10^4$ years in Figure 6.1). Turbulence acts to greatly concentrate these particles in small regions, which can undergo gravitational collapse, resulting in the formation of planetesimals with a range of sizes (the vertical spike in Figure 6.1). According to recent models, centimeter- or meter-sized particles can be sufficiently concentrated to form small planetary embryos (curve A). In contrast, the standard picture of embryo formation relies on a stage of runaway growth from kilometer-sized planetesimals, during which the growth rate is a strong function of mass (curve B). The two curves rejoin during the "oligarchic" phase, and the growth of planetary embryos is slowed because planetesimal velocities are increased (the flattening of the curve before 10^6 years). Once a critical threshold is reached at $\sim 10^6$ years, giant collisions between embryos occur. This phase of growth lasts 10^7–10^8 years and is influenced by giant planets, which are constrained to form in the few megayears' lifetime of the gas disk. Giant impacts become even larger as embryos grow, and culminate at about 10^8 years – in the solar system; this is the approximate timing of the Moon-forming impact. Planets continue to accrete small bodies at a decreased rate for long timescales.

In this chapter, the stage is first set by very roughly and quickly introducing protoplanetary disks, the birthplace of planets (Section 6.2). Next, the stages of terrestrial planet growth are reviewed (Section 6.3). Then, the compositions of terrestrial planets are discussed and simple accretion models to the known sample of extrasolar giant planets are applied (Section 6.4). Finally, conclusions and avenues for future study are presented (Section 6.5).

Formation and Evolution of Exoplanets. Edited by Rory Barnes

ISBN: 978-3-527-40896-2

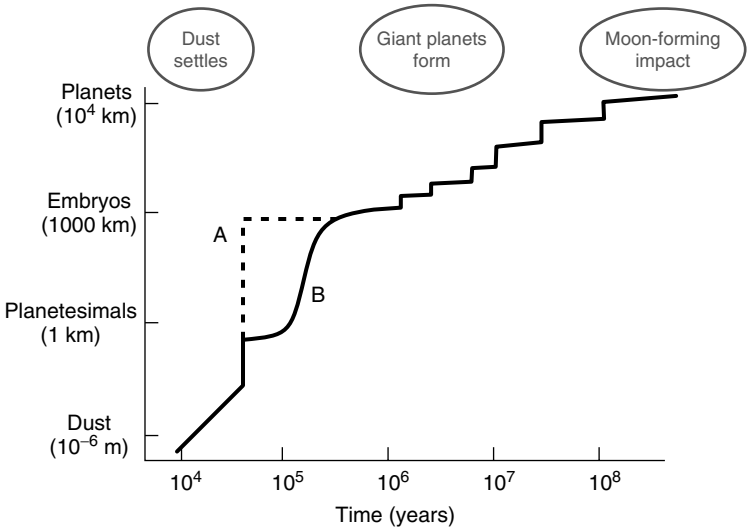


Figure 6.1 An illustration of the stages of terrestrial planet growth (not to scale). See the text for an explanation.

6.2
Setting the Stage: Protoplanetary Disks

Planet formation starts when a parcel of gas within a molecular cloud becomes gravitationally unstable, possibly triggered by a shock wave from, for example, a supernova [1, 2]. A protostar forms at the center of the collapsing clump of gas and a disk is spun from a portion of the remaining gas [3]. Disks have been observed around young stars via both direct imaging and via the spectroscopic detection of an infrared excess [4–7]. This excess is due to emission from dust grains, as direct detection of the gas is currently very difficult [8].

Protoplanetary disks undergo viscous evolution, which causes the disks to spread in radius such that most of the disk is eventually accreted onto the star [9–11]. The source of viscosity in protoplanetary disks is thought to be the magneto-rotational instability (MRI); [12, 13]. Mass accretion onto young stars is indeed observed to decrease in time [14, 15]. The surface density of material decreases steadily as the disk spreads, and the disk eventually dissipates via accretion onto the central star and through a combination of photoevaporation from both the central star and external sources [16–18].

The lifetimes of gaseous disks observed around young stars are relatively short, only a few megayears [19–21]. Gaseous planets such as Jupiter and Saturn must have formed within this timescale, which is 1–2 orders of magnitude shorter than the timescale for Earth's formation [22]. Thus, giant planets are fully formed during the final phases of terrestrial planet growth and certainly play an important role (see below).

One well-known model for the structure of the solar nebula is the ''minimum-mass solar nebula'' (MMSN) model of Weidenschilling [23] and Hayashi [24]. To generate the MMSN model, the mass of each planet is inflated to solar composition and spread into concentric contiguous annuli. The assumption is made that the surface density of solid material Σ in disks can be roughly approximated by a power law with radial distance r:

$$\Sigma(r) = \Sigma_1 r^{-\alpha} \tag{6.1}$$

where Σ_1 is the surface density at 1 AU and α controls the radial distribution of solids (not to be confused with the viscosity parameter, often represented by the same symbol). In the MMSN model, $\alpha = 3/2$ and $\Sigma_1 = 6$ g cm^{-2} and 1700 g cm^{-2} for the solid and gaseous components of the disk, respectively [24]. However, new analyses of the MMSN model derive values for α between 1/2 and 2 [25, 26]. In addition, current models and observations generally favor $\alpha \approx 1$ [27–29]. The total mass in the MMSN disk is $\sim 0.01 M_\odot$, although estimates for the solar nebula's actual mass range up to $0.1 M_\odot$ or even larger.

The masses of protoplanetary disks are difficult to measure. Recent studies have used millimeter and submillimeter observations to estimate outer disk masses for several young star clusters [5, 30, 31]. Typical disk masses appear to be somewhat less massive than the MMSN; Andrews and Williams [32, 33] measure a log-normal distribution for the disk mass with a mean of $0.001–0.005 M_\odot$ and a standard deviation of 1.3 dex. There is some evidence for a roughly linear disk mass–stellar mass correlation, although this is still debated [33–35]. Since the most important parameter for planet formation is probably the disk mass [36, 37], this is of great importance to planet formation.

Real disks are far more complicated than simple power laws. Disks are flared, contain zones with varying surface density and temperature profiles, and are turbulent [29, 38–40]. In terms of the solids, there exist condensation fronts for a range of molecular species corresponding to critical disk temperatures [41]. These fronts move around in time as the disk evolves [42, 43]. Terrestrial planets form from the $\sim$1% of the total disk mass that is condensable, so that in a given region of the disk the composition of solid bodies depends on a combination of the local temperature and temperature history, as well as the detailed composition of the disk, which can be inferred from a star's metallicity [44, 45].

Studies of terrestrial planet formation generally assume that disks are exceedingly simple and static; this is clearly not the case. It is certainly not tractable to include all the relevant physical effects when modeling the growth of solid bodies. However, it is important to remember that the setting may be bumpier than advertised.

6.3 Stages of Terrestrial Planet Growth

Planet formation takes place within evolving disks of roughly 99% gas and 1% dust, and completes on a $\sim 10^8$-year timescale. The discussion below starts with

a protoplanetary disk in which dust has settled to the thin "midplane" of the disk, a process which requires $\sim 10^4$ years in a nonturbulent disk [46–48]. Three distinct dynamical stages of terrestrial planet growth are reviewed: (i) the formation of kilometer-sized planetesimals starting from dust grains; (ii) the formation of ~1000-km-sized planetary embryos (also known as *protoplanets*); and (iii) the formation of full-sized terrestrial planets. In addition, some unresolved issues are also pointed out.

6.3.1 From Dust to Planetesimals

Planetesimals represent the transition from the gas-dominated to the gravity-dominated growth regimes. They are generally defined to be the smallest rocky bodies that are decoupled from the gaseous disk. The most commonly assumed planetesimal size is 1 km, corresponding to a mass of the order of 10^{16} g. However, kilometer-sized bodies are not completely decoupled from the gas, and they do have their orbits significantly altered by gas drag via relatively rapid ($\sim 10^3$–10^4 years) damping of their eccentricities and inclinations, and much slower ($\sim 10^6$ years) decay of their semimajor axes [49]. Thus, the actual size of planetesimals is unknown, as this is determined by their formation mechanism, which remains uncertain (see below). The planetesimal size is therefore used as a parameter in some models of later stages of planetary growth [50].

Modeling planetesimal formation requires a detailed treatment of the structure of the gaseous disk, including turbulence, local pressure gradients, magnetic processes, and vortices. Models can be constrained by observations of dust populations in disks around young stars, although interpretation of observations remains difficult [51]. Currently there exist two qualitatively different theories for planetesimal formation: collisional growth from smaller bodies [52] and local gravitational instability of smaller bodies [53, 54].

Collisional growth of micron-sized grains, especially if they are arranged into fluffy aggregates, appears efficient for relatively small particle sizes and impact speeds of ~ 1 m s^{-1} or slower [55–58]; see review by Dominik [59]). However, there is a constant battle between disk turbulence, which increases random velocities, and drag-induced settling, which reduces them [60, 61]. Growth of particles in such collisions appears effective until they reach roughly 1 cm to 1 m in size. At that point, continued growth may be suppressed by collision velocities of ≥ 10 m s^{-1} [51, 62].

Meter-sized bodies are the bottleneck of planetesimal formation. As an object in the gaseous disk grows, it becomes less strongly coupled to the gas such that its orbital velocity transitions between the gas velocity, which is slowed by partial pressure support, and the local Keplerian velocity. This increases the relative velocity between the object and the local gas such that the object feels a head wind, which acts to decrease its orbital energy and cause the body to spiral toward the star. Large (greater than or equal to tens to hundreds of meters) objects have enough inertia so that orbital decay occurs slowly, but there exists a critical size for which orbital decay is fastest. For the case of rocky bodies in a gaseous disk, this critical

size is roughly 1 m, and the timescale for infall for meter-sized bodies can be as short as 100 years at 1 AU [63]. This has been called the *meter-sized catastrophe*, because the infall timescale is far shorter than typical growth timescales. Collisional growth models must therefore quickly cross the barrier at meter sizes if they are to reach planetesimal sizes [52, 58, 64].

Alternatively, the gravitational instability model for planetesimal formation suggests that a large number of small patches of particles could become locally gravitationally unstable and form planetesimals [53, 54]. This process requires a concentration of meter-sized or smaller particles. If the density of solids in a small patch exceeds a critical value, then local gravitational instability can occur, leading to top-down formation of planetesimals. Concentration of small particles compared with the gas by a large factor is the key to the process.

Models for the concentration of small particles often rely on structure within the gaseous component of the disk, generated by turbulence or self-gravity [66, 67]. If the disk is even weakly turbulent, a size-dependent concentration of small particles can occur, leading to the formation of chondrules [66, 68, 69]. Self-gravitating clumps of chondrules may end up as 10- to 100-km sized planetesimals; in this case, particles do not collapse rapidly on the dynamical timescale but slowly contract into planetesimals [70]. Turbulence can also help concentrate larger, meter-sized particles by producing local pressure maxima, which can act as gathering points for small bodies. As for the meter-sized catastrophe, boulder-sized objects are the fastest to drift toward pressure maxima [71, 72].[1)] The concentration in these regions can be further increased via a streaming instability between the gas and solids [73, 74], and gravitational collapse of the clumps can occur in these dense regions. Johansen *et al.* [65] showed that planetesimals can form via this process and that the particle clumps (i.e., the rubble-pile planetesimals) have a distribution of sizes that ranges up to 1000 km or larger. Figure 6.2 shows the surface density of boulder-sized particles in a nonturbulent disk from a simulation of Johansen *et al.* [65] in which four 1000-km-scale objects have formed. An alternate location for planetesimal formation via gravitational instability are regions with an increased local density of solids [75]. Other ways to concentrate solids include drag-induced inspiralling to disk edges [76], vortices [77], or photoevaporative depletion of the gas layer [78].

6.3.2
From Planetesimals to Planetary Embryos

Planetary embryos are roughly lunar-sized objects that are generally thought to form by collisional growth of planetesimals [79]. However, given the efficiency with which turbulent concentration can form 1000-km bodies [65], it is possible that

1) In fact, the idea of the meter-sized catastrophe assumes that the disk has a smooth pressure gradient [63]. For disks with small-scale pressure fluctuations, small particles do not necessarily spiral inward but simply follow the local pressure gradient [71].

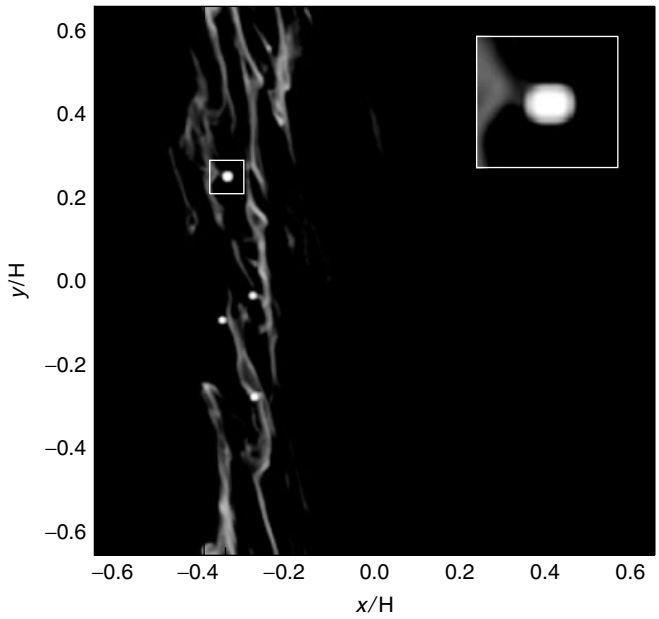


Figure 6.2 Concentration of boulder-sized particles in MRI-turbulent structures in a simulation by Johansen *et al*. [65]. The *x* and *y* axes are shown in units of the disk's vertical scale height H, and this snapshot is from seven orbital times after a clumping event occurred. The grayscale represents the local density of particles (lighter is denser), and the solid circles show the location of four clumps, each of which is more massive than Ceres (i.e., they correspond to ∼1000 km or larger "planetesimals" (or small embryos) in the overdense filament). The inset focuses on one clump as shown. Image credit: Anders Johansen.

the early, runaway growth stage of embryo formation may be bypassed. Indeed, several lines of evidence point to early seeding of the disk with 100- to 1000-km sized objects (see below). At the end of the later oligarchic growth stage, perhaps a few hundred embryos existed in the inner solar system. These embryos were the primary foodstuff of the terrestrial planets [80].

In the following discussion, we assume that kilometer-scale planetesimals have formed and are the primary component of the inner disk. While gas is still present in the disk, eccentricities of planetesimals are damped quickly because of aerodynamic gas drag, albeit it happens much faster for smaller planetesimals [49]. While velocities remain low, bodies that are slightly larger than the typical size can increase their collisional cross sections because of gravitational focusing and thereby accelerate their growth [81, 82]:

$$\frac{dM}{dt} \approx \frac{\pi R^2 v_{\mathrm{rand}} \Sigma}{2\,H} \left(1 + \frac{v_{\mathrm{esc}}^2}{v_{\mathrm{rand}}^2}\right) \tag{6.2}$$

where R represents the body's physical radius, Σ is the local surface density of planetesimals, v_{esc} is the escape speed from the body's surface ($v_{\mathrm{esc}} = \sqrt{2GM/R}$),

$v_{\rm rand}$ represents the velocity dispersion of planetesimals, and H is the vertical scale height of planetesimals. While random velocities are small, gravitational focusing can increase the growth rates of bodies by a factor of a few hundreds, such that $dM/dt \sim M^{4/3}$, leading to a phase of rapid ''runaway growth'' [82–87]. The length of this phase depends on the timescale for $v_{\rm rand}$ to increase, which depends on a combination of eccentricity growth via interactions with large bodies and eccentricity damping. For small (∼100-m-sized) planetesimals, gas drag is stronger such that runaway growth can be prolonged and embryos may be larger and grow faster [50, 88].

As large bodies undergo runaway growth, they gravitationally perturb nearby planetesimals. The random velocities of planetesimals are therefore increased by the larger bodies in a process called *viscous stirring* [85]. During this time, the random velocities of large bodies are kept small via dynamical friction with the swarm of small bodies [89]. As random velocities of planetesimals increase, gravitational focusing is reduced, and the growth of large bodies is slowed to the geometrical accretion limit, such that $dM/dt \sim M^{2/3}$ [88, 90]. Nonetheless, large bodies continue to grow and jostle each other such that a characteristic spacing of several mutual Hill radii $R_{H,m}$ is maintained ($R_{H,m} = 0.5[a_1 + a_2][M_1 + M_2/3M_*]^{1/3}$, where a_1 and M_1 denote the orbital distance and mass of object 1, etc; [91]). This phase of growth is often referred to as ''oligarchic growth'', as just a few large bodies dominate the dynamics of the system, with reduced growth rates and increased interactions between neighboring embryos [79, 92–94].

Figure 6.3 shows snapshots in time of a simulation of the formation of planetary embryos from planetesimals near 1 AU by Kokubo and Ida [93]. Accretion proceeds faster in the inner disk such that the outer disk is still dominated by planetesimals when embryos are fully formed in the inner disk. Oligarchic growth tends to form systems of embryos with roughly comparable masses and separations of 5–10 mutual Hill radii [79, 92, 95]. The details of the embryo distribution depend on the total mass and surface density distribution of the disk [93]. Typical embryo masses in a solar nebula model are a small percentage of an Earth mass, i.e., roughly lunar- to Mars-sized [79, 96]. Figure 6.4 shows nine distributions of embryos with a range in surface density exponents α and surface densities Σ_1 (see Eq. (6.1); [94]). For surface density profiles steeper than r^{-2}, the embryo mass decreases with orbital distance. Embryo masses scale roughly linearly with the local disk mass, and formation times are much faster for more massive disks.

The process of embryo formation via runaway and oligarchic growth has very recently come into question for three reasons. First, disk turbulence increases the random velocities of planetesimals, often above the critical disruption threshold for kilometer-sized planetesimals. Planetesimals have little tensile strength and are relatively easily disrupted, as represented in terms of Q_D^*, the specific energy required to gravitationally disperse half of the object's mass [97, 98]. For collisions more energetic than Q_D^*, collisions are erosive rather than accretionary, making it difficult for embryos to grow. In the presence of MRI-driven turbulence, accretionary growth of large bodies appears to require the presence of larger bodies with higher Q_D^* [99]. The critical size of these large bodies is 300–1000 km. Second, new collision

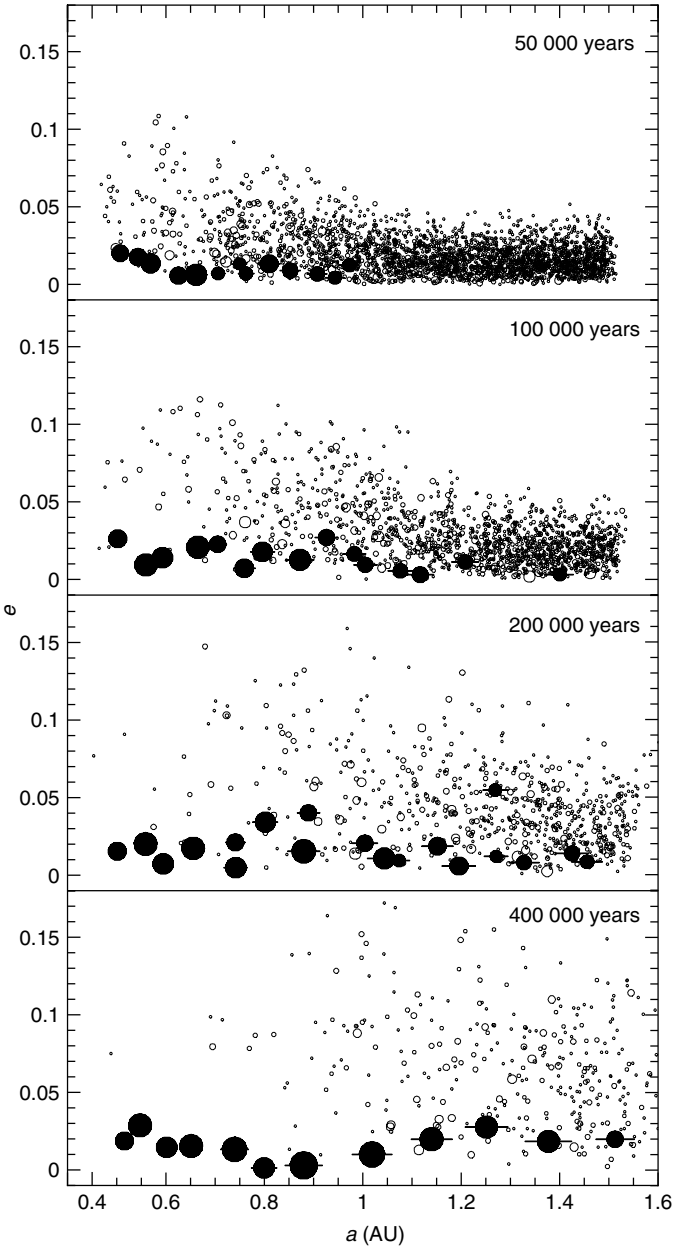


Figure 6.3 Snapshots in orbital eccentricity *e* versus semi-major axis *a* in simulations of the growth of planetary embryos by Kokubo and Ida [93]. The radius of each particle is proportional to the simulation radius but is not to scale on the *x* axis. Image credit: Eiichiro Kokubo.

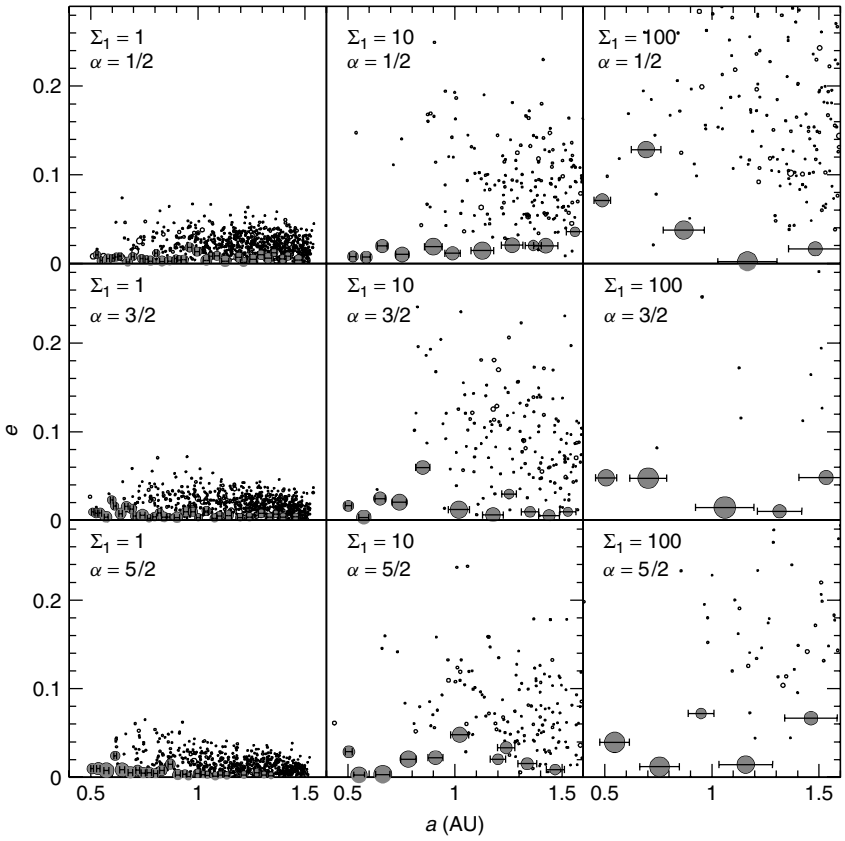


Figure 6.4 Orbital eccentricities e versus semimajor axes a for nine sets of planetary embryos formed in simulations of oligarchic growth [94], labeled by the surface density at 1 AU Σ_1 (in g cm^{-2}) and the surface density exponent α (see Eq. 6.1). Dot sizes are proportional to the sizes of particles in the simulations, and the horizontal error bars represent distances of 10 Hill radii. All simulations are shown at 500 000 years except $\Sigma_1 = 100$, $\alpha = 1/2$, which is shown at 110 000 years, and $\Sigma_1 = 100$, $\alpha = 5/2$, which is shown at 225 000 years. Image credit: Zoe Leinhardt.

models suggest that planetesimals are weaker than previously estimated, such that accretion requires either very slow collisions or preseeding of the disk with larger objects [100]. Third, statistical models that attempt to reproduce the asteroid belt's observed size distribution must also resort to seeding the region with large objects of at least 100 km in size [101]. These three lines of evidence all suggest that large, 100- to 1000-km bodies may have been required for the accretionary growth of large bodies. This paradox could be resolved if planetesimals form via the turbulent concentration plus gravitational collapse model of Johansen *et al.* [65], who inevitably formed 1000-km-scale bodies in MRI-turbulent disks.

Given the relatively short formation times of planetary embryos, interactions between embryos and the gaseous protoplanetary disk can be important. Embryos

more massive than roughly $0.1 M_{\oplus}$ can excite spiral density waves in the gaseous disk and undergo inward type 1 migration on a 10^{5-7} year timescale, depending on the embryo mass and the local surface density ([102–106]; see also Chapter 8). This can reduce the efficiency of terrestrial planet growth by allowing a significant portion of material to migrate all the way into the star [107, 108]. However, type 1 migration is affected by disk turbulence [40, 109], radiative effects [110, 111], and localized transitions in the gaseous disk's surface density or temperature [112]. The speed and even the direction of type 1 migration are, in fact, still debated.

It is important to realize that planetary embryos are the same objects as giant planet cores, assuming that giant planets form via the bottom-up ''core-accretion'' scenario ([41, 113–116]; see also Chapter 4). To form a gas giant, it is of vital importance to form a core of at least $1–10 M_{\oplus}$ before the dissipation of the gaseous disk [41, 117]. Models of core-accretion typically estimate the growth rate of an isolated core, then calculate the accretion of gas onto the core, often neglecting the potentially important effects of nearby cores [118].

6.3.3 From Planetary Embryos to Terrestrial Planets

Once the local mass in planetesimals and planetary embryos (i.e., ''small'' and ''large'' bodies, respectively) is comparable, feeding zones of neighboring embryos overlap, and giant collisions between embryos begin to occur [119, 120]. This is the start of the final phase of terrestrial planet growth, sometimes called *late-stage accretion* [80, 121]. During this stage, planets grow by accreting other embryos as well as planetesimals [122–124]. Eccentricities of embryos are kept smaller than those of planetesimals by dynamical friction, but not small enough to prevent embryo collisions. Growing planets clear their nearby zones of material, and their feeding zones widen and move outward in time [123]. During the very late stages of accretion, there can be substantial mixing across the inner planetary system, with important implications for planetary compositions (see below).

Giant collisions between embryos are not always accretionary: high-speed or off-center collisions can actually erode the target mass [125, 126]. High-speed collisions may alter the planetary composition by preferentially removing more volatile materials or materials closer to the surface; this is the proposed explanation for Mercury's large iron core [127–129]. In addition, low-speed, off-center collisions have the potential to create a circumplanetary disk of debris from which a large moon may accrete [130–134]. Note, however, that such low-speed collisions are rare, especially at later times when a planet's nearby zone has been cleared out [135].

Giant planets have a strong influence on late-stage terrestrial accretion. Gas giants are constrained to form in the 1- to 10-megayear lifetimes of the gaseous component of protoplanetary disks [19, 21]; Briceno *et al.* [20]. In contrast, Hf/W isotopes indicate that the last giant impact experienced by the Earth, presumably the Moon-forming impact, occurred at roughly 50–150 megayears [22]. Thus, if giant planets form in a planetary system, as Jupiter and Saturn did in the solar system, then they play a role in shaping the terrestrial planets during late-stage accretion

[124, 136, 137, 138]. The giant planets accelerate terrestrial accretion by increasing eccentricities of nearby objects, thereby increasing collision rates, but also ejecting bodies that undergo close encounters with the giant planet. In contrast, mutual gravitational interactions between embryos tend to increase eccentricities faster in the inner disk and cause an "accretion front" to move from the inside outward (see, e.g., Figure 6.3 or Raymond *et al.* [123]. The properties of a giant planet system systematically affect the terrestrial planets that can form (see Section 6.4). For example, systems with less massive giant planets will generate smaller embryo eccentricities, reducing the width of feeding zones and inducing the formation of more, smaller terrestrial planets as compared with a system with more massive giant planets [136, 137].

Figure 6.5 shows snapshots in time of the accretion of terrestrial planets from a simulation by Raymond *et al.* [123], including $9.9M_{\oplus}$ of material from 0.5 to 5 AU (roughly twice the mass of the MMSN model). The simulation included a Jupiter-mass giant planet on a circular orbit at 5.5 AU, which is consistent with the "Nice" model of the giant planets' orbital evolution [139]. Several mean motion resonances are clearly visible as vertical spikes in eccentricity after 100 000 years. As expected, eccentricities are driven in the inner disk via interactions between embryos and in the outer disk via secular and resonant perturbations from the giant planet. Large

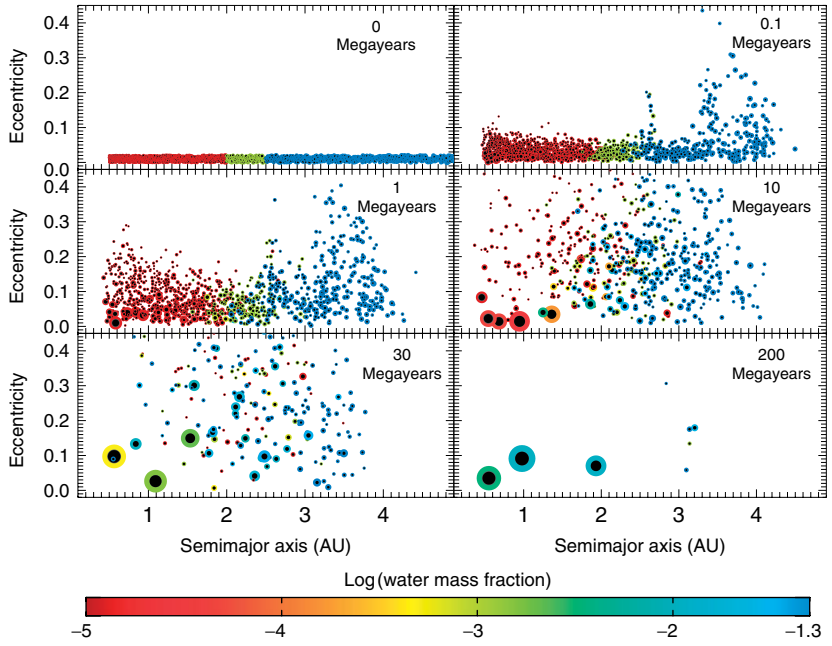


Figure 6.5 Snapshots in time from a simulation of the late-stage accretion of terrestrial planets, starting from 1885 subisolation mass objects [123]. The size of each body is proportional to its mass$^{1/3}$, the dark circle represents the relative size of each body's iron core, and the color corresponds to its water content (red = dry, blue = 5% water; see color bar).

bodies grow more quickly closer to the Sun, because of the faster orbital timescales; by 10 megayears the planet at ~1 AU has reached $1M_{\oplus}$. However, large-scale mixing between zones does not happen until about 20 megayears, when the feeding zones of all three final planets overlap and encompass the entire terrestrial zone, up to ~4 AU [140]. The final configuration of three planets contains a $1.5M_{\oplus}$ planet at 0.55 AU, a $2M_{\oplus}$ planet at 0.98 AU, and a $0.95M_{\oplus}$ planet at 1.93 AU (see [123] for details). The orbits of the planets in Figure 6.5 have slightly higher eccentricities than the current-day terrestrial planets. In similar simulations, O'Brien *et al.* [124] formed terrestrial planet systems with eccentricities even lower than those of Earth, Venus, and Mars.

The Earth's formation is generally considered to have concluded with the last giant impact at $\sim 10^8$ years [22]. After formation, terrestrial planets continue to be impacted by a decreasing flux of small bodies, also called a *late veneer* [141]. Once a system of planets has formed, it is stable for long timescales (at least as long as its formation time). Nonetheless, the orbital motion of the solar system has been shown to be chaotic on long timescales [142, 143]. In addition, dynamical instabilities on long timescales may be quite common [144, 139].

6.4 Planetary Compositions and Habitability

6.4.1 Compositions of Terrestrial Planets

The final compositions of terrestrial planets are determined by a combination of condensation of solid material in the disk and dynamical mixing of material from different radial zones. The ability of a given molecule to condense depends on the local pressure and temperature such that the only solids available to form planets in the inner disk are refractory materials such as rock and iron, whereas the outer disk contains more volatile species such as water and ammonia in solid form [45, 145]. However, given that both the stellar flux and the disk structure evolve in time [146], a planet's composition is largely determined by its feeding zone, i.e., the sum of all the material incorporated during formation (as well as physical processes during and after accretion). If planets form locally, then their compositions are a simple reflection of the composition of the local building blocks. Thus, eccentricity-driven mixing between different radial zones during accretion is a key process that determines the final planetary composition.

The Earth's water content is anomalously high: nebular models suggest that the local temperature at 1 AU was too hot to allow for hydration of planetesimals [147]. It is therefore thought that Earth's water was ''delivered'' from more distant regions in the form of hydrated asteroids [137, 148, 149] or comets [150]. The D/H ratio of Earth's water is virtually identical to that of carbonaceous chondrites, which are linked to C-type asteroids in the outer main belt ([151, 152]; see Table 1 of [148]). Comets, though poorly sampled, appear to have D/H ratios two times higher than

Earth [153–155], suggesting that primitive outer asteroid material may be the best candidate for the source of Earth's water. In Figure 6.5, material is given a starting water content that matches the values for primitive meteorites ([156]; see Figure 6.4 of [137]). Radial mixing during accretion delivers water from the primitive outer asteroid belt (beyond 2.5 AU) to the growing terrestrial planets. The amount of water delivered in planetesimals versus embryos is of the same order [140]. Thus, it appears that water delivery from planetesimals is statistically robust, while the amount of additional water from a small number of water-rich embryos can vary significantly from system to system.

6.4.2
Prospects for Terrestrial and Habitable Planets in Exoplanet Systems

Currently, the smallest known exoplanets orbit neutron stars and have masses as small as $0.02M_{\oplus}$ [157, 158]. However, smaller and smaller planets are being discovered orbiting main sequence stars using the radial velocity and transit methods [159, 160]. In addition, some low-mass planets are being found close to the circumstellar habitable zone, i.e., the annular zone that permits liquid surface water assuming certain atmospheric conditions [161–163]. In order to apply accretion models to situations that may be significantly different than the conditions that existed in the solar nebula, we need to understand the effect of external parameters on the accretion process. In terms of predictive power, the most important such parameters are quantities that are potentially observable (see [164] for a review).

The most important external factors that shape the growth of terrestrial planets are as follows:

- The properties of the solid disk. More massive disks form more massive planets, with a slightly stronger than linear correlation because more massive disks promote increased gravitational scattering and wider feeding zones [80, 137, 140, 165]. Because wider feeding zones allow for increased radial mixing, more massive disks have increased water delivery and form more water-rich terrestrial planets [140]. Disks with steeper density profiles form planets in the regions of highest mass concentration; for profiles flatter than $\Sigma \sim r^{-2}$ (i.e., $\alpha < 2$), the planet mass increases with orbital distance [140, 165, 166]; Lissauer [167]; this is the case in most disk models. The disk mass may control the frequency of giant planet formation, which would link the disk mass with the giant planet properties [36, 37]. In addition, there is some evidence that the disk mass may scale with stellar mass [32–34]. If that is indeed the case then there should exist a further correlation between terrestrial planets and the stellar mass [149].
- The properties of stellar companions or giant planets if they exist. More massive companions increase the width of feeding zones and result in systems of fewer, more massive terrestrial planets [136, 137]. For binary stars, there exists a simple factor of $\sim$5 scaling between the perihelion of the binary's orbit and the outermost terrestrial planet that is accreted [168, 169]. For Jupiter-like orbital distances, the giant planet orbital eccentricity is a key factor in controlling the

water contents of terrestrial planets. More eccentric giant planets strongly perturb and preferentially eject water-rich outer asteroidal material, resulting in systems of drier terrestrial planets as compared with systems with circular giant planets [137, 170]. Given that close-in giant planets are thought to form at large distances and migrate inward [171], hot Jupiter systems offer quite a different testbed for terrestrial planet formation. It has been shown that hot Jupiter systems should contain both very close-in terrestrial planets and more distant, very water-rich terrestrial planets ([172, 173]; see also [174, 175]).

- The dissipation of the disk. The gravitational potential of the disk affects the location of secular resonances between planets [176, 177]. As the disk dissipates, secular resonances can sweep across a given zone and induce collisions between planetary embryos [178, 179]. Tidal damping from small amounts of remnant gas can lower the eccentricities of embryos and also allow shepherding to occur [179–183]. However, secular resonance sweeping requires a smooth dispersal of the disk, which may be unlikely, given the punctuated nature of disk photoevaporation [18, 184]. In addition, the final phases of dissipation of the gaseous disk are thought to occur very rapidly [185, 186] and the possibility of a long-lived, low-density gaseous disk component remains speculative.

Given a basic understanding of terrestrial planet formation as a function of observable quantities such as the giant planet mass and orbital eccentricity, accretion models have been applied to the known sample of extrasolar planets [138, 172, 173]. These studies have focused on the potential for terrestrial planets to exist in the habitable zones (HZs) of other stars, and have derived inner and outer limits on the orbital distance of a giant planet that will allow a terrestrial planet to form in the HZ. Essentially, a giant planet must lie far enough from the HZ that it does not disrupt the accretion process. The outer limit corresponds to the closest a giant planet could lie to the HZ and still permit a planet to form in the region, 2.5 AU for a Sun-like star [138]. The inner limit corresponds to the typical separation between hot Jupiters and terrestrial planets, which have formed from material scattered outward; this separation varies significantly from simulation to simulation and is therefore a weaker constraint [172, 173].

Figure 6.6 shows the results of applying the derived limits to the sample of giant planets as of August 1, 2006 [187]. Roughly 37% (65 out of 178) of the known systems of planets met the criteria of having a known planet sufficiently detached from the HZ to have allowed a terrestrial planet to form in that region. Of these 65 systems, 17 cases satisfied the outer limit by having a distant giant planet and could have solar system-like architectures, i.e., terrestrial planets in the inner planetary system and giant planets farther out. The remaining 49 systems have giant planet limits interior to the HZ with small enough eccentricities to allow a terrestrial planet to have formed. For a list of known potentially habitable systems, see Mandell *et al*. [173]. Since this list was compiled, two systems on that list were found to contain additional planets in, or very close to, the HZ: Gliese 581 [162] and 55 Cancri (Fischer *et al*. [188]). Although these do not constitute true predictions (as for the case of HD 74156 d; see [189], they do lend credence to the model of Mandell *et al*. [173]).

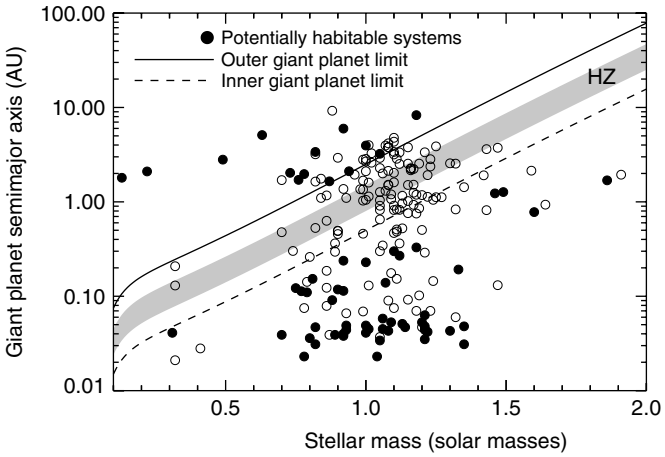


Figure 6.6 Extrasolar planetary systems that could harbor an Earth-like planet in the habitable zone. The habitable zone is shaded, and our inner and outer giant planet limits, listed in Figure 6.1 for specific values of the stellar mass, are shown by the dashed (inner) and solid (outer) lines. Each circle represents a known planet; those that fulfill either the inner or outer limit are filled in black. Note that many planets that appear to meet the limits are not filled; this is because their orbital eccentricities were too large. From Mandell *et al.* [173].

6.5 Conclusions

This chapter described the current state of knowledge about the formation of terrestrial planets. Figure 6.1 illustrates the stages of growth from micron-sized dust grains to full-sized planets. It is interesting to note that the field appears to be close to a paradigm shift regarding runaway growth as the first stage of planetary embryo growth from planetesimals. Several lines of evidence suggest that the growth of kilometer-sized planetesimals is very inefficient because of their weak internal structure and that protoplanetary disks must be ''seeded'' in advance with large, 100- to 1000-km bodies [99–101]. It is of great interest that the most recent models for the formation of planetesimals via concentration of centimeter- or meter-sized particles in turbulent clumps and subsequent gravitational collapse or agglomeration can readily form bodies in the 100- to 1000-km size range [65, 70]. Thus, while the current paradigm of terrestrial planet growth follows curve B in Figure 6.1, it may soon shift to curve A.

Several issues remain to be resolved in each stage of accretion and also in the structure and evolution of protoplanetary disks. For instance, the origin and role of low-viscosity ''dead zones'' is poorly understood [190–192]. The origin and strength of turbulence in disks is also poorly constrained [39, 40]. The details and consequences of giant impacts are very difficult to model, in terms of the fate of collisional debris and compositional changes induced by the impacts [126, 193, 194]. In addition, models readily reproduce Earth and Venus but not Mars or

Mercury. Although Mercury's small mass and large iron content may be related to frequent erosive collisions due to high-speed collisions [128], Mars' small mass remains a mystery [195, 196]. In addition, only recently, by taking dynamical friction into account, have models succeeded in reproducing the very small eccentricities of Venus, Earth, and Mars [124].

The effects of external parameters on accretion are perhaps less well understood than we would like to think. For example, the giant planet eccentricity–terrestrial planet water content correlation [137, 170] inherently assumes that eccentric giant planets acquire their eccentricities early, before accretion. If this assumption is wrong, then one can imagine a scenario in which terrestrial planets tend to form with relatively circular giant planets [139]; this is a beneficial situation in terms of water delivery. However, a late, impulsive eccentricity increase can destabilize the orbits of terrestrial planets and remove them from the system entirely [197]. Indeed, there are many known extrasolar planets that would be favorable for terrestrial planet formation if their orbits were more circular (see Figure 6.6). Thus, if giant planet eccentricity is acquired late, many systems may undergo ''planetary system suicide,'' forming an Earth-like planet and subsequently destroying it.

The first terrestrial planets to be found around other stars will most likely be very close to their host stars because of the biases of transit and radial velocity signals (see Chapter 1). There exist at least six unique scenarios for the formation of such ''hot Earth'' planets: (i) *in situ* accretion; (ii) inward type-1 migration; (iii) shepherding by giant planet migration; (iv) shepherding by secular resonance sweeping; (v) inward migration of a Neptune-like planet plus photoevaporation; and (vi) tidal circularization after a dynamical instability ([198]; see also [183, 199, 200]). Each of these models has a unique signature in terms of two variables: the inner planetary architecture and the planetary composition [198]. If both radial velocity (RV) and transit information are available, then both a mass and radius, and therefore a density, can be calculated. In theory, such information could yield the planetary composition in broad strokes [201–204]. However, given observational and model uncertainties, it is unlikely that the unambiguous characterization of an extrasolar terrestrial planet's composition will be possible for several years [205, 206]. In fact, the least unambiguous case may be the detection of the highest possible density planet; however, pure iron planets are likely to be rare and their small radii will surely render them difficult to detect. Nonetheless, when measurements of planetary composition do become possible, we will uncover the true nature of terrestrial planet formation.

Acknowledgments

The author is grateful to the CNRS and the Observatoire de Bordeaux, where he spent 3 months in the Fall of 2008 and where he wrote this chapter. He is especially grateful to Franck Selsis and Valentine Wakelam for their hospitality during his stay. The author also thanks Anders Johansen, Zoe Leinhardt, and Eiichiro Kokubo for the figures contributed by them.

References

1 Cameron, A.G.W. and Truran, J.W. (**1977**) *Icarus*, **30**, 447.
2 Boss, A.P. (**1995**) *The Astrophysical Journal*, **439**, 224.
3 Lin, D.N.C. and Pringle, J.E. (**1990**) *The Astrophysical Journal*, **358**, 515.
4 Smith, B.A. and Terrile, R.J. (**1984**) *Science*, **226**, 1421.
5 Beckwith, S.V.W., Sargent, A.I., Chini, R.S., and Guesten, R. (**1990**) *The Astronomical Journal*, **99**, 924.
6 Kenyon, S.J. and Hartmann, L. (**1995**) *The Astrophysical Journal Supplement Series*, **101**, 117.
7 O'Dell, C.R. (**2001**) *Annual Review of Astronomy and Astrophysics*, **39**, 99.
8 Eisner, J.A. (**2007**) *Nature*, **447**, 562.
9 Lynden-Bell, D. and Pringle, J.E. (**1974**) *Monthly Notices of the Royal Astronomical Society*, **168**, 603.
10 Lin, D.N.C. and Papaloizou, J. (**1980**) *Monthly Notices of the Royal Astronomical Society*, **191**, 37.
11 Lin, D.N.C. and Pringle, J.E. (**1987**) *Monthly Notices of the Royal Astronomical Society*, **225**, 607.
12 Balbus, S.A. and Hawley, J.F. (**1991**) *The Astrophysical Journal*, **376**, 214.
13 Gammie, C.F. (**1996**) *The Astrophysical Journal*, **457**, 355.
14 Gullbring, E., Hartmann, L., Briceno, C., and Calvet, N. (**1998**) *The Astrophysical Journal*, **492**, 323.
15 Hartmann, L., Calvet, N., Gullbring, E., and D'Alessio, P. (**1998**) *The Astrophysical Journal*, **495**, 385.
16 Hollenbach, D.J., Yorke, H.W., and Johnstone, D. (**2000**) *Protostars and Planets IV*, SAO/NASA Astrophysics Data System, p. 401.
17 Clarke, C.J., Gendrin, A., and Sotomayor, M. (**2001**) *Monthly Notices of the Royal Astronomical Society*, **328**, 485.
18 Adams, F.C., Hollenbach, D., Laughlin, G., and Gorti, U. (**2004**) *The Astrophysical Journal*, **611**, 360.
19 Haisch, K.E., Lada, E.A., and Lada, C.J. Jr. (**2001**) *The Astrophysical Journal*, **553**, L153.
20 Briceno, C., Vivas, A.K., Calvet, N. *et al.* (**2001**) *Science*, **291**, 93.
21 Pascucci, I., Gorti, U., Hollenbach, D. *et al.* (**2006**) *The Astrophysical Journal*, **651**, 1177.
22 Touboul, M., Kleine, T., Bourdon, B. *et al.* (**2007**) *Nature*, **450**, 1206.
23 Weidenschilling, S.J. (**1977**) *Astrophysics and Space Science*, **51**, 153.
24 Hayashi, C. (**1981**) *Progress of Theoretical Physics Supplement*, **70**, 35.
25 Davis, S.S. (**2005**) *The Astrophysical Journal*, **627**, L153.
26 Desch, S.J. (**2007**) *The Astrophysical Journal*, **671**, 878.
27 Dullemond, C.P., Hollenbach, D., Kamp, I., and D'Alessio, P. (**2007**) *Protostars and Planets V*, SAO/NASA Astrophysics Data System, 555.
28 Andrews, S.M. and Williams, J.P. (**2007**) *The Astrophysical Journal*, **659**, 705.
29 Garaud, P. and Lin, D.N.C. (**2007**) *The Astrophysical Journal*, **654**, 606.
30 Andre, P. and Montmerle, T. (**1994**) *The Astrophysical Journal*, **420**, 837.
31 Eisner, J.A. and Carpenter, J.M. (**2003**) *The Astrophysical Journal*, **598**, 1341.
32 Andrews, S.M. and Williams, J.P. (**2005**) *The Astrophysical Journal*, **631**, 1134.
33 Andrews, S.M. and Williams, J.P. (**2007**) *The Astrophysical Journal*, **671**, 1800.
34 Scholz, A., Jayawardhana, R., and Wood, K. (**2006**) *The Astrophysical Journal*, **645**, 1498.
35 Eisner, J.A., Plambeck, R.L., Carpenter, J.M. *et al.* (**2008**) *The Astrophysical Journal*, **683**, 304.
36 Greaves, J.S., Fischer, D.A., Wyatt, M.C. *et al.* (**2007**) *Monthly Notices of the Royal Astronomical Society*, **378**, L1.
37 Wyatt, M.C., Clarke, C.J., and Greaves, J.S. (**2007**) *Monthly Notices of the Royal Astronomical Society*, **380**, 1737.
38 Chiang, E.I. and Goldreich, P. (**1997**) *The Astrophysical Journal*, **490**, 368.
39 Papaloizou, J.C.B. and Nelson, R.P. (**2003**) *Monthly Notices of the Royal Astronomical Society*, **339**, 983.
40 Laughlin, G., Steinacker, A., and Adams, F.C. (**2004**) *The Astrophysical Journal*, **608**, 489.
41 Pollack, J.B., Hubickyj, O., Bodenheimer, P. *et al.* (**1996**) *Icarus*, **124**, 62.

42 Sasselov, D.D. and Lecar, M. (**2000**) *The Astrophysical Journal*, **528**, 995.
43 Dodson-Robinson, S.E., Willacy, K., Bodenheimer, P. *et al.* (**2009**) *Icarus*, **200**, 672.
44 Gaidos, E.J. (**2000**) *Icarus*, **145**, 637.
45 Lodders, K. (**2003**) *The Astrophysical Journal*, **591**, 1220.
46 Weidenschilling, S.J. (**1980**) *Icarus*, **44**, 172.
47 Nakagawa, Y., Nakazawa, K., and Hayashi, C. (**1981**) *Icarus*, **45**, 517.
48 Nomura, H. and Nakagawa, Y. (**2006**) *The Astrophysical Journal*, **640**, 1099.
49 Adachi, I., Hayashi, C., and Nakazawa, K. (**1976**) *Progress of Theoretical Physics*, **56**, 1756.
50 Chambers, J. (**2006**) *Icarus*, **180**, 496.
51 Dullemond, C.P. and Dominik, C. (**2005**) *Astronomy and Astrophysics*, **434**, 971.
52 Weidenschilling, S.J. and Cuzzi, J.N. (**1993**) *Protostars and Planets III*, SAO/NASA Astrophysics Data System, 1031.
53 Goldreich, P. and Ward, W.R. (**1973**) *The Astrophysical Journal*, **183**, 1051.
54 Youdin, A.N. and Shu, F.H. (**2002**) *The Astrophysical Journal*, **580**, 494.
55 Dominik, C. and Tielens, A.G.G.M. (**1997**) *The Astrophysical Journal*, **480**, 647.
56 Wurm, G. and Blum, J. (**2000**) *The Astrophysical Journal*, **529**, L57.
57 Poppe, T., Blum, J., and Henning, T. (**2000**) *The Astrophysical Journal*, **533**, 454.
58 Benz, W. (**2000**) *Space Science Reviews*, **92**, 279.
59 Dominik, C., Blum, J., Cuzzi, J.N., and Wurm, G. (**2007**) *Protostars and Planets V*, SAO/NASA Astrophysics Data System, 783.
60 Cuzzi, J.N., Dobrovolskis, A.R., and Champney, J.M. (**1993**) *Icarus*, **106**, 102.
61 Cuzzi, J.N. and Weidenschilling, S.J. (**2006**) *Meteorites and the Early Solar System II*, SAO/NASA Astrophysics Data System, 353.
62 Dullemond, C.P. and Dominik, C. (**2004**) *Astronomy and Astrophysics*, **421**, 1075.
63 Weidenschilling, S.J. (**1977**) *Monthly Notices of the Royal Astronomical Society*, **180**, 57.
64 Weidenschilling, S.J. (**2000**) *Space Science Reviews*, **92**, 295.
65 Johansen, A., Oishi, J.S., Low, M.-M.M. *et al.* (**2007**) *Nature*, **448**, 1022.
66 Cuzzi, J.N., Hogan, R.C., Paque, J.M., and Dobrovolskis, A.R. (**2001**) *The Astrophysical Journal*, **546**, 496.
67 Rice, W.K.M., Lodato, G., Pringle, J.E. *et al.* (**2006**) *Monthly Notices of the Royal Astronomical Society*, **372**, L9.
68 Cuzzi, J.N., Dobrovolskis, A.R., and Hogan, R.C. (**1996**) *Chondrules and the Protoplanetary Disk*, SAO/NASA Astrophysics Data System, 35.
69 Cuzzi, J.N. and Alexander, C.M.O. (**2006**) *Nature*, **441**, 483.
70 Cuzzi, J.N., Hogan, R.C., and Shariff, K. (**2008**) *The Astrophysical Journal*, **687**, 1432.
71 Haghighipour, N. and Boss, A.P. (**2003**) *The Astrophysical Journal*, **598**, 1301.
72 Johansen, A., Klahr, H., and Henning, T. (**2006**) *The Astrophysical Journal*, **636**, 1121.
73 Youdin, A.N. and Goodman, J. (**2005**) *The Astrophysical Journal*, **620**, 459.
74 Johansen, A. and Youdin, A. (**2007**) *The Astrophysical Journal*, **662**, 627.
75 Goodman, J. and Pindor, B. (**2000**) *Icarus*, **148**, 537.
76 Youdin, A.N. and Chiang, E.I. (**2004**) *The Astrophysical Journal*, **601**, 1109.
77 Tanga, P., Babiano, A., Dubrulle, B., and Provenzale, A. (**1996**) *Icarus*, **121**, 158.
78 Throop, H.B. and Bally, J. (**2005**) *The Astrophysical Journal*, **623**, L149.
79 Kokubo, E. and Ida, S. (**2000**) *Icarus*, **143**, 15.
80 Wetherill, G.W. (**1996**) *Icarus*, **119**, 219.
81 Safronov, V.S. (**1969**), http://adsabs.harvard.edu/abs/1969QB981.S26 1969.
82 Greenberg, R., Hartmann, W.K., Chapman, C.R., and Wacker, J.F. (**1978**) *Icarus*, **35**, 1.
83 Wetherill, G.W. and Stewart, G.R. (**1989**) *Icarus*, **77**, 330.
84 Wetherill, G.W. and Stewart, G.R. (**1993**) *Icarus*, **106**, 190.
85 Ida, S. and Makino, J. (**1992**) *Icarus*, **96**, 107.

