ISSI Scientific Report SR-006

Planetary Systems and Planets in Systems

S. Udry, W. Benz, and R. von Steiger

INTERNATIONAL SPACE SCIENCE INSTITUTE

SR-006
December 2006

Planetary Systems and Planets in Systems

Editors

Stephane Udry
Observatoire de Genève, Sauverny, Switzerland

Willy Benz
Physikalisches Institut, University of Bern, Switzerland

Rudolf von Steiger
International Space Science Institute, Bern, Switzerland

Dedicated to Michel Mayor on the occasion of his 60th birthday

Cover: Artists impression of a transiting extrasolar planet. From obswww.unige.ch/~udry/planet/planet.html

The International Space Science Institute is a Foundation under Swiss law. It is funded by the European Space Agency, the Swiss Confederation, the Swiss National Science Foundation, and the University of Bern.

Published for: The International Space Science Institute
Hallerstrasse 6, CH-3012 Bern, Switzerland

by: ESA Publications Division
Keplerlaan 1, 2200 AG Noordwijk, The Netherlands

Publication Manager: Bruce Battrick
Cover Design: N.N.

ISSN: ZZZZ-YYYX
Price: XX Euros

Contents

Foreword

Some text...

Bern, December 2006 Stephane Udry, Willy Benz, Rudolf von Steiger

Michel Mayor, Science with Enthusiasm and Humour

Michel Mayor joined the Geneva Observatory as an assistant in 1965. Throughout his career he has always shown the juvenile curiosity that we know for interesting questions in astrophysics and science in general. All kinds of questions, from scientific to social and political ones, with him are a matter for lively discussions, not without a good touch of humour. These interests go along with a visionary intuition and perseverance, necessary to make great achievements. Scientific questions do not prevent Michel to like climbing on mountains and travelling around the world. I must also add that Michel has the pleasant conviviality which allowed people to have nice lunches together, full of laughters and jokes, at the Observatory most weekdays during forty years.

Back in 1969, in the first volume of the newborn journal Astronomy & Astrophysics, there was an article by Michel Mayor and Louis Martinet. The subject was already "motions", not yet motions of planets, but of nearby stars in relation with ages and stellar properties. These researches in galactic dynamics culminated in a study with Laurent Vigroux in 1981 on the effects of matter infalling on the Galaxy with zero angular momentum. Indeed, the studies of motions of astronomical objects is the basic "red line" of all Michel's researches. Like all astronomers who wanted to have some chance to stay at Geneva Observatory, Michel had to perform some mandatory photometric missions in Haute-Provence. Around 1968, everything which might more or less resemble to student contests would be regarded with suspicion at Observatory. I remember when Michel, with new familial responsibilities, once asked the administrator what his wife and children would receive as indemnity in case of decease. The administrator, who was nevertheless a nice man, went furious through the institute shouting "These assistants! They want all advantages for themselves".

At the end of the seventies, Michel Mayor had developed, in collaboration with André Baranne, the Coravel spectrograph with the basic idea to study the stellar dynamics in the Galaxy by means of radial velocities. Coravel turned out to be most useful for many original studies, like stellar rotation of low mass stars with Willy Benz and later with José de Medeiros for evolved stars, Cepheids and binaries with Gilbert Burki and later with Antoine Duquennoy and globular clusters with Georges Meylan. In order to improve the accuracy of radial velocities, the Elodie spectrograph was developed in the early nineties. The main objective of the observing programme was brown dwarfs and planetary companions. Today, it may look unbelievable that only ten years ago the subject of planets was not a topical one, even worse it was even a "dusty" and often not well perceived topic. Thus, when they realized that the accuracy was perhaps enough to try to find planets,

Michel Mayor and Didier Queloz in an application for observing time only speak about brown dwarfs and possible "sub-stellar" objects in the title of their proposal.

The historical discovery of the first extra-solar planet around 51 Pegasi in 1995 by Michel and Didier immediately had a great impact in the astronomical community and by the public as well. As often in Switzerland, the discovery became well recognized locally by the media when it was subsequently confirmed by Americans. The finding of extra-solar planets brought a kind of "Planetary Revolution" in Astronomy and placed the planets among the most topical scientific researches and the most popular ones. Exobiology and the search for extraterrestrial life have then become much more respectable. This deep change of minds, which leads to both new theoretical developments and major space programmes, is most impressive. All started out or was strongly boosted by the discovery of 51 Pegasi.

The golden age of discoveries continues today, new extrasolar planets are announced regularly which has stirred international collaboration and sometimes also competition. Michel and his team, with the efficient collaboration of Stéphane Udry, Nuno Santos, Dominique Naef and others not only go on at a leading place in planetary search, but also develops great projects for the future. The new spectrograph Harps, achieved in cooperation with ESO thanks to the competence of Francesco Pepe, reaches the amazing accuracy of about 30 cm/s. It opens up fascinating perspectives in planetary search and asteroseismology.

Since about forty years, Michel and myself have been colleagues and nevertheless good friends, going through various steps in the responsibilities at the Geneva Observatory, as assistants under the chairmanship of Marcel Golay, research associates, professors and even more recently as successive directors of the Geneva Observatory. In the name of all colleagues and friends in Geneva and abroad, I address to Michel our warm congratulations and best wishes for the future. Our compliments are also for Françoise Mayor, who among other qualities has a great sense of the real relative importance of things in life.

At the age of sixty, Michel, keep on going like you did over all this time, as a friendly colleague, keeping good health and scientific enthusiasm!

Geneva Observatory, October 20, 2003 André Maeder

Part I

Giant Planets: General View

Chemical Properties of Stars with Giant Planets

N. C. Santos and M. Mayor
Observatoire de Genève, CH-1290 Sauverny, Switzerland

G. Israelian
Instituto de Astrofísica de Canarias, E-38200 La Laguna, Tenerife, Spain

Abstract. In this paper we review the current situation regarding the study of the metallicity-giant planet connection. We show that planet-host stars are more metal-rich than average field dwarfs, a trend that remains after the discovery of now more than about 100 planets. Current results also point out that the planetary frequency is rising as a function of the [Fe/H]. Furthermore, the source of this high metallicity is shown to have most probably an "original" source, although there are also some evidences for stellar "pollution". These facts seem to be telling us that the metallicity is playing a very important role in the formation of the now discovered giant planets. We further compare the orbital properties (minimum mass, period and eccentricity) of planet hosts with their chemical abundances. Some slight trends are seen and discussed.

1. Introduction

The last few years have been very prolific regarding the research for new planetary systems. Following the discovery of the planetary companion to the solar-type star 51 Peg (Mayor and Queloz, 1995), more than 100 extra-solar giant planets are known today[1]. These results opened a wide range of questions regarding the understanding of the mechanisms of planetary formation.

To find a solution for the many problems risen, we need observational constraints. These can come, for example, from the analysis of the orbital parameters of the known planets, like the distribution of planetary masses, eccentricities, or orbital periods (see contribution from Mayor et al. in this volume). But further evidences are coming from the study of the planet host stars themselves. Precise spectroscopic studies have revealed that stars with planets seem to be particularly metal-rich when compared with "single" field dwarfs (Gonzalez, 1997; Santos et al., 2001). Furthermore, the frequency of planets seems to be a strong function of [Fe/H] (Santos et al., 2001). These facts, that were shown not to result from any sampling bias, are most probably telling us that the metallicity plays a key role in the formation of a giant planet, or at least of the giant planets like the ones we are finding now.

The nature of the metallicity-giant planet connection has been strongly debated in the literature. Some authors have suggested that the high metal content of the

[1] See `http://obswww.unige.ch/exoplanets`. Before 1995 only planets around a pulsar had been detected (Wolszczan and Frail, 1992); these are probably second generation planets. Also, the previously discovered radial-velocity companion around HD 114762 (Latham et al., 1989) has a minimum mass above $10 M_{\rm Jup}$, and is likely a brown-dwarf

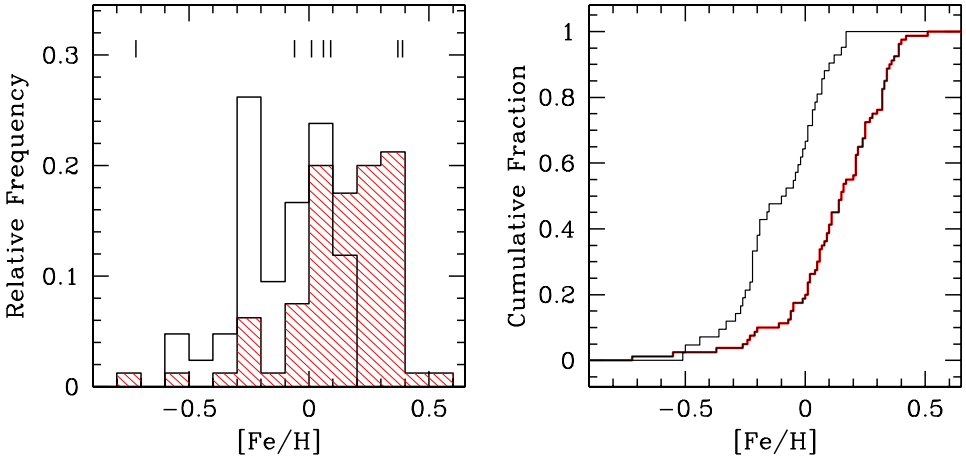

Figure 1. *Left*: metallicity distribution for stars with planets (hashed histogram) compared with the same distribution for the field dwarfs (empty histogram). The vertical lines represent stars with brown dwarf candidate companions. *Right*: The cumulative functions of both samples. A Kolmogorov-Smirnov test shows the probability for the two populations being part of the same sample is around 10^{-7}. From Santos et al. (2002b).

planet host stars may have an external origin: it results from the addition of metal-rich (hydrogen poor) material into the convective envelope of the star, a process that could result from the planetary formation process itself (Gonzalez, 1998). Evidences for the infall of planetary material have in fact been found for a few planet host stars (Israelian et al., 2001; Laws and Gonzalez, 2001; Israelian et al., 2002), although not necessarily able to change considerably their overall metal content. In fact, most evidences today suggest that the metallicity "excess" as a whole has a "primordial" origin (Pinsonneault et al., 2001; Santos et al., 2001; Santos et al., 2003), and thus that the metal content of the cloud giving birth to the star and planetary system is indeed a key parameter to form a giant planet.

In this article we will make a review of the current situation regarding the results obtained in the context of the study of the metallicity of planet-host stars. Most of the results presented here were published in Santos et al. (2001) and Santos et al. (2003), for which we refer for more details.

2. The Metallicity of Planet Host Stars

In Figure 1 we plot the metallicity distribution for all the stars known to have companions with minimum masses lower than $\sim 18 M_{\rm Jup}$ (hashed histogram) when compared to the same distribution for a volume limited sample of stars with no (known) planetary companions (open histogram) – see Santos et al. (2001). For the planet host stars, most of the metallicity values ([Fe/H]) were taken from Santos et al. (2001) and Santos et al. (2003). It is important to remember that the metallicity

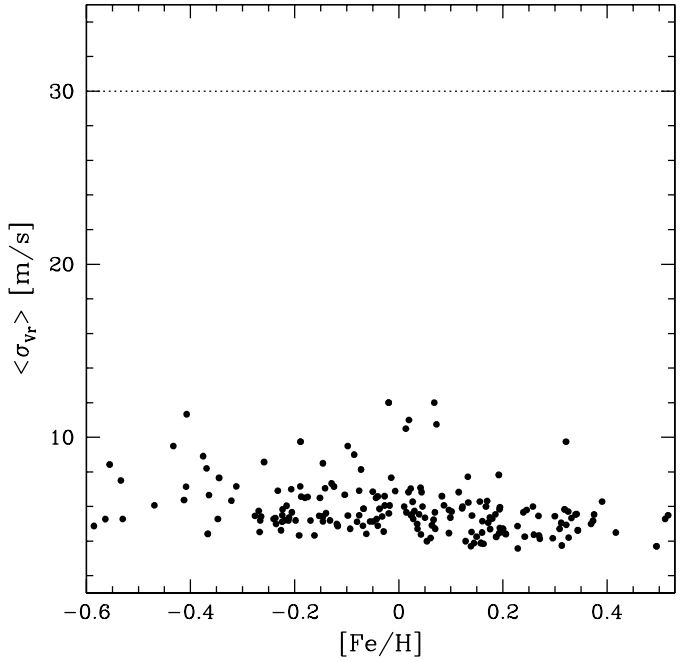

Figure 2. Plot of the mean-photon noise error for the CORALIE measurements of stars having magnitude V between 6 and 7, as a function of the metallicity. Most planet host stars present radial-velocity variations with an amplitude larger than $30\,\mathrm{m\,s^{-1}}$ (dotted line). From Santos et al. (2002).

for the two samples of stars plotted in Figure 1 was derived using exactly the same method, and are thus both in the same scale.

As is clearly visible in the figures, stars with planetary companions are more metal-rich (in average) than field dwarfs. The average metallicity difference between the two samples is about 0.22 dex, and the probability that the two distributions belong to the same sample is of the order of 10^{-7}.

In the figure, stars having planetary companions with masses higher than $10M_{\mathrm{Jup}}$ are denoted by the vertical lines. Given the still low number, it is impossible to do any statistical study of this group. However, up to now there does not seem to exist any clear difference between the two populations.

2.1. No Metallicity Bias

Particular concern has been shown by the community regarding the fact that a higher metallicity will imply that the spectral lines are better defined. This could mean that the final precision in radial-velocity could be better for the more "metallic" objects, leading thus to an increasing detection rate as a function of increasing [Fe/H].

In Figure 2 we plot the mean photon noise error for stars with different [Fe/H] having V magnitudes between 6 and 7. As we can see, there is no particularly important trend in the data. The very slight tendency (metal-rich stars have, in average, measurements with only about 1–2 m s^{-1} better precision than metal poor stars) is definitely not able to induce the strong tendency seen in the [Fe/H] distribution of planet host stars, specially when we compare it with the usual velocity amplitude induced by the known planetary companions (a few tens of meters-per-second).

In fact, in the CORALIE survey we always set the exposure times in order to have approximately the same photon-noise error. This also seems to be the case concerning the Lick/Keck planet search programs (D. Fischer, private communication).

2.2. The Source of the Metallicity

Two different interpretations have been given to the [Fe/H] "excess" observed for stars with planets. One suggests that the high metal content is the result of the accretion of planets and/or planetary material into the star (e.g. Gonzalez, 1998). Another simply states that the planetary formation mechanism is dependent on the metallicity of the proto-planetary disk: according to the "traditional" view, a gas giant planet is formed by runaway accretion of gas by a $\sim$10 earth masses planetesimal (Pollack et al., 2001). The higher the metallicity (and thus the number of dust particles) the faster a planetesimal can grow, and the higher the probability of forming a giant planet before the gas in the disk dissipates.

There are multiple ways of deciding between the two scenarios. Probably the most clear and strong argument is based on stellar internal structure, in particular on the fact that material falling into a star's surface would induce a different increase in [Fe/H] depending on the stellar mass, i.e. on the depth of its convective envelope (where mixing can occur). However, the data shows no such trend (Figure 3) – (Santos et al., 2001; Santos et al., 2003). In particular, a quick look at the plot indicates that the upper envelope of the points is quite constant. A similar conclusion was also recently taken by Pinsonneault et al. (2001) that showed that even non-standard models of convection and diffusion cannot explain the lack of a trend and sustain "pollution" as the source of the high-[Fe/H].

Together with the similarity of [Fe/H] abundances for dwarf and sub-giant planet host stars[2] (Santos et al., 2001; Santos et al., 2003), the facts presented here strongly suggest a "primordial origin" to the high-metal content of stars with giant planets. This result implies that the metallicity is a key parameter controlling planet formation and evolution, and may have enormous implications on theoretical models.

[2] These latter are supposed to have diluted already any strong metallicity excess present, since their convective envelopes have deepened as a result of stellar evolution

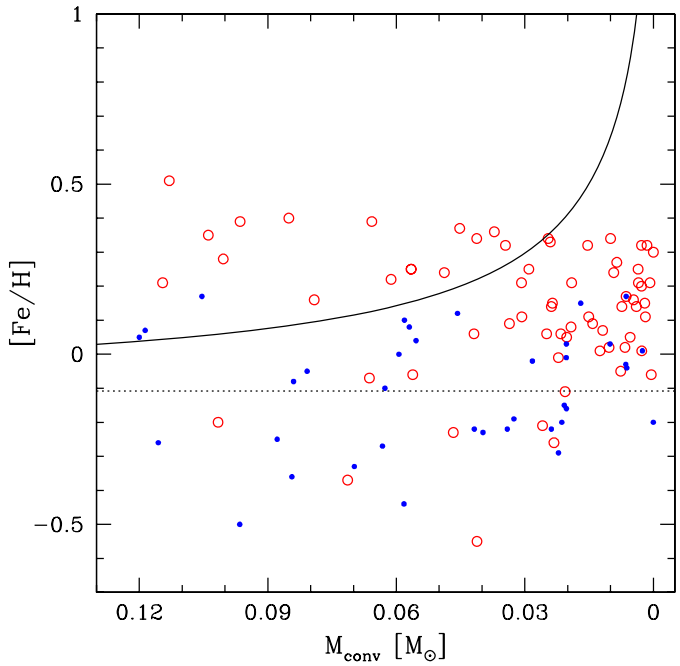

Figure 3. Metallicity vs. convective envelope mass for stars with planets (open symbols) and field dwarfs (points). The [Fe/H]=constant line represents the mean [Fe/H] for the non-planet hosts stars of Figure 1. The curved line represents the result of adding 8 earth masses of iron to the convective envelope of stars having an initial metallicity equal to the non-planet hosts mean [Fe/H]. The resulting trend has no relation with the distribution of the stars with planets. From Santos et al. (2002b).

2.3. The Frequency of Planetary Formation

In Figure 4 (left panel) we plot the [Fe/H] distribution for a large volume limited sample of stars included in the CORALIE survey (Udry et al., 2000)[3] when compared with the same distribution for planet host stars that belong to the CORALIE planet-search sample itself.

The knowledge of the metallicity distribution for stars in the solar neighborhood (and included in the CORALIE sample) permits us to determine the percentage of planet host stars per metallicity bin. This analysis is presented in Figure 4 (right panel). As we can perfectly see, the probability of finding a planet host is a strong function of its metallicity. For example, about 7% of the stars in the CORALIE sample having metallicity between 0.3 and 0.4 dex have been discovered to harbor a planet. On the other hand, less than 1% of the stars having solar metallicity seem to have a planet. This result is thus probably telling us that the probability of forming a giant planet, or at least a planet of the kind we are finding now, depends strongly on the metallicity of the gas that gave origin to the star and planetary system.

[3] The metallicities for this sample were computed from a precise calibration of the CORALIE Cross-Correlation Function (Santos et al., 2002a)

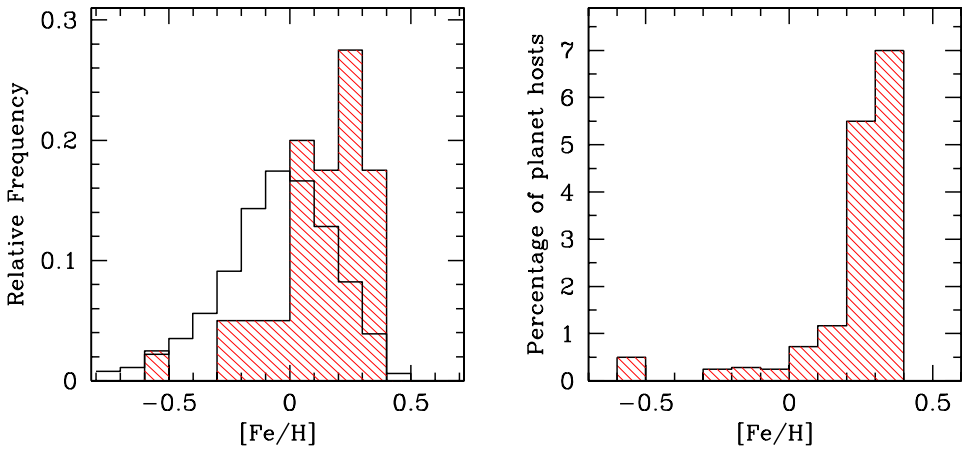

Figure 4. Left: metallicity distribution of stars with planets making part of the CORALIE planet search sample (shaded histogram) compared with the same distribution for the about 1000 non binary stars in the CORALIE volume-limited sample (see text for more details). *Right*: the result of correcting the planet hosts distribution to take into account the sampling effects. The vertical axis represents the percentage of planet hosts with respect to the total CORALIE sample. From Santos et al. (2002b).

This result seems to be telling us that the planetary formation efficiency (at least of the kind of planets we are finding now) is a strong function of the metallicity.

It is important to discuss the implications of this result on the planetary formation scenarios. Boss (2002) has shown that in the core accretion scenario we can expect the efficiency of planetary formation to be strongly dependent on the metallicity, contrarily of what could be expected from a mechanism of disk instability. The results presented here, suggesting that the probability of forming a planet is strongly dependent on the metallicity of the host star, can thus be seen as an argument for the former (traditional) core accretion scenario (Pollack et al., 2001). We note that here we are talking about a probabilistic effect: the fact that the metallicity enhances the probability of forming a planet does not mean one cannot for a planet in a lower metallicity environment. This might be due to the fact that other important and unknown parameters, like the proto-planetary disk mass and lifetime, do control the efficiency of planetary formation as well. Furthermore, these results do not exclude that an overlap might exist between the two planetary formation scenarios.

3. Metallicity and Orbital Parameters

Some hints of trends between the metallicity of the host stars and the orbital parameters of their planets have been discussed already in the literature. The usually low number of points involved in the statistics did not permit, however, to extract

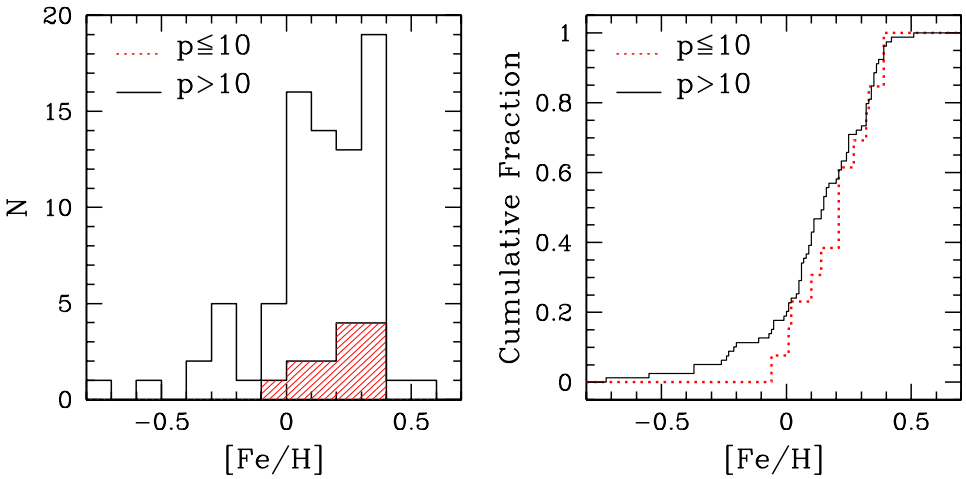

Figure 5. Left: [Fe/H] distributions for planets orbiting with periods shorter and higher than 10 days (the hashed and open bars, respectively). *Right*: cumulative functions of both distributions. A Kolmogorov-Smirnov test gives a probability of ∼0.3 that both samples make part of the same distribution. From Santos et al. (2002b).

any major conclusions (Santos et al., 2001). Today, we dispose of about 80 high-precision and uniform metallicity determinations for planet host stars (Santos et al., 2003), a sample that enables us to look for possible trends in [Fe/H] with planetary mass, semi-major axis or period, and eccentricity with a higher degree of confidence. Let us then see what is the current situation.

3.1. Orbital Period

Gonzalez (1998) and Queloz et al. (2000) have shown evidences that stars with short-period planets (i.e. small semi-major axes) may be particularly metal-rich, even amongst the planetary hosts. The number of planets that were known by that time was, however, not enough to arrive at a definitive conclusion.

In Figure 5 we can see the metallicity distributions for two populations of extra-solar planet hosts. The hashed histogram corresponds to stars hosting planets with orbital periods shorter than 10 days, while the open bars represent stars hosting planets with periods longer than this value. As we can see from the plot, shorter period planets do show a tendency to orbit mostly metal-rich stars. However, the figure also tells us that this tendency looks like a result of the lower number of points corresponding to the short period systems. This lack of difference is in fact supported by a Kolmogorov-Smirnov test, that gives a probability ∼30% that both populations belong to the same sample. In other words, there is no statistically significant difference between the two groups of stars.

It is worth noticing that changing the limits in the orbital period does not bring any new clear trend.

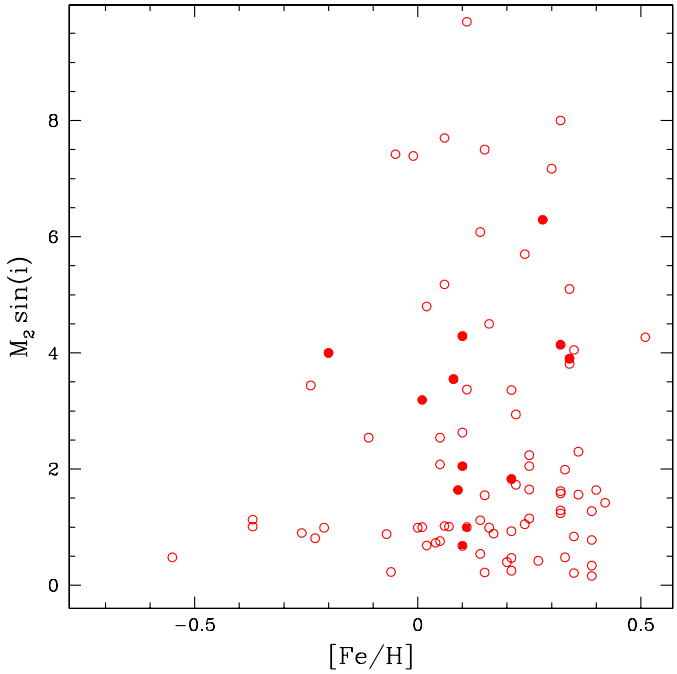

Figure 6. Minimum mass of the planetary companions to solar-type stars as a function of the metallicity of the host stars (only companions with minimum masses lower than $10M_{\rm Jup}$). Filled symbols represent companions around stars that are members of binary systems. From Santos et al. (2002b).

3.2. PLANETARY MASS

In Figure 6 we plot the minimum mass for the "planetary" companions having masses lower than $10M_{\rm Jup}$ as a function of the metallicity. A simple look at the plot tells us that there is a clear lack of high mass companions to metal-poor stars. There seems to be an upper envelope for the mass of a planet as a function of the stellar metallicity. Although not statistically significant, this result deserves some further discussion

As discussed in Udry et al. (2002) and Santos et al. (2003), this result can be seen as an evidence that to form a massive planet (at least up to a mass of $\sim 10M_{\rm Jup}$) we need more metal-rich disks. This tendency might have to do with the time needed to build the planet seeds before the disk dissipates (if you form more rapidly the cores, you have more time to accrete gas around), or with the mass of the "cores" that will later on accrete gas to form a giant planet (the higher the dust density of the disk, the higher will be the core mass you might be able to form (Kokubo and Ida, 2002); does this mass influence the final mass of the giant planet?).

In the plot, the filled symbols represent planets in multiple stellar systems. We do not see any special trend for these particular cases.

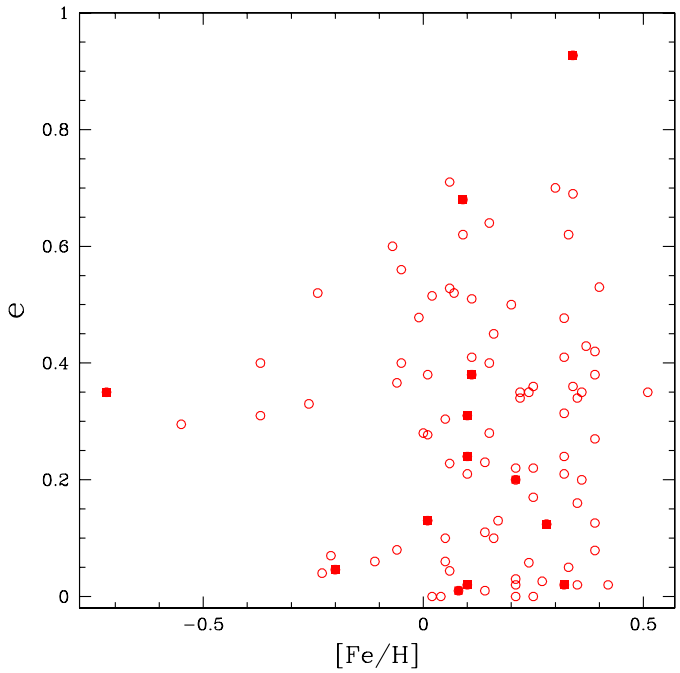

Figure 7. Same as Figure 6 but for the eccentricity.

3.3. ECCENTRICITY

Another interesting result can be found if we plot the eccentricity as a function of the metallicity. As we can see from Figure 7, two main holes in the figure seem to exist. First, the more eccentric planets seem to orbit only stars with metallicity higher or comparable to solar. Furthermore, on the opposite side of the eccentricity distribution, there seems also to be a lack of low eccentricity planets around metal-poor stars.

However, none of this results seems to be statistically significant. We should thus look at this as a trend that needs more data to be confirmed or infirmed.

4. Signs of Planetary Accretion

Although the results presented above seem to rule out pollution as the key parameter inducing the high metallicity of planet hosts stars, some evidences of infall of planetary material have been discussed in the literature (Gonzalez, 1998; Laws and Gonzalez, 2001; Gratton et al., 2001). Perhaps the strongest result concerning this idea came recently from the detection of an "anomalous" ^{6}Li/^{7}Li ratio on the star HD 82943 (Israelian et al., 2001; Israelian et al., 2002), a late F dwarf known to have two orbiting planets.

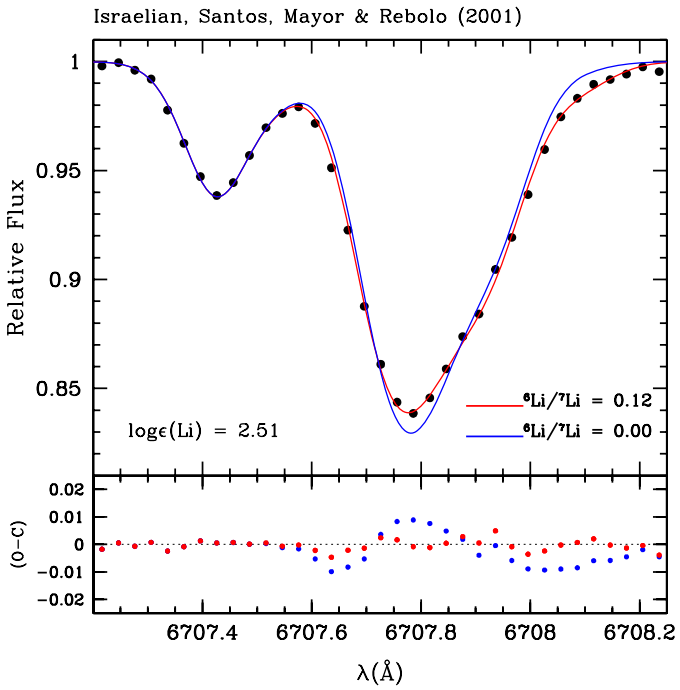

Figure 8. ^{6}Li signature on the spectrum of HD 82943 (dots) and two spectral synthesis, one with no ^{6}Li and the other with a ^{6}Li/^{7}Li ratio compatible with the meteoritic value. The O-C residuals of both fits are shown. From Israelian et al. (2001).

The rare ^{6}Li isotope represents an unique way of looking for traces of "pollution". So far, this isotope had been detected in only a few metal-poor halo and disc stars, but never with a high level of confidence in any metal-rich or even solar-metallicity star. Standard models of stellar evolution predict that ^{6}Li nuclei are efficiently destroyed during the early evolution of solar-type (and metallicity) stars and disappear from their atmospheres within a few million years. Planets, however, do not reach high enough temperatures to burn ^{6}Li nuclei, and fully preserve their primordial content of this isotope. A planet engulfed by its parent star would boost the star's atmospheric abundance of ^{6}Li.

In fact, as discussed in Israelian et al. (2001) and Israelian et al. (2002), planet (or planetesimal) engulfment following e.g. planet-planet (Rasio and Ford, 1996) interactions seems to be the only convincing and the less speculative way of explaining the presence of this isotope in the atmosphere of HD 82943.

Recently, some authors have casted some doubts into the reality of the ^{6}Li detection in HD 82943 (Reddy et al., 2002): they suggested that the observed feature is not from ^{6}Li, but is rather a previously unidentified line of Ti. The most recent analysis seem, however, to show that the feature does not seem to belong to this latter element, and thus confirm that the presence of ^{6}Li is the best way of explaining the observed spectral signature (Israelian et al., 2002). Although not conclusive,

these facts give strong support to the idea that ^{6}Li is present in the atmosphere of HD 82943.

It is important to note, however, that the quantity of material we need to add to the atmosphere of HD 82943 in order to explain the lithium isotopic ratio would not be able to change the [Fe/H] of the star by more than a few cents of a dex. Furthermore, it it not clear how often this kind of events occur, but some studies based on the analysis of other light elements seem to support that at least the quantity of material engulfed cannot be very high (Santos et al., 2002b). In any case, the conclusions presented above supporting a "primordial" source for the high [Fe/H] of planet host stars, are not dependent on these cases of "self-enrichment".

Note also that we are referring to the fall of planets or planetary material after the star has reached the main-sequence phase and fully developed a convective envelope; if engulfment happens before that, all planetary material will be deeply mixed, and no traces of self-enrichment might be found. It is interesting to note that the whole giant-planetary formation phase must take place when a disk of gas (and debris) is present. Massive gas disks may not exist at all when a star like the Sun reaches the main-sequence phase (although still not clear, inner disks seem to disappear after $\sim$10 Myr – e.g. Haisch et al., 2001). Thus, all the "massive" infall that would be capable of changing the measured elemental abundances if the star were already at the main-sequence might simply occur too early, explaining why we do not see strong traces of pollution (in particular, concerning iron).

5. Concluding Remarks

The main results discussed in this review go as follows:

- Planet host stars are metal-rich when compared to the average local field dwarfs, and this excess metallicity seems to reflect the higher metallicity of the gas clouds that gave origin to the star and planetary systems. Furthermore, this result cannot be related to any observational bias.
- The probability of finding a planet increases with the metallicity of the star. About 7% of the stars in the CORALIE planet-search sample having [Fe/H] between 0.3 and 0.4 dex have a planetary companion. This frequency falls to a value of less then 1% for stars with solar metallicity.
- No clearly significant trends between the metallicity of the host stars and the orbital parameters of the planets are seen. However, there are some evidences for an increase of the upper limit mass of a planet with increasing stellar metallicity. Other less strong trends are seen also for the orbital eccentricity and period.
- There are some hints of stellar pollution seen though the analysis of the lithium isotopic ratio.

The results presented above seem to support a scenario where the formation of giant planets, or at least of the type we are finding now, is particularly dependent on the [Fe/H] of the primordial cloud. In other words, and as already discussed in Santos et al. (2001) and Santos et al. (2003), the metallicity seems to play a crucial role in the formation of giant planets, as it is indicated by the shape of the metallicity distribution. This result does not exclude, however, that a giant planet might be formed around a lower metallicity star, but rather that the probability of planetary formation is lower in that case.

Somehow, our results further support the core accretion scenario (Pollack et al., 2001) against the disk instability model (Boss, 2000) as the "main" mechanism of giant-planetary formation. In fact, (Boss, 2002) has shown that contrarily to the core-accretion models, the efficiency of planetary formation in the case of the disk instability should not strongly depend on the metallicity of the disk. In other words, if that were the case, we should probably not see an increase in the frequency of planets as a function of the metallicity: such a trend is clearly seen in our data. We note, however, that the current data does not discard that both situations can occur; the disk instability model could be, for example, responsible for the formation of the more massive "planets", or of the companions (brown-dwarfs of planets) around the more metal-deficient stars.

Acknowledgements

We would like to thank the members of the Geneva extra-solar planet search group, D. Naef, F. Pepe, D. Queloz, S. Udry, our French colleagues J.-P. Sivan, C. Perrier and J.-L. Beuzit, as well as R. Rebolo and R.G. Lopez (from the IAC), who have largely contributed to the results presented here. We wish to thank the Swiss National Science Foundation (Swiss NSF) for the continuous support to this project. Support from Fundação para a Ciência e Tecnologia, Portugal, to N.C.S. in the form of a scholarship is gratefully acknowledged.

References

Boss, A., 2002: 'Stellar Metallicity and the Formation of Extrasolar Gas Giant Planets', *ApJL* **567**, 149.

Boss, A., 2000: 'Possible Rapid Gas Giant Planet Formation in the Solar Nebula and Other Protoplanetary Disks', *ApJL* **536**, 101, 2000ApJ...536L.101B.

Gonzalez, G.: 1998, 'Spectroscopic analyses of the parent stars of extrasolar planetary system candidates', *A&A* **334**, 221.

Gonzalez, G.: 1997, 'The stellar metallicity-giant planet connection', *MNRAS* **285**, 403.

Gratton R.G., Bonanno G., Claudi R.U., Cosentino R., Desidera S., Lucatello S., Scuderi S.: 2001, 'Non-interacting main-sequence binaries with different chemical compositions: Evidences of infall of rocky material?', *A&A* **377**, 123.

Haisch, K.E., Jr., Lada, E.A., Lada, C.J., 2001: 'Disk Frequencies and Lifetimes in Young Clusters', *ApJL* **553**, 153.

Israelian, G., Santos, N. C., Mayor, M., Rebolo, R.: 2002, 'New measurement of the ^{6}Li/^{7}Li isotopic ratio in the extrasolar planet host star HD 82943 and line blending in the Li 6708Å region', *A&A* **403**, 753.

Israelian, G., Santos, N. C., Mayor, M., Rebolo, R.: 2001, 'Evidence for planet engulfment by the star HD82943', *Nature* **411**, 163.

Kokubo, E., Ida, S.,: 2002, 'Formation of proto-planetary systems and diversity of planetary systems', *ApJ* **581**, 666.

Latham D.W., Stefanik R.P., Mazeh T., Mayor M., Burki G.: 1989 'The unseen companion of HD 114762 - A probable brown dwarf', *Nature* **339**, 38.

Laws, C., Gonzalez, G.: 2001, 'A Differential Spectroscopic Analysis of 16 Cygni A and B', *ApJ* **553**, 405.

Mayor, M., and Queloz, D.: 1995, 'A Jupiter-Mass Companion to a Solar-Type Star', *Nature* **378**, 355.

Pinsonneault, M. H., DePoy, D. L., Coffee, M.: 2001, 'The Mass of the Convective Zone in FGK Main-Sequence Stars and the Effect of Accreted Planetary Material on Apparent Metallicity Determinations', *ApJL* **556**, 59.

Pollack, J.B., Hubickyj, O., Bodenheimer, P., Lissauer, J.J., Podolak, M., Greenzweig, Y.: 1996, 'Formation of the Giant Planets by Concurrent Accretion of Solids and Gas', *Icarus* **124**, 62.

Queloz, D., Mayor, M., Weber, L., Blécha, A., Burnet, M., Confino, B., Naef, D., Pepe, F., Santos, N., Udry, S.: 2000, 'The CORALIE survey for southern extra-solar planets. I. A planet orbiting the star Gliese 86', *Icarus* **124**, 62.

Rasio, Frederic A., Ford, Eric B., 1996: 'Dynamical instabilities and the formation of extrasolar planetary systems', *Science* **274**, 954.

Reddy, B.E., Lambert, D.L., Laws, C., Gonzalez, G., Covey, K.: 2002: 'A search for ^{6}Li in stars with planets', *MNRAS* **335**, 1005.

Santos, N.C., Garcia Lopez, R.J., Israelian, G., Mayor, M., Rebolo, R., Garcia-Gil, A., Perez de Taoro, M.R., Randich, S.: 2002b, 'Beryllium abundances is stars hosting giant planets', *A&A* **386**, 1028.

Santos, N.C., Israelian, G., Mayor, M., Rebolo, R., Udry, S.: 2003, 'Statistical properties of exoplanets II. Metallicity, orbital parameters, and space velocities', *A&A* **398**, 363.

Santos, N. C., Mayor, M., Naef, D., Pepe, F., Queloz, D., Udry, S., Burnet, M., Clausen, J. V., Helt, B. E., Olsen, E. H., Pritchard, J. D.: 2002a, 'The CORALIE survey for southern extra-solar planets. IX. A 1.3-day period brown dwarf disguised as a planet', *A&A*, **392**, 215.

Santos, N.C., Israelian, G., Mayor, M.: 2001, 'The metal-rich nature of stars with planets', *A&A* **373**, 1019.

Wolszczan, A., Frail, D.A.: 1992, 'A planetary system around the millisecond pulsar PSR 1257+12', *Nature* **355**, 145.

Udry, S., Mayor, M., Naef, D., Pepe, F., Queloz, D., Santos, N. C., Burnet, M.: 2002, 'The CORALIE survey for southern extra-solar planets. VIII. The very low-mass companions of HD141937, HD 162020, HD 168443 and HD 202206: Brown dwarfs or "superplanets"?', *A&A* **390**, 267.

Udry, S., Mayor, M., Naef, D., Pepe, F., Queloz, D., Santos, N.C., Burnet, M., Confino, B., Melo, C.: 2000, 'The CORALIE survey for southern extra-solar planets. II. The short-period planetary companions to HD 75289 and HD 130322', *A&A* **356**, 590.

Current Challenges Facing Planet Transit Surveys

D. Charbonneau
California Institute of Technology, 105-24 (Astronomy),
1200 E. California Blvd., Pasadena, CA 91125 USA

Abstract. The initial task that confronted extrasolar-planet transit surveys was to monitor enough stars with sufficient photometric precision and complete phase coverage. Numerous searches have been pursued over the last few years. Among these projects are shallow, intermediate, and deep surveys of the Galactic plane, and monitoring of open clusters, and a globular cluster. These projects have all defeated the initial technical challenge, but a new obstacle has risen in its place: Single-color photometric time series are not sufficient to identify uniquely transiting planet systems, as eclipsing binary stars can mimic the signal. Multicolor photometric time series and multi-epoch spectroscopic monitoring are required to cull the list of candidates prior to high-precision radial velocity monitoring. I also discuss the prospects for detecting another transiting system among the planets found by the radial-velocity method, as well as review the recent announcement of OGLE-TR-56 b, the first extrasolar planet detected by the transit method.

1. Introduction

Among the many great changes invoked by the Mayor and Queloz (1995) detection of the planet orbiting 51 Pegasi were those concerning the potential use of photometric transits to detect and characterize extrasolar planets. Prior to 51 Peg b, several papers (Rosenblatt, 1971; Borucki and Summers, 1984; Borucki et al., 1985; Giampapa et al., 1995) had outlined the chief obstacle facing the transit method: Ground-based photometry was likely to succeed only for gas-giant planets, yet such planets were expected only at large distances. Even if all Sun-like stars had a Jupiter, transiting systems would be rare (only 1 in 1000 systems would have the inclination to transit), and, worse yet, the transit event in such systems would occur only once per 12-year orbital period.

In contrast to this scenario, there are now more than 20 active transit surveys (see Horne, 2003 for a complete listing[1]), nearly all of which are focussed on detecting analogs of 51 Peg (which I shall refer to here as hot Jupiters). I remind the reader briefly of the characteristics of the signal the transit searchers seek: The amplitude of the flux decrement is roughly $(R_p/R_*)^2 \simeq 0.01$, where R_p is the radius of the planet, and R_* is that of the star. Transits occur once per 3–7 day orbital period, and last 2–4 hours. The rate of occurrence of hot Jupiters for Sun-like stars is currently estimated at $r = 0.0075$ (Butler et al., 2001), and the likelihood of a hot Jupiter system with a semi-major axis a presenting a transiting inclination is $p \simeq (R_*/a) \simeq 0.1$ (for a uniform distribution of orbital inclinations).

[1] See also `http://star-www.st-and.ac.uk/~kdh1/transits/table.html`, maintained by K. Horne

Assuming that complete phase coverage is achieved, the number of stars that must be examined to find one transiting hot Jupiter system is $n = 1/(r\ p\ g) \simeq 1300/g$, where g is the fraction of stars examined that are "good" targets, i.e. Sun-like and not members of a close binary.

2. Photometry of Stars with Doppler-Detected Planets

Numerous groups have performed high-precision photometry on the known planetary systems with the smallest semi-major axes, and transits by gas-giant objects have been ruled out (in order of increasing semi-major axis) for HD 83443 (S. Udry, personal communication), HD 46375 (Henry, 2000), HD 179949 (Tinney et al., 2001), HD 187123 (Castellano, 2000), τ Boo (Baliunas et al., 1997; Henry et al., 2000), BD-10°3166 (Butler et al., 2000), HD 75289 (S. Udry, personal communication), 51 Peg (Henry et al., 1997; Henry et al., 2000), υ And b (Baliunas et al., 1997; Henry et al., 2000), HD 49674 (Butler et al., 2003), HD 168746 (Pepe et al., 2002), HD 108147 (Pepe et al., 2002) & 55 Cnc b (Baliunas et al., 1997; Henry et al., 2000).

It is worthwhile to consider whether these non-detections are consistent with our expectations. If the probability of a transiting configuration is p, then the probability of finding k transiting systems of n stars examined is $P = n!/[k!(n - k)!](1 - p)^{n-k} p^k$. If we assumed that each of the 14 systems that were monitored (HD 209458 and the 13 listed above) had a probability of presenting transits of $p \simeq R_*/a \simeq 0.1$ (for the Sun, this corresponds to a semi-major axis of 0.047 AU), then the chance of finding one (and only one) transiting system is 0.36. The chance of having found no such systems is 0.23. On the other hand, the chances of having found 2 or 3 such systems are 0.26 and 0.11 respectively. Any of these outcomes would have been roughly consistent with a uniform distribution of orbital inclinations (and the scenario that materialized was the single most likely one).

We needn't assume a uniform value of $p = 0.1$ for all systems; we can estimate R_* from parallax measurements (or stellar modeling), and a is calculated from the radial velocity period and an estimate of the stellar mass M_*. Excluding HD 209458, one can then ask what the probability was of examining the other 13 systems and finding no transits. This value is given by $P = \prod_{i=1}^{13} (1 - R_{*,i}/a_i) \simeq 0.26$ (where I have assumed values for R_* gathered from the literature). This number is also consistent with a uniform distribution of orbital inclinations.

It is reasonable to ask whether photometric monitoring of additional systems is worthwhile. Such observing campaigns are increasingly difficult for longer periods, as the transits are more infrequent, and the uncertainties in the predicted times of the events are greater. With these considerations in mind, G. Laughlin has established a project[2] to motivate amateur astronomers to pursue the most

[2] See `http://www.transitsearch.org`, maintained by G. Laughlin.

promising of the remaining extrasolar planet systems. The requisite photometric precision can be achieved with amateur-grade, commercially-available CCD cameras. Amateur telescopes provide more than enough aperture to gather the requisite flux (recall that Charbonneau et al., 2000 used a 10 cm aperture Schmidt camera to record the first transits of HD 209458), although the paucity of sufficiently bright calibrator stars in the typical field-of-view of such instruments can be a problem. This network has recently claimed to rule out transits for HD 217107 (Fischer et al., 1999). There are 24 extrasolar planets with periods less than 200 days that have not yet been examined for transits. For this sample, the probability of at least one transiting system is $P = 1 - \prod_{i=1}^{24} (1 - R_{*,i}/a_i) \simeq 0.62$ (using values for R_* as listed at the project website).

Photometric monitoring of Doppler-detected planet systems is also a useful check that the radial velocity variations are due to an orbiting planet, and not intrinsic stellar variability. Recently, Henry et al. (2002) presented photometry of HD192263 showing variability at the RV period, and casting the planet interpretation (Santos et al., 2000; Vogt et al., 2000) into doubt.

Although there may be a few more transiting planets in the current radial velocity sample, the ongoing Doppler surveys will not provide a substantial population of such objects. The primary goal of these surveys is to characterize the planet population at large semi-major axes. As a result, these surveys will continue to monitor the current target list (comprising some 1500 stars) for many years to come, but will not add many new targets. Since hot Jupiters are the most quickly and easily detected, the current Doppler precision of 3 m s^{-1} has likely revealed the majority of such objects with masses greater than 0.2 $M_{\rm Jup}$ in the current target list. The desire for a large number of transiting hot Jupiters, and the realization that the current Doppler surveys are unlikely to provide this sample, motivates the various transit searches that are the subject of the remainder of this paper.

3. Radial Velocity Follow-Up of Transit Candidates

Transiting extrasolar planets are of substantial value only if the radial velocity orbit can be measured. Since the current transit surveys target stars ranging in brightness from $9 \leq V \leq 21$, it is worthwhile to consider the resources that will be required to accomplish this follow-up measurement once candidate systems have been identified.

For typical Sun-like stars with rotational velocities at or below a spectrograph resolution R, the Doppler precision may be roughly estimated by (Brown, 1990):

$$\delta v_{\rm rms} \simeq \frac{c}{R\ d\ (N_{\rm lines}\ N_{\rm pix}\ I_c)^{1/2}}, \tag{1}$$

where c is the speed of light, d is the typical fractional line depth (relative to the continuum), I_c is the continuum intensity per pixel, $N_{\rm lines}$ is the number of

spectral lines, and $N_{\rm pix}$ is the number of pixels per line. As an example, consider the precision that may be expected from the HIRES spectrograph on the Keck I telescope: Assuming $d = 0.4$, $N_{\rm pix} = 2$, and $N_{\rm lines} = 100$, and a resolution $R = 70\,000$ and count level $I_c = 90\,000$ (corresponding to a 5 minute exposure of a late G star at $V = 8$), then the formula above yields $\delta v_{\rm rms} = 2.5\ {\rm m\,s^{-1}}$. G. Marcy (personal communication) reports typical photon-limited Doppler errors of $3\ {\rm m\,s^{-1}}$ for 5 minute integrations on V=8 G-dwarf stars, in keeping with this estimate. (In practice, the achieved Doppler precision results from the use of more lines than that assumed here, at a variety of line depths and signal-to-noise ratios).

For radial-velocity follow-up of stars with candidate transiting planets, we are interested to know how much telescope time $t_{\rm obs}$ will be required to detect (or exclude) a secondary mass M_p orbiting a star of mass M_* with a period P. To derive this relation, I start with amplitude of the radial velocity signature induced on the primary,

$$K_* = \left(\frac{2\,\pi\,G}{P}\right)^{\frac{1}{3}} M_p \sin i\ \ M_*^{-\frac{2}{3}}. \tag{2}$$

Equation 1, and the experience of the radial velocity observers, tells us that a precision of $3\ {\rm m\,s^{-1}}$ is obtained with Keck/HIRES in 5 minutes on a $V = 8$ star. I assume that the Doppler precision is photon-noise-limited, and that the amplitude must exceed 4 times the precision to be secure (in keeping with the rule-of-thumb suggested by Marcy et al., 2000). Equating these requirements,

$$\frac{K_*}{4} = \delta v_{\rm rms} = 3\ {\rm m\,s^{-1}} \left[\left(\frac{t_{\rm obs}}{5\ {\rm min}}\right) 10^{-\frac{V-8}{2.5}}\right]^{-\frac{1}{2}}, \tag{3}$$

I solve for the required integration time,

$$t_{\rm obs} = 0.0363\ {\rm min} \left(\frac{M_p}{M_{\rm Jup}}\right)^{-2} \left(\frac{M_*}{M_{\odot}}\right)^{\frac{4}{3}} \left(\frac{P}{3\ {\rm days}}\right)^{\frac{2}{3}} 10^{\frac{V-8}{2.5}}. \tag{4}$$

Equation 4 makes it clear that the required integration time is very sensitive to the planetary mass and the stellar brightness (as one would expect). It is important to avoid the temptation to assume $M_p = 1\ M_{\rm Jup}$ in calculating $t_{\rm obs}$. Of the 17 planets with $a \leq 0.1$ AU, 70% have masses below 1 $M_{\rm Jup}$, and the median mass is 0.5 $M_{\rm Jup}$. Thus if one adopts the value of $r = 0075$ (see §1) in planning a transit search, one should also adopt a mass of 0.2 $M_{\rm Jup}$ in calculating the required $t_{\rm obs}$ with equation 4. The result is to increase the predicted integration times by a factor of 25.

4. A Selection of Transit Surveys

While it is not possible here to review all current transit searches, I have selected four surveys which straddle the range of current efforts. I briefly review the status of the projects, before summarizing the prospects for Doppler (and additional) follow-up observations of candidates yet-to-be identified by these surveys.

4.1. SHALLOW, WIDE-ANGLE SURVEYS

Working with J. T. Trauger's team at NASA/JPL, I have assembled a small-aperture wide-field transit-search instrument (Figure 1) that typifies many extant systems, such as STARE (Brown and Charbonneau, 2000) and Vulcan (Borucki et al., 2001). The system (with the exception of the primary CCD camera) is built entirely of commercially-available items typically intended for amateur use. I describe the basic features here for the benefit of those readers contemplating the fabrication of a similar system. The main optic is an f/2.8 280mm camera lens imaging a $5.7^\circ \times 5.7^\circ$ patch of the sky onto an 2k×2k thinned CCD. A micrometer allows for automated focus adjustments. Each CCD pixel is 13.5μm, corresponding to 10 arcsec. A filter wheel houses the SDSS g', r', i', z' and Bessell R filters. For guiding, an f/6.3 440mm lens feeds a commercial prepackaged CCD guide camera. The system is mounted in an equatorial fork mount, and housed in a refurbished clamshell enclosure at Mt. Palomar in southern California. All systems are operated by a single Linux-based workstation, and a command script guides the instrument through each night's observing.

This system is the third in a network: The other two instruments are STARE (PI: T. M. Brown), located in the Canary Islands, and PSST (PI: E. W. Dunham), located in northern Arizona. Each telescope produces a time series of R-band images (with typical integration times of 2 minutes), and only one field is monitored for a typical observing campaign of 2 months. We perform weighted-aperture photometry on these images to produce a photometric time series for each star. (Image subtraction methods, such as those as described by Alard and Lupton, 1998, are unlikely to result in a significant increase in precision because the 10 arcsec pixels yield slightly undersampled and seeing-independent images). In a typical field-of-view centered on the Galactic plane, roughly 6000 stars ($9 \leq V \leq 11$) are monitored with sufficient accuracy to detect periodic transit-like events with an amplitude of 1%.

For a single telescope, the primary losses in efficiency are due to the day-night cycle, and weather. Currently the telescopes are operated as stand-alone systems monitoring the same field-of-view. Once networked so that the time series are combined prior to performing the search for transit-like signals, we expect a substantial increase in efficiency. Specifically, we anticipate that the 2 months required by a single instrument to achieve 85% completion (for orbital periods less than 4.5 days) can be reduced to only 3 weeks (Figure 2) with the current longitudes afforded by

Figure 1. South-facing (left panel) and north-facing (right panel) views of the Palomar planet search instrument. The primary imaging lens is on the left side of the camera as shown in the right panel; a 280mm f/2.8 commercial lens images a $5.7^\circ \times 5.7^\circ$ patch of the sky onto an 2k×2k thinned CCD. The system is assembled almost entirely from amateur-grade, commercially-available parts, with the exception of the primary CCD camera.

the three sites. Since fields are exhausted much more quickly, the number of stars monitored in a year of operation is greatly increased.

4.2. Intermediate Galactic Plane Surveys – OGLE III

The team operating the Optical Gravitational Lensing Experiment has conducted a search for low-amplitude transits in three fields in the direction of the Galactic center (Udalski et al., 2002a; Udalski et al., 2002b). They obtained 800 I-band epochs per field spanning 32 nights. More than 5 million stars were monitored, to which they applied a substantial cut in color-magnitude space to reduce the number of targets to 52 000 disk stars ($14 \leq V \leq 18$) with photometry better than 1.5%. Of these, 59 candidates were identified with flat-bottomed eclipses with depths less than 8%.

Dreizler et al. (2002) presented spectroscopy of some of these candidate stars, with the goal of obtained secure spectral classifications. The stellar radii they inferred allowed them to calculate more accurate radii of the transiting objects. In many cases, the OGLE-III objects had radii that were too large for planet-mass bodies.

After a similar (but independent) spectroscopic reconnaissance, Konacki et al. (2003) found that 6 of the candidates had solar-type spectra with no radial-velocity

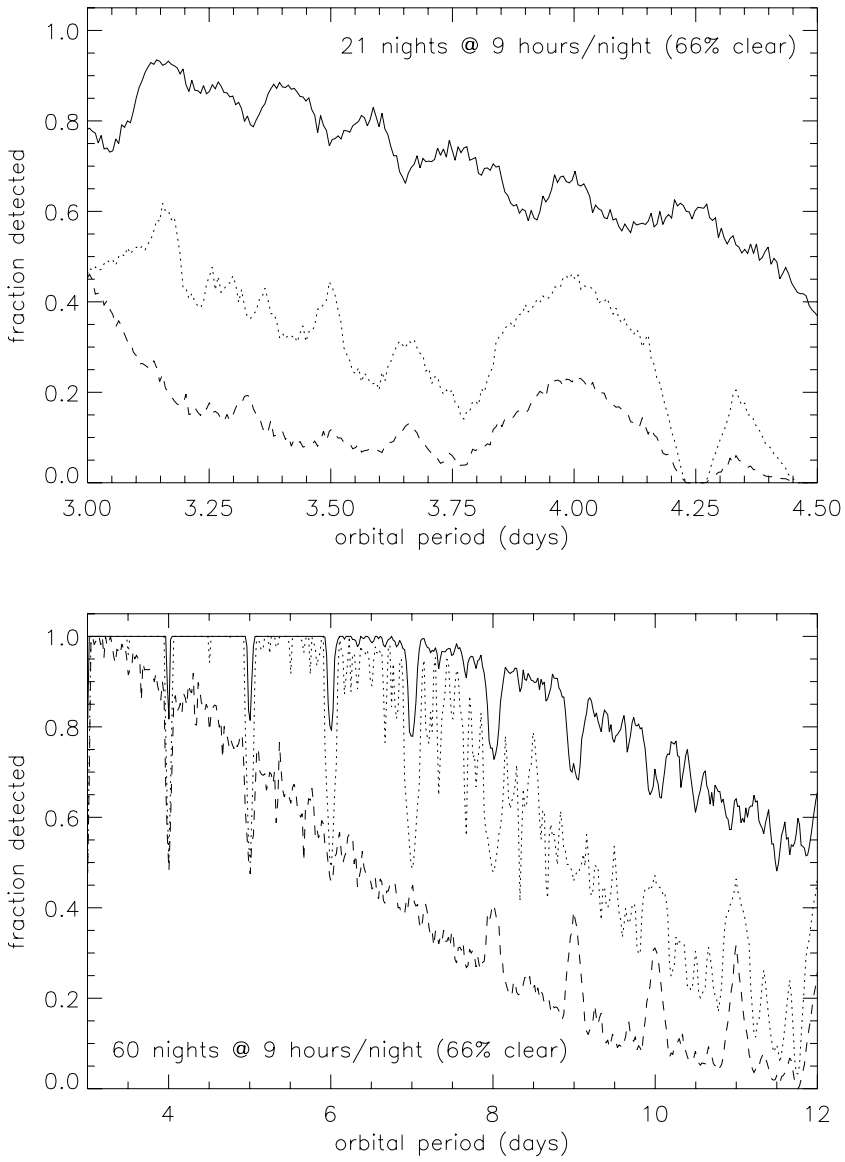

Figure 2. The upper panel shows the recovery rate for a 3-week campaign to find transiting planets (assuming 9 hours per night, and 66% clear weather), as a function of orbital period. Three half transits are required. The dashed line is the result for a single telescope, the dotted line is that for three telescopes at the same longitude, and the solid line is that for a 3-element network with a telescope in each of the Canary Islands, Arizona, and California. The lower panel shows the corresponding recovery rates for a 2-month campaign with the same night length and weather. In both panels, the network has nearly exhausted the field (and hence a new field can be monitored, increasing the total number of targets), whereas both the single-element and single-longitude systems are still lacking significant phase coverage.

variation at the level of a few km s^{-1}. Subsequent radial velocity monitoring with Keck/HIRES revealed that one candidate, OGLE-TR-56, showed a velocity change with an amplitude of 167 m s^{-1} and consistent with the 1.2-day photometric variation. The team performed numerous modeling tests to rule out spectral blends (a significant concern, as the OGLE-III fields are very crowded), leading them to announce the first detection of an extrasolar planet by the transit method. The newfound planet has a mass $M_p = 0.9 \pm 0.3\ M_{\rm Jup}$ and radius $R_p = 1.3 \pm 0.15\ R_{\rm Jup}$, and hence a density of $\rho = 0.5 \pm 0.3$ g cm^{-3}, similar to the more precise estimates for HD 209458 b (Brown et al., 2001). The current radial-velocity phase coverage is sparse, but will likely be filled in during the 2003 bulge season. The most surprising result is the 1.2-day orbital period. The Doppler surveys have found no orbital periods below 2.99 days, and a large pile-up of objects at this value (8 planets with periods less than 4 days).

4.3. Deep Galactic Plane Surveys – EXPLORE

The deepest transit search currently underway is the EXPLORE project (Mallen-Ornelas et al., 2003; Yee et al., 2002). The EXPLORE team conducts observing campaigns (lasting typically 2 weeks) using the CTIO 4-m and CFHT 3.6-m telescopes to continuously monitor fields in the Galactic plane in I-band. EXPLORE-I (southern hemisphere) received 11 nights in 2001 and delivered a photometric precision of better than 1% on 37 000 stars with $14.5 \leq I \leq 18.2$. EXPLORE-II (northern hemisphere) received 14 nights in 2001 and delivered a similar photometric precision of better than 1% on 9500 stars. Both surveys acquired data with a similar sampling interval of 2.7 minutes. The team has identified several candidates, and spectroscopic follow-up has been conducted on VLT/UVES and Keck/HIRES.