86 Kokubo, E. and Ida, S. (**1996**) *Icarus*, **123**, 180.
87 Goldreich, P., Lithwick, Y., and Sari, R. (**2004**) *The Astrophysical Journal*, **614**, 497.
88 Rafikov, R.R. (**2003**) *The Astronomical Journal*, **125**, 942.
89 Ida, S. and Makino, J. (**1992**) *Icarus*, **98**, 28.
90 Ida, S. and Makino, J. (**1993**) *Icarus*, **106**, 210.
91 Kokubo, E. and Ida, S. (**1995**) *Icarus*, **114**, 247.
92 Kokubo, E. and Ida, S. (**1998**) *Icarus*, **131**, 171.
93 Kokubo, E. and Ida, S. (**2002**) *The Astrophysical Journal*, **581**, 666.
94 Leinhardt, Z.M. and Richardson, D.C. (**2005**) *The Astrophysical Journal*, **625**, 427.
95 Weidenschilling, S.J., Spaute, D., Davis, D.R. *et al.* (**1997**) *Icarus*, **128**, 429.
96 Collins, B.F. and Sari, R. (**2009**) *The Astronomical Journal*, **137**, 3778.
97 Melosh, H.J. and Ryan, E.V. (**1997**) *Icarus*, **129**, 562.
98 Benz, W. and Asphaug, E. (**1999**) *Icarus*, **142**, 5.
99 Ida, S., Guillot, T., and Morbidelli, A. (**2008**) *The Astrophysical Journal*, **686**, 1292.
100 Stewart, S.T. and Leinhardt, Z.M. (**2009**) *The Astrophysical Journal*, **691**, L133.
101 Morbidelli, A., Nesvorny, D., Bottke, W.F., and Levison, H.F. (**2008**) *LPI Contributions*, **1405**, 8042.
102 Goldreich, P. and Tremaine, S. (**1980**) *The Astrophysical Journal*, **241**, 425.
103 Ward, W.R. (**1986**) *Icarus*, **67**, 164.
104 Ward, W.R. (**1997**) *Icarus*, **126**, 261.
105 Lin, D.N.C. and Papaloizou, J. (**1986**) *The Astrophysical Journal*, **309**, 846.
106 Tanaka, H., Takeuchi, T., and Ward, W.R. (**2002**) *The Astrophysical Journal*, **565**, 1257.
107 McNeil, D., Duncan, M., and Levison, H.F. (**2005**) *The Astronomical Journal*, **130**, 2884.
108 Ida, S. and Lin, D.N.C. (**2008**) *The Astrophysical Journal*, **673**, 487.
109 Nelson, R.P. (**2005**) *Astronomy and Astrophysics*, **443**, 1067.
110 Paardekooper, S.-J. and Mellema, G. (**2006**) *Astronomy and Astrophysics*, **459**, L17.
111 Kley, W. and Crida, A. (**2008**) *Astronomy and Astrophysics*, **487**, L9.
112 Masset, F.S., D'Angelo, G., and Kley, W. (**2006**) *The Astrophysical Journal*, **652**, 730.
113 Mizuno, H. (**1980**) *Progress of Theoretical Physics*, **64**, 544.
114 Bodenheimer, P. and Pollack, J.B. (**1986**) *Icarus*, **67**, 391.
115 Ida, S. and Lin, D.N.C. (**2004**) *The Astrophysical Journal*, **604**, 388.
116 Alibert, Y., Mordasini, C., Benz, W., and Winisdoerffer, C. (**2005**) *Astronomy and Astrophysics*, **434**, 343.
117 Ikoma, M., Nakazawa, K., and Emori, H. (**2000**) *The Astrophysical Journal*, **537**, 1013.
118 Thommes, E.W., Matsumura, S., and Rasio, F.A. (**2008**) *Science*, **321**, 814.
119 Wetherill, G.W. (**1985**) *Science*, **228**, 877.
120 Kenyon, S.J. and Bromley, B.C. (**2006**) *The Astronomical Journal*, **131**, 1837.
121 Chambers, J.E. and Wetherill, G.W. (**1998**) *Icarus*, **136**, 304.
122 Chambers, J.E. (**2001**) *Icarus*, **152**, 205.
123 Raymond, S.N., Quinn, T., and Lunine, J.I. (**2006**) *Icarus*, **183**, 265.
124 O'Brien, D.P., Morbidelli, A., and Levison, H.F. (**2006**) *Icarus*, **184**, 39.
125 Agnor, C. and Asphaug, E. (**2004**) *The Astrophysical Journal*, **613**, L157.
126 Asphaug, E., Agnor, C.B., and Williams, Q. (**2006**) *Nature*, **439**, 155.
127 Wetherill, G.W. (**1988**) *Mercury*, University of Arizona Press, p. 670.
128 Benz, W., Slattery, W.L., and Cameron, A.G.W. (**1988**) *Icarus*, **74**, 516.
129 Benz, W., Anic, A., Horner, J., and Whitby, J.A. (**2007**) *Space Science Reviews*, **132**, 189.
130 Benz, W., Slattery, W.L., and Cameron, A.G.W. (**1986**) *Icarus*, **66**, 515.
131 Benz, W., Slattery, W.L., and Cameron, A.G.W. (**1987**) *Icarus*, **71**, 30.
132 Ida, S., Canup, R.M., and Stewart, G.R. (**1997**) *Nature*, **389**, 353.
133 Canup, R.M. and Asphaug, E. (**2001**) *Nature*, **412**, 708.
134 Canup, R.M. (**2004**) *Icarus*, **168**, 433.

135 Agnor, C.B., Canup, R.M., and Levison, H.F. (**1999**) *Icarus*, **142**, 219.
136 Levison, H.F. and Agnor, C. (**2003**) *The Astronomical Journal*, **125**, 2692.
137 Raymond, S.N., Quinn, T., and Lunine, J.I. (**2004**) *Icarus*, **168**, 1.
138 Raymond, S.N. (**2006**) *The Astrophysical Journal*, **643**, L131.
139 Tsiganis, K., Gomes, R., Morbidelli, A., and Levison, H.F. (**2005**) *Nature*, **435**, 459.
140 Raymond, S.N., Quinn, T., and Lunine, J.I. (**2007**) *Astrobiology*, **7**, 66.
141 Bottke, W.F., Levison, H.F., Nesvorný, D., and Dones, L. (**2007**) *Icarus*, **190**, 203.
142 Laskar, J. (**1989**) *Nature*, **338**, 237.
143 Laskar, J. (**1990**) *Icarus*, **88**, 266.
144 Levison, H.F., Lissauer, J.J., and Duncan, M.J. (**1998**) *The Astronomical Journal*, **116**, 1998.
145 Pollack, J.B., Hollenbach, D., Beckwith, S. *et al.* (**1994**) *The Astrophysical Journal*, **421**, 615.
146 Ciesla, F.J. and Cuzzi, J.N. (**2006**) *Icarus*, **181**, 178.
147 Boss, A.P. (**1998**) *Annual Review of Earth and Planetary Sciences*, **26**, 53.
148 Morbidelli, A., Chambers, J., Lunine, J.I. *et al.* (**2000**) *Meteoritics and Planetary Science*, **35**, 1309.
149 Raymond, S.N., Scalo, J., and Meadows, V.S. (**2007**) *The Astrophysical Journal*, **669**, 606.
150 Owen, T. and Bar-Nun, A. (**1995**) *Icarus*, **116**, 215.
151 Robert, F. and Epstein, S. (**1982**) *Geochimica et Cosmochimica Acta*, **46**, 81.
152 Kerridge, J.F. (**1985**) *Geochimica et Cosmochimica Acta*, **49**, 1707.
153 Balsiger, H., Altwegg, K., and Geiss, J. (**1995**) *Journal of Geophysical Research*, **100**, 5827.
154 Meier, R., Owen, T.C., Matthews, H.E. *et al.* (**1998**) *Science*, **279**, 842.
155 Bockelee-Morvan, D., Gautier, D., Lis, D.C. *et al.* (**1998**) *Icarus*, **133**, 147.
156 Gradie, J. and Tedesco, E. (**1982**) *Science*, **216**, 1405.
157 Wolszczan, A. and Frail, D.A. (**1992**) *Nature*, **355**, 145.
158 Wolszczan, A. (**1994**) *Science*, **264**, 538.
159 Gillon, M., Pont, F., Demory, B.-O. *et al.* (**2007**) *Astronomy and Astrophysics*, **472**, L13.
160 Mayor, M., Udry, S., Lovis, C. *et al.* (**2009**) *Astronomy and Astrophysics*, **493**, 639.
161 Kasting, J.F., Whitmire, D.P., and Reynolds, R.T. (**1993**) *Icarus*, **101**, 108.
162 Udry, S., Bonfils, X., Delfosse, X. *et al.* (**2007**) *Astronomy and Astrophysics*, **469**, L43.
163 Selsis, F., Chazelas, B., Bordé, P. *et al.* (**2007**) *Icarus*, **191**, 453.
164 Raymond, S.N. (**2008**) *IAU Symposium*, **249**, 233.
165 Kokubo, E., Kominami, J., and Ida, S. (**2006**) *The Astrophysical Journal*, **642**, 1131.
166 Raymond, S.N., Quinn, T., and Lunine, J.I. (**2005**) *The Astrophysical Journal*, **632**, 670.
167 Lissauer, J. J. (**1987**) *Icarus*, **69**, 249.
168 Quintana, E.V., Adams, F.C., Lissauer, J.J., and Chambers, J.E. (**2007**) *The Astrophysical Journal*, **660**, 807.
169 Haghighipour, N. and Raymond, S.N. (**2007**) *The Astrophysical Journal*, **666**, 436.
170 Chambers, J.E. and Cassen, P. (**2002**) *Meteoritics and Planetary Science*, **37**, 1523.
171 Lin, D.N.C., Bodenheimer, P., and Richardson, D.C. (**1996**) *Nature*, **380**, 606.
172 Raymond, S.N., Mandell, A.M., and Sigurdsson, S. (**2006**) *Science*, **313**, 1413.
173 Mandell, A.M., Raymond, S.N., and Sigurdsson, S. (**2007**) *The Astrophysical Journal*, **660**, 823.
174 Fogg, M.J. and Nelson, R.P. (**2005**) *Astronomy and Astrophysics*, **441**, 791.
175 Fogg, M.J. and Nelson, R.P. (**2007**) *Astronomy and Astrophysics*, **472**, 1003.
176 Heppenheimer, T.A. (**1980**) *Icarus*, **41**, 76.
177 Ward, W.R. (**1981**) *Icarus*, **47**, 234.
178 Nagasawa, M., Lin, D.N.C., and Thommes, E. (**2005**) *The Astrophysical Journal*, **635**, 578.
179 Thommes, E., Nagasawa, M., and Lin, D.N.C. (**2008**) *The Astrophysical Journal*, **676**, 728.

180 Agnor, C.B. and Ward, W.R. (**2002**) *The Astrophysical Journal*, **567**, 579.

181 Kominami, J. and Ida, S. (**2002**) *Icarus*, **157**, 43.

182 Kominami, J. and Ida, S. (**2004**) *Icarus*, **167**, 231.

183 Zhou, J.-L., Aarseth, S.J., Lin, D.N.C., and Nagasawa, M. (**2005**) *The Astrophysical Journal*, **631**, L85.

184 Hollenbach, D., Johnstone, D., Lizano, S., and Shu, F. (**1994**) *The Astrophysical Journal*, **428**, 654.

185 Wolk, S.J. and Walter, F.M. (**1996**) *The Astronomical Journal*, **111**, 2066.

186 Chiang, E. and Murray-Clay, R. (**2007**) *Nature Physics*, **3**, 604.

187 Butler, R.P., Wright, J.T., Marcy, G.W. *et al.* (**2006**) *The Astrophysical Journal*, **646**, 505.

188 Fischer, D. A. *et al.* (**2008**) **675**, 790.

189 Barnes, R., Goździewski, K., and Raymond, S.N. (**2008**) *The Astrophysical Journal*, **680**, L57.

190 Fromang, S., Terquem, C., and Balbus, S.A. (**2002**) *Monthly Notices of the Royal Astronomical Society*, **329**, 18.

191 Matsumura, S. and Pudritz, R.E. (**2003**) *The Astrophysical Journal*, **598**, 645.

192 Turner, N.J., Sano, T., and Dziourkevitch, N. (**2007**) *The Astrophysical Journal*, **659**, 729.

193 Genda, H. and Abe, Y. (**2005**) *Nature*, **433**, 842.

194 Canup, R.M. and Pierazzo, E. (**2006**) *7th Annual Lunar and Planetary Science Conference in Houston*, vol. **37**, p. 2146.

195 Wetherill, G.W. (**1991**) *Lunar and Planetary Institute Science Conference Abstracts in Houston*, vol. **22**, 1495.

196 Raymond, S.N., O'Brien, D.P., Morbidelli, A., and Kaib, N.A. (**2009**) *Icarus*, **203**, 644.

197 Veras, D. and Armitage, P.J. (**2006**) *The Astrophysical Journal*, **645**, 1509.

198 Raymond, S.N., Barnes, R., and Mandell, A.M. (**2008**) *Monthly Notices of the Royal Astronomical Society*, **384**, 663.

199 Terquem, C. and Papaloizou, J.C.B. (**2007**) *The Astrophysical Journal*, **654**, 1110.

200 Gaidos, E., Haghighipour, N., Agol, E. *et al.* (**2007**) *Science*, **318**, 210.

201 Fortney, J.J., Marley, M.S., and Barnes, J.W. (**2007**) *The Astrophysical Journal*, **659**, 1661.

202 Valencia, D., Sasselov, D.D., and O'Connell, R.J. (**2007**) *The Astrophysical Journal*, **665**, 1413.

203 Sotin, C., Grasset, O., and Mocquet, A. (**2007**) *Icarus*, **191**, 337.

204 Seager, S., Kuchner, M., Hier-Majumder, C.A., and Militzer, B. (**2007**) *The Astrophysical Journal*, **669**, 1279.

205 Adams, E.R., Seager, S., and Elkins-Tanton, L. (**2008**) *The Astrophysical Journal*, **673**, 1160.

206 Selsis, F., Kasting, J.F., Levrard, B. *et al.* (**2007**) *Astronomy and Astrophysics*, **476**, 1373.

Further Reading

Benz, W., Kallenbach, R., and Lugmair, G.W. (**2000**) *From Dust to Terrestrial Planets*, Kluwer Academic Publishers.

Papaloizou, J.C.B. and Terquem, C. (**2006**) *Reports on Progress in Physics*, **69**, 119.

Weidenschilling, S.J. (**1984**) *Icarus*, **60**, 553.

7
Brown Dwarfs

Kevin L. Luhman

7.1
Introduction

The first extrasolar giant planet around a Sun-like star and the first brown dwarf were discovered almost simultaneously [1–3]. Since that time, a great deal of progress has been made in studying the origin, evolution, and physical characteristics of these two classes of substellar objects. Because of the expected similarities in their atmospheres and interiors, brown dwarfs are frequently used to test theoretical models of giant planets. In fact, surveys for giant planets and free-floating brown dwarfs have demonstrated that they overlap in their mass distributions [4–6]. As a result, it is necessary to carefully examine the possible definitions for giant planets and brown dwarfs and how to distinguish between them observationally. In addition to representing a distinct class of substellar objects from giant planets, brown dwarfs may also harbor planetary systems of their own. Brown dwarfs are important in this regard because they offer an opportunity to test theories of planet formation in an extreme environment. As with stars, planet formation around brown dwarfs can be investigated by studying their circumstellar disks as well as searching for evidence of planetary companions.

7.2
Defining Brown Dwarfs and Giant Planets

Two competing schemes for defining brown dwarfs and giant planets have been widely discussed. Some studies have adopted the deuterium burning mass limit ($M \sim 0.014 M_{\odot}$ [7, 8]) as the boundary between planets and brown dwarfs [9]. These mass-based definitions are applicable to both companions and isolated objects. As a result, the least massive members of young clusters occasionally are referred to as *free-floating planets* [10]. Alternatively, brown dwarfs and giant planets can be defined according to their formation mechanisms. In this case, planets are born in circumstellar disks, while brown dwarfs arise from the fragmentation of molecular clouds.

Formation and Evolution of Exoplanets. Edited by Rory Barnes

ISBN: 978-3-527-40896-2

Applying either of these two schemes to substellar objects can be challenging because of uncertainties in the mass and the formation mechanism of a given substellar object. However, the formation-based definitions of brown dwarfs and giant planets do seem to be preferable. Unlike the fusion of ^{1}H, the presence or absence of deuterium fusion has little impact on the evolution of substellar objects [11]. Adopting the deuterium burning limit as a boundary between brown dwarfs and planets would be equivalent to assigning a separate name to high-mass stars because they undergo the CNO cycle. In fact, the study that is frequently cited for proposing the mass-based definitions did not intend to do so, as indicated when it stated that ''we have arbitrarily classed as 'planets' those objects that do not burn deuterium and as 'brown dwarfs' those that do burn deuterium, but not light hydrogen. While this distinction is physically motivated, we do not advocate abandoning the definition based on origin'' [9]. Therefore, brown dwarfs and giant planets are defined according to their formation in the following discussion.

7.3
Distinguishing between Brown Dwarfs and Giant Planets

The prospects for distinguishing between brown dwarfs and giant planets under the formation-based definitions are now examined. Free-floating objects and wide companions (>100 AU) that are found in star-forming regions and young clusters (τ < 10 megayears) are considered first. The two primary theories of giant planet formation, core accretion and disk instability, do not produce companions at distances greater than ~10 and ~100 AU, respectively [12–15]. Thus, young substellar companions beyond 100 AU probably did not form in their present locations within disks. It is unlikely that enough time has passed in these young populations for giant planets to form and to undergo dynamical ejection to wide orbits or out of their systems. In addition, many of the free-floating substellar members of young clusters possess circumstellar disks, which indicates that they were not born in the disks surrounding other stars or brown dwarfs (Section 7.4). Instead, according to various observations, young substellar objects found in isolation represent an extension of the star formation process down to at least ~5$M_{\rm Jup}$ [6]. Therefore, in general, free-floating objects and wide companions in young clusters can be confidently classified as brown dwarfs.

Among free-floating objects that have evolved beyond the epoch of giant planet formation (τ > 10 megayears), it is more difficult to distinguish between ejected giant planets and brown dwarfs. However, in practice, it has not been necessary to do so to date, given that the least massive free-floating members of the solar neighborhood [16, 17] have masses that are near the observed maximum of ~15$M_{\rm Jup}$ for close-in giant planetary companions [4]. When less massive objects are found in the field, atmospheric abundances may offer signatures of their formation mechanisms. For instance, the gas giants in our solar system appear to be enriched in heavy elements [18, 19], possibly due to the core-accretion process. Such enrichment is not expected for brown dwarfs that formed in isolation. Thus,

spectroscopy may constrain the formation mechanisms of both free-floating objects and companions [11].

Additional diagnostics are available for classifying brown dwarfs and giant planets among companions, such as the distribution of companion masses for given ranges of separations and primary masses. At separations less than 5 AU from solar-type stars, the distribution of companions exhibits a deficit of objects at $20–60 M_{\rm Jup}$, which has been referred to as the *brown dwarf desert* [20]. This feature implies the presence of two populations that formed through two distinct mechanisms, presumably giant planet formation in disks and binary star formation via the fragmentation of cloud cores. Thus, for close companions to stars like the Sun, the companion mass provides an indication of whether it is likely to be a giant planet or a brown dwarf. It is unknown whether two distinct populations appear in the distribution of companions at wider separations, which has been measured down to $\sim 10 M_{\rm Jup}$ [21]. The products of core accretion (or disk instability) and core fragmentation may overlap in mass and separation. For instance, it is possible that both mechanisms can make $8 M_{\rm Jup}$ companions at 10 AU. Thus, mass alone may not be sufficient to distinguish between giant planets and brown dwarfs for many ranges of separations and primary masses.

As with companion masses, the distributions of luminosities and radii of substellar companions may provide clues to their formation histories. For instance, giant planets produced by core accretion may be fainter than brown dwarfs at early stages of their lives because of details concerning the accretion shock for a planet forming by core accretion [22]. Another phenomenon in giant planets, namely, layered convection caused by abundance gradients, may also cause them to be fainter than fully convective brown dwarfs [23]. Thus, giant planets and brown dwarfs may inhabit distinct ranges of luminosity, temperature, and radii at young ages. Investigating this possibility will require companion searches that reach Jovian masses for large samples of young stars that are comparable in size to the samples of solar-type stars that have been previously observed with radial velocity monitoring.

7.4
Protoplanetary Disks around Brown Dwarfs

As discussed in Chapter 2, observations of primordial circumstellar disks provide fundamental constraints on the process of planet formation. In recent years, studies of disks have been extended into the substellar regime. This section reviews the properties of disks around brown dwarfs that are relevant to planet formation.

7.4.1
The Least Massive Objects with Disks

The most readily observed signature of a circumstellar disk is infrared (IR) emission in excess above that expected from a stellar photosphere. Most disks

around brown dwarfs produce significant excess emission only longward of ~4μm [24]. Using ground-based telescopes, it has been possible to obtain photometry at these wavelengths for only a few of the brightest and most massive young brown dwarfs [25, 26]. The best available sensitivity to mid-IR emission is currently offered by the *Spitzer Space Telescope* [27]. In the nearest star-forming regions ($d < 400$ pc), *Spitzer* has detected evidence of disks for objects down to masses of $\sim 8 M_{\rm Jup}$ [24, 28–31]. An example of the IR spectral energy distribution for one of these objects is shown in Figure 7.1. These data demonstrate that the raw materials for planet formation exist across the full range of brown dwarf masses observed to date.

7.4.2 Disk Fractions and Lifetimes

The amount of time available for the formation of giant planets is limited by the lifetime of a circumstellar disk. The typical lifetimes of disks can be estimated by measuring how common disks are among clusters that span a range of ages ($\tau \sim 1$–10 megayears). Because the *Spitzer Space Telescope* has a relatively large field of view (~5′) in addition to its sensitivity to faint sources, it has been possible to efficiently obtain accurate IR photometry for the substellar members of young clusters. *Spitzer* images for the nearest star-forming regions have been used to measure disk fractions of large samples of stars and brown dwarfs. The disk fractions for three clusters near an age of 3 megayears, IC 348, Chamaeleon I, and σ Ori, are shown in Figure 7.2 [24, 30–34]. These data are plotted as a function of

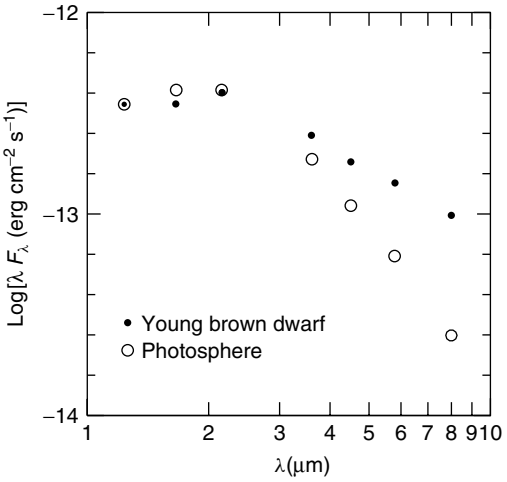


Figure 7.1 IR spectral energy distribution for the young brown dwarf Cha J11070768−7626326 (*filled circles*). This object exhibits excess emission at longer wavelengths relative to a stellar photosphere (*open circles*), making it one of the least massive brown dwarfs known to harbor a circumstellar disk ([24], $M \sim$ 4-10$M_{\rm Jup}$).

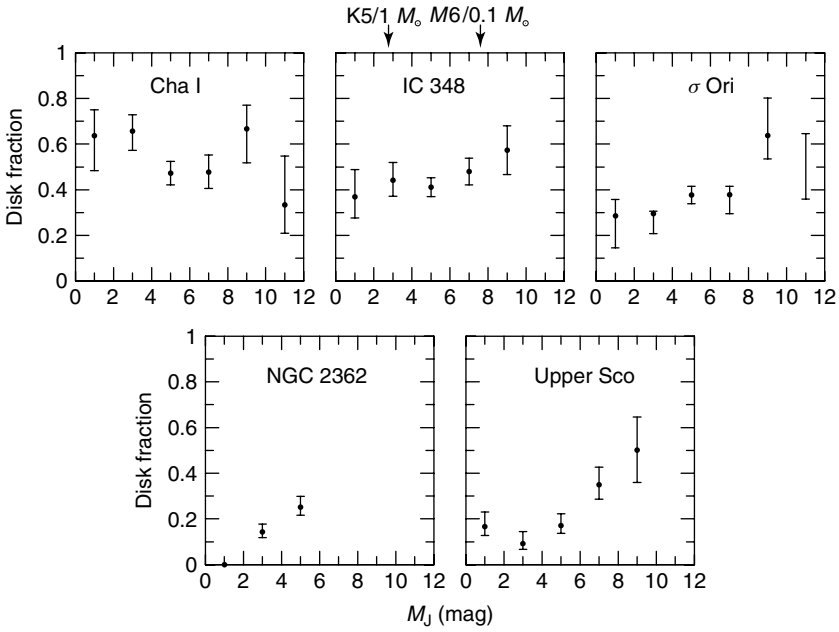


Figure 7.2 Disk fraction as a function of absolute *J*-band magnitude for three clusters near an age of 3 megayears (IC 348, Cha I, σ Ori) and two clusters near an age of 5 megayears (NGC 2362 and Upper Sco) [24, 30–37]. The typical magnitudes for spectral types of K5 ($\sim 1 M_\odot$) and M6 ($\sim 0.1 M_\odot$) are indicated.

absolute magnitude, which acts as a proxy for stellar mass. The disk fractions for the brown dwarfs are similar to or slightly higher than those of the low-mass stars. The disk fractions are also similar from across clusters, except for the solar-mass stars, which have decreasing disk fractions from Chamaeleon I to IC 348 to σ Ori. It is possible that this trend is related to the fact that the most massive star in each cluster has an earlier spectral type from the former to the latter (B9 to O9). To estimate the lifetimes of disks, these data must be compared to disk fractions of clusters at older ages. Thus, two populations with ages near 5 megayears, NGC 2362, and Upper Scorpius, are included in Figure 7.2 [35–37]. Although these older low-mass samples are small, particularly at substellar masses, this comparison between clusters at 3 and 5 megayears indicates that disk fractions decline less rapidly for low-mass stars and brown dwarfs than for solar-mass stars. This trend implies that the lifetimes of disks are longer for lower stellar masses. Quantitatively, estimating the lifetimes of disks at low masses requires large samples of low-mass stars and brown dwarfs at ages of 5–10 megayears, but unfortunately they do not exist. Only a few objects of this kind have been identified, which are members of small associations such as η Cha and TW Hya. For instance, the latter contains four known brown dwarfs, three of which have mid-IR measurements, and two of those

objects have disks [38–41]. These data demonstrate that brown dwarf disks can survive as long as disks around stars, but the sample size is too small for an estimate of the typical disk lifetime.

7.4.3 Disk Compositions

As in disks around stars, planet formation via core accretion in disks around brown dwarfs should begin with the growth of dust grains and their settling to the midplane of the disk (see Chapter 4). One of the primary diagnostics for characterizing dust processing is the silicate emission feature at 10 μm because its strength and shape depend on the size, composition, and structure of dust grains. The *Spitzer Space Telescope* has provided the sensitivity necessary for spectroscopy of this feature in disks around low-mass objects. A sample of *Spitzer* spectra for young low-mass stars and brown dwarfs is shown in Figure 7.3. The profiles of the silicate emission features in these data exhibit evidence of crystalline silicates

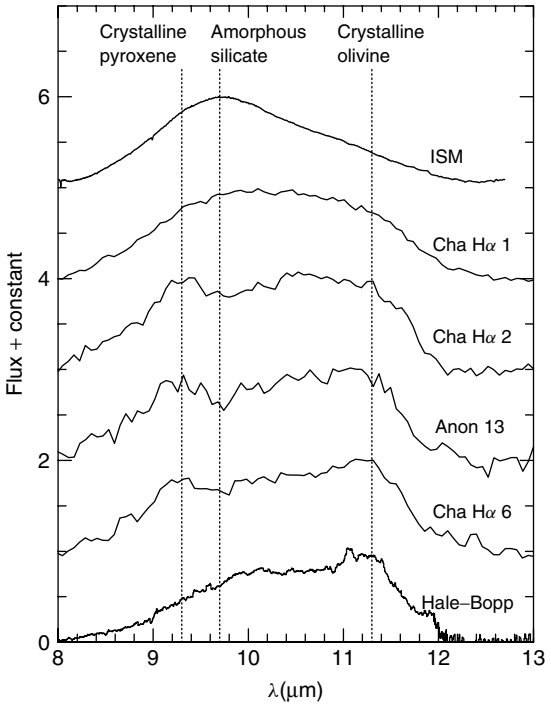


Figure 7.3 IR spectra of disks around young low-mass stars and brown dwarfs [42, 43]. Data for amorphous silicate-dominated interstellar medium and the crystalline-rich comet Hale-Bopp are included for comparison. The spectra have been continuum-subtracted and normalized to the peak emission in the range between 7.6 and 13.5 μm.

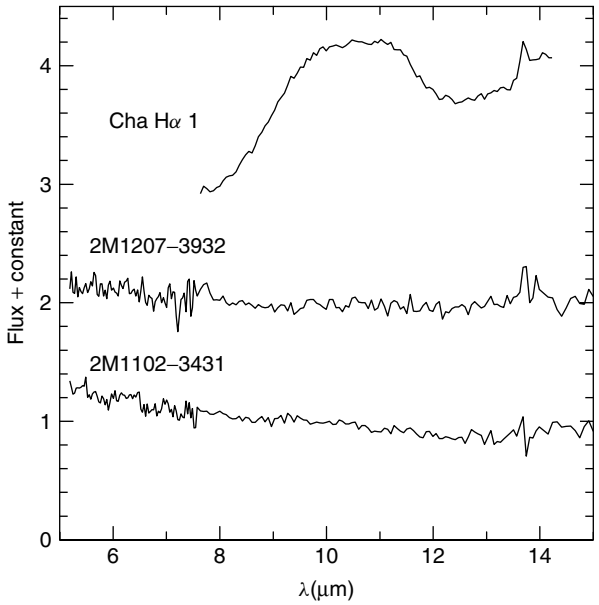


Figure 7.4 IR spectra of disks around brown dwarfs at ages of 3 megayears (Cha Hα 1) and 10 megayears (2M1207-3932, 2M1102-3431) [41, 43]. Silicate emission at 10 μm is absent in the older disks, indicating that the dust grains have grown larger than ∼5 μm. The spectra have been normalized at 8 μm.

and grain growth [42, 43]. In general, silicate emission is weaker for stars at lower masses, which may indicate that low-mass objects experience more advanced grain evolution, or that the disk radius at which silicate emission is produced depends on stellar luminosity [44, 45]. The strength of silicate emission also declines with age among low-mass objects in Taurus ($\tau \sim 1$ megayear), Chamaeleon I ($\tau \sim 3$ megayears), Upper Scorpius ($\tau \sim 5$ megayears), and TW Hya ($\tau \sim 10$ megayears) [37, 41–43], which probably reflects the growth of grains to larger sizes. The *Spitzer* spectra for the oldest known brown dwarf disks are shown in Figure 7.4. Silicate emission is absent from the spectra of these older disks, indicating that grains in the upper disk layers have grown to sizes larger than ∼5 μm. In comparison, silicate emission from small grains is still present in the spectra of disks around stars in TW Hya [46, 47]. Therefore, it appears that grains may grow faster in disks around brown dwarfs than in disks around stars.

7.4.4
Transitional Disks

The growth of dust grains and planetesimals may lead to an inside-out clearing of a circumstellar disk. A forming giant planet can eventually clear the dust from

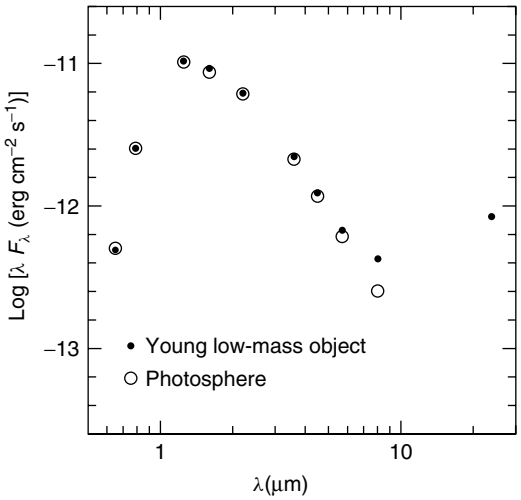


Figure 7.5 IR spectral energy distribution for a young low-mass object in the IC 348 cluster (*filled circles*) [53]. Relative to a stellar photosphere (*open circles*), this object exhibits excess emission only at wavelengths longward of ∼8 μm, indicating the presence of a disk with an inner hole.

the inner few astronomical units of the disk. The absence of dust close to the star results in a distinctive spectral energy distribution that exhibits photospheric colors at shorter IR wavelengths followed by a sudden onset of excess emission from the disk at longer wavelengths. Disks that produce this signature of an inner hole, known as *transitional disks*, have been found around several young stars [47–52]. Data from the *Spitzer Space Telescope* have proven to be particularly valuable for identifying transitional disks and studying them in detail. Indeed, observations with *Spitzer* have uncovered a transitional disk around an object near the hydrogen burning mass limit in the IC 348 cluster [53]. The spectral energy distribution for this source is shown in Figure 7.5. The absence of disk emission at shorter wavelengths for this object indicates the presence of an inner hole. On the basis of modeling of the data in Figure 7.5, the hole appears to have a radius of 0.5–1 AU. Several possible causes for the inner clearing of this disk have been investigated, and it appears that a planetary companion is the most promising candidate. Such a planet could have a mass in the range of $2.5–25 M_\oplus$ [53].

7.4.5 Disk Radii and Masses

The radii and masses of circumstellar disks represent two of the most fundamental initial conditions of planet formation given that they limit the sizes and total masses of planetary systems. Unfortunately, few measurements of these properties are available in the substellar regime because the small sizes and masses of

brown dwarf disks make them difficult to resolve and detect. Only one radius measurement for a brown dwarf disk has been performed to date, which was made possible by the fact that the system is close to edge-on [54]. A comparison of the extended emission detected from this disk with the *Hubble Space Telescope* to simulated images from disk models indicated that the disk radius is between 20 and 40 AU. That brown dwarf disk and a few others have been detected at millimeter wavelengths, providing measurements of their dust masses [55, 56]. Assuming a standard gas-to-dust ratio of 100, these millimeter data imply disk masses of $0.4–6 M_{\rm Jup}$. The ratio of disk mass to stellar mass is a few percent for these systems, which is similar to the typical fraction estimated for stars with disks.

7.4.6
Implications of Disk Properties for Planet Formation

Measurements of disk fractions in young stellar populations across a range of ages indicate that disk lifetimes may be longer at lower stellar masses. As a result, disks around low-mass stars and brown dwarfs may offer more time for planet formation to take place than disks around stars like the Sun. In addition, there is tentative evidence to suggest that grain growth occurs faster in brown dwarf disks than in stellar disks. These two trends bode well for planet formation around brown dwarfs. Indeed, a possible signature of a newly formed planet has been detected around a low-mass object in the form of a disk with an inner hole. The few millimeter measurements available for young brown dwarfs demonstrate that disk masses scale roughly with stellar masses, which is consistent with the observation that gas giants are less common around low-mass stars than around solar-type stars [57, 58]. However, brown dwarf disks do have sufficient masses for producing rocky planets. For instance, when the core-accretion model is applied to disks around brown dwarfs, planets with masses up to $5 M_{\oplus}$ can form at $\sim$1 AU [59] for some initial conditions. The distributions of planetary masses and separations depend on the assumed distributions of disk masses and surface densities, which are poorly constrained for brown dwarfs. For the one brown dwarf disk that has constraints of both mass and radius [54], the surface density is higher than assumed in the core-accretion models recently applied to brown dwarfs [59], which should correspond to more massive planets.

7.5
Searches for Planets around Brown Dwarfs

Surveys for planetary companions obviously have targeted brown dwarfs less extensively than stars, but the former are beginning to receive more attention as a part of efforts to fully explore the circumstances in which planets form. Because of their low luminosities, brown dwarfs are particularly amenable to detections of low-mass companions through direct imaging. This method has been applied to brown dwarfs in the solar neighborhood and the nearest young clusters [60],

uncovering a few companions that appear to have masses near $15M_{\rm Jup}$ and below [61–66]. Some of these objects have been referred to as *giant planets*, but their mass estimates are occasionally underestimated [67], and their separations and masses are probably too large for formation in disks around the primaries (Section 7.3, [59, 68]). Indeed, formation of these companions in a star-like fashion rather than in disks is perfectly plausible given that free-floating objects have been found in the same mass range [6]. Thus, it is likely that no bona fide giant planetary companions to brown dwarfs have been directly imaged, to date.

In addition to direct imaging, most of the other techniques developed for finding planets around stars can be extended to substellar primaries. Relative to a Sun-like star, a low-mass star or brown dwarf exhibits a larger reflex motion from a planetary companion at a given mass. To exploit this feature of low-mass primaries, radial velocity surveys have included low-mass stars in their samples, resulting in the discovery of some of the least massive known extrasolar planets [69, 70]. However, low-mass stars and (particularly) brown dwarfs are faint at the optical wavelengths considered by conventional Doppler surveys. As a result, planet searches are beginning to target young, relatively luminous primaries [71, 72] and employ spectroscopy at IR wavelengths where cool objects are brightest [73, 74]. For instance, a recent IR survey of late-M and L dwarfs achieved a precision of 300 m s^{-1}, corresponding to companions with $M \sin i > 2M_{\rm Jup}$ and $P < 3$ days [75]. Microlensing observations have discovered super-Earths around M dwarfs [76, 77] as well as a brown dwarf [78], which represents the first clear detection of a planet around a substellar primary. Finally, the application of transit surveys to low-mass stars [79–81] and brown dwarfs [82] has been investigated in detail. Because of the small diameters of these primaries, they offer the potential for detecting transiting Earth-mass planets.

References

1 Mayor, M. and Queloz, D. (**1995**) *Nature*, **378**, 355.

2 Marcy, G.W. and Butler, R.P. (**1996**) *The Astrophysical Journal*, **464**, L147.

3 Oppenheimer, B.R., Kulkarni, S.R., Nakajima, T., and Matthews, K. (**1995**) *Science*, **270**, 1478.

4 Marcy, G., Butler, R.P., Fischer, D. *et al.* (**2005**) *Progress of Theoretical Physics Supplement*, **158**, 24.

5 Udry, S. and Santos, N.C. (**2008**) *Annual Reviews of Astronomy and Astrophysics*, **45**, 397.

6 Luhman, K.L., Joergens, V., Lada, C. *et al.* (**2007**) *Protostars and Planets V*, University of Arizona Press, Tucson, p. 443.

7 Hubbard, W.B., Burrows, A., Guillot, T. *et al.* (**1996**) *The Astrophysical Journal*, **460**, 993.

8 Chabrier, G., Baraffe, I., Allard, F., and Hauschildt, P. (**2000**) *The Astrophysical Journal*, **542**, L119.

9 Burrows, A., Marley, M., Hubbard, W.B. *et al.* (**1997**) *The Astrophysical Journal*, **491**, 856.

10 Zapatero Osorio, M.R., Béjar, V.J.S., Martín, E.L. *et al.* (**2000**) *Science*, **290**, 103.

11 Chabrier, G., Baraffe, I., Selsis, F. *et al.* (**2007**) *Protostars and Planets V*, University of Arizona Press, Tucson, p. 623.

12 Pollack, J.B., Hubickyj, O., Bodenheimer, P. *et al.* (**1996**) *Icarus*, **124**, 62.
13 Inaba, S., Wetherill, G.W., and Ikoma, M. (**2003**) *Icarus*, **166**, 46.
14 Boss, A. (**2003**) *The Astrophysical Journal*, **599**, 577.
15 Boss, A. (**2006**) *The Astrophysical Journal*, **637**, L137.
16 Kirkpatrick, J.D., Barman, T.S., Burgasser, A.J. *et al.* (**2006**) *The Astrophysical Journal*, **639**, 1120.
17 Cruz, K.L., Reid, I.N., Kirkpatrick, J.D. *et al.* (**2007**) *The Astronomical Journal*, **133**, 439.
18 Lodders, K. (**2003**) *The Astrophysical Journal*, **591**, 1220.
19 Lodders, K. (**2004**) *The Astrophysical Journal*, **611**, 587.
20 Marcy, G.W. and Butler, R.P. (**2000**) *Publications of the Astronomical Society of the Pacific*, **112**, 137.
21 Metchev, S. and Hillenbrand, L. (**2009**) *The Astrophysical Journal Supplement Series*, **181**, 62.
22 Marley, M.S., Fortney, J.J., Hubickyj, O. *et al.* (**2007**) *The Astrophysical Journal*, **655**, 541.
23 Chabrier, G. and Baraffe, I. (**2007**) *The Astrophysical Journal*, **661**, L81.
24 Luhman, K.L., Allen, L.E., Allen, P.R. *et al.* (**2008**) *The Astrophysical Journal*, **675**, 1375.
25 Apai, D., Pascucci, I., Sterzik, M.F. *et al.* (**2004**) *Astronomy and Astrophysics*, **426**, L53.
26 Mohanty, S., Jayawardhana, R., Natta, A. *et al.* (**2004**) *The Astrophysical Journal*, **609**, L33.
27 Werner, M.W., Roellig, T.L., Low, F.J. *et al.* (**2004**) *The Astrophysical Journal Supplements*, **154**, 1.
28 Luhman, K.L., Adame, L., D'Alessio, P. *et al.* (**2005**) *The Astrophysical Journal*, **635**, L93.
29 Luhman, K.L. and Muench, A.A. (**2008**) *The Astrophysical Journal*, **684**, 654.
30 Scholz, A. and Jayawardhana, R. (**2008**) *The Astrophysical Journal*, **672**, L49.
31 Luhman, K.L., Hernandez, J., Downes, J.J. *et al.* (**2008**) *The Astrophysical Journal*, **688**, 362.
32 Lada, C.J., Muench, A.A., Luhman, K.L. *et al.* (**2006**) *The Astronomical Journal*, **131**, 1574.
33 Hernández, J., Hartmann, L., Megeath, T. *et al.* (**2007**) *The Astrophysical Journal*, **662**, 1067.
34 Luhman, K.L., Lada, C.J., Hartmann, L. *et al.* (**2005**) *The Astrophysical Journal*, **631**, L69.
35 Carpenter, J.M., Mamajek, E.E., Hillenbrand, L.A., and Meyer, M.R. (**2006**) *The Astrophysical Journal*, **651**, L49.
36 Dahm, S.E. and Hillenbrand, L.A. (**2007**) *The Astronomical Journal*, **133**, 2072.
37 Scholz, A., Jayawardhana, R., Wood, K. *et al.* (**2007**) *The Astrophysical Journal*, **660**, 1517.
38 Jayawardhana, R., Ardila, D.R., Stelzer, B., and Haisch, K.E. (**2003**) *The Astromical Journal*, **126**, 1515.
39 Sterzik, M.F., Pascucci, I., Apai, D. *et al.* (**2004**) *Astronomy and Astrophysics*, **427**, 245.
40 Riaz, B., Gizis, J.E., and Hmiel, A. (**2006**) *The Astrophysical Journal*, **639**, L79.
41 Morrow, A.L. Luhman, K.L., Espaillat, C. *et al.* (**2008**) *The Astrophysical Journal*, **676**, L143.
42 Furlan, E., Calvet, N., D'Alessio, P. *et al.* (**2005**) *The Astrophysical Journal*, **621**, L129.
43 Apai, D., Pascucci, I., Bouwman, J. *et al.* (**2005**) *Science*, **310**, 834.
44 Kessler-Silacci, J.E., Dullemond, C.P., Augereau, J.-C. *et al.* (**2007**) *The Astrophysical Journal*, **659**, 680.
45 Sicilia-Aguilar, A., Hartmann, L.W., Watson, D. *et al.* (**2007**) *The Astrophysical Journal*, **659**, 1637.
46 Uchida, K.I., Calvet, N., Hartmann, L. *et al.* (**2004**) *The Astrophysical Journal Supplements*, **154**, 439.
47 Furlan, E., Sargent, B., Calvet, N. *et al.* (**2007**) *The Astrophysical Journal*, **664**, 1176.
48 Calvet, N., D'Alessio, P., Hartmann, L. *et al.* (**2002**) *The Astrophysical Journal*, **568**, 1008.
49 Calvet, N., D'Alessio, P., Watson, D.M. *et al.* (**2005**) *The Astrophysical Journal*, **630**, L185.
50 D'Alessio, P., Hartmann, L., Calvet, N. *et al.* (**2005**) *The Astrophysical Journal*, **621**, 461.

51 Espaillat, C., Calvet, N., D'Alessio, P. *et al.* (**2007**) *The Astrophysical Journal*, **664**, L111.

52 Espaillat, C., Calvet, N., D'Alessio, P. *et al.* (**2007**) *The Astrophysical Journal*, **670**, L135.

53 Muzerolle, J., Adame, L., D'Alessio, P. *et al.* (**2006**) *The Astrophysical Journal*, **643**, 1003.

54 Luhman, K.L., Adame, L., D'Alessio, P. *et al.* (**2007**) *The Astrophysical Journal*, **666**, 1219.

55 Klein, R., Apai, D., Pascucci, I. *et al.* (**2003**) *The Astrophysical Journal*, **593**, L57.

56 Scholz, A., Jayawardhana, R., and Wood, K. (**2006**) *The Astrophysical Journal*, **645**, 1498.

57 Endl, M., Cochran, W.D., Kürster, M. *et al.* (**2006**) *The Astrophysical Journal*, **649**, 436.

58 Johnson, J.A., Butler, R.P., Marcy, G. *et al.* (**2007**) *The Astrophysical Journal*, **670**, 833.

59 Payne, M.J. and Lodato, G. (**2007**) *Monthly Notices of the Royal Astronomical Society*, **381**, 1597.

60 Burgasser, A.J., Reid, I.N., Siegler, N. *et al.* (**2007**) *Protostars and Planets V*, University of Arizona Press, Tucson, p. 427.

61 Chauvin, G., Lagrange, A.-M., Dumas, C. *et al.* (**2004**) *Astronomy and Astrophysics*, **425**, L29.

62 Allers, K.N. (**2006**) PhD thesis. Disks and dissociation regions: The interaction of young stellar objects with their environments, University of Texas, Austin.

63 Close, L.M., Zuckerman, B., Song, I. *et al.* (**2007**) *The Astrophysical Journal*, **660**, 1492.

64 Jayawardhana, R. and Ivanov, V.D. (**2006**) *Science*, **313**, 1279.

65 Béjar, V.J.S., Zapatero Osorio, M.R., Pérez-Garrido, A. *et al.* (**2008**) *The Astrophysical Journal*, **673**, L185.

66 Luhman, K.L., Mamajek, E.E., Allen, P.R. *et al.* (**2009**) *The Astrophysical Journal*, **691**, 1265.

67 Luhman, K.L., Allers, K.N., Jaffe, D.T. *et al.* (**2007**) *The Astrophysical Journal*, **659**, 1629.

68 Lodato, G., Delgado-Donate, E., and Clarke, C.J. (**2005**) *Monthly Notices of the Royal Astronomical Society*, **364**, L91.

69 Rivera, E.J., Lissauer, J.J., Butler, R.P. *et al.* (**2005**) *The Astrophysical Journal*, **634**, 625.

70 Udry, S., Bonfils, X., Delfosse, X. *et al.* (**2007**) *Astronomy and Astrophysics*, **469**, L43.

71 Joergens, V. (**2006**) *Astronomy and Astrophysics*, **446**, 1165.

72 Joergens, V. and Müller, A. (**2007**) *The Astrophysical Journal*, **666**, L113.

73 Martín, E.L., Guenther, E., Zapatero Osorio, M.R. *et al.* (**2006**) *The Astrophysical Journal*, **644**, L75.

74 Ramsey, L.W., Barnes, J., Redman, S.L. *et al.* (**2008**) *Publications of the Astronomical Society of the Pacific*, **120**, 887.

75 Blake, C.H., Charbonneau, D., White, R.J. *et al.* (**2007**) *The Astrophysical Journal*, **666**, 1198.

76 Beaulieu, J.-P., Bennett, D.P., Fouqué, P. *et al.* (**2006**) *Nature*, **439**, 437.

77 Gould, A., Udalski, A., An, D. *et al.* (**2006**) *The Astrophysical Journal*, **644**, L37.

78 Bennett, D.P., Bond, I.A., Udalski, A. *et al.* (**2008**) *The Astrophysical Journal*, **684**, 663.

79 Gould, A., Pepper, J., and DePoy, D.L. (**2003**) *The Astrophysical Journal*, **594**, 533.

80 Nutzman, P. and Charbonneau, D. (**2008**) *Publications of the Astronomical Society of the Pacific*, **120**, 317.

81 Plavchan, P., Jura, M., and Kirkpatrick, J.D. (**2008**) *The Astrophysical Journal Supplement Series*, **175**, 191.

82 Blake, C.H., Bloom, J.S., Latham, D. *et al.* (**2008**) *Publications of the Astronomical Society of the Pacific*, **120**, 860.

8
Exoplanet Chemistry

Katharina Lodders

8.1
A Goodly Gallery of Planets

The current state of observations on the chemical nature of exoplanets is comparable to that of the early days of solar-system planetary science, when not much more than the orbits and densities of planets were known. Information about exoplanet chemical compositions from direct observations is still limited. However, chemical and physical observations and models for planets in our solar system developed since the 1950s should provide the tools to understand other planets, if models for solar-system planets are indeed correct. As such, the observational data to come in the following years undoubtedly will test the fine-tuned planetary formation and evolution models that seem to work well for understanding the planets in our solar system.

There are four broad basic types in our gallery of planets in the solar system and very likely in most of the other planetary systems. The reason for this is the first-order similarity of the relative elemental abundances of most stars in the galactic disk and the compounds that the elements can form. Planets originate from the materials in a stellar accretion disk that had a composition similar to that of the final star itself. This limits the possible chemical nature of the building blocks of planets, despite the wide range of possible physical properties in the different accretion disks around other stars. Different condensed phases in an accretion disk are stable within different temperature limits; refractory phases are stable up to high temperatures, and volatile phases are only stable at relatively low temperatures. Thus, the formation or thermal destruction of solids at different temperatures can lead to a fractionation of refractory elements in the solids from the more volatile elements that are in the gas.