The primary benefit of the EXPLORE campaign is that they will probe a significant number of low-mass (K & M) main sequence stars. This is in contrast to the wide-field surveys, which will probe mostly F & G stars (as less massive stars are not a significant fraction of the $V < 11$ population). The principal challenge facing EXPLORE is that even preliminary follow-up radial velocity monitoring (i.e. with the goal of ruling out eclipsing binary star systems) is very resource intensive, requiring 8-m class telescopes.

4.4. Open Cluster Surveys – PISCES

The identification of open clusters as good targets for transit searches was pointed out by Janes (1996). Open clusters offer ideal laboratories in which to study the characteristics of hot Jupiters, since the stars share a common metallicity and age. The search for transiting planets in open clusters is motivated in part by the conclusion by Gilliland et al. (2000) that the population of hot Jupiters in the globular cluster 47 Tuc is greatly depleted relative to the local solar neighborhood. Two of the many contending explanations for this result are (1) the low-metallicity

environment results in a reduced formation and/or migration rate of Jupiter-mass planets, or (2) protoplanetary disks are tidally disrupted by the close passage of stars in the crowded environment of a globular cluster. Open clusters offer environments at a range of metallicities, yet are not crowded enough for disruption by stellar encounters to be significant. There is an additional benefit for searching for transiting planets in open clusters: interpretation of candidates is greatly simplified, since the stellar radius and mass can be reliably assumed from the cluster color-magnitude diagram.

Mochejska et al. (2002) present results of a month-long campaign on the open cluster NCG 6791. They demonstrate adequate photometric precision for the detection of hot Jupiters, and present lightcurves for 62 variable stars (but no planet candidates). The challenges facing their survey are representative of those for other open-cluster searches (e.g. Street et al., 2002): Open clusters typically contain several thousand stars, and thus would yield only a handful of detections even if complete phase coverage can be obtained. Specifically for Mochejska et al. (2002), NGC 6791 has roughly 10 000 member stars, of which 4110 (59%) and 2053 (29%) have sufficient photometric precision to detect transits for planets with radii of 1.5 and 1.0 $R_{\rm Jup}$, respectively. Furthermore, these stars are very faint ($17 \leq R \leq 21$), and thus even preliminary follow-up spectroscopy (i.e. with the goal of ruling out eclipsing binary star systems) will likely require 8-m class telescopes.

5. The Task at Hand

Several years ago, the principal technical challenge facing proposed transit surveys could be summarized in the following question: Could enough stars ($\sim$10 000) be surveyed with sufficient precision ($\sim$ 3mmag) and for enough nights to obtain complete phase coverage? The good news is that the answer to this question is a definite yes, as evidenced by the diversity of projects described in §4. However, a new challenge has arisen, which I will broadly describe as sorting out the "false positives". Single-color photometric time series alone are not sufficient to identify uniquely the transiting-planet systems. There are at least three general forms of apparent variability that can mimic these signals, all of which involve an eclipsing binary star system. The first is grazing incidence equal-size stellar binaries, such that the occulted area is roughly 1% (and the period is underestimated by a factor of 2). In general, this contaminant can be ruled out by sufficiently precise and rapid photometric observations during transit, as grazing incidence binaries will present a V-shaped (as opposed to a flat-bottomed) eclipse. The second contaminant is a central transit by a smaller star in front of a larger one. Multi-epoch low-precision ($\sim$ km s^{-1}) radial-velocity monitoring would identify such systems, as the amplitude of the radial velocity orbit would be orders of magnitude larger than that expected for a hot Jupiter. The third and most insidious contaminant is a stellar blend, where an eclipsing binary (with central transits) has its eclipse depth

diluted down to $\sim$1% by the light of a third star (either physically associated, or simply lying along the line of sight). In general, multi-color photometry of such a system should reveal a color-dependent transit depth, whereas the transit depth for a hot-Jupiter system should vary only slightly (due to the effects of stellar limb-darkening). For examples of such systems, and follow-up spectroscopy revealing their true (non-planetary) nature, see Borucki et al. (2001).

In summary, most contaminating systems can be rejected by either (1) rapid-cadence, multi-color photometry of the transit curve revealing either a V-shaped eclipse, or a large color-dependence to the transit depth, or (2) multi-epoch spectroscopy revealing large Doppler variations, consistent with a massive secondary.

Transiting planets are of significant value only if both their mass and radius can be estimated. Moreover, the reality of putative planets discovered by transit photometry but without a measured radial-velocity orbit may be doubted. Conversely, the detection of a radial-velocity orbit at the same period and phase as the ones derived from photometry is strong evidence supporting the planet interpretation.

As an illustration of the dramatic effects of target brightness on the prospects for detecting the radial velocity orbit, consider Keck/HIRES observations of stars with $V = 10.5$ (typical brightness for a target star in the wide-angle surveys), $V = 15.3$ (the brightness of OGLE-TR-56) and $V = 18$ (faint star in a cluster survey or deep Galactic plane search). Using equation 4, at $V = 10.5$, a detection of a 0.2 $M_{\rm Jup}$ planet can be achieved with integrations of only 10 minutes. For OGLE-TR-56 ($M_p = 0.9\ M_{\rm Jup}$, $P = 1.2$ d), integrations of 20 minutes are required (indeed, this is close to what Konacki et al., 2003 used). However, less massive planets at longer orbital periods ($M_p = 0.5\ M_{\rm Jup}$, $P = 3$ d) are a challenge, requiring 2 hours of integration per measurement. At $V = 18$, the situation is very difficult indeed: A 1-hour integration yields a detection limit of only 2.5 $M_{\rm Jup}$ for a 3-day orbital period. A planet with a mass of 1 $M_{\rm Jup}$ would require more than 6 hours of integration per observation (and recall than most of the known hot Jupiters have masses below 1 $M_{\rm Jup}$). In summary, radial velocity measurements are straightforward for the typical targets in the wide-angle surveys, challenging but feasible for the intermediate galactic plane surveys (targets brighter than $V = 16$), and unlikely to succeed for targets toward the faint end of the deep Galactic plane searches, and some open cluster surveys. Finally, it is important to note another strong reason to favor bright stars: Many follow-up measurements of HD 209458 are now being vigorously pursued (see Charbonneau, 2003 for a summary), and most of these are photon-noise limited. While some of these techniques may be feasible for candidates emerging from the wide-angle surveys (with stars typically 10 times fainter than HD 209458), they are unlikely to approach a useful precision for fainter stars.

References

Alard, C. and Lupton, R. H.: 1998, 'A Method for Optimal Image Subtraction', *Astrophysical Journal* **503**, 325.

Baliunas, S. L., Henry, G. W., Donahue, R. A., Fekel, F. C., and Soon, W. H.: 1997, 'Properties of Sun-like Stars with Planets: ρ^1 Cancris, τ Bootis, and υ Andromedae', *Astrophysical Journal* **474**, L119.

Borucki, W. J., Caldwell, D., Koch, D. G., Webster, L. D., Jenkins, J. M., Ninkov, Z., and Showen, R.: 2001, 'The Vulcan Photometer: A Dedicated Photometer for Extrasolar Planet Searches', *Publications of the Astronomical Society of the Pacific* **113**, 439–451.

Borucki, W. J., Scargle, J. D., and Hudson, H. S.: 1985, 'Detectability of Extrasolar Planetary Transits', *Astrophysical Journal* **291**, 852–854.

Borucki, W. J. and Summers, A. L.: 1984, 'The Photometric Method of Detecting Other Planetary Systems', *Icarus* **58**, 121–134.

Brown, T. M., Charbonneau, D., Gilliland, R. L., Noyes, R. W., and Burrows, A.: 2001, 'Hubble Space Telescope Time-Series Photometry of the Transiting Planet of HD 209458', *Astrophysical Journal* **552**, 699–709.

Brown, T. M.: 2000, 'The STARE Project: A Transit Search for Hot Jupiters', in *ASP Conf. Ser. 219: Disks, Planetesimals, and Planets*, F. Garzon, C. Eiroa, D. de Winter, and T. J. Mahoney (eds.), ASP, San Franciso, 584–589, `astro-ph/0005009`.

Brown, T. M.: 1990, 'High Precision Doppler Measurements via Echelle Spectroscopy', in *ASP Conf. Ser. 8: CCDs in Astronomy*, G. H. Jacoby (ed.), ASP, San Franciso, 335–344.

Butler, R. P., Marcy, G. W., Vogt, S. S., Fischer, D. A., Henry, G. W., Laughlin, G., and Wright, J.: 2003, 'Seven New Keck Planets Orbiting G and K Dwarfs', *Astrophysical Journal* **582**, 455.

Butler, R. P., et al.: 2001 'Statistical Properties of Extrasolar Planets', in *Planetary Systems in the Universe: Observation, Formation, and Evolution*, A. J. Penny, P. Artymowicz, A.-M. Lagrange, and S. S. Russell (eds.), ASP, San Francisco.

Butler, R. P., Vogt, S. S., Marcy, G. W., Fischer, D. A., Henry, G. W., and Apps, K.: 2000, 'Planetary Companions to the Metal-rich Stars BD -10°3166 and HD 52265', *Astrophysical Journal* **545**, 504–511.

Castellano, T.: 2000, 'A Search for Planetary Transits of the Star HD 187123 by Spot Filter CCD Differential Photometry', *Publications of the Astronomical Society of the Pacific* **112**, 821–826.

Charbonneau, D.: 2003, 'HD 209458 and the Power of the Dark Side', in *ASP Conf. Ser.: Scientific Frontiers in Research on Extrasolar Planets*, D. Deming and S. Seager (eds.), ASP, San Francisco, `astro-ph/0209517`.

Charbonneau, D., Brown, T. M., Latham, D. W., and Mayor, M.: 2000, 'Detection of Planetary Transits Across a Sun-like Star', *Astrophysical Journal* **529**, L45–L48.

Dreizler, S., Rauch, T., Hauschildt, P., Schuh, S. L., Kley, W., and Werner, K.: 2002, 'Spectral Types of Planetary Host Star Candidates: Two New Transiting Planets?', *Astronomy and Astrophysics* **391**, L17–L20.

Fischer, D. A., Marcy, G. W., Butler, R. P., Vogt, S. S., and Apps, K.: 1999, 'Planetary Companions around Two Solar-type Stars: HD 195019 and HD 217107', *Publications of the Astronomical Society of the Pacific* **111**, 50–56.

Giampapa, M. S., Craine, E. R., and Hott, D. A.: 1995, 'Comments on the Photometric Method for the Detection of Extrasolar Planets.', *Icarus* **118**, 199–210.

Gilliland, R. L. and 23 colleagues: 2000, 'A Lack of Planets in 47 Tucanae from a Hubble Space Telescope Search', *Astrophysical Journal* **545**, L47–L51.

Henry, G. W., Donahue, R. A., and Baliunas, S. L.: 2002, 'A False Planet around HD 192263', *Astrophysical Journal* **577**, L111–L114.

Henry, G. W.: 2000, 'Search for Transits of a Short-period, Sub-saturn Extrasolar Planet Orbiting HD 46375', *Astrophysical Journal* **536**, L47–L48.

Henry, G. W., Baliunas, S. L., Donahue, R. A., Fekel, F. C., and Soon, W.: 2000, 'Photometric and Ca II H and K Spectroscopic Variations in Nearby Sun-like Stars with Planets. III.', *Astrophysical Journal* **531**, 415–437.

Henry, G. W., Baliunas, S. L., Donahue, R. A., Soon, W. H., and Saar, S. H.: 1997, 'Properties of Sun-like Stars with Planets: 51 Pegasi, 47 Ursae Majoris, 70 Virginis, and HD 114762', *Astrophysical Journal* **474**, 503.

Horne, K.: 2003, 'Status and Prospects of Transit Searches: Hot Jupiters Galore', in *ASP Conf. Ser.: Scientific Frontiers in Research on Extrasolar Planets*, D. Deming and S. Seager (eds.), ASP, San Francisco.

Janes, K.: 1996, 'Star Clusters: Optimal Targets for a Photometric Planetary Search Program', *Journal Geophysical Research* **101**, 14853–14860.

Konacki, M., Torres, G., Jha, S., and Sasselov, D. D.: 2003, 'A New Transiting Extrasolar Giant Planet', *Nature* **421**, 507, astro-ph/0301052.

Mallen-Ornelas, G., Seager, S., Yee, H. K. C., Minniti, D., Gladders, M. D., Mallen-Fullerton, G., and Brown, T. M.: 2003, 'The EXPLORE Project I: A Deep Search for Transiting Extrasolar Planets', *Astrophysical Journal* **582**, 1123, astro-ph/0203218.

Marcy, G. W., Cochran, W. D., and Mayor, M.: 2000, 'Extrasolar Planets around Main-sequence Stars', in *Protostars and Planets IV*, V. Mannings, A. P. Boss, and S. S. Russell (eds.), Univ. of Arizona Press, Tucson, 1285.

Mayor, M. and Queloz, D.: 1995, 'A Jupiter-mass Companion to a Solar-type Star', *Nature* **378**, 355.

Mochejska, B. J., Stanek, K. Z., Sasselov, D. D., and Szentgyorgyi, A. H.: 2002, 'Planets in Stellar Clusters Extensive Search. I. Discovery of 47 Low-amplitude Variables in the Metal-rich Cluster NGC 6791 with Millimagnitude Image Subtraction Photometry', *Astronomical Journal* **123**, 3460–3472.

Pepe, F., Mayor, M., Galland, F., Naef, D., Queloz, D., Santos, N. C., Udry, S., and Burnet, M.: 2002, 'The CORALIE Survey for Southern Extra-solar Planets VII. Two Short-period Saturnian Companions to HD 108147 and HD 168746', *Astronomy and Astrophysics* **388**, 632–638.

Rosenblatt, F.: 1971, 'A Two-color Photometric Method for Detection of Extra Solar Planetary Systems', *Icarus* **14**, 71.

Santos, N. C., Mayor, M., Naef, D., Pepe, F., Queloz, D., Udry, S., Burnet, M., and Revaz, Y.: 2000, 'The CORALIE Survey for Southern Extra-solar Planets. III. A Giant Planet in Orbit around HD 192263', *Astronomy and Astrophysics* **356**, 599–602.

Street, R. A. and 9 colleagues: 2002, 'Variable Stars in the Field of Open Cluster NGC 6819', *Monthly Notices of the Royal Astronomical Society* **330**, 737–754.

Tinney, C. G., Butler, R. P., Marcy, G. W., Jones, H. R. A., Penny, A. J., Vogt, S. S., Apps, K., and Henry, G. W.: 2001, 'First Results from the Anglo-Australian Planet Search: A Brown Dwarf Candidate and a 51 Peg-like Planet', *Astrophysical Journal* **551**, 507–511.

Udalski, A. and 8 colleagues: 2002, 'The Optical Gravitational Lensing Experiment. Search for Planetary and Low-luminosity Object Transits in the Galactic Disk. Results of 2001 Campaign', *Acta Astronomica* **52**, 1–37.

Udalski, A., Zebrun, K., Szymanski, M., Kubiak, M., Soszynski, I., Szewczyk, O., Wyrzykowski, L., and Pietrzynski, G.: 2002, 'The Optical Gravitational Lensing Experiment. Search for Planetary and Low-luminosity Object Transits in the Galactic Disk. Results of 2001 Campaign – Supplement', *Acta Astronomica* **52**, 115–128.

Vogt, S. S., Marcy, G. W., Butler, R. P., and Apps, K.: 2000, 'Six New Planets from the Keck Precision Velocity Survey', *Astrophysical Journal* **536**, 902–914.

Yee, H. K. C., et al.: 2002, 'The EXPLORE Project: A Deep Search for Transiting Extra-solar Planets', in *Proceedings of the SPIE Conference: Astronomical Telescopes and Instrumentation* **4834**, 150, astro-ph/0208355.

Angular Momentum Constraints on Type II Planet Migration

William R. Ward
Dept. of Space Studies, Southwest Research Institute, Boulder, CO 80302, USA

Abstract. Type II migration, in which a newly formed large planet opens a gap in its precursor circumstellar nebula and subsequently evolves with it, has been implicated as a delivery mechanism responsible for close stellar companions. Large scale migration is possible in a viscously spreading disk when most of its mass is sacrificed to the primary in order to promote a small portion of the disk to much higher angular momentum orbits. There are, however, processes such as photodissociation that could truncate a disk at a finite distance. Recent numerical modeling has illustrated that planets can survive in this case. To date, this problem has been treated primarily by numerical case studies and has not been examined in the context of angular momentum conservation arguments. We show here that much of these results can be anticipated by a simple mapping procedure based on such arguments.

1. Introduction

Large scale type II planet migration can occur in a viscously spreading disk of surface density $\sigma(r, t)$ (Lin and Papaloizou, 1986). Embedded planets that have opened gaps generally follow the disk behavior unless their own angular momentum is comparable to or exceeds that of the disk. Numerical studies of type II migration have been conducted of late; in particular, Trilling *et al.* (2002) has performed a survey of 9000 model runs to obtain statistics of planet survival and ultimate location. Each run consisted of a single planet initially formed at 5 AU and an accretion disk, with a range of planet masses, disk masses, and viscosities considered.

If the disk is allowed to expand indefinitely, virtually all of the disk will fall into the primary in order to send a vanishingly small portion to infinity. For such a case, all planets that could cleanly segment the disk would eventually migrate inward, although those that start near the initial outer edge begin with a transient outward migration (*e.g.*, Lin and Papaloizou, 1986). Thus, it would be difficult to explain the survival of *any* planets if the evolution proceeded indefinitely. Of course, the evolution slows down as the disk expands; the characteristic time for the disk to reach a size $r_{\rm d}$ is of order $t \sim r_{\rm d}^2/\nu$, where ν is the viscosity. For instance, expanding the radius of a $\nu \propto r$ disk by a factor of 10 increases the characteristic time scale by the same factor. Trilling *et al.* (2002) avoid complete planet loss by terminating the model at a finite time. Realistically, there are processes that could truncate a disk at a distance, $r_{\rm d}$, and allow for its dissipation with a finite time span. Shu *et al.* (1993) have suggested that UV radiation from the early Sun could have truncated the solar nebula at $r_{\rm d} \sim O(10)$ AU by photoevaporation. Hollenbach *et al.* (2000) review a number of processes that could remove the disk. However, if

the disk has a finite outer boundary instead of a freely expanding edge, the fraction of its starting mass that is consumed by the star depends on the boundary distance at which the remaining material escapes.

Armitage *et al.* (2002; here after A02) extended the work of Trilling *et al.* by incorporating photo-evaporation into their numerical models, assuming this mechanism progressively removes disk material past $r_g \sim 5$ AU. With the existence a fixed outer edge, it is no longer true that all planets will ultimately migrate inward. A02 find that in some cases, planets can migrate outward permanently, and predict that in 10–15% of the cases a planet survives without significant net migration from its point of origin.

To date, this problem has been treated primarily by numerical case studies and has not been examined in the context of angular momentum conservation arguments, much of which can be done analytically (*e.g.*, Ward, 2003). Although numerical modeling can treat more complicated situations, an analytic approach can often yield complimentary insight into functional dependencies of the problem, providing an auxiliary tool to predict distribution statistics without the need for large numbers of simulations.

2. Conservation Constraints

Consider a disk of mass $M_d = 2\pi \int_0^{r_d} r\sigma(r,0)dr$, where r_d is an effective outer edge created by (say) photo-dissociation. In the limit of complete viscous dissipation of the disk, the amount of mass M_0 lost through an outer boundary is simply that necessary to remove all of the disk's original angular momentum[1],

$$L_d = 2\pi \int_0^{r_d} r\sigma\sqrt{GM_\star r}\, dr \quad ; \quad M_0 = \frac{L_d}{\sqrt{GM_\star r_d}} \tag{1}$$

where $M_\star$ is the primary's mass. For a power low profile, $\sigma \propto r^{-k} (k < 2)$,

$$L_d = C_k M_d \sqrt{GM_\star r_d} \quad ; \quad \frac{M_0}{M_d} = C_k \equiv \frac{2-k}{5/2 - k}. \tag{2}$$

Thus, the fraction $M_{acc}/M_d = 1 - C_k$ of the disk is accreted by the star, and the initial boundary between inward and outward destined material is

$$r_c = r_d\,[1 - C_k]^{1/(2-k)} = \frac{r_d}{(5-2k)^{1/(2-k)}} \tag{3}$$

as shown in Figure 1. Note that a significant fraction of the disk exits the system through the outer boundary.

Now add to the disk a planet of mass M_p with an initial orbit radius a_0. If the planet forms an effective tidal barrier to disk flow across its orbit, it segments

[1] We neglect the small angular momentum of gas being accreted by the star.

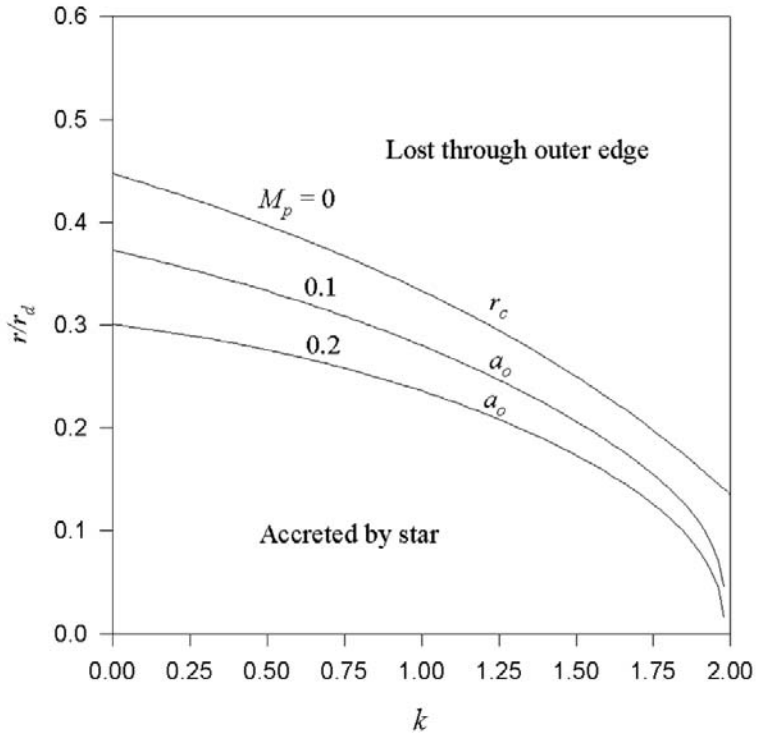

Figure 1. The critical radius r_c/r_d as a function of the surface density power index k. Disk gas below the curve is accreted by the star; gas above the curve is lost through the outer edge. Also shown are the minimum initial semimajor axes for survival of protoplanets of mass $M_p/M_d = 0.1, 0.2$.

the disk into inner and outer components. The disk mass external to the planet is $M_{ext} = M_d\left[1 - (a_0/r_d)^{2-k}\right]$; if the planet remains in the disk ($0 < a < r_d$), it will act as an angular momentum repository:

$$M_p\sqrt{GM_\star a} = L_d + M_p\sqrt{GM_\star a_0} - M_{ext}\sqrt{GM_\star r_d} \tag{4}$$

where a is the final position of the planet when the disk's mass is exhausted. This can be written in terms of the critical radius as

$$\left(\frac{a}{r_d}\right)^{1/2} = \left(\frac{a_0}{r_d}\right)^{1/2} + \left(\frac{M_d}{M_p}\right)\left[\left(\frac{a_0}{r_d}\right)^{2-k} - \left(\frac{r_c}{r_d}\right)^{2-k}\right] \tag{5}$$

It is clear that a planet initially located at r_c will not ultimately migrate. A planet inside (outside) the critical radius will migrate inward (outward). Setting $a = 0$ defines the minimum initial semi-axis a_0 for which a planet will survive. For small planets $M_p \ll M_d$, this approaches r_c, for large planets $M_p \gg M_d$ it approaches zero. Example curves for $M_p/M_d = 0.1, 0.2$ are included in Figure 1.

3. Disk Evolution

In general, however, disks with finite boundaries will not retain a power law profile as they evolve, even if they start that way. Accordingly, the critical point r_c may vary with time. Consequently, model calculations which form a planet at different points in a disk's life, but always at the same distance (*i.e.*, ~5 AU), may sometimes start it inside and sometimes outside $r_c(t)$. For such cases, the mapping can be written more generally as

$$\left(\frac{a}{a_0}\right)^{1/2} = 1 + \frac{M_0(t) - M_{\rm ext}(t)}{M_{\rm p}} \left(\frac{r_{\rm d}}{a_0}\right)^{1/2}$$
$$= 1 + \frac{2\pi}{M_{\rm p}} \left(\frac{r_{\rm d}}{a_0}\right)^{1/2} \int_{r_{\rm c}}^{a_0} r\sigma\ (r, t_{\rm form})\ {\rm d}r \tag{6}$$

where σ is the disk surface density profile at the time $t_{\rm form}$ the planet is formed.

Loss of disk material at both the star and the outer edge can be simulated by solving the diffusion equation

$$\frac{\partial^2 g}{\partial t^2} = \frac{4}{\nu \ell \Omega} \frac{\partial g}{\partial t} \tag{7}$$

for the viscous couple $g = 2\pi\sigma\nu r^3 \partial\Omega/\partial r = -3\pi\sigma\nu\ell$, requiring it to vanish at both boundaries (Lynden-Bell and Pringle, 1974). In eq. 7, $\ell \equiv r^2\Omega = \sqrt{GM_\star r}$ is the specific angular momentum of disk material for keplerian motion and ν is the viscosity. For a power law viscosity of the form $\nu = C_\nu r^n$, the solutions can be written

$$g = \sum_{j=1}^{\infty} A_j (k_j x)^l J_l(k_j x) e^{-k_j^2 t/\kappa^2} \tag{8}$$

where $1/l = 2(2-n), \kappa^{-2} = (3/4)(2-n)^2 (GM_\star)^{2-n} C_\nu, x \equiv \ell^{1/2l}$ and k_j is an eigenvalue found from the zeros of the Bessel function $J_l(k_j x_d) = 0$.

As a particularly simple example, consider a disk with a viscosity that increases linearly with r, *i.e.* $n = 1$. Then, $l = 1/2$, $x = \ell, \kappa^{-2} = (3/4) G M_\star C_\nu$, and $J_{1/2}(k_j x) = (2/\pi k_j \ell)^{1/2} \sin(k_j \ell)$. Requiring $J_{1/2}$ to vanish at $\ell_{\rm d} \equiv (GM_\star r_{\rm d})^{1/2}$ then means that $k_j = j\pi/\ell_{\rm d}$ and

$$g(t) = \sum_{j=1}^{\infty} A_j \left(\frac{2}{\pi}\right)^{1/2} \sin\left(\frac{j\pi\ell}{\ell_{\rm d}}\right) e^{-j^2 t/\tau} \tag{9}$$

where $\tau \equiv (\kappa \ell_{\rm d}/\pi)^2 = (4/3\pi^2) r_{\rm d}^2/\nu_{\rm d}$ is a characteristic viscous timescale with $\nu_{\rm d} \equiv \nu(r_{\rm d})$. Next assume that at $t = 0$, the initial disk surface density varies as $\sigma = \sigma_{\rm d}(r_{\rm d}/r)$ so that the couple reads

$$g(0) = -3\pi\sigma_{\rm d}\nu_{\rm d}\ell = 6\sigma_{\rm d}\nu_{\rm d}\ell_{\rm d} \sum_{j=1}^{\infty} \frac{(-1)^j}{j} \sin\left(\frac{j\pi\ell}{\ell_{\rm d}}\right) \tag{10}$$

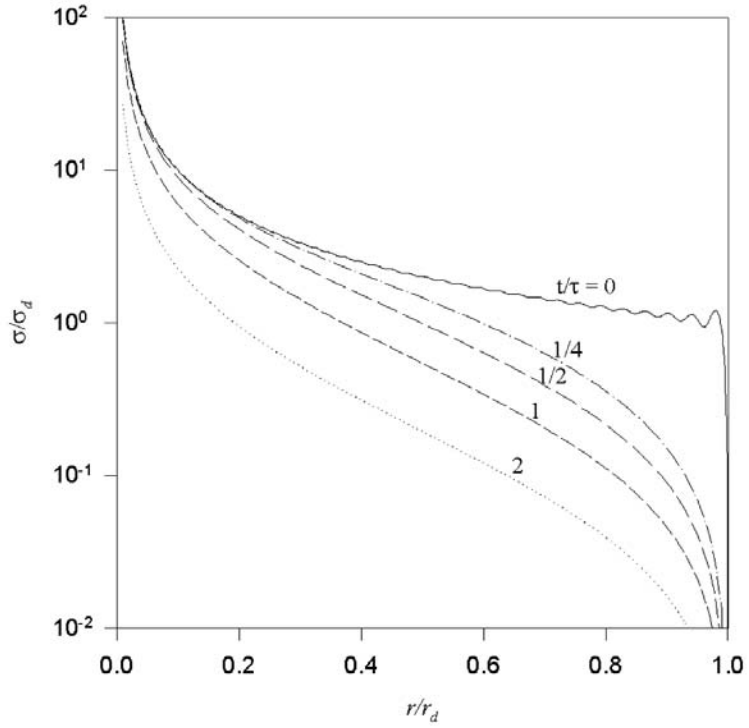

Figure 2. Surface density profile of a $n = 1$ disk are various times. The initial density profile is taken to be $1/r$. The spurious oscillations near $r/r_{\rm d} = 1$ are due to taking a finite (100) number of terms.

where a Fourier expansion for ℓ over the range $-\ell_{\rm d} < \ell < \ell_{\rm d}$ has been employed (*e.g.*, Weast and Astle, 1980). Comparison of eqns. 9 and 10 evaluated at $t = 0$ reveals that $(2/\pi)^{1/2} A_j = (-1)^j 6\sigma_{\rm d}\nu_{\rm d}\ell_{\rm d}/j$.

The surface density as a function of time is then found from

$$\sigma(t) = -\frac{g}{3\pi\nu\ell} = \frac{2}{\pi}\sigma_{\rm d}\left(\frac{\ell_{\rm d}}{\ell}\right)^3 \sum_{j=1}^{\infty} \frac{(-1)^j}{j} \sin\left(\frac{j\pi\ell}{\ell_{\rm d}}\right) e^{j^2 t/\tau} \tag{11}$$

and is displayed in Figure 2. The time varying mass and angular momentum of the disk are

$$\begin{aligned} M_{\rm d} &= 2\pi \int_0^{r_{\rm d}} \sigma r \, {\rm d}r = -\frac{4}{3}\frac{r_{\rm d}}{GM_\star\nu_{\rm d}} \int_0^{\ell_{\rm d}} g \, {\rm d}\ell \\ &= \frac{8}{\pi}\sigma_{\rm d} r_{\rm d}^2 \sum_{j=1}^{\infty} \left[1 - (-1)^j\right] \frac{e^{-j^2 t/\tau}}{j^2} \end{aligned} \tag{12}$$

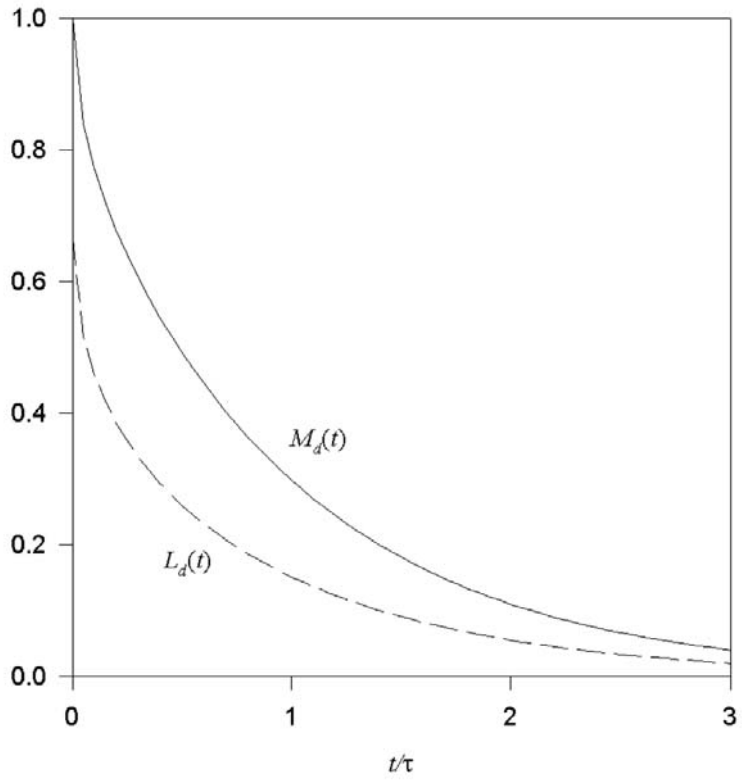

Figure 3. Time varying mass and angular momentum of the $n = 1, k = 1$ disk, normalized to $M_{\rm d}(0)$ and $M_{\rm d}(0)\ell_{\rm d}$, respectively.

and

$$L(t) = 2\pi \int_0^{r_{\rm d}} \sigma r \ell \, {\rm d}r = -\frac{4}{3}\frac{r_{\rm d}}{GM_\star \nu_{\rm d}} \int_0^{\ell_{\rm d}} g\ell \, {\rm d}\ell = \frac{8}{\pi}\sigma_{\rm d} r_{\rm d}^2 \ell_{\rm d} \sum_{j=1}^{\infty} \frac{e^{-j^2 t/\tau}}{j^2} \quad (13)$$

and are shown in Figure 3. The necessary mass to remove L by exiting at $r_{\rm d}$ is $M_0(t) = L(t)/\ell_{\rm d}$. On the other hand, the mass of the disk external to the semimajor axis a_0 at which the planet forms is

$$M_{\rm ext} = 2\pi \int_{a_0}^{r_{\rm d}} \sigma r \, {\rm d}r = \frac{8}{\pi}\sigma_{\rm d} r_{\rm d}^2 \sum_{j=1}^{\infty} \left[1 - (-1)^j \cos\left(\frac{j\pi \ell_{a_0}}{\ell_{\rm d}}\right)\right] \frac{e^{-j^2 t/\tau}}{j^2} \quad (14)$$

where $\ell_a \equiv \sqrt{GM_\star a}$. Substituting M_0 and $M_{\rm ext}$ into eq. 6 yields the mapping,

$$\left(\frac{a}{a_0}\right)^{1/2} = 1 + \frac{M_0 - M_{\rm ext}}{M_{\rm p}} \left(\frac{r_{\rm d}}{a_0}\right)^{1/2}$$

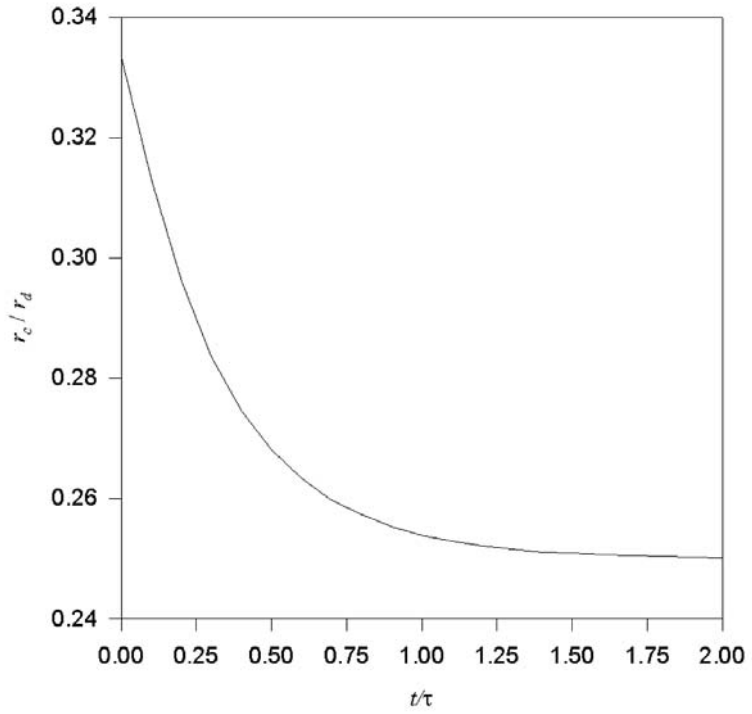

Figure 4. Time variation of the critical distance during the evolution of the $n = 1, k = 1$ disk.

$$= 1 + \frac{8}{\pi}\left(\frac{\sigma_{\rm d} r_{\rm d}^2}{M_{\rm p}}\right)\left(\frac{r_{\rm d}}{a_0}\right)^{1/2} \sum_{j=1}^{\infty} (-1)^j \cos\left(\frac{j\pi \ell_a}{\ell_{\rm d}}\right) \frac{e^{-j^2 t/\tau}}{j^2} \qquad (15)$$

For the power law disk in section 2 with $k = 1$ and $r_0 = r_{\rm d}$, we have $M_{\rm d} = 2\pi\sigma_{\rm d} r_{\rm d}^2$, $C_1 = 2/3$, $L_{\rm d} = 2/3 M_{\rm d} \ell_{\rm d}$, $M_0 = 2/3 M_{\rm d}$, $M_{\rm ext} = [1 - (a_0/r_{\rm d})] M_{\rm d}$, and $r_{\rm c} = r_{\rm d}/3$. In this case, eq. 5 reads

$$\begin{aligned}\left(\frac{a}{r_{\rm d}}\right)^{1/2} &= \left(\frac{a_0}{r_{\rm d}}\right)^{1/2} + \left[\frac{a_0}{r_{\rm d}} - \frac{r_{\rm c}}{r_{\rm d}}\right]\left(\frac{M_{\rm d}}{M_{\rm p}}\right) \\ &= \left(\frac{a_0}{r_{\rm d}}\right)^{1/2} + \left(\frac{2\pi\sigma_{\rm d} r_{\rm d}^2}{M_{\rm p}}\right)\left[\frac{a_0}{r_{\rm d}} - \frac{1}{3}\right] \end{aligned} \qquad (16)$$

Multiplying both sides by $(r_{\rm d}/a_0)^{1/2}$ reproduces the $t = 0$ value of eq. 15 since (*e.g.*, Weast and Astle, 1980)

$$\sum_{j=1}^{\infty} \frac{(-1)^j}{j^2} \cos\left(\frac{j\pi \ell_{a_0}}{\ell_{\rm d}}\right) = \frac{\pi^2}{4}\left(\frac{a_0}{r_{\rm d}} - \frac{1}{3}\right). \qquad (17)$$

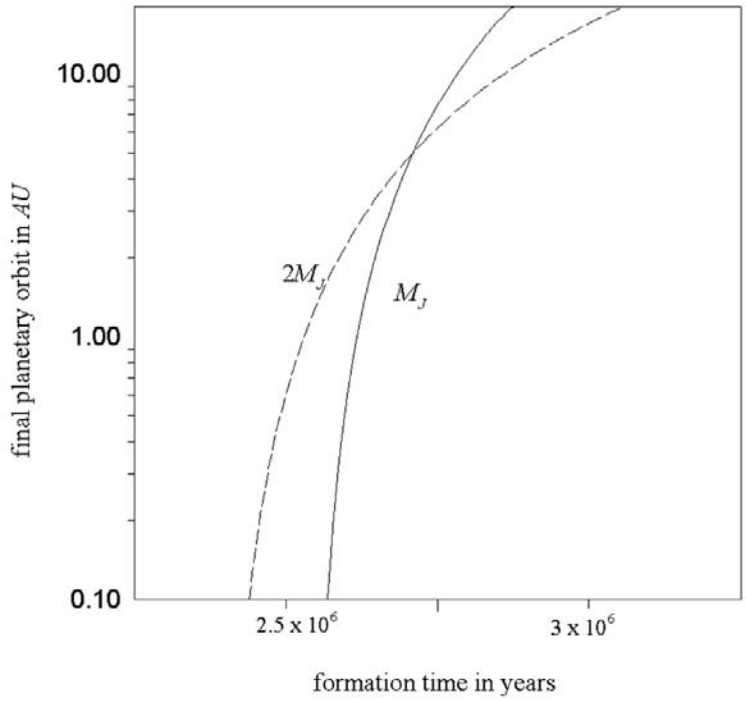

Figure 5. Final semimajor axis of planets following migration as a function of the planet formation time.

The location of the critical distance r_c is found from $2\pi \int_0^{r_c} \sigma r \, dr = M_d - M_0$ which leads to the condition

$$\sum_{j=1}^{\infty} \frac{(-1)^j}{j^2} \cos\left(\frac{j\pi \ell_c}{\ell_d}\right) e^{j^2 t/\tau} = 0 \tag{18}$$

From eq. 17 we conclude that at $t = 0$, $r_c/r_d = 1/3$ as required, but as $t \gg \tau$, $\ell_c/\ell_d \longrightarrow 1/2$, $r_c/r_d \longrightarrow 1/4$. The time variation of $r_c(t)$ is shown in Figure 4. Note for instance, if $a_0 = 5$ AU and $r_d = 18$ AU, an early forming planet will be inside $r_c \sim 6$ AU and suffer orbital decay, while a late forming one will be outside $r_c \sim 4.5$ AU and migrate outward.

4. Comparison With Numerical Experiments

A02 assumed that photoevaporation removes disk material at a rate $\dot{\sigma} \propto 1/r$ starting at $r_1 = 5$ AU out to $r_2 = 33.3$ AU. This implies a mass loss rate $\dot{M} \propto r_2 - r_1$ and an angular momentum loss rate $\dot{L} \propto (2/3)(r_2^{3/2} - r_1^{3/2})$. To replace this with an edge preserving the same relative rates, we set $r_d^{1/2} = (2/3)(r_2^{3/2} - r_1^{3/2})/(r_2 -$

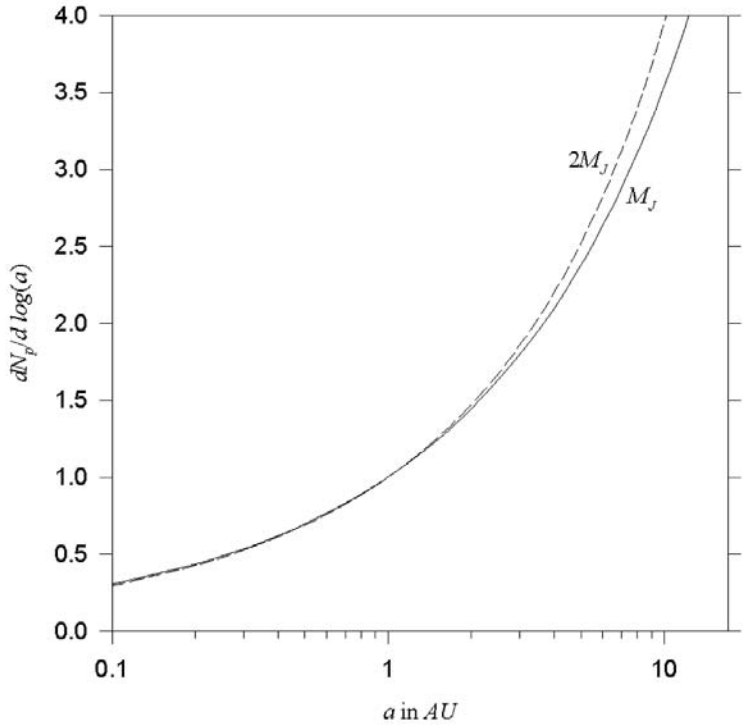

Figure 6. Predicted number of extra-solar planets per logarithmic interval in semimajor axis. The absolute number of planets is arbitrary, and has been normalized to unity at 1 AU.

r_1) for which $r_{\rm d} \sim 18$ AU. We set the initial disk mass to $M_{\rm d}(0) = 0.1 M_\odot$ and the boundary viscosity to $\nu_{\rm d} = 4.16 \times 10^{13}$ cm^2/s, which in our $n = 1$ model[2] implies $\tau = 7.52 \times 10^6$ years. Figure 5 shows the final positions of planets with one Jupiter mass (solid line) and two Jupiter masses (dashed line) as a function of their formation time, assuming all formed initially at 5 AU. Off the top of the figure, planets exit the disk at 18 AU; off the bottom are planets that migrate closer than 0.1 AU of the star. The results are qualitatively similar to that of A02, with the curve steepening at small radii/early formation times. Following their discussion further, if the additional assumption is made of a uniform rate of planet formation, the predicted number of planets per logarithmic interval in radius is $\mathrm{d}N_{\rm p}/|\mathrm{d}\log(a)| \propto \mathrm{d}t/|\mathrm{d}\log(a)| = (a/a_0)/|\mathrm{d}(a/a_0)/\mathrm{d}t| = 2(a/a_0)^{1/2}/|\mathrm{d}(a/a_0)^{1/2}/\mathrm{d}t|$. The time derivative of eq. 15 is simply

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{a}{a_0}\right)^{1/2} = -\frac{8}{\pi}\left(\frac{\sigma_{\rm d} r_{\rm d}^2}{M_{\rm p}}\right)\left(\frac{r_{\rm d}}{a_0}\right)^{1/2} \sum_{j=1}^{\infty}(-1)^j \cos\left(\frac{j\pi \ell_a}{\ell_{\rm d}}\right)\frac{e^{-j^2 t/\tau}}{\tau} \qquad (19)$$

[2] Armitrage *et al.* (2002) employ an $n = 3/2$ viscosity law.

Dividing eq. 15 by the absolute value of eq. 19 gives the form of the predicted distribution aside from a constant as shown in Figure 6. Again this result is similar to the distribution obtained by A02 and indicates that much of the behavior is indeed captured by the simple conservation arguments presented here. The mapping procedure outlined can be economically applied to a large number of planet-disk models, including variations in the planet-disk mass ratio, initial position of the planet, non-uniform formation frequencies, the position of the photo-dissociation edge, and the starting density profile of the disk. Such predictions can help us to explore the linkage between observed extra-solar planet statistics and their precursor circumstellar disks. Of course, there are complications that have so far been omitted in this approach that may ultimately require numerical modeling to explore, such as multiple planets (Kley *et al.*, 2004), time changes the photodissociation boundary (*e.g.*, Matsuyama *et al.*, 2003), gas leakage through a gap (Bryden *et al.*, 1999), and/or gas accretion by the migrating planets (e.g., Kley, 1999; and Lubow *et al.*, 1999). Nevertheless, it may be possible to incorporate simplified versions of these in a mapping approach as well, and work to that end is underway.

Acknowledgements

The author is grateful for the support and hospitality of the conference organizers. This work was also supported in part by NASA's Planetary Geology and Geophysics Program.

References

Armitage, P. J., Livio, M., Lubow, S. H., and Pringle, J. E.: 2002, *M.N.R.A.S.* **334**, 248.
Bryden, G., Chen, X., Lin, D. N. C., Nelson, R.,P., Papaloizou, J. C. B.: 1999, *Astrophys. J.* **514**, 344.
Hollenbach, D. J., Yorke, H. W., and Johnstone, D.: 2000, *in* V. Mannings, A. P. Boss, and S. S. Russell (eds.), *Protostars and Planets IV*, Univ. Arizona Press, Tucson, pp. 401–428.
Kley, W.: 1999, *M.N.R.A.S.* **303**, 696.
Kley, W., Peitz, J., Bryden, G.: 2004, *Astron. & Astrophys.* **414**, 735.
Lin, D. N. C., and Papaloizou, J. C. B.: 1986, *Astrophys. J.* **309**, 846.
Lubow, S. H., Seibert, M., Artymowicz, P.: 1999, *Astrophys. J.* **526**, 100.
Lynden-Bell, D., and Pringle, J. E.: 1974, *M.N.R.A.S.* **168**, 603.
Matsuyama, I., Johnstone, D., Murray, N.: 2003, *Astrophys. J. Lett.* **585**, L143.
Shu, F. H., Johnstone, D., and Hollenbach, D.: 1993, *Icarus* **106**, 92.
Trilling, D. E., Lunine, J. I., and Benz, W., 2002, *Astron.&Astrophys.* **394**, 241.
Ward, W. R.: 2003, *Lunar & Planetary Sci.* **XXXIV**, 1736.
Weast, R. C., and Astle, M. J.: 1980, *CRC Handbook of Chemistry and Physics, Chemical Rubber Publishing Co.*, Boca Raton, p. A-104.

On Multiple Planet Systems

Theory of Planets in resonant Orbits

W. Kley
Astronomie und Astrophysik, Abt. Computational Physics, Universität Tübingen, D-72076 Tübingen, Germany

Abstract. We study the evolution of a system consisting of two protoplanets still embedded in a protoplanetary disk. Results of two different numerical approaches are presented. In the first kind of model the motion of the disk material is followed by fully viscous hydrodynamical simulations, and the planetary motion is determined by N-body calculations including exactly the gravitational potential from the disk material. In the second kind we only solve the N-body part and add additional analytically given forces which model the effect of the torques of the disk. This type of modeling is of course orders of magnitudes faster than the full hydro-model. Another advantage of this two-fold approach is the possibility of adjusting the otherwise unknown parameters of the simplified model.

The results give very good agreement between the methods. Using two different initial setups for the planets and disk, we obtain in the first case a resonant trapping into the 3:1 resonance, and in the second case a trapping into the 2:1 resonance. Resonant capture leads to a rise in the eccentricity and to an alignment of the spatial orientation of orbits. The characteristics of the numerical results agree very favorably with those of 3 observed planetary systems (Gl 867, HD 82943, and 55 Cnc) known to be in mean motion resonances.

1. Introduction

Since the first discovery in 1995, during the last years the number of detected extrasolar planets orbiting solar type stars has risen up to about 100 (for an always up-to-date list see eg. `http://www.obspm.fr/encycl/encycl.html` by J. Schneider). It has been found that among those there are 11 systems with 2 or more planets. With further observations to come, this number may still increase, as for some systems trends in the radial velocity curve have been found. Among these multiple planet extrasolar systems there are now 3 confirmed cases, Gl 876 (Marcy et al., 2001), HD 82943 (the *Coralie Planet Search Programme*, ESO Press Release 07/01), 55 Cnc (Marcy et al., astro-ph/0207294, 2002) where the planets orbit their central star in a low order *mean motion resonance*, where the orbital periods have nearly exactly the ratios 2:1 or 3:1. The parameters of these planetary systems are displayed below in Table I. This implies that about 1/4 of planetary systems, or even more, may be in resonance, a fraction which is even higher if secular resonances, as for example present in υ And (Butler et al., 1999), are also taken into account.

The formation of such resonant planetary systems can be understood by considering the joint evolution of proto-planets together with the protoplanetary disk from which they formed. By local linear analysis it was shown that the gravitational interaction of a protoplanet with the disk may lead to torques resulting in

the migration of a planet (Goldreich and Tremaine, 1980; Lin and Papaloizou, 1986; Ward, 1997; Tanaka et al., 2002). Additionally, for planetary masses of around one Jupiter mass, gap formation as result of angular momentum transfer between the (viscous) disk and the planet was considered (Lin and Papaloizou, 1993). Fully non-linear hydrodynamical calculations (Kley, 1999; Bryden et al., 1999; Lubow et al., 1999; Nelson et al., 2000) for Jupiter sized planets confirmed the expectations and showed clearly that this interaction leads to: *i)* the formation of spiral shocks waves in the disk, whose tightness depends on the sound-speed in the disk, *ii)* an annular gap, whose width is determined by the equilibrium between gap opening tidal torques and gap closing viscous and pressure forces. *iii)* an inward migration on a timescale of 10^5 y for typical disk parameter in particular disk masses corresponding to that of the minimum mass solar nebula, *iv)* a possible mass growth after gap formation up to about 10 $M_{\rm Jup}$ when finally gravitational torques overwhelm, and finally *v)* a prograde rotation of the planet.

Recently, these single planet calculations were extended to calculations with multiple planets. Those have shown already (Kley, 2000; Snellgrove et al., 2001; Nelson and Papaloizou, 2002) that during the early evolution of protoplanetary systems, when the planets are still embedded in the disk, different migration speeds may lead an approach of the planets and eventually to resonant capture.

Here we present new numerical calculations treating the evolution of a pair of two embedded planets in disks. We consider both, fully hydrodynamic and simplified N-body calculations to model the evolution. In the first approach, the motion of the disk is followed by solving the full time dependent Navier-Stokes equations simultaneously with the motion of the planets. Here the motion of the planet is determined by the gravitational potential of the other planet, the star, and that of the disk. In the latter approach we take a simplified assumption and perform 3-body (star and 2 planets) calculations augmented by additional (damping) forces which take the gravity of the disk approximately into account. This approach was first adopted by (Snellgrove et al., 2001) and (Nelson and Papaloizou, 2002), and later used by (Lee and Peale, 2002). These two approaches allow a direct comparison of the methods, and will enable us to determine in detail the damping constants required for the simpler (and much faster) second model.

2. The Observations

The basic orbital parameters of the 3 systems in mean motion resonance are presented in Table I. Two of them, Gl 876 and HD 82943 are in a nearly exact 2:1 resonance. In both cases we note that the outer planet is the more massive one by a factor of about two (HD 82943), and more than three (Gl 876). The eccentricity of the inner (less massive) planet is larger than that of the outer one. For the system Gl 876 the alignment of the orbits is such that the two periastrae are pointing in nearly the same direction. For the system HD 82943 these data have not been

Table I. The orbital parameters of the 3 system known to be in a mean motion resonance. P denotes the orbital period, $M \sin i$ the mass of the planets, a the semi-major axis, e the eccentricity, ω the direction of periastron, and M_* the mass of the central star. It should be noted, that the orbital elements for shorter period planets are due secular time variations. Thus in principle one should always state the epoch corresponding to these osculating elements, see eg. (Laughlin and Chambers, 2001).

System	Nr	Per [d]	$M \sin i$ $M_{\rm Jup}$	a [AU]	e	ω [deg]	M_* $M_\odot$
Gl 876 (2:1)	c	30.1	0.77	0.13	0.24	159	0.32
	b	61.02	2.4	0.21	0.04	163	
HD 82943 (2:1)	b	221.6	0.88	0.73	0.54	138	1.05
	c	444.6	1.63	1.16	0.41	96	
55 Cnc (3:1)	b	14.65	0.84	0.11	0.02	99	0.95
	c	44.26	0.21	0.24	0.34	61	
	d	5360	4.05	5.9	0.16	201	

clearly identified, due to the much longer orbital periods, but do not seem to very different from each other. The last system, 55 Cnc, is actually a triple system. Here the inner two planets orbit the star very closely and are in a 3:1 resonance, while the additional, more massive planet orbits at a distance of several AU.

3. The Models

It is our goal to determine the evolution of protoplanets still embedded in their disks. To this purpose we employ two different methods which supplement each other. Firstly, a fully time-dependent hydrodynamical model for the joint evolution of the planets *and* disk is presented. Because the evolutionary time scale may cover several thousands of orbits these computations require sometimes millions of time-steps, which translates into an effective computational times of up to several weeks.

Because often the main interest focuses only on the orbital evolution of the planet and not so much on the hydrodynamics of the disk, we perform additional simplified 3-body computations. Here the gravitational forces are augmented by additional damping terms designed in such a way as to incorporate in a simplified way the gravitational influence of the disk. Through a direct comparison with the hydrodynamical model it is then possible to infer directly the necessary damping forces.

Table II. Planetary and disk parameters of the models. The mass of the Planet is given in Jupiter masses ($M_{\rm Jup} = 10^{-3} M_\odot$), the viscosity in dimensionless units, the disk mass located between $r_{\rm min}$ and $r_{\rm max}$ in solar masses, and the minimum and maximum radii in AU.

Model	Mass1	Mass2	Viscosity	$M_{\rm disk}$	$r_{\rm min}$	$r_{\rm max}$
A	3	5	$\alpha = 10^{-2}$	0.01	1	30
B	1 (Var)	1 (Var)	$\nu =$ const.	0.01	1	20

3.1. FULL HYDRODYNAMICS

The first set of coupled hydrodynamical-N-body models are calculated in the same manner as described in detail in (Kley, 1998), (Kley, 1999) for single planets and in (Kley, 2000) for multiple planets, and the reader is referred to those papers for more details on the computational aspect of the simulations. Other similar models, following explicitly the motion of single and multiple planets in disks, have been presented by (Nelson et al., 2000; Bryden et al., 2000; Snellgrove et al., 2001). During the evolution material is taken out from the centers of the Roche-lobes of the two planets, which is monitored and assumed to have been accreted onto the two planets. We present two runs: one (model B) where the mass is added to the planet, and another one (model A) where this mass is not added to the dynamical mass of the planets, i.e they always keep their initial mass. They are allowed to migrate (change their semi-major axis) through the disk according to the gravitational torques exerted on them. This assumption of constant planet mass throughout the computation is well justified, as the migration rate depends only weakly on the mass of the planet (Nelson et al., 2000). The initial hydrodynamic structure of the disk, which extends radially from $r_{\rm min}$ to $r_{\rm max}$, is axisymmetric with respect to the location of the star, and the surface density scales as $\Sigma(r) = \Sigma_0 r^{-1/2}$. The material orbits initially on purely Keplerian orbits $v_r = 0$, $v_\varphi = GM_*/r^{1/2}$, and the fixed temperature law follows from the constant vertical height $H/r = 0.05$ and is given by $T(r) \propto r^{-1}$. The kinematic viscosity ν is given by an α-description $\nu = \alpha c_s H$, with the sound speed c_s. We present two models:

i) Model A, having a constant $\alpha = 0.01$, which may be on the high side for protoplanetary disks but allows for a sufficiently rapid evolution of the system to identify clearly the governing physical effects. The two embedded planets have a mass of 3 and 5 $M_{\rm Jup}$ and are placed initially at 4 and 10 AU, respectively.

ii) Model B with a constant ν, equivalent to an $\alpha = 0.004$ at 1 AU. Here the two embedded planets each have an initial masses of 1 $M_{\rm Jup}$, and are placed initially at 1 and 2 AU, respectively. This model is in fact a continuation of the one presented in (Kley, 2000).

All the relevant model parameters are given in Table II.

3.2. DAMPED N-BODY

As pointed out, the full hydrodynamical evolution is computationally very time consuming because ten-thousands of orbits have to be calculated. Hence, we perform also pure N-body calculations of a planetary system, consisting only of a star and two planets. The influence of the surrounding disk is felt here only through additional (damping) forces. For those we assume, that they act on the semi-major axis a and the eccentricity e of the two planets through explicitly specified relations $\dot{a}(t)$ and $\dot{e}(t)$, which may depend on time. Here, as customary in solar system dynamics, it is assumed that the motion of the two planets can be described at all times by approximate Kepler ellipses where the time-dependent parameters $a(t)$ and $e(t)$ represent the values of the osculating orbital elements at the epoch t.

This change of the actual semi-major axis a and the eccentricity e caused be the gravitational action of the disk can be translated into additional forces changing directly the position $\dot{\mathbf{x}}$ and velocity $\dot{\mathbf{u}}$ of the planets. In our implementation we follow exactly (Lee and Peale, 2002), who give the detailed explicit expressions for these damping terms in their appendix. As a first test of the method we recalculated their model for GJ 876 and obtained identical results.

Using the basic idea of two planets orbiting inside of a disk cavity (see Fig. 1 below), we only damp a and e of the outer planet. Here we choose a general given functional dependence of the form $g(t) = g_0 \exp[-(t/t_0)^p]$, with $g \in (a, e)$, where a_0 and e_0 are just the initial values. The values of the exponent p and the timescale t_0 are adjusted to match the results of the full hydrodynamic calculations. Comparative results of the two methods are displayed in Section 4.3.

4. Results

The basic evolutionary sequence of two planets evolving simultaneously with the disk has been calculated and described by (Kley, 2000) and (Bryden et al., 2000). Before concentrating on details of the resonant evolution we first summarize briefly the main results.

4.1. FULL HYDRODYNAMIC EVOLUTION: OVERVIEW

At the start of the simulations both planets are placed into the axisymmetric disk, where the density is initialized such that in addition to the radial density profile partially opened gaps are superimposed. Upon starting the evolution the two main effects are:

a) As a consequence of the accretion of gas onto the two planets the radial regime in between them will be depleted in mass and finally cleared. This phase typically takes only a few hundred orbital periods. At the same time the region

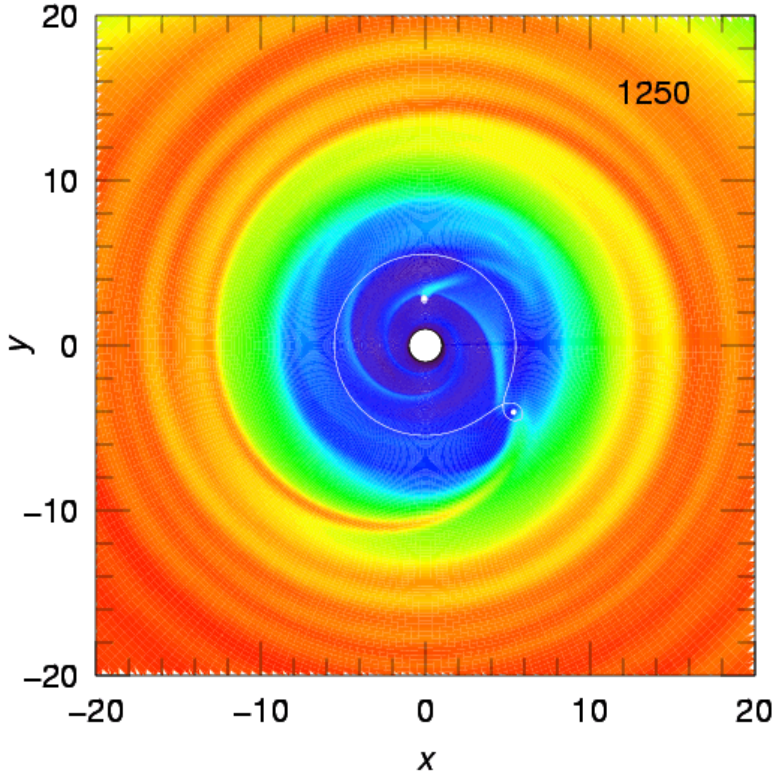

Figure 1. Overview of the density distribution of model A after 1250 orbital periods of the inner planet. Higher density regions are brighter and lower ones are darker. The star lies at the center of the white inner region inside of $r_{\rm min} = 1$. The location of the two planets is indicated by the white dots, and the Roche-lobe sizes are also drawn. Clearly seen are the irregular spiral wakes generated by the planets. Only outside of the 2nd planet the regular inter-twined two spiral arms are visible.

interior of the inner planet will lose material due to accretion onto the central star. Thus, after an initial transient phase we typically expect the configuration of two planets orbiting inside of an inner cavity of the disk, see Fig. 1, and also (Kley, 2000).

b) After initialization the planets quickly (within a few orbital periods) create non-axisymmetric disturbances, the spiral features, in the disk. In contrast to the single planet case these are no longer stationary in time, because there is no preferred rotating frame. The gravitational torques exerted on the two planets by those density perturbations induce a migration process for the planets.