Figure 8.1 shows how the four types can be categorized by their chemical nature. Terrestrial-like planets, defined in the following, are composed mainly of rocky substances, and are represented in our solar system by Mercury, Venus, Earth, Mars, Ceres and, although not planets, by a large number of asteroids. Terrestrial planets can have atmospheres which are, however, insignificant to the total planetary mass budget. The second type consists of planets largely made of rocky and icy materials

Formation and Evolution of Exoplanets. Edited by Rory Barnes

ISBN: 978-3-527-40896-2

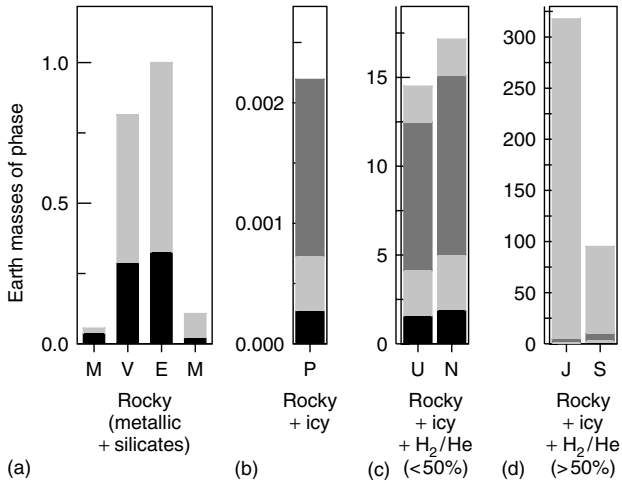


Figure 8.1 The four chemical types of planets represented by solar-system objects. The shading indicates the contributions (in Earth masses) of different phases: black = metallic, light gray = silicate, dark gray = various ices, second light gray on top of ices is H_2+ He gas. (a) Rocky (terrestrial) planets mainly consist of silicates and metallic FeNi(S). (b) Rock plus ice (Plutonian) planets primarily consist of rocky and icy solids. (c) Giant (Neptunian) planets consist mainly of rocky and icy substances and less than half is He and H_2. (d) Gas-giant (Jovian) planets are dominated in mass by He and H_2. Data from Lodders and Fegley [1].

(e.g., H_2O, CH_4, N_2 ices), which may have tenuous atmospheres composed of icy evaporates. In the solar system, such *Plutonian* planets, as one would call them, are only represented by IAU-defined dwarf planets such as Pluto, Eris, Haumea, and Makemake, but most objects beyond Neptune should qualify for this chemical distinction. Some of the asteroids in the outer asteroid belt may also belong to the rocky–icy objects. Giant (Neptunian) planets consist mainly of compounds derived from rocky and icy substances in the accretion disk and contributions from He and H_2 to their total masses are considerable but remain below 50%. Uranus and Neptune could only accrete about 10–20% of their total mass as H and He gas, which makes them ''gas-poor'' giant planets, the third type. The fourth type, the Jovian gas-rich giant planets in which H and He dominate, is represented by Jupiter and Saturn, and they have the closest similarity in elemental composition to their host star, the Sun.

Silicates and metallic phases define the rocky portion of a planet, however, their proportions are reasonably well known only for the terrestrial planets. For all other objects in Figure 8.1, a silicate/metallic mass ratio according to solar elemental abundances and geochemical considerations was applied to the ''rocky'' portion. For the giant planets, similar considerations were used to divide the heavy element fractions into ''metallic, silicate, and icy'' components.

Most known exoplanets have masses between that of Saturn and 10 times that of Jupiter ($1M_{\rm Jup} = \sim 318$ Earth masses, M_E); but smaller Neptune size objects ($\sim 0.05M_{\rm Jup} = \sim 17M_E$), and objects of a few Earth masses ($1M_E = 0.003M_{\rm Jup}$) are

also known (see the frequent exoplanets updates on Jean Schneider's webpage at http://exoplanet.eu). A few companions to stars reach masses above $\sim 13 M_{\rm Jup}$, where theoretical models predict that interior deuterium burning is possible. This mass limit is usually taken to divide ''planets'' from ''brown dwarfs'' that are also observed to orbit stars. However, this definition may be arbitrary (see Chapter 7). To make this point, consider the object HD41004B b, which orbits its early M dwarf host star at 0.0177 AU with an orbital period of ~1.3 days. This is one of the closest orbits of a nonstellar object around a star currently known. However, HD41004B b has a mass of $18.4 M_{\rm Jup}$–clearly above the D-burning mass limit, and is not really counted as a planet.

The atmospheric gas and cloud chemistry in brown dwarfs is not much different from that in gas-giant exoplanets if these objects have similar effective temperatures (see below). The major difference would be the absence of deuterated gases such as CH_3D, HDO, NH_2D, and HD in brown dwarfs that masquerade as huge planets. If observationally feasible, the presence or absence of D-bearing gases may be a means to distinguish large planets from brown dwarfs, but more importantly, it would finally provide observational measures of the minimum mass required for D-burning.

There is no other practical means to distinguish these objects other than by mass. Differences in possible formation modes could also distinguish these larger objects. Gas-giant planets and brown dwarfs can both form by gravitational instabilities in a protostellar accretion disk (Chapter 5). Another prominent formation model for giant planets is ''core accretion'' (Chapter 5). This is accretion of larger protocores and subsequent gravitational capture of large quantities of gas (see below), but this mechanism may not work for making massive brown dwarfs in a timely fashion in a short-lived accretion disk. In principle, the different modes of formation can imprint different chemical compositions onto a gas-giant planet (see below). However, even the interpretation of the observed chemistry of the well-studied gas-giant planets Jupiter and Saturn is not unique. Currently, a large number of researchers seem to favor the core-accretion model for Jupiter's and Saturn's origins, but the case is not clear-cut. Hence, the much more limited chemical information that is available so far for gas-giant planets outside the solar system cannot be used to distinguish planets and brown dwarfs by their formation mode.

8.2
Elemental Ingredients of Planets

The densities of exoplanets can be found if their radii and masses can be determined. Radial velocity measurements provide masses (or upper mass limits M sin *i*) for exoplanets, and radius determinations are available for planets observable during transits of their host star (see Chapter 1). Once planet densities are available, planetary overall compositions may be estimated from the chemical and physical properties of plausible components. Several models for the evolution of gas-giant

exoplanets and their interiors and spectra have been described by, e.g., Burrows *et al.* [2–4], Charbonneau *et al.* [5], Chabrier *et al.* [6], Fortney *et al.* [7–9], Guillot [10], Guillot *et al.* [11, 12], Sudarsky *et al.* [13]. A wide range of mass–radius relations for terrestrial to gas-giant planets is discussed by Fortney *et al.* [14]. There are also several different structure types of terrestrial-like planets that have been modeled, based on different components that may result from solar-composition element mixtures, e.g., Fortney *et al.* [14], Seager *et al.* [15].

Ultimately, it is the abundance of the chemical elements and the physical properties of their compounds that govern the constitution of a planet. As already mentioned, most stars in the solar neighborhood have relative elemental abundances similar to that of the Sun. Hence, the compounds and their relative amounts available in the accretion disks around other stars should be similar to what was present in the solar planetary accretion disk (i.e., the solar nebula). There are, of course, limitations to this first-order approach because the *distribution* of these compounds in a disk is expected to vary as physical disk properties vary. Planets not only develop around forming G2V dwarf stars like the Sun but are also found around lower mass F, K, and M dwarfs, and around stars with up to ~4.5 solar masses that have evolved into G and K giants. Detections of planets around very massive and hotter dwarf stars are not yet reported. Planet formation around early type (O, B, and A-type) dwarfs may be difficult to achieve as estimated planetary formation and evolution times of a few to tens of million years may coincide with stellar lifetimes, and early mass loss from massive stars may interfere with planet formation in an accretion disk.

The estimated minimum mass of the solar nebula is about $0.01–0.02 M_{sun}$, depending on the solar abundance compilation used and the adopted heavy element fractions for the giant planets. Accretion rates through the disk onto the young Sun are estimated as $\sim 10^{-5}$ to $10^{-7} M_{sun}$ per year, depending on the accretion stage. However, accretion disks around other stars may have had total masses different from that of the solar nebula, accretion rates different from that of their parental molecular clouds, and thus different thermal gradients and evolution timescales of their disks. Disk lifetimes are important because the formation of gas-giant planets requires that H_2 and He-rich gas is accreted before the disk dissipates. Meteoritic evidence from short-lived radioactive nuclides and observations of accretion disks in young stellar system suggest that disk lifetimes are at least two to three million years, with an upper limit of about 10 million years, which is then the maximum time available for gas-giant planet formation. On the other hand, the disk lifetimes and the presence of nebula gas may not be as critical for the formation of terrestrial-type planets, which can form as long as rocky planetesimals are abundantly available. The formation of the moon by a giant impact occurred after core formation on the Earth was complete, which was within 30 million years after the oldest known components in meteorites had been processed in solar nebula gas ~4600 million years ago [16]. The craters on Moon, Mars, and Mercury are witness to the late bombardment period which lasted until ~3900 million years ago. Clearly, formation of the Moon and late planetary

bombardments are the larger finalizing events in terrestrial planet accretion, but their timing is outside any reasonable lifetimes of the solar nebula.

Depending on the location of a star in the galactic disk and on the timing of star formation, a star and its accretion disk may also have had metallicity different from that of the Sun (i.e., elements heavier than He have different abundances relative to H). This additional factor influences the outcome of a planet's final constitution. The metallicity of a star is important because it is observed that metal-rich stars are more likely to harbor planets than stars of lower metallicity [17]. Most of the known exoplanets are likely to be gas giants, but one should expect that there is a similar preference for the occurrence of rocky planets around metal-rich stars. The formation of planets is plausibly fostered by a metal-rich disk because a higher mass density of the ''metallic'' elements (in the astronomer's sense) provides more rocky and icy building blocks, which should increase the probability of planet accretion.

8.2.1
Diagnostics from Elemental Abundance Fractionations

The *elemental* composition of a gas-giant planet compared to the composition of its host star may reveal more about its mode of formation, and again the relative amounts of the heavy elements are of particular interest here. A gas-giant planet that is essentially identical in elemental composition to its host star is likely to have formed through gravitational instability in the protostellar accretion disk ([18, 19]; Chapter 4). On the other hand, in the core-accretion model for giant planet formation, a solid protocore of $\sim 5-10 M_E$ has to form first, which is large enough to gravitationally capture surrounding nebular H_2 and He gas ([20–22], see also [23], and Chapter 5). This leads to a gas-giant planet rich in elements heavier than He and H when compared to the composition of the primary star if formation of the initial core required accumulation of solids from wider regions in the disk but gas was only accreted locally. (A similar result is obtained when incomplete accretion of He and H_2 fails to add to the rocky core abundances to get back abundances of the host star.) This model also leads to the notion that all giant planets formed by this mechanism should have ''cores,'' or more precisely, the equivalent mass of heavy elements distributed somewhere in their interior [24].

In addition to various accretion scenarios that separate solid components from the He–H_2 gas, chemical fractionations among heavy elements themselves in the disk can cause nonsolar abundance ratios of the heavy elements found in planets. For example, solar composition has a mass ratio of C/Si $\sim$ 3, whereas the entire Earth has a ratio of $\sim$0.0003. This is a consequence of the different volatilities of the chemical compounds formed by C and Si. The most likely place of fractionation of C and Si was in the solar nebula, and the Earth could not accrete much C and other volatile-element-bearing compounds to begin with.

8.3 Planetary Building Blocks

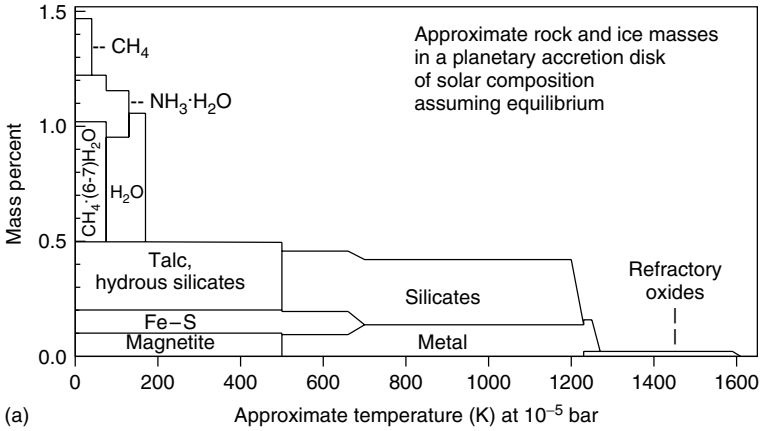
The basic results of chemistry in a H- and He-rich gas, such as the solar nebula, remain applicable for systems that evolve with solar or close-to-solar overall elemental compositions, if temperatures and densities (or total pressures) are not dramatically different. It is well known from condensation calculations that over a wide range of total pressures (i.e., $P < 10^{-3}$ bars), the chemistry in a very H_2 and He-rich, solar-like gas leads to quite similar minerals and thermal stability sequences ([25–27] and references therein). For example, usually the most thermally stable solid containing most of the Fe is an FeNi metal alloy; for Mg, it is the Mg-silicate forsterite (Mg_2SiO_4), and for sulfur, it is troilite, FeS. Hence, the available planetary building blocks in different disks should be comprised of quite similar oxides, silicates, metals, sulfides, and ices. Aside from the thermal stabilities, the quantities of the building blocks are limited by the abundances of the elements that make them. This simplifies the modeling of possible exoplanet compositions, even if the exact conditions in a former accretion disk and the distribution of the solids are not known.

The radial distribution of the rocky, icy, and gaseous ingredients for the formation of planets is determined by the temperature and total pressure gradients in accretion disks. Therefore, even in the absence of any possible turbulent redistribution of constituents, the available proportions of the solid planetary building blocks vary with radial location, which will affect the locations, masses, types, and number of resulting planets.

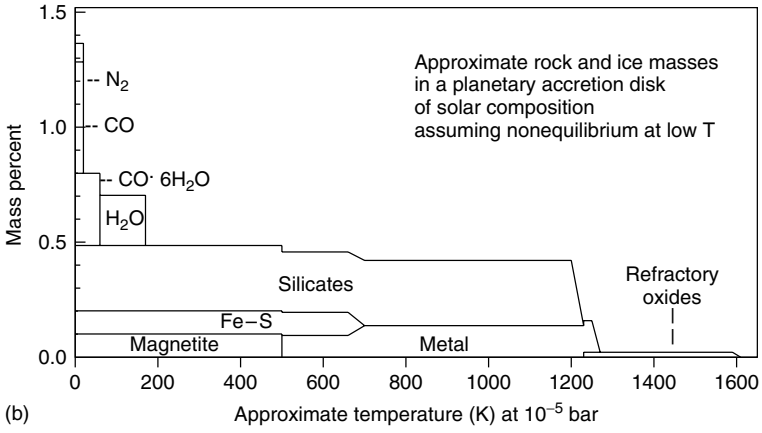
Three principal types of compounds, loosely referred to as *rock*, *ice*, and *gas*, are available in accretion disks, and can wind up in planets. When discussing planetary compositions, the terminology rock, ice, and gas is often kept, but it may relate to the nature of the phases that have accreted, and not necessarily to the nature of phases that occurs in a planet today. The relative proportions of gas, rock, and ice in the accretion disk determine a planet's possible mass but the equations of state determine the planet's size and which stable phases will be present.

The gas component is mainly H_2 and He and includes all the remaining mass that is left after rocky and icy condensates are removed. Depending on temperature and the details of the accretion disk chemistry of ices, inert gases (noble gases, N_2) can contribute a tiny mass to the gas. Since H and He make ~98.5% of all mass in solar-system material, these elements must make up the majority of mass in any gas-giant planet (this assumes that there were no changes in the H- and He-content from postaccretion processing of a giant planet that has migrated too close to its host star).

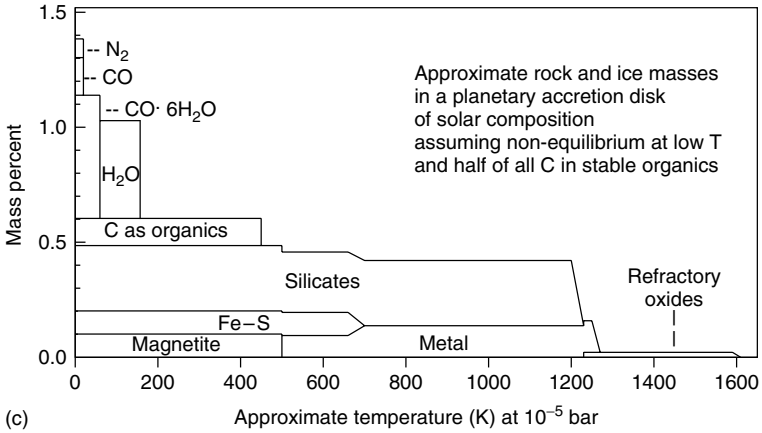
The rocky and icy components (noble gases inclusive) make up ~1.5% of the total mass. Figure 8.2 shows the mass percentages in a solar elemental composition system for various rocky and icy substances as a function of temperature. The three panels correspond to different assumptions of thermodynamic equilibrium, which mainly affect the amount and type of phases present at low temperatures.


Approximate rock and ice masses
in a planetary accretion disk
of solar composition
assuming equilibrium
-- CH_4
-- $NH_3 \cdot H_2O$
$CH_4 \cdot (6\text{-}7)H_2O$
H_2O
Talc,
hydrous silicates
Fe–S
Magnetite
Silicates
Metal
Refractory
oxides
Mass percent
1.5
1.0
0.5
0.0
0
200
400
600
800
1000
1200
1400
1600
Approximate temperature (K) at 10^{-5} bar

(a)


Approximate rock and ice masses
in a planetary accretion disk
of solar composition
assuming nonequilibrium at low T
-- N_2
-- CO
-- $CO \cdot 6H_2O$
H_2O
Silicates
Fe–S
Magnetite
Metal
Refractory
oxides
Mass percent
1.5
1.0
0.5
0.0
0
200
400
600
800
1000
1200
1400
1600
Approximate temperature (K) at 10^{-5} bar

(b)


Approximate rock and ice masses
in a planetary accretion disk
of solar composition
assuming non-equilibrium at low T
and half of all C in stable organics
-- N_2
-- CO
-- $CO \cdot 6H_2O$
H_2O
C as organics
Silicates
Fe–S
Magnetite
Metal
Refractory
oxides
Mass percent
1.5
1.0
0.5
0.0
0
200
400
600
800
1000
1200
1400
1600
Approximate temperature (K) at 10^{-5} bar

(c)

The high-temperature rock includes elements like Si, Mg, Ca, Al, Ti, etc., which form oxides and silicate minerals. Here, FeNi metal alloys and iron sulfide also belong to the "rock" in the environs of planetary accretion disks. Formulas and names of major minerals commonly encountered in planetary science and some mineral systematics are listed in Table 8.1. A good working definition of "rock" is a component containing the approximate elemental abundances and consisting of the mineral phases that are observed in undifferentiated meteorites called *chondrites*. All rocky compounds are relatively refractory and require high temperatures for evaporation or they readily condense out of a gas of solar-system composition at relatively high temperatures (see below for condensation). Since rocky material is the last material to evaporate or the first to condense, it is always present among possible solids in an accretion disk. Planets primarily made of such rocky materials – the terrestrial(-like) planets – characterize the chemical makeup, but not necessarily the size of such planets. The solar-system abundance of the elements limits the amount of rocky material to ∼0.5% of all mass in an accretion disk.

At temperatures below ∼600 K, gas–solid reactions become sluggish and equilibrium may not be reached. As discussed in detail by Fegley [28], under equilibrium conditions, one may expect the formation of sulfide and magnetite from metal, and hydrated silicates from pyroxenes and olivine. However, considering reaction kinetics, only sulfide formation is quantitative within the lifetimes of solar nebula, whereas magnetite formation is a borderline case. The hydration of silicates, usually associated with hydrothermal, high-pressure environs, in the "vacuum" of accretion disks can be ruled out since this requires reaction times that are many orders of magnitude above estimated lifetimes of these disks. Therefore, the nonequilibrium cases in Figure 8.2 do not show hydrous silicates as possible planet-building phases.

Icy materials become stable below ∼200 K, and encompass substances containing C, N, and O. The solids that more or less quantitatively sequester C, N, and O only become stable at low temperatures in low-pressure solar-composition systems. Hence, these low-temperature phases are collectively called "*ice*." Water ice is the most important ice because O is the third most abundant element in the solar system, and water ice is the most refractory ice among these ices. (Oxygen is also present in rocky materials but the other elements in rock are only abundant enough to bind ∼20% of all available solar-system oxygen).

In addition to water ice, other possible abundant ices are solid methane, CH_4; methane clathrate $CH_4 \cdot nH_2O$; carbon monoxide, CO; carbon monoxide clathrate, $CO{\cdot}nH_2O$; carbon dioxide CO_2; carbon dioxide clathrate, $CO_2 \cdot nH_2O$; ammonia, NH_3, ammonia monohydrate $NH_3{\cdot}H_2O$; molecular nitrogen ice, N_2; and nitrogen clathrate, $N_2 \cdot nH_2O$. In the clathrate hydrates, with the generic formula $X{\cdot}nH_2O$,

Figure 8.2 The distribution of solid phases as a function of temperature at 10^{-5} bar total pressure. The amounts shown are mass percent for solar-system composition. (a) Thermodynamic equilibrium case, which is limited by slow gas–gas and gas–solid reaction kinetics at low temperatures. (b and c) Possible (and nonunique) low-temperature nonequilibrium phases and distributions.

Table 8.1 Typical minerals made of more abundant elements.

Mineral group	End-member mineral	Ideal formula
Olivine $(Mg,Fe)_2SiO_4$	Forsterite	Mg_2SiO_4
	Fayalite	Fe_2SiO_4
Pyroxene $(Mg,Fe,Ca)SiO_3$	Enstatite	$MgSiO_3$
	Ferrosilite	$FeSiO_3$
	Wollastonite	$CaSiO_3$
Feldspar	Anorthite	$CaAl_2Si_2O_8$
	Albite	$NaAlSi_3O_8$
	Orthoclase	$KalSi_3O_8$
Metal alloys	Iron–nickel	FeNi
Sulfides	Troilite	FeS
	Pyrrotite	$Fe_{1-x}S$
Oxides	Magnetite	Fe_3O_4
Hydrous silicates	Talc	$Mg_3(Si_4O_{10})(OH)_2$

$n = 5–7$ is the number of water molecules that characterize a unit structure into which another molecule "X" can be retained as a "guest", which is why clathrates are also called *cage compounds*.

One may add solid organic substances to the group of "ices." High-polymer organics can be stable up to temperatures of 400–600 K [29], which makes them more refractory than water ice.

Gas reactions involving very stable molecules such as CO and N_2 are unlikely to reach thermodynamic equilibrium (see [30, 31]). This influences the type of low-temperature C and N gases, which, in turn, determines the type and amount of the ices that can form. In addition, the available proportions of the different ices depend on the relative thermal stability of these ices.

The top graph in Figure 8.2 is the equilibrium case where all CO converts to CH_4 gas, which then condenses as methane clathrate and CH_4 ice. Both phases form because the amount of water is insufficient to clathrate all CH_4 gas as methane clathrate (solar abundances give C : O about 1 : 2 but a ratio of 1 : (6–7) is needed for methane clathrate). The thermodynamically stable low-temperature gas for nitrogen is ammonia gas, which condenses as ammonia hydrate, $NH_3 \cdot H_2O$.

The formation of methane ice(s) requires that the very stable CO gas, dominant at high temperatures (e.g., >1500 K at 10^{-5} bars), was efficiently converted into methane gas at low temperatures, which is unfeasible within typical lifetimes of the disks (see [30]). With inefficient conversion, CO gas remains, and CO clathrates and CO ice condense (Figure 8.2b). Some CO (several percent) is also converted to CO_2 since this reaction is kinetically favored and can proceed down to lower temperatures (800–1000 K) at low total pressures. Thus, smaller amounts of CO_2-bearing ices (equivalent to ~10% of all O) can be present [30, 31], which, for clarity, are not shown in the nonequilibrium case panels in Figure 8.2. Slow gas-phase kinetics also prevents the conversion of N_2 to NH_3 and N_2 gas remains. This leads to N_2-clathrate ice (see [30, 31]).

Another possibility is that interstellar low-temperature organic substances remained unprocessed or that such substances formed in the outer planetary disks through catalyzed reactions such as in the Fischer–Tropsch-type process, where Fe-metal acts as a catalyst to convert CO and H_2 to hydrocarbons. This reduces the amounts of other C-bearing gases and ices.

To first approximation, the exact C-, N-, and O-bearing phases are irrelevant for mass balance as long as all C, N, and nonsilicate O are condensed as ices. Condensation of all C, N, and nonsilicate O into ices makes about ~1% of all material in a solar-composition system. Water ice alone makes up to ~0.6%, depending on whether all remaining O after rock condensation is in H_2O gas or remains in CO gas. However, full retention of the icy phases requires temperatures less than ~30 K, and one can expect that the most volatile ices such as CO, CH_4, and N_2 ice are easily evaporated during formation and evolution of the accretion disk.

Planets may form by accreting varying amounts of icy and rocky materials but this does not mean that rocks and ices remain in a state of ''ice'' or ''rock'' as one typically associates with it. Occasionally, it is convenient to collectively equate the amounts of C, N, and O as ''ice'' in planetary materials, which reflects that these elements may have been brought to the planet in low-temperature phases. Similarly, ''rock'' is a convenient summary of all other condensed phases other than ''ice''. However, temperature and pressure conditions during planetary accretion and subsequent planetary evolution can be quite different from those under which these phases are stable in a low-temperature, low-pressure accretion disk.

Only small planetary objects such as Kuiper belt objects (dormant comets?) in the outer solar system are most likely to have preserved their originally accreted icy phases. However, photochemical reactions are known to produce chemical and physical changes on surfaces of ''icy'' objects. On the other hand, icy phases and rock can thermally decompose during the larger scale differentiation of a planet, or react with other substances once they accreted into larger planetary objects (e.g., reactions of rocky materials with H_2 upon accretion into a gas-giant atmosphere).

During accretion of smaller kilometer-size objects to Mars- and Earth-sized objects, enough gravitational energy is released to allow melting and (partial) evaporation of accreting solids, even if about half the gravitational energy is lost through radiation. The point is that phases present in the accreting bodies may break down, and the elements from such phases react and redistribute into phases that are stable under conditions of the planetary environment. The separation of ''rocky'' elements into the silicate and metal portions in the Earth and other terrestrial planets is a well-known example. Another result of planetary differentiation is the redistribution of rocky elements within the Earth's silicate portion into a crust, the upper mantle of mainly olivine, and a lower mantle containing larger amounts of a high-pressure modification of $(Mg,Fe)SiO_3$ called *silicate perovskite* (not to be confused with the Ca–Ti mineral perovskite which has a similar structure at normal pressure). By analogy to meteorites, the prime candidates for constituents of the metallic cores in terrestrial planets are the ''rocky'' elements Fe, Ni, and S, which transform into high-pressure metal and sulfide phases in planetary interiors,

again phases in states quite different from those encountered in the low-pressure mineralogy of chondritic meteorites.

Another example for the likely differentiation processes is the accretion of rock and ice to form protocores of 5–20 Earth masses, in the core-accretion model for Jupiter and other gas-giant (exo)planets. These protocores may differentiate in a manner similar to the terrestrial planets. If heated by gravitational energy release or radioactive decay, icy materials begin to outgas and to form an atmosphere. With estimated formation times of up to about eight million years for Jupiter [20], at least some partial outgassing can be expected before rapid H_2 and He gas capture occurs. In the absence of larger amounts of H and He, such an atmosphere should contain the same gases that made the initially accreted ices, adjusted for local temperature and pressure equilibria. Thus, an icy + rocky protocore may show an atmosphere made of H_2O, CO, CO_2, N_2, and possibly some traces of methane and ammonia before evolving into a gas-giant planet. If so, hot differentiated and outgassed protocores may be detectable since these gases have characteristic absorption bands.

In the disk instability model, an extended protoplanet (approximately few 100 Jupiter radii) may develop a rocky core. Recently, Helled and Schubert [32] addressed the survival of grains and the sedimentation efficiency in the protoplanetary atmospheres of gas-giant planets that form through disk instability. They find that protoplanets of five Jupiter masses or larger cannot form silicate cores because temperatures become too high. Smaller protoplanets allow grain settling and silicate ''core'' formation if contraction rates are low and low internal temperatures ($<$1300 K) are maintained. However, the temperatures reached during accretion are still too high to allow the accretion of ices to the protoplanet. Thus, the protocore formation process fractionates refractory rock from more volatile elements (C, N, O) that form ices.

There is, however, no evidence that such ''cores'' formed during disk instability or in the core-accretion model would remain intact as cores in giant planets, presumably still composed of the heavy elements originally present in the initial core structures. The destruction of rocky and icy protocores during the gradual growth of a giant planet cannot be excluded. On the other hand, the interior structures (moments of inertia) and thermal evolution models of Jupiter and Saturn require that denser cores are present today [24, 33]. The amount of heavy elements in Jupiter may be up to $42M_E$, depending on the mass distribution of these elements between the gaseous envelope and the core. Jupiter's interior structure can be consistently modeled not only without a heavy element core but also with a core of up to $10M_E$. Saturn's structure seem to require core masses made of heavy elements between 10 and $25M_E$, where up to $10M_E$ of heavy elements may be located in the envelope. One larger uncertainty in these interior models is that the equations of state for H and He at the extreme temperatures and pressures within Saturn, Jupiter, and the more massive giant exoplanets are not yet well known. Thus, whether cores are remnants of the original heavy element protocores that never eroded, or are results of interior phase separations is not fully clear (see also [11, 34]).

The destruction of rock and ice certainly happens during the later stages of accretion when smaller planetesimals plunge into a hot and H_2-rich atmosphere of a more mature gas-giant planet. For example, accreting water ice evaporates as the growing planet is gravitationally heated. Then water vapor becomes a more prominent constituent of the gaseous atmosphere. Another example is solid organic (C-bearing) compounds that react with H_2 to form methane, a major component in cool gas-giant planets and also found in gas-giant exoplanets. Similarly, iron sulfide contained in rocky accreting material can easily break down and react with H_2 to H_2S at temperatures above 600 K. Hydrogen sulfide remains as a gas in a hot giant planet atmosphere (below the NH_4SH cloud base at ~200 K) although it may also be photolyzed by UV light ($\lambda < 280$ nm) at higher altitudes.

Overall, the content of heavy elements in the observable atmospheres of gas-giant planets is from three principal sources:

1) The first source consists of elements present in gaseous form in the planetary accretion disk. In the case of the core-accretion model, the stage of gravitational H_2 and He capture may also include volatile elements in noncondensed phases CO, N_2, CH_4, etc., depending on temperature in the protoplanetary accretion disk. In the disk instability model of giant planet formation, all elements accrete in the same proportions from the accretion disk as present in the host star.
2) If giant planets form through core accretion, the relative heavy element to hydrogen ratio in the outer atmosphere may be higher than that in the host star. There are two possibilities for this. First, the rock and ice protocores erode during accretion of H_2-rich gas and subsequent planetary evolution. However, transport and integration of material from eroding cores into the overall planet may become limited through the metallic and liquid H_2/He layers if the planet is large, and if the planet is not fully convective during its entire evolution. The second possibility is inefficient H_2 and He accretion, which also causes a higher metal to hydrogen ratio.
3) The third source of rocky and icy elements in outer planetary atmospheres comes from the late accretion of rocky and icy planetesimals such as asteroidal or cometary bodies. These contributions affect elemental abundances in giant planets irrespective of their formation mode (core accretion or disk instability mechanism). In the solar system, the time of heavy bombardment of the inner planets lasted until about ~3.9 billion years ago, so accretion of larger rocky and icy objects by the giant planets may have proceeded for ~0.7 billion years after the start of solar-system formation. However, given the mass currently observed in asteroids (less than one Earth mass) and in Kuiper Belt objects (around one Earth mass), the absolute amount contributed to giant planets during late-stage evolution may only have been in the Earth-mass range, and may not be a dominant source for most heavy elements in the observable atmospheres of Jupiter and the other outer planets.

8.4
Similarities between Gas-Giant Planets and Brown Dwarfs

The densities for the known exoplanets suggest that planets with Jupiter masses and above are H_2- and He-rich gas-giant planets. A first-order assumption for modeling is that such planets have similar elemental compositions as their host stars, that is, near-solar or somewhat enriched in elements heavier than helium. This assumes that there were no physical fractionations of the icy and rocky phases through radial redistributions so that the formation location of a gas-giant planet indeed has the same abundances of the elements as the host star. It further assumes that there were no chemical gas–solid fractionations and no fractionations between the solid icy and rocky phases during planetary accretion or by the differentiation process in the planet's interior. A planet's mass and the overall elemental composition are then taken as inputs to model the planet's interior density and temperature structure. The temperature–pressure structure of the outer atmosphere is influenced by the opacity of the gases and condensed phases that are present, and therefore chemistry becomes important. The chemistry is also dependent on T and P, hence the construction of models for the pressure–temperature (P–T) structure of a planet is an iterative process [35–38]. The atmospheric P–T structure also depends on the age of the object and its radiative cooling rates. Yet another variable is the distance of a planet to its host star and the radiation that it receives from it.

The outer atmospheric chemistry of most gas-giant planets should be dominated by molecular and atomic chemistry so that they resemble either cool methane T dwarfs, hotter brown dwarfs of type L, or even late M dwarfs. A major factor for the astonishing range in the chemical variations is the different distances of gas-giant exoplanets to their host stars (see below). The type of chemistry causes differences in the optical, far-red and near-infrared spectra, and guidelines of what we can expect for exoplanet spectra are already available from the many spectra of ultracool stars and brown dwarfs. The appearance and disappearance of certain molecular bands and atomic lines in optical and near IR spectra are applied in the classifications of hot L and cool T dwarfs, which as ''brown dwarfs'' are the links between real dwarf stars of type M and planets like Jupiter [39]. The atmospheric chemistry that characterizes M stars (strong absorption bands of gaseous TiO, VO, CO), L dwarfs (strong alkali lines, strong CO, H_2O, FeH, CrH bands), and T dwarfs (CH_4, H_2O bands) should be quite similar to what could be seen for exoplanets of different ages and locations from their primaries. The following compares brown dwarf and exoplanet atmospheres.

The P–T structure for Jupiter, two T dwarfs (Gl570D, Gl229B), two generic L dwarfs, two M dwarfs, and M giant stars are shown in Figure 8.3 to give some orientation of the T and P range that we are concerned with (data from [40, 41]). Note that the P and T axes are reversed so that T and P increase toward the origin (''into the object''). The T and L dwarfs have masses above $\sim 13 M_{\rm Jup}$, the minimum mass required for D-burning, but are below ~ 60–$70 M_{\rm Jup}$ needed for H-burning. The T dwarf Gl229B has an effective temperature $T_{\rm eff} \sim 960$ K and a mass of $\sim 58 M_{\rm Jup}$. It is a binary companion to the M dwarf star Gl229A at a separation of

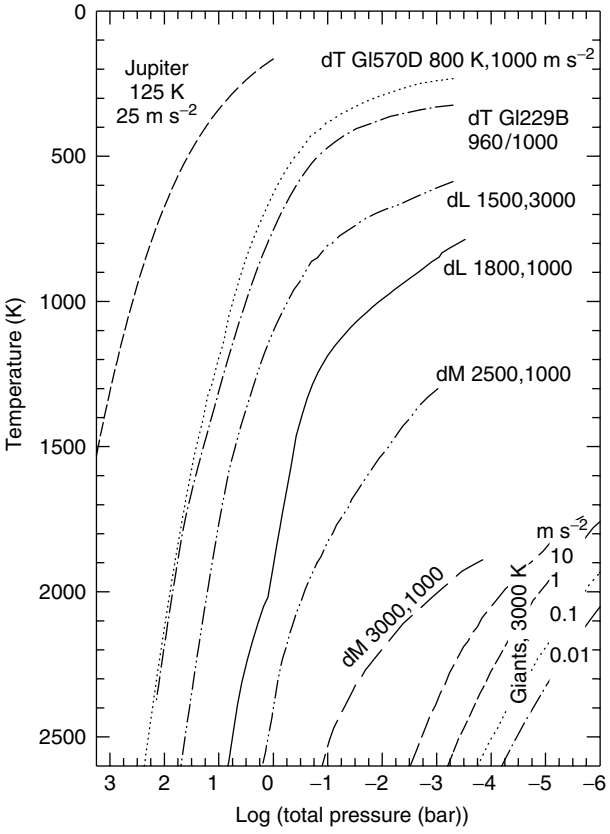


Figure 8.3 Pressure and temperature structures in Jupiter, brown dwarfs, and cool dwarf and giant stars. Numbers after labels indicate effective temperature (K) and surface gravity (in m s^{-2}). Data from Marley *et al*. [40], and Marley and Saumon [41], except for late type stars from Gustafsson *et al*. [44].

~44 AU, so one could think of this brown dwarf as a very massive planet. The T dwarf Gl570D has similar properties $T_{\rm eff} = 800$ K and $34M_{\rm Jup}$ [42]. The P–T models for the T dwarfs have a surface gravity of ~1000 m s^{-2} (Jupiter's is ~25 m s^{-2}) and the radii of T and L dwarfs are comparable to that of Jupiter and gas-giant exoplanets. Models for L dwarfs with relatively similar surface gravities and effective temperatures of 1500 and 1800 K are included for comparison. Figure 8.3 shows that objects with higher effective temperatures have higher temperatures at the same total pressure level (e.g., at 1 bar: Jupiter 165 K, Gl229B ~750 K, L dwarf ~1900 K). The age of an object is also an issue because without an internal heat source cooling occurs over time. Thus, objects of similar mass may show different spectral characteristics if they differ in age. The L dwarf model is for an object of ~$33M_{\rm Jup}$ and has $T_{\rm eff} = 1800$ K, whereas the methane T dwarf Gl570D with about

the same mass ($34M_{\rm Jup}$) only has $T_{\rm eff} = 800$ K and different spectral characteristics. Thus, if not externally heated, the effective temperature of objects of similar mass depends on age (see, e.g., [3, 43] for P–T evolution of brown dwarfs with time). Since gas-giant planets, like brown dwarfs after D-burning, lack strong internal energy sources and are of similar overall elemental composition, the chemistry of their outer atmospheres should change similarly with age, if they are not heated by radiation of their hot stars. Figure 8.4 shows characteristic P–T profiles for Jupiter and Jupiter-mass exoplanets at different separations from their primary star.

The model atmospheres of gas-giant exoplanets with about the same gravity as Jupiter illustrate how the atmospheric P–T structure changes with a gas-giant planet's position from its primary. For comparison, the brown dwarf P–T models from the previous figure are shown using thin lines. The P–T structures for young (like Jupiter 4.6 billion years ago) planets are shown at the top and plotted parallel to the Jupiter profile [9]. These Jupiter-mass planets have an intrinsic $T_{\rm eff}$ of ∼100 K if they would not receive any flux from the host star. However, when

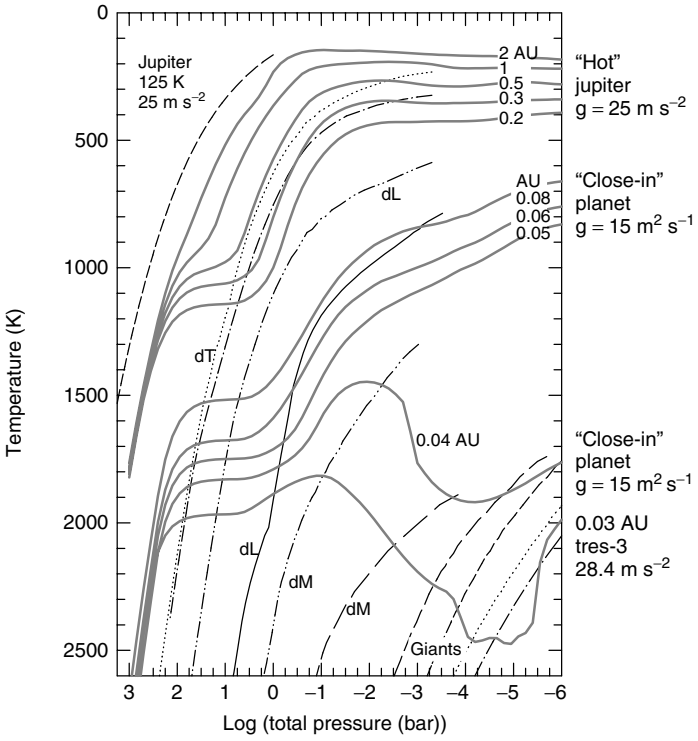


Figure 8.4 Pressure–Temperature structures of Jupiter-mass planets at different distances (indicated by AU) from their primary stars (planet data kindly provided by J. Fortney; see also [8, 9]). The P–T structures of some of the more massive objects from the previous figure are shown with thin lines for comparison.

exposed to radiation from a star like the Sun, the P–T profiles change as shown, from top to bottom, for planetary orbital distances of 2, 1, 0.5, 0.3, and 0.2 AU, respectively. These P–T profiles become similar to that of T dwarfs, as seen from the coinciding profile of the T dwarf Gl 229B and that of a Jupiter-mass planet at 0.3 AU. However, because Jupiter-mass objects have lower surface gravities than brown dwarfs do, they will have lower pressures in their photospheres if they have the same effective temperatures. Conversely, the depth level from the top of the atmosphere at which the same pressure and temperature is reached in a giant planet is larger than in brown dwarfs because a larger overlying mass is required to build up this pressure. Therefore, there is more opacity and the *effective* temperature of such giant planets will be lower; e.g., compare the T dwarf Gl580D ($T_{\rm eff} \sim 800$ K) and the hot Jupiter at 0.5 AU with $T_{\rm eff} \sim 370$ K.

The second set of profiles in Figure 8.4 is for planets at smaller orbital distances (0.05–0.08 AU). These profiles (provided by J. Fortney) are for a somewhat lower gravity of 15 m s^{-2}, corresponding to a Saturn-like planet ($\sim$10 m s^{-2}) instead of Jupiter ($\sim$25 m s^{-2}). However, the change in gravity does not change the P–T structures as much as does the closer proximity to the primary (taken to be a Sun-like star). The upper parts of these atmospheric P–T profiles resemble those of L dwarfs as radiative heating from the primary leads to stronger adiabatic expansion of the atmospheres (less dense; shift to lower total pressure) when compared to the planets located at 0.2 AU and beyond. Many known exoplanets are extremely close to their host stars with orbits of semimajor axis $<$0.1 AU and with orbital periods of a few days. The outer atmospheres of these heated ''Pegasi planets'' (after the prototype 51 Peg b), ''Roaster Planets'' or ''Hot-Jupiters'' can reach temperatures exceeding 1500 K. Without absorbed flux from the primary, such planets may only have effective temperatures of a few hundred kelvin (depending on age). However, because of the high temperatures, the ongoing chemistry should be comparable to that in hot brown dwarfs (L dwarfs), hence spectra of such exoplanets should be comparable to ''free-floating'' brown dwarfs of type L. TrES-1b and HD 209458 b are well-studied, transiting, hot, Pegasi-type exoplanets with outer atmospheres heated to temperatures that correspond to those of brown dwarfs near the L to T dwarf transition ($T_{\rm eff} \sim 1200$ to $\sim$1400 K).

The two planetary thermal profiles at the highest temperatures are for planets at 0.04 and 0.03 AU, the latter a characteristic of the planet Tres-3 b. Exoplanets that are in extremely close orbits with orbital periods of a few hours may develop atmospheric thermal inversions [8, 37, 45] where high temperatures and tenuous pressures that are comparable to regimes in giant and supergiant star atmospheres; and hence the comparison to these much more massive objects here.

8.4.1 Chemistry in Gas-Giant Planets

A useful start of the chemistry discussion is with the more abundant reactive elements (e.g., C, N, O) because their chemistry influences that of less-abundant elements. For example, many elements present as monatomic gases can form

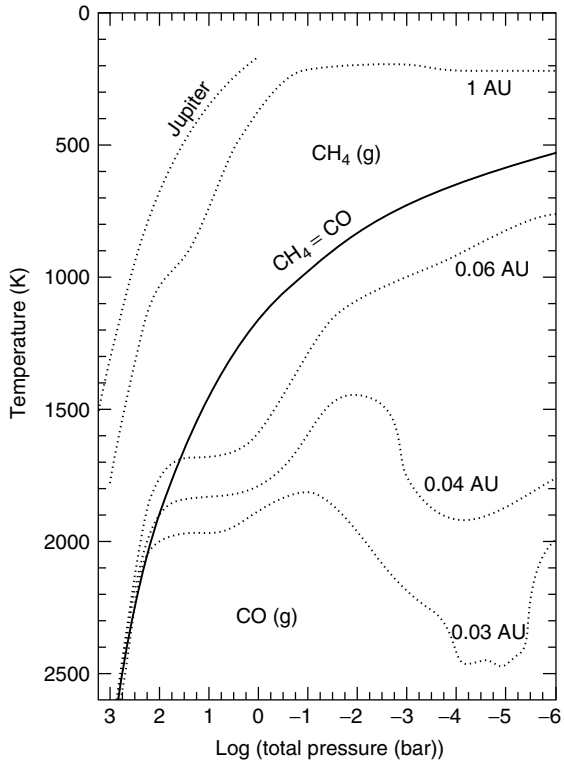


Figure 8.5 Carbon chemistry in low-mass objects is dominated by the gases CO (at high T and low P) and CH_4 (at low T and high P). The solid line shows where CO and CH_4 have equal abundances. The dotted lines are P–T profiles for Jupiter and Jupiter-mass objects at different separations (see previous figure).

oxides in cooling gases, depending on the availability of oxygen. In cooler planets, there are several coupled reactions between molecular gases and condensates, but many of these reactions are regulated by the C and O chemistry.

One major chemical difference between T dwarfs and L dwarfs is in their carbon and oxygen chemistry, which is the same as for cool (old and distant) and hot (young/and or heated) Jupiter-mass planets. In T dwarfs and planets like Jupiter, methane is the major C-bearing gas. In L and M dwarfs and close-in gas-giant planets, it is CO gas. The chemistry of C and O as functions of T and P are shown in Figures 8.5 and 8.6, which are in the same format as the previous figures. The curves indicate where the major C and O gases are equal in abundance and define ''fields'' where a certain gas is dominant. Many details of the C, N, and O distribution and speciation are discussed in Lodders and Fegley [46]. Over a large P–T range, the equilibrium distribution of carbon is controlled by the reaction

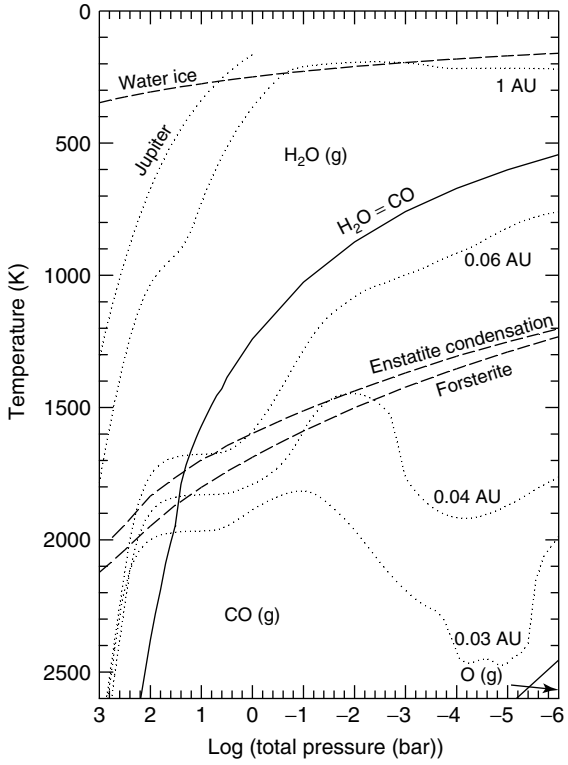


Figure 8.6 Oxygen chemistry in low-mass objects is dominated by the gases CO and water. The solid line shows where CO and water have equal abundances. The dotted line are P–T profiles for Jupiter and Jupiter-mass objects at different separations (see also Figure 8.5).

between CO and CH_4 through the net reaction:

$$CO + 3\ H_2 = CH_4 + H_2O \tag{8.1}$$

This reaction, which depends on total P and temperature, describes the equimolar abundance curve of CO and CH_4 that bisects the C distribution diagram (Figure 8.5). At higher temperatures and lower total pressures, CO is the major C-bearing gas, and methane abundances gradually decrease when moving away from the equimolar abundance curve into the CO field. For comparison, P–T profiles of Jupiter, and Jupiter-like planets at 1, 0.06, and 0.04 AU from Figure 8.4 are overlaid. The region dominated by CO coincides with the upper atmospheric P–T profiles of hot Pegasi-type planets (and L and M dwarfs). The major C-bearing gas in Jupiter's outer observable atmosphere is CH_4, and moving Jupiter to 1 AU or 0.2 AU would not change this. By moving a planet closer to the primary, the internal P–T structure shifts closer to the CO=CH_4 curve. The CO abundances increase exponentially as one moves more toward the CO field, where CO reaches

its maximum abundance. The boundary is crossed if a Jupiter-like planet moves to 0.08 AU or closer, and CO becomes the major gas and CH_4 abundances continue to drop toward higher T and lower P.

Objects whose upper atmospheric P–T profiles fall into the CH_4-rich field are expected to show methane absorption bands and only little CO absorption, if any at all. However, this only applies if the CO to methane conversion reaches equilibrium values. This is the same problem as noted before for the C chemistry in protoplanetary disks. However, due to the higher total pressures in planetary atmospheres, the CO to methane conversion can proceed down to ∼1000 K, whereas in the solar nebula, the temperature for freezing the conversion reaction was about ∼1500 K. At quench-level temperatures in planetary or brown dwarf atmospheres, the chemical reaction timescale becomes too long when compared to atmospheric mixing timescales, and the CO to methane conversion will not proceed to completion. Thus, the amount of CO is frozen-in or quenched, and CO is overabundant when compared to its expected equilibrium abundance at low temperatures. The overabundance of CO is observed in Jupiter, Saturn, and Neptune and is plausibly explained by quenching of the CO to CH_4 conversion due to rapid vertical mixing (see [47] for a discussion). Given the similarity in chemistry, an overabundance in CO was also predicted for the first discovered methane T dwarf Gl 229b [48], which was subsequently observed [49]. The overabundance of CO is now established for several T dwarfs [50]. Hence quenching of the CO to methane conversion should indeed be a universal phenomenon in atmospheres with thermal profiles in the CO-rich field at higher temperatures and in the CH_4 field at lower temperatures. If so, observations and quantization of CO abundances in methane-rich objects provide a tracer of the mixing processes in such objects.

The oxygen distribution is shown in Figure 8.6. The major O-bearing gases are CO and H_2O. Under conditions where CO is the major C-bearing gas, most O is also in CO but H_2O is always quite close in abundance. With a bulk solar $C/O = 0.5$ and all carbon tied to CO, oxygen should be equally distributed between CO and H_2O. However, SiO and other metal oxides reduce the H_2O abundances and the more stable CO becomes the most abundant gas. The conversion reaction of CO to methane produces equimolar amounts of H_2O. The H_2O abundance increases proportionally to the decrease in CO. Therefore, the curve where CO and H_2O are equal in abundance is at a similar location as the $CO{=}CH_4$ abundance curve in Figure 8.5.

Abundances of O-bearing gases are affected by the formation of condensates in the atmospheres. Forsterite (Mg_2SiO_4) and enstatite ($MgSiO_3$) are the two most important sinks for oxygen at high temperatures because Mg and Si are the most abundant rock-forming elements. All high-temperature oxides and silicates consume about 20% of total oxygen. The remaining oxygen at lower temperatures is mainly in water vapor, which condenses into a cloud layer around 200–300 K, depending on details of the P–T structure. It is interesting to note that the P–T profile for the hot Jupiter (at 1 AU) in Figure 8.6 coincides closely with the condensation curve for water ice (at solar metallicity) over an extended range. Such planets would be prime candidates to show spectral signatures of water ice,

whereas in planets like Jupiter, water clouds are buried too deep in the atmosphere to be visible.

The atmospheric P–T profiles of the more strongly heated planets at 0.04 and 0.06 AU cross the Mg-silicate stability curves several times. This implies different possible locations within the atmosphere to trap Mg-silicate cloud material from cooler atmospheric regions. However, the atmosphere of the planet at 0.06 AU should be free of Si- and Mg-bearing gases at temperatures below ~1600 K and an enstatite/forsterite cloud layer is expected between ~1600 and ~2000 K. The planet at 0.04 AU has a more complicated Mg-silicate cloud distribution, and the behavior of condensates in such strongly heated planet with thermal inversions is not fully understood; see, e.g., the discussion by Fortney *et al.* [8] on TiO and removal of Ti-bearing condensates.

One can look at all elements in the periodic table and check which gases and condensates are formed at a given temperature and pressure to sort out conditions where important opacity sources (e.g., atomic lines, TiO, VO, CaH, CO, H_2O, CH_4, FeH, CrH, ...) appear, what their abundance are, and when these gases are removed by formation of other gases and condensates so that their spectral signatures become lost. This evaluation for all stable elements was done for the Jovian and Saturnian atmospheric pressure and temperature profiles [47]. It is an impractical approach to model the chemistry for all of the newly discovered exoplanets this way. However, these results, listed in abbreviated form in Table 8.2 for the 20 most abundant reactive elements, should be useful to give an indication of possible major gases and condensates. Where appropriate, differences in the major gas abundances expected for hot and cooler objects are indicated.

8.4.2 Condensate Clouds

Thermodynamic equilibrium and kinetic calculations are useful to find which elements and chemical species are present as functions of T, P, and metallicity. The computation of gas-phase equilibria is straightforward, but condensate formation in gas-giant planets must consider that initially formed condensates sediment into cloud layers and do not react with gases at higher (cooler) altitudes. The chemical condensate treatment requires the distinction of two possible end-member settings: condensate formation in a strongly gravitationally bound atmosphere and condensate formation in a low-gravity environment such as in protoplanetary disks. Figure 8.7 illustrates the differences in condensates for some elements. In a bound atmosphere, condensates forming directly from the gas at high temperatures (''primary condensates'') settle due to the influence of gravity and form cloud layers. These primary condensates do not react with the cooler gas at altitudes above the condensate clouds so they are out of equilibrium with the gas in the overlying atmosphere. In contrast, condensates in protoplanetary disks can remain dispersed in the gas and react with it to form secondary condensates during cooling. For example, in both cases, Fe-metal is the first Fe-bearing condensate. However, in planetary atmospheres, Fe_3P and FeS cannot form because Fe is in a deep cloud.

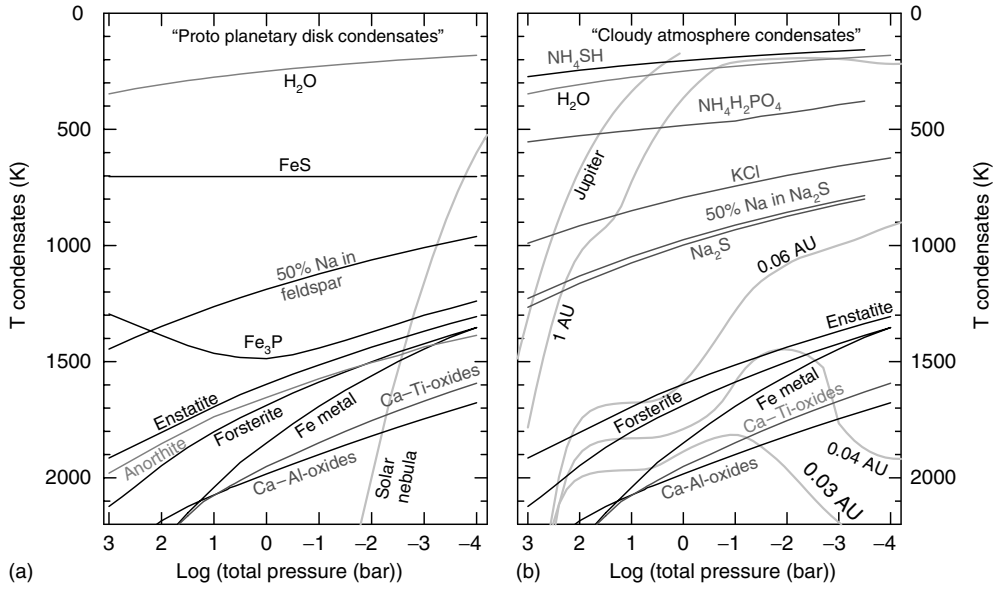


Figure 8.7 (a) Some of the condensates expected in protoplanetary disks. Primary, high T condensates (e.g., Fe-metal) can continue to react with gas to form secondary condensates (e.g., FeS) at lower temperatures. (b) Condensates that appear if primary condensates settle into cloud layers. There are no secondary Fe-bearing condensates and different types of condensates appear at low T. Note that high T and high P are toward the bottom left in the graphs.

Then P condenses mainly into $NH_4H_2PO_4$, and S condenses into sulfides such as Na_2S, MnS, and mainly into NH_4SH.

The condensate cloud layer concept (often called *rainout*) was initially developed for the Jovian planets [51, 52], and it has been found that only models including cloud sedimentation are capable of matching observed brown dwarf color diagrams and spectra. Hence what works for Jupiter and brown dwarfs should work for exoplanets with masses in between.

The cloud masses that will be available depend on the availability of the constituent elements in the condensates. Table 8.2 shows that the most abundant rock-forming elements Mg, Si, and Fe produce massive clouds of forsterite, and enstatite, and liquid iron, respectively. The other less-abundant rock-forming elements such as Ca, Al, Ti, Mn, Cr, and the alkalis will produce less-massive clouds. In addition to the more refractory (high thermal stability) condensates, O, S, N, and C form abundant condensates at temperatures below ~200 K. Giant planets with low-temperature upper atmospheres can have massive icy cloud layers of water (ice or liquid), NH_4SH (s), NH_3 ice, and CH_4 ice. However, they will also have the full sequence of deeper seated cloud layers containing the more refractory elements in the interior.

Once a condensate becomes stable, the abundance of the constituent gases exponentially decreases with decreasing temperature in the atmosphere. With

Table 8.2 Condensates and gases in solar-composition giant planets.

	Abundance Si = 10^6 atoms	Condensate type	Major gases
H	2.88E+10	Minor fraction in H_2O	H_2, (H_2O, CH_4)
O	15 500 000	Oxides, silicates, water	H_2O; CO high T
C	7 080 000	CH_4	CH_4 low T, CO high T
N	2 090 000	NH_4SH, NH_3	NH_3 low T, N_2 high T
Mg	1 020 000	Mg_2SiO_4, $MgSiO_3$	$Mg(OH)_2$, MgOH, MgH, Mg
Si	1 000 000	Mg_2SiO_4, $MgSiO_3$	SiH_4, SiO, SiS
Fe	838 000	Fe-alloy	Fe, $Fe(OH)_2$, FeH
S	445 000	NH_4SH; minor MnS, ZnS, Na_2S	H_2S
Al	84 100	Al_2O_3, Ca–Al-oxides	Al_2O, AlH, AlOH, $HAlO_2$
Ca	62 900	Ca–Al-oxides, Ca-titanates	Ca, $Ca(OH)_2$, CaOH, CaH, CaS
Na	57 500	Na_2S	Na, NaH, NaCl, NaOH
Ni	47 800	NiFe-alloy	Ni, NiH, NiS
Cr	12 900	Cr-metal, Cr_2O_3	Cr, CrH, CrS, CrO
Mn	9170	MnS	Mn, MnH, MnS
P	8060	$NH_4H_2(PO_4)_3$, minor Cu_3P	PH_3, PH_2, P_2, PO
Cl	5240	NaCl, KCl, NH_4Cl	HCl, NaCl, KCl
K	3690	KCl	K, KCl, KOH
Ti	2420	Ca-titanates, e.g., $CaTiO_3$	TiO, TiO_2, Ti
Co	2320	FeNi alloy	Co, CoH
Zn	1230	ZnS	Zn, ZnH, ZnS
F	841	NaF, KF, NH4F	HF
Cu	527	Cu_3P	Cu, CuH
V	288	V-oxides into Ca–Ti and Ca–Al-oxides	VO, VO_2, V
Ge	121	Ge, minor GeTe	GeH_4, GeS, GeSe, GeTe
Se	65.8	PbSe, other selenides	H_2Se
Li	55.5	LiF, Li_2S	Li, LiF, LiCl, LiOH, LiH
Ga	36.0	GaS	GaOH, Ga_2S, GaS
Sc	34.2	Sc_2O_3	ScO
Sr	23.6	SrS	$Sr(OH)_2$, SrOH, Sr
B	17.3	H_3BO_3	H_3BO_3, $NaBO_2$, KBO_2, HBO_2
Br	11.3	NaBr, KBr, RbBr, NH_4Br	HBr, alkali bromides
Zr	11.2	ZrO_2	ZrO, ZrO_2, ZrS
Rb	6.57	RbCl, RbBr	Rb, RbF, RbCl

the removal of the gases, their opacity also disappears. Therefore, the cloud-layer structure in the atmosphere becomes traceable through the changes in gas chemistry. The expected spectral features caused by the removal of gases into condensate clouds in giant planets should be similar to that known for the hot-to-cool brown dwarf sequence (see the discussion in [53]).

Extremely heated gas-giant planets will have very strong TiO and VO bands in the optical, and H_2O, CO, and FeH bands in the near-infrared. Near ~1800–2200 K,

refractory ceramics such as corundum (Al_2O_3) or Ca-aluminates (such as hibonite, $CaAl_{12}O_{19}$) and Ca-titanates (such as perovskite, $CaTiO_3$) become stable. This leads to a gradual disappearance of TiO and VO bands, since VO follows TiO by forming solid solutions with the Ti-bearing condensates. Similarly, Ca- and Al-bearing gases are removed. The exact nature of the refractory condensate is sensitive to the total pressure [54]. The first Ca and Al condensate at high total pressures in the deeper interiors of cool Jupiter-like giant planets is melilite (a solid solution of gehlenite $Ca_2Al_2SiO_7$ and akermanite $Ca_2MgSi_2O_7$). However, it is hibonite below ~1 bar, and Al_2O_3 below 0.01 bar, which is more relevant to Pegasi-type planets or extremely heated exoplanets. The initial Al-bearing condensate in Pegasi planets is Ca-poor, which affects Ca gas abundances. The Ca (and Al) chemistry determines which Ca-titanate forms, which, in turn, is important for the TiO gas abundances.

Since the temperature of gas-giant planets in close orbits is regulated by the irradiation from their host stars, one can expect that the closest planets (around the same type of star) are the most likely to be heated above temperatures necessary for the condensation of ceramics. In that case, TiO and VO remain, and their opacity determines whether a Pegasi-type planet can develop a stratospheric temperature inversion [8, 37, 55]. The photometric data and spectra of HD 209458b [56, 57] seem to be consistent with the presence of TiO and VO and a low-pressure inverted temperature structure in this exoplanet.

Molecular bands of CrH and FeH become stronger in cooler atmospheres but decrease in strength when Fe-metal and, depending on total pressure, Cr or Cr_2O_3 condensation occurs. This can remove all Fe and Cr gases, respectively, around ~1200–1500 K. The removal of Fe into a cloud layer is also the reason why H_2S remains in the atmospheric gas at lower temperatures. If Fe cloud settling did not occur, H_2S would be completely absent from the upper atmospheres because formation of FeS (troilite) from Fe-metal grains with H_2S gas starting at ~700 K consumes all H_2S gas (solar Fe/S ~ 2). However, formation of secondary FeS and removal of H_2S at 600–700 K is at odds with the *Galileo* probe observations on Jupiter. The *Galileo* entry probe mass spectrometer (GPMS) detected H_2S at approximately three times the solar S/H_2 ratio in Jupiter's atmosphere [58] which is only possible if H_2S is not removed. Because H_2S is observed in the Jovian and Saturnian tropospheres at altitudes below the NH_4SH cloud condensation level, the condensate formation models must call for the depletion of Fe-metal into a cloud deep in the Jovian and Saturnian atmospheres.

In the temperature range of Fe-metal condensation, Mg-silicate (forsterite and enstatite) cloud formation is expected as well. This will remove all Mg (e.g., Mg, MgH, MgOH) and Si (e.g., Si, SiO, SiS, SiH_4) gases from the cooler atmosphere. As for sulfur, Jupiter and Saturn are test cases for the cloud-layer approach. The absence of silane (SiH_4) and the presence of germane (GeH_4) in Jupiter and Saturn is due to depletion of refractory Si, but not of volatile Ge, by condensate formation deep in their atmospheres [47]. Silicon is much more abundant than Ge in a solar-composition gas (atomic $Si/Ge \sim 8300$) but there are only observational upper

limits for $SiH_4/H_2 \sim 1 \times 10^{-9}$ (1 ppb) on Jupiter and Saturn. For comparison, the protosolar Si/H_2 molar ratio is 8.23×10^{-5}, which is ~82 000 times larger than the observational upper limit on the silane abundance. In contrast, the observed GeH_4/H_2 is ~0.7 ppb on Jupiter and ~0.4 ppb on Saturn [1]. These values are closer to the protosolar Ge/H_2 molar ratio of 9.9 ppb and the difference arises because not all Ge in Jupiter and Saturn is present as GeH_4 [47]. The presence of Mg-silicate clouds should be testable by searching for SiH_4 (which has strong IR bands) in the atmospheres of methane-rich planets, or for SiO and SiS in hot, CO-rich planets. These gases should be absent or depleted above the Mg_2SiO_4 and $MgSiO_3$ clouds if the cloud condensation models are correct.

Enstatite is the lowest temperature condensate of the major element (Al, Ca, Ti, Mg, Si, Fe, Ni) condensates. Thus, all of the major rock-forming elements are out of the gas above the enstatite cloud. However, lines of monatomic alkali metals K, Na, Cs, and Rb should remain prominent and persist to lower temperatures that correspond to hotter methane-rich objects (Figure 8.5). Monatomic K gas is observed in T dwarfs such as Gl 229B and Gl 570D [2, 42]. Monatomic Na, at less than the solar Na abundance, is inferred for the exoplanet HD 209458b [59]. The monatomic gases are dominant until they gradually convert to NaCl, KCl, and other gases such as oxides, hydroxide, and hydrides, before being removed as sulfide and halide condensates. Sodium is much less abundant than S, so all Na can be sequestered into a Na_2S cloud but only 6.5% of all sulfur would be lost, which does not lead to a significant loss in total H_2S.

The presence of Na and K in cooler objects requires that refractory, rock-forming elements such as Al, Ca, and Si are depleted by condensate cloud formation deep in the atmosphere. Otherwise, Na- and K-gases would condense into silicate minerals such as $XAlSi_3O_8$ (albite (X = Na) and orthoclase (X = K)) at high temperatures, and Na and K gas would be depleted from the observable atmosphere. However, if Al, Ca, and Si are in deep condensate clouds, $XAlSi_3O_8$ condensates cannot form and Na and K stay in the gas. Only model spectra that consider Al, Ca, and Si removal into deep cloud layers show good agreement with the higher observed K I line strengths for Gl 570D and Gl 229B [42]. The chemistry of Li and the other alkalis Rb, and Cs is similar. Instead of being removed from the gas by reaction with high-temperature feldspar, they only form their own sulfide (e.g., Na_2S) and halide condensates at temperatures 500 K lower or more than, e.g., Mg-silicates.

The switch from CO to methane should cause an onset of absorption in the 1.6, 2.2, and 3.3-μm methane bands; similar to what was expected and found in the L dwarf spectral series [60]. With methane becoming more abundant, the infrared colors will change because of the strong IR methane absorption bands, as is known for the change in L and T dwarfs. In cool methane-dominated objects, the alkali lines have vanished. Methane, water, and ammonia characterize the cooler objects until condensation of liquid water or water ice and solid NH_3 leaves only methane behind in an otherwise He- and H_2-rich atmosphere. Removal of methane from the atmosphere into clouds requires very low temperatures (<50 K).

8.4.3
The Effects of Varying the C/O Ratio on Gas-Giant Planet Chemistry

8.4.3.1 Gas Chemistry Variations through Changes in C/O Ratio

Variations in the C/O ratio affect the chemical composition of the atmosphere because the CO, methane, and water equilibria are changed. This affects the interior evolution and spectra of gas-giant planets. For example, Fortney *et al.* [36, 61] show that changes in metallicity and C/O ratios from solar values strongly influence the emergent spectra of brown dwarfs and hot exoplanets. Changes in C/O mainly affect abundances of gases containing C and/or O. To a first approximation, abundances of gases such as NH_3 and N_2 are not affected much because the most abundant N-bearing gases contain neither C nor O. However, the increasing abundances of gases such as HCN that contain C are of minor importance in a solar-composition gas.

Figure 8.8 shows the gas abundances as a function of C/O ratio at 1100 K and 0.01 bar to illustrate trends. At this temperature, C is mainly distributed between CO and CH_4, and O between CO and H_2O in a solar-composition gas. The gas abundances are calculated for C/O ratios from solar to unity by increasing the C abundance (closed symbols in Figure 8.8) or by decreasing O from the system (open symbols). The partial pressures of CO, CH_4, C_2H_2, HCN, and CN increase, whereas that of H_2O declines over orders of magnitude at otherwise constant conditions. The increase in the CO fugacity is relatively modest (a factor of 2), which reflects increasing the C/O from 0.5 to unity and that most C is in the very

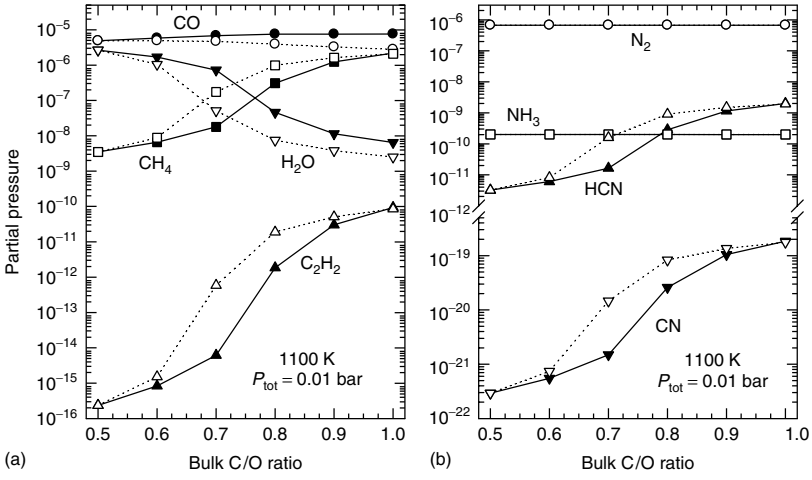


Figure 8.8 Variation of gas chemistry of carbon and oxygen (a) and nitrogen (b) as a function of C/O ratio at 1100 K and 0.01 bar. The C/O ratio was increased from solar by increasing the total C abundance (closed symbols) or decreasing the total oxygen (open symbols). Note the break in the concentration scale for nitrogen gases.

stable CO molecule. Therefore, changes in CO opacity, important for the 4.5-μm spectral window, decreases. A particularly strong increase is for C_2H_2 and HCN, which can become new opacity sources in the infrared. On the other hand, the H_2O fugacity drops by several orders of magnitude if the C/O ratio is increased from solar to unity because with a relatively high C-content, more O is in CO and less O is available for water and other O-bearing molecules.

Figure 8.8 also illustrates that it matters how the C/O ratio is altered. An increase in C/O ratio by O removal yields a larger effect than adding C because the absolute amount of O-bearing molecules decreases. This is the reason why the H_2O abundance decreases much more strongly when O is removed. Since water is a major opacity source in gas-giant exoplanets, substantial decreases in water abundances should change the spectra in the 8-μm region of these objects [36]. Together with observations of CO, this may constrain planetary C/O ratios, which has possible implications for planetary formation scenarios. If a gas-giant planet has a different C/O ratio from its host star, the variation of the C/O ratio must have to do with planet formation itself, if chemical fractionations (such as water-ice condensation) in the planet can be ruled out.

8.4.3.2 Possible Scenarios to Alter C/O Ratios during Planetary Formation

There are two possibilities to change the C/O ratio in a giant planet from the value of ~0.5 in an otherwise solar-composition gas. For an increase in C/O in an accretion disk, the plausible mechanisms are to increase the carbon abundance or decrease the oxygen abundance. For a decrease in C/O, an increase in O is the most plausible possibility.

Assuming that gas-giant planets form by core accretion, an increase in the C abundance occurs if abundant interstellar or locally produced organics are incorporated into the rocky protocore of a giant planet. As discussed earlier, tarry organics could exist in larger amounts in planetary accretion disks beyond an organic condensation/evaporation front (a tarline). This hypothesis is analogous to assuming that planets like Jupiter grew from a protocore composed of rock and water ice, at an orbital distance near or beyond the water-ice condensation/evaporation front (the snowline). Organic substances may be similarly suitable to increase the mass density of the disk to aid rapid formation of a core and by ''gluing'' rocky and carbonaceous substances (see, [62]).

The compositional outcome of these two cases is quite different. If a planet forms with excess water ice, the resulting C/O ratio is lower than the host star's C/O ratio. On the other hand, accretion and growth from a carbonaceous and rocky core leads to a C/O ratio above the host star's ratio.

Organics constituting a tarline require higher temperatures to evaporate than water ice; hence, there can be regions in accretion disks that are rich in carbonaceous material, but poor in water ice. The stability of the organics would be enhanced if water was cold-trapped in a more distant region from the central star, because then water vapor, which is a potential oxidant for the carbonaceous material, is also less abundant.

The cold-trapping of water ice also provides a possibility to raise the C/O ratio because it removes oxygen in the planet-forming region. Radial diffusive redistribution of gas and water cold-trapping at and beyond the snowline would dehumidify the region inward of the snowline [63]. This leads to increased C/O ratios between the star and the snowline. Even if half of all O is in CO as expected from the solar C/O = 0.5, removal of other oxygen moves the C/O ratio toward unity. If planets then acquire water-depleted gas, they may have a higher C/O ratio. However, if gas and water ice are not physically fractionated before accretion to a planet, the *bulk* C/O ratio remains solar.

Thus, a planet with a higher C/O ratio than its host star formed either in a water-depleted region or in a region that was enriched in carbonaceous material. Since fast growth of a massive rocky protocore may be required for making gas-giant planets, the first option, accretion in a water-depleted region could cause a problem because the surface mass density of condensed rocky mass alone is too low for fast growth. In this case, the presence of abundant organic materials could help to increase the mass density. This limits the origin of protocore formation for gas-giant planets with higher C/O ratios to regions in protoplanetary disks below ~400 K, where carbonaceous material remains stable.

The other extreme is a gas-giant planet with a C/O ratio smaller than in its primary star, which is formed by preferential accretion of water ice. Water ice trapping from the inner planetary system must lead to ice pileup at the snowline and beyond, which would be an ideal source. Such a pileup of water is probably recorded by the water-rich planets Uranus and Neptune. If these planets formed near or beyond the snowline with abundant water ice, water enhancements, and smaller C/O ratios would result [64]. This formation scenario applies to disk regions with temperatures below ~180 K to stabilize water ice.

8.5 Outlook

The chemistry of exoplanets is somewhat difficult to summarize since chemical information is mainly available in indirect form through density measurements, and direct spectroscopic observations so far provide limited results. Most known exoplanets are gas-giant planets, and most modeling is thus available for gas-giant exoplanets. This will also undoubtedly change in the coming years, when the upcoming missions begin to discover terrestrial-like planets. However, even without too many observations of chemistry yet, exoplanet chemistry is predictable: it is constrained by the abundances of the elements and the physical conditions in planet formation environments as well as in the planet itself. The observable chemistry of brown dwarfs, the exoplanets' larger ''cousins,'' also has provided valuable guidelines to what can be expected in gas-giant exoplanet spectra, and the known chemistry of the terrestrial planets in the solar system will be equally valuable in interpreting the observations of exoplanets to come.

Acknowledgments

The author thanks Jonathan Fortney and Mark Marley for kindly providing model atmospheres, and Bruce Fegley for discussions. This work was supported in part by NSF grants AST 07-07377 and AST 08-07356 and NASA grant NNG06GC26G.

References

1 Lodders, K. and Fegley, B. (**1998**) *The Planetary Scientist's Companion*, Oxford University Press, New York, pp. 384.
2 Burrows, A., Hubbard, W.B., Lunine, J.I., Marley, M.S., and Saumon, D. (**2000**) in *Protostars and Planets IV* (eds. V. Mannings, A.P. Boss, and S.S Russell), University of Arizona Press, Tucson, pp. 1339.
3 Burrows, A., Sudarsky, D., and Lunine, J.I. (**2003**) *The Astrophysical Journal*, **596**, 587–596.
4 Burrows, A., Sudarsky, D., and Hubeny, I. (**2004**) *The Astrophysical Journal*, **609**, 407.
5 Charbonneau, D., Brown, T.M., Burrows, A., and Laughlin, G. (**2007**) in *Protostars and Planets V* (eds. B. Reipurth, D. Jewitt, and K. Keil), University of Arizona Press, Tucson, pp. 701.
6 Chabrier, G., Baraffe, I., Selsis, F., Barman, T.S., Hennebelle, P., and Alibert, Y. (**2007**) in *Protostars & Planets V* (eds. B. Reipurth, D. Jewitt, and K. Keil), University of Arizona Press, Tucson, pp. 623.
7 Fortney, J.J., Marley, M.S., Hubickyj, O., Bodenheimer, P., and Lissauer, J.J. (**2004**) *Astronomische Nachrichten*, **326**, 925.
8 Fortney, J.J., Lodders, K., Marley, M.S., and Freedman, R.S. (**2008a**) *The Astrophysical Journal*, **678**, 1419.
9 Fortney, J.J., Marley, M.S., Saumon, D., and Lodders, K. (2008b) *The Astrophysical Journal*, **683**, 1104.
10 Guillot, T. (**2008**) *Physica Scripta 2008*, **130**, 014023.
11 Guillot, T., Stevenson, D.J., Hubbard, W.B., and Saumon, D. (**2004**) in *Jupiter: The Planet, Satellites, and Magnetosphere* (eds. F. Bagenal, T.E. Dowling, and W.B. McKinnon), Cambridge University Press, Cambridge, pp. 35.
12 Guillot, T., Santos, N.C., Pont, F., Iro, N., Melo, C., and Ribas, I. (**2006**) *Astronomy and Astrophysics*, **453**, L21.
13 Sudarsky, D., Burrows, A., and Hubeny, I. (**2003**) *The Astrophysical Journal*, **588**, 1121.
14 Fortney, J.J., Marley, M.S., and Barnes, J.W. (**2007**) *The Astrophysical Journal*, **659**, 1661.
15 Seager, S., Kuchner, M., Hier-Majumder, C.A., and Militzer, B. (**2007**) *The Astrophysical Journal*, **669**, 1279.
16 Kleine, T., Muencker, C., Mezger, K., and Palme, H. (**2002**) *Nature*, **418**, 952–955.
17 Gonzalez, G. (**2003**) *Reviews of Modern Physics*, **75**, 101.
18 Boss, A.P. (**1997**) *Science*, **276**, 1836.
19 Boss, A.P. (**2001**) *The Astrophysical Journal*, **563**, 367.
20 Pollack, J.B., Hubickyj, O., Bodenheimer, P., Lissauer, J.J., Podolak, M., and Greenzweig, Y. (**1996**) *Icarus*, **124**, 62.
21 Hubickyj, O., Bodenheimer, P., and Lissauer, J.J. (**2005**) *Icarus*, **179**, 415.
22 Lissauer, J.J. and Stevenson, D.J. (**2007**) in *Protostars and Planets V* (eds. B. Reipurth, D. Jewitt, and K. Keil), University of Arizona Press, Tucson, pp. 591.
23 Chambers, J.E. (**2003**) in *Treatise on Geochemistry*, vol. 1 (ed. A.M. Davis), Elsevier, Amsterdam, pp. 461.
24 Saumon, D. and Guillot, T. (**2004**) *The Astrophysical Journal*, **609**, 1170.
25 Grossman, L. and Larimer, J.W. (**1974**) *Review of Geophysics and Space Research*, **12**, 71.

26 Lewis, J.S. (**1974**) *Annual Review of Physical Chemistry*, **24**, 339.
27 Lodders, K. (**2003**) *The Astrophysical Journal*, **591**, 1220.
28 Fegley, B. (**2000**) *Space Science Reviews*, **92**, 177.
29 Kouchi, A., Kudo, T., Nakano, H., Arakawa, M., Watanabe, N., Sirono, S.I., Higa, M., and Maeno, N. (**2002**) *The Astrophysical Journal*, **667**, L121.
30 Fegley, B. and Prinn, R.G. (**1989**) in *Formation and Evolution of Planetary Systems* (eds. H. Weaver and L. Danly), Cambridge University Press, Cambridge, p. 171.
31 Fegley, B. (**1993**) in *The Chemistry of Life's Origins*, vol. 416 (eds. M. Greenberg, C.X. Mendoza-Gomez, and V. Pirronello), NATO Advanced Science Institute, Series C, Kluwer Academic Publishers, Dordrecht, p. 75.
32 Helled, R. and Schubert, G. (**2008**) *Icarus*, **198**, 156.
33 Wuchterl, I.G., Guillot, T., and Lissauer, J.J. (**2000**) in *Protostars and Planets IV* (eds. V. Mannings, A.P. Boss and S.S. Russell), University of Arizona Press, Tucson, p. 1081.
34 Stevenson, D.J. (**1985**) *Icarus*, **62**, 4.
35 Burrows, A., Sudarsky, D., and Hubeny, I. (**2006**) *The Astrophysical Journal*, **650**, 1140.
36 Fortney, J.J., Saumon, D., Marley, M.S., Lodders, K., and Freedman, R.S. (**2006**) *The Astrophysical Journal*, **642**, 495.
37 Hubeny, I., Burrows, A., and Sudarsky, D. (**2003**) *The Astrophysical Journal*, **594**, 1011.
38 Marley, M.S., Seager, S., Saumon, D., Lodders, K., Ackerman, A.S., Freedman, R., and Fan, X. (**2002**) *The Astrophysical Journal*, **568**, 335–342.
39 Kirkpatrick, J.D. (**2005**) *Annual Review of Astronomy and Astrophysics*, **43**, 195.
40 Marley, M.S., Saumon, D., Guillot, T., Freedman, R.S., Hubbard, W.B., Burrows, A., and Lunine, J.J. (**1996**) *Science*, **272**, 1919.
41 Marley, M.S. and Saumon, D. (**2008**) *The Astrophysical Journal*, **689**, 1327.
42 Geballe, T.R., Saumon, D., Leggett, S.K., Knapp, G.R., Marley, M.S., and Lodders, K. (**2001**) *The Astrophysical Journal*, **556**, 373–379.
43 Burrows, A., Hubbard, W.B., Lunine, J.I., and Liebert, J. (**2001**) *Reviews of Modern Physics*, **73**, 719.
44 Gustafsson, B., Edvardsson, B., Eriksson, K., Jørgensen, U.G., Nordlund, A.A., and Plez, B. (**2008**) *Astronomy and Astrophysics*, **486**, 951.
45 Knutson, J.A., Charbonneau, D., Burrows, A., O'Donovan, F.T., and Mandushev, G. (**2009**) *The Astrophysical Journal*, **691**, 866.
46 Lodders, K. and Fegley, B. (**2002**) *Icarus*, **155**, 393.
47 Fegley, B. and Lodders, K. (**1994**) *Icarus*, **110**, 117.
48 Fegley, B. and Lodders, K. (**1996**) *The Astrophysical Journal*, **472**, L37.
49 Noll, K.S., Geballe, T.R., and Marley, M.S. (**1997**) *The Astrophysical Journal*, **489**, L87–L91.
50 Saumon, D., Marley, M.S., Leggett, S.K., Geballe, T.R., Stephens, D., Golimowski, D.A., Cushing, M.C., Fan, X., Rayner, J.T., Lodders, K., and Freedman, R.S. (**2007**) *The Astrophysical Journal*, **656**, 1136.
51 Lewis, J.S. (**1969**) *Icarus*, **10**, 393.
52 Barshay, S.S. and Lewis, J.S. (**1978**) *Icarus*, **33**, 593.
53 Lodders, K. and Fegley, B. (**2006**) in *Astrophysics Update 2* (ed. J.W. Mason), Springer Verlag, Heidelberg, p. 1.
54 Lodders, K. (**2002**) *The Astrophysical Journal*, **577**, 974.
55 Burrows, A., Hubeny, I., Budaj, J., Knutson, H.A., and Charbonneau, D. (**2007**) *The Astrophysical Journal*, **668**, L171.
56 Knutson, H.A., Charbonneau, D., Allen, L.E., Burrows, A., and Megeath, S.T. (**2008**) *The Astrophysical Journal*, **673**, 526.
57 Désert, J.M., Vidal-Madjar, A., Lecavelier Des Etangs, A., Sing, D., Ehrenreich, D., Hébrard, G., and Fegley, B. (**1993**) in *The Chemistry of Life's Origins*, NATO Advanced Science Institute, Series C, vol. 416 (eds. M. Greenberg, C.X. Mendoza-Gomez, and V. Pirronello), Kluwer Academic Publishers, Dordrecht, p. 75.
58 Niemann, H.B., Atreya, S.K., Carignan, G.R., Donahue, T.M., Haberman, J.A., Harpold, D.N., Hartle, R.E.,

Hunten, D.M., Kasprzak, W.T., Mahaffy, P.R., Owen, T.C., and Way, S.H. (**1998**) *Journal of Geophysical Research*, **103**, 22831–22845.
59 Charbonneau, D., Brown, T.M., Noyes, R.W., and Gilliland, R.L. (**2002**) *The Astrophysical Journal*, **568**, 377.
60 Noll, K.S., Geballe, T.R., Leggett, S.K., and Marley, M.S. (**2000**) *The Astrophysical Journal*, **541**, L75–L78.
61 Fortney, J.J., Marley, M.S., Lodders, K., Saumon, D., and Freedman, R. (**2005**) *The Astrophysical Journal*, **627**, L69.
62 Lodders, K. (**2004**) *The Astrophysical Journal*, **611**, 587.
63 Stevenson, D.J. and Lunine, J.I. (**1988**) *Icarus*, **75**, 146.
64 Lodders, K. and Fegley, B. (**1994**) *Icarus*, **112**, 368.

9
Migration and Multiplicity Effects During Giant Planet Formation

Edward W. Thommes

9.1
Introduction

Ever since the discovery of the first exoplanet orbiting a Sun-like star, 51 Pegasi b [1], observational evidence has been mounting that planetary orbits are strongly influenced by interaction with the protostellar disk that gives birth to them. The ''hot Jupiter'' part of the exoplanet population, giant planets with orbital periods measured in days, point to a mechanism that efficiently shrinks orbits, since their formation *in situ* is very unlikely [2]. Instead, gas giants are thought to actually form at orbital radii much closer to those of Jupiter and Saturn in our own solar system [3, 4]. The discovery of hot Jupiters did not come as a complete surprise, as theoretical and early computational work had already demonstrated the possibility that planets in disks could possess significant radial mobility [5, 6]. The theoretical and computational developments since then have reinforced the importance–indeed, the inevitability–of disk-driven planet migration. Thus we now find ourselves in the curious situation that our own solar system, which seemingly underwent no major migration (at least of its two most massive members) seems to make less cosmogonical sense than the average member of the ensemble of exoplanets discovered thus far. It is only a matter of time until the fundamental question of whether the solar system is, in fact, ''typical'' is decisively answered by observations; for now, the true planet distribution remains obscured by strong selection effects [7]; see also Chapter 1. However, as theorists strive to construct comprehensive models of how a protostellar disk converts itself into a planetary system (models that have both ourselves and our extrasolar counterparts as possible outcomes), predictions are beginning to emerge.

Observations tell us that protostellar gas disks have lifetimes of 10 megayears at most, thus setting the upper time limit for gas giants to form [8]. Rather than regarding the phenomenon of planet migration in isolation, the focus of this chapter is the interplay among formation, migration, and planet–planet interaction over the lifetime of a protostellar gas disk. We begin with a summary of the different modes of migration: ''type I'' for low-mass, fully embedded bodies in Section 9.2, ''type II'' for bodies massive enough to open a disk gap in Section 9.3, as well as

Formation and Evolution of Exoplanets. Edited by Rory Barnes

ISBN: 978-3-527-40896-2

other, more exotic possibilities in Section 9.4. In Section 9.5, we look at past work on simulating how planet–disk interactions can drive planet–planet interactions in multiplanet systems. Finally, in Section 9.6, we look at past work on developing a global model of planet formation and migration. We focus, in particular, on a recent model, which, for the first time, puts together accretion, planet–disk and planet–planet effects, simulating these over the entire lifetime of a gas disk. This model suggests an answer to the question posed above: giant planet formation is typically a fairly violent process, with the shrunken orbits and large eccentricities of the discovered exoplanets a typical outcome. Planetary systems resembling our own, with primordial, nearly circular, giant planet orbits, represent a subset with a more quiescent formation history. Section 9.7 provides a summary.

9.2 Type I Migration

When the mass of a body embedded within a gas disk is sufficiently small, the disk's response is linear. A body orbiting in a gas disk launches density waves at Lindblad and corotation resonances [5], exchanging energy and angular momentum with the disk; the disk thus modifies the body's orbit. The net effect on the body can be obtained by summing over all resonances ([5]; see also, e.g., [9, 10] for a comprehensive review). The result, in a nutshell, is that the parts of the disk inside and outside the planet's orbit, respectively, add and remove angular momentum and energy from the planet's orbit. Thus, there is effectively a repulsion between the the planet's orbit and the inner and outer parts of the disk: the inner disk pushes the planet outward, the outer disk pushes the planet inward. Under most circumstances, the outer part of the disk is thought to dominate slightly, so that the planet undergoes a net orbital decay.

To perform the summation over resonances, one begins by expanding the gravitational potential of the perturbing body in a double Fourier series. The l, m component has a pattern speed

$$\Omega_{l,m} = \Omega_p + (l - m)\kappa_p/m \tag{9.1}$$

where Ω_p and κ_p are the azimuthal and epicyclic frequencies, respectively, of the planet. There is a Lindblad resonance between this potential component and disk material at a radius r in the disk where

$$\Omega(r) \pm \frac{\kappa(r)}{m} = \Omega_{l,m} \tag{9.2}$$

Also, a corotation resonance occurs where

$$\Omega(r) = \Omega_{l,m} \tag{9.3}$$

The radial pressure gradient in the disk causes both the epicyclic and the azimuthal frequency to differ slightly from their Keplerian values:

$$\Omega^2(r) = \frac{GM_*}{r^3} + \frac{1}{r\rho}\frac{\partial}{\partial r}(\rho c_s^2) \tag{9.4}$$

and

$$\kappa^2(r) = \frac{1}{r^3}\frac{\partial}{\partial r}\left(r^4\Omega^2\right) \tag{9.5}$$

In the case of a planet on a circular orbit, all components with $l \neq m$ are zero; $\Omega_{m,m} = \Omega_p$, so all components have the pattern speed of the planet. Furthermore, there is no torque contribution at all from corotation resonances. For simplicity, we will restrict our discussion to circular orbits here. In summing the effect of all Lindblad resonances, one can define a torque density, or torque per unit disk radius. Following Ward [11], the torque density experienced by the disk due to a planet of mass $M = \mu M_*$ ($\mu < 1$) on a circular orbit about a primary of mass M_* at radius r_p, is

$$\left[\frac{\mathrm{d}T}{\mathrm{d}r}(r)\right]_{\mathrm{LR}} = \mathrm{sgn}(r - r_p)\frac{2\mu^2\Sigma(r)r_p^4\Omega_p^4}{r(1+4\xi^2)\kappa^2}m^4\psi^2 \tag{9.6}$$

where $\xi \equiv mc_\mathrm{s}/r\kappa$, c_s is the gas sound speed, and Ψ is the dimensionless satellite forcing function,

$$\Psi = \frac{\pi}{2}\left[\frac{1}{m}\left|\frac{\mathrm{d}b_{1/2}^\mathrm{m}}{\mathrm{d}\beta}\right| + 2\frac{\Omega}{\kappa}\sqrt{1+\xi^2}\beta_{1/2}^\mathrm{m}(\beta)\right] \tag{9.7}$$

with $\beta \equiv r/r_\mathrm{p}$ and

$$b_{1/2}^\mathrm{m}(\beta) = \frac{2}{\pi}\int_0^\pi \frac{\cos m\theta\,\mathrm{d}\theta}{\sqrt{1-2\beta\cos\theta+\beta^2}} \tag{9.8}$$

being the Laplace coefficient.

One can account in a simple, approximate way [12] for the finite thickness of the disk by replacing the Laplace coefficient with a softened approximation,

$$b_{1/2}^\mathrm{m}(\beta) \approx \frac{2}{\pi\beta^{1/2}}K_0\left(m\sqrt{\beta - 2 + \frac{1}{\beta} + \frac{(BH)^2}{rr_\mathrm{p}}}\right) \tag{9.9}$$

with K_0 the modified Bessel function of the second kind, order 0, $H \approx c_\mathrm{s}/\Omega$ the disk scale height, and B a dimensionless scaling factor controlling the degree of softening. In going from a summing of torques at discrete resonances to a torque density, the wavenumber m is turned into a continuous function of radius

$$m(r) = \left[\frac{\kappa^2}{(\Omega-\Omega_p)^2 - c_\mathrm{s}^2/r^2}\right]^{1/2} \tag{9.10}$$

The basic procedure to evaluate the net torque received by a planet embedded in a disk is to perform a sum over the resonant torque contributions ([11, 13] and references therein). The calculations of Tanaka *et al.* [14] take into account three-dimensional effects as well as the corotation torque contribution. They obtain a radial drift timescale

$$\tau_I \equiv \frac{a}{\mathrm{d}a/\mathrm{d}t} = (2.7+1.1\alpha)^{-1}\frac{M_*^2}{M\Sigma a^2}\left(\frac{H}{a}\right)^2\Omega_p \tag{9.11}$$

for a planet of mass M and with semimajor axis a, in a disk having an azimuthally averaged surface density of gas that falls off as a power law with radius, $\Sigma \propto r^{-\alpha}$.

An example, a giant planet core, $\sim$10 $M_\oplus$, at 1 AU in a MMSN disk (1700 g cm^{-}2) with aspect ratio $H/a = 0.05$ would have $\tau_I \sim 10^4$ years. The very short type I drift time has long been considered a serious problem for the core instability model of giant planet formation, since core growth timescales are closer to 10^5–10^6 years [4, 15]. To prevent bodies from falling into the star (or at least off the inner edge of the gas disk) long before they reach core mass, studies of concurrent growth and migration of giant planets have, in the past, either set the type I migration rate to zero (e.g., [16, 17]) or reduced it by a large amount (e.g., [18]). However, it was demonstrated by Thommes and Murray [19] and Thommes *et al.* [20] that the ''unmitigated'' type I migration described by Eq. (9.11) is not nearly as fatal as had been previously thought. Gas disks do not disappear abruptly, rather their density is continuously dropping over their lifetime as they viscously evolve (spreading at their outer edge, accreting at their inner edge). Since $\tau \propto \Sigma^{-1}$, this lengthens the migration timescale everywhere as the disk dissipates. Meanwhile, the protoplanet growth timescale also lengthens as the gas dissipates–gas drag aids accretion of planetesimals by damping their random velocities (e.g., [4, 21]) but the dependence is weaker. Thus, at a sufficiently low disk density, the growth and migration timescales become comparable; from this point on, cores can grow to completion without being lost to migration. Furthermore, it was demonstrated by Thommes *et al.* [20] that for a broad range of plausible disk parameters, this crossover of timescales occurs when the gas disk still contains enough mass to comfortably grow gas giants. It is worth noting that some protoplanets–those that reach substantial mass while the gas disk is also still massive–will still be lost to type I migration. It is the relatively ''late blooming'' protoplanets that undergo only moderate migration, by being born into a moderate-mass gas disk, that survive. Very similar findings were made by Ida and Lin [22].

Although the model of Alibert *et al.* [18] does in fact include the viscous evolution of the disk over time, all of their protoplanets have an initial mass of 0.6 $M_\oplus$, and thus all migrate quite rapidly starting at $t = 0$; this appears to be the major factor leading them to the conclusion that core survival is only possible when type I migration is at least an order of magnitude less rapid than the standard result of Tanaka *et al.* [14].

Several modifications to the standard picture of type I migration have been proposed. One possibility is that the gas disk surface density profile is not smooth as is usually assumed. Laughlin *et al.* [23] and Nelson and Papaloizou [24] show that density fluctuations due to MHD turbulence can overwhelm the intrinsic torque asymmetry of a smooth disk, turning Type I migration from a monotonic orbital decay into a random walk. Menou and Goudman [12] combine calculations of Type I migration with a more detailed disk model, and find that at locations of opacity transition in a disk, variations in surface density can increase the type I migration timescale of a $\sim$10 $M_\oplus$ body to more than 10^6 years, thus creating local migration bottlenecks. Another issue is that an almost universal simplification in the analysis of type I migration–the neglect of the disk self-gravity–likely leads to an overestimate of type I migration rates, as pointed out by Baruteau and Masset [25]. The basic idea is that the dominance of the outer over the inner planet–disk

torques in the standard analysis stems to a large degree from the slightly sub-Keplerian rotation of the gas disk due to pressure support; this systematically shifts the locations of Lindblad resonances inward relative to a perfectly Keplerian disk. However, in reality, the disk's self-gravity slightly *raises* the rotation velocity at every location, partially countering this effect. Baruteau and Masset [25] argue that neglecting this effect leads to an overestimation of type I rates by as much as a factor of 2.

9.3 Type II Migration

When a planet orbiting a disk becomes sufficiently massive, the disk's response is no longer linear, and density waves induced by the planet shock close to where they are launched. The repulsive interaction between planet and disk (see Section 9.2) becomes so strong that an annular gap in the gas begins to open about the planet's orbit. Once a significant gap forms, the planet is said to be in the *"type II" mode of migration*: its orbit essentially acts as a barrier to the gas (though some may still leak through, to pass by the planet and/or be accreted by it), so that the planet becomes locked into the viscous evolution of the gas disk. A detailed review is given in [10].

Planets enter into the type II regime when two conditions are met ([26] and references therein): First, the planet–disk torque must exceed the intrinsic angular momentum transport rate of the disk. This amounts to

$$M \gtrsim M_* \frac{40\nu}{\Omega r^2} \tag{9.12}$$

where ν is the disk viscosity and r the planet's orbital radius. Second, the planet mass must be large enough that its Hill radius exceeds the scale height of the disk, thus preventing disk gas from spilling over/under the planet's orbit:

$$M \gtrsim 3M_* \left(\frac{H}{r}\right)^3 \tag{9.13}$$

This is often referred to as the *thermal condition*, and is also approximately the condition for density waves to shock as soon as they are launched, meaning that the back-reaction on the disk of the disk–planet torques is deposited right at the corresponding resonance locations.

In the simplest approximation, once a planet is firmly in the type II regime, its orbital radius r would shrink at the advection speed of the viscous disk; in other words, the planet would behave just like a representative fluid element in the disk. However, this is only a reasonable approximation as long as the mass of the planet is sufficiently small compared to the mass of the disk. Syer and Clarke [27] define a dimensionless parameter

$$B \equiv \frac{4\pi \Sigma r^2}{M} \tag{9.14}$$

where Σ is the gas surface density at the planet's location; when $B \lesssim 1$, the gap-opening planet starts to act as an impediment to the disk's viscous evolution,

moving inward at less than the disk's intrinsic viscous advection speed, and causing gas to pile up at the planet gap's outer edge. Thus, as a protostellar disk gradually depletes, the "tug-of-war" between planets and disk increasingly favors the former. We revisit this idea in Section 9.6.

9.4 "Exotic" Migration Scenarios

Beyond types I and II, some other scenarios for planet migration have been proposed. One of these has come to be called *type III*, or sometimes, *runaway*, migration. In essence, it arises when the dominant source of planet–disk torques is the interaction of the planet with disk material flowing through the coorbital region, specifically as this material executes a U-turn in the planet's horseshoe-orbit region. This torque scales with migration rate, hence the potential for runaway migration [28–30]. Numerical studies of this regime have proven challenging, since high numerical resolution in the vicinity of the planet is required to model the complex coorbital flows. In particular, because disk material is treated as non-self-gravitating, subtleties arise in handling the gravitational interaction between material within the planet's Hill sphere and the rest of the system. These complications have made it difficult for different groups to converge on a common picture of the exact conditions under which type III migration would take place. It is generally agreed, though, that if such migration does happen, it will do so in the transition between type I and II migration, when a planet orbits within a partially opened gap. Thus, for plausible disk viscosities, it would be applicable to roughly Saturn-mass planets, provided also that the disk mass is still large compared to the planet mass.

Assuming a planet manages to enter this runaway migration regime, it is also unclear how long such an episode will last [10]. Indeed, if one factors in gas accretion onto the planet, it seems difficult to avoid an upper limit: planets in partially open gaps are typically at their peak gas accretion rate [31], thus the planet ought to rapidly grow to a mass sufficient to open a deep gap and change over to the type II regime. Of course, even a transient episode of runaway migration can still move a planet a significant distance, if the migration is fast enough. On the other hand, recent work suggests that accretion itself may prevent the onset of type III migration altogether: D'Angelo and Lubow [31] find that type III migration only commences under rather artificial conditions, key among them a fixed-mass, nonaccreting planet.

A scenario for altering type II migration was first identified by Masset and Snellgrove [32]. It requires a pair of planets, locked into nearby stable orbits by a low-order dynamical mean-motion resonance (see Section 9.5). Specifically, the planets must be in a close enough resonance that they share a common gap within the gas disk. If the outer planet's mass is now low enough that it only partially clears a gap, Masset and Snellgrove [32] found that *outward* migration would tend to take place. It seems that this phenomenon occurs because the pair of planets essentially acts as a sort of one-way valve within the disk: gas from the outer disk can get into

the outer planet's leaky gap. Partial cancellation of the positive planet–disk torque from the inner planet's outer Lindblad resonances, by the negative planet–disk torque from the outer planet's inner Lindblad resonances, then makes it possible for some of this gas to get past the inner planet, ultimately ending up in the inner disk. However, because of the inner planet's larger mass, its gas is deeper, so that little or no gas can flow the other way. If the flow rate from outer to inner disk through this two-planet valve is faster than the overall inward advection rate of the disk gas, then the pair of planets moves outward through the disk.

As is usual with hydrodynamic simulations, only a relatively small number of orbits could be simulated, so it remains unclear, again, whether or not such outward migration is a transient phenomenon. Also, as with type III migration, transience seems hard to avoid if one takes accretion into account: if the smaller outer planet were free to accrete gas, it would rapidly grow to fully gap-opening mass itself, and the pair of planets would revert to conventional type II inward migration. This is a problem for the model of Morbidelli and Crida [33], which proposes that Jupiter and Saturn were originally locked into a 2 : 3 mean-motion resonance, and thus avoided inward migration (or even moved outward) via the above mechanism. However, if this is to have occurred in the presence of a substantial gas disk, it is difficult to see how a partially gap-opening Saturn could have avoided swiftly growing to Jupiter mass.

9.5
Planet–Planet Interactions During Migration

When disk-driven migration operates on multiple planets in a growing system, the potential for "planet–planet–disk" dynamics arises. Like the original work on planet migration itself, this provides a nice example of theoretical–in this case numerical–work anticipating observations: using hydrodynamic simulations, Bryden *et al*. [34] and Kley [35] found that sufficiently closely spaced pairs of gap-opening planets planets in a gas disk would dissipate the intervening gas and migrate convergently, eventually ending up inside a common gap. This raised the possibility that such planets would end up locked in mean-motion resonances (MMRs) (see [36] for a review of MMR dynamics). Indeed, later that same year the first extrasolar MMR was found, the 1 : 2 resonant pair of GJ876b and c. Subsequent work [37–40] looked in more detail at the dynamics of capture and evolution of gas giant planets in low-order MMRs. Given the numerical expense of hydrodynamic simulations, much of this work was actually performed using *N*-body simulations, with simple forces added to approximate the effect of disk-driven migration. The usual setup for such simulations was to presume that only the outer planet was interacting with the disk, and thus to only impose the disk forces on the outer planet. It is probably fair to say that this approach was chosen mainly for simplicity. The physical situation it corresponds to is of a pair of planets orbiting in a central hole in the disk, rather than sharing a common gap. As will be seen in Section 9.6 though, this approach may have turned out to be quite prescient.

Pairs of planets induced to migrate while locked in a mean-motion resonance gain eccentricity. More accurately–since interactions can ''swap'' eccentricity between planets–the resonant two-planet system monotonically increases its angular momentum deficit (relative to the maximum angular momentum state of two circular orbits) as it migrates. Planet–disk interactions are generally also thought to damp eccentricity (however, the presence of a gap complicates the picture; see [9]) but even so some eccentricity growth will occur. Numerical experiments suggest that the eccentricity damping timescale $\tau_e \equiv e/\dot{e}$ is roughly proportional to the type II migration timescale τ_{II}, with $\tau_e \sim 0.1 - 1\tau_{II}$. Thus, in principle, migration in resonance provides a means to generate the sort of large eccentricities seen amongst the discovered exoplanets. The problem is that for the majority of eccentric exoplanets, a resonant companion of a mass sufficiently large to have caused the eccentricity can be ruled out. However, if the eccentricities of the resonant planets grow sufficiently during migration, it is possible for one member of the pair to ultimately be lost, either through ejection or collision with the central star ([41–43], Ch. 11).