Now, the different radial location of the planets within the cavity has a distinct influence on their subsequent evolution. As a consequence of the clearing process the inner planet is no longer surrounded by any disk material and thus cannot grow any further in mass. In addition it cannot migrate anymore, because there is no torquing material in its vicinity. The outer planet on the other hand still has all

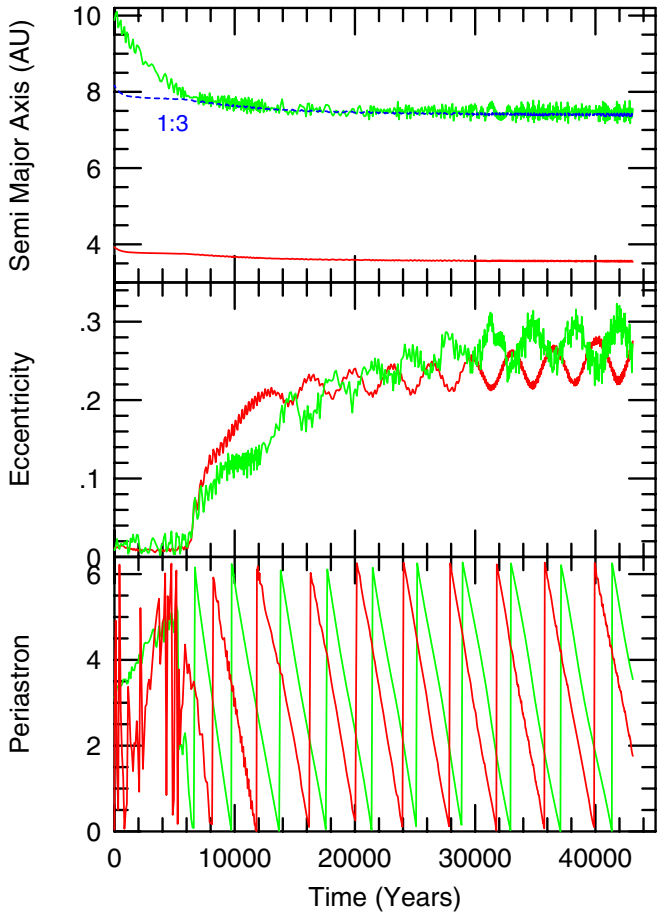

Figure 2. The semi-major axis (AU), eccentricity and position of the periastron of the orbit versus time for Model A. In this example, the planets have fixed masses of 3 and 5 $M_{\rm Jup}$, and are placed initially at 4 and 10 AU, respectively. In the beginning, after the inner gap has cleared, only the outer planet migrates inward, and the eccentricities of both planets remain relatively small, less than ≈ 0.02. After about 6000 years the outer planet has reached a radius with a period exactly three times that of the inner planet. The periodic gravitational forcing leads to a capture of the inner planet into a 3:1 resonance by the outer one. This is indicated by the dark reference line (labeled 3:1), which marks the location of the 3:1 resonance with respect to the inner planet. Upon resonant capture the eccentricities grow, and the orbits librate with a fixed relative orientation of $\Delta\omega = 110°$.

the material of the outer disk available, which exerts negative (Linblad) torques on the planet. Hence, in the initial phase of the computations we observe an inwardly migrating outer planet and a stalled inner planet with a constant semi-major axis, see the first 5000 y in the top panel of Fig. 2.

This decrease in separation causes an increase of the gravitational interaction between the two planets. When the ratio of the orbital periods of the planets has

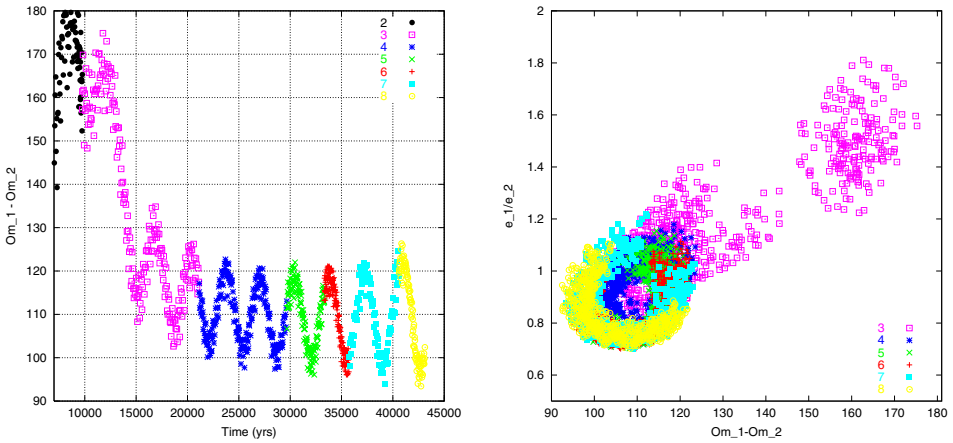

Figure 3. Left: The difference in the direction of the periastron, $\Delta\omega = \omega_1 - \omega_2$, of the two planets vs. time. Right: Ratio of eccentricities e_1/e_2 versus periastron difference. The indices 1,2 refer to the inner and outer planet, respectively. The color and symbol coding is identical for the left and right panel. During the evolution into resonance the dots are 'captured' to eventually circle around the equilibrium value $\Delta\omega = 110^\circ$ and $e_1/e_2 = 0.9$.

reached the fraction of two integers, i.e. they are in a mean motion resonance, this may lead to a resonant capture of the inner planet by the outer one. Whether this happens of not depends on the physical conditions in the disk (eg. viscosity) and the orbital parameters of the planets. If the migration speed is too large for example, there may not be enough time to excite the resonance, and the outer planet just continues its migration process, see eg. (Haghighipour, 1999). Also, if the initial eccentricities are too small, then there may be no capture as well.

4.2. 3:1 Resonance: Model A

In model A this capture happens at $t \approx 6000$ when the outer planet captures the inner one in a 3:1 resonance (see dark line in top panel of Fig. 2). From that point on, the outer planet, which is still driven inward by the outer disk material, will also be forcing the inner planet to migrate inwards. The typical time evolution of the orbital elements, semi-major axis (a), eccentricity (e) and direction of the periastron (ω), of such a case are displayed in Fig. 2 for model A.

We summarize the following important features of the evolution after resonant capture:

a) The inner planet begins to migrate inward as well, forced in by the outer planet. Thus both planets migrate inward simultaneously, always retaining their resonant configuration. As a consequence, the migration speed of the outer planet slows down, and their radial separation declines.

b) The eccentricities of both planets grow initially very fast and then settle to oscillating quasi-equilibrium values which change slowly on a secular time

scale. This slow increase of the eccentricities on the longer timescale is caused by the growing gravitational forces between the planets, due to the decreasing radial distance of the two planets on their inward migration process.

c) The ellipses (periastrae) of the planets librate at a constant angular speed. Caused by the resonance, the speed of libration for both planets is identical, which can be inferred from the parallel lines in the bottom panel of Fig. 2. The orientation of the orbits is phase locked with a constant separation of the periastrae by a fixed phase-shift $\Delta\omega$.

A mean motion resonance arises when the resonant angles that can be defined for the problem themselves librate, see eg. (Nelson and Papaloizou, 2002). This is indeed the case in the present simulations. More detail of the capture into resonance and the subsequent libration of the orbits is illustrated in Fig. 3 for model A. It is seen (left panel) that the difference of the periastrae settles to the fixed average value of $\Delta\omega = 110°$, a libration amplitude of about 15°, and libration period of about 3000 y. The right panel shows the evolution in the e_1/e_2 vs. $\Delta\omega$ using the same color and symbol coding coding. During the initial process of capturing the points (open squares) approach the final region from the top right region of the diagram. At later times the points circle around the equilibrium point. We note that additional models, not displayed here, with different planet masses and viscosities all show approximately the same shift in $|\Delta\omega|$, **if** capture occurs into 3:1 resonance.

4.3. 2:1 Resonance: Model B

The second model setup is taken directly from (Kley, 2000). Here we continued exactly that model for a little longer time, to infer some more characteristics of the intrinsic dynamics of that planetary system. The evolution of the orbital elements a and e is displayed in Fig. 4. Here the planets are placed on initially tighter orbits with an semi-major axis ratio of only 2. The initial orbital evolution is similar to the first model, i.e. an inwardly migrating outer and a stalled inner planet. However, caused by the reduced initial radial distance higher resonances are not available, and the resonant capture occurs into the 2:1 resonance, see reference line in left panel. The eccentricities of both planets rise again upon capture but this time the mass of the inner planet $m_1 = 1M_{\rm Jup}$ is, due to the explained starvation, much smaller than that of the outer one $m_2 = 3.1M_{\rm Jup}$. This leads to to much larger rise in eccentricity, yielding a ratio $e_1/e_2 \approx 4$. In Fig. 5 the alignment of the orbits is indicated. This time, as seems typical for this type of 2:1 resonances (Snellgrove et al., 2001; Lee and Peale, 2002), the separation in the periastrae is centered around zero, $\Delta\omega = 0$, with a libration amplitude of up to about 20°.

For comparison and test, we modeled the evolution of the model B also using the 3-body method, which is briefly outlined above and compared this to the full hydrodynamic evolution. As outlined in Sect. 3.2, only the outer planet is damped

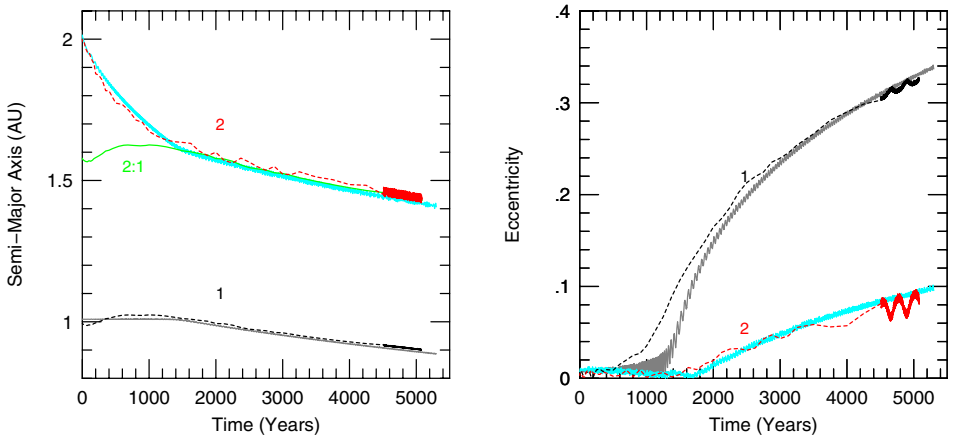

Figure 4. The evolution of semi-major axis (left) and eccentricity (right) for model B. The planets had an initial radial location of 1 and 2 AU, and masses of 1 $M_{\rm Jup}$ each, which were allowed to increase during the computation. The results of the full hydrodynamic evolution are shown by the dashed lines. A reference line, with respect to the inner planet, indicating the location of the 2:1 resonance is shown. Before $t = 4500$ only very few data points are plotted, thereafter they are spaced much more densely, which explains the different looking curves. The solid curves are obtained using the simplified damped 3-body evolution as described in the text.

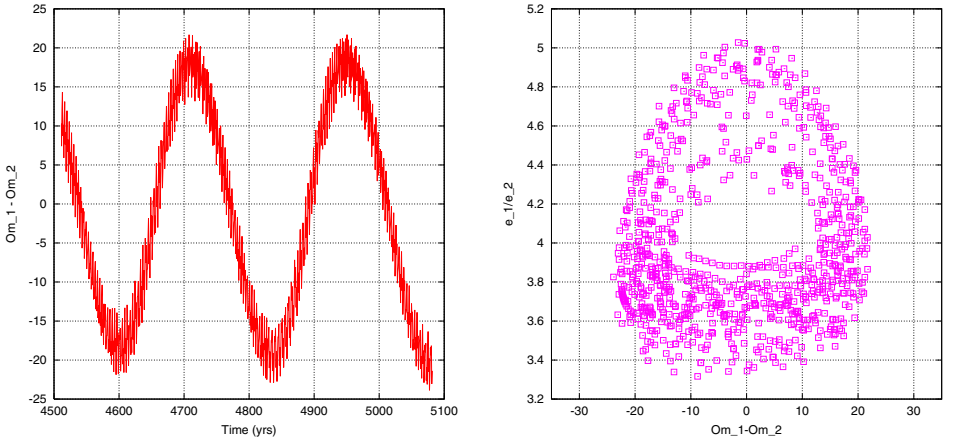

Figure 5. Left: The difference in the direction of the periastrae of the two planets vs. time. Right: Ratio of eccentricities versus periastron difference. The data points are spaced equally in time with a distance of approximately $\delta t = 3/4$ years. Shown is only the very last section of the evolution of model B, from 4500 to about 5100 y, which covers nearly two and a half libration periods. In this case of a 2:1 resonance, the capture leads to a complete alignment of the orbits with $\Delta\omega = 0$.

in its semi-major axis a and eccentricity e. It turned out that for a good agreement the damping time scale t_0 is identical for a and e. Here in model B, we used $t_0 = 75,000$ y and $p = 0.75$ to obtain the displayed results, and to reach the agreement with the full hydrodynamical evolution.

These additional results are also displayed in Fig. 4 with the solid lines. For the semi-major axis (left panel) the agreement is very good indeed. The obtained eccentricities are in good agreement as well. The only difference is the lack in eccentricity oscillations in the simplified damped 3-body model, which may be caused by a change in eccentricity damping in the full model in contrast to the simplified version.

5. Summary and Conclusion

We have performed full hydrodynamical calculations simulating the joint evolution of a pair of protoplanets together with their surrounding protoplanetary disk, from which they originally formed. The focus lies on massive planets in the range of a few Jupiter masses. For the disk evolution we solve the Navier-Stokes equations, and the motion of the planets is followed using a 4th order Runge-Kutta method, considering their mutual interaction, the stellar and the disk's gravitational field. These results were compared to simplified (damped) N-body computations, where the gravitational influence of the disk is modeled through analytic damping terms applied to the semi-major axis and eccentricity.

We find that both method yield comparable results, if the damping constants in the simplified models are adjusted properly. These constant should be obtained from the full hydrodynamical evolution.

Two types of resonant situations are investigated. In the first case the initial radial separation of the two planets was sufficiently far that they were captured into a 3:1 resonance. In this case, the capture leads to an orbit alignment of $\Delta\omega = 110°$. In the second case (model B) the capture is in 2:1 with $\Delta\omega = 0°$. These difference in $\Delta\omega$ for varying types of resonances seems to be a robust and a generic feature, which is supported by the observations of all three resonating planets. Additionally, we find that the inner planet preferably has lower mass caused by the inner disk gap, and finally the lower mass planet should have a larger eccentricity. These findings are indeed seen in the observed 2:1 resonant planetary systems Gl867 and HD 82943.

References

Bryden, G., X. Chen, D. N. C. Lin, R. P. Nelson, and J. C. B. Papaloizou: 1999, 'Tidally Induced Gap Formation in Protostellar Disks: Gap Clearing and Suppression of Protoplanetary Growth'. *Astrophys. J.* **514**, 344–367.

Bryden, G., M. Różyczka, D. N. C. Lin, and P. Bodenheimer: 2000, 'On the Interaction between Protoplanets and Protostellar Disks'. *Astrophys. J.* **540**, 1091–1101.

Butler, R.P., Marcy, G.W., Fischer, D.A., Brown, T.M., Contos, A.R., Korzennik, S.G., Nisenson, P., and Noyes, R.W.: 1999, *Astrophys. J.* **526**, 916.

Goldreich, P. and S. Tremaine: 1980, 'Disk-satellite interactions'. *Astrophys. J.* **241**, 425–441.

Haghighipour, N.: 1999, 'Dynamical friction and resonance trapping in planetary systems'. *Mon. Not. R. Astron. Soc.* **304**, 185–194.

Kley, W.: 1998, 'On the treatment of the Coriolis force in computational astrophysics'. *Astron. & Astrophys.* **338**, L37–L41.

Kley, W.: 1999, 'Mass flow and accretion through gaps in accretion discs'. *Mon. Not. R. Astron. Soc.* **303**, 696–710.

Kley, W.: 2000, 'On the migration of a system of protoplanets'. *Mon. Not. R. Astron. Soc.* **313**, L47–L51.

Laughlin, G. and J. E. Chambers: 2001, 'Short-Term Dynamical Interactions among Extrasolar Planets'. *Astrophys. J. Letter* **551**, L109–L113.

Lee, M. H. and S. J. Peale: 2002, 'Dynamics and Origin of the 2:1 Orbital Resonances of the GJ 876 Planets'. *Astrophys. J.* **567**, 596–609.

Lin, D. N. C. and J. Papaloizou: 1986, 'On the tidal interaction between protoplanets and the protoplanetary disk. III - Orbital migration of protoplanets'. *Astrophys. J.* **309**, 846–857.

Lin, D. N. C. and J. C. B. Papaloizou: 1993, 'On the tidal interaction between protostellar disks and companions'. In: *Protostars and Planets III*. pp. 749–835.

Lubow, S. H., M. Seibert, and P. Artymowicz: 1999, 'Disk Accretion onto High-Mass Planets'. *Astrophys. J.* **526**, 1001–1012.

Marcy, G. W., R. P. Butler, D. Fischer, S. S. Vogt, J. J. Lissauer, and E. J. Rivera: 2001, 'A Pair of Resonant Planets Orbiting GJ 876'. *Astrophys. J.* **556**, 296–301.

Nelson, R. P. and J. C. B. Papaloizou: 2002, 'Possible commensurabilities among pairs of extrasolar planets'. *Mon. Not. R. Astron. Soc.* **333**, L26–L30.

Nelson, R. P., J. C. B. Papaloizou, F. . . Masset, and W. Kley: 2000, 'The migration and growth of protoplanets in protostellar discs'. *Mon. Not. R. Astron. Soc.* **318**, 18–36.

Snellgrove, M. D., J. C. B. Papaloizou, and R. P. Nelson: 2001, 'On disc driven inward migration of resonantly coupled planets with application to the system around GJ876'. *Astron. & Astrophys.* **374**, 1092–1099.

Tanaka, H., T. Takeuchi, and W. R. Ward: 2002, 'Three-Dimensional Interaction between a Planet and an Isothermal Gaseous Disk. I. Corotation and Lindblad Torques and Planet Migration'. *Astrophys. J.* **565**, 1257–1274.

Ward, W. R.: 1997, 'Protoplanet Migration by Nebula Tides'. *Icarus* **126**, 261–281.

A New Constraint on the Formation of Jupiter

Tobias Owen
Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive Honolulu, Hawaii 96822 USA

Abstract. Data from the mass spectrometer on the Galileo Probe indicate an approximately uniform enrichment (relative to hydrogen) of 3 ± 1 times the solar value for Ar, Kr, Xe, C, S and N on Jupiter. These data also reveal that $^{15}N/^{14}N = 2.3 \pm 0.3 \times 10^{-3}$ in Jupiter's nitrogen, indicating that this element reached the planet primarily in the form of N_2 . The ensuing requirements on the nature and total mass of the icy planetesimals needed to produce these results strongly suggest that a disk instability in the solar nebula could not have produced this giant planet.

1. Introduction

There are currently two principal schools of thought regarding the formation of giant planets, exemplified by the disk instability model and the core accretion model. Both models are well represented by papers at this conference. The rapidity with which giant planets can form from disk instability models has been further emphasized by subsequent work (Mayer et al., 2002). However, speed of formation is not the only criterion that a successful model must satisfy. It is the purpose of this note to point out that the predictions of the disk instability model are not consistent with the present composition of Jupiter's atmosphere as revealed by the Galileo Probe Mass Spectrometer (GPMS). The core accretion model can explain the observed atmospheric composition of Jupiter if appropriate planetesimals were available to enhance the heavy elements. Hence core accretion seems to provide a better explanation for the formation of Jupiter, the most massive of our own giant planets. Disk instabilities may well be the source of giant planets in other planetary systems.

2. The Composition of Jupiter's Atmosphere

A comprehensive review of all composition measurements has been published by Atreya et al. (2003). Measurements of isotopic ratios have been tabulated and discussed by Owen and Encrenaz (2003). The salient point for the present discussion is that all of the heavy elements whose abundances could be measured by the GPMS except helium and neon reveal an enrichment relative to solar values (as tabulated by Anders and Grevesse, 1989) by a factor of 3 ± 1 when expressed as a ratio relative to H (Owen et al., 1999). Helium and neon are depleted because of the condensation of helium into droplets that "rain out" in Jupiter's deep interior,

and the solution of neon in those droplets (Stevenson and Salpeter, 1977a, 1977b; Roulston and Stevenson, 1995).

The probe did not survive to levels sufficiently deep in the Jovian atmosphere to permit a measurement of the global value of O/H. The dominant carrier of oxygen is H_2O , and the entry point of the probe was a dessicated region of the atmosphere (a "hot spot"). All condensable species were depleted relative to expected values in the upper part of the probe trajectory. As the probe descended, it reached levels where H_2S and NH_3 were well mixed, but the H_2O mixing ratio was still increasing at the last data transmission at 19.8 bar (Niemann et al., 1998). Hence, the determination of the global value of O/H on Jupiter remains a critical measurement.

3. The Origin of the Enrichment of Heavy Elements

The factor of 3 ± 1 enrichment in all measured elements regardless of the volatility of their presumed primordial carriers was unexpected. Comets have generally been considered to be representative of the planetesimals causing any enrichment of heavy elements on the giant planets. Yet comets are notoriously deficient in nitrogen (Geiss, 1988, 2003; Krankowsky, 1991; Wyckoff et al., 1991). Hence it was assumed that nitrogen would also be deficient on Jupiter (Pollack and Bodenheimer, 1989; Owen and Bar-Nun, 1995). New upper limits of 1/10 and 1/13 solar Ar/O in comets C/2001 A2 (LINEAR) and C/2000 WM1 (LINEAR) (Weaver et al., 2002) severely challenge the reported detection of a solar value of Ar/O in Comet C/1995 01 by Stern et al. (2000), indicating that comets (at least those from the Oort cloud) are deficient in Ar as well. Yet both N and Ar exhibit the same factor of 3 ± 1 enrichment on Jupiter as S and C.

Instead of the comets we know, we are therefore led to postulate the existence of solar composition icy planetesimals (SCIPs) that brought in the observed excess of heavy elements. If Jupiter's atmosphere is indeed representative of the bulk composition of the planet, this three-fold enrichment implies the presence of at least 12 Earth-masses (M_E) of these SCIPs, in addition to the 6 M_E of heavy elements expected from a solar mixture, relative to Jupiter's hydrogen. If this material has also enriched the other giant planets, it must have been the most abundant solid in the early solar system (Owen and Encrenaz, 2003).

The origin of these unusual planetesimals is difficult to understand. If they formed originally by the trapping of volatiles in amorphous ice, low gas densities and temperatures less than 30 K are required, suggesting formation at the outer fringes of the solar system or in the pre-solar interstellar cloud (Owen et al., 1999). The problem then is to determine how the gas-laden ice was transported to Jupiter's orbit where it could participate in the formation of the planet.

If instead, crystalline ice condensed to form clathrate hydrates in grains at Jupiter's orbit, the ice would have to stay in the form of microscopic grains for at least 2 million years to allow the nebula to cool below 38 K so argon could

form clathrates (Gautier et al., 2001a, 2001b). This requirement appears to conflict with more rapid time scales for accretion of planetesimals [growth to ∼80 km size objects in 0.25 My is predicted by Weidenschilling (1997)] and models proposing growth of a 10 Earth mass core in just 0.5 My (Pollack et al., 1996; Hersant et al., 2001). This conflict implies in turn a large inefficiency in the clathration process, leading to a requirement for a hugely over-solar abundance of water in Jupiter today (Owen and Encrenaz, 2003).

Given the importance of solid planetesimals for building giant planets, it is obviously critical to understand how these constraints on the formation of SCIPs can be satisfied. Measurement of the global water abundance on Jupiter would provide a clear test of the clathration hypothesis. If clathration was 100% efficient, O/H on Jupiter would be 9 times the solar value, while amorphous trapping would predict the same enrichment factor of 3 ± 1 exhibited by the other elements.

4. The Nature of Jovian Nitrogen

The isotope ratio ^{15}/N^{14}N on Jupiter provides additional support for the role of SCIPs in producing the observed heavy element enrichement. Initial difficulties in measuring the NH_3 abundance on Jupiter with the GPMS (Niemann et al., 1998) have been overcome by additional laboratory calibrations (Wong and Mahaffy, 2001), so that the value of N/H on Jupiter now reflects both the mass spectrometer result and the indirect determination by attenuation of the probe's radio signal (Folkner et al., 1998). Thus it is clear that N is enriched by the same factor as the other elements that were measured, viz., C, S, Ar, Kr, and Xe.

However, it is important to note that whereas ^{38}Ar/^{36}Ar and ^{13}C/^{12}C are the same in Jupiter and the Earth, ^{15}N/^{14}N is distinctly lower (Fouchet et al., 1998; Owen et al., 2001). This can be understood if the nitrogen on Jupiter originally reached the planet in the form of N_2 rather than NH_3 or other nitrogen compounds (Owen and Bar-Nun, 1995; Owen et al., 2001). The reason is the same one that distinguishes the D/H in H_2 from that in H_2O: in each case the heavier isotope is bound more tightly in compounds than in the homonuclear molecule. In the interstellar medium, the dominant form of nitrogen is N or N_2 , with a fractional abundance of 70 to 90% (van Dishoek et al., 1993). Calculations by Terzieva and Herbst (2000) demonstrate that ion molecule reactions in the interstellar medium at 10 K will produce, ^{15}N/^{14}N in compounds with 1.3 times the value in N_2 . On Jupiter, $^{15}\mathrm{N}/^{14}\mathrm{N} = 2.3 \pm 0.3 \times 10^{-3}$, compared with $3.1^{+0.5}_{-0.4} \times 10^{-3}$ in HCN measured in Comet Hale-Bopp (Jewitt et al., 1997). Both the cometary and Jovian isotope ratios are presumably primordial; the difference then results from the different molecular carriers, in good agreement with the calculations. The terrestrial value of $^{15}\mathrm{N}/^{14}\mathrm{N} = 3.66 \times 10^{-3}$ may ultimately allow us to identify the compound(s) that brought nitrogen to Earth. N_2 was evidently not a major player in

this delivery, whereas it was apparently the dominant carrier of nitrogen to Jupiter and presumably to the Sun as well (Owen et al., 2001).

5. The Constraint on Models for Jupiter's Formation

The results from the GPMS are very robust. The value of He/H determined by this instrument agrees exactly with the measurement by the probe's Helium Abundance Detector (von Zahn et al., 1998). The enrichment of carbon by a factor 3 compared to solar C/H agrees with earlier remote sensing determinations summarized by Gautier and Owen (1989). Thus the measurements of additional abundances and isotopes made by the GPMS are on a sound footing.

This is important, because these measurements provide a useful constraint on models for the formation of Jupiter. Gravitational instabilities in the solar nebula will produce planets with solar composition (Boss, 1998, 2000, 2001; Mayer et al., 2002) that should therefore have no more than 6 M_E of heavy elements. The proposal to enrich the atmospheres of such planets after their formation by means of impacts from small bodies (Boss, 1998) is no longer viable, as there are no small bodies we know of that exhibit solar ratios of noble gases, nitrogen, carbon and sulfur. The deduction that nitrogen was delivered to Jupiter in the form of N_2 adds considerable force to this conclusion. Where it has been measured, N_2 in comets is a minor fraction of the nitrogen inventory, which is consequently depleted relative to solar abundances, as mentioned above. A. Cochran and her collaborators (2000, 2002) were unable to detect N_2^+ at all in three comets. Hence the comets we know could not have delivered the nitrogen we now find on Jupiter.

SCIPs are apparently no longer with us in the abundances required to produce the observed effect. Mixing within the Jovian atmosphere is demonstrated by the depletion of He and Ne mentioned earlier. Thus it is unreasonable to expect that the addition to the planet early in its history of just enough solar composition material to enrich the outermost layers would survive unfractionated in the sensible atmosphere until the present epoch.

The GPMS results therefore provide a strong argument against disk gravitational instability models for forming Jupiter. This obviously does not exclude the possibility that such models could explain the formation of the giant planets found around other stars.

Core accretion models are still viable if SCIPs were available to build the cores and if mixing from the cores into the envelopes was efficient. The same mission that measures the deep water abundance on Jupiter can measure ammonia as well, thereby enabling a good determination of the global abundances of CNO, the major fraction of the heavy elements in the envelope. Accurate tracking of the spacecraft will constrain the mass of the core. Comparable information about the other giant planets (including abundances measured by atmospheric probes) will determine just how uniform the giant planet formation process was in our own solar system.

Acknowledgements

I am grateful for helpful discussions with Sushil Atreya, Therese Encrenaz, Daniel Gautier and Jack Lissauer.

References

Anders, A., and Grevesse, N.: 1989, 'Abundances of the Elements: Meteoritic and Solar', *Geochim et Cosmochim. Acta* **53**, 197–214.

Atreya, S. K., Mahaffy, P. R., Niemann, H. B., Wong, M. H., and Owen, T. C.: 2003, 'Composition and Origin of the Atmosphere of Jupiter–An Update, and Implications for the Extrasolar Giant Planets', *Planet. Space Sci.* **51**, 105–112.

Boss, A. P.: 1998, 'Evolution of the Solar Nebula. IV. Giant Gaseous Protoplanet Formation', *Astrophys. J.* **503**, 923–937.

Boss, A. P.: 2000, 'Possible Rapid Gas Giant Planet Formation in the Solar Nebula and Other Protoplanetary Disks', *Astrophys. J.* **536**, L101–L104.

Boss, A. P.: 2001, 'Gas Giant Protoplanet Formation: Disk Instability Models with Thermodynamics and Radiative Transfer', *Astrophys. J.* **563**, 367–373.

Cochran, A.: 2002, 'A Search for N_2^+ in Spectra of Comet C/2002 C1 (Ikeya-Zhang)', *Astrophys. J.* **576**, L165–L168.

Cochran, A. L., Cochran, W. D., and Barker, E. S.: 2000, 'N_2^+ and CO^+ in Comets 122P/1995 S1 (deVico) and C/1995 O1 (Hale-Bopp)', *Icarus* **146**, 583–593.

Folkner, W. M., Woo, R., and Nandi, S.: 1998, 'Ammonia Abundance in Jupiter's Atmosphere Derived from the Attenuation of the Galileo Probe's Radio Signal', *J. Geophys. Res.* **103**, 22847–22856.

Fouchet, Th., Lellouch, E., Drossart, P., Encrenaz, Th., Bezard, B., and Feuchtgruber, H.: 1998, 'Jupiter Troposphere as Seen with ISO/SWS', *Bull. Amer. Astron. Soc.* **30**, 1061.

Gautier, D., Hersant, F., Mousis, O., and Lunine, J.: 2001a, 'Enrichments in Volatiles: A New Interpretation of the Galileo Measurements', *Astrophys. J.* **550**, L227–l230.

Gautier, D., Hersant, F., Mousis, O., and Lunine, J.: 2001b, 'Erratum', *Astrophys. J.* **559**, L183.

Gautier, D., and Owen, T.: 1989, 'The Composition of Outer Planet Atmospheres'. In: S. K. Atreya, J. B. Pollack, and M. S. Mattthews (eds.): *Origin and Evolution of Planetary and Satellite Atmospheres*. Univ. Arizona Press, Tucson, pp. 487–512.

Geiss, J.: 1988, 'Composition in Halley's Comet; Clues to Origin and History of Cometary Matter', *Rev. Mod. Astron.* **1**, 1–127.

Geiss, J.: 2003, 'Update of the Results from Ulysses' Tranversal of the Tail of Hyakutake', *Nature* **404**, 576–578.

Hersant, F., Gautier, D., and Huré, J.: 2001, ' Two-Dimensional Model for the Primordial Nebula Constrained by D/H Measurements in the Solar System: Implications for the Formation of Giant Planets', *Astrophys. J.* **554**, 391–407.

Jewitt, D. C., Matthews, H. E., Owen, T., and Meier, R.: 1997, 'Measurements of $^{12}C/^{13}C$, $^{14}N/^{15}N$ and $^{32}S/^{34}S$ Ratios in Comet Hale-Bopp (C/1995 01)', *Science* **278**, 90–93.

Krankowsky, D.: 1991,'The Composition of Comets'. In: R. L. Newburn, M. Neugebauer, and J. Rahe (eds.): *Comets in the Post-Halley Era*. Kluwer, Dordrecht, pp. 855–877.

Mahaffy, P. R., et al.: 2000, 'Noble Gas Abundances and Isotope Ratios in the Atmosphere of Jupiter from the Galileo Probe Mass Spectrometer', *J. Geophys. Res.* **105**, 15061–15071.

Mayer, L., Quinn, T., Wadsley, J., and Stadel, J.: 2002, 'Formation of Giant Planets by Fragmentation of Protoplanetary Disks', *Science* **298**, 1756–1759.

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, 'The Composition of the Jovian Atmosphere As Determined by the Galileo Probe Mass Spectrometer', *J. Geophys. Res.* **103**, 22831–22846.

Owen, T., and Bar-Nun, A.: 1995, 'Comets, Impacts and Atmospheres', *Icarus* **116**, 215–226.

Owen, T., Mahaffy, P., Niemann, H. B., Atreya, S., Donahue, J., Bar-Nun, A., and de Pater, I.: 1999, 'A Low Temperature Origin for the Planetesimals That Formed Jupiter. *Nature* **402**, 269–270.

Owen, T., Mahaffy, P. R., Niemann, H. B., Atreya, S., and Wong, M.: 2001, 'Protosolar Nitrogen', *Astrophys. J.* **553**, L79–L81.

Owen, T., and Encrenaz, Th.: 2003, 'Element Abundances and Isotopic Ratios in the Giant Planets and Titan', *Space. Sci. Rev.* **106**, 121–138.

Pollack, J. B., and Bodenheimer, P.: 1989, 'Theories of the Origin and Evolution of the Giant Planets', In: S. K. Atreya, J. B. Pollack, and M. S. Mattthews (eds.): *Origin and Evolution of Planetary and Satellite Atmospheres*. Univ. Arizona Press, Tucson, pp. 564–602

Pollack, J. B., Hubickyj, O., Bodenheimer, P., Lissauer, J. J., Podolak, M., and Greenzweig, Y.: 1996, 'Formation of the Giant Planets by Concurrent Accretion of Solids and Gas,' *Icarus* **124**, 62–85.

Roulston, M. S., and Stevenson, D. J.: 1995, 'Prediction of Neon Depletion in Jupiter's Atmosphere (abstract)', *EOS Trans. AGU* **76** (46) Fall Mtg. Suppl. F343.

Stern, S. A., Slater, D. C., Festou, M. C., Parker, J. W., Gladstone, G. R., A'Hearn, M. F., and Wilkinson, E.: 2000, 'The Discovery of Argon in Comet C/1995 O1 (Hale-Bopp),' *Astrophys. J.* **544**, L69–L72.

Stevenson, D. J., and Salpeter, E. E.: 1977a, 'The Phase Diagram and Transport Properties of Hydrogen-Helium Fluid Planets,' *Astrophys. J. Suppl.* **35**, 221–237.

Stevenson, D. J., and Salpeter, E. E.: 1977b, 'The Dynamics and Helium Distribution Properties for Hydrogen-Helium Fluid Planets', *Astrophys. J. Suppl.* **35**, 239–261.

Terzieva, R., and Herbst, E.: 2000, 'The Possibility of Nitrogen Isotopic Fractionation in Interstellar Clouds', *Mon. Not. Royal Astron. Soc.* **317**, 563–568.

van Dishoeck, E. F., Blake, G. A., Draine, B. T., and Lunine, J. I.: 1993, 'The Chemical Evolution of Protostellar and Protoplanetary Matter'. In: E. H. Levy and J. I. Lunine (eds.): *Protostars and Planets III*. Univ. Arizona Press, Tucson, pp. 163–241.

von Zahn, U., Hunten, D. M., and Lehmacher, G.: 1998, 'Helium in Jupiter's Atmosphere: Results from the Galileo Probe Helium Interferometer Experiment', *J. Geophys. Res.* **103**, 22815–22830.

Weaver, H. A., Feldman, P. D., Combi, M. R., Krasnopolsky, V., Lisse, C. M., and Shemansky, D. E.: 2002, 'A Search for Argon and O VI in Three Comets Using the *Far Ultraviolet Spectroscopic Explorer*', *Astrophys. J.* **576**, L95–L98.

Weidenschilling, S. J.: 1997, 'The Origin of Comets in the Solar Nebula: A Unified Model', *Icarus* **127**, 290–306.

Wong, M. H., and Mahaffy, P. R.: 2001, Revised Deep Water Mixing Ratio in the Galileo Probe Entry Site from GPMS Measurements', *Bull. Amer. Astron. Soc.* **33**, 11.21.

Wyckoff, S., Tegler, S. C., and Engel, L.: 1991, 'Nitrogen Abundance in Comet Halley', *Astrophys. J.* **367**, 641–648.

Part II

Multi-Planetary Systems: Observations

Isotopic Constraints on the Formation of Earth-Like Planets

Alex N. Halliday
Department of Earth Sciences, ETH Zentrum, NO, Sonneggstrasse 5, CH-8092, Zürich, Switzerland

Abstract. The terrestrial "rocky" planets of our solar system are depleted in major volatiles as well as moderately volatile elements like potassium. Isotopic chronometry provides strong evidence that accretion of the terrestrial planets took time-scales on the order of 10^7–10^8 years, consistent with some dynamic simulations. Mars appears to have formed within the first 15 Myrs of the solar system. Growth of the Earth was relatively protracted, the Moon most probably forming >45 Myrs after the start of the solar system, although the exact age is still poorly constrained. There is evidence that significant compositional change occurred during the growth of the Earth-Moon system. The isotopic compositions of Sr and W in lunar samples and Xe in the Earth provide evidence that the proto-planets that contributed to the growth of the Earth were more volatile rich than the present Earth or Moon. Time-integrated radioactive parent / radiogenic daughter elemental ratios Rb/Sr and Hf/W can be deduced from Sr and W isotope data respectively. These show that the materials that eventually formed the Moon had, on average, an order of magnitude higher volatile / refractory element budget and were significantly more oxidizing. Similarly, Xe isotope data for the Earth shows that it grew from material with heavy noble gas budgets that were at least two orders of magnitude higher than today. Therefore, mechanisms that result in such changes, presumably as a result of accretion need to be incorporated into future dynamic models for the growth of the terrestrial planets. A major unsolved question remains the origin of Earth's water.

1. Introduction

The rocky, inner or terrestrial planets are distinct from the gaseous and icy giant, outer or jovian planets, as well as the currently detectable extrasolar planets, in their much smaller size and higher uncompressed density. Just as we had no convincing evidence for extrasolar planets until the mid-90s (Mayor and Queloz, 1995), we currently lack any evidence for extrasolar terrestrial planets. As of the time of Michel Mayor's 60^{th} birthday we still do not know if our solar system is unusual in this respect. Resolving this issue and finding habitable worlds is a prime objective in astrophysics and observational astronomy (Seager, 2003). When we reconvene for Michel Mayor's 80^{th} there may well be a host of new kinds of such planets to discuss. All our current understanding is theoretical or based on observations of our own solar system. From this we can say that the formation of the terrestrial planets was stochastic and strongly influenced by late stage events that resulted in a huge spectrum of surface environments. A good example is the Earth itself, the origin of whose water currently is unclear, but may have been added at a late stage by chance bolides (Owen and Bar-Nun, 2000). Water has had a huge effect on the Earth, not just in its surface environments, but also in reducing the viscosity of the mantle thereby facilitating efficient convection. The water content of the Earth has a direct effect on the products and styles of volcanism and has ultimately led to

the long-term growth of the low density masses of rock we call continents (Taylor, 1992). Terrestrial planets with a geological history and recent surface environments like that of the Earth probably are quite distinctive but could also turn out to be particularly hard to replicate.

Terrestrial planets are of course distinct because they are relatively depleted in major volatile elements like hydrogen and helium. However, they are also relatively depleted in moderately volatile elements like potassium that are still incompletely condensed at $\sim$ 1,000K (Gast, 1960; Wasserburg et al., 1964; Cassen, 1996). The shortage of these volatile constituents has come to be viewed as a consequence of growth from a portion of the inner solar system that was already swept clear of nebular gas and was dominated by dust that had been heated by T-Tauri radiation close in to the Sun (e.g. Hayashi et al., 1985) or had been heated during collapse of the solar nebula (Boss, 1990). Therefore, the most widely accepted dynamic models for the growth of the terrestrial planets assume accretion after the nebula was largely removed. The absence of nebular gas has a significant effect on the calculated time-scales for the growth of the terrestrial planets. Dynamic simulations indicate that the growth of the terrestrial planets would then take 10^7–10^8 years in the absence of Jupiter (Safronov, 1954; Wetherill, 1986). In the presence of a minimum mass solar nebula this would shorten to a few times 10^6 years (Hayashi et al., 1985).

There now exists good evidence of solar noble gases in the Earth and also Mars (Bogard et al., 2001). For the Earth this includes He, Ne and Xe (Caffee et al., 1999; Pepin and Porcelli, 2002). A solar Ar component is not yet well resolved and little data exist for Kr. Although others have been proposed (Trieloff et al., 2000), the simplest and most efficient mechanism for incorporating solar noble gases into the Earth would be via in-gassing (Porcelli et al., 2001) – dissolving a small fraction of the constituents of a hot thick blanketing nebular proto-atmosphere into the silicate and metallic liquids of a magma ocean (Hayashi et al., 1979; Mizuno et al., 1980). This is consistent with new dynamic simulations that indicate that at least some small amount of nebular gas would appear necessary to dampen eccentricities during planetary growth (Agnor and Ward, 2002). If the Earth and Mars trapped significant nebular gases they must have accreted a fraction of their mass quickly i.e. before dispersal of the solar nebula.

These models can be tested with isotopic studies of early solar system objects such as meteorites and lunar samples. There now exist sufficient data to show that different isotopic methods do not always yield the same time-scales (Halliday, 2003). This is most probably caused by differences in behaviour between the elements during the history of accretion (Halliday, 2004). Therefore, it is anticipated that in future we will be able to go far beyond determining time-scales. We will increasingly be able to utilize these chronometers to determine the physical and chemical environment in which the terrestrial planets grew.

All isotopic approaches make certain assumptions and in this chapter some of these assumptions are briefly explained. More detailed discussion can be found in

Halliday (2003, 2004). For example, the calculated time-scale of planetary accretion depends on the ratio of the radioactive parent to radiogenic daughter element. If this ratio were different in the early Earth or its precursor components because of different volatile budgets, the calculated time-scales for Earth accretion would be incorrect. Some lines of evidence for such changes in volatile budgets are discussed here. Equilibration during accretion is also an issue. The most widely used approaches are tungsten and lead isotopes. These methods assume that, as the planet grows, accreting metal and silicate isotopically equilibrates (or mixes) completely with the silicate portion of the growing planet. There is evidence that this is unlikely. The isotopic time-scales deduced would then be too short.

2. Isotopic Chronometry of Accretion

Isotope geochemistry has expanded dramatically in the past few years thanks to technological advances and is now a very broad subject encompassing virtually the entire periodic table. Much of this expansion relates to mass dependent and mass independent isotopic fractionations. However, the development of new methods has also facilitated the study of new radiogenic isotope systems. A variety of isotopic chronometers have been particularly useful for deducing early solar system time-scales (Table I). These range from very short-lived nuclides like ^{26}Al ($T_{1/2} = 0.7$ Myrs) to almost stable isotopes like ^{238}U with a half-life similar to the age of the Earth. The age of the solar system is 4.57×10^9 years. Therefore, a number of early solar system radionuclides in Table I are, by now, effectively extinct. They were present in the early solar system because of production in other stars, many of them massive, shortly before (within a few half-lives of) the formation of the sun. Being extinct, the initial abundance can only be inferred. This is achieved by using differences in the atomic abundance of the daughter isotope in early objects (meteorites) of known absolute age.

To utilize any of these nuclides for deducing early solar system time-scales it is necessary to precisely measure the present day isotopic composition of the daughter element. Therefore, the development of these chronometers is strongly linked to breakthroughs in mass spectrometry. In the simplest system the present day isotopic composition of the daughter element is a function of the initial abundance of the parent and daughter nuclide at the start of the solar system, the parent / daughter element ratio, and how much time elapsed. If you like, the isotopic composition of the daughter represents a time-integrated parent / daughter elemental ratio. If an object forms without a change in parent / daughter ratio, its time of formation will be undetectable with isotopic chronometry. It is the large change in parent / daughter ratio that accompanies certain processes associated with planetary accretion that renders isotopic systems useful. In fact it is this change that is being dated. If there is no change that accompanies accretion there is no way to distinguish when an object formed relative to the start of the solar system.

Table I. Main isotopic decay systems in use in studying the origin and early evolution of the terrestrial planets. * Spontaneous fissionogenic product

Parent	Daughter(s)	Half-life (years)	Applications
Long-lived			
^{40}K	^{40}Ca, ^{40}Ar	1.25×10^9	volcanic rocks, uplift ages, lunar bombardment
^{87}Rb	^{87}Sr	4.88×10^{10}	mantle and crust evolution, marine carbonates, early solar system
^{138}La	^{138}Ba, ^{138}Ce	1.05×10^{11}	very limited applications
^{147}Sm	^{143}Nd	1.06×10^{11}	mantle and crust evolution, early solar system
^{176}Lu	^{176}Hf	3.57×10^{10}	mantle and crust evolution, early solar system
^{187}Re	^{187}Os	4.23×10^{10}	mantle and crust evolution, early solar system
^{232}Th	^{208}Pb	1.40×10^{10}	limited applications
^{235}U	^{207}Pb	7.04×10^8	mantle and crust evolution, early solar system
^{238}U	^{206}Pb	4.47×10^9	mantle and crust evolution, early solar system
Extinct			
^{22}Na	^{22}Ne	2.605	early solar system
^{26}Al	^{26}Mg	7.3×10^5	early solar system
^{53}Mn	^{53}Cr	3.7×10^6	early solar system
^{60}Fe	^{60}Ni	1.5×10^6	early solar system
^{92}Nb	^{92}Zr	3.6×10^7	early solar system
^{107}Pd	^{107}Ag	6.5×10^6	early solar system
^{129}I	^{129}Xe	1.57×10^7	early solar system, terrestrial degassing
^{135}Cs	^{135}Ba	2.3×10^6	early solar system
^{146}Sm	^{142}Nd	1.03×10^8	early solar system, crustal evolution
^{182}Hf	^{182}W	9×10^6	early solar system, core formation
^{205}Pb	^{205}Tl	1.5×10^7	early solar system
^{244}Pu	^{136}Xe*	8.2×10^7	early solar system, terrestrial degassing

The age of an early rock or mineral can therefore be dated if the parent / daughter ratio is fractionated during formation. On this basis a number of early solar system objects have been dated and provide a framework within which to assess the time-scales for overall growth of the terrestrial planets. The earliest precisely dated objects that are thought to have formed within our solar system are the calcium aluminum refractory inclusions (CAIs) found in many chondrites (Table II). How these formed is unclear but they are thought to be recrystallized refractory condensates. Chondrules are also found in chondrites and very common but enigmatic round globules of what is thought to represent dust balls that were melted. It would appear that at least some chondrules formed about 2 million years

Table II. Recent estimates of the ages of early solar system objects

Object	Sample(s)	Method	Reference	Age (Ga)
Earliest Solar System	Allende CAIs	U-Pb	Göpel et al., 1991	4.566 ±0.002
Earliest Solar System	Efremovka CAIs	U-Pb	Amelin et al., 2002	4.5672±0.0006
Chondrule formation	Acfer chondrules	U-Pb	Amelin et al., 2002	4.5647±0.0006
Angrites	Angra dos Reis & LEW 86010	U-Pb	Lugmair and Galer, 1992	4.5578±0.0005
Early eucrites	Chervony Kut	Mn-Cr	Lugmair and Shukulyokov, 1998	4.563±0.001
Mars accretion	Minimum age	Sm-Nd	Harper et al., 1995	$\geq$ 4.54
Mars accretion	Mean age	Hf-W	Lee et al., 1997	4.560
Mars accretion	Minimum age	Hf-W	Lee et al., 1997	$\geq$ 4.54
Mars accretion	Minimum age	Hf-W	Halliday et al., 2001	$\geq$ 4.55
Mars accretion	Minimum age	Hf-W	Kleine et al., 2002	$\geq$ 4.55
Earth accretion	Mean age	U-Pb	Halliday, 2000, 2004	$\leq$ 4.55
Earth accretion	Mean age	U-Pb	Halliday, 2000	$\geq$ 4.49
Earth accretion	Mean age	Hf-W	Yin et al., 2002	$\geq$ 4.55
Lunar highlands	Ferroan anorthosite 60025	U-Pb	Hanan & Tilton, 1987	4.50±0.01
Lunar highlands	Ferroan anorthosite 60025	Sm-Nd	Carlson & Lugmair, 1988	4.44±0.02
Lunar highlands	Norite from breccia 15445	Sm-Nd	Shih et al., 1993	4.46±0.07
Lunar highlands	Ferroan noritic anorthosite in breccia 67016	Sm-Nd	Alibert et al., 1994	4.56±0.07
Moon	Best estimate of age	U-Pb	Tera et al., 1973	4.47±0.02
Moon	Best estimate of age	U-Pb, Sm-Nd	Carlson & Lugmair, 1988	4.44-4.51
Moon	Best estimate of age	Hf-W	Halliday et al., 1996	4.47±0.04
Moon	Best estimate of age	Hf-W	Lee et al., 1997	4.51±0.01
Moon	Maximum age	Hf-W	Halliday, 2000	$\leq$ 4.52
Moon	Maximum age	Rb-Sr	Halliday and Porcelli, 2001	$\leq$ 4.55
Moon	Best estimate of age	Hf-W	Lee et al., 2002	4.51±0.01
Moon	Best estimate of age	Hf-W	Kleine et al., 2002	4.54±0.01
Moon	Best estimate of age	Hf-W	Halliday, 2004	$<$ 4.52

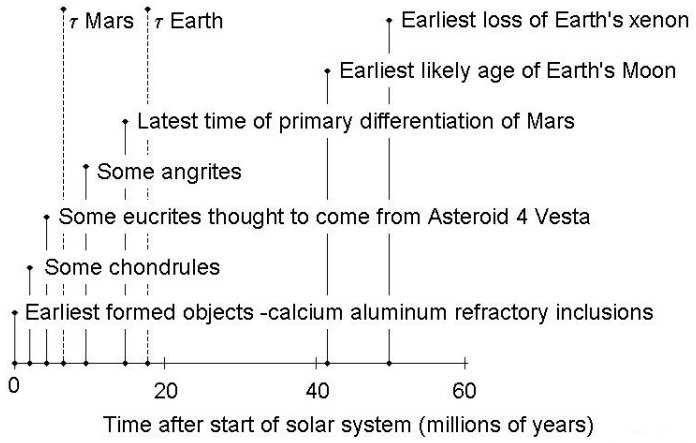

Figure 1. The current best estimates for the time-scales over which very early inner solar system objects and the terrestrial planets formed. The approximated mean life of accretion (τ) is the time taken to achieve 63% growth at exponentially decreasing rates of growth. The dashed lines indicate the mean lives for accretion deduced for Mars based on W isotopes (Lee et al., 1997) and the Earth based on Pb isotopes (Halliday, 2000, 2003, 2004). Estimates for the mean life of accretion of the Earth and the age of the Moon that are based solely on W isotopes (Kleine et al., 2002; Yin et al., 2002) are shorter than the values given here. See Table II for details of other sources.

after the start of the solar system. How chondrules formed is unclear. The most widely accepted models involve dust in the solar nebula. Therefore, the isotopic dating provides evidence that the disk from which the planets grew was still a dusty environment with rapid melting events, 2 million years after the start of the solar system (Fig. 1). Differentiated asteroids and their planetary cores and silicate reservoirs, as represented by some iron and basaltic achondrite meteorites appear to have formed in the first 10 Myrs of the solar system (Table II). Exactly how early is unclear. The ages of meteorites provide a maximum constraint on how slowly the parent bodies formed.

For many larger planetary objects, the timescaleas are still harder to ascertain. The parent / daughter fractionation is more complex because it takes place as the planet grows and this may take some significant period of time ($>10^6$ years). For an object like the Earth there is no such thing as a precise age when all of the constituents were amalgamated. It is more practical to instead assume a style of growth and use the isotopic data to define a rate of growth of a planet from the integrated history of the chemical fractionation of the parent / daughter ratio provided by the isotopic composition of the daughter.

Such a major parent / daughter fractionation clearly affected the noble gases. Xenon isotopes provide powerful evidence that the noble gases were vastly more abundant in the early earth and that a major fraction has been lost, probably by

shock-induced blow off of the atmosphere. The Xe constraints are based on the decay of formerly live ^{129}I ($T_{1/2}$ = 16 Myrs) and ^{244}Pu (fission $T_{1/2}$ = 82 Myrs) (Table I). Put simply, the small difference in composition between the isotopic composition of xenon in the Earth relative to its initial starting composition is consistent with low parent / daughter ratios. The amount of xenon in the Earth today is thought to be at least two orders of magnitude lower than is necessary to satisfy the Xe isotope data (Porcelli et al., 2001). That is, the ratio of elements like I and Pu to Xe is huge; at least two orders of magnitude higher. Therefore, there must have been a major fractionation in parent / daughter ratio at a relatively late stage. Because Xe is a gas that has a strong preference for residing in the atmosphere and because similar effects are found in both the I-Xe and Pu-Xe system, it is clear that there was a significant amount of xenon loss from the Earth, as opposed to selective I or Pu gain. The best estimate for the age of this loss is 50 to 80 Myrs (Fig. 1) (Ozima and Podosek, 1999; Porcelli and Pepin, 2000). It has been proposed that this may have been associated with the Moon-forming giant impact. The Xe constraints represent just one piece of isotopic evidence that accretion took 10^7–10^8 years and that there were significant changes in composition during planetesimal and planet growth in the terrestrial planet forming region.

The most powerful chronometers for determining the growth rates of the terrestrial planets are ^{182}Hf–^{182}W and ^{235,238}U–^{207,206}Pb (Halliday, 2003). In both cases the parent / daughter ratio is strongly fractionated during core formation. All of the terrestrial planets have a metallic iron core formed by segregation of dense iron-rich liquids. Core formation separates the metal-loving elements (including the daughter elements W, Pb) from the silicate-loving elements (in this case the radioactive parents Hf, U) (Fig. 2). As such, the W and Pb isotopic compositions in the silicate portion of the planet reflect how fast it grew and internally fractionated by metal segregation. Although it was once thought that accretion was rapid (Hanks and Anderson 1969) and formation of the core was protracted (Solomon, 1979), there exists strong evidence that core formation is in fact early and rapid (Sasaki and Nakazawa, 1986), such that the core grows proportionally in response to the growth of the planet. On this basis the isotopic data yield constraints on the rates of growth of the planet, provided certain assumptions are valid.

To determine accretion rates one needs to know the parent / daughter ratio in the total planet as well as in the reservoir, the formation of which is being dated. This is sometimes complex because we do not know independently how much resides in the inaccessible core. Some elements like Hf and W are refractory and were not affected by heating in the inner portions of the disk. Therefore, the Hf/W ratio of the Earth is well known from independent measurements of chondrites – meteorites representing average accumulations of early solar system debris (Newsom, 1995). However, Pb is moderately volatile and therefore variably depleted in the Earth, Moon, Vesta, Mars and almost certainly in Venus and Mercury. Therefore, two processes have produced depletions of Pb in the silicate Earth – core formation and volatile loss. Two approaches are mainly relied upon to determine the amount of Pb

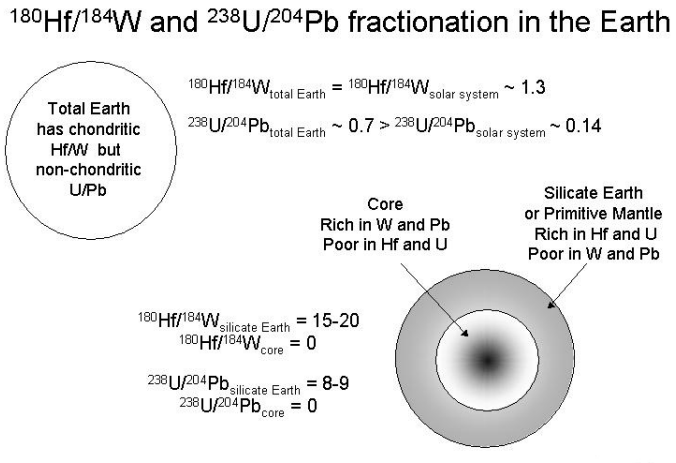

Figure 2. The Hf/W ratio of the total Earth is chondritic (average solar system) because Hf and W are both refractory elements. The U/Pb ratio of the Earth is enhanced relative to average solar system because approximately 80% of the Pb was lost by volatilization or incomplete condensation mainly at an early stage of the development of the circumstellar disk. The fractionation within the Earth for Hf/W and U/Pb is similar. In both cases the parent (Hf or U) prefers to reside in the silicate portion of the Earth. In both cases the daughter (W or Pb) prefers to reside in the core.

in the total Earth. The first is a comparison between the amount of Pb depletion in chondrites and that of other volatile elements that are not partitioned into planetary cores (Allègre et al., 1995). The second is based on estimates of the amount of volatile depletion expected given the average condensation temperature (Galer and Goldstein, 1996). These two approaches are in broad agreement for Pb.

It is also necessary to have a well-defined value for the isotopic composition of the silicate Earth relative to the rest of the solar system. The W isotopic composition of the silicate Earth is indeed extremely well defined because after 4.5 billions of convective mixing the Earth's silicate reservoir has homogenized any early isotopic variability. The W isotopic age of the Earth could in principle be determined from the composition of a tungsten carbide drill bit! However, with W the problem has been to get the correct value for the average solar system. Earliest Hf-W papers assumed a certain value (Jacobsen and Harper, 1996). Early measurements of chondrites were variable (Lee and Halliday, 1996, 2000). Most importantly it is now recognized that the early W isotopic measurements for carbonaceous chondrites reported in Lee and Halliday (1996) are incorrect and the time-scales have had to be reassessed in the light of this (Kleine et al., 2002; Schoenberg et al., 2002; Yin et al., 2002; Halliday, 2004).

The problem for U-Pb is almost the opposite. The present day value for the silicate Earth is very hard to constrain. The reason is that ^{235}U and ^{238}U are still alive in the Earth and capable of producing isotopic variations in ^{207}Pb/^{204}Pb and

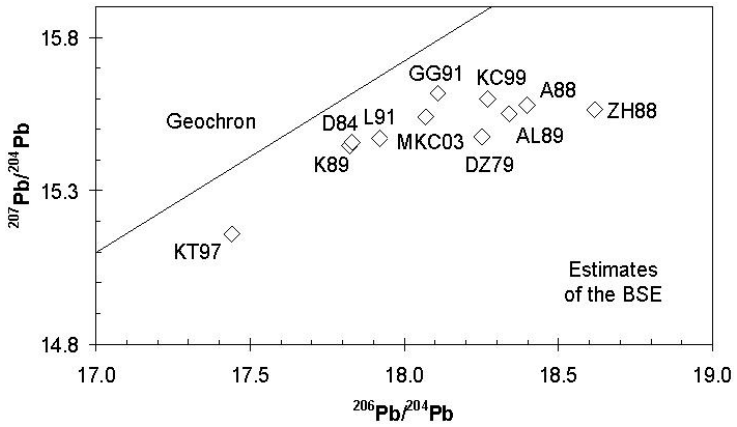

Figure 3. Estimates of the lead isotopic composition of the bulk silicate Earth (BSE) plotted relative to the Geochron defined as the slope corresponding to the start of the solar system. All estimates plot to the right of this line, indicating protracted accretion and / or core formation. Data from Doe and Zartman (1979), Davies (1984), Zartman and Haines (1988), Allègre *et al.* (1988), Allègre and Lewin (1989), Kwon *et al.* (1989), Liew *et al.* (1991), Galer and Goldstein (1991), Kramers and Tolstikhin (1997), Kamber and Collerson (1999) and Murphy *et al.* (2003).

$^{206}Pb/^{204}Pb$. Therefore, the Earth is very heterogeneous in its Pb isotopic composition (Galer and Goldstein, 1996). This is shown in Fig. 3 in which 11 estimates of the average Pb isotopic composition of the silicate Earth are presented. Each of these would imply a different time-scale for earth accretion, as discussed below and shown in Fig. 5.

The simplest kind of age calculation just deduces a time of fractionation or model age assuming the fractionating process was instantaneous. More complex models calculate an accretion rate (Jacobsen and Harper, 1996; Halliday, 2000, 2004) by making assumptions about the style of growth. Generally speaking this is an exponentially decreasing rate of growth to emulate what has arisen from dynamic simulations. To quantify the accretion time-scale with an exponentially decreasing rate of growth requires the use of a time constant for accretion. The inverse of this is the mean life, which corresponds to the time taken to achieve 63% growth. Some of the most recent models (Halliday, 2004) also try and simulate the effect of a more stochastic style of accretion with growth from sporadic large (Moon to Mars-sized) impactors (Fig. 4). The staggered growth model shown in Fig. 4 assumes that the Earth grew from 90 to 99% of its current mass as a result of a collision with an impacting Mars-sized planet called Theia that produced the Moon. The changes in the rate of accretion with time are calculated from the timing of this event. Therefore, the model time of the Giant Impact is directly related to the mean life of accretion in these models (Fig. 5).

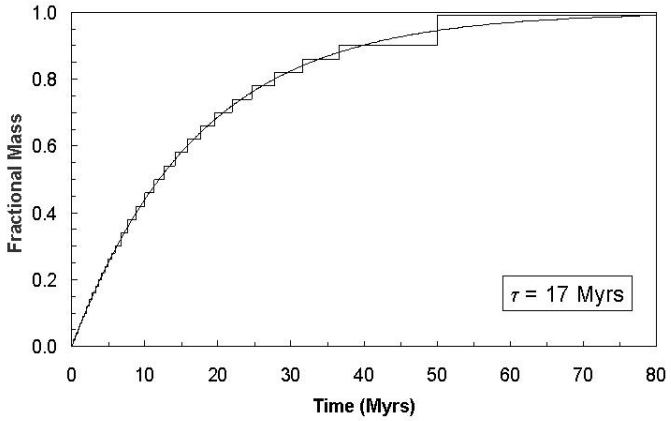

Figure 4. Change in mass fraction of the Earth as a function of time in the model used in Halliday (2004) illustrated with an accretion scenario calculated from a giant impact at 50Myrs after the start of the solar system. The approximated mean life of accretion (τ) is the time taken to achieve 63% growth. Both this and the timing of each increment are calculated from the timing of the giant impact (t_{GI}) to achieve an overall exponentially decreasing rate of growth for the Earth broadly consistent with dynamic simulations. The smooth curves show the corresponding exponentially decreasing rates of growth. The step function curves define the growth used in the isotopic calculations. Growth of the Earth is modelled as a series of collisions between differentiated objects. The overall rate of accretionary growth of the Earth may have decreased in some predictable fashion with time, but the growth events would have become more widely interspersed and larger. Therefore, the model simulates further growth by successive additions of 1% Earth mass objects from 1 to 10% of the current mass, then by 2% objects to 30% and then by 4% objects to 90%. The Moon-forming giant impact is modelled to take place when the Earth was $\sim$ 90% of its current mass and involved an impactor planet Theia that was $\sim$10% of the (then) mass of the Earth. Therefore, in the model the giant impact contributes a further 9% of the current Earth mass. There is evidence against large amounts of accretion after the giant impact.

3. Tungsten and Lead Isotope Differences and Disequilibrium Accretion

The rates of accretion of the Earth based on the model in Fig. 4 and the various estimates of the bulk silicate Earth Pb isotopic composition (Fig. 3) are shown in Fig. 5. Also shown are the results for W. It is clear that the kinds of time-scales involved are on the order of 10^7 to 10^8 years, providing strong vindication of the Safronov-Wetherill style of protracted accretion via major impacts (Halliday, 2003).

It is also immediately apparent that a discrepancy exists between the W results and every estimate based on Pb isotope data (Halliday, 2004). Either all of these Pb isotope estimates are incorrect or there is some basic difference in the way these clocks operate. There appears to be better agreement between the Pb and Xe isotope constraints on the time-scales. There are a number of parameters that play into these models that have a significant level of uncertainty associated.

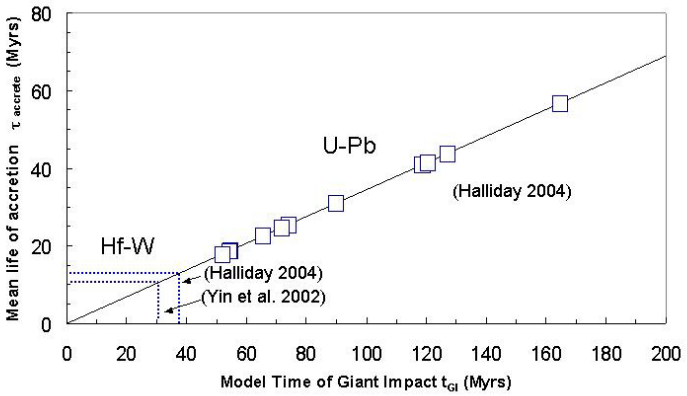

Figure 5. Calculated values for the Earth's mean life of accretion (τ) and time of the giant impact (T_{GI}) given in Myrs as deduced from the different estimates of the Pb isotopic composition of the bulk silicate earth (BSE) shown in Fig. 4 using the type of accretion model shown in Fig. 3 (see Halliday, 2004). All calculations assume continuous core formation and total equilibration between accreted material and the BSE. The $^{238}U/^{204}Pb$ values assumed for the Earth = 0.7 (Allègre et al., 1995). The Hf-W timescales using the same style of model are significantly shorter.

Currently, the most likely explanation for the apparent discrepancy is that there is a difference in the level of equilibration of refractory W as opposed to moderately volatile Pb as metal in the impactor mixes with the silicate Earth during accretion. In all these models it is assumed that growth takes place via collisions. The incoming planetesimals and planets are likely to be differentiated into silicate and metal like Vesta, Mars and the Moon-forming planet Theia (Taylor and Norman, 1990). Many isotopic models assume that all of the metal and silicate in the impactor mixes and equilibrates with the silicate portion of the Earth before further growth of the Earth's core (Kleine et al., 2002; Schoenberg et al., 2002; Yin et al., 2002). In reality the situation could have been very different however (Yoshino et al., 2003). The likelihood of metal from the impactor mixing and equilibrating with the silicate portions of the Earth depends on whether the metal in the incoming planetary core becomes fragmented into small droplets (Rubie et al., 2003). The fact that the silicate Earth has a W isotopic composition that is so similar to average solar system or chondrites means that in general terms equilibration must indeed be the norm. However, the small difference that is now known to exist could reflect an event like the giant impact that involved particularly rapid amalgamation of large fragments of impactor metal with the Earth's core with little chance for equilibration with the silicate portion of the Earth (Halliday, 2004).

4. Loss of Moderately Volatile Elements During Accretion

The depletion of the terrestrial planets in moderately volatile elements has been recognized for about 40 years. This was long considered a result of incomplete condensation because it was believed that the inner solar system formed from a hot solar nebular gas. Now it is understood that accretion of most of the material in the planets came from dust inherited from a presolar molecular cloud.

Oxygen isotopes provide a monitor of the degree of mixing of this presolar material (Clayton and Mayeda, 1975; Wiechert et al., 2001). The oxygen isotopic composition varies among different classes of meteorites and is therefore assumed to be highly heterogeneous in general among inner solar system objects (Clayton et al., 1973; Clayton, 1986, 1993). This is not a normal mass dependent fractionation so much as a mass independent effect, the origin of which is poorly understood or is at least controversial at the present time. Regardless of the reason behind the heterogeneity it is thought that the oxygen isotopic composition provides a monitor of the origins of the material in a planetary object. The provenance of material in an object the size of the Earth is expected to have been quite broad (Wetherill, 1994). As such, the oxygen isotopes merely provide some average. The Earth is distinct from all classes of chondrites except enstatite chondrites, which may have formed in the more reducing inner regions of the solar system. The provenance should be much broader than this. Similarly, the provenance appears to be distinct from that of the material forming the Asteroid 4 Vesta, or Mars, as judged from studies of meteorites thought to be derived from these objects (Clayton, 1986, 1993). The Moon however, has an oxygen isotopic composition that is identical to that of the Earth (Clayton and Mayeda, 1975), recently demonstrated to persist to extremely high precision (Wiechert et al., 2001). This is a striking finding that provides good evidence that the Earth and Moon were formed from material of similar provenance and presumably similar composition (Wiechert et al., 2001).

The Moon is widely considered to be the product of a collision between the proto-Earth and a Mars-sized planet sometimes referred to as "Theia" (Cameron and Benz, 1991; Canup and Asphaug, 2001). (Theia was the mother of Selene, the goddess of the Moon.) Roughly 80% of the Moon forms from the debris derived from Theia (Canup and Asphaug, 2001). If this is correct, the oxygen isotopic evidence provides evidence that Theia and the Earth had similar average provenance.

Yet the Moon is far more depleted in volatile elements than the Earth (Taylor and Norman, 1990). A likely explanation for this is that there were major reductions in volatile constituents during the Moon-forming giant impact (O'Neill, 1991). Support for this is found in the initial $^{87}Sr/^{86}Sr$ of early lunar rocks from the Highlands (Halliday and Porcelli, 2001). These yield an initial $^{87}Sr/^{86}Sr$ for the Moon that is clearly resolvable from the initial composition of the solar system, consistent with radiogenic growth in an environment with relatively high Rb/Sr. This time-integrated Rb/Sr can be determined to be $\sim$0.07, similar to the Rb/Sr

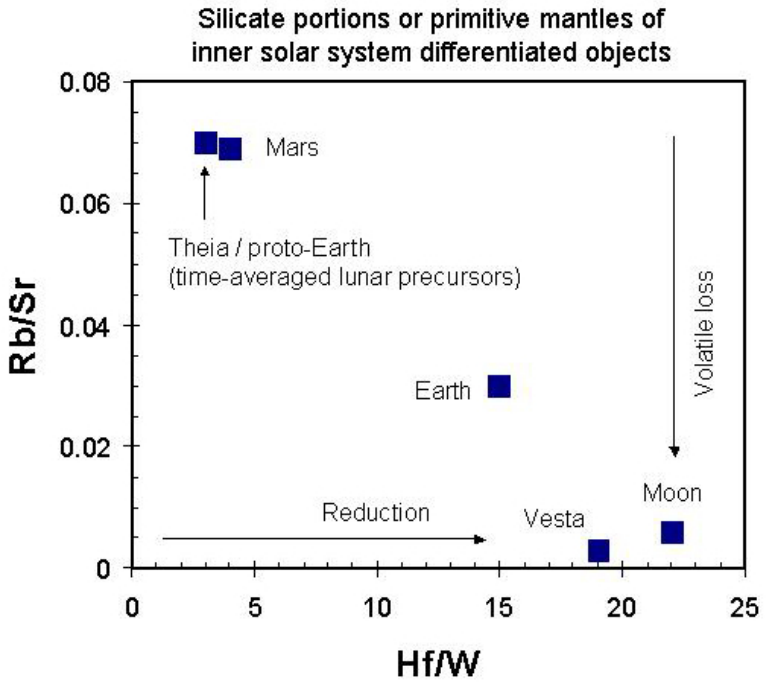

Figure 6. Hf/W appears to be negatively correlated with Rb/Sr in the primitive mantles of inner solar system planetesimals and planets. A possible explanation for this is that the loss of moderately volatile elements was linked to loss of other volatiles during planetary collisions such that the mantles changed from more oxidizing to more reducing. The Moon is an extreme example of this. The fact that the mixture of material from the proto-Earth and Theia that is calculated to have formed the Moon is so like Mars provides evidence that such volatile-rich objects may have been common in the inner solar system during the early stages of planetary accretion. See Halliday and Porcelli (2001) and Halliday (2004) for further details.

of Mars but an order of magnitude less depleted in moderately volatile Rb than the Moon today (Rb/Sr$\sim$0.006). Therefore, it appears as though Mars-like objects once existed in the vicinity of the Earth.

Tungsten isotopes paint a strikingly similar picture (Halliday, 2004). The Hf/W ratio of silicate reservoirs in the terrestrial planets is variable and is thought to depend on the level of oxidation of the mantle because this controls how "metal loving" W is during core formation. If the mantle is relatively reducing the vast majority of the W is segregated into the core and vice versa. There is a reasonably clear relationship between the degree of depletion in moderately volatile elements (like Rb) and the degree of depletion of W in the silicate portions or primitive mantles of different planets and planetesimals (Halliday, 2004). Just as the time-integrated Rb/Sr can be deduced from Sr isotopes, the time integrated Hf/W can be deduced from W isotopes. Tungsten isotope data for the Moon provide evidence that the precursor objects that provided the raw materials for the Moon had a time-

averaged Rb/Sr and Hf/W that was within error identical to the actual Rb/Sr and Hf/W of Mars (Fig. 6).

Therefore, the lunar data provide evidence of planetesimals and planets in the vicinity of the Earth that were more like Mars in terms of chemical composition. The fact that the oxygen isotopic composition of the Earth and Moon are identical provides evidence that such proto-planetary compositions may well have been a normal feature of inner solar system or terrestrial proto-planetary compositions. This however, assumes that oxygen isotopes provide a comparative monitor of the mix of raw ingredients in the inner solar system. The isotopic evidence for impact-driven chemical changes in the terrestrial planets during growth is however well supported by evidence that the planet Mercury lost a major fraction of its silicate by impact erosion (Benz et al., 1987).

Acknowledgements

This research is supported by ETH and SNF. The paper benefited from very fruitful discussions with Willy Benz and Don Porcelli. I'd like to thank Ruedi von Steiger for his patience and Michel Mayor for letting geochemists into Switzerland to dabble in planets.

References

Agnor, C. B. and Ward W. R.: 2002, 'Damping of terrestrial-planet eccentricities by density-wave interactions with a remnant gas disk', *Astrophys. J.* **567**, 579–586.

Alibert, C., Norman M. D. and McCulloch M. T.: 1994, 'An ancient Sm-Nd age for a ferroan noritic anorthosite clast from lunar breccia 67016', *Geochim. Cosmochim. Acta* **58**, 2921–2926.

Allègre, C. J., Manhès G. and Göpel C.: 1995, 'The age of the Earth', *Geochim. Cosmochim. Acta* **59**, 1445–1456.

Allègre, C. J. & Lewin, E.: 1989, 'Chemical structure and history of the Earth: Evidence from global non-linear inversion of isotopic data in a three box model', *Earth Planet. Sci. Lett.* **96**, 61–88.

Allègre, C. J., Lewin, E. & Dupré, B.: 1988, 'A coherent crust-mantle model for the uranium-thorium-lead isotopic system', *Chem. Geol.* **70**, 211–234.

Amelin, Y., Krot A. N., Hutcheon I. D. and Ulyanov A. A.: 2002, 'Lead isotopic ages of chondrules and calcium-aluminum-rich inclusions', *Science* **297**, 1678–1683.

Benz, W., Cameron A. G. W. and Slattery W. L.: 1987, 'Collisional stripping of Mercury's mantle', *Icarus* **74**, 516–528.

Bogard, DD, Clayton RN, Marti K, Owen T, Turner G: 2001, 'Martian volatiles: Isotopic composition, origin, and evolution', *Space Sci. Rev.* **96**, 425–458.

Boss, A. P.: 1990, '3D Solar nebula models: implications for Earth origin'. In*: Origin of the Earth*, H. E. Newsom and J. H. Jones, eds., pp. 3–15, Oxford University Press, Oxford.

Caffee, M. W., Hudson G. B., Velsko C., Huss G. R., Alexander E. C. Jr. and Chivas A. R.: 1999, 'Primordial noble gases from Earth's mantle: Identification of a primitive volatile component', *Science* **248**, 2115–2118.

Cameron, A. G. W. and Benz W.: 1991, 'Origin of the Moon and the single impact hypothesis IV', *Icarus* **92**, 204–216.

Canup, R. M. and Asphaug E.: 2001, 'Origin of the moon in a giant impact near the end of the Earth's formation', *Nature* **412**, 708–712.

Carlson, R.W. and Lugmair G. W.: 1988, 'The age of ferroan anorthosite 60025: oldest crust on a young Moon?' *Earth Planet. Sci. Lett.* **90**, 119–130.

Cassen, P.: 1996, 'Models for the fractionation of moderately volatile elements in the solar nebula', *Meteoritics Planet. Sci.* **31**, 793–806.

Clayton, R. N.: 1986, 'High temperature isotope effects in the early solar system'. In *Stable Isotopes in High Temperature Geological Processes*, eds. J. W. Valley, H. P. Taylor and J. R. O'Neil (Washington D.C.: *Mineral. Soc. Am.*), pp. 129–140.

Clayton, R. N.: 1993, 'Oxygen isotopes in meteorites', *Ann. Rev. Earth Planet. Sci.* **21**, 115-149.

Clayton, R. N. and Mayeda T. K.: 1975, 'Genetic relations between the moon and meteorites', *Proc Lunar Sci Conf* **XI**, 1761–1769.

Clayton, R. N., Grossman, L. and Mayeda T. K.: 1973, 'A component of primitive nuclear composition in carbonaceous meteorites', *Science* **182**, 485–487.

Galer, S. J. G. and Goldstein S. L.: 1996, 'Influence of accretion on lead in the Earth'. In: *Isotopic studies of crust-mantle evolution*, A.R. Basu and S.R. Hart, eds., 75-98, AGU.

Galer, S. J. G. and Goldstein, S. L.: 1991, 'Depleted mantle Pb isotopic evolution using conformable ore leads', *Terra Abstr.* **3,** 485–486.

Gast, P.W.: 1960, 'Limitations on the composition of the upper mantle', *J. Geophys. Res.* **65** 1287–1297.

Göpel, C., Manhès G. and Allègre C. J.: 1991, 'Constraints on the time of accretion and thermal evolution of chondrite parent bodies by precise U-Pb dating of phosphates', *Meteoritics* **26**, 73.

Halliday, A. N.: 2000, 'Terrestrial accretion rates and the origin of the Moon', *Earth Planet. Sci. Lett.* **176**, 17–30.

Halliday, A. N. and Porcelli D.: 2001, 'In search of lost planets – the paleocosmochemistry of the inner solar system', *Earth Planet. Sci. Lett.* **192**, 545–559.

Halliday, A. N., Rehkämper M., Lee D-C. and Yi W.: 1996, 'Early evolution of the Earth and Moon: new constraints from Hf-W isotope geochemistry', *Earth Planet. Sci. Lett.* **142**, 75–89.

Halliday, A. N., Wänke H., Birck J-L. and Clayton R. N.: 2001, 'The accretion, bulk composition and early differentiation of Mars', *Space Sci. Rev.* **96**, 197–230.

Halliday, A.N.: 2003, 'The origin and earliest history of the Earth', pp 509–557. In: *Meteorites, comets and planets* (ed. A.M. Davis) Vol. 1 *Treatise on Geochemistry* (eds. H.D. Holland and K.K. Turekian), Elsevier-Pergamon, Oxford.

Halliday, A.N.: 2004, 'Mixing, volatile loss and compositional change during impact-driven accretion of the Earth', *Nature* in press

Hanan, B.B. and Tilton G.R.: 1987, '60025: Relict of primitive lunar crust?' *Earth Planet. Sci. Lett.* **84**, 15–21.

Hanks, T. C. and Anderson D. L.: 1969, 'The early thermal history of the earth', *Phys. Earth Planet. Int.* **2**, 19–29.

Harper, C. L., Nyquist L. E., Bansal B., Wiesmann H. and Shih C.-Y.: 1995, 'Rapid accretion and early differentiation of Mars indicated by ^{142}Nd/^{144}Nd in SNC meteorites', *Science* **267**, 213–217.

Hayashi, C., Nakazawa K and Nakagawa Y.: 1985, 'Formation of the solar system'. In *Protostars and Planets II*, ed. D. C. Black and M. S. Matthews, 1100–1153 (Tucson: University of Arizona Press).

Hayashi, C., Nakazawa K. and Mizuno H.: 1979, 'Earth's melting due to the blanketing effect of the primordial dense atmosphere', *Earth Planet. Sci. Lett.* **43**, 22–28.

Jacobsen, S. B. and Harper C. L. Jr.: 1996, 'Accretion and early differentiation history of the Earth based on extinct radionuclides'. In *Earth Processes: Reading the Isotope Code*, eds. A. Basu and S. Hart. Washington D. C.: AGU, pp. 47–74.

Kamber, B. S. & Collerson, K. D.: 1999, 'Origin of ocean-island basalts: a new model based on lead and helium isotope systematics', *J. Geophys. Res.* **104,** 25479–25491.

Kleine, T., Münker, C., Mezger, K., Palme, H.: 2002, 'Rapid accretion and early core formation on asteroids and the terrestrial planets from Hf-W chronometry', *Nature* **418**, 952–955

Kramers, J. D. and Tolstikhin I. N.: 1997, 'Two terrestrial lead isotope paradoxes, forward transport modeling, core formation and the history of the continental crust', *Chem. Geol.* **139**, 75–110.

Kwon, S.-T., Tilton, G. R. & Grünenfelder, M. H.: 1989, in:*Carbonatites—Genesis and Evolution*, ed. Bell, K.. Unwin-Hyman, London, 360–387.

Lee, D.-C. and Halliday A. N.: 1996, Hf-W isotopic evidence for rapid accretion and differentiation in the early solar system', *Science* **274**, 1876–1879.

Lee, D.-C. and Halliday A. N.: 1997, 'Core formation on Mars and differentiated asteroids', *Nature* **388**, 854–857.

Lee, D-C., Halliday A. N., Leya I., Wieler R. and Wiechert U.: 2002, ' Cosmogenic tungsten and the origin and earliest differentiation of the Moon', *Earth Plan. Sci. Lett*. **198**, 267–274.

Lee, D-C. and Halliday, A.N.: 2000, 'Accretion of primitive planetesimals: Hf-W isotopic evidence from enstatite chondrites', *Science* 288, 1629-1631.

Liew, T. C., Milisenda, C. C. & Hofmann, A. W.: 1991, 'Isotopic constrasts, chronology of element transfers and high-grade metamorphism: the Sri Lanka Highland granulites, and the Lewisian (Scotland) and Nuk (S.W. Greenland) gneisses', *Geol. Rundsch.* **80,** 279–288.

Lugmair, G. W. and Galer S. J. G.: 1992, 'Age and isotopic relationships between the angrites Lewis Cliff 86010 and Angra dos Reis', *Geochim. Cosmochim. Acta* **56**, 1673–1694.

Lugmair, G. W. and Shukolyukov A.: 1998, 'Early solar system timescales according to ^{53}Mn-^{53}Cr systematics', *Geochim. Cosmochim. Acta* **62**, 2863–2886.

Mayor, M. and Queloz, D.: 1995, 'A jupiter-mass companion to a Solar-Type star', *Nature* **378**, 355–359.

Mizuno, H., Nakazawa K. and Hayashi C.: 1980, 'Dissolution of the primordial rare gases into the molten Earth's material', *Earth Planet. Sci. Lett.* **50**, 202–210.

Murphy, D. T., Kamber, B. S. & Collerson, K. D.: 2003, 'A refined solution to the first terrestrial Pb-isotope paradox', *J. Petrol.* **44,** 39–53.

Newsom, H. E.: 1995, 'Composition of the Solar System, Planets, Meteorites, and Major Terrestrial Reservoirs'. In *Global Earth Physics, A Handbook of Physical Constants,* AGU Reference Shelf 1, American Geophysical Union.

O'Neill, H. St. C.: 1991, 'The origin of the Moon and the early history of the Earth - a chemical model; Part I: The Moon', *Geochim. Cosmochim. Acta* **55**, 1135–1158.

Owen, T. and Bar-Nun A.: 2000, 'Volatile contributions from icy planetesimals'. In: *Origin of the Earth and Moon.* R. M. Canup and K. Righter, eds. Univ. of Arizona Press, Tucson, 459–471.

Ozima, M. and Podosek, F. A.: 1999, 'Formation age of Earth from ^{129}I/^{127}I and ^{244}Pu/^{238}U systematics and the missing Xe', *J. Geophys. Res.* **104, B11**, 25493–25499.

Pepin, R. O. and Porcelli, D.: 2002, 'Origin of noble gases in the terrestrial planets', In: *Noble gases in geochemistry and cosmochemistry* (eds. D. Porcelli, C.J. Ballentine and R. Wieler) *Reviews in Mineralogy and Geochemistry* 47, 191–246.

Porcelli, D. and Pepin R. O.: 2000, 'Rare gas constraints on early earth history'. In: *Origin of the Earth and Moon.* R.M.Canup and K.Righter, eds., 435–458. Univ.of Arizona Press, Tucson.

Porcelli, D., Cassen P. and Woolum D.: 2001, 'Deep Earth rare gases: Initial inventories, capture from the solar nebula and losses during Moon formation', *Earth Planet. Sci. Lett* **193**, 237–251.

Rubie, D. C., Melosh H. J., Reid J. E., Liebske C. and Righter K.: 2003, 'Mechanisms of metal-silicate equilibration in the terrestrial magma ocean', *Earth Planet. Sci Lett.,* **205**, 239–255.

Safronov, V.S.: 1954, 'On the growth of planets in the protoplanetary cloud', *Astron. Zh.* **31**, 499–510.

Sasaki, S. and Nakazawa K.: 1986, 'Metal-silicate fractionation in the growing Earth: energy source for the terrestrial magma ocean', *J. Geophys. Res.* **91**, B9231–9238.

Schönberg, R., Kamber B. S., Collerson K. D. and Eugster O.: 2002, 'New W isotope evidence for rapid terrestrial accretion and very early core formation', *Geochim. Cosmochim. Acta* **66,** 3151–3160.

Seager, S.: 2003, 'The search for Earth-like extrasolar planets', *Earth Planet. Sci. Lett.* 208, 113–124.

Shih, C.-Y., Nyquist L. E., Dasch E. J., Bogard D. D., Bansal B. M. and Wiesmann H.: 1993, 'Age of pristine noritic clasts from lunar breccias 15445 and 15455', *Geochim. Cosmochim. Acta* **57**, 915–931.

Solomon, S. C.: 1979, 'Formation, history and energetics of cores in the terrestrial planets', *Earth Planet. Sci. Lett.* **19**, 168–182.

Taylor, S. R.: 1992, *Solar System Evolution A New Perspective*, ed. S. R. Taylor. New York, Cambridge Univ. Press.

Taylor, S. R. and Norman M. D.: 1990, 'Accretion of differentiated planetesimals to the Earth', in: *Origin of the Earth*, H. E. Newsom and J. H. Jones, eds. Oxford University Press, Oxford, pp. 29–43.

Tera, F., Papanastassiou D.A. and Wasserburg G.J.: 1973, 'A lunar cataclysm at ∼3.95 AE and the structure of the lunar crust', *Lun. Planet. Sci. IV*, 723–725.

Trieloff, M., Kunz J., Clague D. A., Harrison D. and Allègre C.J.: 2000, 'The nature of pristine noble gases in mantle plumes', *Science* **288**, 1036–1038.

Wasserburg, G.J., MacDonald F., Hoyle F. and Fowler W. A.: 1964, 'Relative contributions of uranium, thorium, and potassium to heat production in the Earth', *Science* **143** 465–467.

Wetherill, G. W.: 1986: 'Accmulation of the terrestrial planets and implications concerning lunar origin', in *Origin of the Moon*, eds W. K. Hartmann, R. J. Phillips and G. J. Taylor. Houston, Lunar Planetary Institute, pp. 519–550.

Wetherill, G. W.: 1994, 'Provenance of the terrestrial planets', *Geochim. Cosmochim. Acta* **58**, 4513–4520.

Wiechert, U., Halliday A. N., Lee D.-C., Snyder G. A., Taylor L. A. and Rumble D. A.: 2001, 'Oxygen isotopes and the Moon-forming giant impact', *Science* **294**, 345–348.

Yin, Q. Z., Jacobsen S. B., Yamashita K., Blicher-Toft J., Télouk P. and Albarède F.: 2002, 'A short time scale for terrestrial planet formation from Hf-W chronometry of meteorites', *Nature* **418**, 949–952.

Yoshino, T., Walter M. J. and Katsura T.: 2003, 'Core formation in planetesimals triggered by permeable flow', *Nature* **422**, 154–157.

Zartman, R. E. and Haines, S. M.: 1988, 'The plumbotectonic model for Pb isotopic systematics among major terrestrial reservoirs—a case for bi-directional transport', *Geochim. Cosmochim. Acta* **52,** 1327–1339.

Planets Around Neutron Stars

Alex Wolszczan
Penn State University, Department of Astronomy & Astrophysics, 525 Davey Laboratory, University Park, PA 16802, USA

Abstract. The most recent developments in the investigations of the known planets around two millisecond radio pulsars, PSR B1257+12 and PSR B1620-26, and in the searches for dusty disks around pulsars are discussed and related to relevant issues in astronomy of planets around normal stars. An unambiguous determination of masses and orbital inclinations of two of the three planets in the PSR B1257+12 system has been achieved by means of modeling their mean motion resonance condition. A direct detection of the white dwarf companion to PSR B1620-26 has resulted in new, tight constraints on the nature of a giant planet orbiting this binary system. An approximate coplanarity and a near 3:2 resonance of the orbits of the terrestrial-mass planets B and C around PSR B1257+12 strongly suggest a disk origin of this unique planetary system. The existence of a Jovian-mass planet around PSR B1620-26, in a metal-poor environment of the globular cluster M4, raises new questions concerning a relationship between stellar metallicity and the occurrence of planets around normal stars. The available upper limits on infrared emission from dust around pulsars do not exclude a possible existence of circumpulsar disks with masses reaching up to a few hundred $M_{\oplus}$.

1. Introduction

The most obvious connection between the discoveries of the first planetary system beyond the Sun orbiting a radio pulsar (Wolszczan and Frail, 1992) and the first planet around a Sun-like star (Mayor and Queloz, 1995) is the demonstration of a dramatic diversity of the extrasolar planets that would be difficult to anticipate without access to a direct observational evidence. This important point, which has remained perfectly valid over the 10-year period of extrasolar planet investigations, has been captured in the concluding remarks of the Nature discovery report on a planet around 51 Pegasi by Michel Mayor and Didier Queloz, who wrote: "*From a complete planetary system detected around a pulsar, to the rather unexpected orbital parameters of 51 Peg b, searches begin to reveal the extraordinary diversity of possible planetary formation sites*".

The existence of planets around pulsars has consequences, whose significance cannot be overstated. First of all, such planets represent a truly intriguing astrophysical phenomenon which emphasizes a universal character of the planet formation process. Second, by creating planets in such an exotic environment (PSR B1257+12; Wolszczan, 1994) or moving them to it from elsewhere (PSR B1620-26; Sigurdsson et al., 2003), Nature offers us unique testing grounds for planet formation theories. This fortunate circumstance arises from the fact that a superb precision of the pulse timing method makes it possible to detect orbiting bodies as small as asteroids (Wolszczan, 1997) and study subtle dynamical effects such as gravitational perturbations between Earth-mass planets (Wolszczan, 1994) or a

microsecond-order contamination of the pulsar's spin characteristics by a distant long-period planet (e.g. Thorsett et al., 1999). In principle, given enough time, one should be able to achieve a reasonably complete dynamical description of nearly any conceivable planetary system orbiting a neutron star that can be timed with a microsecond precision. Moreover, as terrestrial-mass planets around normal stars cannot be detected with the currently available observing methods, pulsar planets provide an unique tool to study low-mass components of the extrasolar planetary systems.

In this paper, we present the most recent results of the neutron star planets research and place them in a broader context of astronomy of the extrasolar planets. The relevant background information can be found in Wolszczan (1997) and Thorsett et al. (1999). In Section 2, we describe the measurements of masses and orbital inclinations of the PSR B1257+12 planets (Konacki and Wolszczan, 2003) and a study of the long-term stability of this planetary system (Goździewski et al., 2003). The recent results on the PSR B1620-26 triple system in the globular cluster M4 (Sigurdsson et al., 2003) are discussed in Section 3. Advances in searches for protoplanetary and debris disks around pulsars are summarized in Section 4. Finally, in Section 5 we discuss the current understanding of the pulsar planets and present our conclusions.

2. The PSR B1257+12 Planetary System

The 6.2-ms radio pulsar, PSR B1257+12, is orbited by three terrestrial-mass planets forming a compact system that is not much larger than the orbit of Mercury (Wolszczan, 1994). Relative sizes of the orbits and a distribution of masses of planets A, B, and C are strikingly similar to those of the three inner planets in the Solar System (Mazeh and Goldman, 1995). The updated parameters of the PSR B1257+12 planets have been recently published by Konacki and Wolszczan (2003) and are summarized in Table 1. Oscillations of the measured pulse arrival times caused by planets B and C are shown in Figure 1a.

The near 3:2 mean motion resonance (MMR) between planets B and C in the pulsar system and the existence of detectable gravitational perturbations between the two planets (Fig. 1b) (Rasio et al., 1992; Malhotra et al., 1992; Peale, 1993; Wolszczan, 1994; Konacki et al., 1999) provide the mechanism to derive their masses without an a priori knowledge of orbital inclinations. An approximate analytical model which includes the effect of gravitational interactions between planets B and C has been published by Malhotra (1993). Konacki et al. (2000) have developed a new semi-analytical model in which perturbations between the two planets are parametrized in terms of the two planetary masses and the mutual orientation of the orbits with a sufficient precision to make a practical application of this approach feasible. Using the simulated data, they have demonstrated that the planet masses and hence their orbital inclinations can be derived from a

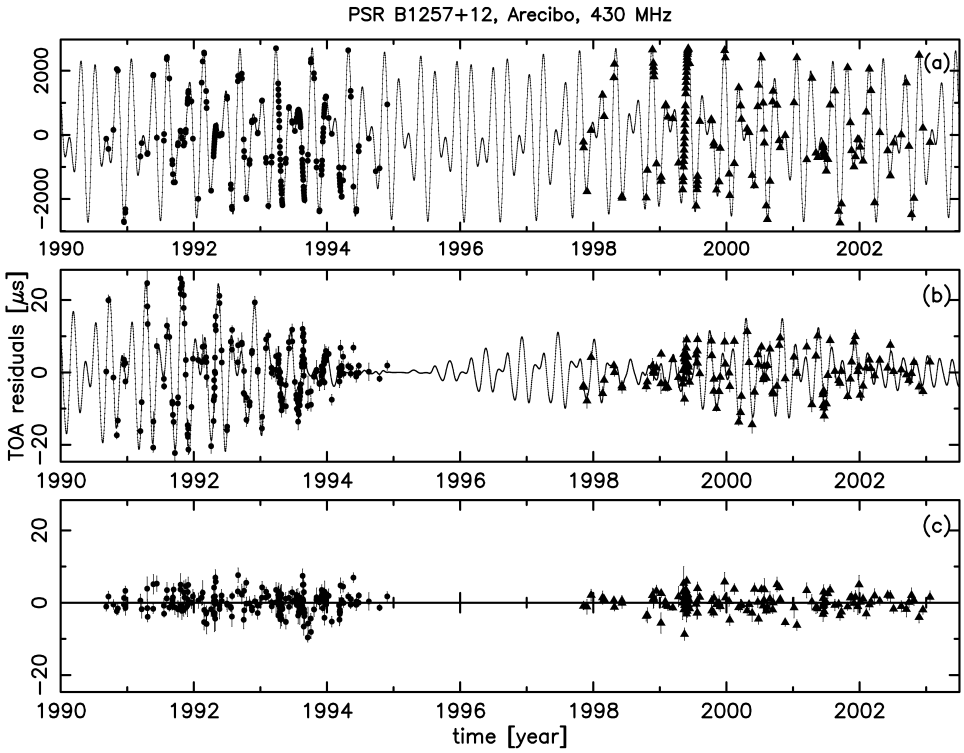

Figure 1. The best-fit, daily-averaged TOA residuals for three timing models of PSR B1257+12 observed at 430 MHz. The solid line marks the predicted TOA variations for each timing model. *(a)* TOA residuals after the fit of the standard timing model without planets. TOA variations are dominated by the Keplerian orbital effects from planets B and C. *(b)* TOA residuals for the model including the Keplerian orbits of planets A, B and C. Residual variations are determined by perturbations between planets B and C. *(c)* Residuals for the model including all the standard pulsar parameters and the Keplerian and non-Keplerian orbital effects.

least-squares fit of this model to the pulse times-of-arrival (TOA) measurements spanning a sufficiently long period of time.

A practical application of this model to the PSR B1257+12 timing data collected with the Arecibo radiotelescope between 1991 and 2003 has been demonstrated by Konacki and Wolszczan (2003). A least-squares fit of the model to data gives the masses of planets B and C of $4.3 \pm 0.2 M_{\oplus}$ and $3.9 \pm 0.2 M_{\oplus}$, respectively, assuming the canonical pulsar mass, $M_{psr} = 1.4 M_{\odot}$. The timing residuals resulting from the best-fit of the perturbation model are shown in Figure 1c. Since the scatter in the known neutron star masses is small (Thorsett and Chakrabarty, 1999), it is unlikely that a possible error in the assumed pulsar mass would significantly affect these results. Because of the $\sin(i)$ ambiguity, there are four possible sets of the orbital inclinations for the planets B and C: (53°, 47°), (127°, 133°) corresponding to the difference in the ascending nodes $\Omega_C - \Omega_B \approx 0°$ (relative inclination $I \approx 6°$), and (53°, 133°), (127°, 47°), corresponding to the difference in the ascending nodes

Table I. Observed and derived parameters of the PSR B1257+12 planets.

Parameter	Planet A	Planet B	Planet C
Projected semi-major axis, x (ms)	0.0030(1)	1.3106(1)	1.4134(2)
Eccentricity, e	0.0	0.0186(2)	0.0252(2)
Epoch of pericenter, T_p (MJD) ..	49765.1(2)	49768.1(1)	49766.5(1)
Orbital period, P_b (d).....	25.262(3)	66.5419(1)	98.2114(2)
Longitude of pericenter, ω (deg) .	0.0	250.4(6)	108.3(5)
Mass ($M_\oplus$).....	0.020(2)	4.3(2)	3.9(2)
Inclination, solution 1, i (deg) ..	...	53(4)	47(3)
Inclination, solution 2, i (deg) ..	...	127(4)	133(3)
Planet semi-major axis, a_p (AU)	0.19	0.36	0.46

$\Omega_C - \Omega_B \approx 180^\circ$ (relative inclination $I \approx 174^\circ$). Obviously, in both cases the planets have nearly coplanar orbits, but in the latter one, their orbital motions have opposite senses. Because our numerical simulations of the system's dynamics show that this situation leads to distinctly different perturbative TOA variations that are not observed, only the first two sets of the orbital inclinations, $53^\circ \pm 4^\circ$ and $47^\circ \pm 3^\circ$ or 127° and 133° are plausible. This implies that the two planets move in nearly coplanar orbits in the same sense.

A precise knowledge of the initial condition of the PSR B1257+12 planetary system derived from the best-fit timing model makes it possible to investigate a dynamical stability of the system by means of a long-term numerical integration of the full equations of motion. Such an integration has been recently performed by Goździewski et al. (2003), who have detected no secular changes of the semi-major axes, eccentricities and inclinations of the planets over a 1 Gyr period. Variations of the orbital elements over the first 500,000 yr are shown in Figure 2. The most notable feature is the presence of a secular apsidal resonance (SAR) between the planets B and C with the center of libration about 180° (Fig. 2d). Curiously, the SAR has been recently found in several extrasolar planetary systems discovered by the radial velocity surveys (Ji et al., 2003). It remains to be established, whether a seemingly common occurrence of the SAR is significant or just coincidental.

In order to fully understand the dynamical stability of the PSR B1257+12 system, Goździewski et al. (2003) have investigated the structure of the phase space in the neighborhood of the initial condition using the Mean Exponential Growth factor of Nearby Orbits (MEGNO; Cincotta and Simó, 2000). MEGNO is the so-called fast indicator that makes it possible to distinguish between regular and chaotic motions during an integration of a system over only 10^4 orbital periods of its outermost planet. This way one can numerically examine a large number of the initial conditions using a reasonable amount of CPU time. As it turns out, the

nominal positions (with respect to the "best-fit" initial condition) of the planets A, B and C are located far from any strong instabilities of the motion in the semi-major axis space. It follows from the full MEGNO analysis that this is also true for the remaining orbital elements which indicates that the system is indeed stable on the Gyr timescale.

3. A Jovian Planet Around PSR B1620-26

The second pulsar orbited by a planet-mass body is a neutron star-white dwarf binary PSR B1620-26 in the globular cluster M4 (Sigurdsson et al., 2003). In this case, the pulsar has a third, substellar-mass companion on a wide, moderately eccentric orbit. The object has been detected, because its dynamical influence on the pulsar induces accelerations that are measurable in the form of higher-order derivatives of its spin period (Backer et al., 1993; Thorsett et al., 1993). The timing analysis of PSR B1620-26 is discussed in detail by Thorsett et al. (1999) and the most recent timing residuals for the pulsar are shown in Figure 3.

A planetary mass of the outer companion to PSR B1620-26 has been recently confirmed by Sigurdsson et al. (2003), who have succeeded in detecting the pulsar's inner, white dwarf companion with the Hubble Space Telescope. A low mass ($0.34 \pm 0.04 M_{\odot}$) and a relatively young age ($4.8 \times 10^8 \pm 1.4 \times 10^8$ years) of the white dwarf confirms a prediction implied by the formation scenario for the PSR B1620-26 system proposed by Sigurdsson (1993). In this scenario, the planet orbiting a main-sequence star in the cluster core survives an exchange interaction with a neutron star binary, in which the original white dwarf companion to the neutron star is replaced by the planet's parent star. Details of a subsequent evolution of the resulting hierarchical triple (Sigurdsson et al., 2003) lead to a prediction that the pulsar's inner companion should be an undermassive, young and relatively bright white dwarf, which is precisely the case.

Given the mass of the pulsar's white dwarf companion one can estimate the orbital inclination of the inner binary (55^{+14}_{-8} degrees). This constrains the semi-major axis and the mass of the outer companion to be 23 AU and $2.5 \pm 1 M_{\rm Jupiter}$, respectively (Sigurdsson et al., 2003). As this mass is obviously in the giant planet range, the PSR B1620-26 planet is not only the oldest one detected so far, but it also represents the only known case of a planet born in a metal-poor environment.

4. Dusty Disks Around Pulsars?

The existence of planet-mass bodies around PSR B1257+12 raises an interesting possibility that at least some pulsars may, like normal stars, be accompanied by protoplanetary or debris disks. All theories of planet formation around neutron stars assume a creation of some sort of a protoplanetary disk out of the material supplied by the pulsar's binary companion or, possibly, from the fallback of supernova

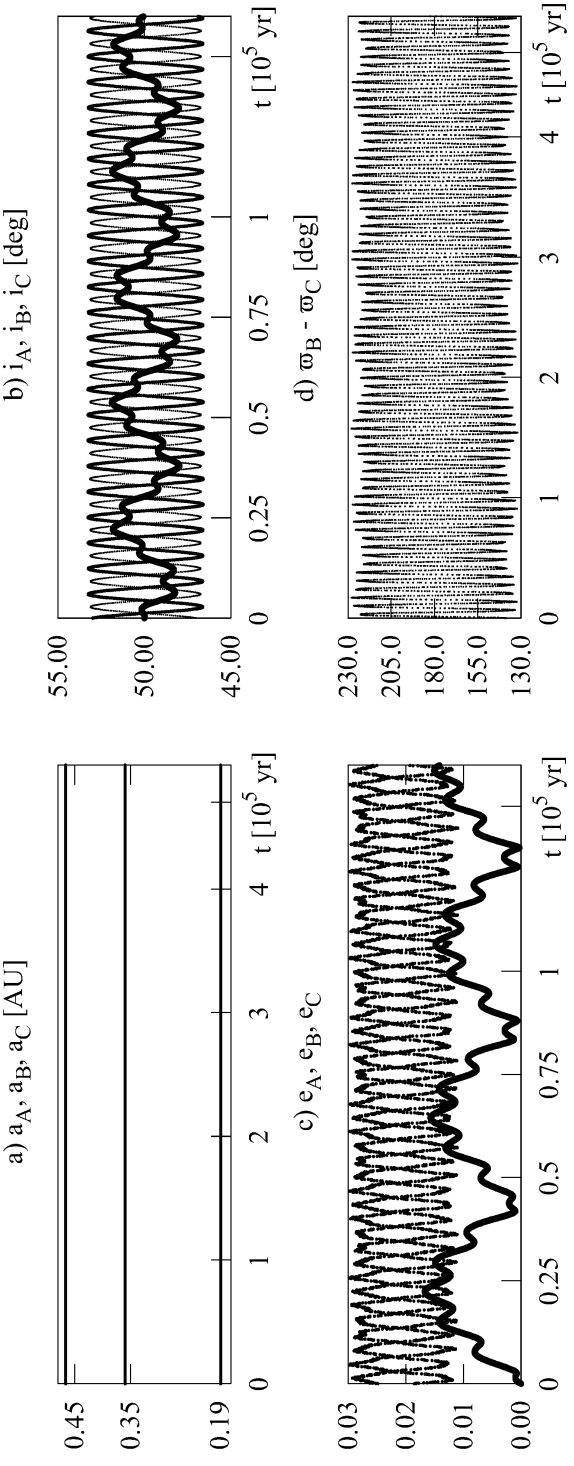

Figure 2. Time variations of *(a)* the semi-major axes, *(b)* the orbital inclinations, and *(c)* the eccentricities of the orbits of the PSR B1257+12 planets over the first 500,000 yr. Panel *(d)* illustrates the secular apsidal resonance between the planets B and C parametrized by the critical argument $\theta = \varpi_B - \varpi_C$. $\varpi = \omega + \Omega$ is the longitude of periastron; Ω and ω are the longitude of ascending node and the argument of periastron of the planet, respectively (courtesy of K. Goździewski).

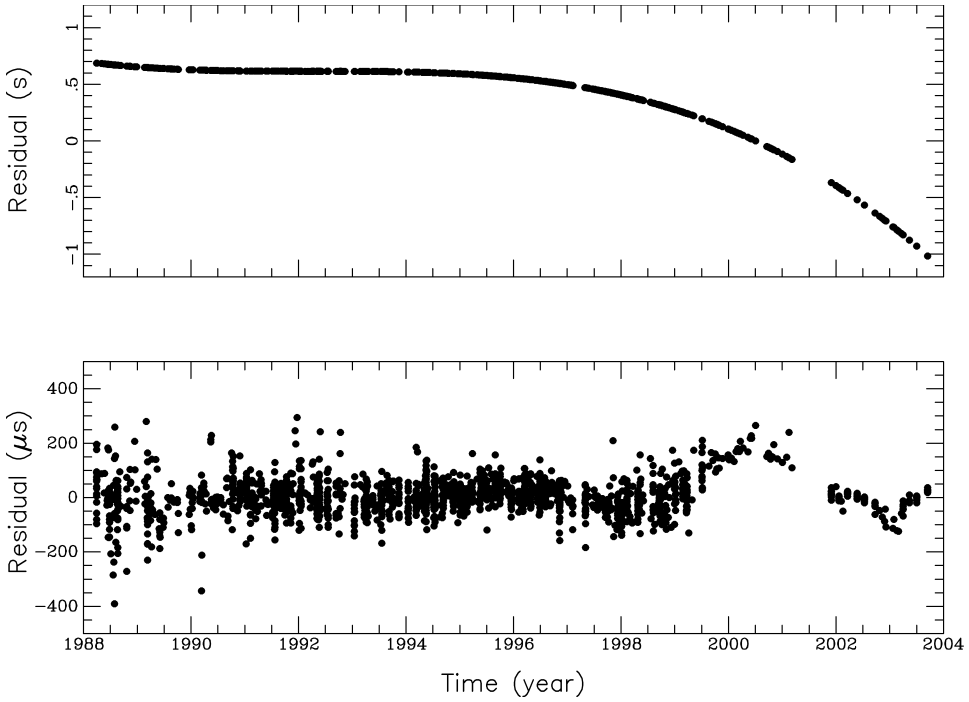

Figure 3. Results of the pulse timing analysis for PSR B1620-26. *(Upper panel)* Residuals after a fit for the spin period, its derivative, standard astrometric parameters, and the inner binary orbit. *(Lower panel)* Residuals after including a fit for five period derivatives to model the outer, planetary orbit (courtesy of A. Lyne, I. Stairs, and S. Thorsett).

ejecta (e.g. Podsiadlowski, Phinney and Hansen, 1993, 1993). Of course, properties and composition of such disks may be significantly different from those of the commonly observed disks around T Tauri stars. In addition, the pulsar emission irradiating these hypothetical disks mostly consists of a wind of ultrarelativistic particles. Although the total spindown luminosities of pulsars are, in principle, more than sufficient to heat the dust grains, the efficiency of this process depends on a number of factors that are difficult to predict in absence of sufficient observational constraints (Miller and Hamilton, 2001).

To set the basic constraints on the circumpulsar disk observability with a minimum set of assumptions, one can apply the approach of Foster and Fischer (1996) by assuming that a fraction f of the pulsar spindown luminosity L heats N dust grains of size a to a temperature T, and calculate the expected infrared flux as a function of disk parameters and pulsar distance from $fL \sim 4\pi a^2 N\sigma T^4$. These estimates can be compared with the upper flux limits derived from the previous attempts to detect the hypothetical circumpulsar disks with both the space- and the ground-based telescopes at wavelengths ranging from 10 μm to 3 mm (van Buren and Terebey, 1993 (IRAF); Zuckerman, 1993 (IRTF); Phillips and Chandler, 1994 (JCMT, OVRO); Foster and Fischer, 1996 (IRTF); Graeves and Holland,

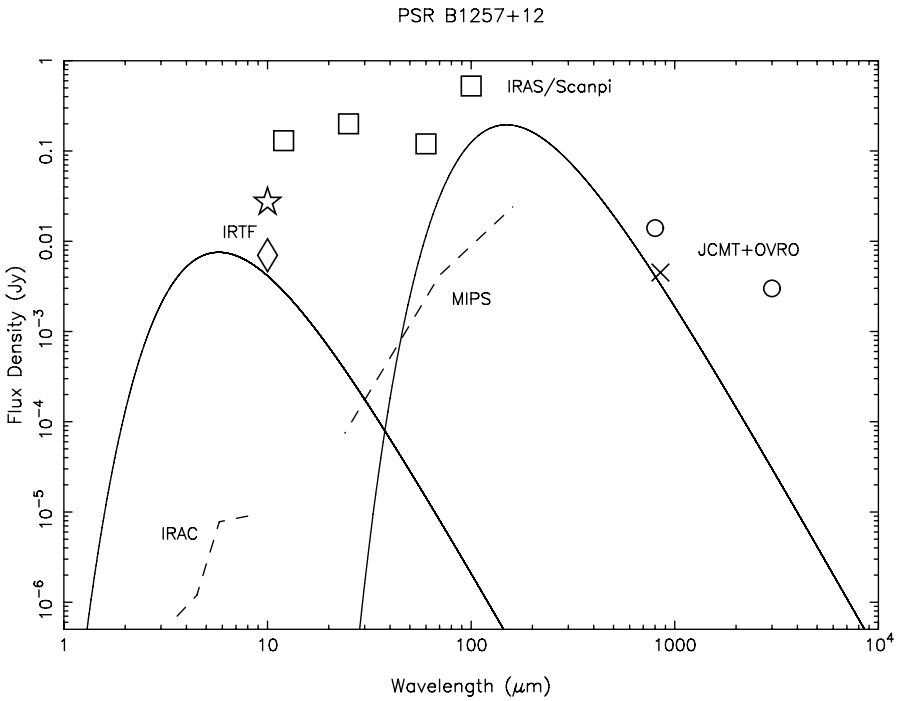

Figure 4. Observational limits and theoretical models for the emission from a hypothetical dusty disk in the PSR B1257+12 planetary system. Symbols mark the flux limits set by observations, and the solid curves denote examples of theoretical models, as explained in the text. Dashed lines correspond to the *Spitzer Space Telescope* 3σ sensitivity limits calculated for the IRAC and the MIPS detectors (http://sirtf.caltech.edu/SSC).

2000 (JCMT); Koch-Miramond et al., 2002 (ISO, IRAF/Scanpi); Lazio et al., 2002 (ISO)).

As an example of typical sensitivities achieved in the previous observations, we summarize the upper flux limits obtained for PSR B1257+12 in Fig. 4. For comparison, we also include theoretical curves for a $T = 20$ K, $300 M_\oplus$ disk consisting of 1 μm grains, and a $T = 500$ K, $10^{-5} M_\oplus$ ($\sim$1/10 solar system asteroid mass) disk with a 0.1 μm grain size, for the PSR B1257+12 spindown luminosity $L = 2 \times 10^{34}$ erg s^{-1} and distance $d = 0.62$ kpc. Neither of the two models violates the constraints set by observations, but, of course, one could achieve the same result with other physically plausible realizations of a circumpulsar disk.

5. Discussion

The existence of planets around the millisecond pulsars PSR B1257+12 and PSR 1620-26 has important consequences for astronomy of extrasolar planets. A near 3:2 MMR between the orbits of planets B and C in the PSR B1257+12 system

and the fact that they are nearly coplanar imply that this pulsar's planetary system has been created as the result of a disk evolution similar to that invoked to describe planet formation around normal stars (Boss, 2003). Together with the true mass measurements of the planets and the long-term system stability calculations described in Section 2, these results offer a fairly complete dynamical characterization of the only known extrasolar planetary system containing terrestrial-mass planets. Its existence provides a demonstration that protoplanetary disks beyond the Sun can evolve Earth-mass planets to a dynamical configuration that is similar to that of the planets in the inner solar system.

The Jupiter-mass planet detected around PSR B1620-26, in an extremely metal-poor environment of the globular cluster M4, represents an outstanding exception to the growing evidence for a positive correlation between the occurence of planets and the metallicity of their parent stars (Santos et al., 2001; Santos et al., 2003; Fischer et al., 2003). Of course, more such detections would be needed to fully understand the implications of this discrepancy. In particular, is must be remembered that no planets have been detected in a transit survey of the cluster 47 Tucanae, which is a factor 5 more metallic than the M4 cluster (Gilliland et al., 2000). One interesting possibility raised by Sigurdsson et al. (2003) is that an enhanced metallicity of protoplanetary disks could encourage planet migration, which would explain the lack of short period planets in 47 Tuc. Nevertheless, the PSR B1620-26 planet provides a strong observational suggestion that the currently postulated evidence for a dependence of the frequency of planets on the metal content of their parent systems may include biases that are not yet understood.

The early theories of the PSR B1257+12-type planet formation have been summarized by Podsiadlowski (1993) and further discussed by Phinney and Hansen (1993). More recently, Miller and Hamilton (2001) and Hansen (2002) have examined the conditions of survival and evolution of pulsar protoplanetary disks. They have concluded that an initially sufficiently massive ($> 10^{28}$ g) disk would be able to resist evaporation by the pulsar accretion flux and create planets on a typical, $\sim 10^7$-year timescale. A quick formation of a massive disk around the pulsar could, for instance, be accomplished by tidal disruption of a stellar companion (Stevens et al., 1992; Phinney and Hansen, 1993) or, possibly, in the process of a white dwarf merger (Podsiadlowski et al., 1991; Livio et al., 1992). Both these processes, although entirely feasible, cannot be very common. In fact, with the exception of PSR B1257+12, no planetary companions have emerged from the precision timing of the known galactic millisecond pulsars (Lorimer, 2001), implying their rarity, independently of the specific formation mechanism.

In the case of the PSR B1620-26 planet, its origin described by Sigurdsson et al. (2003) is dynamically entirely different from the scenarios envisioned for the PSR B1257+12 system. According to Sigurdsson et al., this Jupiter-sized planet has originally formed around a main-sequence progenitor in the core of the M4 cluster and, after an exchange interaction with a neutron star- white dwarf binary, it has been ejected from the core, assuming a wide, eccentric orbit around a newly

formed inner binary, now including the planet's parent star. Obviously, as such interactions must be quite infrequent, even in the clusters with very dense cores (Sigurdsson, 1993), it will be very difficult to find additional examples of planets around globular cluster pulsars. In any case, it will be worthwhile to continue the precision timing programs of pulsars in globular clusters and search for possible planet signatures in the timing residuals. Similarly, it is possible that the statistics of planets around the field millisecond pulsars will eventually improve as the result of continuing pulsar surveys and the followup timing observations.

As shown in Fig. 4 and discussed, for example, by Graeves and Holland (2000) and Koch-Miramond et al. (2002), the existing IR flux limits for PSR B1257+12 (and a few other pulsars) do rule out massive, $\sim 0.01 M_{\odot}$ disks similar to those thought to give rise to planets around normal stars (Boss, 2003). However, these limits do not contradict a possibility that some pulsars may be accompanied by much less massive disks with masses ranging from a fraction of the asteroid belt mass to a few hundred $M_{\oplus}$ over a wide range of temperatures and grain sizes. This conclusion is consistent with the masses of the PSR B1257+12 planets and their disk origin deduced by Konacki and Wolszczan (2003) from timing observations of the pulsar, and with the theoretical constraints on neutron star planet formation scenarios discussed above.

In the absence of new detections of planets around millisecond pulsars further searches for dust around these objects, covering the physically plausible parameter space at a possibly high sensitivity level, is the logical way to meaningfully constrain the models of a creation and evolution of pulsar protoplanetary disks. As shown in Fig. 4, the *Spitzer Space Telescope*, with its factor of 10^2–10^3 improvement in sensitivity, compared to the instruments previously used to search for dust emission from around pulsars, is an obvious choice for further exploration of the physics of neutron star planetary systems.

Acknowledgements

This work was supported by the NASA grant NAG5-13620 and the NSF grant PHY99-07949. The author whishes to thank Drs A. Lyne, I. Stairs and S. Thorsett for permission to display their unpublished timing data on PSR B1620-26 in Figure 3, and Drs K. Goździewski and M. Konacki for their contribution to the results presented in this paper. Arecibo Observatory is part of the National Astronomy and Ionosphere Center, which is operated by Cornell University under contract with the National Science Foundation.

References

Backer, D. C., Foster, R. S. and Sallmen, S.: 1993, 'A Second Companion of the Millisecond Pulsar 1620-26', *Nature* **365**, 817–818.

Boss, A. P.: 2003, 'Progress in Giant Planet Formation' ASP Conf. Ser.: *Scientific Frontiers in Research on Extrasolar Planets* **294**, 269-276.
Cincotta, P. M. and Simó, C. 2000, 'Simple Tools to Study Global Dynamics in Non-Axisymmetric Galactic Potentials - I', *A&AS* **147**, 205–228.
Fischer, D., Valenti, J. A. and Marcy, G.: 2003, 'Spectral Analysis of Stars on Planet-Search Surveys', IAU Symp. 219: *Stars as Suns: Activity, Evolution and Planets*, in press.
Foster, R. S. and Fischer, J.: 1996, 'Search for Protoplanetary and Debris Disks Around Millisecond Pulsars', *ApJ* **460**, 902–905.
Gilliland, R. L., Brown, T. M., Guhathakurta, P., Sarajedini, P. et al.: 2000, 'A Lack of Planets in 47 Tucanae from a Hubble Space Telescope Search', *ApJ* **545**, L47–L51.
Goździewski, K., Konacki, M., and Wolszczan, A.: 2003, 'Long-Term Stability and Dynamical Environment of the PSR 1257+12 Planetary System', astro-ph/0310750.
Graeves, J. S. and Holland, W. S.: 2000, 'A Search for Protoplanetary Discs Around Millisecond Pulsars', *MNRAS*, **316**, L21–L23.
Hansen, B. M. S.: 2002, 'Stellar Collisions and Pulsar Planets', ASP Conf. Ser.: *Stellar Collisions, Mergers and their Consequences* **263**, 221–227.
Ji, J., Liu, L., Kinoshita, H., Zhou, J., Nakai, H., and Li, G.: 2003, 'The Librating Companions in HD 37124, HD 12661, HD 82943, 47 Ursa Majoris, and GJ 876: Alignment or Antialignment?', *ApJ* **591**, L57–L60.
Koch-Miramond, L., Haas, M., Pantin, E., Podsiadlowski, Ph., Naylor, T. and Sauvage, M.: 2002, 'Determination of Limits on Disc Masses Around Six Pulsars at 15 and 90 μm', *AA* **387**, 233–239.
Konacki, M., Maciejewski, A. J., and Wolszczan, A.: 1999, 'Resonance in PSR B1257+12 Planetary System', *ApJ* **513**, 471–476.
Konacki, M., Maciejewski, A. J., and Wolszczan, A.: 2000, 'Improved Timing Formula for the PSR B1257+12 Planetary System', *ApJ* **544**, 921–926.
Konacki, M. and Wolszczan, A.: 2003, 'Masses and Orbital Inclinations of Planets in the PSR B1257+12 System', *ApJ* **591**, L147-L150.
Lazio, T. J. W., Fischer, J. and Foster, R. S.: 2002, 'Far-Infrared ISO Limits on Dust Disks Around Millisecond Pulsars', *ISOPHOT Workshop on P32O Oversampled Mapping*, ESA SP-482, 83–85.
Livio, M., Pringle, J. E., and Saffer, R. A.: 1992, 'Planets Around Massive White Dwarfs', *MNRAS* **257**, 15P–16P.
Lorimer, D. R.: 2001, 'Binary and Millisecond Pulsars at the New Millenium', *Living Reviews in Relativity* **4**, 5–5.
Malhotra, R., Black, D., Eck, A., and Jackson, A.: 1992, 'Resonant Orbital Evolution in the Putative Planetary System of PSR 1257+12', *Nature* **356**, 583.
Malhotra, R.: 1993, 'Three-Body Effects in the PSR 1257+12 Planetary System', *ApJ* **407**, 266–275.
Mayor, M. and Queloz, D.: 1995, 'A Jupiter-Mass Companion to a Solar-Type Star', *Nature* **378**, 355–357.
Mazeh, T. and Goldman, I.: 1995, 'Similarities Between the Inner Solar System and the Planetary System of the PSR 1257+12', *PASP* **107**, 250–250.
Miller, M. C. and Hamilton, D. P.: 2001, 'Implications of the PSR 1257+12 Planetary System for Isolated Millisecond Pulsars', *ApJ* **550**, 863–870.
Peale, S. J.: 1993, 'On the Verification of the Planetary System Around PSR 1257+12', *AJ* **105**, 1562–1570.
Phillips, J. A, and Chandler, C. J.: 1994, 'A Search for Circumstellar Material Around Pulsars', *ApJ*, **420**, L83-L86.
Phinney, E. S. and Hansen, B. M. S.: 1993, 'The Pulsar Planet Production Process', ASP Conf. Ser.: *Planets Around Pulsars* **36**, 371–390.