9.6 From Disk to Planets: Putting the Pieces Together

Ultimately, our objective is to understand how these various manifestations of planet–disk and planet–planet interactions actually mesh together as a protostellar disk converts itself into planets. This presents a major modeling challenge; detailed 2D (to say nothing of 3D) hydro simulations of even a single planet in a disk are limited to typically thousands of orbits, while gas disks persist for up to ten million years [8], so clearly this approach can only give us brief snapshots of the process. The trick is to extrapolate from these, using less computationally costly methods. One approach is that taken in the series of papers [16, 22, 44, 45]. Using prescriptions for solids and gas accretion fit to the core accretion model of giant planet formation [46], together with prescriptions for type II migration, they performed Monte Carlo simulations of giant planets growing and migrating in an evolving gas disk. Later versions added type I migration [21] demonstrating, in agreement with the findings of Thommes and Murray [18] and Thommes *et al.* [19], that this was not in fact ''fatal'' to giant planet formation. By adjusting model parameters to fit properties of the ensemble of observed exoplanets, the researchers were able to make a number of interesting insights and predictions. For one thing, their findings supported the idea that the observed correlation of discovered planet frequency with host star metallicity [47, 48] was a consequence of giant planet formation taking place via core accretion (thus coupling the solids complement of a protostellar disk to its ability to give birth to gas giants). The models also predicted a planet ''desert'': the planet mass range from 10 to 100 Earth masses ought to be underpopulated, due to the rapidity of gas accretion in this mass range. Data from the *COROT* and *Kepler* missions will allow this prediction to be tested in detail. Using an α-prescription for the disk, initial iterations of the model obtained the best match to observations

for gas disks with a relatively low viscosity corresponding to $\alpha = 10^{-4}$. In later versions, this was raised to 10^{-3} in order for the disk dissipation time to be more consistent with the observational constraint of 10 megayears or less.

A key limitation in Monte Carlo models of this sort is that each planet is modeled in isolation. The approach of Thommes *et al.* [49], on the other hand, was designed to explicitly include multiplicity effects, while at the same time being computationally less expensive than a full hydrodynamic calculation, so that modeling the disk's entire lifetime remained feasible (albeit quite a bit slower than with a Monte Carlo approach). Details of the code are given in the supplementary online material to Thommes *et al.* [49].

The simulations performed by Thommes *et al.* [49] using this code suggested that, indeed, multiplicity effects play a pivotal role in the birth of planetary systems. A typical simulation involves a complex and stochastic interplay of planet–disk interaction, planet–planet interaction, and competitive accretion. As found by Ikoma *et al.* [3], gas-giant formation begins at some time τ_{giant}, at a radius in the disk typically comparable to the Jupiter–Saturn region of the solar system, and spreads out from there. The results of Thommes *et al.* [49] show that, typically, this happens after the gas disk has significantly evolved; its surface density everywhere having dropped by an order of magnitude or more below its original value. As a result, the type I migration rate of the progenitor cores is usually modest (see Section 9.2), and most of the time, migration actually *speeds up* when these cores become gas giants, open a deep gap and enter into the type II regime. At the same time, the reduced disk masses mean that type II migration almost always proceeds at less than the disk advection speed. As a result, gas builds up on the outside of a planet's gap, and drains away on the inside, so that gaps soon grow into holes.

In a sufficiently massive and/or viscous disk, the global evolution is, at least initially, in a ''disk-dominated'' regime: a gas giant is born and migrates to the inner disk edge before the next gas giant is formed, in other words, the disk clears the planets. If the disk mass is sufficiently low, on the other hand, type II migration is slow enough that the the evolution is ''planet-dominated'': the next planet forms while the previous one is still in the process of migrating and (if its gap is not yet clean) accreting gas. When this happens, the gas supply to the previous planet–the inner of the two–is reduced or cut off altogether. This throttles both its growth (if any) and its migration. In essence, the planets clear the disk from the inside out, with an ever-increasing number of planets left in an ever-growing inner disk hole. Whether planets accumulate in a hole, or, as in the disk-dominated case, all the way down at the original inner disk edge, inward migration tends to produce crowded systems of planets. This, in turn, leads to the excitation of eccentricities as planets lock into and migrate in mean-motion resonances (e.g., [37, 41, 43]), or simply scatter each other [50, 51]. Unfortunately, Thommes *et al.* [49] are unable to make a direct comparison to the observed exoplanet eccentricity distribution: for computational reasons–to prevent the need for an unfeasibly small timestep–their simulations have an inner boundary at 0.1–0.25 AU, with every body crossing it considered to have collided with the star. Thus bodies with high eccentricities, i.e., small pericenters, are preferentially removed.

The simulations of Thommes *et al.* [49] span the range of disk masses and viscosities consistent with observations, and reveal that, despite the large stochastic variations between individual outcomes, there are clear trends with the properties of the birth disk (Figure 9.1): At one extreme, low disk mass combined with high viscosity results in systems that produce no gas giants at all. At the other extreme, high disk mass combined with low viscosity results in the production of numerous gas giants; most undergo considerable inward migration, and many acquire large eccentricities.

We can understand these results in terms of two fundamental timescales in a planet-forming disk: one is the gas disk depletion time $\tau_{\rm disk}$, shorter for higher disk viscosities. The other is the time to form the first gas giant, the $\tau_{\rm giant}$. Note that the time to form a gas giant at a given location in the disk is the sum of the time to form the core, $\tau_{\rm core}$, and the time for that core to undergo runaway gas accretion, $\tau_{\rm KH}$[1)]; $\tau_{\rm giant}$ is then the minimum gas-giant formation time anywhere in the disk. As shown in Figure 9.2 there is an initial burst of planets born in rapid succession beginning at $\tau_{\rm giant}$, which gradually slows down. The crucial factor determining how a given system's formation history will play out is the timing of this burst. In cases with $\tau_{\rm giant} > \tau_{\rm disk}$ (Figure 9.1, lower right), the gas is removed before any gas giant has a chance to form, leaving behind systems consisting solely of rocky–icy bodies. In cases with $\tau_{\rm giant} < \tau_{\rm disk}$ (upper-left region of Figure 9.1), planets are born into a substantial gas disk, and such systems generally produced a number of gas giants that migrate a considerable distance inward.

The results of Thommes *et al.* [49] also suggest how the solar system fits into the picture. In systems with $\tau_{\rm giant} \approx \tau_{\rm disk}$, gas giants do form but undergo only modest migration and eccentricity growth; thus, it is here that one would most naturally expect to find a solar system-like outcome. Figure 9.1 shows that these cases occupy a relatively narrow region within the parameter space, roughly a diagonal line extending from ($\alpha = 10^{-3}, M_{\rm disk} = 0.03 M_{\rm Sun}$) to ($\alpha = 10^{-2}, M_{\rm disk} = 0.08 M_{\rm Sun}$). Thus, whatever the true distribution of disks within Figure 9.2–unless disks with $\tau_{\rm giant} \approx \tau_{\rm disk}$ are somehow preferred–it is likely that only a minority will lie within this region. Furthermore, even within this subset, there are large stochastic variations, as evidenced by Figure 9.1, in only one of the outcomes ($\alpha = 3 \times 10^{-3}, M_{\rm disk} = 0.05 M_{\rm Sun}$) do the gas giants bear a reasonable resemblance to Jupiter and Saturn.

In all these simulations, the formation of a gas giant brings with it violent scattering of neighboring smaller bodies, including other cores about to undergo runaway gas accretion themselves. Such scattering has been proposed as the origin of Uranus and Neptune [54], with dynamical friction from the remnant outer planetesimal disk (not modeled here) serving to prevent their ejection and ultimately recircularize their orbits. Thus, though Jupiter and Saturn analogs may not be common, it is likely that the opposite is true for Uranus and Neptune analogs.

1) We approximate the envelope collapse time as its Kelvin–Helmholtz time $\tau_{\rm KH}$, i.e., the time needed to radiate away a significant part of its gravitational potential energy.

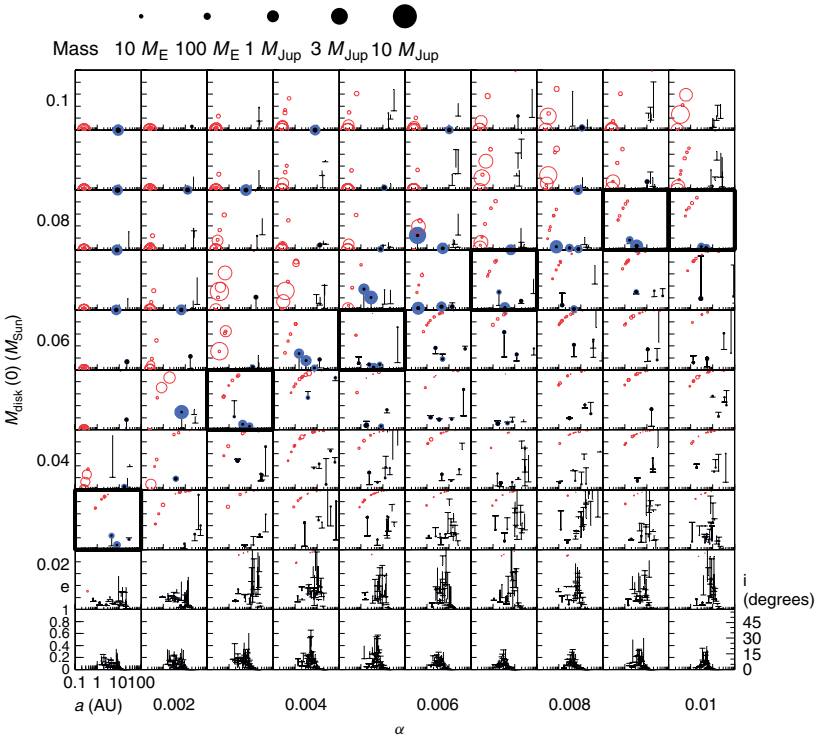


Figure 9.1 Final outcomes of a set of 100 simulations, spanning 10^{-3} to 10^{-2} in viscosity parameter α, and $0.01M_{Sun}$ to $0.1M_{Sun}$ in initial disk mass, $M_{disk}(0)$. Simulations are ended after 0.5 gigayear has elapsed, or if they fail to produce any gas giants within the gas disk's lifetime. The semimajor axes, eccentricities, and inclinations of planets are plotted, as well as the relative solid and gas content of each planet (solid core: black; gaseous envelope: dark blue; size $\propto$ mass$^{1/3}$, see mass scale at top). To keep computational cost reasonable, simulations have an inner boundary at 0.25 AU, beyond the initial inner edge of the gas disk; any body that crosses the boundary is removed, and a red circle is plotted showing its orbital elements and mass at the time of removal. Toward high M_{disk} and low α, planets form early and often during the gas disk's lifetime, most migrate extensively, and many acquire high eccentricities in the process. Toward low M_{disk} and high α, planet formation is too slow to produce any gas giants during the disk lifetime. Between these two extremes is a relatively narrow boundary region (thick borders) in which gas giants migrate little and remain at low eccentricity, thus producing some outcomes similar to the solar system. No gas giants form in disks of $0.02M_{Sun}$ or less; for comparison, this is the approximate lower limit on the solar system's birth disk, called the *minimum mass solar nebula* (MMSN) [52]. This figure has been adapted from [49].

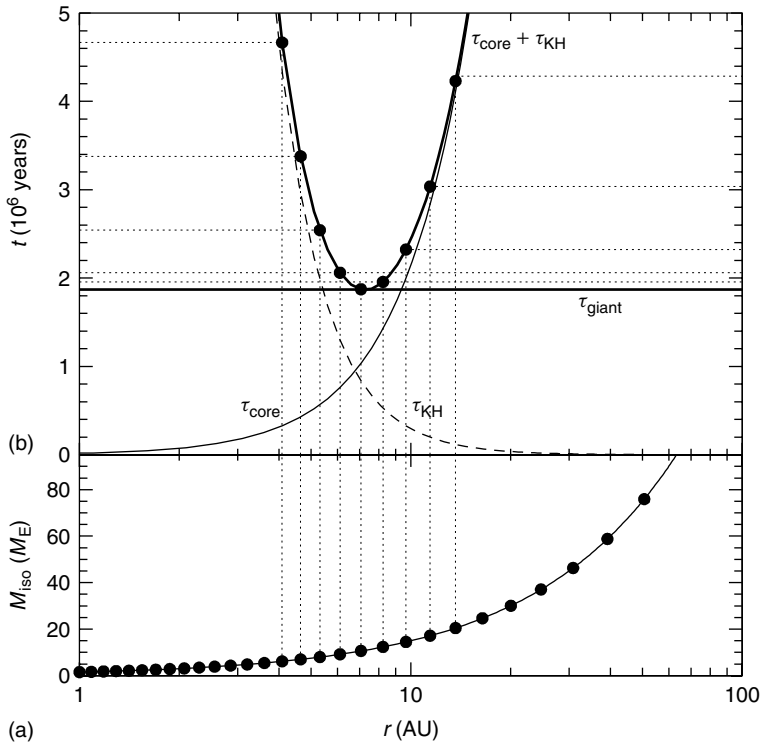


Figure 9.2 Approximate timing and location of gas-giant formation in a protoplanetary disk. (a) The final or "isolation" mass of solid cores (black dots) M_{iso}, with spacing between successive cores taken from planet formation simulations [20]. (b) The time (thick solid curve) for a core (black dots; vertical dotted lines connect to corresponding core in (a)) to become a gas giant (horizontal dotted lines show times for individual protoplanets). This is approximated as the sum of the time for the core to reach its final mass, τ_{core} (thin solid curve), and the time for the core to undergo runaway gas accretion, taken to be its Kelvin–Helmholtz time [53], τ_{KH} (dashed curve). As in more detailed calculations [3], it is found that gas-giant formation commences at one particular radius, which, for typical parameters, lies in or near the Jupiter–Saturn region; in this case at 7 AU, and at time τ_{giant} just under 2 megayears. Giant formation begins in a burst, with several planets growing in rapid succession, then slows down as it spreads to larger and smaller radii. In practice, once an inner hole forms in the gas disk, formation is constrained to progress only outwards. This figure has been adapted from [49].

The results of Thommes *et al.* [49] also provide qualitative agreement with some of the key features observed among the ensemble of exoplanets, if observational bias is taken into account (see the supplementary online material of Thommes *et al.* [49] for details). Like Ida and Lin [44], they find a strong correlation of stellar metallicity with gas-giant formation likelihood, for fundamentally the same reason: a higher metallicity disk contains more solids for a given amount of gas, and thus solid cores grow faster and larger. With reference to Figure 9.2, this shortens τ_{giant}

for a given gas disk mass, thus increasing the area of the giant-planet-forming region at the expense of the barren region.

The observed distribution of planet orbital periods is also reproduced, in the sense that there is a peak toward both the longest and the shortest periods. The model provides a clear interpretation: the peak at the long end corresponds to the periods with which giant planets are born, whereas the peak at the short end corresponds to the smallest periods to which migration can deliver them–in other words, where they ''fall off'', the inner edge of the disk. The long- and short-period peaks thus reflect the higher probability of a planet ending up in either the source or sink region at the time the disk dissipates, rather than in transit between the two.

Another trend reproduced in the simulations of Thommes *et al.* [49] is that more eccentric planets tend to also be more massive. An inspection of the simulations shows two causes for this. First, and most generally, the more massive planets tend to be born from the more massive disks; such disks also produce larger numbers of gas giants as well as more extensive migration, increasing the chances for eccentricity-raising planet–planet interactions. Systems consisting only of moderate-mass gas giants, in contrast, tend to also have less dramatic interactions, leaving them with more modest eccentricities. Secondly, there is also a specific mechanism that systematically correlates eccentricity and mass: As a planet becomes more eccentric, it increases the width of its accretion zone. This allows it to grow beyond its zero-eccentricity gap-opening mass. The frequently repeated configuration of multiple planets in a centrally cleared disk, locked together into successive mean-motion resonances, provides the perfect setting for this scenario to take place. As the outer planet is driven inward by the disk, the eccentricities of all planets increase, allowing the outermost one, which is still accreting from the disk, to grow to larger and larger mass. Thus, the most massive planets in the overall population tend not to have low eccentricities. It should be pointed out, though, that in this scenario, the most massive planet *within a given multiplanet system* is not the highest-eccentricity one, since its resonantly locked inner neighbors are driven to even higher eccentricities (though some of these may eventually be ejected).

9.7 Summary and Discussion

We have presented here a summary of the current state of progress in understanding planet migration due to planet–gas disk interactions. Our emphasis has been on ultimately trying to fit these findings together into a ''big picture'' of how a planetary system evolves during its birth: how the concurrent processes of planet growth, migration, interaction of planets with their neighbors, and the evolution of the disk all play off each other. In the end, rather than complicating the picture, this approach actually gives us a clearer view.

Planet migration has long presented a challenge to understanding planet formation, making it appear that protostellar disks essentially kill planets (or at least

turn them all into ultrahot Jupiters/Neptunes) faster than they produce them. As discussed in Section 9.2 and in [49], the resolution of this problem appears to be simply that migration–both type I and type II–eventually trails off. The results of Thommes *et al.* [49] demonstrate that multiplicity effects help: the more planets there are in a system, the greater is the angular momentum that the disk must extract to drive them all toward the star, and thus the more the migration slows down. And of course, all the gas mass locked up in planets comes at the expense of disk mass. All this makes the planet mortality rate via migration significantly lower than in the classic so-called "last of the Mohicans" scenario, [15, 55–58], wherein all planets, except the last few, fall into the star.

Our own solar system appears to fit into this picture as an instance of a system that grew two gas giants barely in time, just before running out of gas, a sort of "last and only of the Mohicans". As a result of being born into a tenuous disk, Jupiter and Saturn simply never underwent much migration; their current orbits are close to their primordial locations and, without any opportunity for close encounters with each other, they retained low eccentricities. More such systems almost certainly exist; however, because some coincidence in timing, $\tau_{\rm giant} \approx \tau_{\rm disk}$, is a necessary (though not sufficient) condition, they are likely to be somewhat uncommon.

References

1 Mayor, M. and Queloz, D. (**1995**) A jupiter-mass companion to a solar-type star. *Nature*, **378**, 355–359.

2 Bodenheimer, P., Hubickyj, O., and Lissauer, J.J. (**2000**) Models of the *in situ* formation of detected extrasolar giant planets. *Icarus*, **143**, 2–14.

3 Ikoma, M., Nakazawa, K., and Emori, H. (**2000**) Formation of giant planets: dependences on core accretion rate and grain opacity. *The Astrophysical Journal*, **537**, 1013–1025.

4 Thommes, E.W., Duncan, M.J., and Levison, H.F. (**2003**) Oligarchic growth of giant planets. *Icarus*, **161**, 431–455.

5 Goldreich, P. and Tremaine, S. (**1980**) Disk-satellite interactions. *The Astrophysical Journal*, **241**, 425–441.

6 Lin, D.N.C. and Papaloizou, J. (**1986**) On the tidal interaction between protoplanets and the protoplanetary disk. III-Orbital migration of protoplanets. *The Astrophysical Journal*, **309**, 846–857.

7 Marcy, G., Butler, R.P., Fischer, D. *et al.* (**2005**) Observed properties of exoplanets: masses, orbits, and metallicities. *Progress of Theoretical Physics Supplement*, **158**, 24–42.

8 Haisch, K.E., Lada, E.A., and Lada, C.J. (**2001**) Disk frequencies and lifetimes in young clusters. *The Astrophysical Journal*, **553**, L153–L156.

9 Goldreich, P. and Sari, R. (**2003**) Eccentricity evolution for planets in gaseous disks. *The Astrophysical Journal*, **585**, 1024–1037.

10 Papaloizou, J.C.B., Nelson, R.P., Kley, W. *et al.* (**2007**) Disk-planet interactions during planet formation, in *Protostars and Planets V* (eds B. Reipurth, D. Jewitt, and K. Keil), University of Arizona Press, pp. 655–668.

11 Ward, W.R. (**1997**) Protoplanet migration by nebula tides. *Icarus*, **126**, 261–281.

12 Menou, K. and Goodman, J. (**2004**) Low-mass protoplanet migration in T tauri α-disks. *The Astrophysical Journal*, **606**, 520–531.

13 Papaloizou, J.C.B. and Larwood, J.D. (**2000**) On the orbital evolution and growth of protoplanets embedded in a gaseous disc. *Monthly Notices of the Royal Astronomical Society*, **315**, 823–833.

14 Tanaka, H., Takeuchi, T., and Ward, W.R. (**2002**) Three-dimensional interaction between a planet and an isothermal

gaseous disk. I. Corotation and lindblad torques and planet migration. *The Astrophysical Journal*, **565**, 1257–1274.

15 Chambers, J. (**2006**) A semi-analytic model for oligarchic growth. *Icarus*, **180**, 496–513.

16 Trilling, D.E., Lunine, J.I., and Benz, W. (**2002**) Orbital migration and the frequency of giant planet formation. *Astronomy and Astrophysics*, **394**, 241–251.

17 Ida, S. and Lin, D.N.C. (**2004**) Toward a deterministic model of planetary formation. I. A desert in the mass and semi major axis distributions of extrasolar planets. *The Astrophysical Journal*, **604**, 388–413.

18 Alibert, Y., Mordasini, C., Benz, W., and Winisdoerffer, C. (**2005**) Models of giant planet formation with migration and disc evolution. *Astronomy and Astrophysics*, **434**, 343–353.

19 Thommes, E.W. and Murray, N. (**2006**) Giant planet accretion and migration: surviving the type I regime. *The Astrophysical Journal*, **644**, 1214–1222.

20 Thommes, E.W., Nilsson, L., and Murray, N. (**2007**) Overcoming migration during giant planet formation. *The Astrophysical Journal*, **656**, L25–L28.

21 Kokubo, E. and Ida, S. (**1998**) Oligarchic growth of protoplanets. *Icarus*, **131**, 171–178.

22 Ida, S. and Lin, D.N.C. (**2008**) Toward a deterministic model of planetary formation. IV. Effects of type I migration. *The Astrophysical Journal*, **673**, 487–501.

23 Laughlin, G., Steinacker, A., and Adams, F.C. (**2004**) Type I planetary migration with MHD turbulence. *The Astrophysical Journal*, **608**, 489–496.

24 Nelson, R.P. and Papaloizou, J.C.B. (**2004**) The interaction of giant planets with a disc with MHD turbulence - IV. Migration rates of embedded protoplanets. *Monthly Notices of the Royal Astronomical Society*, **350**, 849–864.

25 Baruteau, C. and Masset, F. (**2008**) Type I planetary migration in a self-gravitating disk. *The Astrophysical Journal*, **678**, 483–497.

26 Bryden, G., Chen, X., Lin, D.N.C. *et al.* (**1999**) Tidally induced gap formation in protostellar disks: gap clearing and suppression of protoplanetary growth. *The Astrophysical Journal*, **514**, 344–367.

27 Syer, D. and Clarke, C.J. (**1995**) Satellites in discs: regulating the accretion luminosity. *Monthly Notices of the Royal Astronomical Society*, **277**, 758–766.

28 Masset, F.S. and Papaloizou, J.C.B. (**2003**) Runaway migration and the formation of hot Jupiters. *The Astrophysical Journal*, **588**, 494–508.

29 Artymowicz, P. (**2004**) Dynamics of gaseous disks with planets, in *Debris Disks and the Formation of Planets*, Astronomical Society of the Pacific Conference Series, vol. 324, (eds L. Caroff, L.J. Moon, D. Backman, and E. Praton), ASP Conference Series, p. 39.

30 Papaloizou, J.C.B. (**2005**) Disk planet interactions and early evolution in young planetary systems. *Celestial Mechanics and Dynamical Astronomy*, **91**, 33–57.

31 D'Angelo, G. and Lubow, S.H. (**2008**) Evolution of migrating planets undergoing gas accretion. *The Astrophysical Journal*, **685**, 560–583.

32 Masset, F. and Snellgrove, M. (**2001**) Reversing type II migration: resonance trapping of a lighter giant protoplanet. *Monthly Notices of the Royal Astronomical Society*, **320**, L55.

33 Morbidelli, A. and Crida, A. (**2007**) The dynamics of Jupiter and Saturn in the gaseous protoplanetary disk. *Icarus*, **191**, 158–171.

34 Bryden, G., Różyczka, M., Lin, D.N.C., and Bodenheimer, P. (**2000**) On the interaction between protoplanets and protostellar disks. *The Astrophysical Journal*, **540**, 1091–1101.

35 Kley, W. (**2000**) On the migration of a system of protoplanets. *Monthly Notices of the Royal Astronomical Society*, **313**, L47–L51.

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

37 Lee, M.H. and Peale, S.J. (**2002**) Dynamics and origin of the 2:1 orbital resonances of the GJ 876 planets. *The Astrophysical Journal*, **567**, 596–609.

38 Murray, N., Paskowitz, M., and Holman, M. (**2002**) Eccentricity evolution of migrating planets. *The Astrophysical Journal*, **565**, 608–620.

39 Nelson, R.P. and Papaloizou, J.C.B. (**2002**) Possible commensurabilities among pairs of extrasolar planets. *Monthly Notices of the Royal Astronomical Society*, **333**, L26–L30.

40 Kley, W., Peitz, J., and Bryden, G. (**2004**) Evolution of planetary systems in resonance. *Astronomy and Astrophysics*, **414**, 735–747.

41 Adams, F.C. and Laughlin, G. (**2003**) Migration and dynamical relaxation in crowded systems of giant planets. *Icarus*, **163**, 290–306.

42 Moorhead, A.V. and Adams, F.C. (**2005**) Giant planet migration through the action of disk torques and planet planet scattering. *Icarus*, **178**, 517–539.

43 Lee, A.T. Thommes, E.W., and Rasio, F.A. (**2008**) Resonance trapping in protoplanetary disks. I. Coplanar systems. *The Astrophysical Journal*, **691**, 1684–1696.

44 Ida, S. and Lin, D.N.C. (**2004**) To a deterministic model of planetary formation. II. The formation and retention of gas giant planets around stars with a range of metallicities. *The Astrophysical Journal*, **616**, 567–572.

45 Ida, S. and Lin, D.N.C. (**2005**) Toward a deterministic model of planetary formation. III. Mass distribution of short-period planets around stars of various masses. *The Astrophysical Journal*, **626**, 1045–1060.

46 Pollack, J.B., Hubickyj, O., Bodenheimer, P. *et al.* (**1996**) Formation of the giant planets by concurrent accretion of solids and gas. *Icarus*, **124**, 62–85.

47 Fischer, D.A. and Valenti, J.A. (**2003**) Metallicities of stars with extrasolar planets, in *Scientific Frontiers in Research on Extrasolar Planets*, vol. 294, Astronomical Society of the Pacific Conference Series (eds D. Deming and S. Seager), ASP Conference Series, pp. 117–128.

48 Santos, N.C., Israelian, G. and Mayor, M. (**2004**) Spectroscopic [Fe/H] for 98 extrasolar planet-host stars. Exploring the probability of planet formation. *Astronomy and Astrophysics*, **415**, 1153–1166.

49 Thommes, E.W., Matsumura, S., and Rasio, F.A. (**2008**) Gas disks to gas giants: simulating the birth of planetary systems. *Science*, **321**, 814–817.

50 Jurić, M. and Tremaine, S. (**2008**) Dynamical origin of extrasolar planet eccentricity distribution. *The Astrophysical Journal*, **686**, 603–620.

51 Chatterjee, S., Ford, E.B., Matsumura, S., and Rasio, F.A. (**2008**) Dynamical outcomes of planet-planet scattering. *The Astrophysical Journal*, **686**, 580–602.

52 Hayashi, C. (**1981**) Structure of the solar nebula, growth and decay of magnetic fields and effects of magnetic and turbulent viscosities on the nebula. *Progress of Theoretical Physics Supplement*, **70**, 35–53.

53 Bryden, G., Lin, D.N.C., and Ida, S. (**2000**) Protoplanetary formation. I. Neptune. *The Astrophysical Journal*, **544**, 481–495.

54 Thommes, E.W., Duncan, M.J., and Levison, H.F. (**2002**) The formation of uranus and neptune among Jupiter and Saturn. *The Astronomical Journal*, **123**, 2862–2883.

55 Gonzalez, G. (**1997**) The stellar metallicity-giant planet connection. *Monthly Notices of the Royal Astronomical Society*, **285**, 403–412.

56 Laughlin, G. and Adams, F.C. (**1997**) Possible stellar metallicity enhancements from the accretion of planets. *The Astrophysical Journal*, **491**, L51+.

57 Lin, D.N.C. (**1997**) Planetary formation in protostellar disks, in *IAU Colloq. 163: Accretion Phenomena and Related Outflows*, vol, Astronomical Society of the Pacific Conference Series (eds D.T. Wickramasinghe, G.V. Bicknell, and L. Ferrario), ASP Conference Series, p. 321.

58 Trilling, D.E., Benz, W., Guillot, T. *et al.* (**1998**) Orbital evolution and migration of giant planets: modeling extrasolar planets. *The Astrophysical Journal*, **500**, 428–439.

10
Planets in Mean-Motion Resonance

Wilhelm Kley

10.1
Introduction

At the time of writing (July 2009), about 30 extrasolar planetary systems containing multiple planets have been detected. Among those, a high fraction (about 30%) seems to be located in a mean-motion resonance (MMR). In an MMR, the orbital periods of the involved planets are commensurable and their ratio can be described by two small integers. The most frequent commensurability found in the extrasolar planetary systems is the 2 : 1 MMR, though others such as 3 : 1, 4 : 1, or 3 : 2 MMR may also exist. Current scenarios of planet formation allow for the creation of planets at arbitrary radii. Hence, it is strongly believed that the special resonant configurations did not form *in situ* but are a result of the evolution of the whole planetary system. Thus, the mere existence of resonant configurations may provide additional information about the evolutionary history of planetary systems.

In the solar system, the 3 : 2 MMR between Neptune and Pluto was caused by an outward migrating Neptune. Similarly, extrasolar planets can be driven into a resonant configuration through the operation of a dissipative mechanism that is able to change the energy of the orbits, that is, the corresponding semimajor axis of the objects. In the context of planet formation of massive planets, the interaction of the protoplanets with the ambient gaseous protoplanetary disk is able to provide a differential migration between two planets, and, upon reaching commensurability, resonant capture may occur. In this chapter, we briefly summarize the current observational evidence for MMRs in extrasolar planetary systems and then outline, in more detail, possible formation scenarios and the particular conditions that may have led to the presently observed states.

10.2
Extrasolar Systems in Mean-Motion Resonance

In this case, two planets are located inside a mean-motion resonance (MMR) when their orbital periods P (or mean motions n) are in a ratio of two small integers.

Formation and Evolution of Exoplanets. Edited by Rory Barnes

ISBN: 978-3-527-40896-2

In the following, we denote the inner planet (of the resonant configuration under consideration) with a subscript "1" and the outer planet with a subscript "2". We restrict all our considerations to the planar case where all planetary orbits lie in one plane. The presence of an MMR is indicated if at least one of the so-called resonant angles librates, that is, its variation does not cover the full $(0:2\pi)$ domain (see [1] for details; see also Chapters 2 and 3). For a general $p:q$ commensurability, the resonant angles are defined (planar case) as

$$\Phi_{p,q,k} = p\lambda_2 - q\lambda_1 - p\varpi_2 + q\varpi_1 + k(\varpi_2 - \varpi_1) \tag{10.1}$$

where λ_1, λ_2 denote the mean longitudes and ϖ_1, ϖ_2 the longitude of periapse for the inner (1) and the outer planet (2). The positive integers p and q satisfy $p > q$, and there are $p - q + 1$ possible k values with $q \leq k \leq p$. A planetary system is said to be in resonance if at least one of the angles $\Phi_{p,q,k}$ is in libration. Of the $p - q + 1$ resonant angles, at most two are linearly independent, that is, for a resonant configuration either all angles librate or only one librates [2]. In the following, we focus mainly on the 2 : 1 MMR, which appears to be the most frequent configuration, well established in four exoplanetary systems. In this case, $p = 2$ and $q = 1$, and we have only two resonant angles ($k = 1, 2$), which are denoted as

$$\Theta_i = 2\lambda_2 - \lambda_1 - \varpi_i \tag{10.2}$$

with $i \in \{1, 2\}$. Often the second angle Θ_2 is replaced by

$$\Delta\varpi = \varpi_2 - \varpi_1 \quad = \Theta_2 - \Theta_1 \tag{10.3}$$

If the second resonant angle $\Delta\varpi$ is in libration, that is, the periapses of the two orbits are basically aligned, the configuration is in an *apsidal corotation Resonance* (ACR).

A recent summary of the properties of all exoplanetary systems with emphasis on resonant systems has been given by Udrey *et al.* [3]. Among the multiplanet systems detected (up to now), possibly 25% reside in an MMR. Table 10.1 gives an overview of the main orbital characteristics of the present candidate systems. The first four well-established cases are in a 2 : 1 MMR, whereas the following systems, which are in 3 : 1, 4 : 1, and 5 : 1 MMR, need more observational data to determine their dynamical state more accurately in the future. The systems below the double lines are very uncertain, but we list them in the table just to indicate the current status of observation and interpretation. In a recent paper by Wright *et al.* [4], the parameters of multiplanet systems are listed.

The best investigated system certainly is GJ 876, which was also the first planetary system where an MMR between two planets was discovered [17]. Because of the short orbital periods of the planets (≈30 and 60 days), many periods have been observed, and the inferred orbital parameters are the most accurate of all the systems quoted in Table 10.1. The large mass ratios (planetary mass over stellar masses) in GJ 876 leads to a strongly interacting system such that resonant interactions can be seen straight away. In this case, a standard two-Keplerian fit does not lead to reliable and stable orbits, and the procedure for fitting has to be based on dynamical n-body calculations [18]. This system is clearly engaged in a 2 : 1 MMR with ACR, where Θ_1 and ϖ are both librating around zero with small amplitudes (7° and 34°) [19].

Table 10.1 Dynamical properties of candidate planetary systems in mean-motion resonance (MMR).

System	No	P (d)	$M \sin i$ ($M_{\rm Jup}$)	a (AU)	e	ϖ (degree)	M_* [$M_\odot$]	References/ comments
GJ 876	c	30.57	0.56	0.13	0.24	159	0.32	[5]
(2 : 1)	b	60.13	1.89	0.21	0.04	163		
HD 82943	b	217.9	1.40	0.74	0.46	135	1.15	[6]
(2 : 1)	c	456.6	1.78	1.19	0.36	140		
HD 128311	b	464	1.56	1.11	0.38	80	0.80	[7]
(2 : 1)	c	910	3.08	1.73	0.21	21		
HD 73526	b	188	2.90	0.66	0.19	203	1.08	[8]
(2 : 1)	c	377	2.50	1.05	0.14	13		
HD 45364	b	227	0.19	0.68	0.17	163	0.82	[9]
(3 : 2)	c	343	0.66	0.90	0.10	7.4		
HD 160691	d	130	0.52	0.92	0.067	189	1.00	[10]
(2 : 1)	b	643	1.68	1.50	0.128	22		(μ Ara; uncertain)
55 Cnc	b	14.67	0.78	0.11	0.02	131	0.95	[11]
(3 : 1)	c	43.93	0.21	0.24	0.44	244		(questionable)
HD 60532	b	201	1.05	0.76	0.28	352	1.44	[12]
(3 : 1)	c	605	2.49	1.58	0.03	136		[13]
HD 108874	b	394	1.36	0.07	1.06	250	1.00	[14]
(4 : 1)	c	1600	1.02	0.25	2.68	20		(uncertain)
GJ 317	b	692	1.20	0.95	0.19	–	0.24	[15]
(4 : 1)	c	2800	0.83	2.35	0.42	–		(to be confir med)
HD 202206	b	255	17.42	0.83	0.43	161	1.15	[16]
(5 : 1)	c	1383	2.43	2.54	0.26	55		(circumbinary; tbc)

In the table, we have included only those planets (denoted chronologically by ''b'', ''c'', etc.) that are involved in the MMR, even though some systems (e.g., GJ 876 or 55 Cnc) may contain additional planets. Here, P denotes the orbital periods, $M \sin i$ the mass of the planets, a the semimajor axis, e the eccentricity, ϖ the position of periastron, and M_* the stellar mass. The planetary mass and the distance a are quantities derived assuming the quoted stellar mass. The references quoted refer to recent publications from which the listed data have been taken. Additional comments indicate the current status of the observations.

Even though the possibility of a 2 : 1 MMR in the system HD 82943 has been suggested by the Geneva group in 2001, the orbital elements are still not very well determined. It has been noted that the published orbital elements [20] lead to dynamically unstable solutions [21–23]. Through dynamical integrations and stability analyses, it is possible to constrain the allowed orbital parameters. The system appears to lie indeed in a 2 : 1 resonance, where either both resonant angles are librating around zero (ACR) or one may be circulating [23]. The most recent analysis by Beauge *et al.* [6] demonstrates again that the scarce observational data make it difficult to constrain the parameters further. Their estimates favor aligned apsidal corotation, and the data listed in Table 10.1 are taken from that paper. Interestingly, as the authors point out, it is also possible to match the observed data set possibly with a three-planet fit in a Laplacian resonance. To constrain

the current state of that particular system, clearly more observational data are required.

The two recently discovered systems, HD 128311 and HD 73526, are dynamically very similar to each other. In both systems, the planets are engaged in 2 : 1 MMRs, with a large temporal variation of at least one of the planetary eccentricities. In the first system, apsidal corotation is "broken" with only θ_1 librating and $\Delta\varpi$ circulating [24]. In the second case (HD 73526), a similar configuration has been suggested by Tinney *et al.* [8], while dynamical fits allow also for apsidal corotation [25]. In Figure 10.1, we contrast the dynamical state of HD 128311 with that of the classic GJ 876 case. In GJ 876, the small libration amplitude of the resonant angles implies small eccentricity oscillations, whereas in HD 128311, the circulation of Θ_2 is accompanied by large eccentricity variations.

Now, the system 55 Cnc is known to harbor five planets [26]. Because of this wealth of planets, the system is compared frequently to the solar system, which is particularly interesting because the fifth new planet lies just within the habitable zone. These new data imply that planets b and c in 55 Cnc are close to, but presently not in, a 3 : 1 MMR. We kept the system nevertheless in the list above, as more refined new data may possibly change this again.

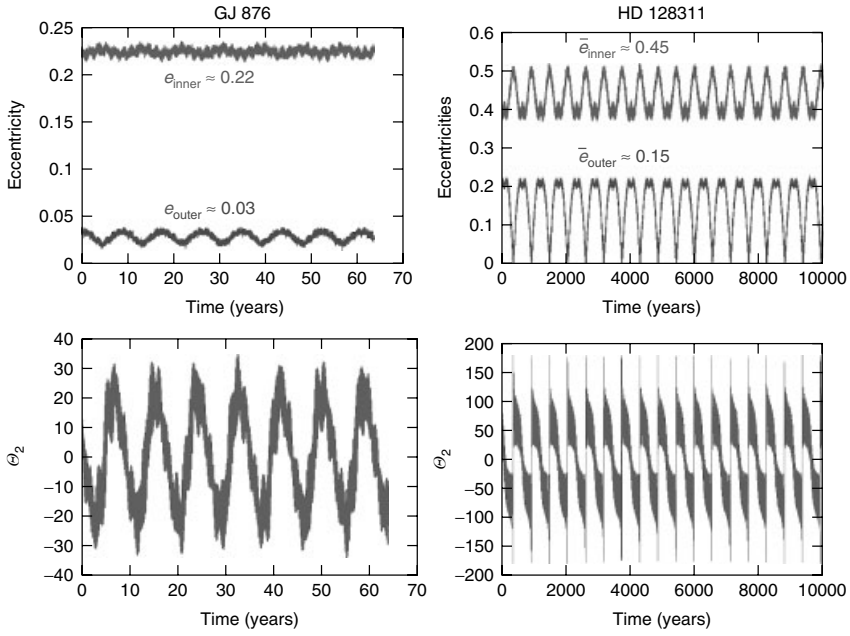


Figure 10.1 Evolution of the orbital elements of GJ 876 and HD 128311 as obtained through three-body integrations using the observed data. This plot illustrates the different dynamical state of a system in apsidal corotation (left panel for GJ 876) and broken apsidal resonance, that is, only one resonant argument librates (right panel for HD 128311).

Several new systems have been suggested to be in resonance, which are also listed in the table. Many of these require more observational data to be definitively confirmed. If the presently tentative status of these resonances is confirmed, this will significantly change the statistical occurrence of resonances. Additionally, in the system 47 UMa, a 5 : 2 MMR has been suggested [27] and a stability analysis has been performed [28]. Recently, the possibility of a system in a 3 : 2 MMR has been suggested [9].

In the following, we concentrate on formation scenarios for systems engaged in a 2 : 1 MMR and also comment briefly on higher order resonances.

10.3 Planetary Migration

Since planets can be assembled at arbitrary locations in the disk, the direct formation of resonant configurations (which require very special locations in the disk) seems very unlikely. Dissipative processes that alter the semimajor axis of embedded planets constitute possible mechanisms to bring planets into resonance. In addition to the observed hot Jupiters and Neptunes orbiting very close to the host star, the existence of resonant configurations is usually taken as proof that planets have indeed migrated through the protoplanetary disk.

Recently, however, the possibility of forming such systems directly through scattering events has been considered [29]. Through direct *N*-body simulations (of different random initial conditions), the authors estimate that the fraction of systems ending up in resonances is in agreement with the fraction of the phase space volume occupied by these resonances. Clearly, it will be difficult to form a system in a deep resonance with apsidal corotation (such as GJ 876) through scattering. Systems with only one angle in libration may form easier, but it is not clear yet whether this scenario can reproduce the observed dynamical properties of resonant configurations. The small likelihood of forming systems, such as HD 82943, through a direct formation scenario has also recently been pointed out by Michtchenko *et al.* [30]. Consequently, in this chapter, we focus on the standard formation mechanism through disk-driven convergent migration.

10.3.1 Planet–Disk Interaction

The mechanism responsible to change the semimajor axis of a forming planet has to alter its orbital energy and hence must be dissipative. In early studies of satellite–disk interaction with application to the Saturnian ring system and embedded moonlets, it has been shown that gravitational interaction and subsequent angular momentum transfer between the embedded moon and the small disk particles can lead to a change in the semimajor axis of the Moon [31]. It was soon realized that a very similar process must have taken place between the gaseous protoplanetary disk and the embedded young protoplanets in the solar system [32,

33]. Similarly, dissipative effects may occur for a planet embedded in a planetesimal disk [34]. Here, we restrict ourselves to the interaction of a forming giant planet with the gaseous protoplanetary disk as all the observed exoplanets in resonant configurations are rather massive. For simplification, we consider only coplanar cases.

The embedded planet disturbs the surrounding disk due to its gravitational action. The disturbances (i.e., sound waves) created in the disk superimpose and create (for small planetary masses) two wakes or spiral arms in the ambient disk, the outer one trailing and the inner one leading [35]. With increase in the planetary mass, the wake amplitudes increase and they transform into shock waves that lead to dissipation. As these are stationary in a coordinate system corotating with respect to the planet, the outer (inner) spiral is faster (slower) than the local disk material. In other words, the outer wave constitutes a positive angular momentum wave and the inner a negative one. Then, upon wave dissipation, positive angular momentum is given to the outer disk and negative to the inner with respect to the local unperturbed disk values. The result is that the material is pushed away from the planetary orbit, which eventually leads to reduction in the density, that is, to the opening of a gap. This gap formation process of embedded planets has been described in many works, and it is not discussed here [36–38]. In a European-wide collaborative effort, a standard setup has been recently used to compare different numerical codes on the same planet–disk problem [39], which may serve as the ideal reference and test bed for all such future studies. To open a significant deep gap with a depth of at least half of the unperturbed density, a Saturn mass planet ($1/3\ M_{\mathrm{Jup}}$) is required. The exact depth and the width of the gap region depend on a detailed balance of gravitational, viscous, and pressure torques [38, 40, 41]. Increasing the planet mass deepens the gap, whereas increased viscosity and pressure tend to close it.

The nonaxisymmetric density enhancements in such a disk, the spiral arms, will pull gravitationally on the planet or, phrased differently, exert a torque on it. The outer spiral pulls the planet back and slows it down, whereas the inner spiral will accelerate it. Hence, the outer disk will push the planet inward and the inner outward. The detailed balance of the two contributions will determine the net force acting on the planet and will also determine its direction of motion through the disk. As a result, the semimajor axis of the planet will change and lead to a radial migration of the planet through the disk [32, 33]. In addition to the torque contribution of the spiral arms, the so-called Lindblad torques, there is a contribution coming from material in the coorbital region of the planet. The magnitude and the direction of this horseshoe drag depend on the local vortensity-density gradient. The coorbital material is naturally more important for determining the migration properties when the gap region is less cleared, that is, for smaller planetary masses. It has been shown that under typical conditions in the protoplanetary nebula, the timescale for a Jupiter-type planet to migrate from about 5 AU all the way to the star is only about 10^5 years [42]. A comprehensive overview of all aspects of planetary migration theory has recently be given by Papaloizou *et al.* [43]. The problem of too rapid migration for low-mass planets may be solved by recent theoretical progress that has been made in realizing that the

thermodynamic properties of the disk can play a crucial role in the determination of the magnitude and the direction of smaller mass planets [44–46].

The observed resonant planetary systems have masses in the range of Jupiter and should have opened a rather deep gap in the disk. In this case, the coorbital matter plays only a minor role in determining the migration properties and we can focus on the action of the Lindblad torques generated through the spiral arms. Quite generally, the outer arm will pull stronger and the planet will migrate inward. In any event, in well-resolved multidimensional hydrodynamic calculations, all contributions will automatically be taken care of.

10.4 Resonant Capture through Convergent Migration

Having determined the acting agent, the protoplanetary disk, that is responsible for moving planets around in their forming environment, we can now study in detail the evolutionary process of a pair of planets embedded in an accretion disk. We analyze the capture process through full hydrodynamic simulations and forced N-body calculations.

10.4.1 Hydrodynamical Studies

Two embedded planets will generate a highly variable flow field in the disk, and two massive Jupiter-type planets will open a wide joint gap in the disk in which both planets will orbit [47, 48]. To demonstrate this effect, we show here the results of numerical, two-dimensional hydrodynamic simulations of two embedded planets in a disk, starting at 1 and 2 AU, respectively, on circular orbits. The masses of the two planets are 2.4 and 2.55 $M_{\rm Jup}$ orbiting a star of $1.08 M_{\odot}$, where the parameters have been chosen to roughly match the system HD 73526. To allow for a sizable evolution during the run-time of the simulations, we have used a viscosity with a relatively large value of $\alpha = 0.01$. In Figure 10.2(a), we present a plot of the two-dimensional density distribution, while in Figure 10.2(b) the azimuthally averaged density is shown at different times of the evolution. Between the planets, a joint deep and wide gap is created after about 200–300 orbital periods of the planets. In contrast to earlier simulations [49, 50], in this case, a more realistic persisting inner disk has been taken into account. In the presented simulation, the disk is allowed to accrete viscously onto the central star. Both planets orbit inside a wide joint gap, which is not cleared entirely due to the complex flow field that allows material from the inner/outer disk to enter and even cross the gap. However, even in this case of no accreting planets, the density inside the gap is at least an order of magnitude smaller than that of the ambient disk. As both planets are surrounded by an inner disk and an outer disk, the inner planet will feel a small positive torque (due to the inner disk) and the outer a negative torque (due to the outer disk). Upon migration, the outer planet is driven inward and the inner (very

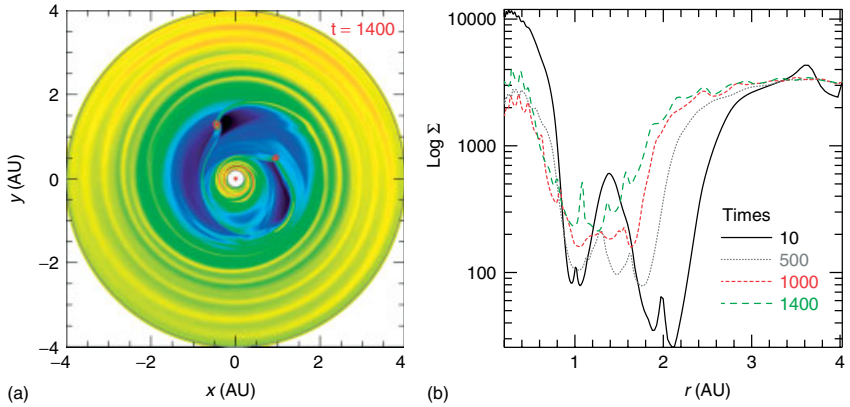


Figure 10.2 (a) Color plot of the density distribution of two embedded planets in a protoplanetary disk after 1400 years, starting with an initially axisymmetric configuration. (b) Azimuthally averaged density distribution (dimensionless units) at four different times (in years).

slowly) outward. We have clearly reached the situation where the relative migration of the two planets is convergent (Figure 10.3). The outward migration of the inner planet is very small because of the little mass contained in the inner disk.

When the outer planet approaches the 2 : 1 commensurability, its gravitational interaction periodically excites large perturbations in the orbit of the inner planet, which leads to an increase in eccentricity and eventually to capture in the 2 : 1 MMR. This is illustrated in Figure 10.3, where we show the change in semimajor axis and eccentricity of the two planets for the above-mentioned hydrodynamic simulation. The resonant capture at $t \approx 400$ years leads to an increase in the eccentricities of the two planets. For comparable planet masses, the eccentricity of the outer planet will be smaller than that of the inner planet. At the end of this simulation, the inner planet reaches $e = 0.32$ and the outer planet $e = 0.07$. The eccentricities stop increasing at $t = 1000$ years due to the increased damping effect of the inner disk. Hence, in this scenario, the final eccentricity of the outer planet (and subsequently of the inner) will be determined by the damping action of the ambient inner and outer disks. In the case of growing eccentricity, the planets will periodically enter the disk material, and the action of the disk will be to circularize the planet and damp the eccentricity.

In fact, the small eccentricities of the resonant planetary system GJ 876 with two planets deep in resonance seem to require an effective damping action of the disk; we discuss this point later. An equilibrium value for the eccentricity of the outer planet has been given by Snellgrove *et al.* [51], which depends, for given planetary masses, only on the ratio of the migration timescale $\tau_{\rm mig}$ to the circularization timescale $\tau_{\rm circ}$. A dispersing disk will stop the migration eventually and basically freeze the final configuration. Knowledge of the present configuration will give a rough estimate of the distance that the planets have been traveling through the disk.

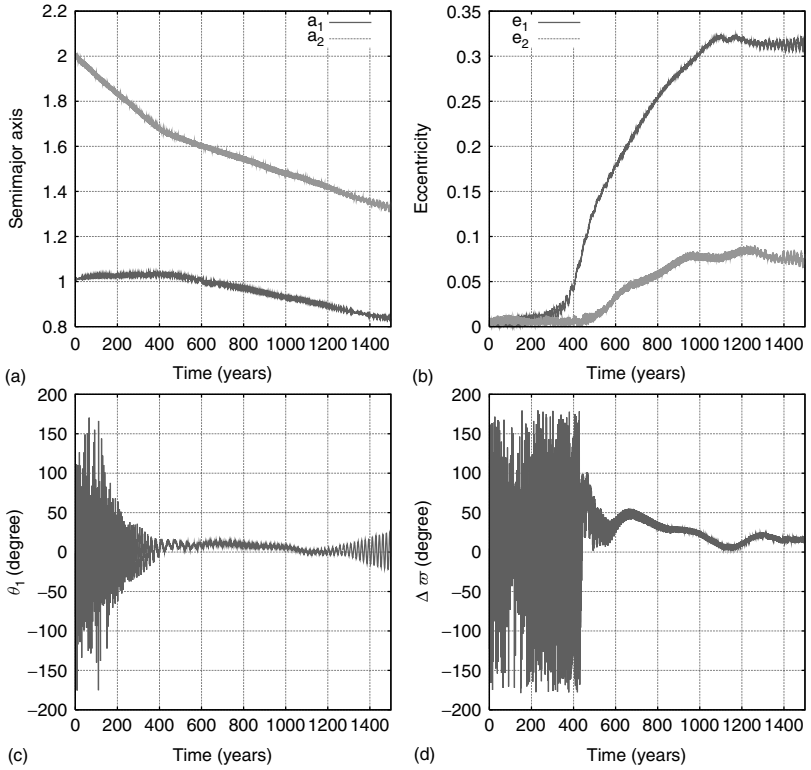


Figure 10.3 The semimajor axis (a) and eccentricity (b) evolution of two embedded planets in a protoplanetary disk. Owing to the presence of an inner disk, planet 1 slowly migrates outward. The resonant angles Θ_1 (c) and $\Delta\varpi$ (d). At $t \approx 400$ years, capture in a 2 : 1 resonance occurs, the eccentricities increase, the orbits align, and the planets migrate inward simultaneously.

If the inward migration proceeds over larger distances, the eccentricities can (for small disk damping) possibly grow to large values such that upon disk dissipation the remaining planetary system is unstable. Subsequent chaotic evolution may then create the observed high eccentricity systems [52].

Upon capture not only the eccentricities will increase but also the orbits will align to minimize dangerous close encounters. This is illustrated in Figure 10.3(c) and (d), where Θ_1 and $\Delta\varpi$ are plotted as a function of time. Already before capture, Θ_1 begins to librate around 0°, followed later by $\Delta\varpi$ (or Θ_2). In the initial phase of capture with only Θ_1 librating, there is some indication that Θ_2 librates transiently around 180°, which would imply an antisymmetric configuration [30, 53]. During the subsequent evolution, Θ_2 begins to librate around 0°.