Podsiadlowski, P., Pringle, J. E., and Rees, M. J.: 1991, 'The origin of the Planet Orbiting PSR 1829-10', *Nature* **352**, 783–784.
Podsiadlowski, P.: 1993, 'Planet Formation Scenarios', ASP Conf. Ser.: *Planets Around Pulsars* **36**, 149–165.
Rasio, F. A., Nicholson, P. D., Shapiro, S. L., and Teukolsky, S. A.: 1992, 'An Observational Test for the Existence of a Planetary System Orbiting PSR 1257+12', *Nature* **355**, 325–326.
Santos, N. C., Israelian, G. and Mayor, M.: 2001, 'The Metal-Rich Nature of Stars with Planets', *AA* **373**, 1019–1031.
Santos, N. C., Israelian, G. and Mayor, M., Rebolo, R. and Udry, S.: 2003, 'Statistical Properties of Exoplanets. II. Metallicity, Orbital Parameters, and Space Velocities', *AA* **398**, 363–376.
Sigurdsson, S.: 1993, 'Genesis of a Planet in Messier 4', *ApJ* **415**, L43-L46.
Sigurdsson, S., Richer, H. B., Hansen, B. M., Stairs, I. H. and Thorsett, S. E.: 2003, 'A Young White Dwarf Companion to Pulsar B1620-26: Evidence for Early Planet Formation', *Science* **301**, 193–196.
Stevens, I. R., Rees, M. J., and Podsiadlowski, P.: 1992, 'Neutron Stars and Planet-Mass Companions', *MNRAS* **254**, 19P–22P.
Thorsett, S. E., Arzoumanian, Z. and Taylor, J. H.: 1993, 'PSR B1620-26–A Binary Radio Pulsar with a Planetary Companion?', *ApJ* **412**, L33–L36.
Thorsett, S. E. and Chakrabarty, D.: 1999, 'Neutron Star Mass Measurements. I. Radio Pulsars', *ApJ* **512**, 288–299.
Thorsett, S. E., Arzoumanian, Z., Camilo, F. and Lyne, A. G.: 1999, 'The Triple Pulsar System PSR B1620-26 in M4', *ApJ* **523**, 763–770.
van Buren, D. and Terebey, S.: 1993, 'IRAS Sources Near Positions of Pulsars', ASP Conf. Ser.: *Planets Around Pulsars* **36**, 327–333.
Wolszczan, A. and Frail, D. A.: 1992, 'A Planetary System Around the Millisecond Pulsar PSR 1257+12', *Nature* **355**, 145–147.
Wolszczan, A.: 1994, 'Confirmation of Earth Mass Planets Orbiting the Millisecond Pulsar PSR B1257+12', *Science* **264**, 538–542.
Wolszczan, A.: 1997, 'Searches for Planets Around Neutron Stars', *Celestial Mechanics and Dynamical Astronomy* **68**, 13–25.
Zuckerman, B.: 1993, 'Dusty Debris Clouds Around Main Sequence and Post-Main Sequence Stars', ASP Conf. Ser.: *Planets Around Pulsars* **36**, 303–315.

Systems of Multiple Planets

G. W. Marcy and D. A. Fischer
University of California, Berkeley, Dept. of Astronomy, Berkeley, CA 94720 USA

R. P. Butler
Department of Terrestrial Magnetism, Carnegie Institution of Washington, 5241 Broad Branch Road, NW, Washington DC 20015, USA

S. S. Vogt
UCO Lick Observatories, University of California, Santa Cruz, Santa Cruz, CA 95064

Abstract. To date, 10 stars are known which harbor two or three planets. These systems reveal secular and mean motion resonances in some systems and consist of widely separated, eccentric orbits in others. Both of the triple planet systems, namely Upsilon And and 55 Cancri, exhibit evidence of resonances. The two planets orbiting GJ 876 exhibit both mean–motion and secular resonances and they perturb each other so strongly that the evolution of the orbits is revealed in the Doppler measurements. The common occurrence of resonances suggests that delicate dynamical processes often shape the architecture of planetary systems. Likely processes include planet migration in a viscous disk, eccentricity pumping by the planet–disk interaction, and resonance capture of two planets.

We find a class of "hierarchical" double–planet systems characterized by two planets in widely separated orbits, defined to have orbital period ratios greater than 5 to 1. In such systems, resonant interactions are weak, leaving high–order interactions and Kozai resonances plausibly important. We compare the planets that are single with those in multiple systems. We find that neither the two mass distributions nor the two eccentricity distributions are significantly different. This similarity in single and multiple systems suggests that similar dynamical processes may operate in both. The origin of eccentricities may stem from a multi–planet past or from interactions between planets and disk. Multiple planets in resonances can pump their eccentricities pumping resulting in one planet being ejected from the system or sent into the star, leaving a (more massive) single planet in an eccentric orbit. The distribution of semimajor axes of all known extrasolar planets shows a rise toward larger orbits, portending a population of gas–giant planets that reside beyond 3 AU, arguably in less perturbed, more circular orbits.

1. Introduction

It is remarkable to recall that only seven years ago, the totality of our knowledge about extrasolar planets resided in a modest plot of velocity versus time for the star 51 Pegasi (Mayor and Queloz 1995). We remain impressed that Michel and Didier abstained from a definitive interpretation of those velocities until the completion of a wide variety of studies of that star. For 1 1/2 years after first detecting the sinusoidal velocity variations, they carried out photometry and line shape measurements to search for stellar pulsations, and they examined the rotation of the star to rule out a stellar companion which would have tidally spun up the star. Thus, in their discovery paper they had fully anticipated and eliminated the two serious alternative explanations that others would lance against the planet interpretation.

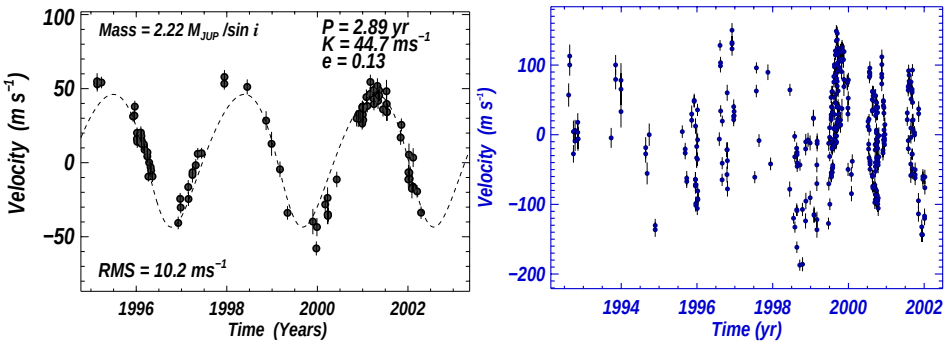

Figure 1. Evidence for multiple planets. Left: A fit to the measured velocities of 47 UMa that invokes only one planet. Coherent departures are obvious in 1997 and 2000, indicating the inadequacy of a single planet. Right: Measured velocities vs time for Upsilon Andromedae, also showing velocity variations on multiple time scales

Our team followed a month later with next two planets, namely around 70 Vir and 47 UMa that created further disturbances (Butler and Marcy, 1996; Marcy and Butler, 1996). For the former, the orbital eccentricity of 0.4 demonstrated unambiguous Keplerian motion, eliminating stellar pulsation as an explanation. But that planet violated our expectations of circular orbits for planets. For 47 UMa the early Keplerian fit, while implying a nearly circular (e = 0.06) orbit, failed to follow the predictions of that orbit in subsequent velocity measurements (Fig. 1). Coherent departures from the predicted Keplerian curve are apparent both in 1997 and 2000 (Fig. 1). Even greater departures were seen in the predicted Keplerian orbit in the next planet discovered, namely around Upsilon And (Butler et al., 1999). Moreover the fitted velocities for the planet around 55 Cancri exhibited a trend in the residuals. Thus, among the first six planets discovered (including Tau Boo b), three exhibited evidence of multiple companions.

To date, 10 stars are known for which the velocities are best interpreted by two or three planets, as listed in Table 1. Among them are five planetary systems apparently trapped in mean–motion and secular resonances.

2. Upsilon Andromedae

Upsilon Andromedae was the first multiple-planet system discovered (Butler et al., 1999). Its velocities (from Lick Obs.) are shown in Fig. 1 obtained from 1993–2002. A power spectrum (Fig. 2) shows a strong peak at 4.6 days, the orbital period of the inner planet. Fitting the velocities with a simple Keplerian orbit leaves residuals that reveal two additional periodicities, visible to the eye (Fig. 2). To find a solution, we collaborated with R.Noyes, T.Brown, and S.Korzennik who had obtained excellent velocity measurements with their AFOE spectrometer (Brown et al., 1998).

Table I. Properties of the 10 known multiple–planet systems.

Star	$M_{\rm Star}$ ($M_\odot$)	P (d)	K (m/s)	ecc.	$M \sin i$ ($M_{\rm JUP}$)	a (AU)	Comment
Ups And b	1.30	4.617	70.15	0.01	0.64	0.058	
Ups And c	1.30	241.16	53.93	0.27	1.79	0.805	apsidal lock
Ups And d	1.30	1276.15	60.62	0.25	3.53	2.543	apsidal lock
55 Cnc b	1.03	14.653	71.5	0.03	0.83	0.115	3:1 Resonance
55 Cnc c	1.03	44.3	11.2	0.40	0.18	0.241	3:1 Resonance
55 Cnc d	1.03	4400	50.2	0.34	3.69	5.2	
GJ 876 b	0.32	61.020	210.0	0.10	1.89	0.207	2:1 Mean Motion &
GJ 876 c	0.32	30.120	81.0	0.27	0.56	0.130	Secular Resonance
47 UMa b	1.03	1079.2	55.6	0.05	2.86	2.077	8:3 commensur.?
47 UMa c	1.03	2845.0	15.7	0.00	1.09	3.968	8:3 commensur.?
HD 168443 b	1.01	58.1	470.0	0.53	7.64	0.295	
HD 168443 c	1.01	1770.0	289.0	0.20	16.96	2.873	
HD 37124 b	0.91	153.3	35.0	0.10	0.86	0.543	
HD 37124 c	0.91	1942.0	19.0	0.60	1.00	2.952	
HD 12661 b	1.07	263.6	74.4	0.35	2.30	0.823	Secular Resonance
HD 12661 c	1.07	1444.5	27.6	0.20	1.57	2.557	Secular Resonance
HD 38529 b	1.39	14.309	54.2	0.29	0.78	0.129	
HD 38529 c	1.39	2174.3	170.5	0.36	12.7	3.68	
HD 82943 b	1.05	444.6	46.0	0.41	1.63	1.159	2:1 Resonance
HD 82943 c	1.05	221.6	34.0	0.54	0.88	0.728	2:1 Resonance
HD 74156 b	1.05	51.6	112.0	0.65	1.61	0.278	
HD 74156 c	1.05	>2650	125.0	0.35	>8.21	>3.82	

Occam's Razor suggests that the first multi–planet system would contain the minimum complexity: two detectable planets. This logic dissuaded us from invoking three planets until the residuals left no alternative. By 1999, with reluctance and many tests (Monte Carlo and comparisons with other F8V stars) we were forced to invoke three planets for Upsilon And (Butler et al., 1999) to explain the velocities within errors. The current fit to all three planets is shown in Fig. 2. The subsequent velocity measurements at Lick observatory have followed the predicted triple-Keplerian orbit within the errors. No planet–planet interactions nor additional planets are revealed in our recent velocities.

Interestingly, the line of apsides of the orbits of the outer two planets are nearly aligned, with current values of ω of 246° and 259°, each with uncertainty of 4°. Theoretical simulations indicates that libration of the apsides likely stems from a secular resonance. Treating the planets as smeared out mass distributions of "wire" along their orbits implies torques that cause oscillation of the orientation of the ma-

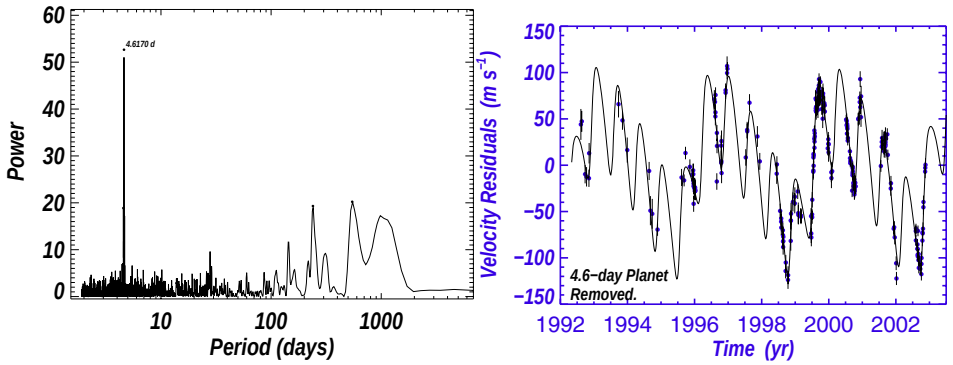

Figure 2. Analysis for three planets orbiting Upsilon And. Left: The power spectrum of velocities shows several significant peaks, notably at 4.6 d, 2/3 yr, and several years. Right: Best two–planet fit to the velocities, after subtracting the effects of the inner (4.6 d) planet. Velocities have followed the prediction from a three–planets model for three years.

jor axes (Chiang and Murray, 2002). Such a secular resonance indicates that some gentle migration from their original orbits delicately cradled them into their current configuration (Chiang and Murray, 2002; Kley 2000; Nelson and Papaloizou, 2002; Kley, 2003).

3. GJ 876

Star 876 in the Gliese-Jahreiss Catalog of Nearby stars has spectral type M4V and mass 0.35 $M_{\odot}$. Its measured velocities can be fit with a model of two planets orbiting in simple (independent) Keplerian motion around the star (Marcy et al., 2001). The orbital periods are 30.1 and 61.0 days, indicating the possibility of a dynamical resonance. A 2:1 mean–motion resonance is indicated by simulations of the two-planet system that include the planet-planet interactions (Lissauer and Rivera, 2001; Rivera and Lissauer, 2001; Lee and Peale, 2002; Nelson and Papaloizou, 2002). The planets will librate about this 2:1 ratio of orbital periods indefinitely.

The two planets perturb each other on a time scale of years, detectable in the existing measurements. Figure 3 shows the best fit to the velocities achievable with a model that includes two independent Keplerian orbits. The value of $\sqrt{\chi^2}$ is 2.4, and the RMS of the residuals is 11 m s^{-1}, both indicating that the residuals are $\sim$2.4$\times$ the known errors and the fit is inadequate.

A model that includes the planet–planet interactions lowers the value of $\sqrt{\chi^2}$ from 2.4 to 1.5 over the model that does not include perturbations (Lissauer and Rivera, 2001; Laughlin and Chambers, 2001). Moreover, the planets cannot be arbitrarily massive, as the perturbations are modest in magnitude (roughly 10 m s^{-1} during several years). Thus the orbit plane cannot be arbitrarily face-on, and simulations limit the value of $\sin i$ to be 0.5–0.8 (Rivera and Lissauer, 2001; Laughlin

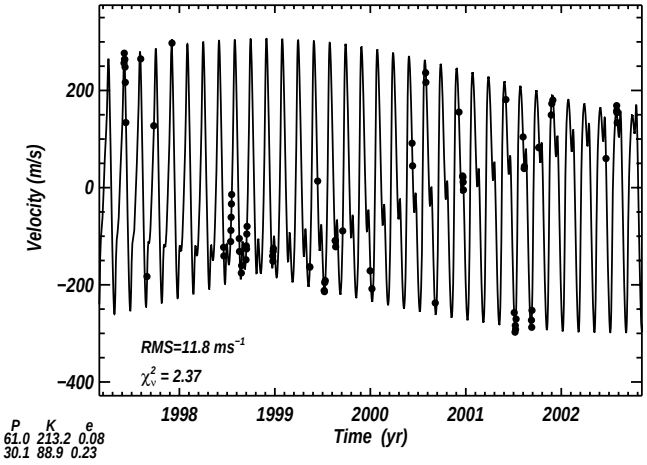

Figure 3. Velocities and two–Keplerian fit for GJ 876. The high value of χ^2 and RMS are a result of planet–planet interactions.

and Chambers, 2001). Recent astrometric measurements of the wobble of GJ 876 find an orbital inclination for the outer planet of $84° \pm 10°$, implying a mass of $1.89 \pm 0.34 M_{\rm JUP}$ (Benedict et al. 2002).

4. 55 Cancri

The G8 main sequence star 55 Cnc revealed a velocity periodicity of 14.6 d and $M \sin i = 0.9$, but also revealed a trend in the velocity residuals to the Keplerian fit (Butler et al., 1997). These velocity residuals rose from 1989 until 1996, declining thereafter. Intense monitoring (especially by D.A.F. during the past four years) reveals a minimum in the velocity residuals, indicating closure of the outer orbit (see Fig. 4 and Marcy et al., 2002).

A two–planet fit to 55 Cancri leaves residuals that exhibit a periodicity of 44 days, possibly caused by a third planetary companion. However, the rotation period of the star is 35–42 days, as shown by Henry et al. (2000) from periodicities in the CaII H&K line. Thus a clear danger exists that the 44–day period in the velocities is caused by surface inhomogeneities on the stellar surface rotating into and out of the view. The slow rotation period of the star and its low level of chromospheric activity suggest that surface inhomogeneities will produce intrinsic velocity "jitter" of no more than 5 m s^{-1}, lower than the observed velocity variations of 14 m s^{-1} associated with the 44-day period. Thus, it is unlikely, but not impossible, that the 44-day period is simply rotation modulation of surface features. The 44-day velocity periodicity represents, most likely, a third planet in the system. Greg Laughlin has shown that this "middle" planet, if real, not only would be stable against the

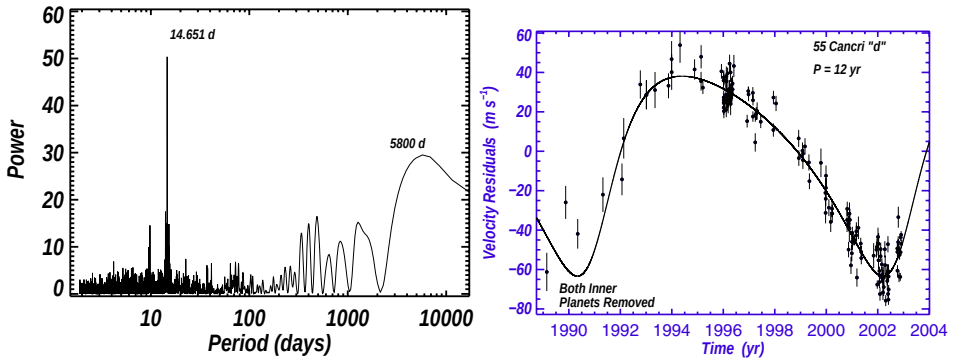

Figure 4. Three planets orbiting 55 Cancri. Left: periodogram of raw velocities versus time, showing peaks at 14.6 d and several longer periods. Right: The best three–planet fit to the velocities, after subtracting the inner two planets. The third planet resides at ∼5.5 AU (Marcy et al., 2002).

perturbations of the other two planets, but that it resides in a near (but not secure) 3:1 resonance with the inner planet of period 14.6 d (Marcy et al., 2002).

The 55 Cancri system has an outer planet at nearly 6 AU, not unlike that of our Jupiter at 5.2 AU. However the architecture of the 55 Cnc system as a whole, with its two inner jupiter–mass planets, clearly differs from that of our Solar System. Some migration may have occurred with those inner planets, but not with the outer one, for reasons not understood.

5. 47 UMa

The G0V star 47 UMa has velocities from Lick Observatory dating from 1988 to 2002 (see Fig. 5). Coherent departures from the single–planet Keplerian fit are apparent in 1997 and 2000, as well as at other epochs. A double–Keplerian model immediately resolves these departures, fitting the velocities within errors (Fischer et al., 2002). The orbital periods are 2.9 and 7.6 yr and the masses are $M \sin i$ = 2.6 and 0.8 $M_{\rm JUP}$, respectively. Both orbits are nearly circular, with the inner orbit having $e = 0.06$ and the outer orbit having $e < 0.2$.

The period ratio of the two planets orbiting 47 UMa is not close to any small–number integer ratio, but one cannot rule out the role of higher–order mean–motion resonances nor secular resonances (Laughlin et al., 2002). Terrestrial–sized bodies in orbits within 1 AU would likely be dynamically stable (Laughlin et al., 2002).

6. Hierarchical Planetary Systems

Some Doppler measurements initially reveal only a single planet, but also exhibit long–term trends in the residuals. The star 55 Cnc (Fig. 4) was the first such case.

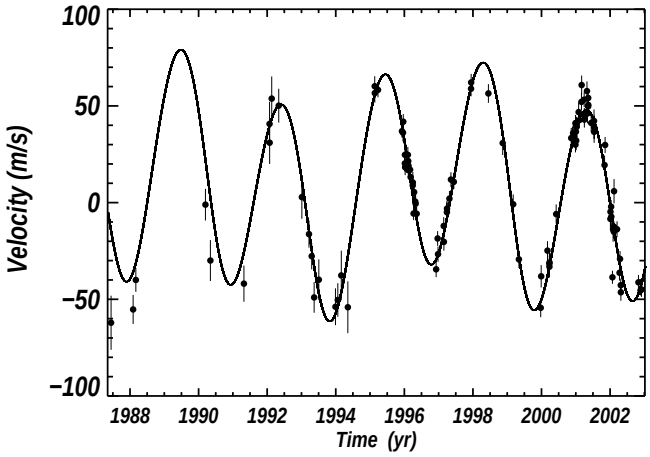

Figure 5. Velocities and two–Keplerian fit for 47 UMa. The two planets have orbital semimajor axes of 2.1 and 4.0 AU and $M \sin i$ = 2.9 and 1.1 $M_{\rm JUP}$, respectively.

Such trends in residuals, some showing curvature, indicate an additional companion and they immediately impose lower limits on the mass and orbital period of the outer companion. With patience, some cases reveal lovely full orbital motion of the outer companion, often planetary in mass and usually having an orbital period much longer than that of the inner planet. We adopt the term "hierarchical" for such widely–spaced multiple planet systems as they are structurally and dynamically reminiscent of triple–star systems (Marchal, 1990; Krymolowski and Mazeh, 1999; Ford, Kozinsky, and Rasio, 2000).

6.1. HD 37124

Velocities from 1997–2000 for HD 37124 revealed a planet with period ∼150 d and $M \sin i$ = 0.9 (Vogt et al., 2000). Subsequent velocity measurements revealed residuals that decline with time, followed by an abrupt rise and then a new decline. The velocities now reveal the full orbit of an outer planet having $M \sin i$ = 1.0 and P = 1940 d in addition to the inner one (see Fig. 6 and Tab. I). The best–fit eccentricity of the outer planet is $e = 0.7$, but nearly acceptable fits are found for e as low as $e = 0.4$, seen in the broad minimum of χ^2. However dynamical simulations by one of us (DAF) and by Greg Laughlin show that the higher eccentricities can be excluded due to the resulting unstable orbit. For $e = 0.7$, the outer planet approaches the inner one too closely resulting in subsequent ejection of one of them. From these dynamical considerations, the outer planet probably has an eccentricity less than ∼0.6 . This system typifies a growing number of multiple–planet systems in which the outer planet is considerably more distant than the inner one and for which dynamical considerations impose constraints on the plausible orbits.

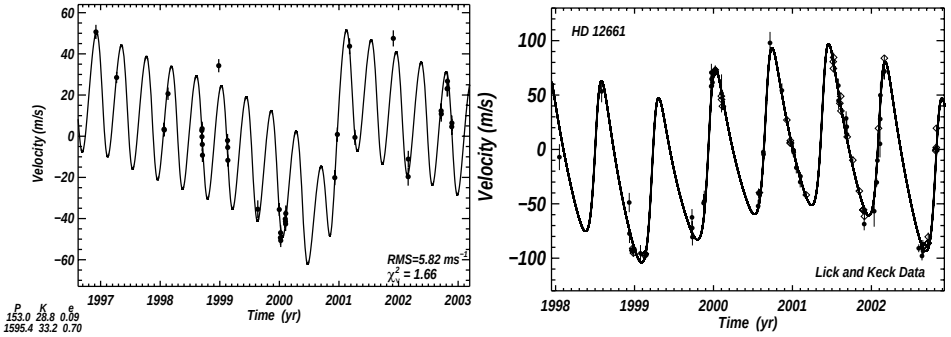

Figure 6. "Hierarchical" double–planet systems HD 37124 and HD 12661. The velocities are fit with a double–Keplerian model (solid line). Velocities are from Keck (diamonds) and Lick (dots). The widely separated periods render these system "hierarchical" (see text).

We introduce the classification "hierarchical" for widely separated double-planet systems, arbitrarily defined as those in which the ratio of orbital periods is greater than 5:1. Such systems are less likely to be actively engaged in mean–motion or secular resonances and even if so are easily perturbed out of such resonances by other planets in the system. Any weak resonances that do persist would act on time scales that are long compared to both the orbital periods and human observations.

A theory of secular interactions for hierarchical planetary systems has been constructed by Lee and Peale (2003). Their octupole approximation shows that the inner and outer planets ("1" and "2") will probably not reside in a secular resonance if the angular momentum ratio is well less than unity. That is, ignoring eccentricities, a secular resonance is unlikely if $(m1/m2)\sqrt{a1/a2} \ll 1$. For HD 37124, and using $M \sin i$ for actual masses, this ratio is only 0.37, under unity. Indeed numerical simulations suggest that the system does not reside in a secular resonance. Moreover, the two planets are definitely stable, as shown by Ji et al. (2003).

For HD 37124, the periods are 155 and 1900 d, meeting the definition of heirarchical. One wonders if the two planets ever interacted strongly in their past. A close encounter, leaving them in their current positions, would have resulting in large orbital eccentricities, especially for the less massive planet. However, such is not the case. These two planets have nearly equal values of $M \sin i$, and the inner planet has $e = 0.1$, not very eccentric, indicating that it did not suffer a close scattering. But the outer planet has a large eccentricity of $e = 0.4$–0.6, leaving a mystery of the origin of its eccentricity as close scattering seems unlikely. Gentle interactions either between the two planets or between the planets and the original disk seem plausible (e.g. Kley 2003; Goldreich and Sari, 2003).

6.2. HD 12661

Another hierarchical double–planet system is HD 12661, with two planets having periods of 260 and 1440 d. Figure 6 shows the current best two–planet fit, using velocity measurements from both Lick and Keck observatories (Fischer et al., 2003). At first glance, a resonance seems unlikely for this system, given the large ratio of orbital periods, $P_2/P_1 = 5.48$.

However, analytic and numerical simulations by Lee and Peale (2003) find that this system may definitely resides in a secular resonance which could be affected by proximity to the 11:2 mean-motion commensurability. Their theory shows that the HD 12661 system is locked in a secular resonance with the apsides being anti-aligned on average. That is, HD12661 is apparently the first extrasolar planetary system found to have $\omega_1 - \omega_2$ librating about 180°. Moreover, the apparent stability of these two planets requires that $\sin i > 0.3$ (Lee and Peale, 2003). For HD 12661, $(m1/m2)\sqrt{a1/a2} \approx 0.83$ making secular resonances possible (Lee and Peale, 2003).

With eccentricities of 0.3 and 0.2, it remains possible that a past encounter scattered the planets into their current orbits, or that an ongoing interaction actively exchanges eccentricities (Holman et al., 1997). But here, as with HD 37124, past interactions with the disk may have contributed to their settling into the secular resonance (e.g., Chiang and Murray, 2002).

6.3. HD 168443

HD 168443 is another hierarchical system (Fig. 7 and Tab. I), with orbital periods of 58 d and 5.0 yr. The two orbiting objects are so massive, with $M \sin i$ = 7.6 and 17 $M_{\rm JUP}$, that the term "planet" must be used prohibitively (see Vogt et al., 2000). The distribution of known planet masses reveals a sharp rise from 8 $M_{\rm JUP}$ toward lower masses (see Figure 11 and the paper by M. Mayor in this volume). The high–mass tail of the planet mass distribution appears to extend weakly out to 10–15 $M_{\rm JUP}$. But less than 1% of FGKM main sequence stars harbor a companion in the mass range 7–30 $M_{\rm JUP}$. The star HD 168443 has two such massive "planets". Thus one wonders whether they formed in a protoplanetary disk and if so whether an initial rocky core served as the gravitational nucleus. Instead some have proposed that they formed in an unusually massive disk by a gravitational instability in the gas (see Boss, 2002; 2003). In either case, the inner and outer "planets" have eccentricities of 0.5 and 0.2, leaving the origin of those eccentricities unknown.

6.4. HD 38529

The star HD 38529 has a mass of 1.39 $M_\odot$, making it the most massive planet–bearing star known. Its velocities (Fig. 7) reveal two planets in an extreme hierarchical configuration (Fischer et al. 2003). The inner and outer planets have minimum masses ($M \sin i$) of 0.67 and 11.3 $M_{\rm JUP}$ and orbital periods of 14.3 d

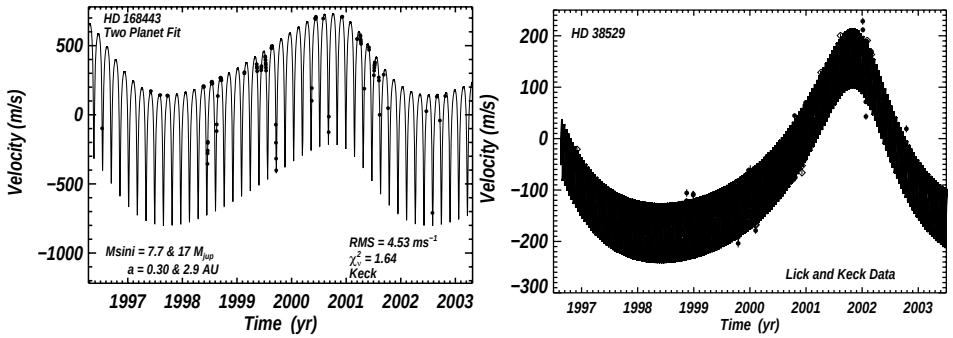

Figure 7. "Hierarchical" double–planet systems HD 168443 and HD 38529. The velocities are fit with a double–Keplerian model (solid line) in both cases. The outer orbit is now better constrained for HD 38529.

and 6.0 yr (Table 1). Because the outer planet is clearly the more massive, it is impossible that the inner planet scattered the outer planet to its current orbit. Thus this system highlights a feature of hierarchical planet systems. The planets could not have settled into the final orbits by virtue of interactions with each other. The significant orbital eccentricities (both have $e = 0.3$) of the two planets around HD 38529 must have arisen by some other mechanism than mutual interactions. The inner planet is simply insignificant. Moreover, it remains difficult to understand the origin of the eccentricity of 0.3 for the outer planet in HD 38529 as it seems unlikely that any other planet would have had comparable mass to perturb it. Obvious possibilities for the origin of the eccentricity are perturbation on the planet by either the protoplanetary disk or by an orbiting stellar companion.

6.5. HD 82943 AND HD 74156

Velocities for the 1.05 $M_\odot$ star HD 82943 reveal two planets with periods of 221 and 444 d, indicating a likely 2:1 mean motion resonance (ESO Press Release 07/01, April 4th 2001). Subsequent measurements by Mayor et al. (2002) confirm the two planets and reveal no significant mutual perturbations detected so far.

The velocities for HD 74156 reveal an inner planet with period of 52 d and an outer companion with a period of at least 6 yr. More velocity measurements are needed to determine the mass and period of the outer companion whose mass is at least 8 $M_{\rm JUP}$.

7. Eccentricities and Orbit Sizes among Single and Multiple Planets

Resonances and high orbital eccentricities are common among multiple–planet systems. However many single planets also reside in highly eccentric orbits, without any massive planets or stars currently orbiting the star. The stars 14 Her and 70 Vir both have massive planets in eccentric orbits. But the velocities show no evidence

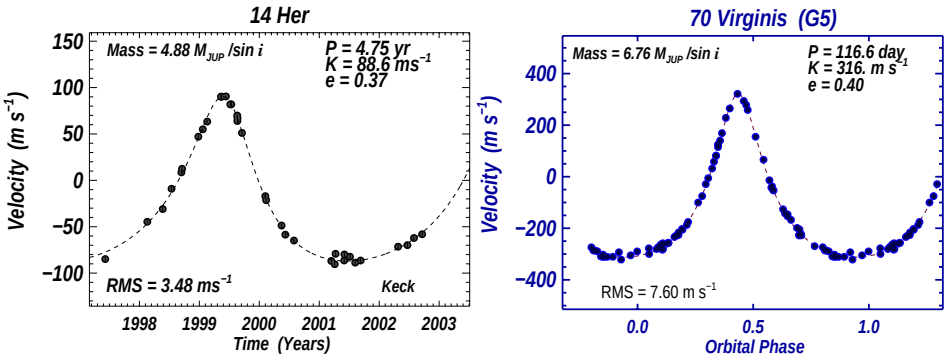

Figure 8. Representative examples of single planets in highly eccentric orbits despite the absence of additional massive planets or stellar companions. Left: 14 Her. Right: 70 Vir.

of additional companions, planetary or stellar (see Fig. 8). High eccentricities occur without benefit of any additional, current companions.

Are eccentricities pumped in single planets by a different mechanism than in multiple–planet systems? To investigate this, one may compare the eccentricity distributions among single and multiple–planet systems.

Figure 9 shows orbital eccentricity vs. semimajor axis for all 97 securely known planet systems. Here we plot only the most reliable planet detections, and thus exclude dubious and falsely–reported planets, such as HD 83443c, HD 223084b, and HD 192263b and a few others. This leaves only the 97 most reliable extrasolar planets known as of 10 Sept 2002, discovered mostly by the Geneva team and ours (see Mayor et al. this volume).

The eccentricities in Fig. 9 show several remarkable features. The eccentricities are distributed from 0 to 0.71, with only case (HD 80606 likely caused by its stellar companion) above that upper limit. An explanation has not been provided for this upper limit to the orbital eccentricities of planetary orbits. Moreover, for semimajor axes between 0.1 and 1.0 AU, there appears to be an upper envelope of eccentricities that increases. In particular, there is a paucity of eccentricities between 0.3 and 0.7 for orbits having $a = 0.1$–1.0 AU. In contrast, for $a > 1.0$ AU, the most probable eccentricity is 0.35–0.45, as seen in Fig. 9. These upper envelopes to eccentricities at $e = 0.7$ and the trend for $a = 0.1$–1.0 require a theoretical explanation.

Figure 9 also shows, with asterisks, those planets that reside in multiple planet systems. There are 22 such planets, residing in 10 multiple systems two of which are triple. *The distribution of orbital eccentricities among the planets in multi–planet systems is indistinguishable from that of the single planets.*

The distribution of semimajor axes for all securely known 97 extrasolar planets is shown in Fig. 10. The distribution shows a population of "hot jupiters" that reside at 0.04–0.06 AU, with a sharp cut–off at a period of 3.0 d. No explanation is

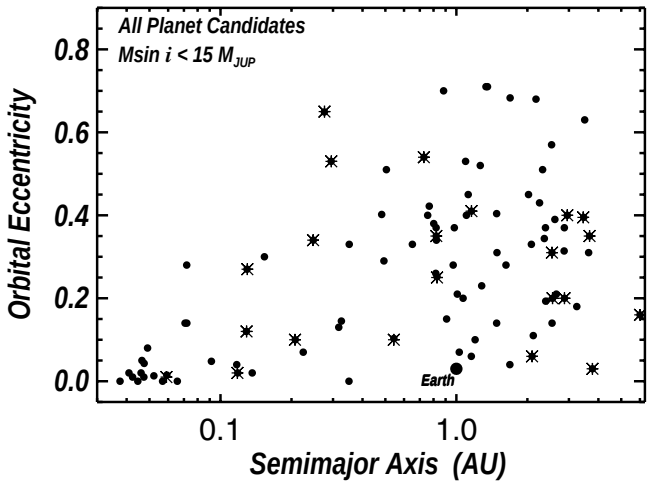

Figure 9. Orbital eccentricity vs. semimajor axis for all 97 securely known extrasolar planets. The maximum eccentricity is $e = 0.71$ for single stars with planets, and there is a trend in the upper envelope of e for orbits between 0.1 and 1.0 AU. The planets in multi–planet systems are overplotted as asterisks. Their distribution is indistinguishable from those of the single planets. (The binary star, HD 80606 is off the plot at $e = 0.93$.)

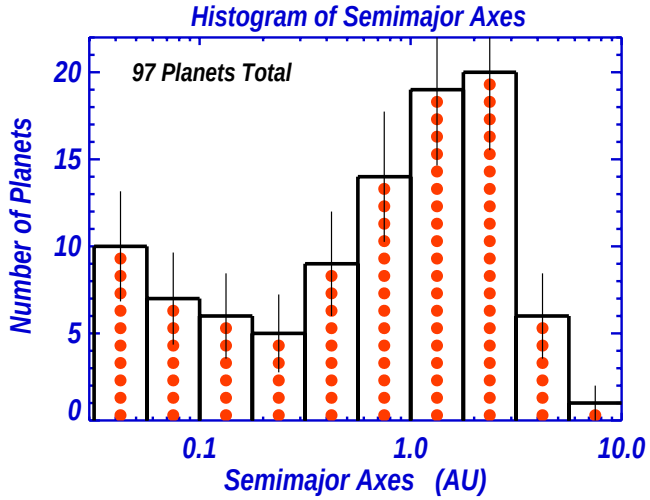

Figure 10. Distribution of semimajor axes for all 97 securely known extrasolar planets, in bins of equal $\Delta \log a = 0.25$. There is an apparent minimum near $a = 0.25$, and a rise toward larger orbits. There are apparently an increasing number of planets at larger orbits. The effect is suppressed due to incompleteness in detectability beyond 2 AU. A population of gas giants beyond 4 AU is implied by extrapolation.

known for the inner edge at $P = 3.0$ d. The distribution reveals a clear minimum at $a = 0.25$, followed by a rise toward larger semimajor axes.

It is apparent that the number of saturn and jupiter–mass planets increases with orbital size from 0.3 AU toward 3 AU. Indeed, for orbits larger than 2 AU, our detectability is relatively poor as periods longer than 5 years are difficult to detect. Our good quality data from Lick, Keck, and the AAT only spans about 5 years. It is thus remarkable that the increasing numbers of planets in large orbits is already apparent. The distribution of semimajor axes of planets within multi–planet systems is indistinguishable from that of the single planets shown.

An extrapolation of Fig. 10 toward larger semimajor axes suggests that a population of planets resides beyond 4 AU where our sensitivity will become robust in the next few years. This reservoir of jupiters at 4–10 AU is consistent with the models of migration by Trilling et al. (2002) and Armitage et al. (2002).

8. Origin of Eccentricities: Migration, Resonance Capture, Pumping

The similarity in the eccentricities of single planets and multiple–planets suggests that the origin of the eccentricities may be the same. Such a similar mechanism is remarkable, because the planets in multi–planet systems are definitely perturbed by each other as seen dynamically in the GJ 876 system (Rivera and Lissauer, 2001; 2002).

Several mechanisms have been proposed to explain the orbital eccentricities, namely planet–disk interactions (Goldreich and Sari, 2003) and planet–planet interactions (Marzari and Weidenschilling, 2002; Ford et al., 1999; Ford, 2003). An important hybrid scenario involves planet migration in a disk leading to resonance capture of the two planets (Lee and Peale, 2002; Chiang and Murray, 2002; Chiang, Fischer, and Thommes, 2002). The two planets in resonance will pump their eccentricities which can lead to such high eccentricities that the periapse distance approaches the stellar radius leading to engulfment of the planet (Thommes, 2002). Alternatively the pumped eccentricities could lead to a close planet–planet scattering and ejection of the lighter planet.

A comprehensive model for orbital eccentricities should include both processes. Two or more planets will migrate in a viscous disk at different rates, allowing them to capture each other in mean–motion resonances. Such resonances can pump the eccentricities of both planets up to values as high as 0.7 (Chiang, 2002), as observed. The growth in eccentricities, especially as the disk dissipates, can render the two orbits unstable as close passages occur. Ford (this volume) has shown that such close passages often leads to ejection of the less massive planet, leaving one behind. This sequential set of processes provides a natural explanation for the resonances and eccentricities observed among multiple–planet systems and for the eccentricities observed in single planets. Single planets may represent the survivors of double–planet progenitors.

A selection effect may play a role in the high eccentricities observed among the *single* extrasolar planets discovered to date. Most known extrasolar planets reside within 3 AU due to the limited duration ($\sim$10 y) of the Doppler surveys. Thus the planets detected to date represent a subset that ended up within 3 AU. Giant planets within 3 AU may systematically represent the survivors of scattering events in which the other planet was ejected and extracted energy from the surviving planet, throwing it inward. We would detect systematically the more massive, surviving planet residing now in an orbit with period less than 10 years.

9. The Future: Beyond 4 AU

It remains plausible that a population of gas–giant planets resides at orbital distances beyond 4 AU, as yet largely undetected (except 55 Cancri d). These giant planets would reside in wider orbits, not very different from the original, circular orbits in which they formed, because they had not suffered significant perturbations. Such a population of jupiters in wide, circular orbits is predicted by models of migration (Trilling et al., 2002; Armitage et al., 2002).

In the next few years we expect to discover, in large numbers, planets that orbit beyond 4 AU, comparable to the orbits of Jupiter and Saturn. Their eccentricities, masses, and semimajor axes, along with any planetary siblings, will reveal the diverse architectures among planetary systems, allowing us to determine how many of them in the Milky Way Galaxy are like our own.

Acknowledgements

We acknowledge the collaboration of Chris Tinney, Jack Lissauer, Eugenio Rivera, Ed Thommes, Greg Laughlin, Greg Henry, Dimitri Pourbaix, Man Hoi Lee, Stan Peale, Hugh Jones, Chris McCarthy, Jason Wright and Brad Carter. We acknowledge support by NSF grant AST-9988087, NASA grant NAG5-12182 and travel support from the Carnegie Institution of Washington (to RPB), NASA grant NAG5-8299 and NSF grant AST95-20443 (to GWM), NSF grant AST-9988358 (to SSV), and by Sun Microsystems. This research has made use of the NASA Astrophysics Data System, and the SIMBAD database, operated at CDS, Strasbourg, France. We made extensive use of the Hipparcos catalog (ESA, 1997). The authors wish to extend special thanks to those of Hawaiian ancestry on whose sacred mountain of Mauna Kea we are privileged to be guests.

References

Armitage, P. J., Livio, M., Lubow, S. H., and Pringle, J. E.: 2002, *Mon. Not. R.A.S.* **334**, 248.

Benedict, G.F., McArthur, B.E., Forveille, T., Delfosse, X., Nelan, E., Butler, R.P., Spiesman, W., Marcy, G., Goldman, B., Perrier, C., Jefferys, W.H., and Mayor, M.: 2002, *Astrophys. J.* **581**, L115.
Boss, A. P.: 2002, *Astrophys. J.* **576**, 462.
Boss, A.P.: 2003, this volume.
Brown, T.M., Kotak, R., Horner, S.D., Kennelly, E.J., Korzennik, S., Nisenson, P., and Noyes, R.W.: 1998, *ApJL* **494**, 85.
Butler, R. P., Marcy, G. W., Williams, E., McCarthy, C., Dosanjh, P., and Vogt, S. S.: 1996, *Pub. Ast. Soc. Pac.* **108**, 500.
Butler, R. P. and Marcy, G.W.: 1996, *Astrophys. J. Let.* **464**, 153.
Butler, R. P., Marcy, G. W., Williams, E., Hauser, H., and Shirts, P. 1997, *Astrophys. J. Let.* **474**, 115.
Butler, R.P., Marcy, G.W., Fischer, D.A., Brown, T., Contos, A., Korzennik, S., Nisenson, P.,Noyes, R.W.: 1999, *Astrophys. J.* **526**, 916.
Chiang, E. and Murray, N.: 2002, *Astrophys. J.* **576**, 473.
Chiang, E., Fischer, D.A. and Thommes, E.: 2002, *Astrophys. J. Let.* **564**, 105.
ESA: 1997, The Hipparcos and Tycho Catalogue, *ESA SP*-**1200**.
Fischer, D.A., Marcy, G.W., Butler, R.P., Laughlin, G. and Vogt, S.S.: 2002, *Astrophys. J.* **564**, 1028.
Fischer, D.A., Marcy, G.W., Butler, R.P., Vogt, S.S., Henry, G.W., Pourbaix, D., Walp, B., Misch, A.A., and Wright, J.T.: 2003, *Astrophys. J.* **586**, 1394.
Ford, E. B., Rasio, F. A., and Sills, A.: 1999, *Astrophys. J.* **514**, 411.
Ford, E. B., Kozinsky,B., and Rasio, F. A.: 2000, *ApJ* **535**, 385.
Ford, E. B.: 2003, this volume.
Goldreich, P. and Sari, R.: 2003, *Astrophys. J.* **585**, 1024.
Henry, G. W., Baliunas, S. L., Donahue, R. A., Fekel, F. C., and Soon, W.: 2000, *Astrophys. J.* **531**, 415.
Holman, M., Touma, J. and Tremaine, S.: 1997, *Nature* **386**, 254.
Ji, J., Kinoshita, H., Liu, L., and Li, G.: 2003, *Astrophys. J.* **585**, L139.
Kley, W.: 2000, *Mon. Not. Roy. Ast. Soc.*, 313, L47.
Kley, W.: 2003, this volume.
Krymolowski, Y. and Mazeh, T.: 1999, *MNRAS* **304**, 720.
Laughlin, G. and Chambers, J. E.: 2001, *Astrophys. J. Let.* **551**, 109.
Laughlin, G., Chambers, J. E. and Fischer, D.A.: 2002, *Astrophys. J. Let.* **579**, 455.
Lee, M. H. and Peale, S. J.: 2002, *Astrophys. J.* **567**, 596.
Lee, M.H. and Peale, S.J.: 2003, *Astrophys. J.* **592**, 1201.
Lissauer, J. J., and Rivera, E.: 2001, *Astrophys J.* **554**, 1141.
Marchal, C.: 1990, *The Three–Body Problem* (Amsterdam: Elsevier).
Marcy, G. W. and Butler, R. P.: 1996 *Astrophys. J. Let.* **464**, 147.
Marcy, G. W., Butler, R. P., Fischer, D., Vogt, S. S., Lissauer, J. J., and Rivera, E. J.: 2001, *Astrophys. J.* **556**, 296.
Marcy, G.W., Butler, R.P., Fischer, D.A., Laughlin, G., Vogt, S.S., Henry, G.W., and Pourbaix, D.: 2002, *Astrophys. J.* **581**, 1375.
Marzari, F., and Weidenschilling, S. J.: 2002, *Icarus* **156**, 570.
Mayor, M. and Queloz, D.: 1995, *Nature* **378**, 355.
Mayor, M., et al.: 2002, personal communication.
Nelson, R.P. and Papaloizou, J.C.B.: 2002, *MNRAS* **333L**, 26.
Rivera, E. J., and Lissauer, J. J.: 2001, *Astrophys. J.* **558**, 392.
Rivera, E. J. and Lissauer, J. J.: 2002, *Astron. Astrophys.*S/Division of Dynamical Astronomy Meeting, 33.
Trilling, D., Lunine, J., and Benz, W.: 2002, *Astron. Astrophys.* **394**, 241
Thommes, E.: 2002, private communication.
Vogt, S. S., Marcy, G. W., Butler, R. P., and Apps, K.: 2000, *Astrophys. J.* **536**, 902.

Debris Disks

Observations and models

A.-M. Lagrange
Laboratoire d'Astrophysique de Grenoble, France

J.-C. Augereau
Sterrewacht Leiden, The Netherlands

Abstract. The understanding of planetary system formation and evolution processes requires to study extrasolar planets as those discovered through radial velocity technics, but also planetary disk material in the form of dust, planetesimals and gas. Indeed, disks are the sites of future or past or on-going planet formation and they tell us about the conditions in which the planetary formation has or will occur. Disks have been identified now in a variety of evolutionary stages, from protoplanetary disks around young (My) stars such as T Tauri to debris disks seen around mature ($\geq$ typ. 10 My and up to Gy) stars. The later disks host second generation dust, i.e. dust with lifetimes shorter than the star age. They are especially important in the context of planetary formation as 1) they provide informations on the probable sources of the dust, namely planetesimals and comets, 2) they may indirectly reveal the presence of unseen planets and 3) they bring clues to the dynamical history of the planetary system. This paper aims at reviewing the current knowledge on these debris disks and summarizing what they tell us concerning planetary system formation, and at addressing pending questions that one may hope to answer to in the forthcoming years.

1. Disks as Tracers of Planetary Formation

Theory predicts that disks and stars form from the collapse of a nebulae made of gas and small, submicronic grains. This nebulae rotates rapidly and grains rapidly settle in the equatorial plane of the system. Within about 10^5 years, kilometer size bodies, so-called planetesimals are supposed to form. These planetesimals are the building blocks of telluric planets and cores of giant planets. Giant planets cores accrete the gas at a relatively early stage (1 My). It is believed that after a few My, most of the gas and hence most of the planetary system mass is either accreted onto planetesimals or giant planets or expelled from the system. Telluric planets are expected to build on a longer timescale, up to several 10^7 years. When mature planetary systems are formed, dust can still be present, in relatively smaller amounts, thanks to collisions among solid bodies or evaporation. From a theoretical point of view, disks are then expected to be present in all stages of planetary systems formation and evolution, with properties however considerably changing depending on their evolution status.

From an observational point of view, several types of disks have been observed in the past years or decades. On the one hand, protoplanetary disks have been observed around young ($\leq$ a few My) T Tauri stars. They are thought to be made of small, primordial grains, as well as large amounts of gas distributed throughout the disk. These disks are optically thick and gas dominated. On the other hand, evolved

disks have been inferred around older (several tens My and up to Gy) stars such as Vega, Fomalhaut or Epsilon Eridani from thermal IR excess. In those disks, the dust has a lifetime against destruction processes such as radiation pressure or Poynting-Roberston drag shorter than the system age. It is thus not primordial, but is rather attributed to collisions or evaporation among larger solid bodies. This defines the second generation, or debris disks. These systems are tenuous ($L_{\rm dust}/L_{\rm star} \leq 10^{-3}$) and their gaseous content, whenever gas is present, is significantly smaller than the dust content.

The transition between primordial disks and debris disks is not sharp and intermediate systems are to be expected, with characteristics very variable from one object to the other. Obviously, the knowledge of all these types of disks, spanning ages between a few My to Gy is crucial to understand the whole process of planetary systems evolution.

To study second generation or transition systems, one must investigate their solid and gaseous (if present) contents. For the solid material, one wants to know its total mass, the size spectrum from sub-micronic bodies up to planets; one also wants to know the nature of the dust (chemical composition) and its spatial distribution. For the gas phase, if present, one wishes to know the total amount of gas, its chemical composition and its the dynamics. An important parameter relevant for the dynamics and evolution of the system is the gas/dust ratio which will tell about the possible coupling between gas and dust and hence put constrains on the dynamics of the material within the disk. Finally and importantly, the system age is of course an important parameter for evolutionary purposes. Unfortunately, unless the stars are known members of stellar associations, the ages are rarely well determined.

2. Disks Observations

By nature, debris disks host little amounts of dust. They are best detected by their albeit small thermal IR excess. However, to ascertain that the systems are second generation, one has to characterize the dust size, which requires measurements at several wavelengths. Also, to determine the dust geometry, resolved images are needed. At optical or near-IR wavelengths, direct imaging of the faint dust scattered light close to the bright star requires high dynamical range, diffraction limited devices. Such devices are now provided by coronagraphs in space or coupled to Adaptive Optics systems on the ground. The resolved optical or near-IR data thus reveal the disks with spatial resolution down to 1–10 AU for typical systems (see e.g. Figure 1). At longer wavelengths, the dust is seen through its thermal emission and the dust to star brightness contrast rapidly decreases. This allows direct detection, at however lower spatial resolution. The low spatial resolution has several times led to attribute excesses to CS dust whereas they were due to background, unresolved sources. Thermal images trace the emission of the inner dust, while

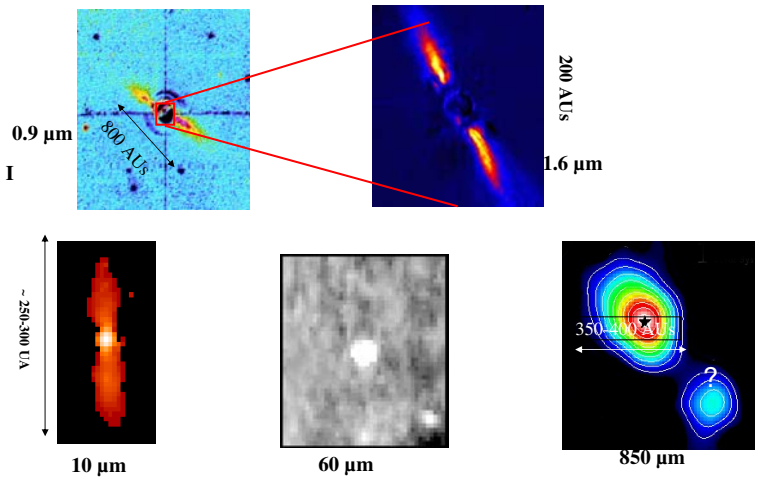

Figure 1. β Pictoris disk observed at several wavelengths. Top left: Smith and Terrile (1984); top right: Mouillet et al. (1997); bottom left: Pantin et al. (1997), bottom right: Holland et al. (1998).

submillimetric observations probe the outer dust. Long wavelength data allow an estimation of the grain size and of the minimum mass of the solid material within the system.

Almost two decades ago, IRAS revealed that a large fraction of nearby Main Sequence (MS) stars (between 15 and 40 %, depending on the stellar type) exhibit 20 to 100 micron excess due to circumstellar (CS) dust. Later-on, ISO also led to the detection of numerous IR excess stars, with ages $\geq$ 1 My, see e.g. Spangler et al. (2001). Available observations show that the amount of dust rapidly decreases with system age, with an t^{-2} law. Among IR excess stars, very few systems have been resolved at near IR and/or optical wavelengths. In all cases the imaged systems are aged between a few My and may-be 200 My (old Herbig stars or young MS stars); no CS disk has been imaged at scattered light around Gy old stars sofar, despite several attempts and a number of unconfirmed detections. Submillimeter observations have allowed to resolve the outer dust for several stars, with a large variety of ages, from young (a few My) systems to old ones such as Vega, α PsA and ε Eri (Holland et al., 1998; Greaves et al., 1998; Wilner et al., 2002).

Concerning the gas, several searches have been made around MS or young stars see (Lagrange et al., 2000) for review of results and references prior to 2000. Generally, except in 2 cases (β Pictoris and HR6519), none of the MS stars with strong IR excess (Vega, α PsA, ε Eri) show evidence for metallic, circumstellar gas down to levels at least ten times lower than β Pictoris. Nor do they show evidence for CO, with still model dependant upper limits for Vega and α PsA. CO searches gave positive detections on young (a few Myrs) stars in the recent years see e.g. (Thi et al., 2002). In some cases of a few myrs Herbig stars (e.g. HD163296, AB Aur), Keplerian CO disks were found (Mannings and Sargent, 2000) up to 300-400

AU. It has to be noticed however that CO life-time around early-type stars is very short against photodissociation or ISM UV flux, unless shielded; at large distances it may be also frozen onto grains. As a consequence, and conversely to what is done in the ISM or in the case of young systems the total amount of gas cannot be derived from CO measurements for debris disks. Other tracers have to be used. C and C^+ (Kamp et al., 2003) for A stars are proposed, but deriving the total gaseous mass is still very model dependant. H2 measurements are absolutely needed and lack sofar, with very few exceptions (Thi et al., 2002; Richter et al., 2002). Also and as underlined in the later, most of the present low spatial resolution data may suffer from possible contamination with background sources.

3. Resolved Debris Disks

3.1. β PICTORIS

Beta Pictoris (A5V, 20–100 My) is the best studied case among these debris disks. Figure 2 shows the panoply of available data on this object. For this object located at 20 pc, scattered light high resolution (typically 0.1 ie 2 AU) data are available between 20 and more than 1000 AU, as well as seeing limited infrared thermal images at 5 AU resolution, probing the 0 to 100 AU region. Together with the longer wavelength unresolved data, they help characterizing the dust and showing that at least 5 micron, short lived dust is present in the disk, which is hence of second generation. Also imaging revealed several asymmetries: butterfly asymmetry seen in the outer disk at low spatial resolution (Kalas et al., 1995) as well as inner warp, detected at high resolution (Mouillet et al., 1997). Presently, detailed dynamical and physical modeling allows to explain most of the disk characteristics: shape, extension, and asymmetries, as well as the Spectral Energy Distribution, starting from a distribution of colliding planetesimals located within 120 AUs from the star plus a planet on an inclined orbit (Augereau et al., 2001). Finally, spectroscopy at thermal IR indicates the presence of silicates in the disks, producing emission features resembling those of comets in our Solar System.

Gas is also present in the CS disk around β Pictoris, some of which stable, mainly ionized, but also neutral and molecular (CO). For a review on the results on gas content prior to 2000 see e.g. (Lagrange et al., 2000). One peculiar feature is the observed infall of ionized gas towards the star with high velocities (100 of km/s). This infall has been attributed to -and modeled as-the evaporation of comets close to the star. A proposed triggering mechanism for this comet infall is the gravitational perturbation of a planet via mean motion resonances (Beust et al., 2000). This is another indirect indication for the presence of a planet around β Pic. Interestingly, the required planet mass and distance to star are compatible with the ones deduced from the modeling of the warp seen in scattered light images (Mouillet et al., 1997). Finally, FUSE spectra revealed the presence of very ionized

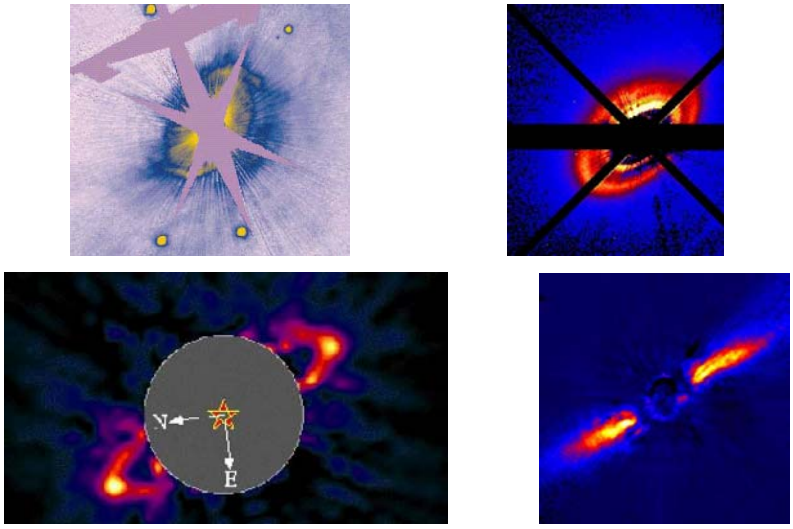

Figure 2. Panoply of resolved disks. Top left: HD 100546 (Grady et al., 2001); top right: HD 141569 (Augereau et al., 1999); bottom left: HR4796 (Schneider et al., 1999); bottom right: β Pic (Mouillet et al., 1997).

lines (CIII, OVI) possibly tracing chromospheric activity, sofar unexpected for such stellar types (Bouret et al., 2002). Whether this peculiarity is connected or not to the planetary system is still an open question, and will probably remain as such until other examples are found.

3.2. HR 4796

Another example of resolved system is that of the young 8 My star HR4796 (A0V, 66 pc), member of the nearby young stellar association TW Hya. The HR4796 disk shapes as a very narrow ring located at about 70 AU from the star (Fig. 1). Modeling shows that hot dust, responsible for the thermal IR emission, is also present closer to the star. Detailed modeling also shows that the dust is at least a few microns, and hence of second generation (Augereau et al., 1999). This is an evidence that planetesimals are present in systems as young as a few My. It has been suggested that a planet orbiting the star could be also present and responsible for the ring shape of the disk. However no detailed modeling is available yet, and alternative explanations such as gas to dust interaction do exist (Takeuchi and Artymowicz, 2001). No gas has been detected sofar in the disk down an upper limit of 1–7$M_\oplus$ for H_2 (Greaves et al., 2000). However, such limits are actually not strong enough to rule out a possible efficient gas to grain interaction.

The HR4796 disk can be compared to the system around HD98800, another member of the TW Hya group. HD98800 is a member of a multiple (quadrupole) system located at 47 pc. HD98800 (K5, 5–20 My) exhibits a very high disk luminosity compared to HR4796 (factor of 1000). Images at near IR revealed a

pole on flared disk between 20 and more than 200 AU (Weinberger et al., 2002). Sub-millimeter and millimeter data indicate the presence of large grains (mm size) (Sylvester et al., 1996), indicating that grain growth is already active. Both the dust and gas appear to be more abundant than in the case of HR4796. It is not clear at this point whether the difference between is due to the star properties (mass, temperature, luminosity).

3.3. HD 141569

Another interesting example for which a disk has been clearly resolved is the 5 My old star HD 141569 (B9.5V, 100 pc). Similarly to HR 4796, the disk is strongly structured with rings or ringlets extending to more than 300 AU. The observed non axi-symmetry of the disk cannot be explained by anisotropic scattering alone and is attributed to the gravitational perturbation of a massive body (Mouillet et al., 2001). Similarly to HR4796, HD141569 is a member of a binary system, which could partly explain the disk shape; alternatively the gravitational perturbation by a planet has also been proposed. Detailed dynamical simulations are needed to further test this point (Augereau et al., in prep). An important question is the presence of gas in this system. Recently, CO and H3+ emissions have clearly been detected (Brittain et al., 2002) towards HD 141569. The location of the gas and hence its origin are still unclear as the data lack of spatial resolution. Nevertheless, Brittain et al. (2002) proposed that part of the gas could be surrounding a giant forming planet. Recent Plateau de Bure observations of CO allowed to spatially resolve a disk of CO in keplerian orbit; these observations are expected to put strong constrains on the gas properties (Augereau et al., 2003).

3.4. HD 100546

A last example is the HD 100546 disk. Eventhough the star is of similar masse and age as the previous one, HD 141569, there are good indications (spectroscopic, images) that the system is less evolved (Augereau et al., 2001). For this object, spectral variability has also been observed and attributed to cometary infall and evaporation (Grady et al., 1997), but detailed modeling clearly questions the cometary hypothesis (Beust et al., 2001): indeed the forces acting on the evaporated gas are much stronger than in the β Pictoris case and could prevent from forming a detectable cloud of gas.

4. Summary and Future Prospects

The observations obtained sofar show that extrasolar planetary disks exist around stars aged a few My and up to Gy. Globally the dust and gas amounts decrease with time over the first millions years. Debris disks, ie made at least partly of second generation dust are found around stars aged at least a few My. They indicate that

large solid bodies are present at such ages. In the case of β Pictoris (20–100 My), there is now a bulk of evidences that a giant planet could be already present in the disk.

Among the still limited number of studied systems, discrepancies in disks properties are seen, indicating that age is not the only parameter governing disk evolution. Probably, binarity plays an important role in disk evolution but this is not yet proved. Also, the gas content in this disks is generally small or not well known, preventing thus from deriving possible constrains on the dynamical histories of the systems.

To progress in the knowledge of these systems, one needs to observe more disks, in order to investigate through statistics, the effect of parameters such as stellar type, system age, binarity. In that frame, forthcoming tools such as SIRTF, ALMA and JWST will certainly provide unprecedent informations.

One also needs to perform detailed and multiwavelength observations of each system, so as to better study its morphology and find possible signs of planet. This will be done thanks to high angular resolution and high contrast devices present on large ground telescopes equipped with adaptive optics, with interferometers or with JWST.