One of the most important results from hydrodynamical evolutions with two embedded planets is the fact that for converging inward migration of two massive planets that capture each other in a 2 : 1 MMR, one always finds an alignment of the apsidal lines, that is, apsidal corotation (ACR), where both resonant angles librate

around zero [2, 51, 49]. As discussed later, this refers exactly to the observed state of GJ 873 and possibly to HD 82943. Another important implication of this scenario refers to the final masses of the two planets. As the outer planet is in contact with a quite massive ambient disk, it can still grow substantially in mass, whereas the inner planet has only a small-mass reservoir available. Hence, in this scenario, we would expect preferentially a more massive outer planet, which appears to be realized in three out of the four resonant 2 : 1 MMR systems.

10.4.2 Forced Migration

Using full multidimensional hydrodynamical simulations, the heavy computational effort does not allow for detailed parameter studies. To perform these studies, typically a simplified approach is used where a standard three-body program (a star and two planets) is extended to include additional (dissipative) forces that mimic the migration and eccentricity damping induced by the disk. As discussed above, the end configuration will be a situation where two planets are orbiting in a large gap surrounded by an inner and an outer disk. Early hydrodynamic simulations assumed that the inner disk can be neglected, which is approximately true if the mass of the inner disk is small or it has been accreted onto the central star. In this case, the disk torque is applied only to the outer planet. Such a forcing can be implemented directly as simple additional forces to the individual velocity components of the outer planets [2, 54] or by specifying the required damping rates $\tau_{\rm mig}$ and $\tau_{\rm ecc}$ of semimajor axis and eccentricity for the outer planet and transforming those into forces [19, 49]. This approach reduces the computational effort from weeks for one model to a few seconds. Typically the migration rate $\tau_{\rm mig}$ is specified and then $\tau_{\rm ecc}$ is given as a fixed fraction of the former as pioneered by Lee and Peale [19]:

$$\tau_{\rm mig} = \frac{a}{\dot{a}} \qquad \text{and} \qquad \tau_{\rm ecc} = \frac{e}{\dot{e}} = \frac{1}{K}\,\tau_{\rm mig} \tag{10.4}$$

that is, the eccentricity damping timescale is a factor K shorter than the migration time. For small-mass planets, where the interactions between the disk and the planets can be treated in the linear approximation, the value of K is roughly given by $(R/H)^2$, where R is the distance to the star and H the vertical thickness of the disk [55, 56].

In our case of high-mass planets, the assumed parameter K can be calibrated by a comparison of such forced N-body models to detailed hydrodynamical simulations. This has been attempted in [49], where it has been found that K has to be chosen to lie between 1 and 10 for a good agreement with the hydrodynamics and forced three-body calculations. However, this low value for K is problematic when N-body results are compared to observations of GJ 876, which is the best analyzed system. Applying a forced migration model using Equation 10.4 *only to the outer planet*, a value of $K \approx 100$ is required to keep the resonantly excited eccentricities sufficiently small to be in agreement with the observations of GJ 876 [19]. Possible solutions to the problem are mentioned in the following section.

10.5
Matching Observed Systems

We now consider the properties of real observed systems. We demonstrate the evidence for a plain adiabatic migration, capture process for GJ 876, and then show that in explaining further systems such as HD 128311 and HD 73526, one has to include additional short-lived perturbations to obtain a good match with the observations.

10.5.1
GJ 876: A Case of Adiabatic Migration

The first extrasolar system to be discovered that showed clear signs of an MMR was GJ 876 [17]. Here, two massive planets orbit a central star with orbits of $\approx$30 and 60 days. Later, an additional close-in third planet (d) has been discovered [57], which does not, however, influence the dynamics of the resonant pair appreciably. The apsidal lines (of planets b and c) are aligned and $\Delta\varpi$ and Θ_1 librate around zero with amplitudes of about 34° and only 7°, respectively [5]. In a detailed theoretical investigation, Lee and Peale [19] analyzed possible formation scenarios of the 2 : 1 resonant configuration in GJ 876. They performed a series of forced three-body simulations, where, following the previously described hydrodynamical scenario, only the outer planet experiences a migration due to the ambient disk and is slowly driven toward the inner planet. They used fixed constant values for the migration rate $\tau_{\rm mig}$ and the eccentricity damping factor K as defined in Equation (10.4) and found that the planets inevitably are captured into a 2 : 1 MMR in a symmetric configuration ($\Delta\varpi = 0$) and apsidal corotation. For sufficiently high values of the eccentricity damping, they find that the eccentricities level off to constant values that are given solely by the magnitude of K and not by the migration timescale $\tau_{\rm mig}$. Interestingly, the rather small observed values of $e_1 = 0.23$ and $e_2 = 0.05$ for the inner and outer planets can only be attained in equilibrium for a quite large value of $K \approx 100$, that is, through strong eccentricity damping of the ambient disk. In contrast, the aforementioned hydrodynamical simulations consistently yield much smaller values of K between 1 and 10 [49].

In a large parameter study using hydrodynamical as well as forced three-body simulations, Kley *et al.* [50] reinvestigated formation scenarios of GJ 876. To possibly increase the eccentricity damping and avoid the problem of too small K values, a range of high viscosities, higher temperatures, and fully nonisothermal radiative disks have been studied. However, none of these mechanisms is able to increase the eccentricity damping substantially; the system is always driven deep into the resonance with large eccentricities and apsidal corotation. This led to the conclusion that only limited radial migration can be allowed after the planets are captured into the 2 : 1 resonance. Assuming that the resonant capture happened at the final stages in the formation process of GJ 876, the inclusion of a disk dispersal occurring on the viscous timescale resulted in values for the eccentricities that are consistent with the observations [50].

This termination of migration via disk dissipation requires a sort of fine-tuning of parameters to match the observations and is not fully satisfactory. A possible self-consistent remedy has been suggested recently by Crida *et al.* [58]. Most of the previous simulations of planets embedded in disks assumed a simplified unrealistic inner boundary condition (at radius $r_{\min}$), which lead to a rapid emptying of the inner disk onto the star [49, 50]. However, as pointed out, for example, by Crida *et al.* [59] and in the hydrodynamic model discussed above, a proper boundary treatment (at $r_{\min}$) will prevent too rapid emptying. Taking the persistent inner disk into account, it could be shown that its presence always damps the eccentricity of the inner planet as seen in Figure 10.4. Fully self-consistent hydrodynamic models including this inner disk allow for an extended radial migration for the GJ 876 planets while matching the observations [58]. As seen in Figure 10.4, this eccentricity damping is accompanied for small eccentricities with an outward migration of the inner planet. Both of these effects can be taken into account again through simple forced N-body models. Using these, Crida *et al.* [58] have also analyzed all the additional known 2 : 1 systems and have shown that for reasonable K_{inner}, a good agreement with observations can be obtained.

10.5.2
Formation of Systems HD 128311 and HD 73526 through Mixed Scenarios

The radial velocity measurements suggest that the systems HD 128311 and HD 73526 contain pairs of giant planets engaged in a 2 : 1 protective MMR [7,8]. Numerical integrations, based on the published orbital data, indicated that the planets in both systems are in a 2 : 1 resonance without apsidal corotation. Moreover, the eccentricities of the giant planets show large-amplitude oscillations, which are regular in the case of HD 128311 (see Figure 10.1 and irregular (chaotic) in the case of HD 73526 [60]). In both the cases, the resonant angle Θ_1 librates around 0° with amplitudes of $\sim 60^\circ$ (HD 128311) and $\sim 90^\circ$ (HD 73526), respectively. The

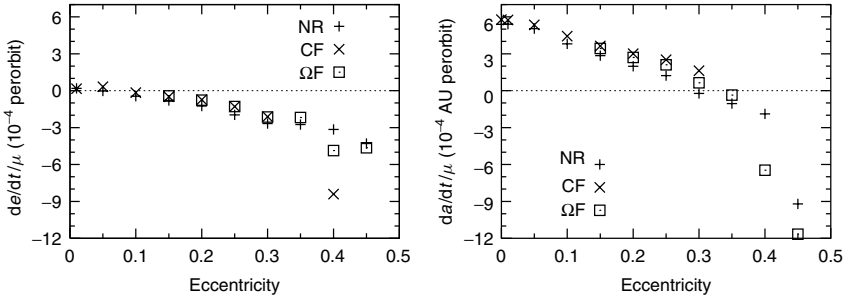


Figure 10.4 Influence of an inner disk on the semimajor axis and eccentricity of an outer planet as a function of planetary eccentricity. The labels refer to simulations performed in different coordinate frames. NF: nonrotating; CF: rotating with a constant speed; and ΩF: rotating with the actual angular velocity of the planet. Adapted from [58].

lack of corotation in these system means that both Θ_2 and $\Delta\varpi$ circulate. Stability investigations performed by means of the relative Lyapunov indicator [60] show that the planets in HD 73526 are located in a weakly chaotic region of the phase space [25]. Through a small variation in the orbital parameter, also for HD 73526 regular orbital solutions can be found with strong resemblance to HD 128311. The only difference between the dynamics of the systems is that in the majority of the orbital fits the giant planets of HD 73526 are in apsidal corotation, however with enlarged amplitudes in Θ_2 and $\Delta\varpi$ [25].

As discussed above, the standard scenario for forming resonant systems is the adiabatic migration scenario, which typically leads to systems in apsidal corotation with small eccentricity variations. Clearly, the orbital behavior of the giant planets around HD 128311 and HD 73526 violate this result, since (i) in the case of HD 128311, the orbits of the giant planets are not in apsidal corotation and (ii) in both systems, the eccentricities exhibit considerable variations. However, the presence of resonance indicates that these systems may have gone through an adiabatic migration process in the past similar to that of GJ 876.

To resolve this discrepancy, *mixed* evolutionary scenarios of migrating planetary systems have been suggested. These combine differential migration with subsequent capture into resonance with other additional *perturbative effects* such as sudden termination of migration or a planet–planet scattering event [24].

10.5.2.1 Planet–Planet Scattering

The behavior of the eccentricities of the giant planets of the systems HD 128311 and HD 73526 is very similarto that observed in the eccenticities of the system around υ Andromedae. Ford *et al.* [61] suggested that this behavior in υ is most likely the result of a planet–planet scattering event. The present status of HD 128311 and HD 73526 can indeed be reached by such an effect, that is, one of the giant planets suffers a close encounter with an already existing smaller mass planet ($m_p = 10M_\oplus$) [24, 25]. The small-mass planet can reside inside (at smaller radii) or outside of the resonant system. In both the cases, scattering of the single planet with the system can lead to breaking the resonance, but an inner planet is more effective in this regard. The small-mass planet attains a chaotic orbit and is often ejected later from the system [25]. A typical example of the eccentricity and $\Delta\varpi$ evolution of the resonant system is shown in Figure 10.5. After resonant capture, the system settles into apsidal corotation with small variations of orbital elements (as described above), and at $t \approx 2600$ the interaction with the small planet yields a modification of the resonance.

10.5.2.2 Sudden Stop of Migration

Another mechanism, which may induce large-amplitude variations of the eccentricities and the resonant angles, is the sudden termination of the migration. Recent *Spitzer* observations of young stars show that the inner part of some protoplanetary disks may practically be emptied, possibly due to photoevaporation induced by the central star [62, 63] or the emission of jets from the inner region of the accretion disk. The inwardly migrating resonant system will eventually reach such an inner

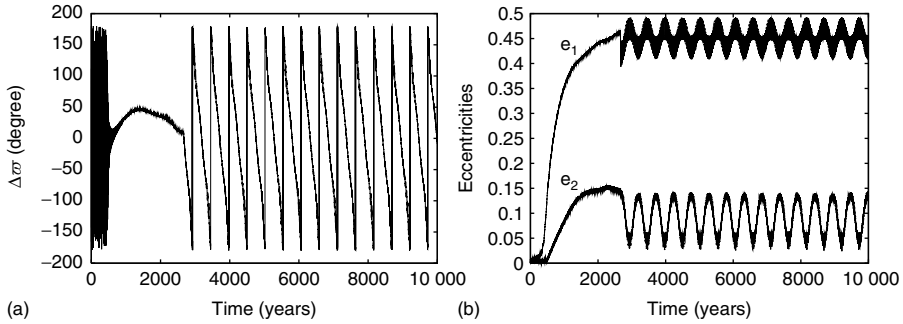


Figure 10.5 The breaking of the apsidal corotation due to a scattering event induced by a small-mass inner planet in the system of HD 128311 (after [24]). The evolution of $\Delta\varpi(t)$ (a) and the eccentricities (b) are shown.

rim of the disk and be halted there. If the stopping time is sufficiently short, orbital variation similar to the above scattering event can be obtained [24, 25]. Pure scattering [29] or stochastic forcing [64] could also explain these systems.

10.5.3
Capture in 3 : 2 Resonance

So far we have dealt primarily with capture in the 2 : 1 resonance as this is by far the strongest resonance and observationally best confirmed. However, there is recent evidence for an extrasolar system engaged in a 3 : 2 resonance [9], not to mention the resonance between Neptune and Pluto in the solar system. The latter did form through outward migration of Neptune, while we believe that many of the extrasolar systems may have formed via convergent inward migration.

One of the first works dealing theoretically with capture in 3 : 2 has been by Masset and Snellgrove [65], who found very interesting dynamical behavior for a planetary system containing two planets, an inner Jupiter mass object and an outer Saturn mass object embedded in a protoplanetary disk. In this case, the outer low-mass planet (Saturn) undergoes a very rapid type III migration [66], crosses the location of the 2 : 1 resonance, and is eventually captured into a 3 : 2 resonance by Jupiter. However, instead of continuing the inward migration, the now coupled system experiences an outward migration. This surprising result is possible because the low-mass outer planet allows for mass to cross its orbit from outside, which is subsequently passed by the inner Jupiter even further toward the star. In turn, Jupiter gains angular momentum and begins to move outward, always maintaining its resonance with Saturn. This result has been confirmed with more simulations by Morbidelli and Crida [67]. Recently, the capture process of a Jupiter mass planet with an outer smaller mass planet has been studied through a more comprehensive parameter study by Pierens and Nelson [68]. The authors confirm that a Saturn mass object is robustly captured into 3 : 2 with an inner Jupiter. This is particularly important with respect to the early solar system history where it is believed that

Saturn may have originally resided inside the 2 : 1 resonance with Jupiter. Its outward crossing might then have led to the late heavy bombardment [69].

10.5.4
Destruction of Resonances

The hydrodynamic studies discussed in the previous section rely on a laminar (viscous) hydrodynamic modeling. However, accretion disks are driven by turbulence, which may possibly influence the dynamics of embedded planets, leading to a new type of ''stochastic'' migration [70]. Question arises about the influence of stochastic fluctuations on the stability of the (forming) resonant configurations. This was first addressed by Adams *et al*. [71], who used a pendulum model for the resonance and described turbulence through a stochastic forcing. Their analysis suggests that systems in MMR should be rare, as their survival rate is very low. This work has been recently extended by a more detailed analysis taking into account possible damping effects due to the disk [72], but the conclusions remain the same. Because of the small number statistics and uncertainty concerning the observed fraction of planetary systems that actually reside in resonance, it is very difficult to quantify ''rarity''. The evolution of resonant systems under stochastic forcing has been looked at recently by Rein and Papaloizou [64]. To more accurately estimate the effect of the turbulence on the resonance, dedicated multidimensional (magneto-)hydrodynamic simulations will be required.

The observed number of resonant configurations may also be a consequence of destruction at later times of the planet formation process. Systems that have experienced a larger range of radial (inward) migration can have relatively high eccentricities. Upon disk dissipation, its stabilizing damping effect is reduced, which can easily lead to a direct destabilization of the orbits, followed possibly by a chaotic evolution with the ejection of one of the planets. In fact, the occurrence of the observed large eccentricities (also in single planet extrasolar systems) has been attributed to this process [52, 73].

10.6
Summary

In this chapter, we have studied possible formation scenarios of the observed MMRs in extrasolar planetary systems. The mere existence of resonant configurations allows us to infer constraints on the physical environment at the time of their formation. As the majority of resonant systems are locked in a 2 : 1 resonance, we concentrate, in particular, on these systems.

We have shown that planet–disk interaction between young embedded protoplanets and the protoplanetary disk will typically lead to a differential migration where the outer planet approaches the inner one. To analyze the outcome of the convergent migration, we first presented results of full hydrodynamical simulations with two embedded planets and then compared these to forced three-body

simulations where additional forces mimicking the planet–disk interaction have been added. We find that resonant capture into the 2 : 1 resonance indeed occurs frequently and that continued driving of the disk will move the system deep into the resonance with significant growth of the eccentricity of the two planets. Additionally, the final configuration of the orbits is symmetric where the periapses are aligned with only small libration. We refer to this state as *apsidal corotation*.

The well-observed system GJ 876 is exactly in this symmetric configuration with only small libration of the two resonant angles. It constitutes the clearest evidence that convergent migration of planets in protoplanetary disks has indeed occurred. The small eccentricities in this system can be explained if the disk acts as a damping agent to the eccentricities of the planets. In cases when the planets orbit within an inner cavity of the disk, the relatively small planetary eccentricities can only be explained if one assumes that after resonant capture occurred the duration of migration was limited by a rapid disk dispersal. The more realistic inclusion of an inner disk allows for large inward migration while maintaining the small eccentricities of resonant planets.

In the other two systems considered (HD 128311 and HD 73526), the eccentricities in the current observed state vary significantly and vanish periodically for one planet. We demonstrate that the dynamical structure of these two systems may be the result of a mixed evolutionary scenario melding together inward migration and a sudden perturbation, induced, for example, by an additional small planet in the system or a sudden termination of the migration process.

In summary, the complex dynamical behavior of the observed resonant planetary systems can be understood in the framework of their formation process. Interaction with the protoplanetary disk and possibly more sudden planet–planet scattering events are important ingredients in shaping planetary systems.

Acknowledgments

This research has been supported in part by the Deutsche Forschungsgemeinschaft (DFG) through grant DFG Forschergruppe 759 *The Formation of Planets: The Critical First Growth Phase* and DFG-grant KL 650/7.

References

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

2 Nelson, R.P. and Papaloizou, J.C.B. (**2002**) Possible commensurabilities among pairs of extrasolar planets. *Monthly Notices of the Royal Astronomical Society*, **333**, L26–L30.

3 Udry, S., Fischer, D., and Queloz, D. (**2007**) A decade of radial-velocity discoveries in the exoplanet domain, in *Protostars and Planets V* (eds B. Reipurth, D. Jewitt, and K. Keil), University of Arizona Press, Tucson, pp. 685–699.

4 Wright, J.T., Upadhyay, S., Marcy, G.W. *et al.* (**2009**) Ten new and updated multiplanet systems and a survey of exoplanetary systems. *The Astrophysical Journal*, **693**, 1084–1099.

5 Laughlin, G., Butler, R.P., Fischer, D.A. *et al.* (**2005**) The GJ 876 planetary system: a progress report. *The Astrophysical Journal*, **622**, 1182–1190.

6 Beaugé, C., Giuppone, C.A., Ferraz-Mello, S., and Michtchenko, T.A. (**2008**) Reliability of orbital fits for resonant extrasolar planetary systems: the case of HD82943. *Monthly Notices of the Royal Astronomical Society*, **385**, 2151–2160.

7 Vogt, S.S., Butler, R.P., Marcy, G.W. *et al.* (**2005**) Five new multicomponent planetary systems. *The Astrophysical Journal*, **632**, 638–658.

8 Tinney, C.G., Butler, R.P., Marcy, G.W. *et al.* (**2006**) The 2:1 resonant exoplanetary system orbiting HD 73526. *The Astrophysical Journal*, **647**, 594–599.

9 Correia, A.C.M., Udry, S., Mayor, M. *et al.* (**2009**) The HARPS search for southern extra-solar planets. XVI. HD 45364, a pair of planets in a 3:2 mean motion resonance. *Astronomy and Astrophysics*, **496**, 521–526.

10 Pepe, F., Correia, A.C.M., Mayor, M. *et al.* (**2007**) The HARPS search for southern extra-solar planets. VIII. μ Arae, a system with four planets. *Astronomy and Astrophysics*, **462**, 769–776.

11 McArthur, B.E., Endl, M., Cochran, W.D. *et al.* (**2004**) Detection of a neptune-mass planet in the ρ^1 cancri system using the hobby-eberly telescope. *The Astrophysical Journal Letters*, **614**, L81–L84.

12 Desort, M., Lagrange, A.-M., Galland, F. *et al.* (**2008**) Extrasolar planets and brown dwarfs around A-F type stars. V. A planetary system found with HARPS around the F6IV-V star HD 60532. *Astronomy and Astrophysics*, **491**, 883–888.

13 Laskar, J. and Correia, A.C.M. (**2009**) HD 60532, a planetary system in a 3:1 mean motion resonance. *Astronomy and Astrophysics*, **496**, L5–L8.

14 Butler, R.P., Wright, J.T., Marcy, G.W. *et al.* (**2006**) Catalog of nearby exoplanets. *The Astrophysical Journal*, **646**, 505–522.

15 Johnson, J.A., Butler, R.P., Marcy, G.W. *et al.* (**2007**) A new planet around an m dwarf: revealing a correlation between exoplanets and stellar mass. *The Astrophysical Journal*, **670**, 833–840.

16 Correia, A.C.M., Udry, S., Mayor, M. *et al.* (**2005**) The CORALIE survey for southern extra-solar planets. XIII. A pair of planets around HD 202206 or a circumbinary planet? *Astronomy and Astrophysics*, **440**, 751–758.

17 Marcy, G.W., Butler, R.P., Fischer, D. *et al.* (**2001**) A pair of resonant planets orbiting GJ 876. *The Astrophysical Journal*, **556**, 296–301.

18 Laughlin, G. and Chambers, J.E. (**2001**) Short-term dynamical interactions among extrasolar planets. *The Astrophysical Journal Letters*, **551**, L109–L113.

19 Lee, M.H. and Peale, S.J. (**2002**) Dynamics and origin of the 2:1 Orbital resonances of the GJ 876 planets. *The Astrophysical Journal*, **567**, 596–609.

20 Mayor, M., Udry, S., Naef, D. *et al.* (**2004**) The CORALIE survey for southern extra-solar planets. XII. Orbital solutions for 16 extra-solar planets discovered with CORALIE. *Astronomy and Astrophysics*, **415**, 391–402.

21 Barnes, R. and Quinn, T. (**2004**) The (In)stability of planetary systems. *The Astrophysical Journal*, **611**, 494–516.

22 Ferraz-Mello, S., Michtchenko, T.A., and Beaugé, C. (**2005**) The orbits of the extrasolar planets HD 82943c and b. *The Astrophysical Journal*, **621**, 473–481.

23 Lee, M.H., Butler, R.P., Fischer, D.A. *et al.* (**2006**) On the 2:1 orbital resonance in the HD 82943 planetary system. *The Astrophysical Journal*, **641**, 1178–1187.

24 Sandor, Z. and Kley, W. (**2006**) On the evolution of the resonant planetary system HD 128311. *Astronomy and Astrophysics*, **451**, L31–L34.

25 Sandor, Z., Kley, W., and Klagyivik, P. (**2007**) Stability and formation of the resonant system HD 73526. *Astronomy and Astrophysics*, **472**, 981–992.

26 Fischer, D.A., Marcy, G.W., Butler, R.P. *et al.* (**2008**) Five planets orbiting 55 cancri. *The Astrophysical Journal*, **675**, 790–801.

27 Naef, D., Mayor, M., Beuzit, J.L. *et al.* (**2004**) The ELODIE survey for northern extra-solar planets. III. Three planetary candidates detected with ELODIE. *Astronomy and Astrophysics*, **414**, 351–359.

28 Psychoyos, D. and Hadjidemetriou, J.D. (**2005**) Dynamics of extrasolar systems at the 5/2 resonance: application to 47 UMa, in *IAU Colloq. 197: Dynamics of Populations of Planetary Systems* (eds Z. Knezevic and A. Milani), Cambridge University Press, pp. 55–62.

29 Raymond, S.N., Barnes, R., Armitage, P.J., and Gorelick, N. (**2008**) Mean motion resonances from planet-planet scattering. *The Astrophysical Journal Letters*, **687**, L107–L110.

30 Michtchenko, T.A., Beaugé, C., and Ferraz-Mello, S. (**2008**) Dynamic portrait of the planetary 2/1 mean-motion resonance - I. Systems with a more massive outer planet. *Monthly Notices of the Royal Astronomical Society*, **387**, 747–758.

31 Goldreich, P. and Tremaine, S. (**1980**) Disk-satellite interactions. *The Astrophysical Journal*, **241**, 425–441.

32 Lin, D.N.C. and Papaloizou, J. (**1986b**) On the tidal interaction between protoplanets and the protoplanetary disk. III - Orbital migration of protoplanets. *The Astrophysical Journal*, **309**, 846–857.

33 Ward, W.R. (**1997**) Protoplanet migration by nebula tides. *Icarus*, **126**, 261–281.

34 Murray, N., Hansen, B., Holman, M., and Tremaine, S. (**1998**) Migrating planets. *Science*, **279**, 69.

35 Ogilvie, G.I. and Lubow, S.H. (**2002**) On the wake generated by a planet in a disc. *Monthly Notices of the Royal Astronomical Society*, **330**, 950–954.

36 Kley, W. (**1998**) On the treatment of the Coriolis force in computational astrophysics. *Astronomy and Astrophysics*, **338**, L37–L41.

37 Kley, W. (**1999**) Mass flow and accretion through gaps in accretion discs. *Monthly Notices of the Royal Astronomical Society*, **303**, 696–710.

38 Bryden, G., Chen, X., Lin, D.N.C. *et al.* (**1999**) Tidally induced gap formation in protostellar disks: gap clearing and suppression of protoplanetary growth. *The Astrophysical Journal*, **514**, 344–367.

39 de Val-Borro, M., Edgar, R.G., Artymowicz, P. *et al.* (**2006**) A comparative study of disc-planet interaction. *Monthly Notices of the Royal Astronomical Society*, **370**, 529–558.

40 Lin, D.N.C. and Papaloizou, J. (**1986a**) On the tidal interaction between protoplanets and the primordial solar nebula. II - Self-consistent nonlinear interaction. *The Astrophysical Journal*, **307**, 395–409.

41 Crida, A., Morbidelli, A., and Masset, F. (**2006**) On the width and shape of gaps in protoplanetary disks. *Icarus*, **181**, 587–604.

42 Nelson, R.P., Papaloizou, J.C.B., Masset, F., and Kley, W. (**2000**) The migration and growth of protoplanets in protostellar discs. *Monthly Notices of the Royal Astronomical Society*, **318**, 18–36.

43 Papaloizou, J.C.B., Nelson, R.P., Kley, W. *et al.* (**2007**) Disk-planet interactions during planet formation, in *Protostars and Planets V* (eds B. Reipurth, D. Jewitt, and K. Keil), University of Arizona Press, Tucson, pp. 655–668.

44 Baruteau, C. and Masset, F. (**2008**) On the corotation torque in a radiatively inefficient disk. *The Astrophysical Journal*, **672**, 1054–1067.

45 Paardekooper, S.-J. and Papaloizou, J.C.B. (**2008**) On disc protoplanet interactions in a non-barotropic disc with thermal diffusion. *Astronomy and Astrophysics*, **485**, 877–895.

46 Kley, W. and Crida, A. (**2008**) Migration of protoplanets in radiative discs. *Astronomy and Astrophysics*, **487**, L9–L12.

47 Kley, W. (**2000**) On the migration of a system of protoplanets. *Monthly Notices of the Royal Astronomical Society*, **313**, L47–L51.

48 Bryden, G., Różyczka, M., Lin, D.N.C., and Bodenheimer, P. (**2000**) On the interaction between protoplanets and protostellar disks. *The Astrophysical Journal*, **540**, 1091–1101.

49 Kley, W., Peitz, J. and Bryden, G. (**2004**) Evolution of planetary systems in resonance. *Astronomy and Astrophysics*, **414**, 735–747.

50 Kley, W., Lee, M.H., Murray, N., and Peale, S.J. (**2005**) Modeling the resonant planetary system GJ 876. *Astronomy and Astrophysics*, **437**, 727–742.

51 Snellgrove, M.D., Papaloizou, J.C.B., and Nelson, R.P. (**2001**) On disc driven inward migration of resonantly coupled planets with application to the system around GJ876. *Astronomy and Astrophysics*, **374**, 1092–1099.

52 Ford, E.B., Havlickova, M., and Rasio, F.A. (**2001**) Dynamical instabilities in extrasolar planetary systems containing two giant planets. *Icarus*, **150**, 303–313.

53 Beaugé, C., Ferraz-Mello, S., and Michtchenko, T.A. (**2003**) Extrasolar planets in mean-motion resonance: apses alignment and asymmetric stationary solutions. *The Astrophysical Journal*, **593**, 1124–1133.

54 Beaugé, C., Michtchenko, T.A., and Ferraz-Mello, S. (**2006**) Planetary migration and extrasolar planets in the 2/1 mean-motion resonance. *Monthly Notices of the Royal Astronomical Society*, **365**, 1160–1170.

55 Tanaka, H. and Ward, W.R. (**2004**) Three-dimensional interaction between a planet and an isothermal gaseous disk. II. Eccentricity waves and bending waves. *The Astrophysical Journal*, **602**, 388–395.

56 Cresswell, P., Dirksen, G., Kley, W., and Nelson, R.P. (**2007**) On the evolution of eccentric and inclined protoplanets embedded in protoplanetary disks. *Astronomy and Astrophysics*, **473**, 329–342.

57 Rivera, E.J., Lissauer, J.J., Butler, R.P. *et al.* (**2005**) A 7.5 M_{earth} planet orbiting the nearby star, GJ 876. *The Astrophysical Journal*, **634**, 625–640.

58 Crida, A., Sandor , Z., and Kley, W. (**2008**) Influence of an inner disc on the orbital evolution of massive planets migrating in resonance. *Astronomy and Astrophysics*, **483**, 325–337.

59 Crida, A., Morbidelli, A., and Masset, F. (**2007**) Simulating planet migration in globally evolving disks. *Astronomy and Astrophysics*, **461**, 1173–1183.

60 Sandor, Z., Erdi, B., Szell, A., and Funk, B. (**2004**) The relative lyapunov indicator: an efficient method of chaos detection. *Celestial Mechanics and Dynamical Astronomy*, **90**, 127–138.

61 Ford, E.B., Lystad, V., and Rasio, F.A. (**2005**) Planet-planet scattering in the upsilon Andromedae system. *Nature*, **434**, 873–876.

62 D'Alessio, P., Hartmann, L., Calvet, N. *et al.* (**2005**) The truncated disk of coku Tau/4. *The Astrophysical Journal*, **621**, 461–472.

63 Calvet, N., D'Alessio, P., Watson, D.M. *et al.* (**2005**) Disks in Transition in the taurus population: spitzer irs spectra of GM aurigae and DM tauri. *The Astrophysical Journal Letters*, **630**, L185–L188.

64 Rein, H. and Papaloizou, J.C.B. (**2009**) On the evolution of mean motion resonances through stochastic forcing: fast and slow libration modes and the origin of HD 128311. *Astronomy and Astrophysics*, **497**, 595–609.

65 Masset, F. and Snellgrove, M. (**2001**) Reversing type II migration: resonance trapping of a lighter giant protoplanet. *Monthly Notices of the Royal Astronomical Society*, **320**, L55–L59.

66 Masset, F.S. and Papaloizou, J.C.B. (**2003**) Runaway migration and the formation of hot jupiters. *The Astrophysical Journal*, **588**, 494–508.

67 Morbidelli, A. and Crida, A. (**2007**) The dynamics of Jupiter and Saturn in the gaseous protoplanetary disk. *Icarus*, **191**, 158–171.

68 Pierens, A. and Nelson, R.P. (**2008**) Constraints on resonant-trapping for two planets embedded in a protoplanetary disc. *Astronomy and Astrophysics*, **482**, 333–340.

69 Gomes, R., Levison, H.F., Tsiganis, K., and Morbidelli, A. (**2005**) Origin of the cataclysmic Late Heavy Bombardment period of the terrestrial planets. *Nature*, **435**, 466–469.

70 Nelson, R.P. and Papaloizou, J.C.B. (**2004**) The interaction of giant planets with a disc with MHD turbulence – IV.

Migration rates of embedded protoplanets. *Monthly Notices of the Royal Astronomical Society*, **350**, 849–864.

71 Adams, F.C., Laughlin, G., and Bloch, A.M. (**2008**) Turbulence implies that mean motion resonances are rare. *The Astrophysical Journal*, **683**, 1117–1128.

72 Lecoanet, D., Adams, F.C., and Bloch, A.M. (**2009**) Mean motion resonances in extrasolar planetary systems with turbulence, interactions, and damping. *The Astrophysical Journal*, **692**, 659–676.

73 Moorhead, A.V. and Adams, F.C. (**2005**) Giant planet migration through the action of disk torques and planet-planet scattering. *Icarus*, **178**, 517–539.

Further Reading

Rivera, E.J. and Lissauer, J.J. (**2000**) Stability analysis of the planetary system orbiting υ Andromedae. *The Astrophysical Journal*, **530**, 454–463.

11
Planet–Planet Gravitational Scattering

F. Marzari

11.1
Introduction

The present population of extrasolar planets have orbits with a wide range of semimajor axes and eccentricities. While the variety of semimajor axes can be explained by different migration mechanisms, the large orbital eccentricities are still matter of debate. In Figure 1.2 – from Chapter 1, the distribution of eccentricity versus period (and, hence, semimajor axis) is shown for all extrasolar planets discovered to date. Most of them are on highly elliptical orbits, much larger than that of any solar system planet. On the basis of the present understanding of planet formation via core-accretion in a protostellar disk ([1], see also Chapter 5), planets are expected to be on nearly circular orbits at the end of their accretion process. A natural guess is that the observed eccentricities are the outcome of a subsequent evolutionary process.

What is the dynamical route that can pump up planetary eccentricities as large as ∼0.93 (HD 80606 b, HD 4113 b)? One of the earliest mechanisms proposed for explaining these large eccentricities is planet–planet scattering [2–4]. In its classical formulation, planet–planet scattering predicts that planetary systems with two or more planets may become unstable because of the mutual gravitational perturbations between the planetary bodies on a timescale that depends on their initial orbital separations and masses. Instability leads to eccentricity growth and orbital crossing that eventually starts a chaotic phase, dominated by close encounters between the planets. This dynamically violent phase ends only when one or more planets are ejected from the system on hyperbolic trajectories. The initial orbital structure of the multiplanet system is dramatically altered and while the escaping planet will contribute to the population of giant free-floating planets, the surviving planets are left on highly eccentric noninteracting orbits possibly inclined respect to the initial orbital plane.

This mechanism may account also for some inward orbital migration of one (or more) of the planets, which releases orbital energy to the ejected ones. Subsequent tidal interaction with the host star could circularize the orbit of the remaining planet and possibly explain the close circular orbits of some of the ''Hot Jupiters''.

Formation and Evolution of Exoplanets. Edited by Rory Barnes

ISBN: 978-3-527-40896-2

In addition, planet–planet scattering may also account for the production of pairs of planets in mean motion resonances [5]. After one planet is ejected, the following orbital rearrangement may insert a pair of planets in a resonant configuration like in GJ 876 or other systems (see Table I in [5]).

11.2
Onset of Instability in Multiplanet Systems: the Gas-Free Scenario

Dynamical instability in a multiplanet system is a basic requirement for the onset of strong gravitational interactions between the planets leading ultimately to close encounters and chaos. Instability can be promoted in two different scenarios of planetary evolution depending on *whether or not the gas component of the protostellar disk is present*.

The classic scenario for planet–planet scattering starts from the assumption that a closely packed system of giant planets emerges from the protostellar disk when the gas dissipates. The conventional core-accretion scenario for planet formation ([6], Chapter 5) predicts that the coagulation of small icy–rocky dust particles in the outer part of the disk leads to the formation of planetesimals either by collisional coagulation [7] or by instability of the dusty layer in the midplane of the disk [8]. Subsequently, planetesimals grow by pairwise accumulation into solid cores with masses up to 10 Earth masses. These cores have a gravitational pull strong enough to capture gas from the surrounding disk reaching masses comparable to that of Jupiter on a timescale of the order of a few megayears, thanks also to inward migration due to tidal interaction with the disk [1]. Crowded systems populated by three or more giant planets can be formed if the circumstellar disk were massive enough. Near the end of the process, the gas is dispersed by a combination of mechanisms including strong stellar winds associated to the T-Tauri phase of the star and UV photoevaporation induced by the central or nearby stars. Once in a gas-free environment, the planets perturb each other more strongly, either because they just reached their final masses by rapid gas infall, which is a very fast process lasting only a few thousand years, or because the damping effect of the gaseous disk prevented, prior to its dissipation, the increase of eccentricity. The mutual gravitational perturbations among the planets in empty space may cause instability and hence chaotic behavior driven by close encounters between the planets.

11.2.1
The Stability Limit

A key concept for understanding the onset of chaos in a multiplanet system is that of dynamical stability. If the initial spacing between the planets is wide enough, the system can be stable over the solar system age (and beyond) and planet–planet scattering will not occur at any time. However, if the initial separations are below a certain threshold, let us call it the *stability limit*, the mutual gravitational perturbations cause a transition from quasi-periodic evolution to chaotic behavior

and subsequent rearrangement of the planetary system. This transition is marked by a sudden growth of the orbital eccentricities and strongly depends on how far the system is initially from the stability limit. Multiplanet orbital configurations farther from the stability limit will survive longer before becoming unstable. This means that when we look at a packed young planetary system, it may be on the verge of a strong chaotic phase.

Analytical formulations of the stability limit for two planets can be found in the literature and they estimate the minimum orbital separation Δ_c preventing the planets from undergoing close encounters for all time. This is the notion of *Hill stability*. If two planets are formed from an accretion disk with semimajor axis a_1 and a_2 whose difference $a_2 - a_1$ is larger than Δ_c, they will never undergo planet–planet scattering. Within the elliptic three-body problem in barycentric coordinates, Gladman [9] derived an analytical estimate for Δ_c given in implicit form:

$$\alpha^{-3}\left(\mu_1 + \frac{\mu_2}{\left(1+\frac{\Delta_c}{a_1}\right)^2}\right)(\mu_1\gamma_1 + \mu_2\gamma_2\left(1+\left(\frac{\Delta_c}{a_1}\right)^{1/2}\right)^2 > 1 + 3^{4/3}\frac{\mu_1\mu_2}{\alpha^{4/3}} \tag{11.1}$$

where $\mu_i = m_i/M, \alpha = \mu_1 + \mu_2, \gamma_i = (1-e_i^2)^{1/2}$, m_i is the mass of planet i, M is the mass of the star, e_i is the eccentricity of planet i. In the case of initially circular (and coplanar) orbits, as expected for planets emerging from a stationary protostellar disk, the above equation can be solved and, to the lowest order in mass, it gives

$$\Delta_c \sim 2.40(\mu_1 + \mu_2)^{1/3} a_1 \tag{11.2}$$

A slightly different expression has been calculated by Chambers *et al.* [10] via direct numerical integration of two-planet systems initially on circular orbits. The value of Δ_c computed over many simulations is

$$\Delta_c \sim 2\sqrt{3} R_H \tag{11.3}$$

where R_H is the planets' mutual Hill radius given by

$$R_H = \left(\frac{m_1 + m_2}{3M_\odot}\right)^{1/3}\left(\frac{(a_1 + a_2)}{2}\right) \tag{11.4}$$

It should be noted that the numerical constant is multiplied by a_1 in Eq. (11.2) and by R_H in Eq. (11.3). The difference is small for a system of two planets and it is related to the different approaches that lead to the two equations.

No general stability criterion is known for a planetary system with more than two planets, so numerical N-body experiments are used to study the evolution of such systems. As a first approximation, one could compute a value of $\Delta_c^{i,j}$ for each adjacent pair of planets in a ladder analogy.

In this context, Eq. (11.3) appears more appealing since it allows a computation of Δ_c for each pair. Chambers *et al.* [10] carried out numerical integrations of three and more small mass ($m_p = 10^{-5}, 10^{-7}$, and $10^{-9} M_\odot$) planets using, as initial

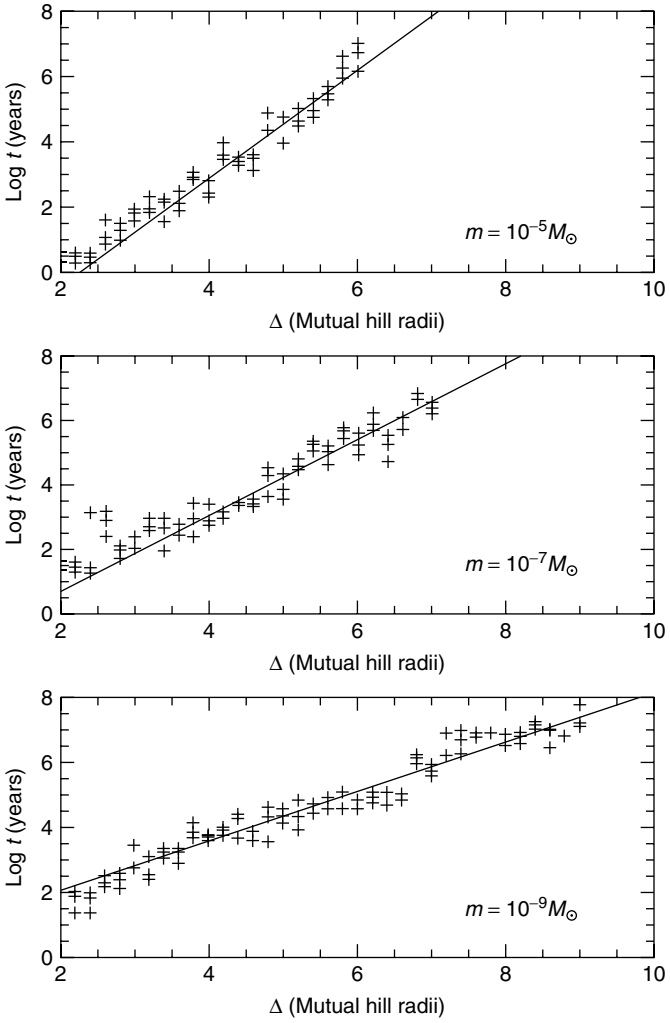


Figure 11.1 Logarithm of the timescale t required by a system of three equal mass planets with $m_p = 10^{-5}, 10^{-7}$, and $10^{-9} M_\odot$, started on initially circular orbits, to become unstable and undergo frequent close encounters. (Adapted from [10]).

spacing between the orbits, a fixed multiple of R_H (see Figure 11.1). They find an exponential relationship (linear in log t) between the time span for the onset of instability and $\Delta = KR_H$ of the form

$$\log t = b\Delta + c \tag{11.5}$$

where b and c are constants that depend on the planet masses and on their number. However, both b and c relax to a fixed value (see Table I in [10]) when the number

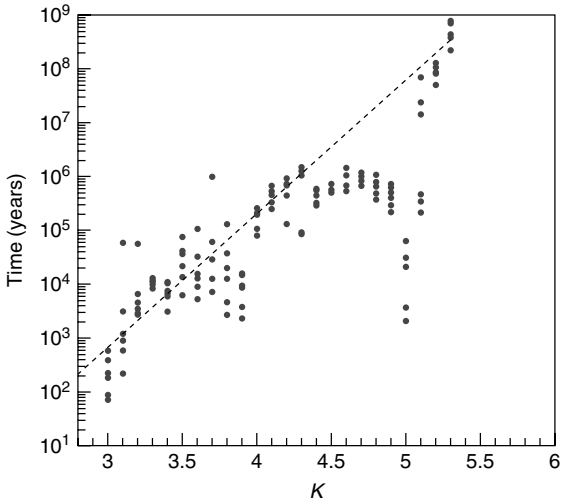


Figure 11.2 Logarithm of the instability growth timescale as a function of the initial separation, expressed as K times the mutual Hill radius, for three equal mass planets ($m = 10^{-3} M_\odot$). The significant drop of t for systems with $K \sim 5$ is due to the chaotic region associated with a 2 : 1 mean motion resonance between either planets 1 and 2, or planets 2 and 3.

of planets is larger than five. It is also manifest from Figure 11.1 that a Δ_c value for the system depends on m_p and ranges from about 7 for $m_p = 10^{-5} M_\odot$ to 10 for $m_p = 10^{-9} M_\odot$. Marzari and Weidenschilling [11] considered a system of three equal mass Jovian planets ($m_p \sim 10^{-3}$, 100 times larger than those considered in [10]) and they found that the dependence of $\log t$ on Δ sensibly deviates from a linear trend. The behavior predicted by Eq. (11.5) is, in fact, strongly perturbed by mean motion resonances between the planets whose influence depends on the planet masses. In Figure 11.2, the instability timescale shows a noticeable reduction of the time the three planets may survive without evolving chaotically for $K \sim 5$. This is due to the 2 : 1 resonance either between planets 1 and 2, or planets 2 and 3.

All these dynamical simulations confirm that a system of planets, from terrestrial to gaseous giants, may be stable for some time after the dissipation of the gaseous component of the protostellar disk before the onset of instabilities that lead to chaotic behavior and planet–planet scattering. The time span before instability depends on the initial orbital spacing and physical structure (masses) of the system.

11.2.2
Planet–Planet Scattering Dynamics

In this section, we explore in detail the dynamics of strong planet–planet interactions. Let us concentrate on systems with three planets since they present a

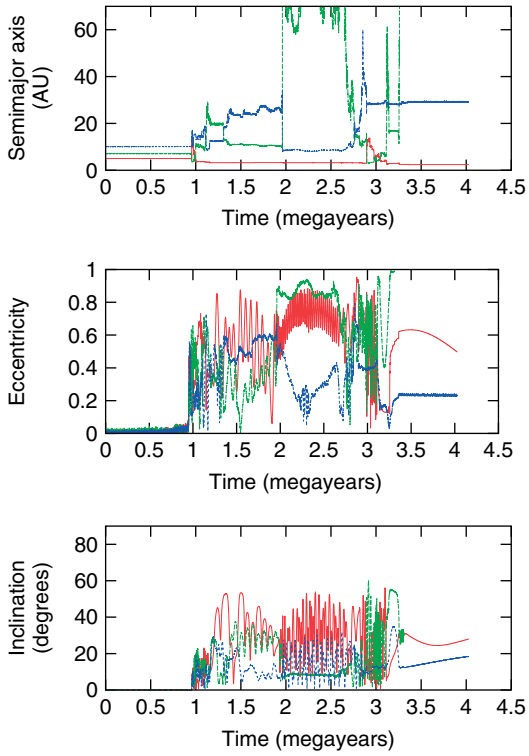


Figure 11.3 Typical time evolution of semimajor axis, eccentricity, and inclination of a system of three Jovian planets. The solid (red), green (dashed) and blue (dotted) lines illustrate the orbital parameters of the initially inner ($a = 5$ AU), middle ($a = 7.1$ AU), and outer ($a = 10.2$ AU) planets. The orbits become crossing after about 1 megayear and, after multiple close encounters, the middle planet is ejected hyperbolically ($t \sim 3.3$ megayears). The two surviving planets are left on possibly stable but highly interacting orbits. The high mutual inclination, larger than 40°, just after the ejection of the middle planet, induces strong secular perturbations of the Kozai type between the two planets with a period of about 2.6 megayears.

significantly more interesting range of dynamical evolutions compared to systems with two planets only. In Figure 11.3, the typical evolution of the semimajor axis and eccentricity of a system of three Jovian mass planets evolving around a solar mass star is illustrated. The trajectories have been computed by numerically solving the equation of motion of the hierarchical four-body problem where the planets interact with each other through gravity and, eventually, physical collisions. The three planets are placed on initially circular orbits with semimajor axes of 5, 7, and 10 AU from the star, respectively, with small inclinations (0.01°, 0.01°, 0.02°). After an initial period of orbital stability, lasting approximately 1 megayear, the system becomes chaotic. The eccentricities grow until the orbits cross and close encounters between the planets occur. Large changes in the orbital elements are induced at each gravitational approach between the planets until at least one planet,

as in the case of Figure 11.3, is ejected out from the system on a hyperbolic orbit contributing to a possible population of Jovian size dark bodies roaming around in the galaxies. The dynamical evolution of the system is very sensitive to the choice of the initial orbital parameters due to the chaotic nature of the evolution. Not only will the timescale for the onset of instability change by two orders of magnitude for close initial conditions but also the subsequent phase characterized by repeated close encounters will lead to rapidly diverging trajectories for small changes of the initial parameters.

At the end of the scattering phase, the planet injected on the inner orbit, which may be any of the initial three, has a semimajor axis of about 2.43 AU. Repeating the numerical simulation with different initial conditions will always lead to a planet on an orbit with a similar semimajor axis. This is due to the conservation of the orbital energy of the system. The initial configuration of three planets has a total energy of

$$E = -\frac{GM_*}{2}\left[\frac{m_1}{a_1} + \frac{m_2}{a_2} + \frac{m_3}{a_3}\right] \tag{11.6}$$

If there is no energy dissipation, either because of collisions between the planets or for interactions with the protostellar disk, the final configuration with one planet in an inside orbit, one in a distant orbit, and one escaping on a barely hyperbolic orbit, must have the same energy. Assuming that the outer surviving planet is on an orbit far enough to have a negligible orbital energy compared to that of the inner survivor, as it happens in most cases, then one can estimate the semimajor axis of the inner planet as

$$a_i = \simeq \frac{GM_* m_i}{2E} \tag{11.7}$$

While it is possible to estimate the final semimajor axis of the inner planet, there is no way to predict analytically a range for the semimajor axis of the outer one, which can be many times larger than the initial distance of the outermost planet. A statistical approach is possible by performing many simulations differing slightly in the initial conditions. Figure 11.4 illustrates the final orbits in the a–e plane of the planets surviving the chaotic phase for a large number of systems [12]. All the semimajor axes are scaled to that of the initially closest planet ($a_{1,i}$). The spreading in the semimajor axis distribution of the inner planet around the value predicted by Eq. (11.7) is due to the neglect of the energy terms due to the outer and ejected planets from Eq. (11.7).

As clearly shown in Figure 11.4, the inner planet's orbit may end up on a very eccentric orbit. This is the more appealing feature of the planet–planet mechanism as it explains, in a very easy and direct way, the present distribution of the extrasolar planet orbital eccentricities (Figure 1.2, Chapter 1). It also predicts that the mutual inclination of the planets in a system, which survived a phase of chaotic evolution dominated by mutual close encounters, may be significant and far from the equatorial plane of the star.

In this basic model of unstable three-planet systems, the two remaining planets at the end of the chaotic phase can be either in a dynamically stable configuration

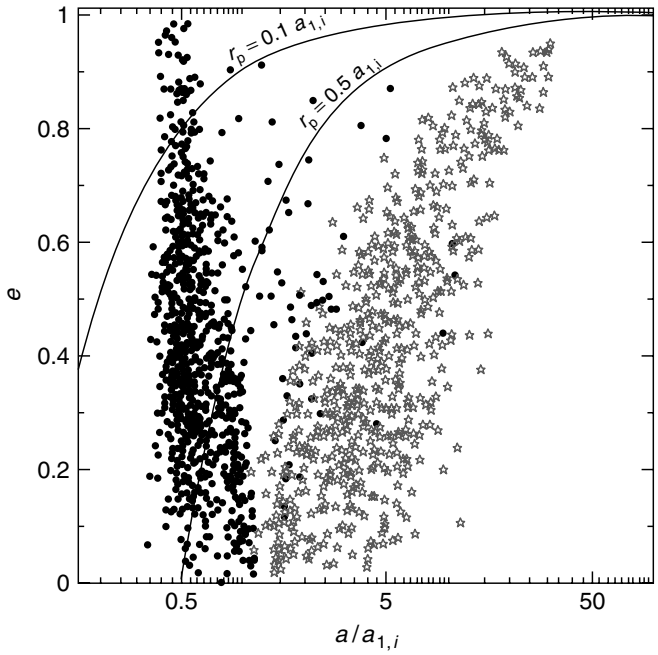


Figure 11.4 Distribution of semimajor axes versus eccentricities of the surviving planets after the ejection of one of the initial three by gravitational scattering. The black solid circles represent the inner survivor while the red empty stars show the orbital elements of the outer one. Adapted from [12].

evolving only secularly (see Chapter 3), or in a mean motion resonance [5] or they can still interact until a further ejection occurs. In this case, a single planet remains in orbit around the star on a close-in eccentric orbit. In a minority of cases, physical collisions may occur between the planets and systems with one or two more massive bodies emerging from the chaotic phase. Equation (11.7) is not anymore predictive of the final semimajor axis of the inner planet because of energy dissipation during the planet merging. The range of possible outcomes of the model can be significantly increased by varying the initial masses of the planets or by increasing the number of planets involved in the chaotic phase. Lin and Ida [4] considered planet–planet scattering scenarios where several (up to nine) equal mass giant planets formed in massive circumstellar disks which, according to Beckwith *et al.* [13], may have masses up to $0.5 M_{\odot}$ (the minimum mass solar nebula is $0.02 M_{\odot}$). They find a significantly higher chance of planet merging among the inner planets and a higher rate of planet ejections. Chatterjee *et al.* [12] used a range of planetary mass distributions in trying to better reproduce the observed orbital properties of extrasolar planets. Their results suggest that high eccentricities might be more common among less-massive giant planets.

11.2.3
Tidal Interaction with the Star and Formation of "Hot" Jupiters

The scattering model described so far predicts that the planet injected in the inner orbit at the end of the scattering phase has, in most cases, a large eccentricity that, combined with the release of orbital energy to the ejected planet(s) and consequent shrinking of the orbit, leads to a small periastron distance. Further evolution of the inner planet orbit caused by energy loss due to tidal interactions with the central star might circularize its trajectory near the periastron distance. The sequence of impulsive tidal interactions at each periastron passage causes a decrease of both semimajor axis a and eccentricity e. If the energy is damped within the planet by radial tides, then tidal circularization preserves the orbital angular momentum J_o. Dissipation within the star could transfer planetary orbital angular momentum to the stellar rotation; however, for most dissipation models, the stellar Q is much larger than that of a planet. It follows that if J_o is constant

$$J = \sqrt{GM_*}M_p\sqrt{a(1-e_p^2)} \tag{11.8}$$

When the orbit is finally circularized, the semimajor axis of the planet will be $a_p^f = q(1+e_p) \sim 2q$. When tidal dissipation halts, the end state of the system will have the inner planet on a circular orbit with a_p approximately equal to twice the initial periastron distance. As a consequence, the combination of planet–planet scattering and tidal circularization would act like conventional mechanisms of planet migration due to the planet interactions with a disk of gas or planetesimals and explain the so-called Hot Jupiters, planets with orbital periods of a few days moving very close to the host star. The computation of the secular variations of the planet's orbit due to tidal dissipation is a complex task and requires detailed calculations of the planet and star reaction to the tidal stresses. Chapter 12 is devoted to the computation of the orbital circularization of planets with periastron close to the star.

11.2.4
Kozai Oscillations

The efficiency of close-in planet production by the combined effects of planet–planet scattering and tidal circularization is significantly increased by Kozai oscillations induced by the outer planet(s) on the inner one. The long-term evolution of two planets orbiting a massive central one is usually described by the traditional Lagrange–Laplace linear classical perturbation theory. However, the large orbital eccentricities and mutual inclinations of the planets after the chaotic phase limit the application of this theory, which is based on a second-order truncation in e and i of the disturbing function. The evolution of the inner planet perturbed by the outer one(s) after the gravitational scattering phase is better approximated by the Kozai secular theory derived from a series development of the disturbing function R in r_1/r_2 where r_1 is the radial distance of the inner body and r_2 that of the outer body. The expansion of R is truncated at the first order in the ratio a_1/a_2 and

can be considered a good approximation for orbital evolution after planet–planet scattering. The Kozai Hamiltonian, after the elimination of short periodic terms depending on the orbital periods of the planets, is given by

$$H_K = \gamma(2 + 3e_1^2)\left(3\cos^2(I_1) - 1\right) + 15e_1^2\sin^2(I_1)\cos(2\omega_1) \tag{11.9}$$

where

$$\gamma = \frac{1}{16}\frac{m_2}{m_2 + M_*}\frac{n_2^2}{(1 - e_2^2)^{3/2}}a_1^2 \tag{11.10}$$

H_K does not depend on ω_1, the node longitude of the inner planet, and the related action variable, proportional to

$$\sqrt{h} = \sqrt{1 - e_1^2}\cos(I_1) \tag{11.11}$$

a constant of motion. For the evolution of the periastron argument ω_1 of the inner planet, there are two possible solutions. For $h > 0.6$, ω_1 circulates, while for $h < 0.6$ (for $e_1 \sim 0$ it implies $I > 40^\circ$), ω_1 librates either around 90° or 270°, a behavior usually referred to as the *Kozai resonance*.

The most interesting aspect of the Kozai regime, in terms of producing close-in planets, is the periodic exchange in magnitude between eccentricity and inclination. Let us consider the inner planet at the end of the chaotic planet–planet scattering phase moving on a low-eccentricity orbit with a large mutual inclination ($i_{\rm mut} \geq 40^\circ$) relative to the outer planet. The secular perturbations of the outer planet(s) cause repeated Kozai cycles of eccentricity and inclination as shown in Figure 11.5. When the large-amplitude eccentricity oscillation takes the planet in a high-eccentricity state the periastron becomes very small. Tidal dissipation occurs during these periods causing a gradual orbital decay. This kind of "Kozai migration" was pointed out by Wu and Murray [14] in the case of a planet perturbed by the companion star in a binary star system. Nagasawa *et al.* [15] have shown that Kozai migration may occur also if the outer planet acts as a perturber. The Kozai cycles coupled to tidal circularization increases the fraction of planetary systems that, at the end of the planet–planet scattering phase, can produce close-in planets. According to Nagasawa *et al.* [15], the Kozai mechanism may also be effective during the scattering phase by favoring the tidal circularization of the planet, which is temporary injected into an inside orbit. The inner planet can become temporarily dynamically isolated from the other gravitationally interacting planets and start to lose orbital energy by tidal interaction with the star. In about 30% of cases, planet–planet scattering coupled to tidal circularization can produce close-in planets.

11.3
Planet–Planet Scattering in Presence of the Gas Disk

The second scenario where planet–planet scattering plays a significant role is in the presence of *gas* in the protostellar disk. Different from the gas-free scenario, it

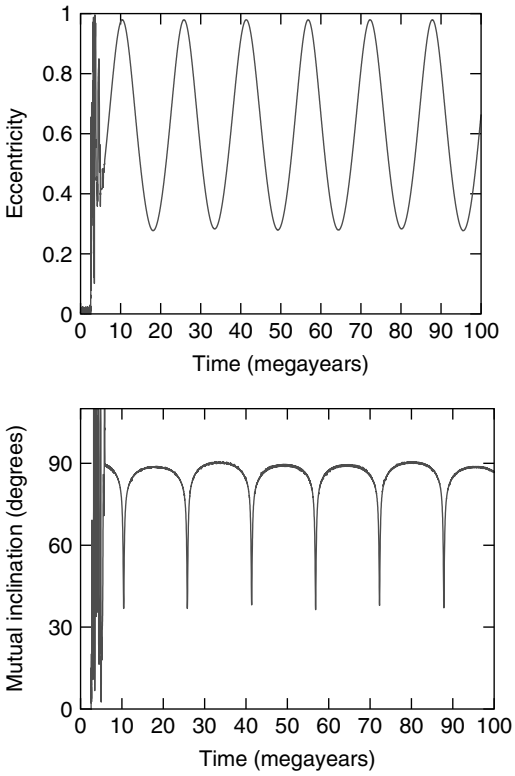


Figure 11.5 Long-term orbital evolution of the inner planet after the end of the chaotic planet–planet scattering phase. Its orbital eccentricity secularly oscillates up to large values due to the perturbations of the outer planet. When in the high-eccentricity state, the mutual inclination is at the minimum, as predicted by Eq. (11.11).

is assumed that the protostellar disk does not dissipate shortly after the formation of Jovian planets but lasts long enough to dynamically interact with the planets embedded in it. Observations of young T-Tauri stars show a wide dispersion in the lifetime of circumstellar disks ranging from 1 to 10 megayears [16,17], while giant planet formation may require 10^5–10^6 years to occur [1], even if many uncertainties are involved in the process. A multiplanet system may reach completion when the bodies are still within the disk, evolving not only under the action of the mutual gravitational perturbations but also influenced by the gas of the disk. Disk–planet interactions provide an additional strong source of mobility to the system, possibly causing any pair of planets to come closer than the stability limit and start a planet–planet scattering phase. Once started, the chaotic evolution caused by close encounters dominates the subsequent evolution of the planets and the outcome is not significantly different from the gas-free scenario. At the end of the scattering phase, if the gas component of the disk is still present, the tidal interaction between

the disk and the surviving planets may still alter the final configuration of the planetary system by moving the planets close to the star or by circularizing their orbits.

In most cases, gravitational torques between each planet and the gas of the disk result in orbital migration of the body. If the mass of the planet is lower than $\sim 30 M_\oplus$, it excites spiral density waves at the Lindblad resonances and the induced differential torque causes an inward drift called *type I migration* at a rate approximately given by

$$\frac{dr_p}{dt} \sim -c_1 q \left(\frac{\Sigma r_p^2}{M_*} \right) (r_p \Omega) \left(\frac{r}{H} \right)^3 \tag{11.12}$$

where r_p is the planet's orbital radius (assuming an almost circular orbit), c_1 is a coefficient accounting for the imbalance between the torques coming from the inner and outer side of the disk with respect to the planetary orbit [18], $q = m_p/M_*$ is the planet/star mass ratio, Σ is the surface mass density of the disk, Ω is the Keplerian angular velocity, and H/r is the disk aspect ratio (H is the disk thickness). When the Roche radius of a planet $(M_p/3M_*)^{1/3} r_p$ is larger than H, i.e., about $30 M_\oplus$ for most disk models, the planet carves a gap in the disk. After gap formation, each planet moves inward being coupled to the viscous evolution of the disk: this is type II migration. If the mass of the planet is less than or comparable to the local disk mass, the migration rate can be easily computed as

$$\frac{dr_p}{dt} \sim \frac{3\nu}{2r_p} \tag{11.13}$$

where ν the kinematic viscosity of the disk [19]. For a typical value of $\nu \sim 10^{-5}$, a planet formed at 5 AU from a solar-type star would migrate 1 AU inward in about 7×10^5 years. According to Eq. (11.13), two planets would migrate at different speeds, with the inner one drifting faster inside, and therefore diverging from each other. This would preclude the possibility of stronger gravitational interactions and close encounters between them. There is, however, a case in which the orbits may become convergent and ultimately lead to planet–planet scattering. Let us consider, as an example, two Jovian mass planets embedded in a viscously evolving disk. If the planets formed sufficiently close, or moved close when they were protoplanets under the action of type I migration (which strongly depends on the planet mass), their combined tidal effects cause the clearing of the region between them because of overlapping Lindblad resonances. According to numerous hydrodynamic simulations [20–23], the timescale for clearing is only a few hundred orbital periods. In most cases, the region inside the orbit of the inner planet quickly loses material due to accretion onto the host star so that the two planets move within a cavity in the protostellar disk, which extends beyond the outer planet [22]. As a consequence, the inner planet does not migrate because there is no torque-exciting disk material left close to it. On the other hand, the outer disk exerts negative torques at the Lindblad resonances on the outer planet, which gradually migrates inward moving closer to the inner one. Depending on

the migration speed of the outer planet and the mass ratio between the two planets, the system may follow three alternative paths.

1) If the migration speed of the outer planet converging toward the inner one is too high, resonant capture will not occur and the two planets will start to have encounters as they approach each other. Planet–planet scattering will be the dominant mechanism after the first close encounter and the evolution will be similar to that described for the gas-free case. At the end of the chaotic phase, additional migration may occur changing the final orbital configuration of the system.
2) For lower migration speeds, a stable mean motion resonance, in most cases, the 3 : 1, 2 : 1, and 3 : 2, can form between the planets. Subsequent to the resonant capture, the two planets migrate together maintaining the commensurability. They move inward or outward depending on their mass ratio and physical parameters of the disk while their eccentricities are confined to low values by the disk material preventing mutual close encounters. For masses comparable to that of Jupiter and Saturn, migration reversal may occur and the system in resonance, instead of drifting all the way to the central star, may move outward [24]. This might explain why the giant planets of the solar system are relatively far from the central star compared to most of the newly discovered extrasolar planets. Once the gas dissipates, either they remain in resonance or additional interactions with a planetesimal disk [25] may disrupt the resonance leaving the planets uncoupled but still far enough to prevent mutual close encounters. This scenario works for quiet disks but it fails if the disk is turbulent, as expected in the presence of magnetorotational instability (MRI) [26], since the fluctuations in the disk can affect the planets and destabilize their resonant locking. According to Adams *et al*. [27], the fraction of planetary systems that survive in resonance in turbulent circumstellar disks decreases approximately as the square root of time and, after about 1 megayear, only a small percentage of resonant systems will still be in resonance. Once out of the resonance, the planets still embedded in the turbulent disk will possibly approach each other and start a chaotic planet–planet scattering phase.
3) The third kind of evolution predicts that the planets, while in resonance, have their eccentricities pumped up by the combination of resonant interactions and dissipative tidal forces of the disk. As a result, when a threshold eccentricity is exceeded, the system becomes unstable and close encounters between the planets are triggered. According to Adams and Laughlin [28], in 60% of the cases one planet is lost by ejection, 20% are accreted by the host star, while 1% end up in collisions between the two planets. Figure 11.6 shows the behavior of a Saturn-like (outer) and Jupiter-like (inner) planet under the action of the tidal force of the disk. The calculation of the orbital evolution has been performed within a three-body model (central stars and the two planets) with an additional dissipative force that reproduces the effects of the disk–planet interaction, like in [29, 30]. The outer planet migrates inward quite rapidly

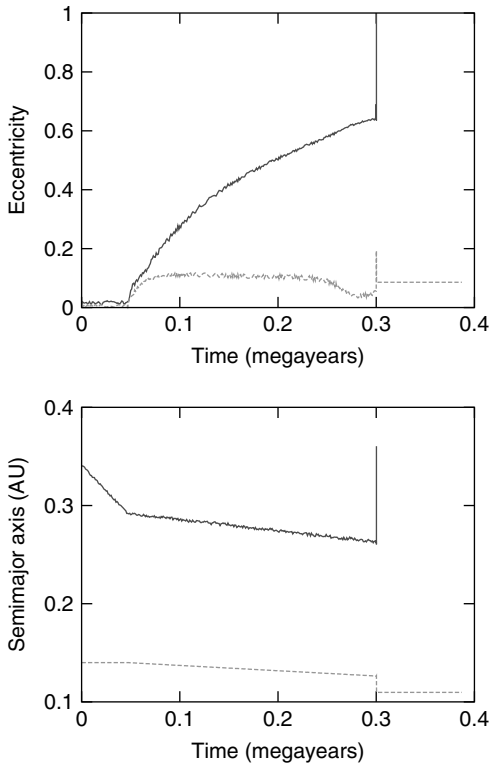


Figure 11.6 Numerical simulations showing the resonant capture of two giant planets (the outer one with the mass of Saturn and the inner one with the mass of Jupiter) migrating toward the star. The migration is simulated with a dissipative term added to the Newtonian equations of motion. Once in resonance, the two planets drift inside together until the eccentricity of the outer planet grows beyond the crossing value. A phase of planet–planet scattering leads to the ejection of the Saturn-like planet, while the inner one is pushed on an orbit closer to the star.

and the two planets are captured in the 3 : 1 mean motion resonance. After a period of coupled migration, during which the eccentricity of the outer planet is pumped up, the two planets enter a short chaotic phase dominated by planet–planet scattering and, finally, the outer planet is ejected on a hyperbolic trajectory. The inner planet is left on a stable orbit, which might further interact with the disk in the future even if the surrounding region of the disk is cleared at the moment of the planet–planet scattering event. The critical condition for this kind of evolution is the rapid dissipation of the inner disk. Crida *et al.* [31] have shown that if the dissipation is delayed, in some cases, the inner disk may damp the inner planet eccentricity and maintain the resonance, preventing the onset of chaotic behavior.

Which path will be followed by each individual planetary system strongly depends on the parameters of the disk like total mass, density and temperature profiles, and

viscosity. Also the mass of the planets and their initial orbits may influence the outcome of resonance trapping during migration.

11.4 Planet–Planet Scattering in Binary Stars

A binary system consists of two stars orbiting around their common barycenter. It is believed that more than 50% of all solar-type stars are in binary (or to a less extent multiple) systems [32]. A large variety of orbital configurations have been observed among binaries, from pairs so close to be almost in contact to pairs so far away that their link is indicated only by their common proper motion. Binary systems with periods longer than ∼3 years generally have very elliptical orbits with eccentricities ranging between 0.3 and 0.9. Those binaries with periods of less than 3 years have lower eccentricities, in between 0.15 and 0.45, possibly circularized by tidal dissipation.

There is also some correlation between the mass ratio and the orbital separation. For binary systems with orbital periods longer than ∼100 years, the secondary (less-massive) star tends to be of very low mass, while for closer binaries the secondary star's mass tends to be close to the mass of the primary.

At present, about 20% of known extrasolar planetary systems have been found around stars possessing one or more companions, but this percentage may hide some observational bias. Anyway, planets in binaries significantly contribute to a comprehensive inventory of extrasolar planets. The presence of a secondary star may alter the planetary formation process and evolution leading to different outcomes than for single stars. Both circumstellar and circumbinary disks develop spiral density and bending waves in response to the gravitational perturbations of the companion star and they may also become eccentric. The behavior of the disk may influence planetary evolution due to tidal interaction with the disk and alter the final outcome of migration. Here, we concentrate on how the presence of a companion star affects the planet–planet scattering dynamics around the primary star. The companion, particularly in those systems with semimajor axes lower than 100 AU, has a strong gravitational influence on planets around the primary by exciting secular changes of their eccentricities and inclinations. These perturbations, which depend on the binary mass ratio μ, eccentricity e_b, and semimajor axis a_b, can even cause wide Kozai-type oscillations if the mean orbital plane of the planetary system is significantly inclined with respect to that of the companion star.

In the following, we outline the major differences between planet–planet scattering in planetary systems around the primary star of a binary and planet–planet scattering around single stars. Our benchmark will be a system with three initial planets with masses comparable to that of Jupiter orbiting a solar-type star.

- The initial spacing between the planets that leads to instability can be much larger than that computed for three planets around a single star. If the initial separation between the planets' semimajor axes is given as K-times R_H, a value of K leading

to a stable system around a single star might give an unstable system if the star is part of a binary. On average, the corresponding value K' granting stability for a three-planet system in a binary will always be $K' > K$. The difference between K' and K grows quickly by increasing e_b or reducing a_b, which corresponds to a stronger perturbing configuration. The dependence on μ is instead less trivial.

- Compared with the single-star case, systems with three initial planets in binaries are more likely to leave a single survivor. According to the long-term numerical investigation of Holman and Wiegert [34], there is a limiting value of the semimajor axis a_c inside of which planets are in stable orbits around the central star of a binary. This value can be easily computed on the basis of their semiempirical formula:

$$\begin{aligned} a_c = {} & [(0.464 \pm 0.006) + (-0.380 \pm 0.010)\mu \\ & + (-0.631 \pm 0.034)e_b + (0.586 \pm 0.061)\mu e_b \\ & + (0.150 \pm 0.041)e_b^2 + (-0.198 \pm 0.074)\mu e_b^2]a_b \end{aligned} \tag{11.14}$$

Beyond a_c, the companion perturbations destabilize the orbit, ejecting the planet in a hyperbolic orbit. After the onset of planet–planet scattering, the chaotically evolving planets may end up, even temporarily, on an orbit with semimajor axis larger than a_c. If this is the case, the subsequent dynamical evolution of the planet is strongly affected by the companion and ejection from the system is very likely. Marzari *et al.* [33] have shown that the fraction of planetary systems retaining two planets in a stable configuration at the end of the chaotic phase is significantly reduced in binaries. As shown in Figure 11.7, the binary eccentricity

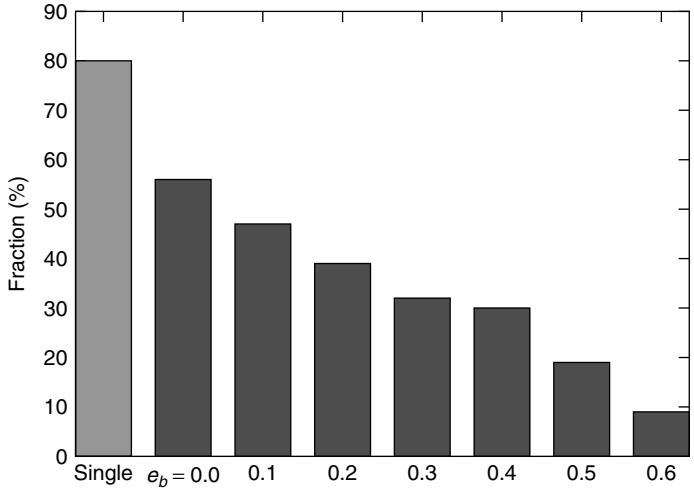


Figure 11.7 Fraction of initially three-planet systems that end up with two planets in stable orbit for single stars (green bar) and binary stars (red bars) with different eccentricities. The mass ratio of the stellar pair is $\mu = 0.4$, while the semimajor axis is fixed at 50 AU. Data taken from [33].

e_b is predictive of the fraction of systems ending up with two planets, which decreases almost linearly with e_b.
- The gravitational scattering phase may build up a significant inclination not only between the two surviving planets but also between the planets and the companion star. When only one planet survives the chaotic phase, it may be in a Kozai resonance with the secondary star. This might explain the orbit of the planet found around 16 Cyg B, part of a binary system, which has an eccentricity of 0.67. This large eccentricity could be either directly the outcome of the planet scattering phase or the high-eccentricity phase of a Kozai cycle [35].

The number of impacts on the primary star of one of the initial planets is much higher than that on the secondary star. If the companion did not develop a planetary system by its own, it might be possible to detect a different degree of metallicity between the two stars of the pair.

11.5 Summary

Planet–planet scattering is a dynamical mechanism, which can significantly alter the orbital configuration of a planetary system after its formation. The individual planetary eccentricities can be increased well beyond those observed in the solar system; the initial semimajor axis architecture can be completely mixed up and some planets can be ejected out of the system. A phase of gravitational scattering may occur either because a newly born planetary system is dynamically unstable or because the tidal interactions of planets with the gaseous component of the disk bring them close enough to start strong mutual gravitational interactions. Modeling the effects of planet–planet scattering in trying to reproduce observed planetary systems is a complex task since the parameter space of multiplanet systems is huge. In addition, the protostellar disk may have an active role in preventing or favoring planet encounters. A statistical approach is needed to fully explore the planet–planet scattering phase based on a detailed sampling of the following parameters:

- The number of planets that emerge from the circumstellar disk, potentially varying from 2 to 10 or more.
- Their initial mass distribution that can range from Saturn-like planets up to bodies with 10–12 Jupiter masses.
- The orbital spacing among the bodies and their eccentricity and inclination. These last two parameters are supposed to be small for planets forming within a disk; however, planet–planet scattering is chaotic and also a small change in the initial conditions may lead to totally different outcomes.
- The amount of the disk gas left over by the planet formation process and the viscosity value. Both these parameters influence the migration rate of the planets before the final gas dispersal. Convergent migration can occur at different rates.

- The mass left in the disk of planetesimals after planet formation. The bodies in the disk gravitationally interact with the planets and may alter the delicate equilibrium of a barely stable multiplanet system. It can cause orbital migration, as hypothesized for the solar system [25], and excite the eccentricity of the planets.
- The parameters describing the tidal interaction with the star.

Also a more detailed numerical model is needed to simulate both the *N*-body and hydrodynamical aspects of the problem. This model must be able to accurately follow the different phases of planet formation and subsequent dynamical evolution in order to theoretically predict the present orbital and mass distribution of extrasolar planets. Such a model is not currently available due to the complexity and numerical load. However, a first attempt in this direction with a limited approach has been performed by Moorhead and Adams [36]. They considered two-planet systems evolving under the action of disk torques and planet–planet scattering. Many simulations have been performed with different initial masses for the planets and different parameters for type II migration. The final distributions of the orbital elements of the surviving planets are in reasonable agreement with those of the observed extrasolar planets. It is noteworthy that the interaction between large planetesimals in the remnant disk and a newly born planetary system may explain the apsidal behavior observed among extrasolar planetary orbits. According to Barnes and Greenberg [37], multiplanet extrasolar systems tend to have their orbits near the boundary (separatrix) between apsidal libration (i.e., the difference in periastron longitudes of two planets, $\Delta\varpi$, oscillates about a fixed value) and circulation ($\Delta\varpi$ oscillates through 360°). One way to produce this "near-separatrix" secular motion is to excite the eccentricity of one planet prior to the onset of the violent planet–planet scattering phase. This eccentricity can be the consequence of repeated encounters of the planet with planetesimals crossing its orbit. Barnes and Greenberg [38] suggested that a single but massive planetesimal (a "rogue protoplanet") passing close to a regular planet also might introduce enough orbital eccentricity in its orbit. The subsequent chaotic evolution of the planetary system would end up mostly with a near-separatrix motion. This mechanism confirms the need of a multicomponent model to be used for reproducing observational data.

References

1 Alibert, Y., Mordasini, C., Benz, W., and Winisdoerffer, C. (**2005**) *Astronomy and Astrophysics*, **434**, 343.

2 Weidenschilling, S.J. and Marzari, F. (**1996**) *Nature*, **384**, 619.

3 Rasio, F.A. and Ford, E.B. (**1996**) *Science*, **274**, 254.

4 Lin, D.N.C. and Ida, S. (**1997**) *The Astrophysical Journal*, **477**, 781.

5 Raymond, S.N., Barnes, R., Armitage, P.J., and Gorelick, N., *The Astrophysical Journal*, **687**(2), pp. L107–L110.

6 Pollack, J.B., Hubickyj, O., Bodenheimer, P. *et al.* (**1996**) *Icarus*, **124**, 62.

7 Wetherill, G.W. and Stewart, G.R. (**1989**) *Icarus*, **77**, 330.

8 Johansen, A., Oishi, J.S., Low, M.M. *et al.* (**2007**) *Nature*, **448**, 1022.

9 Gladman, B. (**1993**) *Icarus*, **106**, 247.

10 Chambers, J.E., Wetherill, G.W., and Boss, A.P. (**1996**) *Icarus*, **119**, 261.
11 Marzari, F. and Weidenschilling, S.J. (**2002**) *Icarus*, **156**, 570.
12 Chatterjee, S., Ford, E.B., Matsumura, S., and Rasio, F.A. (**2008**) *The Astrophysical Journal*, **686**, 580.
13 Beckwith, S.V.W., Sargent, A.I., Chini, R.S., and Guesten, R. (**1990**) *The Astronomical Journal*, **99**, 924.
14 Wu, Y. and Murray, N. (**2003**) *The Astrophysical Journal*, **589**, 605.
15 Nagasawa, M., Ida, S., and Bessho, T. (**2008**) *The Astrophysical Journal*, **678**, 498.
16 Strom, K.M., Strom, S.E., Edwards, S., *et al*. (**1989**) *The Astronomical Journal*, **97**, 1451.
17 Haisch, K.E., Lada, E.A., and Lada, C.J. (**2001**) *The Astrophysical Journal*, **553**, L153.
18 Ward, W.R. (**1997**) *Icarus*, **126**, 261.
19 Nelson, R.P., Papaloizou, J.C.B., Masset, F., and Kley, W. (**2000**) *Monthly Notices of the Royal Astronomical Society*, **318**, 18.
20 Snellgrove, M.D., Papaloizou, J.C.B., and Nelson, R.P. (**2001**) *Astronomy and Astrophysics*, **374**, 1092.
21 Papaloizou, J.C.B. (**2003**) *Celestial Mechanic and Dynamical Astronomy*, **87**, 53.
22 Kley, W., Peitz, J., and Bryden, G.A. (**2004**) *Astronomy and Astrophysics*, **414**, 735.
23 Kley, W., Lee, M.H., Murray, N., and Peale, S.J. (**2005**) *Astronomy and Astrophysics*, **437**, 727.
24 Masset, F. and Snellgrove, M. (**2001**) *Monthly Notices of the Royal Astronomical Society*, **320**, L55.
25 Tsiganis, K., Gomes, R., Morbidelli, A., and Levison, H.F. (**2005**) *Nature*, **435**, 439.
26 Balbus, S.A. and Hawley, J.F. (**1991**) *The Astrophysical Journal*, **376**, 214.
27 Adams, F.C., Laughlin, G., and Bloch, A.M. (**2008**) *The Astrophysical Journal*, **683**, 1117.
28 Adams, F.C. and Laughlin, G. (**2003**) *Icaurs*, **163**, 290.
29 Lee, M.H. and Peale, S.J. (**2002**) *The Astrophysical Journal*, **567**, 596.
30 Beaugé, C., Michtchenko, T.A., and Ferraz-Mello, S. (**2006**) *Monthly Notices of the Royal Astronomical Society*, **365**, 1160.
31 Crida, A., Sandor, Z., and Kley, W. (**2008**) *Astronomy and Astrophysics*, **483**, 325.
32 Duquennoy, A. and Mayor, M. (**1991**) *Astronomy and Astrophysics*, **248**, 485.
33 Marzari, F., Weidenschilling, S.J., Barbieri, M., and Granata, V. (**2005**) *The Astrophysical Journal*, **618**, 502.
34 Holman, M.J. and Wiegert, P.A. (**1999**) *The Astronomical Journal*, **117**, 621.
35 Holman, M.J., Touma, J., and Tremaine, S. (**1997**) *Nature*, **386**, 254
36 Moorhead, A.V. and Adams, F.C. (**2005**) *Icarus*, **178**, 517.
37 Barnes, R. and Greenberg, R. (**2006**) *The Astrophysical Journal*, **638**, 478.
38 Barnes, R. and Greenberg, R. (**2007**) *The Astrophysical Journal*, **659**, L53.

Further Reading

Kozai, Y. (**1962**) *The Astronomical Journal*, **67**, 591.

12
Tides and Exoplanets

Brian Jackson, Rory Barnes, and Richard Greenberg

12.1
Introduction

The discoveries in recent years of hundreds of exoplanets have dramatically expanded our theories of planet formation and evolution. The wide range of orbital and physical properties of extrasolar systems have shown that planetary systems can be unlike our own, forcing a revision of canonical knowledge. For example, the orbital eccentricities e of exoplanets span a broad range, from 0 to nearly 1, and orbital semimajor axes a run the gamut from a few hundredths of an AU to more than 100 AU. The radii of exoplanets, as determined from transit observations, can also be very different from Jupiter and Saturn, even if the planets share similar masses and compositions. These characteristics provide important clues about the formation and evolution of exoplanets.

Even among exoplanets, close-in planets stand out as unusual. As shown in Figure 1.2 in Chapter 1, the close-in planets do not have the same eccentricity distribution as the planets in more distant orbits. In fact, Jackson *et al.* [1] showed that planets beyond 0.2 AU share a remarkably similar e distribution that is very different from close-in planets. This result suggests that whatever mechanism endowed the distant exoplanets with large eccentricities either did not operate at small astrocentric distances, or that some postformation process circularized the orbits.

Being so close to their host stars, close-in planets have undoubtedly been affected by tides, which modified their orbits and physical characteristics. Indeed within 1 year of the discovery of the planet 51 Peg b, it was suggested that tides could circularize the orbits of close-in planets on a timescale shorter than the host star ages [2]. This hypothesis remains the explanation for the disparate e distributions, and has many important consequences.

As the planets and stars rotate and orbit one another, the shape of the tide on each shifts in response to the changing gravitational potential, driving dissipation of energy, which in turn reduces e and a over billions of years. The dissipation of energy within observed close-in planets may affect their internal structures and

Formation and Evolution of Exoplanets. Edited by Rory Barnes

ISBN: 978-3-527-40896-2

inflate the radii of many transiting planets. Consideration of these effects must be included in models of formation and evolution.

These effects may also dominate the orbital and physical properties of rocky exoplanets. Because rocky exoplanets tend to have masses much smaller than their gaseous counterparts, initially they may only be detected in close-in orbits. Tides may determine the rotation rates of these planets, affecting the atmospheric transport of heat; tidal evolution of their orbits may move them too close to their host stars to be habitable; and dissipation of tidal energy within the planets may drive vigorous geophysical processes. Although criteria for planetary habitability remain poorly understood, the potential magnitudes of these tidal effects suggest they may largely determine whether these rocky planets are hospitable to life.

In this chapter, we discuss applications of tidal models to exoplanetary evolution, dynamics, and habitability. In Section 12.2, we acquaint the reader with tidal theory. In Section 12.3, we review recent work on the effects of tides on the orbital and physical properties of observed close-in planets. In Section 12.4, we discuss work on the effects of tides on rocky exoplanets and on their habitability.

12.2 Tidal Physics

Tides are important in a variety of astrophysical contexts, including binary star systems [3–8], satellite systems of our own solar system [9–17], and in extrasolar systems [1, 2, 18–28]. In extrasolar systems, tides can modify the rotation rates and orbits of planets and stars, and the wide range of orbital and physical properties for close-in exoplanets has prompted the development of new tidal theories. Although the accuracy of tidal mode may be limited by our poor understanding of the tidal response of planets and stars, most models provide similar predictions for the effects of tides.

The gravitational potential on a body A due to another mass B can vary considerably across the diameter of A and will distort its shape to some degree; see Figure 12.1. The point on the surface of A that is closest to B experiences a stronger gravitational pull than the point that is farthest. The result is that A becomes elongated along the line connecting the centers of the two bodies and shortened about the two orthogonal directions. Body A can have a similar effect on B.

When measured from the center of A, the tidal potential from B can be expressed as a spherical harmonic of second order [29–31]. The height of the bulge (relative to A's shape in isolation) depends on the self-gravity, internal structure, and rigidity of A and is traditionally encapsulated in the Love numbers k_2 and h_2 [29, 30, 32, 33]. To first order, deformations due to other fluid or elastic motions within A (due, for example, to convection) can usually be neglected relative to the tidal deformation. However, in some cases, these deformations may also play important roles [34, 35].

As body A rotates on its axis and revolves in its orbit, its shape continually adjusts to the changing tidal potential. In the case of a circular orbit that is not inclined relative to A's equator, the maximum tidal force passes periodically through a point

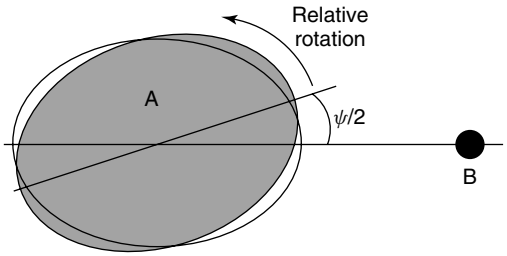


Figure 12.1 Tidal deformation of a planetary or stellar body, A, caused by another body, B. Body A's bulge tries to point toward B (unshaded figure), but its rotation pulls the lag ahead (shaded figure). The angle between the tidal bulge and the line connecting A's and B's centers is $\epsilon/2$ [31, 37, 38].

fixed on its surface with a frequency $2(\Omega - n)$, where Ω is A's rotation rate and n is the orbital mean motion. Note that the bulge is not fixed in longitude, rather A's shape must constantly change to keep up with the changing potential.

The above picture is idealized, as it assumes that A can react instantaneously to the evolving tidal potential. Real materials cannot do this, and so the body's shape lags the tidal potential by some small phase angle ϵ. At time t, dissipation within A causes the body to take the shape of the tidal potential from a short time in the past, $t - \frac{\epsilon}{2(\Omega - n)}$. (Note that ϵ may depend on frequency, see below). As a result, the line of symmetry in A is offset by a small angle $\epsilon/2$ from the line connecting the centers of the two bodies (Figure 12.1). The bulge on A may point ahead or behind the tide raiser, depending on the relative values of Ω and n. The misalignment results in a gravitational torque that can change rotation rates and orbits. More dissipation leads to a larger lag angle, which leads to a larger torque. The torque drives the rotation rate to an equilibrium value (the rate for which the tidal torque averaged over an orbit is zero), and the timescale to reach this equilibrium rate scales with ϵ [1, 2, 31, 36].

The tidal response of planets and stars is usually framed by an analogy with the damped, driven harmonic oscillator [39]. The tidal potential of B drives deformation of A, just as a motor may drive periodic displacement in a harmonic oscillator. Body A's internal elasticity and self-gravity act to restore A's shape, like the spring in the oscillator, while dissipation within A acts to delay A's tidal response to the potential, like a dashpot. In keeping with this analogy, the phase lag is conventionally cast in terms of the quality factor or tidal dissipation parameter Q, which is equal to the ratio of the maximum energy stored in the tidal deformation to the energy dissipated over one cycle [3, 40–42]. Thus smaller Q indicates more dissipation and a larger ϵ. For planets and stars Q is usually $\gtrsim$100. In this limit of large Q, Q is equal to $1/\epsilon$ [38].

If the orbital eccentricity and/or inclination relative to A's equator i is nonzero, the situation is more complicated, and the response of A can be thought of as a harmonic oscillator driven by motors operating at multiple frequencies [39].

In this case, it is common practice to expand the tidal potential as a Fourier series involving frequencies related to the rotational rate and orbital mean motion [29–31, 43]. Terms in the series have amplitudes proportional to various powers of e and i, and the higher the order of the expansion in e and i, the more frequencies appear in the tidal potential. For example, when $e = i = 0$ (the example above), there is only one tidal component with a frequency $2(\Omega - n)$. At first order in e and i, the other relevant frequencies are $2\Omega - 3n$, $2\Omega - n$, $\Omega - 2n$, and n. The tidal response to each of these components is usually assumed to be independent of the others, so the total tidal response of body A is the linear sum of the individual responses, just as the total displacement of the harmonic oscillator is the linear sum of its responses to each component (motor) of the forcing [30, 31, 39]. The sum is used to determine the effect of the total tidal bulge on the orbital semimajor axis, eccentricity, and inclination [31].

It is important to make clear the distinction between the mathematical description above and the actual deformation. The shape of body A changes with time owing to the apparent angular position, velocity, and size of the perturber, which is modeled as waves that sweep over the surface with different frequencies and amplitudes. The tidal perturbation does not actually set up such waves on the surface, rather, we presume the shape can be represented by the superposition of these ''waves''.

The dependence of the tidal response and dissipation rates on frequency plays an important role in determining the effects of tides on orbits and rotations in current models of tidal evolution. Unfortunately, tidal response and dissipation within planets and stars is probably complex and still poorly understood. Direct observational evidence for the effects of tides is lacking, in part because they are too small to measure for bodies other than the Earth and Moon [44] and Jupiter and Io [16]. No clear consensus exists for the frequency dependence of dissipation, which is needed to determine the contribution of each component in the tidal response. As discussed below, various relationships have been considered: a phase lag proportional to frequency, or constant with respect to frequency, or more complex relationships. Thus interpretations of tidal models, particularly where a wide range of frequencies is involved, should be regarded with caution [39]. Nevertheless, tidal models have provided powerful insight into the origins and properties of exoplanets.

Among studies of the tidal evolution of binary star systems, it is common to assume the phase lag is proportional to the frequency [3, 8, 45, 46]. For the case of a circular orbit (with only a single frequency), the lag angle ϵ is $\tau \times 2(\Omega - n)$, where τ is a constant (which for stars, is probably related to time for convective overturn in the stellar atmospheres [3, 47]). For eccentric or inclined orbits, each component of the tidal response will lag by the same amount of time τ, resulting in a different phase lag for each component. In effect, each component has its own Q. This approach is sometimes called the *constant-time-lag model*.

In studies of planetary systems, some studies [30, 48] suggest that the phase lag has a magnitude independent of frequency [1, 30, 37, 39, 48]. Tidal response components with positive frequencies have positive phase lag angles, and components with negative frequencies have negative phase lag angles, but in both cases they

have the same magnitude, $1/Q$ [1, 30, 37, 39, 48]. Thus, each component has a different lag time. This approach is sometimes called the *constant-phase-lag model*.

The assumptions of constant time lag or constant phase lag permit useful estimates of the effects of tides, but are undoubtedly simplifications of the true tidal response of celestial bodies. Recent theoretical studies of the hydrodynamic response of tidally perturbed fluid bodies suggest a strong dependence on the internal stratification of the bodies, giving rise to a complex dependence of the phase lag on frequency [41, 42, 49–52]. For rocky bodies, a recent study suggests that the phase lag may scale as the tidal frequency to some fractional power [38]. However, the observational and experimental data required to determine conclusively the tidal response of celestial bodies are lacking. Moreover, the assumption that the total tidal response is simply the sum of the various components may not be valid for such complex frequency dependence [30, 39].

Even without detailed knowledge of tidal response, we can apply tidal models to extrasolar systems. The properties of extrasolar systems allow us to make several simplifying assumptions: (i) The masses of exoplanet-hosting stars are much larger than the planetary masses, so we assume n is independent of planet mass. (ii) When they have been measured, the orbital inclinations relative to their host star's equator for most (but not all) planets are negligible [53–56], so we assume they are zero. (iii) The timescales for bringing planetary rotation rates to equilibrium values are less than a few million years, so the tide raised on the close-in planets reduces orbital energy but not the orbital angular momentum [1, 2, 57]. (iv) When they have been measured, stellar rotation rates are usually much smaller than orbital mean notion, so the tide raised on the stars, however, reduces the orbital angular momentum and energy [57, 58]. However, this latter assumption may break down in a small number of cases [27].

In the parameter space occupied by extrasolar systems, all tidal models give similar results, so we can understand, qualitatively at least, tidal evolution in extrasolar systems. Tides tend to reduce a and e and can dissipate significant energy within the planets. Rather than present a comprehensive comparison of all tidal models, we focus on the constant-phase-lag model, noting the important differences between models wherever they are relevant.

For this model, the rates of change of e and a, to second order in e, are

$$\frac{1}{e}\frac{de}{dt} = -\left(\frac{63}{4}(GM_*)^{1/2}\frac{R_p^5}{Q_p' M_p} + \frac{225}{16}(G/M_*)^{1/2}\frac{R_s^5 M_p}{Q_*'}\right) a^{-13/2} \tag{12.1}$$

$$\frac{1}{a}\frac{da}{dt} = -\left(\frac{63}{2}(GM_*)^{1/2}\frac{R_p^5}{Q_p' M_p} + \frac{9}{2}(G/M_*)^{1/2}\frac{R_s^5 M_p}{Q_*'}\left(1+\frac{57}{4}e^2\right)\right) a^{-13/2} \tag{12.2}$$

Here G is the gravitational constant, and subscripts p and s refer to the planet and star, respectively. R is a body's radius, M is its mass and $Q' = 3Q/2k_2$, encapsulating both the unknown Love number and tidal dissipation parameter. The effects of the tide raised on the star by the planet are reflected in the terms involving Q_*', while the terms involving Q_p' reflect the tide raised on the planet by

the star. The tide raised on the planet reduces the orbital energy without affecting the angular momentum, while the tide on the star reduces both.

Note that initial investigations into tidal evolution of exoplanets made numerous simplifications to Eqs. (12.1 and 12.1), most notably the neglect of the tide raised on the star and the coupling between a and e [2, 20, 22, 24, 57]. These assumptions led to an inappropriately short estimate for the timescale for circularization, and predicted no evolution in a. As a result, a misconception emerged that close-in planets more than a few billion years old should be on circular orbits, or that additional planets must be present to pump up the eccentricity via planet–planet interactions ([59]; see also Chapter 3). A more comprehensive analysis showed that additional planets are not necessary in all cases [1], but cannot be ruled out either. The results shown in Sections 12.3 and 12.4 incorporate all terms in Eqs. (12.1 and 12.1).

The dissipation of energy within a close-in exoplanet due to tides reduces the planet's orbital energy (except very early on while the planetary spin is quickly brought to equilibrium [60]). The rate of energy dissipation in the planet H is

$$H = \left(\frac{63}{4}\right) \frac{(GM_*)^{3/2} M_* R_p^5}{Q_p'} a^{-15/2} e^2 \tag{12.3}$$

H is therefore often called the *tidal heating rate*. Note that since a and e evolve with time, H will also do so. In some cases, it is also useful to consider the heat flux through the surface, which is $h = H/4\pi R_p^2$.

The above equations assume that, over the course of an orbit, there is no net transfer of angular momentum between the planetary rotation and orbit, a situation called *tidal locking*. The time for locking to occur depends on the initial orbit and rotation, but is probably on the order of millions of years for close-in exoplanets [2, 61]. The constant-phase-lag model predicts the equilibrium rotation rate of the planet $\Omega_{\rm equil}$ will be

$$\Omega_{\rm equil} = n(1 + \frac{19}{2} e^2) \tag{12.4}$$

Note that for $e \neq 0$, the equilibrium rate rate is greater than the orbital mean motion [31, 32, 48, 61].

Some confusion has arisen in the literature regarding the distinction between tidal locking and synchronous rotation. Kasting *et al.* [62] pointed out that rocky planets inside the habitable zone (HZ) of low-mass stars will be tidally locked, and, as they considered only circular orbits, tidal locking meant the planets would be synchronously rotating ($\Omega_{\rm equil} = n$). They believed that synchronous rotation has potentially deleterious effects on habitability, leading to the belief that the planets orbiting M stars are not good places to look for life. In actuality, synchronous rotation is only one type of tidal locking that may occur when $e = 0$. We return to this point in Section 1.4.3.

Planets can become locked in a spin–orbit resonance other than synchronous rotation. In these cases, as the eccentricity damps, and $\Omega_{\rm equil}$ drops, it is possible for the rotation frequency to become resonantly locked with the orbital frequency. This generally only occurs because of inhomogeneities in the body, and hence is

more likely to occur in solid bodies. For example in the solar system, Mercury is in a 3 : 2 spin–orbit resonance, and the Moon is in a 1 : 1 resonance. In the latter case, the Moon's eccentricity is 0.055, and Eq. (12.4) would predict that the Moon would make one additional rotation every 3 years, but because of internal mass asymmetries, the Moon actually librates. So, although the Moon appears to be in synchronous rotation, it is better described as in a 1 : 1 spin–orbit resonance.

Other tidal models have included terms involving higher powers of e, which may be particularly important for extrasolar systems where e can be near unity [21, 45, 63]. However, for these higher-order terms, the range of frequencies that contribute to the tidal evolution expands and so the frequency dependence of the tidal response becomes ever more important. However, the true frequency dependence of the tidal response remains elusive, so it is not clear that these models are any more accurate than the lower-order one shown here [31, 39]. In any case, the higher-order models generally predict somewhat faster orbital evolution [64].

Interactions between a tidally evolving planet and other planets or stars in the system may also affect a tidally evolving planet's orbit. If the interactions are of a secular or resonant nature, they can cause the orbital eccentricity to oscillate over short timescales ($\leq$10 000 years; see Chapter 3), and they can slow the rate of tidal eccentricity damping [21, 65]. This effect may be responsible for substantial eccentricities observed for some close-in planets [64–66]. Kozai interactions (see Chapter 3) between a tidally evolving planet and a distant stellar or substellar companion may also cause the orbital eccentricity and inclination to oscillate [67] and may be partly responsible for producing close-in planets in the first place [68]. Currently, most close-in planets are observed in single-planet, single-star systems, but, as detection techniques improve, more multiplanet, multistar systems will be detected, and the interplay between tidal- and planet–planet interactions will become more apparent.

12.3 Tidal Effects on Gaseous Exoplanets

In this section, we review research into the orbital evolution and tidal heating of known, gaseous exoplanets.

12.3.1 Orbital Evolution

The dichotomy between eccentricities of close-in planets and planets farther out (see Figure 1.2 – in Chapter 1) is very likely the result of tidal evolution [1, 2]. Because the magnitude of tidal effects falls off very rapidly with increasing a, the action of tides could plausibly have reduced e for close-in planets, but not for those farther out. In this scenario, close-in exoplanets begin on more eccentric orbits that circularize over long timescales until they are observed (i) to be effectively circular, (ii) with a residual eccentricity that has not fully damped out, or (iii) with an eccentricity maintained by a companion.

By integrating the tidal evolution equations backward in time for close-in exoplanets for the age of the system, [1] showed that many close-in planets may have formed with eccentricities similar to planets beyond the reach of tides. This result suggests many planets' orbits are still circularizing. They considered a range of Q'_p and Q'_* values and determined the set that produced an initial e distribution for tidally evolving orbits that best matched that of nontidally evolved planets. Figure 12.2 shows some of the results of this experiment. We see that both e- and a-values for close-in planets were significantly larger in the past. The same study also showed that if $Q'_p \sim 10^{6.5}$ and $Q'_* \sim 10^{5.5}$ (plausible Q'-values [11, 14, 16, 65]), the planets could have begun with an e distribution statistically similar from the other planets. However, it is important to note that for any individual system, the Q' values could be different, and/or additional companions could affect the evolution. Nonetheless, this study corroborates the tidal hypothesis as the explanation for the observed low eccentricities of close-in planets.

At this point, we may only speculate whether nonzero eccentricities are a result of slow damping of the primordial eccentricity or pumping by an external agent. One planet that has attracted considerable attention in this regard is GJ 436 b

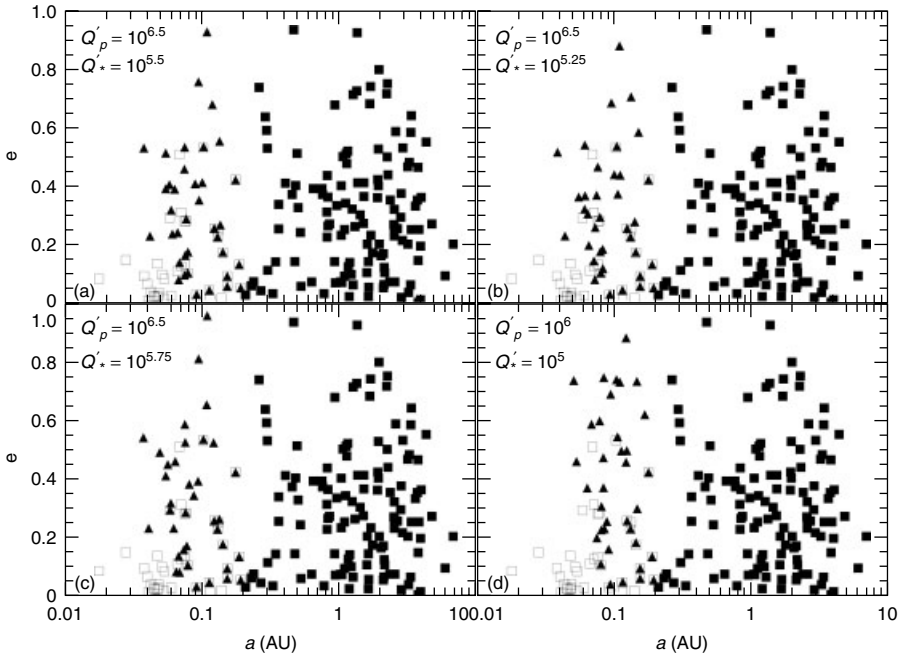


Figure 12.2 (a)–(d) Distributions of orbital elements before and after tidal evolution. Each panel assumes all planets and stars share the same value of Q'_p and Q'_*, respectively. Squares (filled and open) represent the currently observed orbital elements, with the open squares (with $a < 0.2$ AU) being candidates for significant tidal evolution. Triangles show the initial orbital elements determined by integrating the equations of tidal evolution backward in time to the formation of the planet. The best match is (a) [1].

with $e = 0.15$. Although no additional companion has been detected [69, 70], many studies have proposed existence of an additional planet [66, 71]. The detection of additional planets by considering tidal and planet–planet interactions would, of course, be exciting, but given the complexity of the interactions, poor constraints on stellar ages [72], and the difficulty in measuring small eccentricities [73–75], caution must used when interpreting these data.

On the other hand, [76] suggested that close-in planets with large eccentricities simply have larger Q'_p-values than planets with small eccentricities. They found that Q'_p-values spanning the range from 10^5 to $>10^9$, with each value chosen on a case-by-case basis, can reproduce the observed distribution. Q'_p-values larger than 10^6 may be inconsistent with constraints on Q'_p-values for some gaseous planets in our solar system [11, 14, 16, 65]. However, Q'_p-values of exoplanets could be different from planets in our solar system.

Recent studies have also investigated how a-values drop with time. This migration can continue long after e has become very small, and many close-in planets we observe today will be destroyed in the next few billion years. Levrard *et al.* [27] considered the known transiting planets (because of their more precisely determined masses and radii) and showed that all but one is doomed to merge with its host star. They used a constant-time-lag model and found that the timescale for merging ranged from tens of millions of years to tens of billions of years. Jackson *et al.* [28] considered the observed distribution of ages and a, and found a paucity of close-in planets around older stars. They showed that the constant-phase-lag model, with plausible Q' values, predicted that many close-in planets would be destroyed before they became too old. These results may invalidate previous suggestions that some process(es) brought close-in planets nearer their current orbits and then stopped them there [18, 24]. Instead, close-in planets have probably been migrating inward since formation.

For example, Jackson *et al.* [28] found that the recently discovered planet CoRoT-7 b may be destroyed in less than a few billion years, as illustrated in Figure 12.3. They assumed $e = 0$ and hence the evolution is due entirely to the tide raised on the star. As transiting planets tend to be close to their star, it is likely that *C*oRoT and *K*epler will find many such ''doomed'' planets.