On top of these observations that will help constraining the dynamics of those systems, spectroscopy in particular at thermal IR and sub-mm will help understanding the chemical status and evolution of these systems, and a great deal has still to be done in that field.

In the forthcoming years, debris disks around young stars but also older MS and late type stars, in particular stars with radial velocity planets may be of reach, providing thus a link with other studies of planetary systems. The panoply of new instruments working from the optical to the millimeter will allow unprecedent modeling of those systems.

References

Augereau, J.-C. et al.: 1999, 'On the HR4796A circumstellar disk', *Astron. Astrophys.* **348**, 557–569.

Augereau, J.-C. et al.: 2001, 'Dynamical modeling of large scale asymmetries in the beta Pictoris dust disk', *Astron. Astrophys.* **370**, 447–455.

Augereau J.-C. et al.: 2001, 'HST/NICMOS2 coronagraphic observations of the circumstellar environment of three old PMS stars: HD 100546, SAO 206462 and MWC 480' *Astron. Astrophys.* **365**, 78–79.

Augereau, J.-C. et al.: 2003, 'CO around HD 141569', in prep.

Besut, H. and Morbidelli, A.: 2000, 'Falling Evaporating Bodies as a Clue to Outline the Structure of the beta; Pictoris Young Planetary System', *Icarus* **143**, 170–188.

Beust H. et al.: 2001, 'Falling Evaporating Bodies around Herbig stars. A theoretical study', *Astron. Astrophys.* **366**, 945–964.

Bouret, J. C. et al.: 2002, 'A chromospheric scenario for the activity of beta Pictoris, as revealed by FUSE', *Astron. Astrophys.* **390**, 1049–1061.

Brittain S. D., et al.: 2002, 'CO and H+3 in the protoplanetary disk around the star HD 141569', *Nature* **418**, 57–59

Grady, C. A., et al.: 1997, 'The star-grazing extrasolar comets in the HD 100546 system', *Astrophys J.* **483**, 449–456.
Grady, C. A., et al.: 2001, 'The Disk and Environment of the Herbig Be Star HD 100546', *Astronom J.* **122**, 3396–3406.
Greaves J. S. et al.: 1998, 'A dust ring around ε Eridani: Analog to the young solar system', *Astrophys J.* **506**, L133–L137.
Greaves, J. S. et al.: 2000, 'The dust and gas content of a disk around the young star HR 4796A', *Icarus* **143**, 155–158.
Holland, W. S., et al.: 1998, 'Submillimetre images of dusty debris around nearby stars', *Nature* **392**, 788–790.
Jayawardhana, R. et al.: 2001, 'Mid-infrared imaging of candidate Vega-like systems', *Astronom J.* **122**, 2047–2054.
Kamp, I. et al.: 2003, 'Line emission from circumstellar disks around A stars', *Astron. Astrophys.* **397**, 1129–1141.
Kalas, P. and Jewitt, D.: 1995, 'Asymmetries in the beta Pictoris dust disk', *Astronom. J.* **110**, 794–804.
Koerner, D. W., et al.: 1998, 'Mid-infrared imaging of a circumstellar disk around HR 4796: Mapping the debris of planetary formation', *Astrophys. J.* **503**, L83–L88.
Lagrange, A.-M., Backman, D. and Artymowicz, P.: 2000, 'Planetary Material around Main-Sequence Stars', In *Protostars and Planets IV*; ed. V. Mannings, A. P. Boss and S. S. Russell (Tucson: The University of Arizona Press), 639–672.
Mannings, V. and Sargent, A.: 2000, 'High-Resolution Studies of Gas and Dust around Young Intermediate-Mass Stars. II. Observations of an Additional Sample of Herbig AE Systems' *Astrophys. J.* **529**, 391–401.
Mouillet, D. et al.: 1997, 'A planet on an inclined orbit as an explanation of the warp in the Beta Pictoris disc', *Month. Not. R. A. S.* **292**, 896–903 Mouillet, D. et al.: 2001, 'Asymmetries in the HD 141569 circumstellar disk', *Astron. Astrophys.* **372**, L61–L64.
Pantin, E. et al.: 1997: 'Mid-infrared images and models of the beta Pictoris dust disk', *Astron. Astrophys.* **327**, 1123–1136.
Richter, M.J. et al.: 2002, 'Looking for Pure Rotational H2 Emission from Protoplanetary Disks' *Astrophys. J.* **572**, L561-L564
Spangler, C. et al.: 2001, 'Dusty debris around solar-type stars: temporal disk evolution', *Astrophys. J.* **555**, 932–944.
Schneider, G. et al.: 1999, 'NICMOS Imaging of the HR 4796A Circumstellar Disk', *Astrophys. J.* **513**, 127–130.
Smith, B. and Terrile, R.: 1984, 'A circumstellar disk around Beta Pictoris', *Science* **226**, 1421–1424.
Sylvester, R.J. and Skinner, C.J.: 1996, 'Optical, infrared and millimetre-wave properties of Vega-like systems -II. Radiative transfer modelling', *Month. Not. R. A. S.* **283**, 457–470.
Takeuchi, T. and Artymowicz, P: 2001, 'Dust migration and morphology in optically thin circumstellar gas disks', *Astrophys. J.* **557**, 990–1006.
Thi, W.-F. et al.: 2002, 'H2 and CO Emission from Disks around T Tauri and Herbig Ae Pre-Main-Sequence Stars and from Debris Disks around Young Stars: Warm and Cold Circumstellar Gas', *Astrophys. J.* **561**, 1074–1094.
Weinberger, A., et al.: 2002, ' Infrared views of the TW Hydra disk', *Astrophys. J.* **556**, 409–418.
Wilner, D. J., et al.: 2002, 'Structure in the dusty debris around Vega', *Astrophys. J.* **569**, L115–L119.

Part III

Multi-Planetary Systems: Theory

Formation of Cores of Giant Planets and an Implication for "Planet Desert"

S. Ida
Tokyo Institute of Technology, Ookayama, Meguro-ku, Tokyo 152-8551, Japan

Abstract. We discuss accretion of cores of giant planets from planetesimals, based on the results of N-body simulations. We derive the core isolation mass, which is a final core mass as a result of runaway/oligarchic growth, as functions of disk mass and semimajor axis. The core accretion rate is well predicted by a theoretical model based on two-body approximation. If cores reach a critical mass of $\sim 10M_\oplus$ during the persistence of disk gas, gas accretion onto the cores occurs and gas giants are formed. Modeling the gas accretion process, we discuss the dependence of formation sites of gas giants on disk mass, assuming a constant dust-gas ratio. In the minimum mass disk for the Solar system, gas giants can be formed only just outside "snow line" at a few AU. On the other hand, giants can be formed well inside the snow line in disks 10 times as massive as the minimum mass disk. In such inner regions, the core accretion is so fast, that formation of gas giants is completed before the gas is dissipated. Since planets grow very rapidly from $10M_\oplus$ to $100M_\oplus$, our model predicts the deficit of planets with intermediate masses ($10M_\oplus$ to $100M_\oplus$) at $a \lesssim$ a few AU, which we call "planet desert". The idea of planet desert may be consistent with the current observational data and will be examined in detail by future observations.

1. Introduction

The increased number of the discovered extrasolar planets may be large enough to allow us to deduce clues on formation of the extrasolar planets from their distributions of orbits and masses. Figure 1 shows the distributions of semimajor axes (a) and masses ($M_{\rm p}$) of the discovered extrasolar planets (`http://www.obspm.fr/encycl/encycl.html`). Radial velocity survey may have found planets with radial velocity amplitude, $v_r \gtrsim 10$ m/s, and relatively small semimajor axes, $a \lesssim$ 4–5 AU. The mass distribution in these regions may follow a weak power law function in $M_{\rm p}$ with an abrupt cut-off at about 10 Jupiter masses ($M_{\rm J}$) (e.g., Jorissen et al., 2001; Tabachnik and Tremaine, 2002) except for some paucities.

The deficit of massive ($M_{\rm p} \gtrsim 3M_{\rm J} \sim 1000M_\oplus$ where $M_\oplus$ is an Earth mass) and close-in ($a \lesssim 0.2$ AU) planets were discussed by Zucker and Mazeh (2002) and Pätzold and Rauer (2002). Here we focus on another paucity in the $M_{\rm p} - a$ distribution: the regions $a \gtrsim 0.2$ AU and $M_{\rm p} \lesssim M_{\rm J} \sim 300M_\oplus$ (Fig. 1; also see Chapter by M. Mayor). Because v_r is well over 10 m/s in this region, the paucity would not be due to observational selection effects but should reflect planetary formation processes. We will show that the deficit of such planets would be naturally accounted for by the following scenario. Gas giant planets grow by accreting solid cores from planetesimals (Wuchterl et al., 2000, and references therein). Cores must reach a critical mass of $\sim 10M_\oplus$ to accrete gas. The core masses and the core

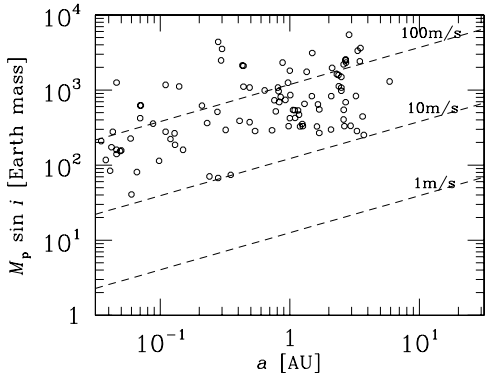

Figure 1. Distributions of semimajor axes (a) and masses ($M_{\rm p}$) of discovered extrasolar planets are shown by open circles. Unit of mass is Earth mass $M_\oplus$ (Jupiter mass is $M_{\rm J} = 318 M_\oplus$). i is the angle between an orbital plane and the line of sight. The ascending lines correspond to radial velocity amplitude of 100 m/s, 10 m/s, and 1 m/s (assuming the host star mass is $1 M_\odot$).

accretion timescales are determined by surface density of the planetesimal swarm and orbital radius (e.g., Kokubo and Ida, 2002). In general, formation of gas giants is favored just outside "snow line" at a few AU, where ice condensation facilitates formation of large cores. At lager distances ($\gtrsim$ 10 AU), the core accretion is so slow that the gas tends to be dissipated before it can be accreted. In very massive protoplanetary disks, however, cores can reach the critical mass even inside the snow line. Since the core accretion is relatively fast in the inner regions, gas accretion actually occurs. The planets grow so rapidly from $10 M_\oplus$ to more than $100 M_\oplus$, that planets of intermediate masses (10–$100 M_\oplus$) should be rare ("planet desert"). Outside the snow line, the intermediate mass planets can accrete after the gas dissipation, the mass distribution should be continuous. As discussed later, type II migration and the Roche lobe overflow can bring the intermediate mass planets to $a \lesssim 0.1$ AU (although we have not included these effects in our calculations). Therefore, the deficit of the intermediate mass planets would remain as "planet desert" at 0.2 AU to a few AU.

2. Accretion of Cores from Planetesimals

In an early stage of planetesimal accretion, the largest planetesimals grow more rapidly, which is called "runaway growth" (e.g., Wetherill and Stewart, 1989; Kokubo and Ida, 1996). Because the runaway growth is self-regulated (Ida and Makino, 1993), similar-sized largest protoplanets are formed with almost equal orbital spacing in a planetesimal swarm ("oligarchic growth"; Kokubo and Ida, 1998, 2000).

Figures 2 show the result of an N-body simulation of planetesimal accretion, demonstrating oligarchic growth. Initially, 10,000 planetesimals of 2×10^{23} g are

distributed in nearly coplanar circular orbits between 0.5 AU and 1.5 AU. The surface density of planetesimals ($\Sigma_{\rm solid}$) is given by

$$\Sigma_{\rm solid} = f_{\rm disk} \times 10 \left(\frac{a}{1\ {\rm AU}}\right)^{-3/2}\ {\rm g\,cm^{-2}}, \tag{1}$$

where $f_{\rm disk} = 1$ in this simulation, so that $\Sigma_{\rm solid}$ here is similar to the minimum mass disk model for the Solar system (e.g., Hayashi et al., 1985). In this calculation, physical radii are artificially enlarged by a factor 6. So, labels of time should be multiplied by a factor 6–36 (depending on the importance of gravitational focusing) to be translated into realistic timescales. For other details of calculations, see the figure caption.

For simplicity, we here assume the same power law a-dependence of $\Sigma_{\rm solid}$, dust-gas ratio and disk size for all disks. Then, $f_{\rm disk}$ indicates a mass of a protoplanetary disk relative to the minimum mass disk model for the Solar system. Observationally inferred disk masses ($M_{\rm disk}$) around T Tauri stars show no clear dependence on the stellar age up to 10^7 years, but have a distribution ranging from $10^{-3} M_\odot$ to $M_\odot$ with a peak at $\sim 0.03 M_\odot$ (Beckwith and Sargent, 1996). Since in the minimum mass disk model with $f_{\rm disk} \sim 1$, $M_{\rm disk} \sim 0.01 - 0.02 M_\odot$, the range of $f_{\rm disk}$ we should consider is from 0.1 to 100.

Figures 2 show that orbital separation distance Δa between protoplanets is $\sim 10 - 15 r_{\rm H}$, where $r_{\rm H}$ is the Hill radius defined with a protoplanet (solid core) mass $M_{\rm sc}$ by

$$r_{\rm H} \equiv \left(\frac{M_{\rm sc}}{3M_\odot}\right)^{1/3} a. \tag{2}$$

Since growth time scale is almost proportional to a^3 (see e.g., Kokubo and Ida, 2002), all the planetesimals quickly accrete to the protoplanets in inner regions. In outer regions, protoplanets are still accreting at 400,000 years.

Given $\Sigma_{\rm solid}$, the relation $2\pi a \Delta a \Sigma_{\rm solid} = M_{\rm sc}$ gives a final mass $M_{\rm iso}$ of a protoplanet as a result of runaway/oligarchic growth. The mass is called "isolation mass" and given by (Kokubo and Ida, 1998, 2000, 2002)

$$M_{\rm iso} \simeq 0.15 f_{\rm disk}^{3/2} f_{\rm ice}^{3/2} \left(\frac{a}{1\ {\rm AU}}\right)^{3/4} \left(\frac{\Delta a}{10 r_{\rm H}}\right)^{3/2} M_\oplus, \tag{3}$$

where $f_{\rm ice}$ expresses the increase of solid materials due to ice condensation beyond the snow line (at $a_{\rm snow}$); $f_{\rm ice} = 1$ for $a \lesssim a_{\rm snow}$ and $f_{\rm ice} = 3$–4 for $a \gtrsim a_{\rm snow}$. In optically thin disks with Solar luminosity, $a_{\rm snow} \simeq 2.7$ AU (Hayashi et al., 1985).

The left (straight line) parts of the two curves in each panel in Fig. 2B are predicted $M_{\rm iso}$ with $\Delta a = 15 r_{\rm H}$ (upper) and $\Delta a = 10 r_{\rm H}$ (lower). The right (curved) parts are a prediction by the analytic growth model based on the two-body approximation (e.g., Safronov, 1969; also see Kokubo and Ida, 2002 for details). The results of the N-body simulation are consistent with the above model with

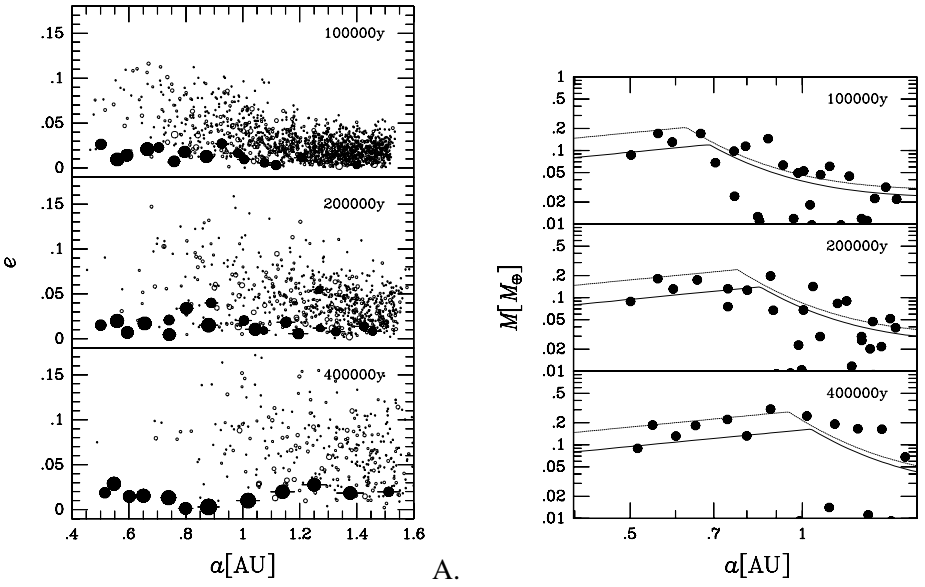

Figure 2. A) Snapshots of an N-body simulation of planetesimal accretion (three-dimensional), starting from 10,000 planetesimals with 2×10^{23}g ($f_{\rm disk} = 1$). e is orbital eccentricity and a is semimajor axis. Radii of circles are proportional to physical sizes. Filled circles are protoplanets (runaway bodies). The length of bars of the protoplanets are $5r_{\rm H}$ in both sides. Orbits are integrated by 4th-order Hermite with individual timesteps and direct summation of gravitational interactions of all bodies. Gas drag is neglected. B) Mass distributions of the same result. The left parts of the two curves are isolation masses predicted by Eq.(3) with $\Delta a = 15r_{\rm H}$ (upper) and $\Delta a = 10r_{\rm H}$ (lower). In the large a regions, masses of protoplanets have not reached $M_{\rm iso}$. The right parts of two curves are a prediction from the analytic growth model.

$\Delta a = 10-15r_{\rm H}$. Even if we start with different initial conditions (but still matching Eq. 1), we obtain similar results. Also in Jovian planet regions, similar results are obtained (Thommes et al., 2002). Kokubo and Ida (2002) did N-body simulations of planetesimal accretion with $f_{\rm disk} = 0.1, 1, 10$ and found that in all cases, the results are in good agreement with Eq. (3).

Eccentricity damping due to disk-planet interaction (Artymowicz, 1993; Ward, 1993) would inhibit protoplanets from starting orbit crossing and the protoplanet masses are kept $\sim M_{\rm iso}$ until $M_{\rm disk}$ decays to $\lesssim 10^{-5}M_\odot$ and the damping becomes weak enough (Iwasaki et al., 2002; Kominami and Ida, 2002). If $M_{\rm iso}$ is large enough ($\gtrsim 10M_\oplus$), rapid gas accretion occurs to form gas giant planets (see Eqs. 4 and 5). Otherwise, protoplanets stay in circular orbits without gas accretion. After the disk dissipation to $\lesssim 10^{-5}M_\odot$, protoplanets start orbit crossing and coalescence. Earth and Venus should have been formed in this way because their masses are considerably larger than $M_{\rm iso} \sim 0.15M_\oplus$ at 1 AU for $f_{\rm disk} \sim 1$. Note that even if the coalescence forms planets large enough for gas accretion, gas accretion cannot occur any longer, because the gas is already dissipated.

3. Gas Accretion

If a core mass ($M_{\rm sc}$) becomes larger than the critical core mass ($M_{\rm crit}$), the gas accretion onto the core starts. The critical core mass is give by (Ikoma et al., 2000)

$$M_{\rm crit} \simeq 10 \left(\frac{\dot{M}_{\rm sc}}{10^{-6} M_\oplus/{\rm year}} \right)^{1/4} M_\oplus. \tag{4}$$

As indicated by Eq. (3), in a disk with $f_{\rm disk} = 1$, $M_{\rm iso}$ can reach $M_{\rm crit}$ only outside the snow line at $\sim$ 3 AU. On the other hand, at $a \gtrsim 10$ AU, protoplanets cannot grow enough for gas accretion during disk life time ($\tau_{\rm disk} \sim 10^7$ years). So, formation of gas giants is favored just outside the snow line. However, Eq. (3) indicates that in very massive disks ($f_{\rm disk} \gtrsim 10$), $M_{\rm iso}$ can reach $M_{\rm crit}$ even inside the snow line.

To see formation sites of gas giants more clearly, Ida and Lin(2004) (Ida and Lin(2004)) did analytical calculations of core and gas accretion. The planetary growth rate is estimated by the two-body approximation with realistic physical sizes and damping of velocity dispersion due to gas drag (for details, see Kokubo and Ida, 2002; Ida and Lin(2004), Ida and Lin(2004)). The effect of realistic sizes increases timescales in Fig. 2 by a factor of 6–36, while the gas drag effect reduces the timescales by a factor of several. Using the planetary growth rate ($\dot{M}_{\rm sc}$), they calculated time evolution of protoplanet masses and their $M_{\rm crit}$.

They modeled that when $M_{\rm sc}$ reaches $M_{\rm crit}$, gas accretion onto a core starts with a rate,

$$\frac{dM_{\rm p}}{dt} \simeq \frac{M_{\rm p}}{\tau_{\rm KH}};\ \tau_{\rm KH} \simeq 10^9 \left(\frac{M_{\rm p}}{M_\oplus} \right)^{-3} \text{ years}, \tag{5}$$

where $M_{\rm p}$ is the planet mass including gas envelope and $\tau_{\rm KH}$ is Kelvin-Helmholtz contraction time of the planetary atmosphere (Ikoma et al., 2000; slightly modified). Note that when $M_{\rm sc}$ reaches $M_{\rm iso}$, $M_{\rm crit} = 0$ and gas accretion starts independent of the value of $M_{\rm sc}$. However, if $M_{\rm sc} \lesssim 5M_\oplus$, the gas accretion is so slow (Eq. 5) that $M_{\rm p}$ is hardly increased by the gas accretion during the disk life time. In order for a gas giant to form, $\tau_{\rm KH}$ must be smaller than $\tau_{\rm disk}$ in addition to $M_{\rm sc} > M_{\rm crit}$.

Equation (5) shows that gas accretion is accelerated with time. When $M_{\rm p}$ becomes massive enough to open up a gap in the disk, gas accretion would be truncated (e.g., Lin and Papaloizou, 1985). Ida and Lin(2004) (Ida and Lin(2004)) adopted a simple prescription for the truncation: gas accretion stops when all the gas in "feeding zone" is captured by the planet, where the feeding zone is defined by the smaller one of the region within $r_{\rm H}$ (where $M_{\rm sc}$ in Eq. (2) is replaced by $M_{\rm p}$) from the planetary orbit and one within 3:2 mean motion resonance. In this case, the truncation mass is $\sim 10 - 100 M_{\rm J}$. Such a simple prescription may be sufficient to show dependence of regions forming gas giants on disk mass.

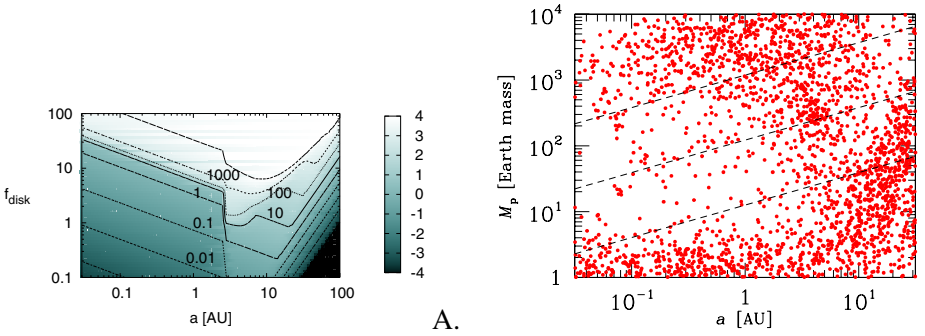

Figure 3. A) Planet masses including gas envelope at 10^9 years as a function of $f_{\rm disk}$ and a. Labels in the contours are $M_{\rm p}/M_\oplus$. (Numbers of the color box are $\log_{10}(M_{\rm sc}/M_\oplus)$.) In this calculation, $\tau_{\rm disk} = 10^7$ years and $a_{\rm snow} = 2.7$ AU. B) Theoretically predicted distribution based on the core accretion model for gas giants and oligarchic growth of the cores (for details, see text).

The disk gas surface density ($\Sigma_{\rm gas}$) is assumed to decay as

$$\Sigma_{\rm gas} \simeq f_{\rm disk} \times 2400 \left(\frac{a}{1\ {\rm AU}}\right)^{-3/2} \exp\left(-\frac{t}{\tau_{\rm disk}}\right)\ {\rm g\,cm^{-2}}. \quad (6)$$

Since the truncation mass assumed in the above prescription decreases with decrease in $\Sigma_{\rm gas}$ on a timescale $\tau_{\rm disk}$, cores that grows on timescales longer than $\tau_{\rm disk}$ cannot capture the gas.

Figure 3A shows the estimated planet masses including gas envelope at 10^9 years as a function of $f_{\rm disk}$ and a. $\tau_{\rm disk} = 10^7$ years and $a_{\rm snow} = 2.7$ AU are adopted. In the case of $f_{\rm disk} \sim 1$, gas giants with $M_{\rm p} \gtrsim 100 M_\oplus$ ($\sim 0.3 M_{\rm J}$) are formed in the limited regions outside the snow line. In massive disks with $f_{\rm disk} \gtrsim 10$, faster core growth and larger $M_{\rm iso}$ allow gas giants to form even inside the snow line. On the other hand, in less massive disks with $f_{\rm disk} \lesssim 0.5$, no gas giants are formed. In larger a regions (in particular, outside the snow line), $M_{\rm iso}$ can be larger than $10 M_\oplus$, while core accretion timescales can be shorter than $\tau_{\rm disk}$ in small a regions. If both regions overlap, gas giants are formed. The fraction of the overlapping regions increases with disk mass (equivalently, $f_{\rm disk}$).

Gas accretion is increasingly accelerated until a gap is opened, once $M_{\rm sc}$ exceeds $M_{\rm crit} \sim 10 M_\oplus$. So, the regions of planets with $M_{\rm p} \sim 10 - 100 M_\oplus$ are very small at $a \lesssim 3$ AU in Fig. 3A. In outer regions ($a \sim 10$ AU), however, icy planets in such mass ranges can be formed. They grow so slowly to miss gas accretion but they eventually reach $\gtrsim 10 M_\oplus$ after the gas is dissipated

Using the same theoretical model of core growth and gas accretion, Ida and Lin(2004) (Ida and Lin(2004)) did Mote Carlo calculations to produce a theoretical prediction of the $M_{\rm p} - a$ distribution of extrasolar planets to compare it with observations (Fig. 1). Planets are assumed to form with equal probability per interval of $\log a$. The assumed distribution of $\log_{10} f_{\rm disk}$ has the range from –1 to 2 peaked at 0.5 and that of $\log_{10} \tau_{\rm disk}$ is 6–8 peaked at 7 to be consistent with

observational data (Beckwith and Sargent, 1996). The range of stellar mass is also considered, since $a_{\rm snow} \propto L_*^{1/2} \propto M_*^2$, where L_* and M_* are stellar luminosity and mass. According to the host stars harboring extrasolar planets so far discovered, the range is $M_* = 0.7 - 1.4 M_\odot$ (correspondingly, $a_{\rm snow} = 1.3 - 5.3$ AU).

The result in Fig. 3B, clearly shows a deficit of intermediate mass ($\sim 10 - 100 M_\oplus$) planets at $a \lesssim 3$ AU. Giant planets may migrate (type II migration) from original formation sites at large a and pile up near the disk inner edge at $a \lesssim 0.1$ AU (Lin et al. 1995), although such an effect have not been included here. In such inner regions, Roche lobe overflow (Trilling et al., 1998) may decrease giant planet masses of $\gtrsim 100 M_\oplus$ to the intermediate level. However, at 0.2–3 AU, the intermediate mass planets are resulted only when gas dissipates just during the migration (Trilling et al., 2002) or rapid gas accretion (such planets can be found in Fig. 3B). The probability should be small. Hence, the result in Fig. 3B predicts the deficit of planets with $M_{\rm p} \sim 10 - 100 M_\oplus$ at $a \simeq 0.2$–3 AU ("planet desert"). In the regions of radial velocity $\gtrsim 10$ m/s, the observations (Fig. 1) seem consistent with the idea of the planet desert predicted by our theoretical model.

4. Discussions

We have discussed dependence of formation sites of gas giants on disk mass, based on the core accretion model of gas giant planets. Observations suggest the disk masses range from 0.1 to 100 times of the minimum mass disk for our Solar system.

Our model predicts the deficit of intermediate mass planets ($\sim 10 - 100 M_\oplus$) at $a \simeq 0.2$–3 AU. Normally, gas giants can be formed outside the snow line at a few AU; inside the snow line, planet masses remain $\lesssim 10 M_\oplus$. In very massive disks, however, cores become massive enough at $a \simeq 0.2$–3 AU within disk the life time, and they can start rapid gas accretion to become $\gtrsim 100 M_\oplus$. Thus, the intermediate mass planets are rare ("planet desert").

Gas giants could also be formed by self-gravitational collapse of a disk (see Chapter by A. Boss). Although the mass distribution produced by it is not clear, if the planet desert is observationally confirmed, the core accretion model may be favored as a formation mechanism of extrasolar gas giants. The observational data up to now with $v_r \gtrsim 10$ m/s and $a \lesssim 4$–5 AU may show a hint of the planet desert. However, in order to clarify whether the planet desert exists or not, we need further accumulation of observational data to find planets with longer periods (larger a) and higher-resolution observations to discover planets with smaller v_r.

Acknowledgements

I wish to thank E. Kokubo and D. N. C. Lin for useful discussions.

References

Artymowicz, P.: 1993, 'Disk-Satellite Interaction via Density Waves and the Eccentricity Evolution of Bodies Embedded in Disks'. *Astrophys. J.* **419**, 166–180.

Beckwith, S. V. W., A. I. Sargent: 1996, 'Circumstellar disks and the search for neighbouring planetary systems'. *Nature* **383**, 139–144.

Hayashi, C., K. Nakazawa. Y. Nakagawa: 1985, 'Formation of the Solar System'. In: D. C. Black, M. S.Matthew (eds.): *Protostars and Planets II*, Univ. of Arizona Press, AZ, pp. 1100–1153.

Ida, S., J. Makino: 1993, 'Scattering of planetesimals by a protoplanet: Slowing down of runaway growth'. *Icarus* **106**, 210–227.

Ida, S. and Lin, D.N.C.: 2004, *Astrophys. J.* **604**, 388–413.

Ikoma, M., K. Nakazawa, H. Emori: 2000, 'Formation of Giant Planets: Dependences on Core Accretion Rate and Grain Opacity'. *Astrophys. J.* **537**, 1013–1025.

Iwasaki, K., H. Emori, K. Nakazawa, H. Tanaka: 2002, 'Orbital Stability of a Protoplanet System under a Drag Force Proportional to the Random Velocity'. *PASJ* **54**, 471–479.

Jorissen, A., M. Mayor, S. Udry: 2001, 'The distribution of exoplanet masses'. *Astron. Astrophys.* **379**, 992–998.

Kokubo, E., S. Ida: 1996, 'On Runaway Growth of Planetesimals'. *Icarus* **123**, 180–191.

Kokubo, E., S. Ida: 1998, 'Oligarchic Growth of Protoplanets'. *Icarus* **131**, 171–178.

Kokubo, E., S. Ida: 2000, 'Formation of Protoplanets from Planetesimals in the Solar Nebula'. *Icarus* **143**, 15–27.

Kokubo, E., S. Ida: 2002, 'Formation of Protoplanet Systems and Diversity of Planetary Systems'. *Astrophys. J.* **581**, 666–680.

Kominami, J., S. Ida: 2002, 'The Effect of Tidal Interaction with a Gas Disk on Formation of Terrestrial Planets'. *Icarus* **157**, 43–56.

Lin, D. N. C., Papaloizou, J. C. B.: 1985, 'On the dynamical origin of the solar system' In: D. C. Black, M. S.Matthew (eds.): *Protostars and Planets II*, Univ. of Arizona Press, AZ, pp. 981–1072.

Lin, D. N. C., P. Bodenheimer, P., D. Richardson: 1996, 'Orbital migration of the planetary companion of 51 Pegasi to its present location'. *Nature* **380**, 606–607.

Pätzold, M., Rauer, H.: 2002, 'Where are the Massive Close-In Extrasolar Planets'. *Astrophys. J.* **568**, L117–L120.

Safronov, V.:1969, *Evolution of thr Protoplanetary Cloud and Formation of the Earth and Planets*. Nauka Press, Moscow.

Tabachnik, S., S. Tremaine: 2002, 'Maximum-likelihood method for estimating the mass and period distributions of extrasolar planets'. *MNRAS* **335**, 151–158.

Thommes, E. W., M. J. Duncan, H. F. Levison: 2002, 'Oligarchic growth of giant planets' *Icarus* **161**, 431–455.

Trilling, D. E., W. Benz, T. Guillot, J. I. Lunine, W. B. Hubbard, A. Burrows: 1998, 'Orbital Evolution and Migration of Giant Planets: Modeling Extrasolar Planets'. *Astrophys. J.* **500**, 428-439.

Trilling, D. E., J. I. Lunine, W. Benz: 2002, 'Orbital migration and the frequency of giant planet formation'. *Astron. Astrophys.* **394**, 241–251.

Ward, W.: 1993, 'Density waves in the solar nebula - Planetesimal velocities'. *Icarus* **106**, 274–287.

Wetherill, G. W., G. R. Stewart: 1989, 'Accumulation of a swarm of small planetesimals'. *Icaurs* **77**, 330–357.

Wuchterl, G., T. Guillot, J. J. Lissauer: 2000, 'Giant Planet Formation'. In: V. Mannings, A. P. Boss, S. Russell, S. (eds.): *Protostars and Planets IV*, University of Arizona Press, AZ, pp. 1081– 1110.

Zucker, S., T. Mazeh: 2002, 'On the Mass-Period Correlation of the Extrasolar Planets'. *Astrophys. J.* **568**, L112–L116.

Chaotic Interactions in Multiple Planet Systems

Eric B. Ford
Department of Astrophysical Sciences, Princeton University, Peyton Hall – Ivy Lane, Princeton, NJ 08544-1001, USA

Frederic A. Rasio and Kenneth Yu
Northwestern University, Dept. of Physics & Astronomy, 2145 Sheridan Rd., Evanston, IL 60208-0834, USA

Abstract. The detection of $\sim$ 100 extrasolar planets has provided many challenges for theories of planet formation. Numerous mechanisms have been proposed to explain the surprising distributions of semi-major axes and eccentricities. Dynamical instabilities and chaotic interactions in planetary systems with multiple giant planets appear to provide a natural mechanism for producing the highly eccentric orbits frequently observed.

In a protoplanetary disk, the semi-major axis of a protoplanet is determined before the mass of the eventual planet. As nearby protoplanets accrete mass, a dynamical instability may lead to close encounters. Alternatively, a dissipative disk may prevent mutual interactions from exciting significant eccentricities. Once the disk clears, the eccentricities may grow and eventually permit close encounters. In any case, once planets begin to undergo close encounters, the planets may collide or be ejected from the system. These processes can significantly alter the orbits of the remaining planets. Numerical orbital integrations of possible planetary systems can determine the frequencies of the final outcomes for such systems and allow for comparison with observations.

Simulations for two equal-mass planets initially on nearly circular coplanar orbits result in a bimodal distribution for the final eccentricity. Initial conditions leading to a collisions between the two planets produce a more massive planet in a nearly circular orbit. Instead when one planet is ejected from the system, the other planet remains on an eccentric orbit with $0.4 \leq e \leq 0.8$. However, if the two planets have different masses, then the eccentricity after an ejection depends on the ratio of the planet masses. The observed distribution of eccentricities can be well reproduced for plausible distributions of the planet mass ratio. While a proper comparison would require careful consideration of observational selection effects and the unknown initial distributions, the two-planet scattering model predicts a maximum eccentricity of $\simeq 0.8$ independent of these complications.

Simulations for three equal-mass planets initially on nearly circular coplanar orbits also result in a broad distribution of final eccentricities which seldom exceeds 0.8. The three-planet scattering model distinguishes itself by predicting that an additional planet typically lies on a longer period orbit. Further, the relative inclination between the two planets is typically increased, up to $\sim 40^\circ$. Future observations, including long-term radial velocity measurements, long-term precision astrometric measurements, and direct imaging, could test these predictions.

1. Introduction

The first planet discovered around a solar type star, 51 Pegasi b, was in a surprisingly short period orbit (Mayor and Queloz, 1995). Other early planets such as 70 Virginis b revealed surprisingly large orbital eccentricities (Marcy and Butler 96). Multiple planet systems have revealed intricate dynamical interactions such as the apsidal lock in Upsilon Andromedae (Butler et al., 1999) and GJ876 (Marcy et al.,

2001). These findings challenge traditional theories of planet formation that were originally developed to explain our own Solar System.

Rising to the challenge, theorists have developed a variety of possible explanations for these findings. The known extrasolar planets (Fig. 1 left) can be roughly divided into two groups: those with short-period, nearly circular orbits ($a \leq 0.07$ AU) and those with wider and more eccentric orbits ($a \geq 0.07$ AU). While the short-period giant planets may have formed in situ (Bodenheimer, Olenka, and Lissauer, 2000), it is generally believed that these planets migrated from the site of their formation beyond the snow line (e.g. Sasselov and Lecar, 2000). Several mechanisms have been proposed to explain the large scale orbital migration presumed for such planets. Possiblities include viscous evolution of a planet embedded in a gaseous disk (e.g., Nelson et al., 2000), scattering of planetesimals by a planet in a massive protoplanetary disk (Murray, Hansen, Holman, and Tremaine, 1998), and dynamical interactions in multiple planet systems (Rasio and Ford, 1996; Weidenschilling and Marzari, 2002; Lin and Ida, 1997). Many of the short-period planets, like their prototype 51 Peg, are so close to their parent star that tidal dissipation would have likely circularized their orbits, even if they were originally eccentric (Rasio *et al.* 1996). Thus, their small observed eccentricities do not provide a good indicator of their dynamical history.

Similarly, the planets in eccentric orbits are generally believed to have formed on nearly circular orbits and evolved to their presently observed large eccentricities. While it is possible for migration in either a gaseous or planetesimal disk to generate significant eccentricities for a planet of at least several Jupiter masses, these mechanisms are unlikely to explain the eccentricities of lower mass planets or large eccentricities (Artymowicz, 1992). Mechanisms for producing large eccentricities include a passing binary star (Laughlin and Adams, 1998), secular perturbations due to a distant stellar or planetary companion (Holman, Touma, and Tremaine, 1997; Ford, Kozinsky, and Rasio, 2000), and strong planet-planet scattering events (Rasio and Ford, 1996; Weidenschilling and Marzari, 1996; Ford, Havlickova, and Rasio, 2001; Marzari and Weidenschilling, 2002). In the planet-planet scattering model, a planetary system with two or more giant planets becomes dynamically unstable, leading to a collision or the ejection of one of the planets from the system.

While several mechanisms have been proposed to explain the surprising orbits of extrasolar planets, it is unclear to what extent each mechanism is relevant in actual planetary systems, or whether completely new mechanisms are required. In this paper we discuss dynamical instabilities, which seem to provide a natural mechanism for producing the highly eccentric orbits seen in many systems.

2. Origin of Instability

Since dynamically unstable planetary systems are highly chaotic, it is not feasible to conduct numerical integrations of a particular system. Instead, investigations

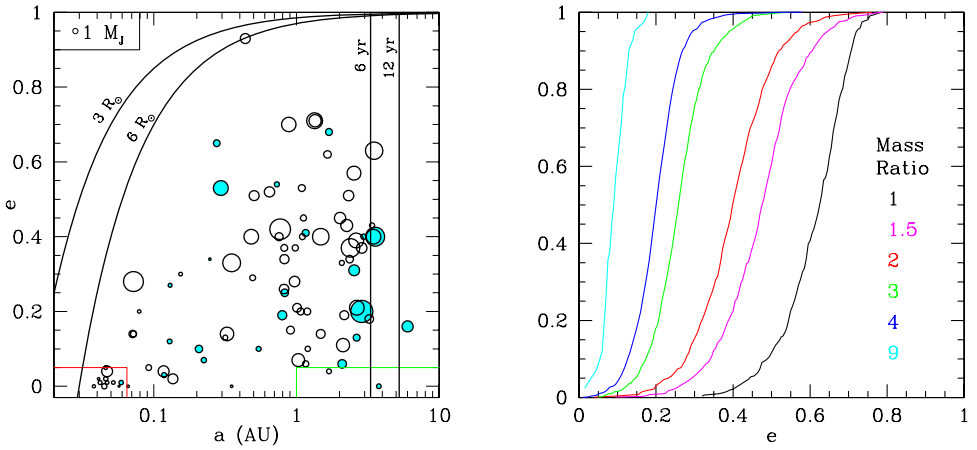

Figure 1. Left panel: Orbital eccentricity versus semi-major axis of known extrasolar planets. The eccentricity of planets in very short period orbits is limited by tidal damping. Nearly all of the remaining planets have large eccentricities. The area of the circle is proportional to $m \sin i$. Planets represented by shaded circles are part of a multiple planet system. *Right panel:* Cumulative distribution of eccentricity. Each line shows the cumulative distribution of the eccentricity of the remaining planet after the other planet has been ejected from the system in our simulations for a particular mass ratio. The distributions are not sensitive to the total mass of the planets or to which planet is more massive.

consider the statistical properties of an ensemble of planetary systems with similar initial conditions. While some authors have simulated multiple planet systems beginning with the planet formation stage, computational limitations have limited such simulations to a small number of initial conditions (Kokubo and Ida, 1998; Levison, Lissauer, and Duncan, 1998). Thus, most investigations of dynamical instabilities in multiple planet systems proceed by simulating planetary systems after planets have formed and perturbations due to the disk are no longer significant. The planets are placed on plausible initial orbits and numerically integrated according to the gravitational potential of the central star and other planets.

Clearly, the choice of initial conditions will determine whether the systems are dynamically stable and will affect the outcome of unstable systems. One potential concern about the relevance of dynamical instabilities is whether the necessary initial conditions will manifest themselves in nature. There are several possible mechanisms which could lead to dynamical instabilities in multiple planet systems, including mass growth, accumulation of significant random velocities, and orbital migration.

According to core accretion models, once a rocky planetary core reaches a critical mass, it rapidly accretes the gas within its radius of influence in a circumstellar disk. Thus, the semi-major axis of a planet is determined by the collisional evolution of protoplanets, while the mass of a giant planet is determined by the state of the gaseous disk when the core reaches the critical mass (Lissauer, 1993). Two

planetary cores could form with a separation sufficient to prevent close encounters while their masses are less than the critical mass for runaway accretion, but then the rapid mass growth due to gas accretion could render their separation insufficient to prevent a dynamical instability.

The accumulation of random velocities provides another possible source of a dynamical instability. Assuming planets form in the presence of a dissipative disk, they are expected for form on nearly circular and coplanar orbits. While the timescale for dissipation in the disk is short remains shorter than the timescales for secular perturbations, eccentricities and inclination will be damped to keep planets on nearly circular and coplanar orbits. As the disk dissipates, planets could acquire significant eccentricities and inclinations due to mutual planetary perturbations or perturbations due to a distant stellar companion. Either source can cause large secular eccentricity oscillations which could eventually lead to close encounters between planets and eventually result in a collisions or ejections.

Finally, the discovery of giant planets at small orbital separations suggests that large-scale orbital migration may be common. In multiple planet systems, the migration mechanism might preferentially affect one planet and result in convergent migration, i.e., a decreasing ratio of semi-major axes. While migration in a gaseous disk typically results in divergent orbital migration, planetesimal scattering could result in convergent migration if the outer planet is more massive and hence more effective at ejecting planetesimals. Observationally, the apsidal lock and small libration amplitude in the υ Andromedae system suggests convergent migration may have occurred in this system (Chiang, Tabachnik, and Tremaine, 2001).

With so many possibilities for triggering dynamical instabilities in multiple planet systems, it seems likely that these instabilities could be common. Further, many of the known planetary systems are near the limit of stability. For example, several studies of the υ Andromedae system found that the system rapidly becomes unstable as the eccentricity of the outer planet increases above 0.25 (e.g., Barnes and Quinn, 2001). While the present best estimate of the eccentricity of υ And d is close to this transition, these stability investigations typically assume that the orbital elements are uncorrelated. Consideration of the correlations between orbital elements may lead to more stable systems. Similarly, the masses, semi-major axes, and eccentricities of the GJ 876 system would render it unstable were it not for the protection of the 2:1 mean motion and secular resonances (Lee and Peale, 2002). Additionally, the observational uncertainties in the orbital elements of several systems (GJ 876, 47 UMa, 55 Cnc) allow dynamically unstable configurations (Marcy *et al.* 2001; Fischer *et al.* 2002; Marcy *et al.* 2002).

While our own Solar System has been dynamically stable for billions of years, some long term integrations result in Venus ejecting Mercury within the next four billion years (Laskar, 1994). Additionally, chemical analyses provide strong evidence that the Earth's moon formed via an oblique collision between the Earth and another massive body and the accretion of the Moon from the resulting debris disk (Newson and Taylor, 1989; Canup and Asphaug 2001). The presence of short-

lived ring systems around the giant planets and particularly the large ring system of Saturn suggest that collisions between giant planets and smaller mass bodies are frequent (Colwell, 1994). Collisions with massive bodies may also explain Uranus's large obliquity (Korycansky et al., 1990) and the peculiar orbits of Neptune's moons, particularly Triton (McKinnon 1984). Together, these observations suggest that even our relatively quiet Solar System may have had a violent early history.

3. Numerical Investigations

3.1. Two Equal-Mass Planets

Shortly after the discovery of 51 Pegasi b, Rasio and Ford (1996) conducted Monte Carlo integrations of planetary systems containing two equal-mass planets initially placed just inside the Hill stability limit (Gladman, 1993). They numerically integrated the orbits of such systems until there was a collision, or one planet was ejected from the system, or some maximum integration time was reached. The two most common outcomes were collisions between the two planets, producing a more massive planet in a nearly circular orbit between the two initial orbits, and ejections of one planet from the system while the other planet remains in a tighter orbit with a large eccentricity. The relative frequency of these two outcomes depends on the ratio of the planet radius to the initial semi-major axis.

While this model could naturally explain how planets might acquire large eccentricities, Ford, Havlickova, and Rasio (2001) compared the results of more extensive simulations to the observed planets and found two important differences. First, for the relevant radii and semi-major axes, collisions are more frequent in our simulations than nearly circular orbits among the presently known extrasolar planets. Second, in systems which led to an ejection, the eccentricity distribution of the remaining planet was concentrated in a narrow range which is greater than the typical eccentricity of the known extrasolar planets (See Fig. 1, right, rightmost curve).

3.2. Two Planets, Unequal Masses

We have conducted new simulations of planetary systems contain two planets with unequal masses (Ford, Rasio, and Yu, 2003). Here we summarize these recent results.

3.2.1. *Initial Conditions*

The new simulations use the mixed variable symplectic algorithm (Wisdom and Holman, 1991) modified to allow for close encounters between planets as implemented in the publicly available code Mercury (Chambers, 1999). The results presented below are based on $\sim 10^4$ numerical integrations.

Throughout the integrations, close encounters between any two bodies were logged, allowing us to present results for any values of the planetary radii using a single set of orbital integrations. We consider a range of radii to allow for the uncertainty in both the physical radius and the effective collision radius allowing for dissipation in the planets. When two planets collided, mass and momentum conservation were used to compute the final orbit of the resulting single planet.

Each run was terminated when one of the following four conditions was encountered: (i) one of the two planets became unbound (which we defined as having a radial distance from the star of $1000\,a_{1,\rm init}$); (ii) a collision between the two planets occurred assuming $R_i/a_{1,\rm init} = R_{\rm min}/a_{1,\rm init} = 0.1\,R_{\rm Jup}/5\,{\rm AU} = 0.95 \times 10^{-5}$; (iii) a close encounter occurred between a planet and the star (defined by having a planet come within $r_{\rm min}/a_{1,\rm init} = 10\,R_\odot/1\,{\rm AU} = 0.06$ of the star); (iv) the integration time reached $t_{\rm max} = 5 \times 10^6 - 2 \times 10^7$ depending on the masses of the planets. These four types will be referred to as "collisions," meaning a collision between the two planets, "ejections," meaning that one planet was ejected to infinity, "star grazers," meaning that one planet had a close pericenter passage, and "two planets," meaning that two bound planets remained in a (possibly new) dynamically stable configuration.

Our numerical integrations were performed for a system containing two planets, with mass ratios $10^{-4} < m_i/M < 10^{-2}$, where m_i is the mass of one of the planets and M is the mass of the central star. A mass ratio of $m/M \simeq 10^{-3}$ corresponds to $m \simeq 1\,M_{\rm Jup}$ for $M = 1\,M_\odot$. The initial semimajor axis of the inner planet ($a_{1,\rm init}$) was set to unity and the initial semimajor axis of the outer planet ($a_{2,\rm init}$) was drawn from a uniform distribution ranging from $0.9a_{1,\rm init}\,(1+\Delta_c)$ to $a_{1,\rm init}\,(1+\Delta_c)$, where $1+\Delta_c$ is the critical ratio above which Hill stability is guaranteed for initially circular coplanar orbits (Gladman, 1993). The initial eccentricities were distributed uniformly in the range from 0 to 0.05, and the initial relative inclination in the range from 0° to 2°. All remaining angles (longitudes and phases) were randomly chosen between 0 and 2π. Throughout this paper we quote numerical results in units such that $G = a_{1,\rm init} = M = 1$. In these units, the initial orbital period of the inner planet is $P_1 \simeq 2\pi$.

3.2.2. *Branching Ratios*

For the range of m_i/M considered (up to 10^{-2}), the frequency of the various outcomes is relatively insensitive to the ratio of the total planet mass to the stellar mass. However the frequencies of the different outcomes are significantly effected by the planetary mass ratio and by the planetary radii. Collisions are most common for large planetary radii and mass ratios near unity, while ejections are most common for small planetary radii and mass ratios far from unity. The frequency of outcomes does not depend on whether the more massive planet is initially in the inner or outer orbit. In Fig. 2 (left), we show the frequency of ejection or collision as a function of the planetary radii with solid and dashed lines, respectively. The lines labeled "1" show the result for equal-mass planets. The other lines are for different mass ratios.

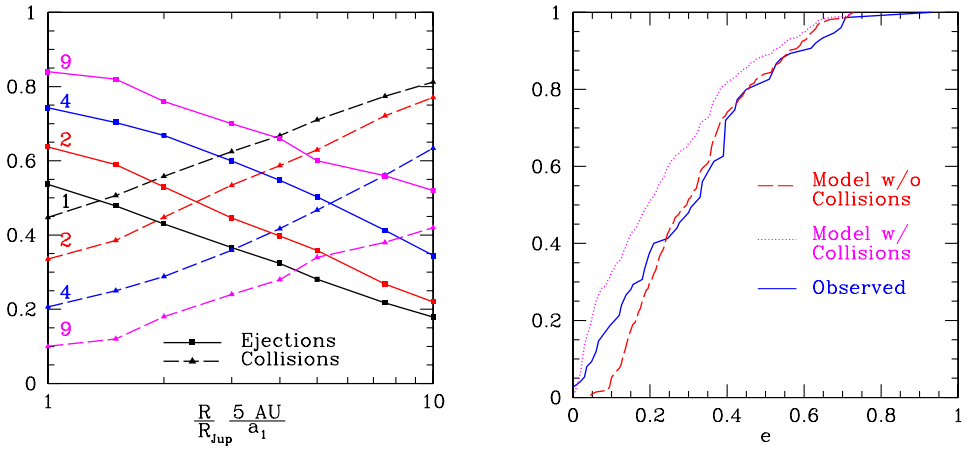

Figure 2. Left panel: Branching ratio versus planet radius. The frequency of collisions (dashed lines) and ejections (solid lines) depends on the planet's radius relative to the semi-major axis (x-axis) and the planetary mass ratio (different lines). The fraction of integrations resulting in one planet being ejected and the other planet remaining in an eccentric orbit is increased when the two planetary masses differ. *Right panel:* Simulation with mass distribution. Here we compare the observed distribution of eccentricities (excluding planets with $P < 10$ d, solid line) to the eccentricity distributions produced by our simulations when each planet's mass is drawn independently from $P(m) \sim m^{-1}$ for $1\,M_{\rm Jup} < m < 10\,M_{\rm Jup}$. The dashed line shows the eccentricity of the planet remaining after an ejection, while the dotted line includes systems resulting in a collision.

Ejections become more common and collisions less frequent as the planetary mass ratio departs from unity.

3.2.3. *Collisions*

Collisions leave a single, larger planet in orbit around the star. The energy in the center-of-mass frame of the two planets is much smaller than the binding energy of a giant planet. Therefore, we model the collisions as completely inelastic and assume that the two giant planets simply merge together while conserving total momentum and mass. Under this assumption, we have calculated the distributions of orbital parameters for the collision products. The final orbit has a semi-major axis between the two initial semi-major axes, a small eccentricity, and a small inclination. While collisions between planets may affect the masses of extrasolar planets, a single collision between two massive planets does not cause significant orbital migration or eccentricity growth if the planets are initially on low-eccentricity, low-inclination orbits near the Hill stability limit.

3.2.4. *Ejections*

The typical time until an ejection scales approximately as $m^{-3/2}$ for equal-mass planets. While the dependance on mass ratio is more complex, the timescale is longest for equal-mass planets, assuming the sum of the planet masses is held

constant. When one planet is ejected from the system, the less massive planet is nearly always ejected if $\frac{m_<}{m_>} \leq 0.8$, where $m_<$ and $m_>$ refer to the masses of the less and more massive planets, respectively. Since the escaping planet typically leaves the system with a very small (positive) energy, energy conservation sets the final semimajor axis of the remaining planet slightly less than

$$\frac{a_f}{a_1} \simeq \frac{a_2\, m_<}{a_1\,(m_1+m_2)+(a_2-a_1)\,m_1}. \tag{1}$$

Thus, the ejection of one of two equal-mass planets results in the most significant reduction in the semi-major axis, but is limited to $\frac{a_f}{a_{1,\mathrm{init}}} \geq 0.5$. The remaining planet acquires a significant eccentricity, but its inclination typically remains small. The eccentricity and inclination distributions for the remaining planet are not sensitive to the sum of the planet masses, but depends significantly on the mass ratio. Both the final eccentricity and inclination are maximized for equal-mass planets.

In Fig. 1 (right) we show the cumulative distributions for the eccentricity after a collision for different mass ratios. While any one mass ratio results in a narrow range of eccentricities, a distribution of mass ratios would result in a broader distribution of final eccentricities. Also note that there is a maximum eccentricity which occurs for equal-mass planets. Thus, the two-planet scattering model predicts a maximum eccentricity of about 0.8 independent of the distribution of planet masses. Of the $\sim$ 100 known extrasolar planets, only one has an eccentricity larger than 0.8, HD80606, which has an eccentricity of 0.93 and a pericenter distance of 0.03 AU (Naef *et al.* , 2001). While the predicted maximum eccentricity compares favorably with the presently known planets, future discoveries of additional extrasolar planets will certainly test this prediction.

3.2.5. *Stargrazers*

In a small fraction of our numerical integrations ($\sim$ 3%) one planet underwent a close encounter with the central star. Due to the limitations of the numerical integrator used, the accuracy of our integrations for the subsequent evolution of these systems cannot be guaranteed. Moreover, these systems could be affected by additional forces (e.g., tidal forces, interaction with the quadrupole moment of the star) that are not included in our simulations and would depend on the initial separation and the radius of the star. Nevertheless, our simulations suggest that for giant planets with initial semimajor axes of a few AU, the extremely close pericenter distances necessary for tidal circularization around a main sequence star (leading to the formation of a 51 Peg-type system) are possible.

3.2.6. *Comparison to Observations*

We now consider whether the the two-planet scattering model could produce a distribution of eccentricities consistent with that of known extrasolar planets with periods longer than 10 days. In Fig. 2 (right) we show the cumulative distribution for the eccentricity of the remaining planet in our simulations. Here we have chosen

a simple distribution of planet masses, $P(m) \simeq m^{-1}$ for $1M_{\rm Jup} \leq m \leq 10M_{\rm Jup}$, which is consistent with the mass distribution of the observed extrasolar planets (Tabachnik and Tremaine, 2002). The dashed line is just for systems which resulted in an ejection, while the dotted line includes systems which resulted in a collision. Both are reasonably close to the observed distribution shown by the solid black line. It should be noted that the mass distribution used for this model has two cut-offs. While the $10M_{\rm Jup}$ cutoff is supported by observations, the lower cutoff is not constrained by present observations and is unlikely to be physical. Alternative mass distributions differing for $m \leq 1M_{\rm Jup}$ could alter this result. A detailed comparison to observations would require the careful consideration of observational selection effects, as well as the unknown initial mass distributions. Nevertheless, we find that the two-planet scattering model is able to reproduce the eccentricity distribution of the known extrasolar planets for plausible mass distributions.

The fraction of the systems producing stargrazers is comparable to the fraction of stars in radial velocity surveys that have very short period extrasolar planets (51 Peg type). However, it is smaller than the fraction of 51 Peg type planets among detected extrasolar planets. Therefore, if most giant planets around nearby solar-type stars have already been found by radial velocity surveys, then two-planet scattering cannot explain the abundance of giant planets in very short period orbits. However, if most nearby solar-type stars have giant planets in long period orbits not yet accessible to radial velocity surveys and the majority of nearby solar-type stars formed multiple giant planets which were subject to a dynamical instability, then the combination of dynamical instabilities and tidal circularization may be able to explain the frequency of giant planets in very short period orbits. The scattering of two giant planets has more difficulty explaining the presence of giant planets at these orbital periods, small enough to require significant migration, but large enough that tidal circularization is ineffective. Still, it might be possible to circularize giant planets at slightly larger distances if the circularization occurs while the star is pre-main sequence and has a larger radius or while a circumstellar disk is still present.

3.3. THREE PLANETS

Shortly after the discovery of the planet around 51 Pegasi, Weidenschilling and Marzari (1996) conducted simulations of planet-planet scattering in three-planet systems. More recently, Marzari and Weidenschilling (2002) presented the results of a larger number of integrations, considering only two sets of planet masses, but exploring a range of initial separations. Additionally, we begun performing numerical integrations of three equal-mass planet systems (Ford, Rasio, and Yu, 2003). Still, the large parameter space for three planet systems remains largely unexplored.

For three-planet systems, a single ejection or collision results in a system containing two planets. The remaining two planets may either be dynamically stable

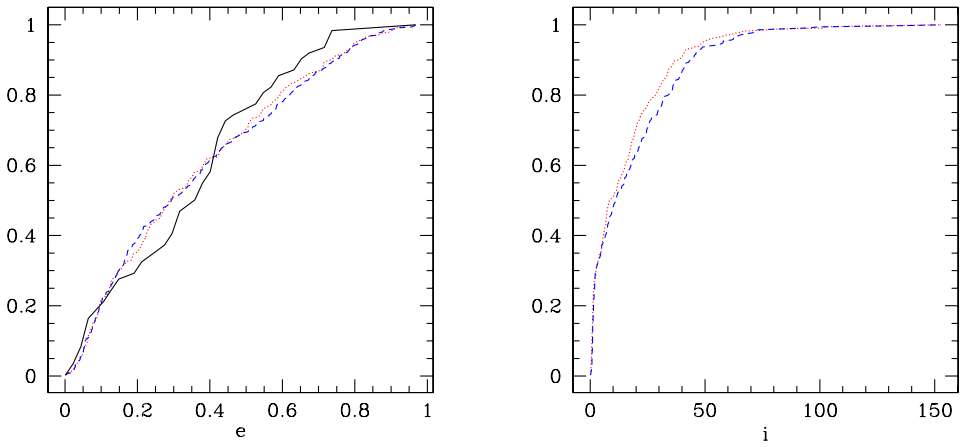

Figure 3. Results from $\sim$ 1000 simulations of 3 equal-mass planet systems (Ford, Rasio, and Yu, 2003). Each planet was given a mass ratio $m/M \simeq 10^{-3}$ and initially placed on a nearly circular orbit and given an inclination of $\pm 1^\circ$. The initial semi-major axes are fixed at 1, 1.39, and 1.92 AU, so that the ratio of orbital periods is $P_3/P_2 = P_2/P_1 = 1.63$. The typical time for instability to develop is $\sim 10^6$ years. The dashed and dotted lines are for simulations with $e_{\rm init} = 0.$ and $e_{\rm init} = 0.05$, respectively. *Left panel:* Cumulative distribution of eccentricity. Here we compare the observed distribution of eccentricities (excluding those with $P < 10$ d, solid black line) to the eccentricity distribution of the remaining inner planet after there has been one ejection or collision for one of two sets of simulations with different initial eccentricities. *Right panel:* Cumulative distribution of inclination. Each line shows the distribution of the inclination of the remaining inner planet (with respect to the initial orbital plane) after there has been one ejection or collision for one of two sets of simulations with different initial eccentricities.

or undergo a subsequent ejection or collision to end with a single planet. Typically, two planets are left in a stable configuration after one planet was ejected or two planets collided. Often, the remaining planets are widely spaced so that the outer planet would be very difficult to detect via a radial velocity survey.

In contrast to the case of two-planet systems, there is no sharp stability criterion for three-planet systems. Three-planet systems can be unstable even for initial orbital spacings significantly greater than would be necessary for similar two-planet systems to be stable. While unstable two-planet systems quickly result in close encounters, widely spaced three-planet systems can survive for several orders of magnitude longer. For example, three Jupiter mass planets with the innermost planet at 5 AU can evolve for $\sim 10^9$ years before the first close encounter (Marzari and Weidenschilling 2002). This longer timescale until close encounters could allow sufficient time for all three planets to form via either the disk instability or core accretion models.

In Fig. 3 we show the distributions of the eccentricity (left) and inclination (right) of the inner planet after a collision or ejection of one planet. Both are in qualitative agreement with those presented by Marzari and Weidenschilling (2002, based on slightly different initial conditions). The cumulative distribution

of the inner planet's eccentricities (Fig. 3, left) is in reasonable agreement with the observed distribution for extrasolar planets. The distribution of resulting eccentricities extends beyond 0.8, in contrast to what was found in two-planet simulations (Fig. 1, right). Another difference is that dynamical instabilities in three-planet systems can produce much larger inclinations than scattering in two-planet systems (Fig. 3, right). We also looked at the cumulative distribution for $\Delta\varpi = |(\Omega_2 + \omega_2) - (\Omega_1 + \omega_1)|$, the angle between the periapses of the two planets remaining after one planet has been ejected from the system. We do not find the two remaining planets to have preferentially aligned periapses (cf. Malhotra, 2002).