What do these results imply for the origins of close-in planets? The agreement between the distributions of initial eccentricities for close-in planets and for planets unaffected by tides suggests that planet–planet scattering probably played a role in the early histories of both classes of planets. In fact, the first studies of the dynamical histories of exoplanetary systems suggested that close-in planets actually resulted from planet–planet scattering [2, 77]; see also Chapter 11. Future models of planet–planet scattering should consider whether it can produce a distribution of eccentricities for close-in planets that resembles the distribution for the other planets. Likewise, the results of studies of the origins of close-in planets through gas-disk migration (Chapter 9, [18, 19]) should be revisited in light of the evidence for tidal destruction, which reproduces the distribution of a-values for observed close-in exoplanets, perhaps without requiring a mechanism to halt the inward migration of the planets [28].

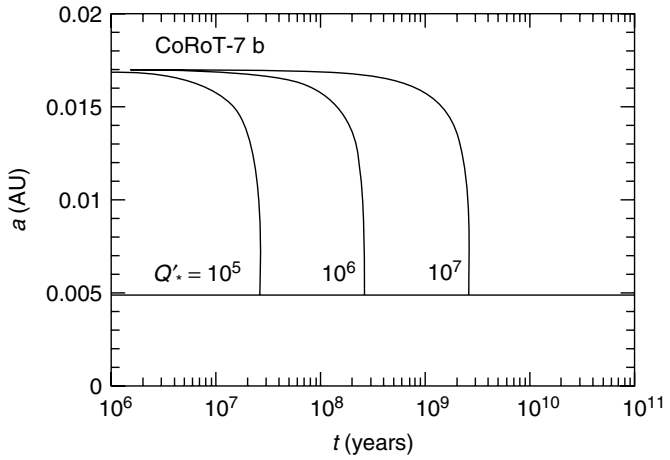


Figure 12.3 Tidal evolution of the 11 $M_\oplus$ CoRoT-7 b planet for a range of Q'_*-values. The horizontal line represents the host star's Roche limit [28].

12.3.2 Tidal Heating

During tidal evolution, orbital energy is transformed into heat inside the planet. The rate of tidal heating in the second-order, constant-phase-lag model is given by Eq. (12.3). Planets on circular orbits will experience no tidal heating, but for even small values of e, the heating rate can be very large ($>10^{18}$ W). Such heating can significantly affect a planet's interior structure, which may be detectable in radius measurements during transit.

Indeed, such heating may help to explain the observed mass–radius relationship of transiting exoplanets, see Figure 1.4 in Chapter 1. Early on it was noted that some transiting planets seemed to have unexpectedly large radii compared to theoretical structure models [22, 78–84], such as HD 209458 b, the first-discovered transiting planet [85]. Although the current estimated value of tidal heating rate can, in some cases, provide a significant amount of energy [86], in most cases it cannot.

However, past tidal heating may be responsible [25], see Figure 12.4. Since it may take a long time for their interiors to cool [78], gas giants will not necessarily respond instantaneously to internal heating. In fact evolutionary models suggest it may take a billion years or more for planets to relax to an equilibrium radius after they form [83]. Thus it may take equally long for the effects of tidal heating to become apparent in the planet's size.

Recent studies have shown that past tidal heating can indeed explain the large radii of some transiting exoplanets [87, 88]. Thus, the inflated radius of a planet may reflect not just current but also past heating, although some planets still remain anomalously large. Tidal effects may therefore be able to explain this long-standing conundrum.

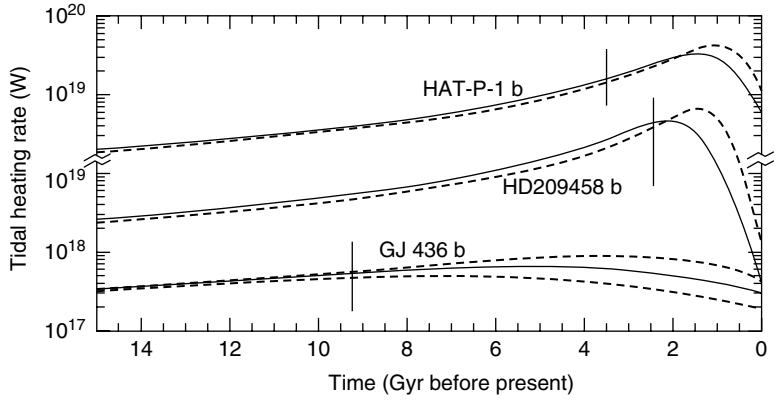


Figure 12.4 The tidal heating rates for planets HAT-P-1 b, HD 209458 b, and GJ 436 b as a function of time. The present time ($t = 0$) is at the right, and the scale indicates the time before the present. Vertical lines indicate the best estimate for the age of the system. For each planet, the vertical scale that corresponds to each curve is the scale intersected by that curve. The solid curve for each planet is based on the current nominal eccentricity value, while the dashed lines assume the maximum and minimum heating consistent with observational uncertainty in the orbital elements. For HAT-P-1 b and HD 209458 b, observations could not exclude a current eccentricity of zero, so the lower bound on heating rates is formally zero. Hence in those cases only one dashed line is shown, representing the upper limit [25].

12.4 Tidal Effects on Rocky Exoplanets

As observational technology and techniques improve, Earth-like planets will be found orbiting stars other than the Sun. Such a discovery would be a milestone in the search for extraterrestrial life. Current observational techniques are most sensitive to Earth-mass planets if they orbit close to inactive, low-mass stars [89–93]. For these planets, tides will play an important role in determining the planets' orbital and physical properties.

Rocky planets respond differently to tides than do gaseous planets. Whereas gaseous planets tend to have Q' values near 10^6, rocky planets like the Earth probably have Q'-values closer to 500 [21, 94]. This difference means that terrestrial-like exoplanets may experience more rapid orbital evolution and a larger surface flux of tidal heating than the gas giants presented in the previous section. As these bodies are largely hypothetical, we focus on how tides may affect habitability of these planets in the following subsections.

12.4.1 Orbital Evolution and Habitability

A planet's orbital eccentricity and semimajor axis help determine the insolation the planet receives during its orbit, so changes in either can change the insolation

[95]. Thus as tides damp e and a over billions of years, the insolation may slowly increase, even causing a planet that starts out in the habitable zone (HZ) to leave it.

Recent work found that rocky planets may spend a short amount of time in the HZ, depending on orbital and physical properties [61], as illustrated in Figure 12.5. In this example, all the planets initially receive the same amount of insolation over an orbit, slightly more than the present-day Earth. In this case, planets with initial eccentricities less than about 0.5 spend more than 10 billion years in the HZ. However, for larger values of initial eccentricity, the amount of time a planet spends in the HZ can be significantly less, perhaps limiting planetary habitability. This result shows that when rocky planets are discovered, tidal evolution models may help determine whether such planets were ever in the HZ, or whether they will be leaving HZ in the future. Barnes *et al.* [61] considered the past evolution of GJ 581 c, and found that it was likely *n*ot habitable any time in the past 10 gigayears.

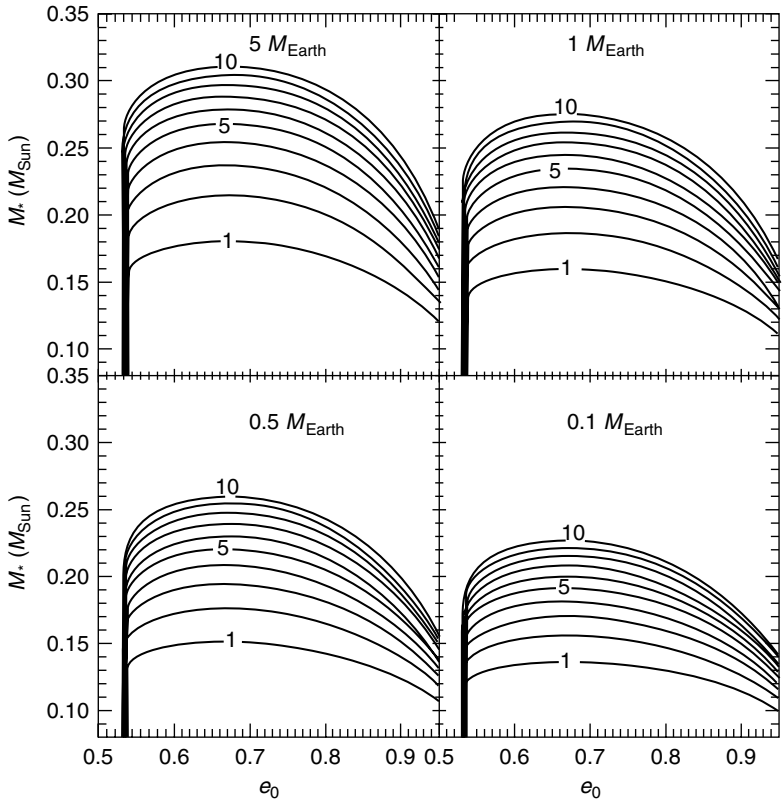


Figure 12.5 Contours show the amount of time in billions of years that a planet spends in the HZ of its host star, for a range of stellar and planetary masses (in solar M_{Sun} and Earth masses M_{Earth}, respectively) and initial eccentricities [61]. All planets begin on an orbit that receives the same orbit-averaged flux, which is slightly more than that the present-day Earth receives. The planet's mass is given at the top of each panel.

12.4.2 Effects of Tidal Heating

As with gaseous exoplanets, tidal evolution of a rocky exoplanet's orbit may be accompanied by dissipation of substantial energy within the planet. Too much heating may drive rapid, global resurfacing through volcanism, as on Jupiter's moon Io [96–99], perhaps precluding the development of life. A moderate amount of heating may enhance a planet's habitability by driving more moderate volcanism and plate tectonics, which could stabilize the climate [100]. Therefore tidal heating could play an important role in determining a planet's potential habitability.

Figures 12.6 and 12.7 show the heating rates as functions of planetary mass M_p, stellar mass M_* and eccentricity e [26]. In this example, the planets were placed in the center of the "eccentric HZ" [61], which assumed that the orbit-averaged flux determines surface temperature [95]. (This choice means the HZ moves to larger distances at high e–hence the decrease in tidal heating for planets on those orbits.) Although the thermal evolution and geophysical processes of a rocky planet may be complicated and cannot be predicted simply on the basis of a heat-flow estimate,

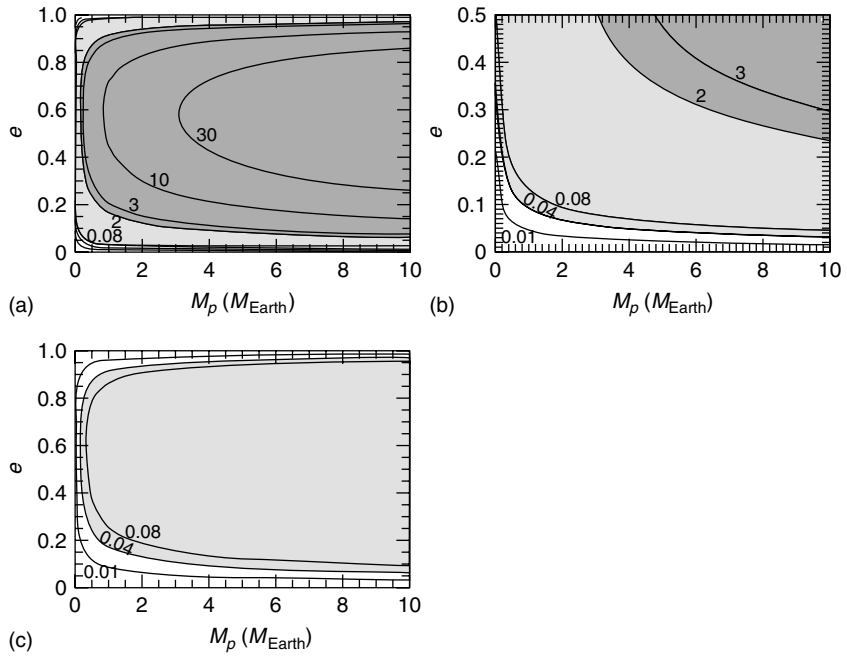


Figure 12.6 Contours of heating flux in watts per meter squared for a range of M_p and e for planets in the center of the "eccentric HZ" [61, 95]. For panels (a) and (b), $M_* = 0.1 M_{Sun}$, and $M_* = 0.2$ (M_{Sun}) for panel (c). Panel (b) shows a close-up of the bottom left corner of panel (a). White areas indicate $h < 0.04$ W/m^2. Light gray areas indicate 0.04 W/m^2 $< h <$ 2 W/m^2 (perhaps adequate for plate tectonics), while dark gray indicates $h >$ 2 W/m^2 (more than Io) [26].

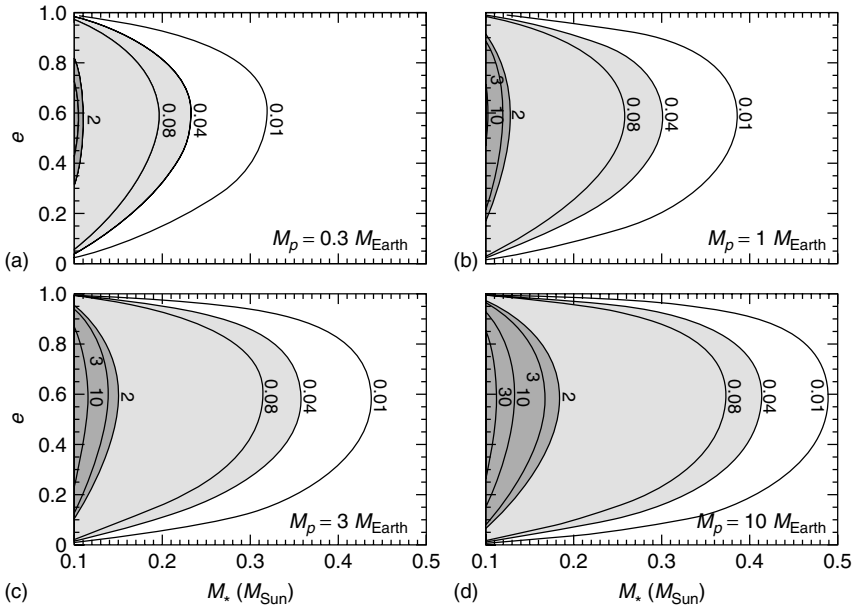


Figure 12.7 Same as Figure 12.6 but for a range of M_* and e. For panels (a)–(d), $M_p = 0.3$, 1, 3, and 10 Earth masses [26], respectively.

such estimates may be compared to heating rates for bodies in our solar system to gauge the range of geophysical processes that might result.

For example, Io emits 2–3 W/m^2 of internally generated heat from tides [97, 99, 101, 102], resulting in planetwide volcanism that resurfaces the planet in less than one megayear [99]. Such rapid resurfacing might challenge the development of life, even if the other ingredients required for life were present. The dark gray regions in Figures 12.6 and 12.7 show the wide range of masses and orbits that may lead to more heating than Io's [26], assuming $Q'_p = 500$.

As another example, Earth emits ~0.08 W/m^2 of internal (radiogenic and primordial) heat [103], which is apparently adequate for plate tectonics. Considerations of Mars' geophysical history suggest a heating rate of 0.04 W/m^2 is a minimum required for tectonic activity and volcanism in a rocky planet [104]. Previous studies suggested that plate tectonics plays a significant role in keeping the Earth's surface habitable by buffering the amount of carbon dioxide in the atmosphere [100]. Consequently, some process may be required to sequester greenhouse gases in the planet's mantle to prevent a runaway greenhouse effect [62], and sufficient internal heating may be required. The white regions in Figures 12.6 and 12.7 show the range of masses and orbits for which tidal heating below 0.01 w/m^2, perhaps insufficient to drive tectonic activity [26]. The light gray regions represent the masses and orbits for which tidal heating lies between 0.04 w/m^2 and 2 w/m^2.

We can compare the range of orbits expected to contain habitable planets from considerations only of insolation (based on [95, 105]) to the range for which tidal heating may allow a planet to be habitable ($0.04\ \mathrm{W/m^2} \leq h \leq 2\ \mathrm{W/m^2}$), as in Figure 12.8 [106]. For planets to the left of the yellow curves, the planets have more heating than Io, to the right, they may require radiogenic and/or primordial heating to drive plate tectonics. We see that for a wide range of orbits and masses the two HZs overlap significantly. These results suggest a wide range of geophysical states for planets in the habitable zones of M stars.

In some cases, close-in exoplanets may lose their atmospheres from impactors [107] or coronal mass ejections (CMEs) [90, 92, 108]. Tidal heating may help some extrasolar terrestrial planets remain habitable despite this phenomenon. It could melt the interior, producing a liquid core that leads to a magnetic field, which could shield the planet's atmosphere from the ionized portion of the CME. Alternatively, persistent outgassing could replenish the atmosphere maintaining a biosphere. Ongoing research is investigating both the mantle outgassing and thermal evolution of large, rocky exoplanets [109] and can shed light on these intriguing possibilities.

Tidal heating could also enhance the habitability of ice-rich planets, by maintaining a liquid ocean just below the surface [110], even if the exposed surface is frozen. Such a planet might be similar to Europa, where $0.19\ \mathrm{W/m^2}$ of tidal heating is sufficient to maintain water under a thin ice shell [111] that could be suitable for life [43, 112, 113]. Whether planets as volatile rich as Europa could actually form around an M star is speculative [114, 115], but if they do form, the same processes that potentially make Europa habitable may also occur on these planets.

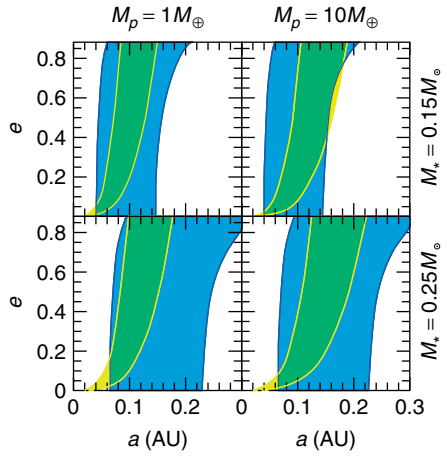


Figure 12.8 Habitable limits from consideration of insolation (blue region) and of tidal heating (yellow region). The blue region is the HZ from [61], the yellow the tidal heating limits from [26] (the left edge corresponds to Io, the right edge to the minimum for plate tectonics). Where the two regions overlap (green), planets may have the right tidal heating and stellar insolation for Earth-like habitability.

As with gaseous exoplanets, tidal heating rates can vary dramatically over time. Figure 12.9 shows the evolution of e, a, and h [26] for different systems. The heating rates for planets in the HZs of larger mass stars do not change that much over 15 billion years. If a planet requires tidal heating to remain habitable, the duration of its habitability will depend on its mass and original orbit in a complex way.

Thus, tidal heating of rocky exoplanets in the HZs of M stars may be large and long lived, may drive important geophysical activity, and may be crucial for habitability. Geophysical and atmospheric modeling of rocky exoplanets will be required to elucidate the properties of rocky exoplanets as they are discovered in the next few years, and tidal heating may play a key role.

Tidal heating may be very important for close-in rocky planets about larger stars, too. In these cases, tidal heating rates may be very large, even for very small

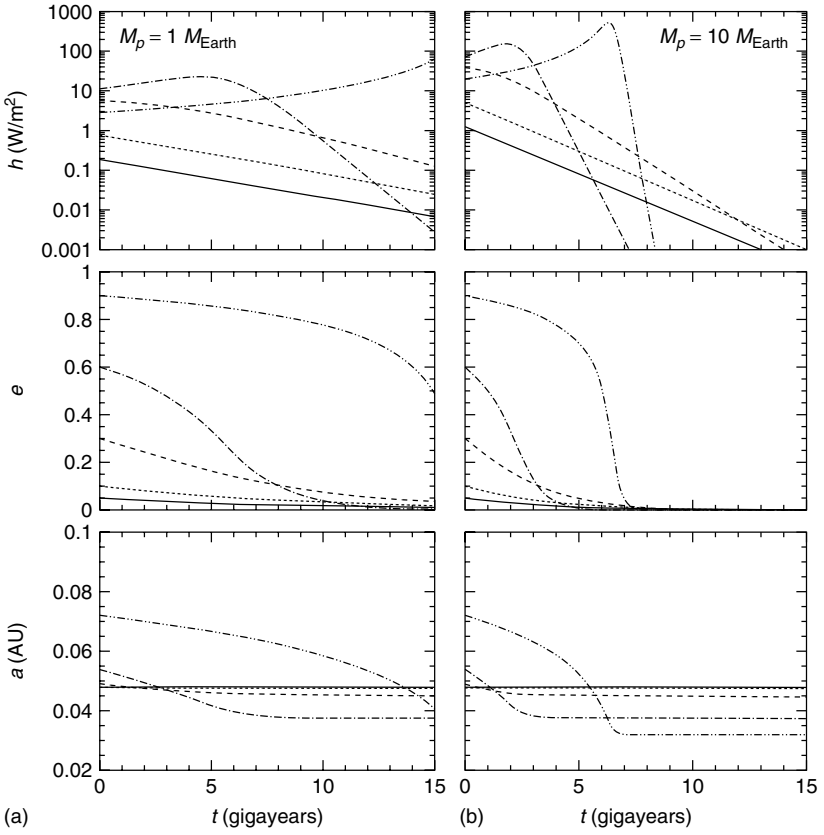


Figure 12.9 Tidal heating histories and tidal evolution of e and a for planets of various masses with various values of initial eccentricity e_0 orbiting a star with mass 0.1 solar masses. Parts (a) and (b) illustrate histories for planets with masses of 1, and 10 M_{Earth}, respectively. These planets are assumed to be in the HZ of a 0.1 M_{Sun} star. Solid lines correspond to $e_0 = 0.05$, dotted lines to 0.1, dashed to 0.3, dashed-dotted to 0.6, and dashed-dot-dot-dot to 0.9 [26].

eccentricities. With *K*epler and *C*oRoT scouring the sky for such planets, we may be on the precipice of discovering many such "super-Ios", of which HD 40307 b and GJ 581 e could be the first [59, 106].

12.4.3 Rotation and Habitability

A rocky planet in a circular orbit close in the HZ of its host star may become synchronously rotating within a few million years [62, 116]. Although such a rotation rate might cause one side of the planet to be too hot and the other too cold for life, recent calculations suggest that atmospheric transport of heat might efficiently transport heat between the planet's two faces, allowing such a planet to remain habitable [117–119].

However, if a rocky exoplanet is on an eccentric orbit, the rotation rate might be very different from synchronous. In the case of an eccentric orbit, the planet's rotation rate is determined by a balance between the tidal torque and the torque arising from any permanent asymmetries in the planet's mass distribution, such as for Mercury [31, 32, 48, 61, 116]. Therefore planets may either be rotating nonsynchronously or be trapped in a spin–orbit resonance, and the effects on a planet's surface temperature may be complicated, requiring consideration of atmospheric transport of heat as well as the rotation state.

As a planet's orbital eccentricity evolves, the tidally locked rotation rate can change, as shown in Figure 12.10 [61]. The planet's rotation state may be further complicated by the exchange of angular momentum between the planet's solid body and atmosphere [35], as may be important for Venus [120] or Saturn's moon Titan [121, 122]. To determine the rotational state of a close-in, rocky exoplanet, sophisticated modeling and detailed observational data may be required [123].

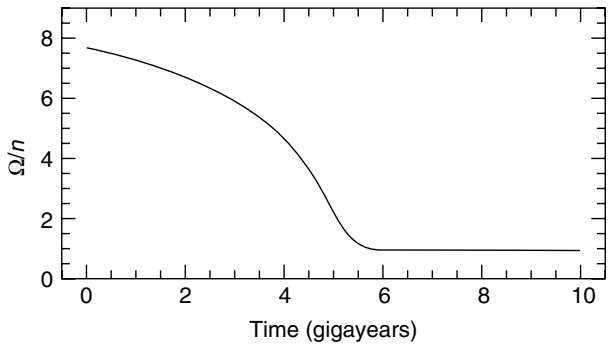


Figure 12.10 Evolution of the equilibrium rotation rate (here given as Ω) in terms of its orbital frequency for a planet undergoing tidal evolution [61]. The planet began on an $e = 0.84$ orbit, which damped to 0 at 6.5 gigayears.

12.5 Conclusions

Tides play a variety of important roles in extrasolar systems. They shape the distribution of orbital elements of close-in planets, and tidal models may provide clues that enable us to find additional exoplanets. Dissipation of tidal energy within close-in planets may warm the planets, or slow their long-term cooling. The same tidal processes may help determine whether rocky exoplanets are suitable for life. The models presented here are relatively simple, and we look forward to an improved understanding of tidal processes, most importantly the frequency dependence(s).

The future is bright for tidal studies of exoplanets. Observational techniques tend to be most sensitive to planets close to their host stars, so we may expect many more tidally evolving planets to be discovered soon by missions such as *K*epler. Moreover, small planets, with the potential to be rocky (and hence habitable), are most easily detected in regions in which tidal forces can be significant. The first planet with a minimum mass less than $10M_{\oplus}$ discovered in the habitable zone, GJ 581 d, is probably tidally locked and tidally heated at a rate of 0.01 W/m^2 [106]. This discovery highlights the importance of a firm understanding of tidal physics, and its subsequent effects on planetary habitability. It may be that the first-discovered inhabited exoplanet also experiences strong tidal forcing.

References

1 Jackson, B., Greenberg, R., and Barnes, R. (**2008**) Tidal evolution of close-in extrasolar planets. *The Astrophysical Journal*, **678**, 1396.

2 Rasio, F., Tout, C., Lubow, S., and Livio, M. (**1996**) Dynamical instabilities and the formation of extrasolar planetary systems. *The Astrophysical Journal*, **470**, 1187.

3 Zahn, J.P. (**1977**) Tidal friction in close binary stars. *Astronomy and Astrophysics*, **57**, 383.

4 Duquennoy, A. and Mayor, M. (**1991**) Multiplicity among solar-type stars in the solar neighbourhood. II - Distribution of the orbital elements in an unbiased sample. *Astronomy and Astrophysics*, **248**, 485.

5 Mathieu, R. (**1994**) Pre-main-sequence binary stars. *Annual Review of Astronomy and Astrophysics*, **32**, 465.

6 Mathieu, R. (**2004**) The Rotation of Low-Mass Pre-Main-Sequence Stars. *Proceedings of IAU Symposium No. 215, 113, Cancun, Yucatan, Mexico*.

7 Mazeh, T. (**2008**) Observational evidence for tidal interaction in close binary systems. *EAS Publications Series*, **29**, 1.

8 Zahn, J.P. (**2008**) Tidal dissipation in binary systems. *EAS Publications Series*, **29**, 97.

9 Peale, S. and Cassen, P. (**1978**) Contribution of tidal dissipation to lunar thermal history. *Icarus*, **36**, 245.

10 Lambeck, K. (**1978**) Tidal dissipation in the oceans, in *Tidal Friction and the Earth's Rotation, Proceedings of a Workshop*, Springer-Verlag, Berlin, p. 95.

11 Yoder, C. and Peale, S. (**1981**) The tides of Io. *Icarus*, **47**, 1.

12 Sagan, C. and Dermott, S. (**1982**) The tide in the seas of Titan. *Nature*, **300**, 731.

13 McKinnon, W. (**1984**) On the origin of triton and Pluto. *Nature*, **311**, 355.

14 Tittemore, W. and Wisdom, J. (**1989**) Tidal evolution of the Uranian satellites. II - an explanation of the anomalously high orbital inclination of Miranda. *Icarus*, **78**, 63.

15 Showman, A. and Malhotra, R. (**1997**) Tidal evolution into the Laplace resonance and the resurfacing of Ganymede. *Icarus*, **127**, 93.

16 Aksnes, K. and Franklin, F. (**2001**) Secular acceleration of Io derived from mutual satellite events. *The Astronomical Journal*, **122**, 2734.

17 Peale, S. and Lee, M. (**2002**) A primordial origin of the Laplace relation among the Galilean satellites. *Science*, **298**, 593.

18 Lin, D., Bodenheimer, P., and Richardson, D. (**1996**) Orbital migration of the planetary companion of 51 Pegasi to its present location. *Nature*, **380**, 606.

19 Trilling, D., Benz, W., Guillot, T. *et al.* (**1998**) Orbital evolution and migration of giant planets: modeling extrasolar planets. *The Astrophysical Journal*, **500**, 428.

20 Bodenheimer, P., Lin, D., and Mardling, R. (**2001**) On the tidal inflation of short-period extrasolar planets. *The Astrophysical Journal*, **548**, 466.

21 Mardling, R. and Lin, D. (**2002**) Calculating the tidal, spin, and dynamical evolution of extrasolar planetary systems. *The Astrophysical Journal*, **573**, 829.

22 Bodenheimer, P., Laughlin, G., and Lin, D. (**2003**) On the radii of extrasolar giant planets. *The Astrophysical Journal*, **592**, 555.

23 Adam, F. and Laughlin, G. (**2006**) Long-term evolution of close planets including the effects of secular interactions. *The Astrophysical Journal*, **649**, 992.

24 Ford, E. and Rasio, F. (**2006**) On the relation between hot Jupiters and the Roche Limit. *The Astrophysical Journal*, **638**, 45.

25 Jackson, B., Greenberg, R., and Barnes, R. (**2008**) Tidal heating of extrasolar planets. *The Astrophysical Journal*, **681**, 1631.

26 Jackson, B., Barnes, R., and Greenberg, R. (**2008**) Tidal heating of terrestrial extrasolar planets and implications for their habitability. *Monthly Notices of the Royal Astronomical Society*, **391**, 237.

27 Levrard, B., Winisdoerffer, C., and Chabrier, G. (**2009**) Falling transiting extrasolar giant planets. *The Astrophysical Journal*, **692**, 9.

28 Jackson, B., Barnes, R., and Greenberg, R. (**2009**) Observational evidence for tidal destruction of exoplanets. *The Astrophysical Journal*, **698**, 779.

29 Jeffreys, H. (**1961**) The effect of tidal friction on eccentricity and inclination. *Monthly Notices of the Royal Astronomical Society*, **122**, 339.

30 Goldreich, P. (**1963**) On the eccentricity of satellite orbits in the solar system. *Monthly Notices of the Royal Astronomical Society*, **126**, 257.

31 Ferraz-Mello, S., Rodríguez, A., and Hussmann, H. (**2008**) Tidal friction in close-in satellites and exoplanets: The Darwin theory re-visited. *Celestial Mechanics and Dynamical. Astronomy*, **101**, 171.

32 Murray, C. and Dermott, S. (**1999**) *Solar System Dynamics*, Princeton University Press.

33 Hurford, T. (**2005**) *Tides and tidal stress: applications to Europa. Ph.D. Dissertation*, The University of Arizona.

34 Dorbovolskis, A. and Ingersoll, A. (**1980**) Atmospheric tides and the rotation of Venus. I. Tidal theory and the balance of torques. *Icarus*, **41**, 1.

35 Correia, A., Levrard, B., and Laskar, J. (**2008**) On the equilibrium rotation of Earth-like extra-solar planets. *Astronomy and Astrophysics*, **488**, 63.

36 Goldreich, P. and Peale, S. (**1966**) Spin-orbit coupling in the solar system. *The Astronomical Journal*, **71**, 425.

37 Macdonald, G. (**1964**) Tidal friction. *Reviews of Geophysics and Space Physics*, **2**, 467–541.

38 Efroimsky, M. and Lainey, V. (**2007**) Physics of bodily tides in terrestrial planets and the appropriate scales of dynamical evolution. *Journal of Geophysical Research*, **112**, E12003.

39 Greenberg, R. (**2009**) Frequency dependence of tidal Q. *The Astrophysical Journal*, **698**, 42.

40 Goldreich, P. and Soter, S. (**1966**) *Icarus*, **5**, 375.

41 Ogilvie, G. and Lin, D. (**2004**) Tidal dissipation in rotating giant planets. *The Astrophysical Journal*, **610**, 477.

42 Ogilvie, G. and Lin, D. (**2007**) Tidal dissipation in rotating solar-type stars. *The Astrophysical Journal*, **661**, 1180.

43 Greenberg, R. (**2005**) *Europa-The Ocean Moon: Search for an Alien Biosphere*, Springer-Praxis, Berlin.

44 Dickey, J. *et al.* (**1994**) Lunar laser ranging: a continuing legacy of the apollo program. *Science*, **265**, 482.

45 Hut, P. (**1981**) Tidal evolution in close binary systems. *Astronomy and Astrophysics*, **99**, 126.

46 Eggleton, P., Kiseleva, L., and Hut, P. (**1998**) The equilibrium tide model for tidal friction. *The Astrophysical Journal*, **499**, 853.

47 Zahn, J. (**1989**) Tidal evolution of close binary stars. I - revisiting the theory of the equilibrium tide. *Astronomy and Astrophysics*, **220**, 112.

48 Goldreich, P. and Peale, S. (**1966**) Spin-orbit coupling in the solar system. *Astronomical Journal*, **71**, 425.

49 Goldreich, P. and Nicholson, P. (**1977**) Turbulent viscosity and Jupiter's tidal Q. *Icarus*, **430**, 301.

50 Ivanov, P. and Papaloizou, J. (**2007**) Dynamic tides in rotating objects: orbital circularization of extrasolar planets for realistic planet models. *Astronomy and Astrophysics*, **476**, 121.

51 Ioannou, P. and Lindzen, R. (**1993**) Gravitational tides in the outer planets. II - Interior calculations and estimation of the tidal dissipation factor. *The Astrophysical Journal*, **406**, 266.

52 Ioannou, P. and Lindzen, R. (**1994**) Gravitational tides on Jupiter. 3: atmospheric response and mean flow acceleration. *The Astrophysical Journal*, **424**, 1005.

53 Ohta, Y., Taruya, A., and Suto, Y. (**2005**) The Rossiter-McLaughlin effect and analytic radial velocity curves for transiting extrasolar planetary systems. *The Astrophysical Journal*, **622**, 1118.

54 Winn, J. *et al.* (**2005**) Measurement of spin-orbit alignment in an extrasolar planetary system. *The Astrophysical Journal*, **631**, 1215.

55 Gaudi, S. and Winn, J. (**2007**) Prospects for the characterization and confirmation of transiting exoplanets via the Rossiter-McLaughlin Effect. *The Astrophysical Journal*, **655**, 550.

56 Cochran, W., Redfield, S., Endl, M., and Cochran, A. (**2008**) The spin-orbit alignment of the HD 17156 transiting eccentric planetary system. *The Astrophysical Journal*, **683**, 59.

57 Trilling, D. (**2000**) Tidal constraints on the masses of extrasolar planets. *The Astrophysical Journal*, **537**, 61.

58 Barnes, S. (**2001**) An assessment of the rotation rates of the host stars of extrasolar planets. *The Astrophysical Journal*, **561**, 1095.

59 Barnes, R., Jackson, B., Raymond, S.N. *et al.* (**2009**) The HD 40307 planetary system: super-earths or mini-Neptunes? *The Astrophysical Journal*, **695**, 1006.

60 Showman, A. and Guillot, T. (**1996**) Atmospheric circulation and tides of ''51 Pegasus b-like'' planets. *Astronomy and Astrophysics*, **385**, 1187.

61 Barnes, R., Raymond, S., Jackson, B., and Greenberg, R. (**2008**) Tides and the evolution of planetary habitability. *Astrobiology*, **8**, 557.

62 Kasting, J., Whitmire, D., and Reynolds, R. (**1993**) Habitable zones around main sequence stars. *Icarus*, **101**, 108.

63 Wisdom, J. (**2008**) Tidal dissipation at arbitrary eccentricity and obliquity. *Icarus*, **193**, 637.

64 Mardling, R. (**2007**) Long-term tidal evolution of short-period planets with companions. *Monthly Notices of the Royal Astronomical Society*, **382**, 1768.

65 Zhang, K., and Hamilton, D. (**2007**) Orbital resonances in the inner neptunian system. I. The 2:1 Proteus Larissa mean-motion resonance. *Icarus*, **188**, 386.

66 Batygin, K. *et al.* (**2009**) A quasi-stationary solution to Gliese 436b's eccentricity. *The Astrophysical Journal*, **699**, 23.

67 Takeda, G. and Rasio, F. (**2005**) High orbital eccentricities of extrasolar planets induced by the Kozai mechanism. *The Astrophysical Journal*, **627**, 1001.

68 Nagasawa, M., Ida, S., and Bessho, T. (**2008**) Formation of hot planets by a combination of planet scattering, tidal circularization, and the Kozai mechanism. *The Astrophysical Journal*, **678**, 498.

69 Alonso, R. *et al.* (**2008**) Limits to the planet candidate GJ 436c. *Astronomy and Astrophysics*, **482**, 21.

70 Bean, J. and Seifahrt, A. (**2008**) Observational consequences of the recently proposed Super-Earth orbiting GJ 436. *Astronomy and Astrophysics*, **487**, 25.

71 Ribas, I., Font-Ribera, A., and Beaulieu, J.P. (**2008**) A $\sim 5M_{\oplus}$ super-earth orbiting GJ 436? The power of near-grazing transits. *The Astrophysical Journal*, **677**, 59.

72 Takeda, G., Ford, E.B., Sills, A. *et al.* (**2007**) Structure and evolution of nearby stars with planets. II. Physical properties of ~1000 cool stars from the SPOCS catalog. *The Astrophysical Journal Supplement Series*, **168**, 297.

73 Lucy, L. and Sweeney, M. (**1971**) Spectroscopic binaries with circular orbits. *Astronomical Journal*, **76**, 544.

74 Knutson, H. *et al.* (**2007**) A map of the day-night contrast of the extrasolar planet HD 189733b. *Nature*, **447**, 183.

75 Shen, Y. and Turner, E. (**2008**) On the eccentricity distribution of exoplanets from radial velocity surveys. *The Astrophysical Journal*, **685**, 553.

76 Matsumura, S., Takeda, G., and Rasio, F. (**2008**) On the origins of eccentric close-in planets. *The Astrophysical Journal*, **686**, 29.

77 Weidenschilling, S. and Marzari, F. (**1996**) Gravitational scattering as a possible origin for giant planets at small stellar distances. *Nature*, **384**, 619.

78 Guillot, T. (**1999**) Interior of giant planets inside and outside the solar system. *Science*, **286**, 72.

79 Burrows, A. *et al.* (**2000**) On the radii of close-in giant planets. *The Astrophysical Journal*, **534**, 97.

80 Baraffe, I. *et al.* (**2003**) Evolutionary models for cool brown dwarfs and extrasolar giant planets. The case of HD 209458. *Astronomy and Astrophysics*, **402**, 701.

81 Guillot, T. (**2005**) The interiors of giant planets: models and outstanding questions. *Annual Review of Earth and Planetary Sciences*, **33**, 493.

82 Laughlin, G., Marcy, G., Vogt, S. *et al.* (**2005**) A comparison of observationally determined radii with theoretical radius predictions for short-period transiting extrasolar planets. *The Astrophysical Journal*, **629**, 121.

83 Burrows, A., Hubeny, I., Budaj, J., and Hubbard, W. (**2007**) Possible solutions to the radius anomalies of transiting giant planets. *The Astrophysical Journal*, **661**, 502.

84 Fortney, J., Marley, M., and Barnes, J. (**2007**) Planetary radii across five orders of magnitude in mass and stellar insolation: application to transits. *The Astrophysical Journal*, **659**, 1661.

85 Charbonneau, D., Brown, T., Latham, D., and Mayor, M. (**2000**) Detection of planetary transits across a sun-like star. *The Astrophysical Journal*, **529**, 45.

86 Liu, X., Burrows, A., and Ibgui, L. (**2008**) Theoretical radii of extrasolar giant planets: the cases of TrES-4, XO-3b, and HAT-P-1b. *The Astrophysical Journal*, **687**, 1191.

87 Ibgui, L. and Burrows, A. (**2009**) Coupled evolution with tides of the radius and orbit of transiting giant planets: general results. *The Astrophysical Journal*, **700**, 1921–1932.

88 Miller, N., Fortney, J.J., and Jackson, B. (**2009**) Inflating and deflating hot jupiters: coupled tidal and thermal evolution of known transiting planets. *The Astrophysical Journal*, **702**, 1413–1427.

89 Khodachenko, M. *et al.* (**2007**) Coronal Mass Ejection (CME) activity of low mass M stars as an important factor for the habitability of terrestrial exoplanets. I. CME impact on expected magnetospheres of earth-like exoplanets in close-in habitable zones. *Astrobiology*, **7**, 167.

90 Lammer, H. *et al.* (**2007**) Coronal Mass Ejection (CME) activity of low mass M stars as an important factor for the habitability of terrestrial exoplanets. II. CME-induced ion pick up of earth-like exoplanets in close-in habitable zones. *Astrobiology*, **7**, 185.

91 Scalo, J. *et al.* (**2007**) M stars as targets for terrestrial exoplanet searches and biosignature detection. *Astrobiology*, **7**, 85.

92 Tarter, J. *et al.* (**2007**) A reappraisal of the habitability of planets around M dwarf stars. *Astrobiology*, **7**, 30.

93 Nutzman, P. and Charbonneau, D. (**2008**) Design considerations for a ground-based transit search for habitable planets orbiting M dwarfs. *Publications of the Astronomical Society of the Pacific*, **120**, 317.

94 Yoder, C.F. and Ahrens, T. (**1995**) Astrometric and geodetic properties of earth and the solar system, in *Global Earth Physics: A Handbook of Physical Constants*, (ed. T. Ahrens), American Geophysical Union, Washington, DC, p. 1.

95 Williams, D. and Pollard, D. (**2002**) Earth-like worlds on eccentric orbits: excursions beyond the habitable zone. *International Journal of Astrobiology*, **1**, 61.

96 Morabito, L., Synnott, S., Kupferman, P., and Collins, S. (**1979**) Discovery of currently active extraterrestrial volcanism. *Science*, **204**, 972.

97 Peale, S., Cassen, P., and Reynolds, R. (**1979**) Melting of Io by tidal dissipation. *Science*, **203**, 892.

98 McEwen, A. *et al.* (**1998**) Active volcanism on Io as seen by Galileo SSI. *Icarus*, **135**, 181.

99 McEwen, A., Keszthelyi, L., Lopes, R. *et al.* (**2004**) *The lithosphere and surface of Io, from Jupiter: The Planet, Satellites, and Magnetosphere*, Cambridge University Press, Cambridge, p. 307.

100 Walker, J., Hays, P., and Kasting, J. (**1981**) A negative feedback mechanism for the long-term stabilization of the earth's surface temperature. *Journal of Geophysical Research*, **86**, 9776.

101 Matson, D., Ransford, G., and Johnson, T. (**1981**) Heat flow from Io. *Journal of Geophysical Research*, **86**, 1664.

102 Veeder, G. *et al.* (**1994**) Io's heat flow from infrared radiometry: 1983–1993. *Journal of Geophysical Research*, **99**, 17095.

103 Davies, G. (**1999**) *Dynamic Earth*, Cambridge University Press.

104 Williams, D., Kasting, J., and Wade, R. (**1997**) Habitable moons around extrasolar giant planets. *Nature*, **385**, 234.

105 Selsis, F., Kasting, J., Levrard, B. *et al.* (**2007**) Habitable planets around the star Gliese 581? *Astronomy and Astrophysics*, **476**, 1373.

106 Barnes, R., Jackson, B., Greenberg, R., and Raymond, S.N. (**2009**) Tidal limits to planetary habitability. *The Astrophysical Journal*, **700**, L30.

107 Melosh, H.J. and Vickery, A. (**1989**) Impact erosion of the primordial atmosphere of Mars. *Nature*, **338**, 487.

108 Hawley, S., Gizis, J., and Reid, I. (**1996**) The palomar/MSU nearby star spectroscopic survey.II.The southern M dwarfs and investigation of magnetic activity. *The Astronomical Journal*, **112**, 2799.

109 Papuc, A. and Davies, G. (**2008**) The internal activity and thermal evolution of Earth-like planets. *Icarus*, **195**, 447.

110 Vance, S., Harnmeijer, J., Kimura, J. *et al.* (**2007**) Hydrothermal systems in small ocean planets. *Astrobiology*, **7**, 987.

111 O'Brien, D., Geissler, P., and Greenberg, R. (**2002**) A melt-through model for chaos formation on Europa. *Icarus*, **156**, 152.

112 Reynolds, R., Squyres, S., Colburn, D., and McKay, C. (**1983**) On the habitability of Europa. *Icarus*, **56**, 246.

113 Chyba, C. (**2000**) Energy for microbial life on Europa. *Nature*, **403**, 381.

114 Lissauer, J. (**2007**) Planets formed in habitable zones of M dwarf stars probably are deficient in volatiles. *The Astrophysical Journal*, **660**, 149.

115 Raymond, S., Scalo, J., and Meadows, V. (**2007**) High-resolution simulations of the final assembly of Earth-Like planets. 2. Water delivery and planetary habitability. *The Astrophysical Journal*, **669**, 606.

116 Peale, S. (**1977**) *Rotation Histories of the Natural Satellites, from Planetary Satellites*, Unversity of AZ Press, Tucson, p. 112.

117 Haberle, R., McKay, C., Tyler, D., and Reynolds, R. (**1996**) *Can Synchronously Rotating Planets Support An Atmosphere? from Circumstellar Habitable Zones, Proceedings of the First International Conference*, Travis House Publications, Menlo Park, p. 29.

118 Joshi, M., Haberle, R., and Reynolds, R. (**1997**) Simulations of the atmospheres of synchronously rotating terrestrial planets orbiting m dwarfs: conditions for atmospheric collapse and the implications for habitability. *Icarus*, **129**, 450.

119 Joshi, M. (**2003**) Climate model studies of synchronously rotating, *Astrobiology*, **3**, 415–427.

120 Pechmann, J. and Ingersoll, A. (**1984**) Thermal tides in the atmosphere of Venus - Comparison of model results with observations. *Journal of the Atmospheric Sciences*, **41**, 3290.

121 Tokano, T., Neubauer, F. (**2005**) Wind-induced seasonal angular momentum exchange at Titan's surface and its influence on Titan's length-of-day. *Geophysical Research Letters*, **32**, L24203.

122 Lorenz, R. *et al.* (**2008**) Titan's rotation reveals an internal ocean and changing zonal winds. *Science*, **319**, 1649.

123 Pallé, E. *et al.* (**2008**) Identifying the rotation rate and the presence of dynamic weather on extrasolar Earth-like planets from photometric observations. *The Astrophysical Journal*, **676**, 1319.

Further Reading

Agol, E., Steffen, J., Sari, R., and Clarkson, W. (**2005**) On detecting terrestrial planets with timing of giant planet transits. *Monthly Notices of the Royal Astronomical Society*, **359**, 567.

Bakos, G. *et al.* (**2007**) *The Astrophysical Journal*, **656**, 552.

Baraffe, I., Chabrier, G., and Barman, T. (**2003**) Structure and evolution of super-Earth to super-Jupiter exoplanets. I. Heavy element enrichment in the interior. *Astronomy and Astrophysics*, **482**, 315.

Barnes, J. (**2007**) *Publications of the Astronomical Society of the Pacific*, **119**, 986.

Beckwith, S., Sargent, A., Chini, R., and Guesten, R. (**1990**) A survey for circumstellar disks around young stellar objects. *Astronomical Journal*, **99**, 924.

Bodenheimer, P. and Pollack, J. (**1986**) Calculations of the accretion and evolution of giant planets: the effects of solid cores. *Icarus*, **67**, 391.

Bryden, G. *et al.* (**1999**) Tidally induced gap formation in protostellar disks: gap clearing and suppression of protoplanetary growth. *The Astrophysical Journal*, **514**, 344.

Carone, L. and Pätzold, M. (**2007**) Constraints on the tidal dissipation factor of a main sequence star: the case of OGLE-TR-56b. *Planetary and Space Science*, **55**, 643.

Charbonneau, D., Brown, T., Burrows, A., and Laughlin, G. (**2007**) *Protostars and Planets, V*, University of AZ Press, Tucson, 701.

Chatterjee, S., Ford, E., Matsumura, S., and Rasio, F. (**2008**) Dynamical outcomes of planet–planet scattering. *The Astrophysical Journal*, **686**, 580.

Deming, D., Harrington, J., Laughlin, G. *et al.* (**2007**) Spitzer transit and secondary eclipse photometry of GJ 436b. *The Astrophysical Journal*, **667**, 199.

Dobbs-Dixon, I., Lin, D., and Mardling, R. (**2004**) Spin-orbit evolution of short-period planets. *The Astrophysical Journal*, **610**, 464.

Ford, E., Kozinsky, B., and Rasio, F. (**2000**) Secular evolution of hierarchical triple star systems. *The Astrophysical Journal*, **535**, 385.

Ford, E. and Rasio, F. (**2008**) Origins of eccentric extrasolar planets: testing the planet–planet scattering model. *The Astrophysical Journal*, **686**, 621.

Geissler, P. *et al.* (**1999**) Galileo imaging of atmospheric emissions from Io. *Science*, **285**, 870.

Gillon, M. *et al.* (**2008**) Detection of transits of the nearby hot Neptune GJ 436 b. *Astronomy and Astrophysics*, **472**, 13.

Giuppone, C., Tadeu dos Santos, M., Beaugé, C. *et al.* (**2009**) Detectability and error estimation in orbital fits of resonant extrasolar planets. *The Astrophysical Journal*, **699**, 1321.

Guillot, T. *et al.* (**2006**) A correlation between the heavy element content of transiting extrasolar planets and the metallicity of their parent stars. *Astronomy and Astrophysics*, **453**, 21.

Holman, M. and Murray, N. (**2005**) The use of transit timing to detect terrestrial-mass extrasolar planets. *Science*, **307**, 1288.

Ida, S. and Lin, D. (**2004**) Toward a deterministic model of planetary formation. I. A desert in the mass and semimajor axis distributions of extrasolar planets. *The Astrophysical Journal*, **604**, 388.

Jurić, M. and Tremaine, S. (**2008**) Dynamical origin of extrasolar planet eccentricity distribution. *The Astrophysical Journal*, **686**, 603.

Lunine, J. (**2005**) *Astrobiology: A Multi-Disciplinary Approach*, Benjamin Cummings.

Mandushev, G. *et al.* (**2007**) *The Astrophysical Journal*, **667**, 195.

Maness, H., Marcy, G., Ford, E. *et al.* (**2007**) The M dwarf GJ 436 and its Neptune-Mass planet. *Publications of the Astronomical Society of the Pacific*, **119**, 90.

Mardling, R. (**2008**) On the long-term tidal evolution of GJ 436b in the presence of a resonant companion, eprint arXiv:0805.1928.

Mardling, R. and Lin, D. (**2004**) On the survival of short-period terrestrial planets. *The Astrophysical Journal*, **614**, 955.

Massoroti, A. (**2008**) *The Astronomical Journal*, **135**, 2287.

Mayor, M. and Queloz, D. (**1995**) A Jupiter-mass companion to a solar-type star. *Nature*, **378**, 355.

Mignard, P. (**1979**) *Moon and the Planets*, **20**, 301.

Mignard, P. (**1980**) *Moon and the Planets*, **23**, 185.

Nelson, R., Papaloizou, J., Masset, F., and Kley, W. (**2000**) The migration and growth of protoplanets in protostellar discs. *Monthly Notices of the Royal Astronomical Society*, **318**, 18.

O'Neill, C. and Lenardic, A. (**2007**) Conditions for the onset of plate tectonics on terrestrial planets and moons. *Geophysical Research Letters*, **34**, L19204.

Pont, F. *et al.* (**2007**) *Astronomy and Astrophysics*, **465**, 1069.

Press, W. *et al.* (**1996**) *Numerical Recipes in C: The Art of Scientific Computing*, Press Syndicate of University of Cambridge, New York.

Rasio, F. and Ford, E. (**1996**) Dynamical instabilities and the formation of extrasolar planetary systems. *Nature*, **180**, 307.

Sandquist, E., Dokter, J., Lin, D., and Mardling, R. (**2002**) *The Astrophysical Journal*, **572**, 1012.

Showman, A., Guillot, T. (**2002**) *Astronomy and Astrophysics*, **385**, 166.

Sozzetti, A. *et al.* (**2007**) *The Astrophysical Journal*, **664**, 1190.

Valencia, D., O'Connell, R., and Sasselov, D. (**2007**) Inevitability of plate tectonics on super-earths. *The Astrophysical Journal*, **670**, 45.

Ward, W. (**1997**) Protoplanet migration by Nebula tides. *Icarus*, **126**, 261.

Wetherill, G. (**1980**) Formation of the terrestrial planets. *Annual Review of Astronomy and Astrophysics*, **18**, 77.

Wolszczan, A. and Frail, D. (**1992**) A planetary system around the millisecond pulsar PSR1257 + 12. *Nature*, **355**, 145.

Index

Formation and Evolution of Exoplanets. Edited by Rory Barnes

ISBN: 978-3-527-40896-2

d