3.4. Many Planets

If giant planets form by disk fragmentation, then several massive planets might develop concurrently and lead to dynamical instabilities in planetary systems with many massive planets (Lin and Ida, 1997; Papaloizou and Terquem, 2001). Simulations of systems with five to one hundred massive planets often evolve to systems with two or three remaining planets. The most massive planet in a system can undergo some migration and typically becomes the innermost planet. The remaining planets can be either widely spaced or tightly packed and have large eccentricities and inclinations, including some retrograde orbits. Unfortunately, the enormous parameter space of possible initial conditions makes it difficult to conduct a comprehensive study of planet-planet scattering in systems with many planets.

4. Conclusion

A planetary system with two or more giant planets may become dynamically unstable, leading to a collision or the ejection of one of the planets from the system. Early simulations of equal-mass planets revealed discrepancies between the results of numerical simulations and the observed orbital elements of extrasolar planets. However, more recent simulations of two planets with *unequal masses* reveal a reduced frequency of collisions as compared to scattering between equal-mass planets and that suggests that the two-planet scattering model can reproduce the observed eccentricity distribution for a plausible distribution of planet mass ratios. Additionally, the two-planet scattering model predicts a maximum eccentricity of ~ 0.8, which is independent of the distribution of planet mass ratios. This predicted eccentricity limit compares favorably with current observations and will be tested by future planet discoveries.

Investigations of dynamical instabilities in systems with three or more planets have only begun to explore the large available parameter space. Dynamical instabilities in systems with three or more planets can produce a broad range of eccentricities, including eccentricities larger than 0.8. While the scattering of two

planets initially on nearly coplanar orbits does not result in large inclinations, dynamical instabilities in systems with three or more giant planets can produce large relative inclinations between the remaining planets, or even retrograde orbits. This prediction could be tested by observations sensitive to orbital inclinations such as astrometry and direct imaging.

The combination of planet-planet scattering and tidal circularization may be able to explain the existence of giant planets in very short period orbits. However, the presence of giant planets at slightly larger orbital periods (small enough to require significant migration, but large enough that tidal circularization is ineffective) is more difficult to explain. Finally, the planet-planet scattering model predicts a significant number of extremely loosely bound and free floating giant planets, which may be observable (Lucas and Roche, 2000; Zapatero Osorio *et al.* , 2002).

Acknowledgements

We are grateful to Eugene Chiang, Renu Malhotra, and Scott Tremaine for valuable discussions. E.B.F. thanks the ISSI, the organizers, and all the participants of this workshop for making it such an interesting and enjoyable experience. E.B.F. acknowledges the support of the NSF graduate research fellowship program and Princeton University, and thanks the Theoretical Astrophysics group at Northwestern University for hospitality. F.A.R. and K.Y. acknowledge support from NSF grant AST-0206182. Our computations were supported by the National Computational Science Alliance under Grant AST 980014N and utilized the SGI/Cray Origin2000 supercomputers at Boston University. Additional support for E.B.F. was provided by NASA through Hubble Fellowship grant HST-HF-01195.01A awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS 5-26555.

References

Canup, R.M., and Asphaug, E.:, 2001 'Origin of the Moon in a Giant Impact Near the End of the Earth's Formation', *Nature* **412**, 708–712.

Chambers, J.E., Wetherill, G.W., and Boss, A.P.: 1996, 'The Stability of Multi-Planet Systems', *Icarus* **119**, 261–268.

Chambers, J.E.: 1999, 'A Hybrid Symplectic Integrator that Permits Close Encounters between Massive Bodies', *M. Notices of Royal Ast. Soc.* **304**, 793–799.

Chiang, E.I., Tabachnik, S., and Tremaine, S.: 2001, 'Apsidal Alignment in υ Andromedae', *Astrophys. J.* **122**, 1607–1615.

Colwell, J.E.: 1994, 'The Disruption of Planetary Satellites and the Creation of Planetary Rings', *Planet and Space Sci.* **42**, 1139–1149.

Fischer, D.A., Marcy, G.W., Butler, R.P., Laughlin, G., and Vogt, S.S.: 2002, 'A Second Planet Orbiting 47 Ursae Majoris', *Astrophys. J.* **564**, 1028–1034.

Ford, E.B., Havlickova, M., and Rasio, F.A.: 2001, 'Dynamical Instabilities in Extrasolar Planetary Systems Containing Two Giant Planets', *Icarus* **150**, 303–313.

Ford, E.B., Rasio, F.A., and Yu, K.: 2003 'Dynamical Instabilities in Extrasolar Planetary Systems', to appear in the ASP Conference Series vol. 28: "Scientific Frontiers in Research on Extrasolar Planets", eds. D. Deming and S. Seager.

Gladman, B.: 1993, 'Dynamics of Systems of Two Close Planets', *Icarus* **106**, 247–263.

Kokubo, E. and Ida, S.: 1998, 'Oligarchic Growth of Protoplanets', *Icarus* **131**, 171–178.

Korycansky, D.G., Bodenheimer, P., Cassen, P., and Pollack, J.B.: 1998, 'One-dimensional Calculations of a Large Impact on Uranus', *Icarus* **84**, 528-541.

Laskar, J.: 1994, 'Large-scale Chaos in the Solar System', *Astron. Astrophys.* **287**, L9–L12.

Lee, M.H. and Peale, S.J.: 2002, 'Dynamics and Origin of the 2:1 Orbital Resonances of the GJ 876 Planets', *Astrophys. J.* **567**, 596–609.

Levison, H.F., Lissauer, J.J., and Duncan, M.J.: 1998, 'Modeling the Diversity of Outer Planetary Systems', *Astronom. J.* **116**, 1998–2004.

Lin, D.N.C. and Ida, S.: 1997, 'On the Origin of Massive Eccentric Planets', *Astrophys. J.* **447**, 781–791.

Lissauer, J.J.: 1993, 'Planet Formation', *Ann. Rev. Astron. & Astrophys.* **31**, 129–174.

Lucas, P.W. and Roche, P.F.: 2000, 'A Population of Very Young Brown Dwarfs and Free-floating Planets in Orion', *M. Notices of Royal Ast. Soc.* **314**, 858–864.

Malhotra, R.: 2002, 'A Dynamical Mechanism for Establishing Apsidal Resonance', *Astrophys. J.* **575**, L33–L36.

Marcy, G.W., Butler, R.P., Fischer, D., Vogt, S.S., Lissauer, J.J., and Rivera, E.J.: 2001, 'A Pair of Resonant Planets Orbiting GJ 876', *Astrophys. J.* **556**, 296–301.

Marcy, G.W., Butler, R.P., Fischer, D., Laughlin, G., Vogt, S.S., Henry, G.W., and Pourbaix, D.: 2002, 'A Planet at 5 AU around 55 Cancri', *Astrophys. J.* **581**, 1375–1388.

Marzari, F. and Weidenschilling, S.J.: 2002, 'Eccentric Extrasolar Planets: The Jumping Jupiter Model', *Icarus* **156**, 570–579.

McKinnon, W.B.: 1984, 'On the Origin of Triton and Pluto', *Nature* **311**, 355–358.

Naef, D., Latham, D.W., Mayor, M., Mazeh, T., Beuzit, J.L., Drukier, G.A., Perrier-Bellet, C., Queloz, D., Sivan, J.P., Torres, G., Udry, S., and Zucker, S.: 2001 'HD 80606 b, A planet on an Extremely Elongated Orbit', *Astron. Astrophys.* **375**, L27–L30.

Newson, H.E. and Taylor, S.R.: 1989, 'Geochemical Implications of the Formation of the Moon by a Single Giant Impact', *Nature* **338**, 29–34.

Papaloizou, J.C.B. and Terquem, C.: 2001, 'Dynamical Relaxation and Massive Extrasolar Planets', *M. Notices of Royal Ast. Soc.* **325**, 221–230.

Rasio, F.A. and Ford, E.B.: 1996, 'Dynamical Instabilities and the Formation of Extrasolar Planetary Systems', *Science* **274**, 954–956.

Rasio, F.A., Tout, C.A., Lubow, S.H., and Livio, M.: 1996, 'Tidal Decay of Close Planetary Orbits', *Astrophys. J.* **470**, 1187–1191.

Tabachnik, S. and Tremaine, S.: 2002, 'Maximum Likelihood Method for Estimating the Mass and Period Distributions of Extra-solar Planets', *M. Notices of Royal Ast. Soc.* **335**, 151–158.

Weidenschilling, S.J. and Marzari, F.: 1996, 'Gravitational Scattering as a Possible Origin for Giant Planets at Small Stellar Distances', *Nature* **384**, 619–621.

Wisdom, J. and Holman, M.: 1991, 'Symplectic Maps for the N-body Problem', *Astronom. J.* **102**, 1528–1538.

Zapatero Osorio, M. R., Bejar, V. J. S., Martin, E. L., Rebolo, R., Navascues, D. Barrado Y., Mundt, R., Eisloffel, J., and Caballero, J. A.: 2002 'A Methane, Isolated, Planetary-Mass Object in Orion' *Astrophys. J.* **578**, 536–542.

Disc Interactions Resonances and Orbital Relaxation in Extrasolar Planetary Systems

J. C. B. Papaloizou and R. P. Nelson
Astronomy Unit, Queen Mary, University of London, Mile End Rd., London, E14Ns.

C. Terquem
Institut d'Astrophysique de Paris, 98 bis Boulevard Arago, 75014 Paris, France.

Abstract. The discovery of extrasolar planets some of which are close to their central stars has stimulated work on orbital migration produced by disc protoplanet interactions. Such interactions may also lead to commensurabilities in many planet systems. The existence of high eccentricities indicates that strong gravitational interactions may also have occurred in them. We review some recent work on protoplanet disc interactions which incorporates MHD turbulence in the disc as well as work on resonances and strong gravitational relaxation in many planet systems.

1. Introduction

The recently discovered 100 or so extrasolar giant planets have masses in the range $0.12 - 11$ Jupiter masses. They may be found at distances of several AU or close to the central star with periods of a few days. High orbital eccentricities are common (eg., Mayor and Queloz, 1995; Marcy and Butler, 1998; Marcy et al., 2000).

It has been suggested that giant planets may form through gravitational instability in a disc at large radii or through the 'critical core mass' model in which a ~ 15 $M_\oplus$ core is formed in a disc by accumulation of solids and then undergoes rapid gas accretion (see eg. Papaloizou, Terquem and Nelson, 1999, for a review and appropriate references). In this case it is expected that the cores of gas giant planets should begin to form beyond a radius of $r \sim 4$ AU, the so–called 'ice condensation radius' where the existence of ices facilitates the accumulation of solids.

In order to explain the existence of the closely orbiting extrasolar giant planets one is then led to propose orbital migration. In principle this may occur through the gravitational interaction between a protoplanet and the protostellar disc or through mutual gravitational interactions among a strongly interacting system of protoplanets. High orbital eccentricities may also be produced. In this paper we review some recent work on these topics.

2. Protoplanet Disc Interactions

A promising mechanism for producing orbital migration is through protoplanet disc interaction. The basic idea is that a protoplanet orbiting in a gaseous disc

excites spiral density waves which have an associated angular momentum flux. Through their propagation away from the orbit and subsequent dissipation, the disc exterior to the protoplanet gains angular momentum from it while the interior disc loses angular momentum to it (Papaloizou and Lin, 1984). The exchange of angular momentum between disc and protoplanet leads to orbital migration. Type I migration occurs when the disc response is linear and the protoplanet is fully embedded in the disc. Type II migration occurs when the radius of the Hill sphere exceeds the disc scale height, h, making the interaction nonlinear and the angular momentum flux associated with waves exceeds the internal angular momentum flux arising from effective viscosity acting in the disc. Then a gap forms (eg. Bryden et al., 1999; Kley, 1999). This happens for commonly used protoplanetary disc models which have a typical effective dimensionless kinematic viscosity $\nu \sim 10^{-5}$, aspect ratio $h/r \sim 0.05$ and Jovian mass (M_J) planets. Then orbital migration occurs on essentially the disc evolution timescale induced by the action of the effective viscosity.

To study the interaction in detail, fully three dimensional simulations of the disc and protoplanet are needed. One should also attempt to model the physical processes leading to the effective disc viscosity that induces angular momentum transport and evolution of the disc mass distribution. The most robust and promising mechanism for providing disc viscosity is through turbulence arising from the Magnetorotational instability (MRI) (Balbus and Hawley, 1991). Although there is still an issue as to whether the ionization is adequate for retaining magnetic fields and, if produced externally by cosmic rays or X rays, whether ionization is adequate only in surface layers (Gammie, 1996), ideal MHD simulations are a reasonable first step.

However, the extensive computational requirements of such simulations are such that they are only beginning to be undertaken (see Papaloizou and Nelson, 2003; Nelson and Papaloizou, 2003; hereafter I and II). Up to now almost all simulations are for two dimensional flat laminar discs with anomalously large viscosity coefficient (eg. Bryden et al., 1999; Kley, 1999). Below we describe some aspects of some recent MHD simulations as well as some two dimensional laminar disc simulations which are still necessary for work involving runs over many orbits, multi-planet problems and extensive parameter space surveys.

Basic equations

The governing equations for MHD written in a frame rotating with a uniform angular velocity $\Omega_p\hat{\mathbf{k}}$ with $\hat{\mathbf{k}}$ being the unit vector in the vertical direction are:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \rho \mathbf{v} = 0, \tag{1}$$

$$\rho\left(\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v}\cdot\nabla\mathbf{v} + 2\Omega_p\hat{\mathbf{k}}\times\mathbf{v}\right) = -\nabla p - \rho\nabla\Phi + \frac{1}{4\pi}(\nabla\times\mathbf{B})\times\mathbf{B}, \tag{2}$$

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{v} \times \mathbf{B}). \tag{3}$$

where $\mathbf{v}$, P, ρ, $\mathbf{B} = (B_z, B_r, B_\phi)$ and Φ denote the fluid velocity, pressure, density, magnetic field, and gravitational potential, respectively.

The angular velocity Ω_p could be taken to be that of a planet in fixed circular orbit, Ω, at radius D which would then appear stationary. The potential Φ contains contributions due to the gravity of the central star, protoplanet and the centrifugal potential $-(1/2)\Omega_p^2 r^2$. Here cylindrical coordinates (z, r, ϕ) based on the centre of mass of the central star plus protoplanet system are adopted. The simulations adopt a locally isothermal equation of state

$$P(r) = \rho c(r)^2, \tag{4}$$

where $c(r)$ denotes the sound speed. In most of the disc, this is taken to be $\propto 1/r$ corresponding to a fixed aspect ratio $h/r = 0.1$ (for these and other details see papers I and II). Because of significantly enhanced computational requirements, vertical stratification is not included, or equivalently, z dependence in the gravitational potential is omitted.

When run on from fairly arbitrary initially specified fields with zero net flux, statistically steady turbulence with ratio of magnetic energy density to pressure ~ 0.01 results (see eg. Steinacker and Papaloizou, 2002) and the typical volume averaged value of the stress parameter $\alpha \simeq 5 \times 10^{-3}$.

Average turbulent stresses, angular momentum transport and external torques
For a global description of the flow, we use quantities that are vertically and azimuthally averaged over the (ϕ, z) domain (e.g. Hawley 2000) together with an additional time average. Thus for any quantity Q

$$\overline{Q(r,t)} = \frac{\int \rho Q(z, r, \phi, \tau) dz d\phi d\tau}{\int \rho dz d\phi d\tau}. \tag{5}$$

We take the time averaging interval or range of τ, which is centred on time t, to be $(t + \Delta, t - \Delta)$ and introduce a mean surface density

$$\Sigma = \frac{1}{4\pi\Delta} \int \rho(z, r, \phi, \tau) dz d\phi d\tau. \tag{6}$$

The averaged Maxwell and Reynolds stresses are respectively defined as follows:

$$T_M(r,t) = 4\pi\Delta\Sigma \overline{\left(\frac{B_r B_\phi}{4\pi\rho}\right)} \tag{7}$$

and

$$T_{Re}(r,t) = 4\pi\Delta\Sigma \overline{\delta v_r \delta v_\phi} \tag{8}$$

The velocity fluctuations δv_r and δv_ϕ are defined through,

$$\delta v_r = v_r - \overline{v_r}, \delta v_\phi = v_\phi - \overline{v_\phi}. \tag{9}$$

The standard Shakura and Sunyaev (1973) α parameter for the total stress is given by

$$\alpha(r,t) = \frac{T_{Re} - T_M}{4\pi \Delta \Sigma \overline{(P/\rho)}}. \tag{10}$$

One can derive an averaged continuity equation in the form (see Balbus and Papaloizou, 1999)

$$\frac{\partial \Sigma}{\partial t} + \frac{1}{r}\frac{\partial \left(r\Sigma \overline{v_r}\right)}{\partial r} = 0, \tag{11}$$

Similarly the averaged azimuthal component of the equation of motion, which expresses global conservation of angular momentum, can be written in the form

$$\frac{\partial \left(\Sigma \overline{j}\right)}{\partial t} + \frac{1}{r}\left(\frac{\partial \left(r\Sigma \overline{v_r}\,\overline{j}\right)}{\partial r} + \frac{\partial \left(\Sigma r^2 \alpha \overline{P/\rho}\right)}{\partial r}\right) = -\int \rho \frac{\partial \Psi}{\partial \phi}\frac{d\phi dz d\tau}{4\pi \Delta}. \tag{12}$$

Here $j = r v_\phi$ is the specific angular momentum and the term on the right hand side derives from the torque acting on the disc due to the internally orbiting protoplanet.

We comment that equations (11) and (12) are obtained identically in a two dimensional flat laminar disc model provided we identify the anomalous kinematic viscosity coefficient $\nu = (2/3)\alpha h^2 \Omega$.

However as indicated in I, such correspondences are meaningful for a turbulent disc only if long enough time averages of up to several orbital periods at the outer disc boundary are considered. On short timescales erratic fluctuations in mean radial velocity and α occur. When a protoplanet is inserted the correspondence remains for those parts of the disc that are relatively unperturbed by it.

A contour plot for the midplane disc density 82 protoplanet orbits after the insertion of a protoplanet of mass $5M_J$ into a fixed circular orbit of dimensionless radius $r = 2.2$ in a disc with $h/r = 0.1$ in which statistically steady turbulence had been set up is given in Fig. 1 (see II). This maintains a gap which deepens with time. The excited spiral waves are clearly visible.

Following equation (12), we consider the non advected part of the averaged angular momentum flux which is proportional to $\Sigma r^2 \alpha \overline{P/\rho}$. We plot these angular momentum fluxes (here the total flow through the circle of radius r) in Fig. 2. These exhibit a stable pattern of behaviour for subsequent times.

The behaviour may be understood from considering equation (12) under the assumption of a disc in a quasi steady state with $\overline{v}_r = 0$. For a completely empty gap containing the protoplanet, equation (12) predicts that the angular momentum flux tends to a constant value at large distances from the planet. The flux is zero in the gap where there is no material and is generated by the tidal torque term at the

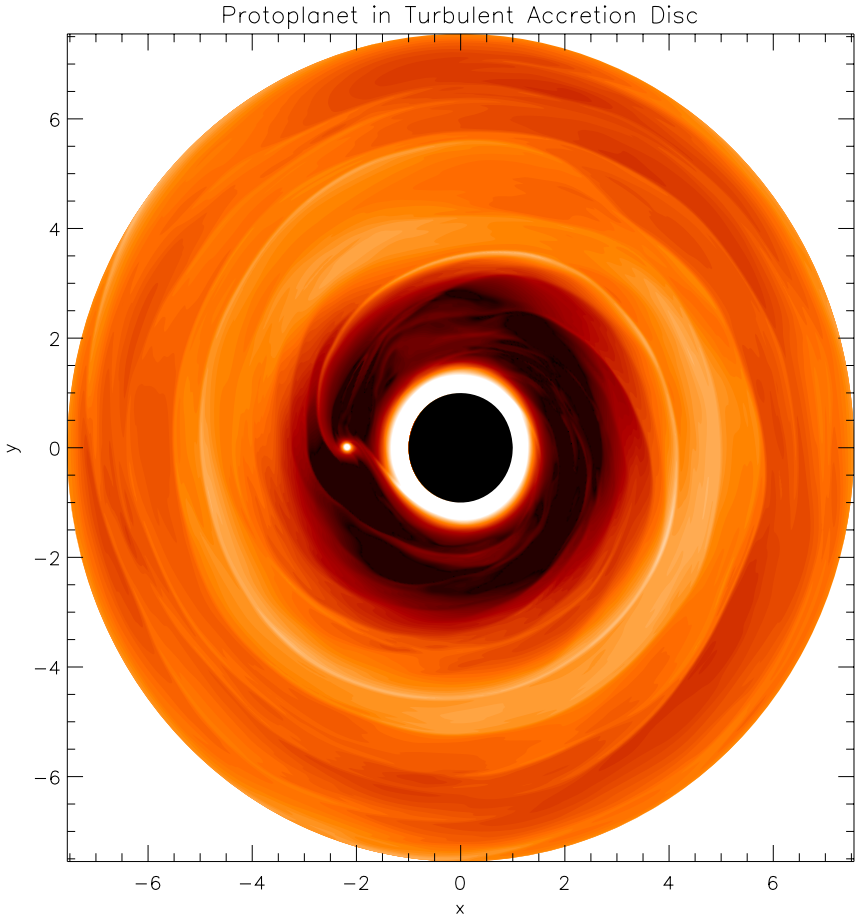

Figure 1. A contour plot for the midplane disc density at 82 protoplanet orbits after protoplanet insertion into the disc. A lighter shade corresponds to a higher density. The total density range between the lightest and darkest regions shown here is a factor of 140. Protolanet, gap and wakes are apparent. The underlying turbulence has the effect of slightly blurring the spiral wave pattern.

disc edge near the planet. In accordance with this we see that the magnetic stress rises from small values near the planet to almost the total unperturbed flux at large distance. Close to the protoplanet the Reynolds' stress rises due to the spiral wave excitation. This then decreases outwards as the waves are damped.

The plots in Fig. 2 indicates that the outward angular momentum flow in the outer disc is produced by the protoplanet as it maintains the gap. Accordingly it must migrate inwards on a timescale corresponding to the viscous evolution of the disc as obtained in 2D laminar disc simulations. In I and II it was explained that because the turbulence does not in fact behave as if it produced an anomalous viscosity, flow details in the neighbourhood of the protoplanet will differ. How-

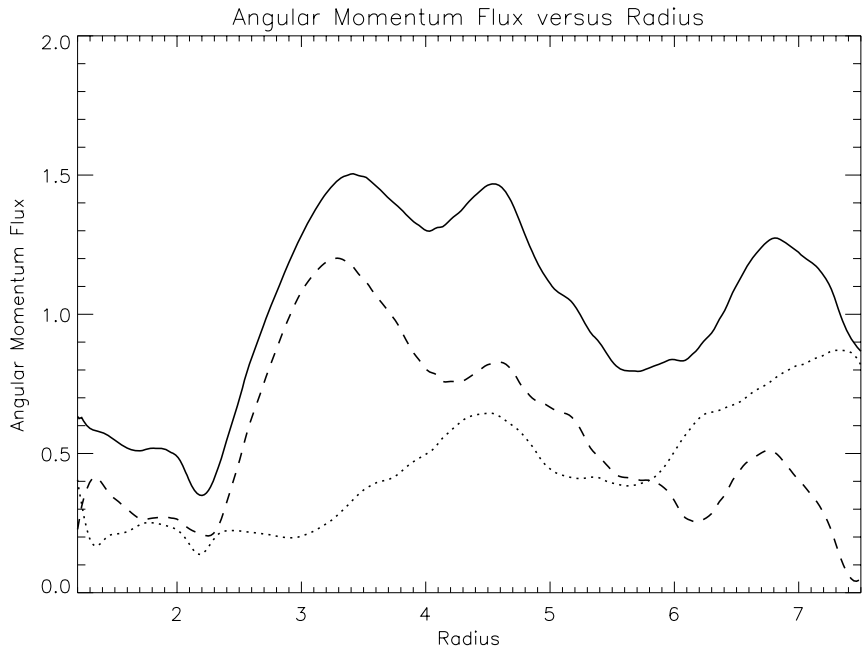

Figure 2. Averaged values of the non advected part of the angular momentum flux in arbitrary units plotted as a function of dimensionless radius. The solid line gives the total contribution from the magnetic plus Reynolds' stress. The dashed line gives the Reynolds' stress contribution and the dotted line gives the magnetic stress contribution. The average is taken between 80.5 and 94 protoplanet orbits after insertion into the disc.

ever, this is probably not very significant for cases with a strong gap and type II migration because of the dearth of material near the protoplanet.

A 2D simulation of the interaction of a Jupiter mass protoplanet inserted into a circular orbit in a laminar disc with $r = 1, h/r = 0.05$ and $\nu = 10^{-5}$ in dimensionless units taken from Nelson et al. (2000) is illustrated in Fig. 3. In this case the orbit was allowed to evolve with time. The disc surface density is plotted after the times measured in units of the orbital period at $r = 1$ indicated. A gap forms,while the inner disc is accreted (this is allowed to flow through the inner boundary). The predominance of the outer disc leads to inward migration on the viscous evolution time. This is typically 10^5 y if the planet starts at 5 AU. The planet may grow up to several Jupiter masses if efficient accretion takes place before reaching the central star. The fact that the migration time is an order of magnitude or so shorter than the disc lifetime indicates that some stopping mechanism may be needed. On the other hand the 6 percent incidence of extra solar planets around solar type stars allows a model in which many planets form continuously and are lost to the central star, with those that remain surviving because the disc disappears while they migrate (see Armitage et al., 2002; Trilling, Lunine and Benz, 2002). This model

also predicts an increase of numbers of giant planets with radius which is supported by observations for masses $> 1M_J$ (Zucker and Mazeh, 2002).

3. Multiple Systems

Currently, several multiple planet systems have been detected including three that show a mean motion resonance or commensurability. An example is the GL876 system which is observed to be in 2:1 resonance (Marcy et al., 2001). The bringing about of such a situation is expected in a multiple planet system in which orbital semi-major axes change through dissipative processes. Tw planets can evolve into a commensurability which is subsequently maintained while they migrate together. This process is known to operate for the Galilean satellites of Jupiter.

A generic situation one may consider is two protoplanets orbiting within an inner disc cavity. The outer protoplanet being closer to the disc migrates inwards as a result of interaction with it. It then approaches the inner protoplanet until a commensurability is reached. Subsequently the orbits lock and migrate together maintaining the commensurability. Nelson and Papaloizou (2002) considered a simple model of this process in which the effect of the disc was simply modelled by adding additional forces to the equation of motion of the outer planet to reduce its angular momentum and damp its eccentricity. They distinguish between four types of orbital evolution in the resonant migration phase. The orbital elements of the outer planet have subscript '1' and the inner planet subscript '2'. For a $p : q$ commensurability, the resonant angles are defined by

$$\phi_{p,q,k} = p\lambda_1 - q\lambda_2 - p\varpi_1 + q\varpi_2 + k(\varpi_1 - \varpi_2). \tag{13}$$

Here λ_1, λ_2, ϖ_1 and ϖ_2 denote the mean longitudes and longitudes of periapse for the planets 1 and 2 respectively. The positive integers p and q satisfy $p > q$, and there are $p - q + 1$ possible values of the positive integer k such that $q \leq k \leq p$. Although there are $p - q + 1$ corresponding angles, no more than two can be linearly independent. This means that if libration occurs, either all librate or only one librates. Depending on initial conditions 2:1, 3:1, 4:1, 5:1 commensurabilities could be generated.

Type A: Eccentricities increase during migration until they reach steady state equilibrium values, with possibly small oscillations superposed. After this migration continues in a self-similar manner and all the resonant angles defined above are in libration.

Type B: Large values of e_1 arise when the eccentricity damping rate is small. As in type A all the resonant angles considered go into libration. The difference is that e_1 exceeds 0.2 before a steady state can be reached. There is then a possibility that the outer planet reenters the disc.

Type C: For higher initial values of e_2, there are cases for which e_1 and e_2 rise continuously during the migration phase, even for efficient applied damping of e_1.

In this case only one of the resonant angles defined above is found to librate so that the evolution differs from A and B.

Type D: This mode of resonant migration corresponds to small values of $e_1 < 0.2$ being maintained for the outer planet, whilst the eccentricity of the inner planet grows to attain values close to unity. All the resonant angles show large amplitude libration.

A number of different situations may arise including generation of high eccentricities and potential orbital instability. However, reentry into and proper interaction with the disc needs to be considered in order to assess likely outcomes. This is computationally demanding and likely to take some time to complete.

Conservation of Energy and Angular Momentum

For type A migration, a steady state for the orbital eccentricities is attained while the system migrates self-similarly. A simple relationship between e_1, assumed not too large, and the circularization and migration rates induced by the disc on the outer planet may be obtained.

Consider the two planets with masses m_1, m_2 and osculating semi-major axes a_1, a_2, orbiting a central mass M_*. The total angular momentum is

$$J = J_1 + J_2 = m_1\sqrt{GM_*a_1(1-e_1^2)} + m_2\sqrt{GM_*a_2(1-e_2^2)} \tag{14}$$

and the energy E is

$$E = -\frac{GM_*m_1}{2a_1} - \frac{GM_*m_2}{2a_2} \tag{15}$$

For resonant self-similar migration in which a_2/a_1, e_1, and e_2 are constant, conservation of angular momentum gives

$$\frac{dJ}{dt} = J_1\frac{1}{2a_1}\frac{da_1}{dt}\left(1 + \frac{m_2\sqrt{a_2(1-e_2^2)}}{m_1\sqrt{a_1(1-e_1^2)}}\right) = -T \tag{16}$$

and conservation of energy gives

$$\frac{dE}{dt} = \frac{GM_*m_1}{2a_1^2}\frac{da_1}{dt}\left(1 + \frac{m_2a_1}{m_1a_2}\right) = -\frac{n_1T}{\sqrt{1-e_1^2}} - D, \tag{17}$$

where $n_1 = \sqrt{GM_*/a_1^3}$ and we suppose there is a tidal torque $-T$ produced by the disc which acts on m_1. In addition we suppose there to be an associated tidally induced orbital energy loss rate which is written as $n_1T/\sqrt{1-e_1^2} + D$, with $D \equiv (GM_*m_1e_1^2)/(a_1(1-e_1^2)t_c)$. Here t_c is the circularization time of m_1 that would apply if the tidal torque and energy loss rate acted on the orbit of m_1 with m_2 being

absent. In that case $de_1/dt = -e_1/t_c$. A migration time $t_{\rm mig}$ can be defined through $T = m_1\sqrt{GM_*a_1(1-e_1^2)}/(3t_{\rm mig})$. This is the time for n_1 to exponentiate if m_2 was absent and the eccentricity e_1 was fixed. Note that t_c and $t_{\rm mig}$ are determined by the disc planet tidal interaction and may depend on e_1.
By eliminating $\frac{da_1}{dt}$ from (17) and (16) one obtains

$$e_1^2 = \frac{t_c\left(1 - e_1^2 - \frac{(1-e_2^2)^{1/2}(1-e_1^2)^{1/2}}{a_2^{-3/2}a_1^{3/2}}\right)}{3t_{\rm mig}\left(\frac{m_1a_2}{m_2a_1} + \frac{a_2^{3/2}(1-e_2^2)^{1/2}}{a_1^{3/2}(1-e_1^2)^{1/2}}\right)} \tag{18}$$

For a 2:1 commensurability $(a_1/a_2)^{3/2} = 2$ and (18) reduces in the case $e_i^2 \ll 1$ to

$$e_1^2 = \frac{t_c m_2 a_1}{3t_{\rm mig}(2m_1a_2 + m_2a_1)} \tag{19}$$

The above determines the eccentricity of the outer planet e_1 as a function of t_c and $t_{\rm mig}$.

Resonant Planets in GL876
Snellgrove, Papaloizou and Nelson (2001) performed 2D laminar disc simulations of the resonant coupling in the GJ876 system induced by tidally induced migration of the outer planet. The disc model is locally isothermal with aspect ratio $h/r = 0.07$, and a kinematic viscosity $\nu = 1 \times 10^{-5}$ in dimensionless units. The two planets start in circular orbits with semi-major axes in the ratio 0.6. Minimum planet masses are assumed (orbital plane inclination to the plane perpendicular to the line of sight $i = 90^\circ$). These are assumed to be no longer accreting from the disc. To minimize computational requirements, it is assumed that both planets start within a tidally truncated cavity with surface density one percent of that in the outer disc.

Inward migration of the outer planet resulted in locking into the 2:1 commensurability such that both planets migrated together maintaining it while the eccentricities grew (see Fig. 4). The resonant angles (with subscripts dropped) $\phi = 2\lambda_1 - \lambda_2 - \varpi_1$, and $\psi = 2\lambda_1 - \lambda_2 - \varpi_2$, go into libration about zero.

A typical contour plot of the surface density distribution is shown in Fig. 5. At the end of the simulation the eccentricities are comparable to the observed ones. But it is difficult to conclude that the migration corresponds to type A given the limited feasible dynamic range. Also, the final situation depends on the disc model parameters and the nature of the initial set up. The exact quantity of material that leaks into the inner cavity may be significant. This may be responsible for limiting the eccentricities and possibly stalling the inward migration. These issues as well as the effects of boundary conditions and the manner of final disappearance of the disc are currently being investigated.

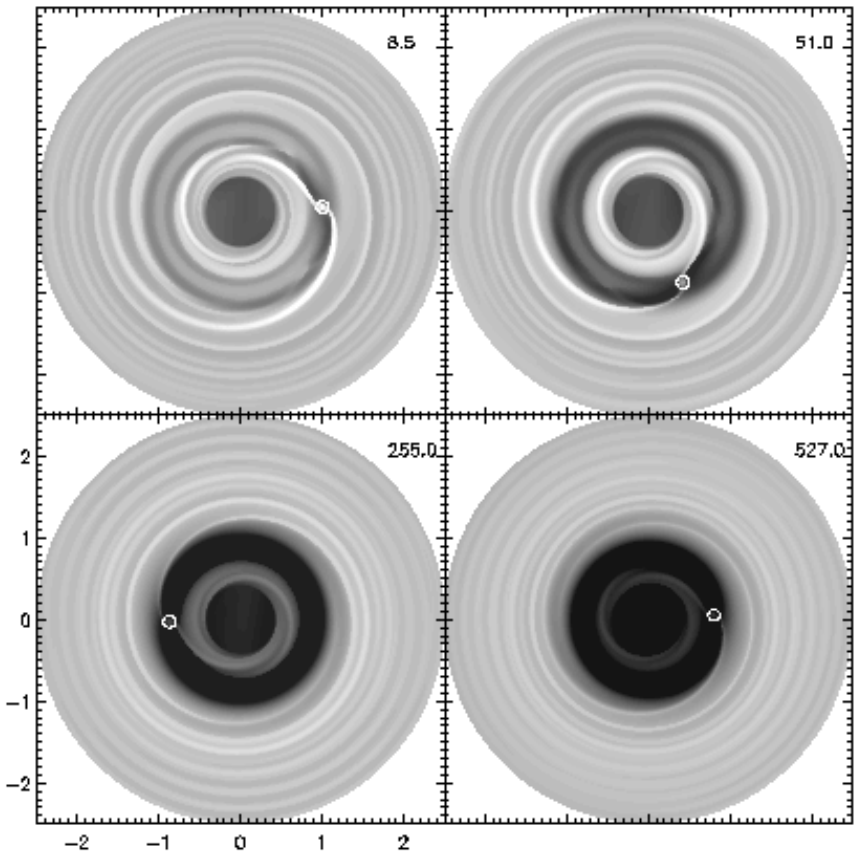

Figure 3. This figure shows the evolution of a protoplanet embedded in a protostellar disc. The relative surface density of disc material is represented by the grey-scale with lighter shades corresponding to larger values. The range of values corresponding to the lightest and darkest regions represented here is a factor of 4000. The white circle represents the position of the protoplanet. The planet is inserted into the disc at time $t = 0$.

4. Apsidal Resonance

This type of resonance is exhibited by the two outer planets in Upsilon Andomeda and possibly by the two planets in 47 UMa (Laughlin, Chambers and Fischer, 2002). In it the apsidal lines are locked together. This may occur independently of orbital separation and so no orbital migration is needed in principle to bring it about. The pair of planets can be considered to be in a single secular normal mode (m=1) in which the eccentricities take on appropriate values to match apsidal precession rates.

When a disc is present that may also participate in a joint normal mode. The action of non conservative processes may then favour a particular mode so leading

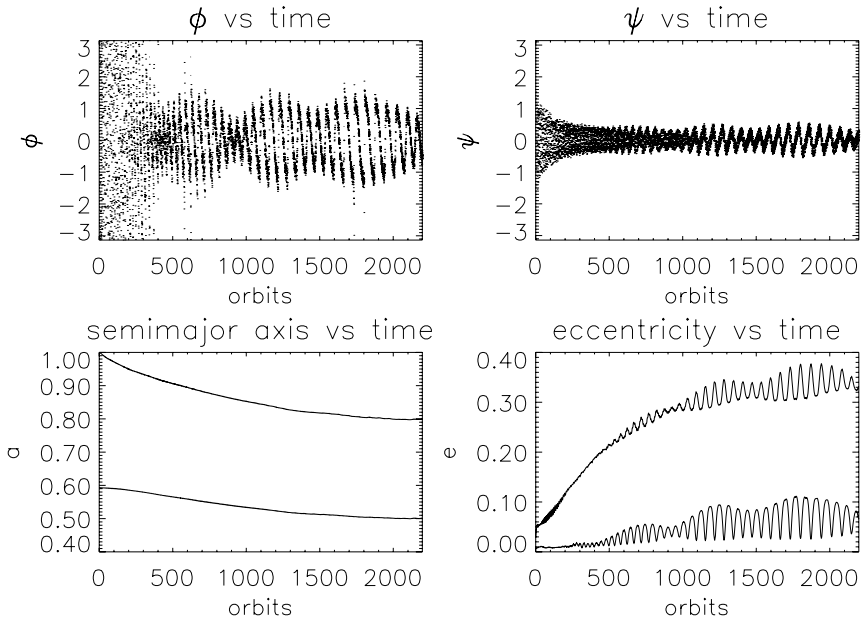

Figure 4. The semi major axes a (lower left) and eccentricities e (lower right) for both planets in GL876. The upper plots show the resonance angles ϕ and ψ. The time unit is measured in orbit periods of the disc material at the initial location of the outer planet.

to an explanation of the origin of the configuration. Such joint normal modes have been considered by Papaloizou (2002) in the linear regime.

Basic Equations

The basic equations (1) and (2) adapted to a 2D non magnetic flat disc may be used. The adaption is made by replacing ρ by the surface density Σ and the pressure by $\Pi = \int_{-\infty}^{\infty} P dz$ the vertically integrated value. In addition $\mathbf{B} = 0$, $\Omega_p = 0$ and vertical dependence is suppressed. The sound speed is then given by $c = \sqrt{d\Pi/d\Sigma}$.

Linearization

Here we are interested in global $m = 1$ modes with slowly varying pattern. To study these, we linearize the basic equations about a circular equilibrium state denoting perturbations to quantities with a prime.

As usual we assume that the dependence of all perturbations on ϕ and t is through a factor $\exp i\,(\phi - \sigma t)$. For the slowly varying modes we consider $|\sigma| << \Omega$.

The gravitational potential perturbation induced by the disc is

$$\Phi'_D = -G \int \Sigma' K_1(r, r') r' dr', \tag{20}$$

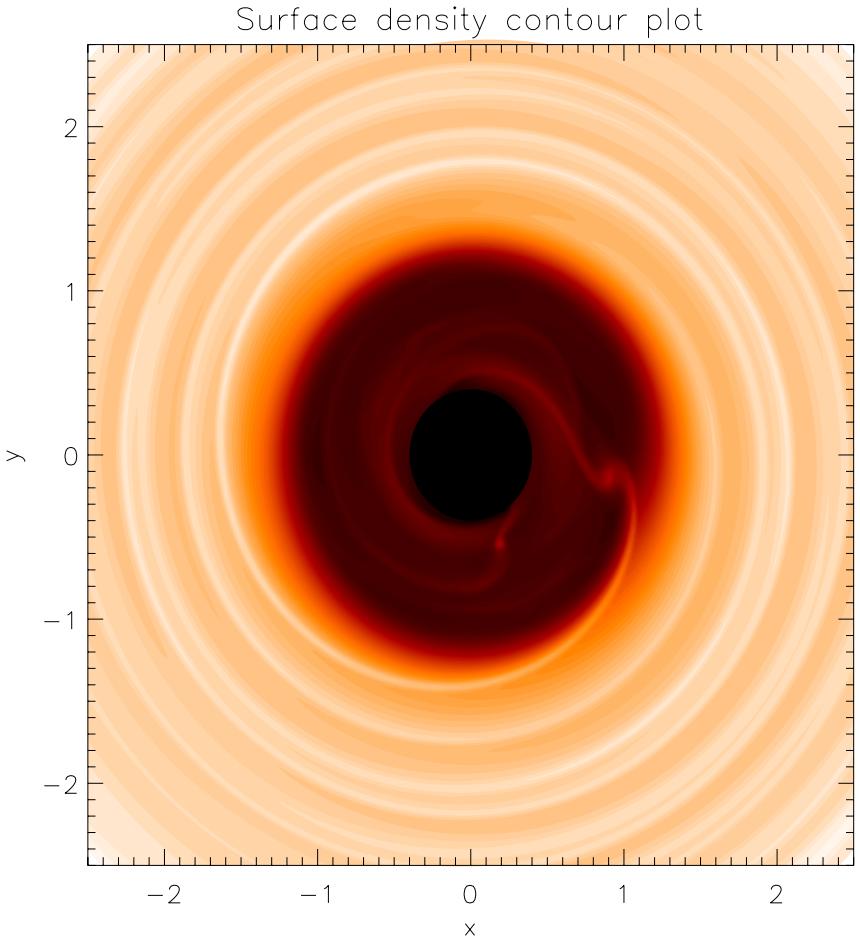

Figure 5. The surface density for the inner disc after 300 time units. Darker areas correspond to lower surface densities. The range of values represented corresponds to a factor 10000. The disc cavity containing the wakes of the two planets and the exterior density waves can be seen.

with

$$K_1(r,r') = \int_0^{2\pi} \frac{\cos(\varphi)}{\sqrt{(r^2 + r'^2 - 2rr'\cos(\varphi))}} d\varphi - \frac{\pi r}{r'^2}, \tag{21}$$

where the last term corresponds to the well known indirect term which for this problem turns out to be ignorable. The surface density perturbation is given, in the low pattern speed limit, in terms of the disc eccentricity $e(r)$ by

$$\Sigma' = -r\frac{d(\Sigma e(r))}{dr}. \tag{22}$$

The existence of a mode with $m = 1$ in the disc causes the orbits of any protoplanets interior to the disc to become eccentric. The secular perturbing potential giving the contribution of such an orbiting protoplanet is

$$\Phi'_p(r) = -\frac{Gm_p e_p}{2\pi r_p}\frac{\partial}{\partial r_p}\left(K_1(r, r_p)r_p^2\right). \tag{23}$$

Here the eccentricity of the protoplanet is e_p and its mass is m_p. The total external potential perturbation is then found by summing over the perturbing protoplanets:

$$\Phi'_{\rm ext}(r) = \sum_p \Phi'_p(r) \tag{24}$$

The total potential perturbation is

$$\Phi' = \Phi'_{\rm ext} + \Phi'_D \tag{25}$$

One obtains a normal mode equation relating $e(r)$, e_p and σ in the form (see Papaloizou, 2002)

$$2\left(\sigma - \omega_p\right)\Omega r^3 e(r) = \frac{d}{dr}\left(\frac{r^3c^2}{\Sigma}\frac{d[\Sigma e(r)]}{dr}\right) - \frac{d\left(r^2\Phi'\right)}{dr}. \tag{26}$$

Here the precession frequency $\omega_p = \Omega - \kappa$, with the square of the epicyclic frequency $\kappa^2 = (2\Omega/r)(d(r^2\Omega)/dr)$.

In addition we have an equation for each protoplanet of the form

$$2\left(\sigma - \omega_p\right)\Omega(r_p)r_p^3 e_p = -\left[\frac{d\left(r^2\Phi'\right)}{dr}\right]_{r=r_p}, \tag{27}$$

where of course for a particular protoplanet there is no self-interaction term in the sum for $\Phi'_{\rm ext}$. The inclusion of self-gravity in the eigenvalue problem is essential if modes with prograde precession frequency are to be obtained. For typical protoplanetary disc models, self-gravity can be strong enough to induce prograde precession for the long wavelength $m = 1$ modes considered here.

An interesting property is the orthogonality of modes with different eigenfrequency σ. Related to this is the existence of an adiabatic invariant for each joint normal mode related to the radial action in the form

$$I = \int (\Omega r^2 e^2) dm, \tag{28}$$

where the integral is one over the whole mass distribution of disc and protoplanets.

This feature may be relevant to the origin of apsidal resonances as it indicates that a mode will survive under slow external changes to the system such as outward expansion of the disc with a consequent change to the magnitude of the protoplanet eccentricities.

5. Orbital Relaxation in Young Planetary Systems

The presence of high orbital eccentricities amongst extrasolar planets is suggestive of a strong orbital relaxation or scattering process. For this to occur formation must occur on a timescale short enough that strong dynamical interactions may occur subsequently. The gaseous environment of a disc may act to inhibit such interactions until it is removed. Gas free dynamical interactions of coplanar protoplanets formed on neighbouring circular orbits have been considered by Weidenschilling and Mazari (1996) and Rasio and Ford (1996). These may produce close scatterings and high eccentricities but the observed distribution of extra solar planets is not reproduced.

Papaloizou and Terquem (2001) investigate a scenario in which $5 \leq N \leq 100$ planetary objects in the range of several Jupiter masses are assumed to form rapidly through fragmentation or gravitational instability occurring in a disc or protostellar envelope on a scale of $R_{\rm max} = 100$ AU. If these objects are put down in circular orbits about a solar mass star, at random in a volume contained within a spherical shell with inner and outer radii of $0.1R_{\rm max}$ and $R_{\rm max}$ respectively, a strong relaxation on a timescale $\sim$ 100 orbits occurs leading to independence of details of initial conditions.

For a range of $100 > N > 5$, the evolution is similar to that of a star cluster. Most objects escape leaving at most 3 bound planets, the innermost with semi–major axis in the range $0.1R_{\rm max} - 0.01R_{\rm max}$. However, close encounters or collisions with the central star occurred for about 10% of cases. Tidal interaction giving orbital circularization at fixed pericentre distance leading to the formation of a very closely orbiting giant planet is then a possibility.

An example of a run with $N = 8$ masses selected uniformly at random in the interval $(0, 5 \times 10^{-3} M_{\odot})$ and central stellar radius $R_* = 1.337 \times 10^{-4} R_{\rm max}$ is illustrated in Fig. 6. At the end of this run only 2 planets remain bound to the central star.

The relaxation processes discussed above are more likely to apply to the more massive extrasolar planets exceeding 1 Jupiter mass M_J. Their observed number increases with distance (Zucker and Mazeh, 2002) as found in the simulations. There are 6 isolated candidate objects with mass $M > 4.5 M_J / \sin i$, around HD 190228, HD 222582, HD 10697, 70 Vir, HD 89744 and HD 114762 with semi–major axes in the range $0.3 - 2.5$ AU and eccentricities $0.12 \leq e \leq 0.71$. Amongst the 'hot Jupiters' detected so far, τ Boo is a possible candidate being unusually massive with $M \sin i \sim 4 M_J$.

Effect on a low mass planet formed in an inner disc

Terquem and Papaloizou (2002) have considered the effects of an outer relaxing distribution of protoplanets on an inner disc in which one inner planet with a mass $m = 0.3$ $M_{\rm J}$ is formed on a timescale of 10^6 y. During the formation process,

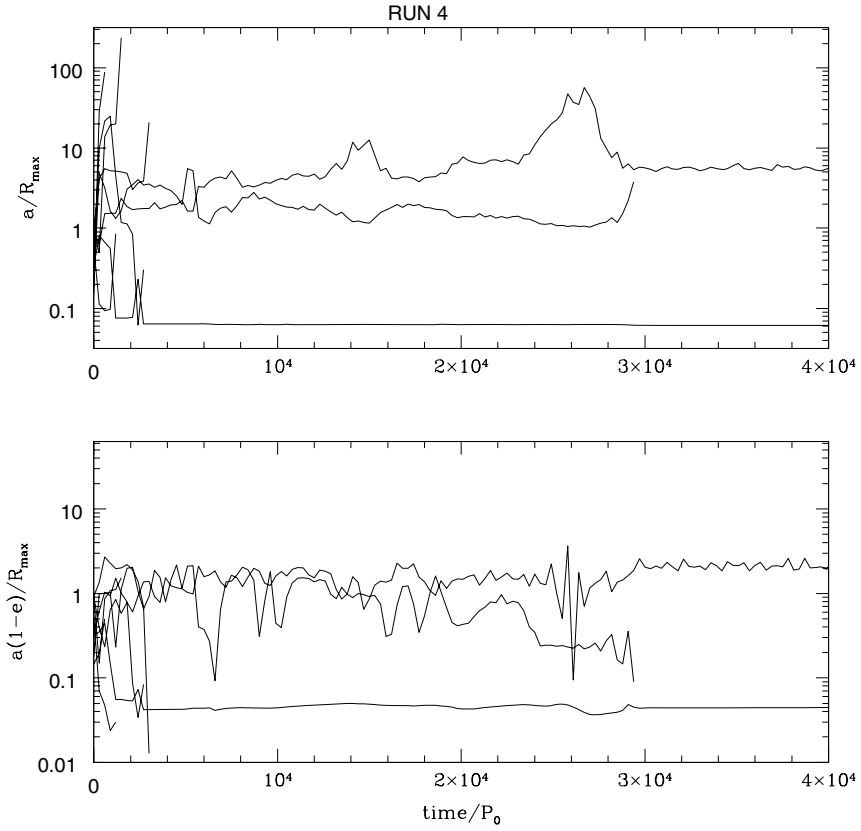

Figure 6. This figure shows the evolution of the semi-major axes (*upper plot*) and pericentre distances (*lower plot*) of the N = 8 planets in the system versus time (measured in units of P_0, being the orbital period at $R_{\max}$). The lines correspond to the different planets. A line terminates just prior to the escape of a planet.

eccentricity is assumed to be damped by tidal interaction with the disc while the planet is built up progressively with orbital radius a in the range 0.3–10 AU.

An example of a run with $N = 9$ outer planets of mass $8M_{\rm J}$, and $R_{\max} = 100$ AU is illustrated in Fig. 7. After $t \sim 10^6$ years, one outer planet with $a \sim 8$ AU and $e \sim 0.6$ remains. The inner planet enters into a cycle in which e varies between 0.1 and 0.24. The mutual inclination oscillates between 0 and 30°. More extreme cycles have been produced in other examples.

Thus an outcome of an outer relaxing system could be an inner lower mass protoplanet with high orbital eccentricity. Candidates are the planets around HD 16141, HD 83443 and HD 108147. In this case there should be an outer massive planet with high eccentricity. Among the candidate systems with large radial velocity drifts potentially due to a companion selected by Fischer et al. (2001) was HD 38529 which has a planet with $m \sin i = 0.76$ M$_{\rm J}$, $a = 0.13$ AU and $e = 0.27$. This is similar to the system illustrated in Fig. 7. Recently the discovery of a companion

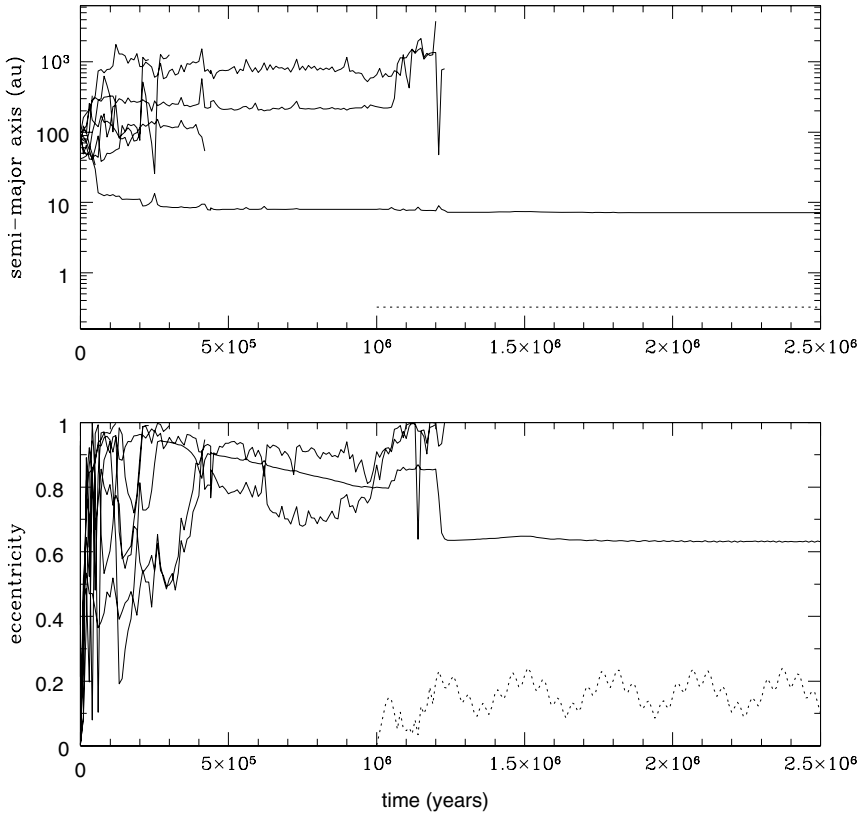

Figure 7. Semi-major axes (in AU, *upper plot*) and eccentricities (*lower plot*) of the $N = 9$ outer planets (*solid lines*) and low mass planet (*dotted lines*) versus time. Here, for the innermost planet, $a = 0.3$ AU initially. The eccentricity of the low mass planet varies between 0.1 and 0.24 while a is almost constant.

with $M \sin i = 12.7\ \mathrm{M_J}$, $a = 3.7$ AU and $e = 0.36$ has been announced (Fischer et al., 2002). If $\sin i = 0.8$ is adopted for this system a significant eccentricity is not excited in an initially circular inner planet orbit if the system is assumed coplanar. But, as illustrated in Fig. 8, for a mutual orbital inclination of 60°, a cycle is found in which the inner eccentricity reaches its observed value.

6. Discussion

Simulations of protoplanet disc interactions incorporating internally generated turbulence rather than anomalous viscosity prescriptions have begun. Disc interactions provide a promising mechanism for producing orbital migration as well as commensurable pairs and apsidal resonances. Strong gravitational interactions amongst a population of distant giant planets formed early in the life of the protostellar disc through gravitational fragmentation may also produce close orbiters and

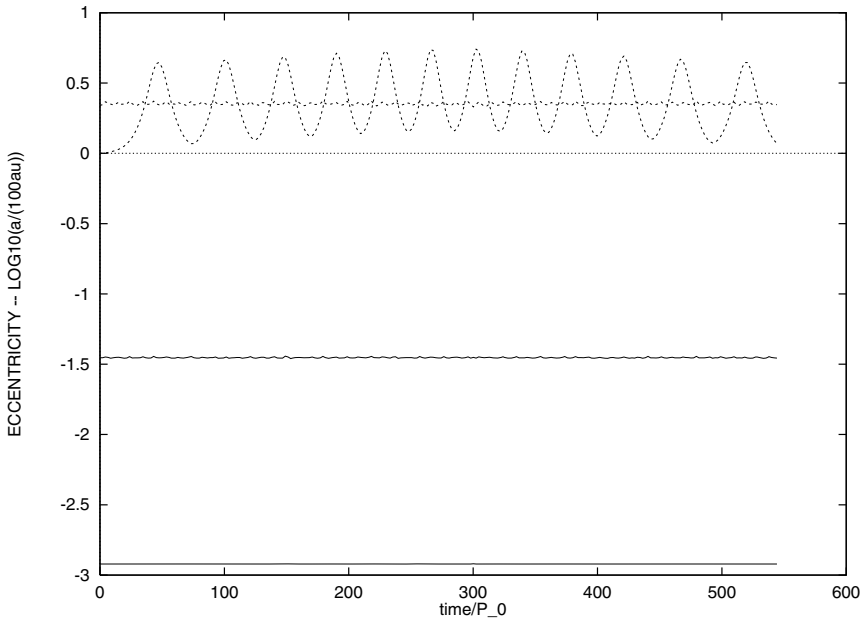

Figure 8. Motion in HD35829 for mutual inclinations of 0 and 60°. The unchanging $\log_{10}(a/R_{\max})$ for the two planets are given by the lower curves. The eccentricities by the upper curves. The oscillatory curve indicates a Kozai like cycle for the inner planet in the high inclination case. In the low inclination case the eccentricity of the inner planet remains close to zero.

planets on highly eccentric orbits. The latter may belong to the original distribution or be formed in an inner disc.

References

Armitage, P.J., Livio, M., Lubow, S. H., Pringle, J. E.: 2002, *MNRAS* **334**, 248.
Balbus, S. A., Hawley, J. F.: 1991, *ApJ* **376**, 214.
Balbus, S. A., Papaloizou, J.C.B.: 1999, *ApJ* **521**, 650.
Bryden, G., Chen, X., Lin, D.N.C., Nelson, R.P., Papaloizou, J.C.B.: 1999, *Astrophys. J.* **514**, 344.
Fischer, D.A., Marcy, G.W., Butler, R.P., Vogt, S.S., Frink, S., Apps, K.: 2001, *ApJ* **551**, 1107.
Fischer, D.A., Marcy, G.W., Butler, R.P., Vogt, S.S., Henry, G.W., Pourbaix, D., Walp, B., Misch, A.A., and Wright, J.T.: 2003, *Astrophys. J.* **586**, 1394.
Gammie, C.F.: 1996, *Astrophys. J.* **462**, 725.
Hawley, J. F.: 2000, *ApJ* **528**, 462.
Kley, W.: 1999, *MNRAS* **303**, 696.
Laughlin, G., Chambers, J., Fischer, D.: 2002, *Astrophys. J.* **579**, 455.
Marcy, G.W., Butler, R.P.: 1998, *ARA&A,* **36**, 57.
Marcy, G.W., Butler, R.P.: 2000, *PASP* **112**, 137.
Marcy, G.W., Butler, R.P., Fischer, D., Vogt, S., Lissauer, J., Rivera, E.: 2001, *Astrophys. J.* **556**, 296.
Mayor, M., Queloz, D.: 1995, *Nature* **378**, 355.
Nelson,R.P., Papaloizou, J.C.B., Masset, F.S., Kley, W.: 2000, *MNRAS* **318**, 18.

Nelson,R.P., Papaloizou, J.C.B.: 2002, *MNRAS* **333**, L26.
Nelson, R.P. and Papaloizou, J.C.B.: 2003 (II), *Monthly Notices of the Royal Astronomical Society* **339**, 993.
Papaloizou, J.C.B, Lin, D.N.C.: 1984, *ApJ* **285**, 818.
Papaloizou, J.C.B. and Nelson, R.P.: 2003 (I), *Monthly Notices of the Royal Astronomical Society* **339**, 983.
Papaloizou, J.C.B., Terquem, C., Nelson, R.P.: 1999, ASP Conference series, 160, 186, edited by J.A. Sellwood and J. Goodman
Papaloizou, J.C.B., Terquem, C.: 2001, *MNRAS* **325**, 221.
Papaloizou, J.C.B.: 2002, *A&A* **338**, 615.
Rasio, F.A., Ford E. B.: 1996, *Science* **274**, 954.
Shakura, N. I., Sunyaev, R. A.: 1973, *A&A* **24**, 337.
Snellgrove, M.D., Papaloizou, J.C.B., Nelson, R.P.: 2001, *A&A* **374**, 1092.
Steinacker, A., Papaloizoi, J.C.B.: 2002, *ApJ* **571**, 413.
Terquem, C., Papaloizou, J.C.B.: 2002, *MNRAS* **332**, L39.
Trilling, D. E., Lunine, J. I., Benz, W.: 2002, *A&A* **394**, 241.
Weidenschilling S.J., Mazari F.: 1996, *Nature* **384**, 619.
Zucker, S. Mazeh, T.: 2002, *Astrophys. J.* **568**, L113.

Part IV

Planets in Systems

Properties of Planets in Binaries

S. Udry, A. Eggenberger and M. Mayor
Geneva Observatory, ch. des Maillettes 51, CH-1290 Sauverny, Switzerland

Abstract. Among the more than 115 extra-solar planets discovered to date, 19 are orbiting a component of a binary system. We discuss the properties of this subsample and compare them with the equivalent characteristics of planets around single stars. Differences in the mass-period-eccentricity distributions are observed: exoplanets with $m_2 \sin i > 2M_{\rm Jup}$ and $P \leq 40 - 100$ days are in binaries and present low eccentricities. In the context of the migration scenario, these characteristics could tentatively be explained in the light of recent simulations of planet-disk interactions showing increased accretion and migration rates of planets in case an additional perturbing close stellar companion is present in the system. Finally, different observational approaches to find planets in stellar binaries, aiming at improving the still poor available statistics, are presented: a search for short-period planets in spectroscopic binaries, and an adaptive optics search for faint companions to stars with planets and without planets for comparison.

1. Introduction

The orbital-parameter and mass distributions of extra-solar planets carry information on the formation and evolution mechanisms of these systems. The observed orbital characteristics of extra-solar giant planets have forced considerable modifications of the standard formation model (Pollack et al., 1996). It is generally believed that planets form within a protoplanetary disk of gas and dust orbiting a central star, but the precise modes by which this formation takes place are still debated, especially for giant planets (e.g. Lin et al., 1996; Boss, 1997; Boss, 2000; Bodenheimer et al., 2000; Wuchterl et al., 2000, or several contributions in this volume). Two major model categories – core accretion or disk instability – have been proposed to explain giant planet formation, each with its advantages and limitations, but there is currently no consistent model that accounts for all the observed characteristics. In addition, in the late stage of their formation, the protoplanets continue to evolve due to their interaction with the disk or with other companions (stars or planets) present in the system. This subsequent evolution can alter significantly the initial orbital parameters of the planets, erasing in this way some signatures of the early stages of planet formation.

Searches for extra-solar planets using the radial-velocity technique or for faint companions to planet-hosting stars with adaptive optics have shown that giant planets do exist in multiple stellar systems. As the majority of stars belong to double or multiple systems, it is of importance to consider the existence and the properties of planets in this kind of environment. The number of such planets is, however, still low, mainly because close binaries are difficult targets for planet radial-velocity surveys and were consequently often rejected from the samples.

The present list of known planets in multiple stellar systems is given in the next section. Emerging statistical trends in their properties are then emphasized. In particular short-period, massive planets are preferentially found in binary systems (Zucker and Mazeh, 2002; Eggenberger et al., 2004a). They moreover move on low-eccentricity orbits. In order to discuss these properties in the light of our theoretical knowledge in the field, models of formation and evolution of giant planets in binaries are also briefly reviewed. Their predictions are then compared to the observations. In the next section, our on-going programmes searching for planets in different types of multiple stellar systems are reviewed, and conclusions are drawn in the last section.

2. Known Exoplanet Candidates in Multiple Systems

Among the more than 115 extra-solar planets discovered to date, 19 are orbiting a component of a multiple stellar system. Planets have been found around stars known to be part of a wide common proper motion pair (CPM), known to be in a visual binary (some with tentative orbits, VO) or in a "closer" spectroscopic binary (SB). Alternatively, searches for faint companions to stars hosting planets have revealed a few new systems. Table I summarizes this information providing interesting planetary and binary orbital properties as well. Note that there is no candidate in the 25 to 100 AU separation range for the binary separation ($a_{\rm bin}$). This interval corresponds to angular separations $\lesssim 1''$ for stars within 100 pc from the Sun, problematic separations for radial-velocity measurements.

In the same way as planets orbiting single stars, planets in binaries present a large variety of characteristics. At this point, it is not clear whether very distant companions do influence much planet formation and evolution.

3. Statistical Properties of Exoplanets in Binaries

Though the sample of planets found in multiple stellar systems is not large, a preliminary comparison between their characteristics and the properties of planets orbiting isolated stars can be made. Here, we will discuss the mass-period and the eccentricity-period diagrams for extra-solar planets, focusing on the possible differences between the two populations.

Our global sample of planets is made of the candidates with minimum masses $M_2 \sin i \ \leq \ 18\, M_{\rm J}$[1]. The candidates orbiting a component of a multiple stellar system are listed in Table I.

[1] see e.g. `http://obswww.unige.ch/Exoplanets/`

Table I. List of planets orbiting a star member of a multiple system with known orbital solution or confirmed common proper motion. CPM stands for common proper motion, SB for spectroscopic binary and VO for visual orbit. Interesting orbital and planetary properties are given as well.

Star	a_{bin} [AU]	P_p [days]	a_p [AU]	$M_p \sin i$ [M_J]	e_p	Notes	References
HD 40979	∼6400	267.2	0.83	3.28	0.25	CPM	Halbwachs (1986), Fischer et al. (2003)
Gl 777 A	∼3000	3902	4.8	1.32	0.48	CPM	Allen et al. (2000), Naef et al. (2003)
HD 80606	∼1200	111.8	0.469	3.90	0.927	CPM	Naef et al. (2001)
55 Cnc	∼1065	14.65	0.115	0.84	0.02	CPM	Duquennoy and Mayor (1991), McGrath et al. (2002), Marcy et al. (2002), Butler et al. (1997)
		44.28	0.241	0.21	0.34		
		5360	5.9	4.05	0.16		
16 Cyg B	∼850	798.5	1.66	1.64	0.68	CPM	Patience et al. (2002), Cochran et al. (1997), Hauser and Marcy (1999)
Ups And	∼750	4.617	0.059	0.69	0.02	CPM	Lowrance et al. (2002), Patience et al. (2002), Butler et al. (1997), Butler et al. (1999)
		241.3	0.829	1.89	0.24		
		1308	2.53	3.75	0.31		
HD 178911 B	∼640	71.49	0.32	6.292	0.124	CPM	Tokovinin et al. (2000), Zucker et al. (2002)
HD 219542 B	∼288	111.0	0.46	0.30	0.32	CPM	Desidera et al. (2003)
Tau Boo	∼240	3.313	0.05	4.08	0.02	VO	Hale (1994), Patience et al. (2002), Butler et al. (1997)
HD 195019	∼150	18.20	0.14	3.50	0.01	CPM	Patience et al. (2002), Allen et al. (2000), Fischer et al. (1999)
HD 114762	∼130	83.89	0.35	11.03	0.34	CPM	Patience et al. (2002), Latham et al. (1989), Marcy et al. (1999)
HD 19994	∼100	535.7	1.3	2.0	0.3	VO	Hale (1994), Queloz et al. (2004), Mayor et al. (2003)
HD 41004 A	∼23	650.0	1.33	2.5	0.25	SB	Udry et al. (2004), Zucker et al. (2003b), Santos et al. (2002)
γ Cep	∼22	903.0	2.03	1.59	0.2	SB	Campbell et al. (1988), Cochran et al. (2002), Hatzes et al. (2003)
Gl 86	∼20	15.78	0.11	4.0	0.046	SB,CPM	Els et al. (2001), Queloz et al. (2000)

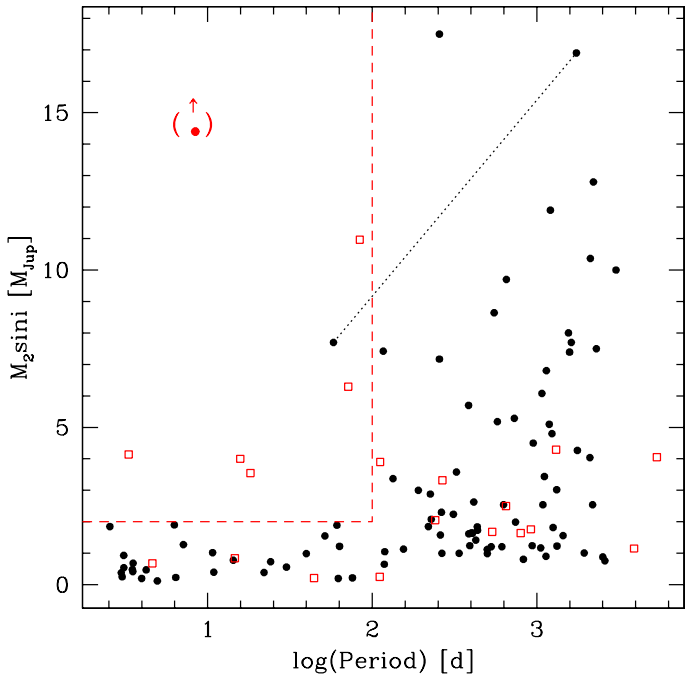

Figure 1. Minimum masses versus periods for the known exoplanet candidates. Squares indicate planets in binaries whereas circles are used for planets around single stars. The components of the possible multi brown-dwarf system HD 168443 (Udry et al., 2002; Marcy et al., 2001) are linked by a dotted line. The point in "()" in the upper-left corner represents HD 162020 for which synchronisation arguments indicate a mass in the brown-dwarf regime (Udry et al., 2002). The dashed lines in the diagram indicate limits at $P = 100$ d (vertical), and at $M_2 \sin i = 2\,M_{\rm Jup}$ (horizontal).

3.1. The Mass-Period Diagram

In Fig. 1 we display the distribution of the extra-solar planetary candidates in the $(M_2 \sin i - \log P)$ plane. Two interesting features emerge from this plot: there are no short-period extra-solar planets with a mass $M_2 \sin i \gtrsim 5\,M_{\rm Jup}$, and the most massive short-period planets are found in multiple stellar systems (Zucker and Mazeh, 2002; Udry et al., 2004; Udry et al., 2003b; Eggenberger et al., 2004a). Indeed, planetary candidates with a mass $M_2 \sin i \gtrsim 2\,M_{\rm Jup}$ and a period $P \lesssim 100$ days are all orbiting a component of a multiple stellar system[2].

The paucity of massive short-period planets cannot be attributed to observational selection effects since these planets are the easiest to detect. Moreover, even

[2] The only exceptions are: i) HD 162020 b (in parentheses) a probable brown dwarf with an actual mass much larger than its minimum mass (Udry et al., 2002), and ii) HD 168443 b member of a 2 very massive "planet" (brown-dwarf?) system (dotted line - see Udry et al. (2002) and Marcy et al. (2001) for further details)

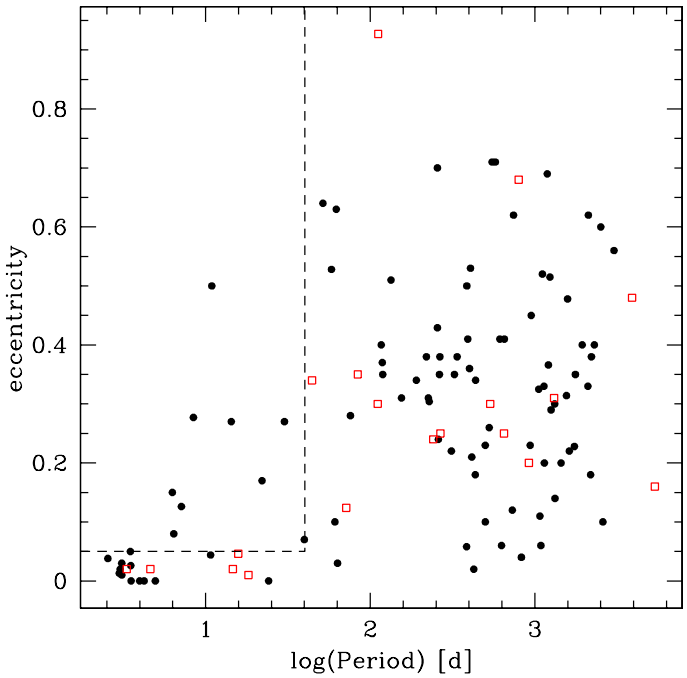

Figure 2. Eccentricity versus orbital period for all the extra-solar planetary candidates. Planets orbiting single stars are represented by filled circles while planets in binaries are represented by open rectangles. The dashed lines represent limits at $P = 40$ d and $e = 0.05$.

if the sample of planets orbiting a component of a multiple stellar system is small and incomplete, the presence of a few candidates in a zone of the diagram where there are no other planets is significant (Zucker and Mazeh, 2002).

For periods longer than ∼100 days, the distribution of the planetary companions in the $(M_2 \sin i - \log P)$ plane is also slightly different for the two samples. The mean $M_2 \sin i$ is smaller for planets in binaries than for planets orbiting single stars. This difference comes from the fact that no very massive planet has been found on a long-period orbit around a component of a multiple stellar system. Again, this cannot be attributed to an observational bias, because several planets with a smaller mass and on long-period orbits have been found in multiple stellar systems. However, due to the sparse data available for planets in multiple systems, it is not yet possible to say whether such a paucity is real or not (Eggenberger et al., 2004a).

3.2. The Eccentricity-Period Diagram

The distribution of the extra-solar planetary candidates in the $(e - \log P)$ plane is illustrated in Fig. 2. In this diagram, we note that all the planets with periods

shorter than ∼40 days members of multiple stellar systems have an eccentricity smaller than 0.05 whereas for longer period planets found in star systems this is not the case anymore. Some of the very short-period planets are so close to their parent star that tidal dissipation in the planet is likely to have circularized their orbit, even if they were originally eccentric (Rasio et al., 1996). This is certainly not the case for planets on longer periods like the three planets with a period between 10 and 40 days orbiting in multiple stellar systems. Their low eccentricities cannot be explained by invoking tidal dissipation alone. The observation that the three candidates have very low eccentricities can then be a clue to their formation and/or subsequent evolution history.

Remarks
1) Even if the orbital parameters of the binaries hosting planets are not exactly known, the projected separations of these systems (see Table I) indicate that the five planets with a period shorter that 40 days reside in very different types of systems according to the binary separations. There is thus no obvious correlation between the properties of these planets and the known orbital characteristics of the binaries or the star masses.
2) The formation/evolution history of planets found in multi-planet systems is probably different from the one of "single" planets. For the analysis presented here, all planets have been considered, but we have checked that our conclusions remain unchanged if the planets belonging to multi-planet systems are removed from the samples. Among the five short-period planets found in multiple stellar systems, two also belong to multi-planet systems.

Can available theoretical studies on planets in binaries help us to understand these observed properties?

4. Planet Formation and Stellar Duplicity

4.1. Modeled Effects of a Stellar Companion

Several points of view were followed by theoreticians to tackle the influence of the presence of a close additional stellar companion in the system on planet formation, at various stages of the formation process. We briefly review here the main approaches of planet formation in binaries to see if their results are in agreement with the observed emerging properties of exoplanets in multiple stellar systems.

Binary-disk interactions. Transfer of angular momentum between the binary and circumbinary or circumstellar disks leads to a truncation of their inner or outer edges, respectively. The location of the truncation radius is determined by the balance between the gravitational and the viscous torques. Artymowicz and Lubow (1994) derived quantitative estimates for the truncation of these disks. For example, in the case of $\mu = M_2/(M_1 + M_2) = 0.3$, the circumprimary disk is truncated at

$r_t = 0.4a_{bin}$ for a nearly circular binary and at $r_t = 0.18a_{bin}$ for $e_{bin} = 0.5$. The disk truncation can influence substantially the planet formation and evolution.

Giant planet formation in binaries. Two main mechanisms are proposed to explain giant planet formation: core accretion and disk instability (Pollack et al., 1996; Boss, 1997; Mayer et al., 2002). The core accretion mechanism involves runaway gas accretion on a solid core formed at a distance $\geq$ 3-5 AU from the star (beyond the "ice boundary") where the dust icy outer layer allows for rapid core agglomeration. In the disk instability model, a gravitationally unstable disk fragments directly into self-gravitating clumps of gas and dust that can contract and become giant gaseous protoplanets.

These two mechanisms share the common characteristic that giant planets form preferentially in the relatively cool outer regions of protoplanetary disks. The planets have then to migrate to the central regions of the disk where they are currently being observed. It is less clear, however, how a secondary star may affect the efficiency of planet formation. According to Nelson (2000), the companion has a negative influence on both mechanisms, slowing or inhibiting altogether giant planet formation. Boss (1998) claims the opposite, namely that giant planet formation via gravitational collapse is favoured in binaries.

Evolution of an embedded planet in a binary. Kley (2001) studied the evolution of a giant planet still embedded in a protoplanetary disk around the primary component of a binary system. A 1 M_{Jup} planet was placed on a circular orbit at 5.2 AU from a 1 $M_\odot$ star. The secondary star had a mass of 0.5 $M_\odot$ and an eccentricity of 0.5. The binary semi-major axis was varied from 50 to 100 AU. The simulations showed that the companion altered the evolutionary properties of the planet: the mass accretion rate was increased and the inward migration time was reduced. In the simulations, the planet eccentricity was also modified: it first grew due to the perturbations induced by the secondary star, but then declined because of the damping action of the disk. The final result was a rapid decay of the planet semi-major axis and a damping of the initial eccentricity.

Long-term stability of orbits. Holman and Wiegert (1999) studied the long-term survival of planets in different regions of phase space near a binary, for different values of the binary eccentricity, separation and mass ratio. For a binary with $\mu = M_2/(M_1+M_2) = 0.3$, the largest stable orbit around the primary is at $r_c = 0.37a_{bin}$ for $e_{bin} = 0$ or at $r_c = .14\,a_{bin}$ for $e_{bin} = 0.5$. A companion star orbiting beyond about 5 times the planetary distance is thus not a serious threat to the long-term stability of inner planetary orbits.

4.2. Discussion

From the different theoretical approaches, we see that a planet can be formed and almost always persist in a binary stellar system. We also expect the secondary star to have an impact on planet formation, at least for close binaries. Kley (2001)

has shown that this remains true for the subsequent evolution (migration) of giant planets in binaries. The changes in the migration process induced by a stellar companion (faster migration, enhanced mass accretion rate, eccentricity damping) seem to be in agreement with the observation that the most massive short-period planets are found in multiple stellar systems and have very small eccentricities. It should, however, be noticed that several of the binaries known to host planets are probably very different from the ones studied by Kley. The five planets with a period shorter than 40 days orbit in binaries with very different separations (from $\sim$20 to $\sim$1000 AU). It seems unlikely that the perturbations produced by a wide companion would influence the evolution of a protoplanet orbiting at or below a few AU. This, however, deserves further study.

5. Enlarging the Sample of Planets in Binaries

As previously stated, our knowledge regarding the existence of extra-solar planets in multiple stellar systems is still far from being complete. To fulfill the urgent need of increasing the available statistics, several programmes aiming at detecting and studying planets in binaries have been recently started. None of them are concluded yet, and only preliminary results are available. We will briefly mention here some of the efforts we are pursuing in this direction, mainly our spectroscopic programme searching for giant planets in spectroscopic and close visual binaries and the reverse approach, searching for stellar companions around stars known to host planets with VLT/NACO.

5.1. Short-Period Planets in Spectroscopic Binaries

Goal. We have started a programme aiming to search for short-period circumprimary giant planets in single-lined spectroscopic binaries (Eggenberger et al., 2004b; Eggenberger et al., 2003; Udry et al., 2004). This programme will enable us to probe the existence of giant planets in close binaries for which circumstellar disks can be seriously truncated. The results will also possibly allow us to establish the minimal separation for a binary to host circumprimary giant planets and provide observational constraints regarding the formation and subsequent evolution of these planets.

Sample. Our sample of binaries is composed of about 100 systems in both hemispheres with periods ranging from approximately 2 to more than 50 years. These binaries have been selected on the basis of different CORAVEL surveys of G and K-dwarfs of the solar neighbourhood (Duquennoy and Mayor, 1991; Udry et al., 1998; Halbwachs et al., 2003). For each binary, 10-15 high-precision radial-velocity measurements are obtained with the CORALIE or the ELODIE spectrographs. Residual velocities around the Keplerian orbit (for relatively short-period binaries) or

around a drift (for longer period systems) are then analyzed looking for short-period radial-velocity variations.

Preliminary results. Preliminary results show for most of the targets a very small dispersion of the high-precision radial velocities (a few m/s), ruling thus out the presence of a short-period giant planet orbiting the primary star. In some cases, however, the measured radial-velocity dispersion is high. The goal is now to reject false alarms among these interesting candidates. Residual velocity variations can be due to different effects:

- There is a planet orbiting the primary.
- The primary is an active star.
- The star is an unrecognized double-lined spectroscopic binary revealed by the higher resolution of the ELODIE/CORALIE spectrographs. In such a case, the secondary spectrum introduces an additional source of noise on the measurement of the radial velocity.
- The spectroscopic binary is also a visual pair. When the binary separation is of the order of the diameter of the spectrograph fiber (projected onto the sky), the relative fraction of light coming from the stars entering the fiber is then variable and depends on the seeing and guiding characteristics. This produces radial-velocity variations due to changes of the gravity center of the combined cross-correlation function (Pepe et al., 2004).
- The binary is a triple system whose secondary component, bright enough to perturb the primary spectrum, is itself a spectroscopic double system (Santos et al., 2002; Zucker et al., 2003a).

Diagnostics such as the analysis of the cross-correlation function shape (bisector inverse slope), photometry and correlation with different sets of lines are used to differentiate between these different effects (Queloz et al., 2001; Santos et al., 2002; Eggenberger et al., 2003).

As mentioned above, typical cases of false alarm are induced by the contamination of the stellar light by the secondary spectrum, when the light ratio of the binary components is larger than $\sim 1\%$. In such cases, the two components have to be disentangled and individual radial velocities obtained. This is possible with 2D cross-correlation schemes such as TODCOR (Zucker and Mazeh, 1994). Such an approach is fundamental in the case of close visual binaries as illustrated by the case of HD 41004. Santos et al. (2002) describe the false detection of a Saturn-mass planet orbiting around HD 41004 A with $P = 1.33$ d. They have shown that the false detection was induced by the Doppler-moving spectrum of a faint close spectroscopic binary system (HD 41004 B) superimposed on the primary spectrum. HD 41004 B was visually detected by Hipparcos ($\Delta m \sim 4$ and $\rho = 0.5''$).

This case was treated with a multi-order version of TODCOR (Zucker, 2003). The two-dimension correlation technique provided individual velocities for the 2 components of the visual binary. The faint binary companion (HD 41004 B) is confirmed to be a system made of an M2 dwarf + a $M_2 \sin i = 19\,M_{\rm Jup}$ companion (Zucker et al., 2003a) and the primary star (HD 41004 A) is now found to host a $2.3\,M_{\rm Jup}$ planet (Zucker et al., 2003b). Similar cases from our CORALIE planet-search programme are under study.

5.2. Searches for Faint Companions to Stars Harboring Planets

Another approach regarding the issue of planets in multiple stellar systems is to look for stellar companions to stars known to harbor planets. Interferometry, adaptive optics and direct imaging have been used to search for faint companions to stars bearing planets over separations ranging from a few AU to a few hundreds of AU (Lloyd et al., 2004; Els et al., 2001; Luhman and Jayawardhana, 2002; Lowrance et al., 2002; Patience et al., 2002). Companion searches are complementary to radial-velocity observations and the combination of different observational techniques is required to have a good characterization of the environment and binary status of stars bearing planets.

The VLT/NACO programme. In this perspective, we have started a systematic AO search with VLT/NACO for faint companions to ~ 100 solar-type stars in two subsamples: stars with known extra-solar planets and "single" stars for comparison. The statistically significant results of the survey will permit to quantify the star multiplicity effect on planet formation and subsequent evolution. Finally, as a by-product, this programme may also reveal new stellar systems hosting planets, allowing us to further precise the emerging characteristics of planets in binaries discussed in Sect. 3.

Observations and detection limits. The observational strategy consists in observing pairs of *planet-comparison* stars, in narrow-band filters within the H or K bands. Whenever a companion is detected, a supplementary observation is made in J for better resolution and in order to obtain a colour index. Given our exposure times and experimental settings, the detection limits (5σ) in H ($\lambda_c = 1.644\,\mu$m) are $\Delta H \sim 3.2$ at $0.1''$ and $\Delta H \sim 9$ for $0.7'' \lesssim \rho \lesssim 3.3''$. In K ($\lambda_c = 2.166\,\mu$m), the detection limits (5σ) are $\Delta K \sim 2.5$ at $0.1''$ and $\Delta K \sim 7$ for $0.6'' \lesssim \rho \lesssim 9''$. Most of our stars being at a distance between 10 and 100 pc, our survey will probe systems with projected separations from a few AU up to about 1000 AU.

Preliminary results. During the first 5 nights granted to this survey, 28 stars with planets have been observed and 1 companion has been found. The binarity rate for the planet-hosting star sample is thus $3 \pm 3\,\%$. For the comparison sample, we have observed 34 stars and found 7 companions. The binarity rate is therefore $20 \pm 7\,\%$. Error bars have been computed using the bootstrap method. These preliminary results already show a significant difference between the two subsamples

which indicates that planets form preferentially around single stars than around the components of binary or multiple stellar systems. Globally, stellar multiplicity seems thus to have a negative effect on planet formation and/or survival.

This result is, however, based on small samples and we will have to wait for the whole survey to be completed before reaching a statistically significant and definitive conclusion. We also need second-epoch observations in order to confirm that the detected companions are physical and not optical, although contamination by background stars is not expected to be important for the detected companions.

6. Conclusion

The characteristics of giant planets found in multiple stellar systems seem to be different from those of planets orbiting around single stars, at least for the short-period massive planets. The major observed differences are that short-period planets can be more massive if they are part of a multiple stellar system and that they moreover move on low-eccentricity orbits. A larger statistics of planets found in multiple systems is, of course, needed to precise this preliminary trend and enable us to quantify in a more definite way the effect of stellar duplicity on giant planet formation and subsequent evolution.

Simulations run by Kley (2001) show that the presence of a companion star affects the orbital properties and the accretion rate of a Jupiter-mass planets still embedded in a disk around the primary component of a binary by increasing the migration and mass accretion rates of the planet and damping the orbital eccentricity. Even though the star systems known to host planets are usually different from the ones modeled by Kley, migration seems to have played an important role in the history of planets found in multiple systems and the conclusions by Kley could be valid more generally. The extend to which a star companion affects the formation and/or the subsequent evolution of a planet is still an open question and further investigations will be needed to precise this point.

From the observational point of view, the characterization of the star systems susceptible to host planets is still to be done and could bring constraints for the models, thus helping our understanding of giant planet formation. The few programmes searching for planets in binaries are then of importance and will hopefully soon start to provide interesting data. A preliminary first census of close faint companions to stars hosting planets compared to equivalent results for stars without planets suggests that stellar multiplicity acts against the occurrence of planets with periods in the range of the spans of the present radial-velocity planet-search surveys.

References

Allen, C., A. Poveda, and M. Herrera: 2000, 'Wide binaries among high-velocity and metal-poor stars'. *A&A* **356**, 529.

Artymowicz, P. and S. Lubow: 1994, 'Dynamics of binary-disk interaction. 1: Resonances and disk gap sizes'. *ApJ* **421**, 651.

Bodenheimer, P., O. Hubickyj, and J. Lissauer: 2000, 'Models of the in Situ Formation of Detected Extrasolar Giant Planets'. *Icarus* **143**, 2.

Boss, A.: 1997, 'Giant planet formation by gravitational instability.'. *Science* **276**, 1836.

Boss, A.: 1998, 'Giant Planet Formation Induced by a Binary Star Companion'. *AAS/Division for Planetary Sciences Meeting* **30**, 1057.

Boss, A.: 2000, 'Formation of Extrasolar Giant Planets: Core Accretion or Disk Instability?'. *Earth Moon and Planets* **81**, 19.

Butler, R., G. Marcy, D. Fischer, T. Brown, A. Contos, S. Korzennik, P. Nisenson, and R. Noyes: 1999, 'Evidence for Multiple Companions to upsilon; Andromedae'. *ApJ* **526**, 916.

Butler, R., G. Marcy, E. Williams, H. Hauser, and P. Shirts: 1997, 'Three New "51 Pegasi–Type" Planets'. *ApJ* **474**, L115.

Campbell, B., G. Walker, and S. Yang: 1988, 'A search for substellar companions to solar-type stars'. *ApJ* **331**, 902.

Cochran, W., A. Hatzes, R. Butler, and G. Marcy: 1997, 'The Discovery of a Planetary Companion to 16 Cygni B'. *ApJ* **483**, 457.

Cochran, W., A. Hatzes, M. Endl, D. Paulson, G. Walker, B. Campbell, and S. Yang: 2002, 'A Planetary Companion to the Binary Star Gamma Cephei'. *AAS/Division for Planetary Sciences Meeting* **34**, 916.

Desidera, S., R. Gratton, M. Endl, M. Barbieri, R. Claudi, R. Cosentino, S. Lucatello, F. Marzari, and S. Scuderi: 2003, 'A search for planets in the metal-enriched binary HD 219542'. *A&A* **405**, 207.

Duquennoy, A. and M. Mayor: 1991, 'Multiplicity among solar-type stars in the solar neighbourhood. II - Distribution of the orbital elements in an unbiased sample'. *A&A* **248**, 485.

Eggenberger, A., S. Udry, and M. Mayor: 2003, 'Planets in Binaries'. In: D. Deming and S. Seager (eds.): *Scientific Frontiers in Research on Extrasolar Planets*, ASP Conf. Ser. **294**, pp. 43–46.

Eggenberger, A., Udry, S., and Mayor, M.: 2004a, 'Statistical properties of exoplanets. III. Planet properties and stellar multiplicity', *Astron. Astroph.* **417**, 353–360.

Eggenberger, A., S. Udry, M. Mayor, J.-L. Beuzit, A.-M. Lagrange, and G. Chauvin: 2004b, 'Detection and properties of extrasolar planets in multiple star systems'. In: C. Terquem, A. Lecavelier-des-Etangs, and J.-P. Beaulieu (eds.): *Extrasolar planets, today and tomorrow*, ASP Conf. Ser. **321**, pp. 93–100.

Els, S., M. Sterzik, F. Marchis, E. Pantin, M. Endl, and M. Kürster: 2001, 'A second substellar companion in the Gliese 86 system. A brown dwarf in an extrasolar planetary system'. *A&A* **370**, L1.

Fischer, D., G. Marcy, R. Butler, S. Vogt, and K. Apps: 1999, 'Planetary Companions around Two Solar-Type Stars: HD 195019 and HD 217107'. *PASP* **111**, 50.

Fischer, D., G. Marcy, R. Butler, S. Vogt, G. Henry, D. Pourbaix, B. Walp, A. Misch, and J. Wright: 2003, 'A Planetary Companion to HD 40979 and Additional Planets Orbiting HD 12661 and HD 38529'. *ApJ* **586**, 1394.

Halbwachs, J.: 1986, 'Common proper motion stars in the AGK 3'. *A&AS* **66**, 131.

Halbwachs, J., M. Mayor, S. Udry, and F. Arenou: 2003, 'Multiplicity among solar-type stars. III. Statistical properties of the F7-K binaries with periods up to 10 years'. *A&A* **397**, 159.

Hale, A.: 1994, 'Orbital coplanarity in solar-type binary systems: Implications for planetary system formation and detection'. *AJ* **107**, 306.

Hatzes, A.P., Cochran, W.D., Endl, M., McArthur, B., Paulson, D.B., Walker, G.A.H., Campbell, B., and Yang, S.: 2003, *Astrophys. J.* **599**, 1383.

Hauser, H. and G. Marcy: 1999, 'The Orbit of 16 Cygni AB'. *PASP* **111**, 321.

Holman, M. and P. Wiegert: 1999, 'Long-Term Stability of Planets in Binary Systems'. *AJ* **117**, 621.

Kley, W.: 2001, 'Planet Formation in Binary Systems'. In: *IAU Symposium 200*. p. 511.

Latham, D., R. Stefanik, T. Mazeh, M. Mayor, and G. Burki: 1989, 'The unseen companion of HD114762 - A probable brown dwarf'. *Nature* **339**, 38.

Lin, D., P. Bodenheimer, and D. Richardson: 1996, 'Orbital migration of the planetary companion of 51 Pegasi to its present location.'. *Nature* **380**, 606.

Lloyd, J.P., Liu, M.C., Graham, J.R., Enoch, M., Kalas, P., Marcy, G.W., Fischer, D., Patience, J., Macintosh, B., Gavel, D.T., Olivier, S.S., Max, C.E., White, R., Ghez, A.M., and McLean, I.S.: 2004, *IAU Symposium* **202**, 462.

Lowrance, P., J. Kirkpatrick, and C. Beichman: 2002, 'A Distant Stellar Companion in the Upsilon Andromedae System'. *ApJ* **572**, L79.

Luhman, K. and R. Jayawardhana: 2002, 'An Adaptive Optics Search for Companions to Stars with Planets'. *ApJ* **566**, 1132.

Marcy, G., R. Butler, D. Fischer, G. Laughlin, S. Vogt, G. Henry, and D. Pourbaix: 2002, 'A Planet at 5 AU around 55 Cancri'. *ApJ* **581**, 1375.

Marcy, G., R. Butler, S. Vogt, D. Fischer, and M. Liu: 1999, 'Two New Candidate Planets in Eccentric Orbits'. *ApJ* **520**, 239.

Marcy, G., R. Butler, S. Vogt, M. Liu, G. Laughlin, K. Apps, J. Graham, J. Lloyd, K. Luhman, and R. Jayawardhana: 2001, 'Two Substellar Companions Orbiting HD 168443'. *ApJ* **555**, 418.

Mayer, L., T. Quinn, J. Wadsley, and J. Stadel: 2002, 'Formation of Giant Planets by Fragmentation of Protoplanetary Disks'. *Science* **298**, 1756.

Mayor, M., Udry, S., Naef, D., Pepe, F., Queloz, D., Santos, N.C., and Burnet, M.: 2004, *Astron. Astroph.* **415**, 391.

McGrath, M., E. Nelan, D. Black, G. Gatewood, K. Noll, A. Schultz, S. Lubow, I. Han, T. Stepinski, and T. Targett: 2002, 'An Upper Limit to the Mass of the Radial Velocity Companion to ρ^1 Cancri'. *ApJ* **564**, L27.

Naef, D., D. Latham, M. Mayor, T. Mazeh, J. Beuzit, G. Drukier, C. Perrier-Bellet, D. Queloz, J. Sivan, G. Torres, S. Udry, and S. Zucker: 2001, 'HD 80606 b, a planet on an extremely elongated orbit'. *A&A* **375**, L27.

Naef, D., M. Mayor, S. Korzennik, D. Queloz, S. Udry, P. Nisenson, R. Noyes, T. Brown, J. Beuzit, C. Perrier, and J. Sivan: 2003, 'The ELODIE survey for northern extra-solar planets. II. A Jovian planet on a long-period orbit around GJ 777 A'. *A&A* **410**, 1051.

Nelson, A.: 2000, 'Planet Formation is Unlikely in Equal-Mass Binary Systems with A $\sim$ 50 AU'. *ApJ* **537**, L65.

Patience, J., R. White, A. Ghez, C. McCabe, I. McLean, J. Larkin, L. Prato, S. Kim, J. Lloyd, M. Liu, J. Graham, B. Macintosh, D. Gavel, C. Max, B. Bauman, S. Olivier, P. Wizinowich, and D. Acton: 2002, 'Stellar Companions to Stars with Planets'. *ApJ* **581**, 654.

Pepe, F., Mayor, M., Queloz, D., and Udry, S.: 2004, *IAU Symposium* **202**, 103; also available at `http://obswww.unige.ch/Preprints/Preprints/cine\_art.html`.

Pollack, J., O. Hubickyj, P. Bodenheimer, J. Lissauer, M. Podolak, and Y. Greenzweig: 1996, 'Formation of the Giant Planets by Concurrent Accretion of Solids and Gas'. *Icarus* **124**, 62.

Queloz, D., G. Henry, J. Sivan, S. L. Baliunas, J.-L. Beuzit, R. Donahue, M. Mayor, D. Naef, C. Perrier, and S. Udry: 2001, 'No planet for HD 166435'. *A&A* **379**, 279.

Queloz, D., M. Mayor, L. Weber, A. Blécha, M. Burnet, B. Confino, D. Naef, F. Pepe, N. Santos, and S. Udry: 2000, 'The CORALIE survey for southern extra-solar planets. I. A planet orbiting the star Gliese 86'. *A&A* **354**, 99.

Queloz, D., Mayor, M., Naef, D., Pepe, F., Santos, N.C., Udry, S., and Burnet, M.: 2004, *IAU Symposium* **202**, 106, also available at `http://obswww.unige.ch/Preprints/Preprints/cine\_art.html`.

Rasio, F., C. Tout, S. Lubow, and M. Livio: 1996, 'Tidal Decay of Close Planetary Orbits'. *ApJ* **470**, 1187.

Santos, N., M. Mayor, D. Naef, F. Pepe, D. Queloz, S. Udry, M. Burnet, J. Clausen, B. E. Helt, E. H. Olsen, and J. Pritchard: 2002, 'The CORALIE survey for southern extra-solar planets. IX. A 1.3-day period brown dwarf disguised as a planet'. *A&A* **392**, 215.

Tokovinin, A., R. Griffin, Y. Balega, E. Pluzhnik, and S. Udry: 2000, 'The Triple System HR 7272'. *Astronomy Letters* **26**, 116.

Udry, S., Eggenberger, A., Mayor, M., Mazeh, T., and Zucker, S.: 2004, 'Planets in multiple star systems: properties and detection', *Revista Mexicana de Astronomia y Astrofisica Conference Series* **21**, 207.

Udry, S., M. Mayor, D. Latham, R. Stefanik, G. Torres, T. Mazeh, D. Goldberg, J. Andersen, and B. Nordstrom: 1998, 'A Survey for Spectroscopic Binaries in a Large Sample of G Dwarfs'. In: *ASP Conf. Ser. 154: Cool Stars, Stellar Systems, and the Sun*, 2148.

Udry, S., M. Mayor, D. Naef, F. Pepe, D. Queloz, N. C. Santos, and M. Burnet: 2002, 'The CORALIE survey for southern extra-solar planets. VIII. The very low-mass companions of HD 141937, HD 162020, HD 168443 and HD 202206: Brown dwarfs or "superplanets"?'. *A&A* **390**, 267.

Udry, S., M. Mayor, and N. Santos: 2003b, 'Statistical properties of exoplanets. I. The period distribution: constraints for the migration scenario'. *A&A* **407**, 369.

Wuchterl, G., T. Guillot, and J. Lissauer: 2000, 'Giant Planet Formation'. *Protostars and Planets IV*, 1081.

Zucker, S.: 2003, 'Cross-correlation and maximum-likelihood analysis: a new approach to combining cross-correlation functions'. *MNRAS* **342**, 1291.

Zucker, S. and T. Mazeh: 1994, 'Study of spectroscopic binaries with TODCOR. 1: A new two-dimensional correlation algorithm to derive the radial velocities of the two components'. *ApJ* **420**, 806.

Zucker, S. and T. Mazeh: 2002, 'On the Mass-Period Correlation of the Extrasolar Planets'. *ApJ* **568**, L113.

Zucker, S., T. Mazeh, N. C. Santos, S. Udry, and M. Mayor: 2003a, 'Multi-order TODCOR: Application to observations taken with the CORALIE echelle spectrograph. I. The system HD 41004'. *A&A* **404**, 775.

Zucker, S., Mazeh, T., Santos, N.C., Udry, S., and Mayor, M.: 2003b, 'Multi-order TODCOR: Application to observations taken with the CORALIE echelle spectrograph. II. A planet in the system HD 41004', *Astron. Astroph.* **404**, 775.

Zucker, S., D. Naef, D. Latham, M. Mayor, T. Mazeh, J.-L. Beuzit, G. Drukier, C. Perrier-Bellet, D. Queloz, J. Sivan, G. Torres, and S. Udry: 2002, 'A Planet Candidate in the Stellar Triple System HD 178911'. *ApJ* **568**, 363.

On the Formation and Evolution of Giant Planets in Close Binary Systems

Richard P. Nelson
Astronomy Unit, School of Mathematical Sciences, Queen Mary, University of London, Mile End Road, London, E1, 4NS, U.K.

Abstract. We present the results of hydrodynamic simulations of Jovian mass protoplanets that form in circumbinary discs. The simulations follow the orbital evolution of the binary plus planet system acting under their mutual forces, and forces exerted by the viscous circumbinary disc. The evolution involves the clearing of the inner circumbinary disc initially, so that the binary plus planet system orbits within a low density cavity. Interaction between disc and planet causes inward migration of the planet towards the inner binary. Subsequent evolution can take three distinct paths: (*i*) The planet enters the 4:1 mean motion resonance, but is gravitationally scattered through a close encounter with the binary; (*ii*) The planet enters the 4:1 mean motion resonance, the resonance breaks, and the planet remains in a stable orbit outside the resonance; (*iii*) When the binary has initial eccentricity $e_{\rm bin} \geq 0.2$, the disc becomes eccentric, leading to a stalling of the planet migration, and apparent stability of the planet plus binary system.

1. Introduction

Extrasolar planets have been observed to exist in binary systems (e.g. γ Cephei, 16 Cygni B). Most solar-type stars appear to be members of binary systems (Duquennoy and Mayor, 1991), and most T Tauri stars, whose discs are thought to be the sites of planet formation, appear to members of binary systems (e.g. Ghez, Neugebauer, and Matthews, 1993; Leinert et al., 1993). A number of circumbinary discs have also been observed (e.g. DQ Tau, AK Sco, UZ Tau, GW Ori). Therefore it is of interest to explore how stellar multiplicity affects planet formation, and post-formation planetary orbital evolution, including formation in circumbinary discs.

Previous work examined the stability of planetary orbits in binary systems (Dvorak, 1986; Holman and Wiegert, 1999). This showed that there is a critical ratio of planetary to binary semimajor axis for stability, depending on the binary mass ratio, $q_{\rm bin}$, and eccentricity $e_{\rm bin}$. A recent paper (Quintana et al., 2002) explored the late stages of terrestrial planet formation in the α Centauri system, concluding that the binary companion can help speed up planetary accumumlation.

Recent work (Kley and Burkert, 2000) examined the effect that an external binary companion can have on the migration and mass accretion of a giant planet forming in a circumstellar disc. In this work we explore the evolution of Jovian mass protoplanets forming in *circumbinary* discs. Previous work has shown that a giant protoplanet embedded in a disc around a single star undergoes inward migration driven by the viscous evolution of the disc (Lin and Papaloizou, 1986;

Nelson et al., 2000). Here we are interested in how this process is affected if the central star is replaced by a close binary system. In particular we are interested in exploring the orbital stability of planets that migrate towards the central binary. For multiple planet systems, migration can induce resonant locking (e.g. Snellgrove, Papaloizou, and Nelson, 2001). An issue explored in this article is whether a planet can become stably locked into resonance with a central binary, with the disc acting as a source of dissipation.

This article is organised as follows. In section 2 we describe the numerical setup. In section 3 we described the results of the simulations. Finally, we discuss the results and draw conclusions in section 4.

2. Numerical Setup

We consider the interaction between a coplanar binary plus planet system and a two-dimensional, gaseous, circumbinary disc within which it is supposed the giant planet formed. The equations of motion are similar to those described in Nelson et al., 2000. Each of the stellar components and the planet experience the gravitational force of the other two, as well as that due to the disc. The disc is evolved using the hydrodynamics code NIRVANA (Ziegler and Yorke, 1997). The planet and binary orbits are evolved using a fifth-order Runge-Kutta scheme (Press et al., 1992)

We adopt a disc model in which the aspect ratio $H/r = 0.05$, and the viscosity parameter $\alpha = 5 \times 10^{-3}$. The surface density is set up to have an inner cavity within which the planet and binary orbit:

$$\Sigma(r) = \begin{cases} 0.01\Sigma_0 & \text{if } r < r_p \\ \Sigma_0 r^{-1/2} \exp\left[(r - r_g)/\Delta\right] & \text{if } r_p \leq r \leq r_g \\ \Sigma_0 r^{-1/2} & \text{if } r > r_g \end{cases} \quad (1)$$

where r_p is the initial planet orbital radius, r_g is the radius where the gap joins the main disc, and Δ is chosen to ensure that the gap profile correctly joins onto the main disc and the inner cavity. Simulations initiated with no inner cavity show that one is formed by the action of the binary system and planet clearing gaps in their local neighbourhood. As the planet migrates in towards the central binary these gaps join to form a single cavity. The disc mass is normalised through the choice of Σ_0 such that a standard disc model with $\Sigma(r) = \Sigma_0 r^{-1/2}$ throughout would contain about 4 Jupiter masses interior to the initial planet radius r_p (assumed in physical units to be 5 AU). Thus the disc mass interior to the initial planet radius would be about twice that of a minimum mass solar nebula model. Calculations were also run with disc masses a factor of three higher. The total mass of the binary plus planet system is assumed to be 1 $M_\odot$. Dimensionless units are used such that the total mass of the binary system plus planet $M_{\rm tot} = 1$ and the gravitational constant $G = 1$. The initial binary semimajor axis is $a_{\rm bin} = 0.4$ in all simulations, and the

Table I. Column 1 gives the run label, column 2 the planet mass (Jupiter masses), column 3 the binary mass ratio, column 4 the binary eccentricity, column 6 the disc mass (in units described in section 2). Column 7 gives the number of grid cells in the r and ϕ direction, column 8 describes the mode of evolution which is defined in the text.

Run label	m_p (M_J)	$q_{\rm bin}$	$e_{\rm bin}$	m_d/m_{mmsn}	$N_r \times N_\phi$	Result
A1	1	0.1	0.1	2	160×320	Mode 2
A2	1	0.1	0.1	2	320×640	Mode 2
B1	3	0.1	0.1	2	160×320	Mode 1
B2	3	0.1	0.1	2	320×640	Mode 1
C1	1	0.1	0.1	6	160×320	Mode 2
C2	1	0.1	0.1	6	320×640	Mode 2
D1	3	0.1	0.1	6	160×320	Mode 1
D2	3	0.1	0.1	6	320×640	Mode 1
E1	1	0.1	0.05	2	160×320	Mode 1
E2	1	0.1	0.05	2	320×640	Mode 1
F1	1	0.1	0.05	6	160×320	Mode 1
G1	1	0.25	0.1	2	160×320	Mode 1
G2	1	0.25	0.1	2	320×640	Mode 1
H1	1	0.25	0.1	6	160×320	Mode 1
H2	1	0.25	0.1	6	320×640	Mode 1
I1	3	0.25	0.1	6	160×320	Mode 1
I2	3	0.25	0.1	6	320×640	Mode 1
J1	1	0.1	0.2	2	160×320	Mode 3
K1	1	0.1	0.3	2	160×320	Mode 3

initial planet semimajor axis $a_p = 1.4$. The unit of time quoted in the discussion of the simulation results below is the orbital period at $R = 1$.

3. Numerical Results

The results of the simulations are shown in Table I. They can be divided into three categories, which are described below, and are most strongly correlated with changes in the binary mass ratio, $q_{\rm bin}$, and binary eccentricity $e_{\rm bin}$. Runs for similar parameters but calculated at different resolutions always gave the same qualitative

results. In some runs the planet enters the 4:1 mean motion resonance with the binary. The associated resonant angles are defined by:

$$\begin{aligned} \psi_1 &= 4\lambda_s - \lambda_p - 3\omega_s & \psi_2 &= 4\lambda_s - \lambda_p - 3\omega_p \\ \psi_3 &= 4\lambda_s - \lambda_p - 2\omega_s - \omega_p & \psi_4 &= 4\lambda_s - \lambda_p - 2\omega_p - \omega_s \end{aligned} \tag{2}$$

where λ_s, λ_p are the mean longitudes of the secondary and planet, respectively, and ω_s, ω_p are the longitudes of pericentre of the secondary and planet, respectively. When in resonance ψ_3 or ψ_4 should librate, or all the angles should librate. Below we give a very brief description of the 3 different modes of evolution obtained in the simulations.

Mode 1: The protoplanet migrates inwards and temporarily enters the 4:1 mean motion resonance with the binary before being gravitationally scattered by the secondary star. This arises for a binary system with low mass ratio (i.e. $q_{\rm bin} = 0.1$), when the planet becomes tightly locked into the resonance, because resonant eccentricity pumping leads to a close encounter between planet and secondary star. Planetary scattering also occurs for central binaries with larger mass ratios (i.e. $q_{\rm bin} = 0.25$), but *via* a different evolutionary path. Here, the resonance locking is weak, and the resonance breaks. The subsequent interaction between planet and binary is strong enough to perturb the planet into a close encounter with the secondary star.

Mode 2: This occurred only for binary mass ratios $q_{\rm bin} = 0.1$. The planet migrates inwards and enters the 4:1 mean motion resonance. The resonance locking is weak, and the resonance breaks, but the planet does not undergo a close encounter and scattering off the central binary. Instead it migrates outwards slightly through interaction with the disc, and becomes 'parked' in an orbit beyond the 4:1 resonance, where it remains for the duration of the simulation.

Mode 3: This applies when the central binary has $e_{\rm bin} \geq 0.2$. The circumbinary disc is driven eccentric by the binary. The interaction between protoplanet and eccentric disc stalls the inward migration, and the planet remains orbitally stable over long times.

Examples of each of these modes of evolution are described in detail below. Note that similar simulations run at different numerical resolution resulted in the same qualitative results, as shown in table I.

3.1. Planetary Scattering – Mode 1

Table I shows that a number of simulations resulted in a close encounter between the planet and binary system, leading to gravitational scattering of the planet to larger radii, or into an unbound state. These runs are labelled as 'Mode 1'. Typically the initial scattering causes the eccentricity of the planet to grow to values $e_p \simeq 0.9$, and the semimajor axis to increase to $a_p \simeq 6$–8. In runs that were continued for significant times after this initial scattering, ejection of the planet could occur after subsequent close encounters. We note, however, that the small disc sizes considered in these models preclude us from calculating the post-scattering evolution

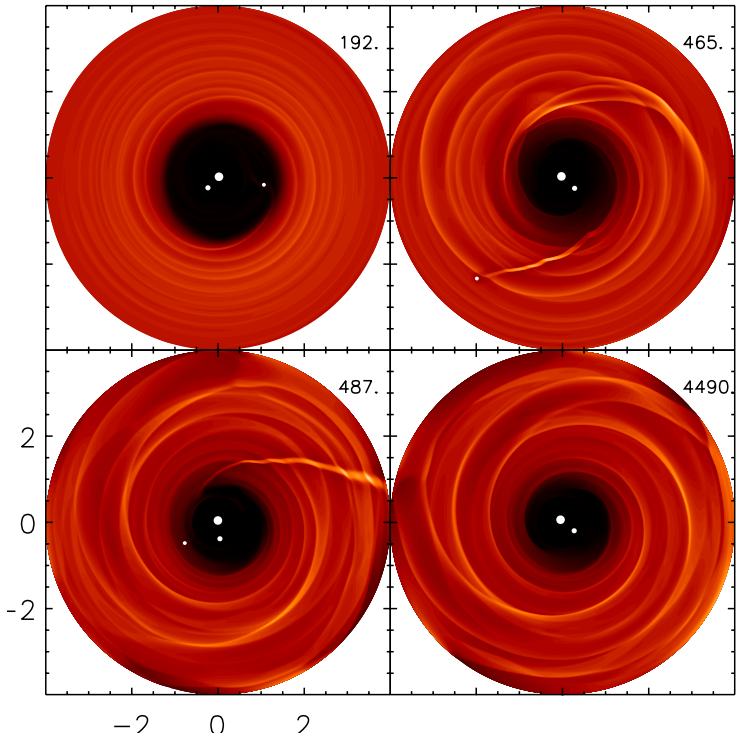

Figure 1. This figure shows the evolution of disc and planet plus binary system for run D2. Time are shown at top right hand corners in orbital periods at $R = 1$.

accurately, since the planet trajectories take them beyond the outer boundary of the disc, usually located at $R_{\rm out} = 4$. The eventual ejection of the planets may be a function of this if continued disc-planet interaction after scattering causes eccentricity damping. Subsequent close encounters may then be prevented.

3.1.1. *Low Binary Mass Ratios*

In all models with $q_{\rm bin} = 0.1$ and $e_{\rm bin} \leq 0.1$, that resulted in the planet being scattered, the evolution proceeded as follows. The protoplanet migrates in towards the central binary and temporarily enters the 4:1 mean motion resonance with the binary. The resonant angle ψ_3 defined in equation 2 librates with low amplitude, indicating that the planet is strongly locked into the resonance. The resonance causes the eccentricity of the planet to increase, until the planet has a close encounter with the secondary star, and is scattered out of the resonance into a high eccentricity orbit with significantly larger semimajor axis.

We use the results of model D2 to illustrate the main points discussed above. Figure 1 shows the evolution of the disc and planet plus binary system. The early migration stage is shown in the first panel. The second panel shows the system just

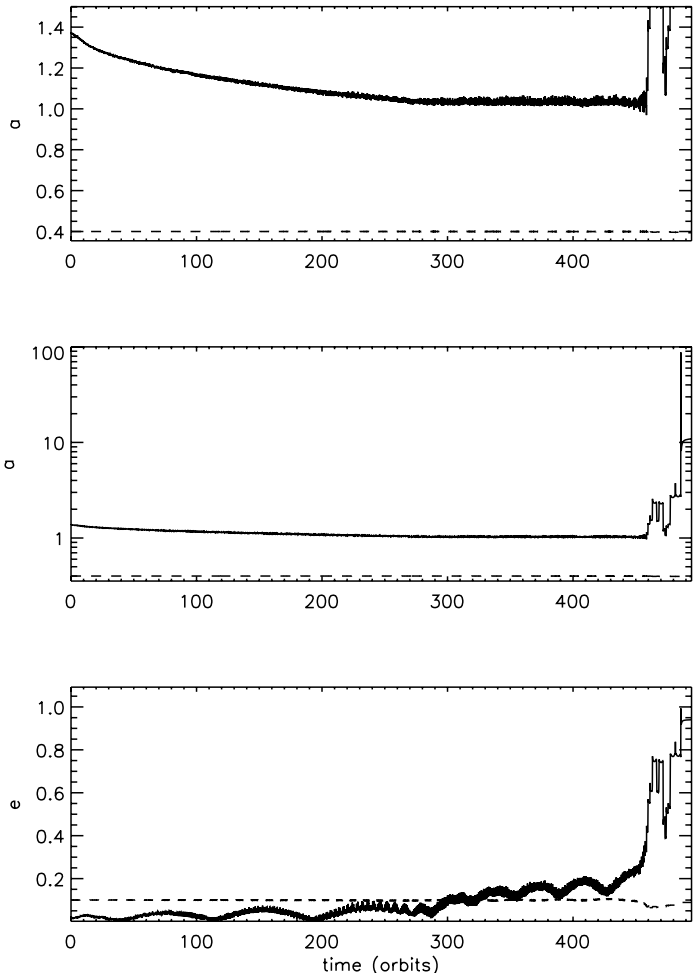

Figure 2. This figure shows the evolution of the semimajor axes and eccentricities of the planet (solid line) and binary system (dashed line) for run D2.

after the initial scattering encounter, with the planet immersed in the main body of the circumbinary disc. The third panel shows the planet approaching the binary during a subsequent pericentre passage, and the fourth panel shows a time when the planet orbit takes it out beyond the main body of the disc modeled here. For those simulations that resulted in the planet being completely ejected, a circumbinary disc remains that eventually returns to a state similar to that which would have existed had no planet been present.

The orbital evolution of the planet and binary for model D2 is shown in Fig. 2. The upper and middle panels show the semimajor axes versus time. The lowest panel shows the evolution of the eccentricities versus time. The time evolution of

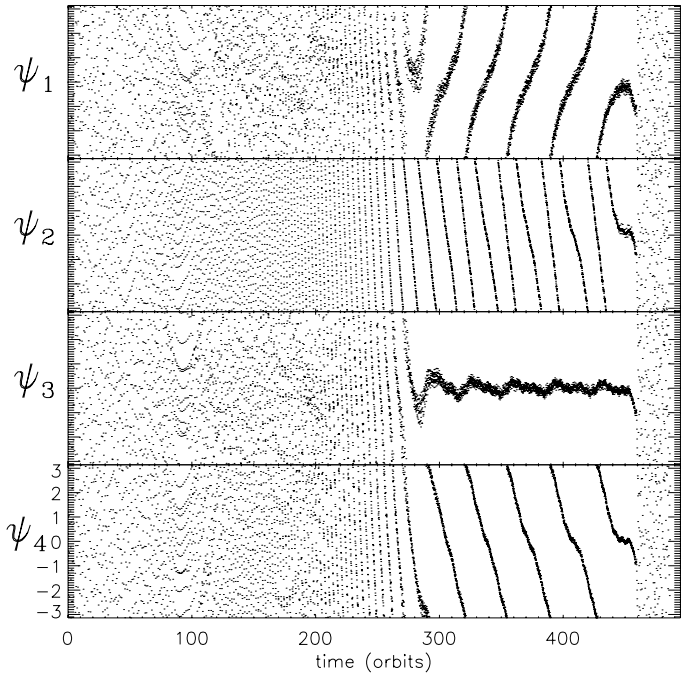

Figure 3. This figure shows the evolution of the resonant angles (ψ_i, $i = 1...4$) for run D2. Note the low amplitude libration of ψ_3 for $t > 300$, indicating that planet is strongly locked in resonance.

the resonant angles ψ_1, ψ_2, ψ_3, and ψ_4 defined in equation 2 is shown in Fig. 3. The planet enters the 4:1 resonance at $t \sim 300$, corresponding to the time when the planet eccentricity e_p starts to grow steadily (Fig. 2). Figures 2 and 3 show that the resonance breaks at $t \sim 460$, and that e_p at this stage has been increased to $e_p \simeq 0.4$. A close encounter with the secondary star excites e_p up to between $e_p = 0.8$–1. The final state at the end of the simulation has $a_p \simeq 10$ and $e_p \simeq 0.95$. It is likely that a longer integration of this system would result in the planet being ejected from the system, leading to the formation of a 'free-floating planetary mass object'.

3.1.2. *Higher Binary Mass Ratios*

For calculations with $q_{\rm bin} = 0.25$, the evolution differed from that just described, although scattering of the planet still occurred. The planet enters the 4:1 mean motion resonance, and large amplitude librations of the resonant angle ψ_3 occur, accompanied by large oscillations of e_p, indicative of weak resonant locking. The resonance becomes undefined and breaks when $e_p = 0$ during these high amplitude librations, and the subsequent interaction between planet and binary causes a close encounter and scattering of the planet. This is facilitated by the larger value of $q_{\rm bin}$,

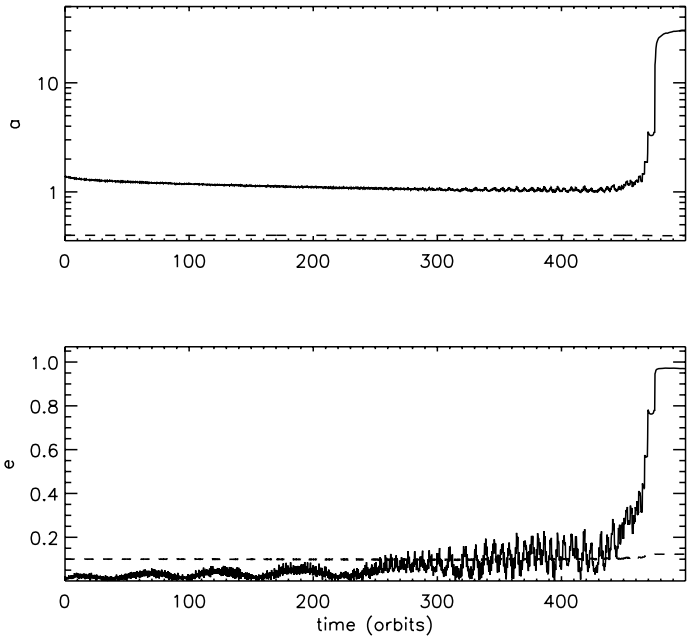

Figure 4. This figure shows the evolution of the semimajor axes and eccentricities for the planet (solid line) and binary system (dashed line) in run I1.

since a similar run involving weak resonant locking with $q_{\rm bin} = 0.1$ does not lead to scattering of the planet (see section 3.2 below).

Results for model I1 are shown in Figs. 4 with 5. Figure 4 shows the evolution of the semimajor axes and eccentricities. Figure 5 shows the time evolution of the resonant angles. Comparing Fig. 4 and Fig. 2 we can see that the eccentricity in this case does not increase to such large values before the resonance breaks. Instead the resonance appears to break because e_p goes to zero momentarily. After resonance breaking, the planet-binary interaction is strong enough to perturb the planet into an orbit that leads to a close encounter with the binary. Comparing Figs. 5 and 3, we see that the resonant angles undergo libration with much greater amplitude in this case, indicating that the resonant locking is weaker in run I1 than in run D2.

3.2. Near-Resonant Planet – Mode 2

A mode of evolution was found in some of the simulations with $q_{\rm bin} = 0.1$ and $e_{\rm bin} = 0.1$ leading to the planet orbiting stably just outside of the 4:1 resonance. These cases are labelled as 'Mode 2' in table I. Here, the protoplanet migrates inwards and becomes weakly locked into the 4:1 resonance, with the resonant angle ψ_3 librating with large amplitude. The evolution in the resonance is similar to that described for run I1 in section 3.1.2. The resonance is undefined and breaks when

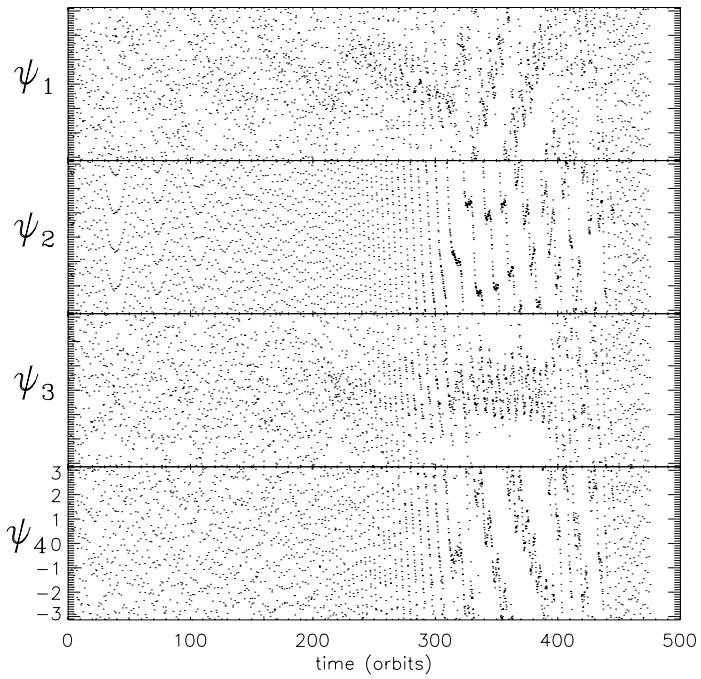

Figure 5. This figure shows the evolution of the resonant angles (ψ_i, $i = 1...4$) for run I1. Note the high amplitude libration of ψ_3 for $t > 300$, showing that the planet is weakly locked in resonance.

$e_p = 0$. However, because $q_{\rm bin}$ is smaller, the planet is not perturbed into an orbit that leads to a close encounter with the binary. Instead it undergoes a period of outward migration through interaction with the disc by virtue of the eccentricity having reattained values of $e_p \simeq 0.15$–0.2 once the resonance is broken. Calculations by Nelson, 2003 have shown that gap-forming protoplanets orbiting in tidally truncated discs undergo outward migration if they are given eccentricities of this magnitude impulsively, due to the sign of the torque exerted by the disc reversing for large eccentricities. The outward migration moves the planet to a safer distance away from the binary, thus avoiding instability.

Once the planet has migrated to just beyond the 4:1 resonance the outward migration halts, since its eccentricity reduces slightly, and the planet remains there for the duration of the simulation. The system achieves a balance between eccentricity damping by the disc and eccentricity excitation by the binary, maintaining a mean value of $e_p \simeq 0.12$. The torque exerted by the disc on the planet is significantly weakened by virtue of the finite eccentricity (Nelson, 2003), preventing the planet from migrating back towards the binary.

We use calculation C2 to illustrate the points discussed above. The orbital evolution of the system is shown in Fig. 6. The upper panel shows the semimajor axes. The planet initially migrates in towards the binary, and halts as it reaches

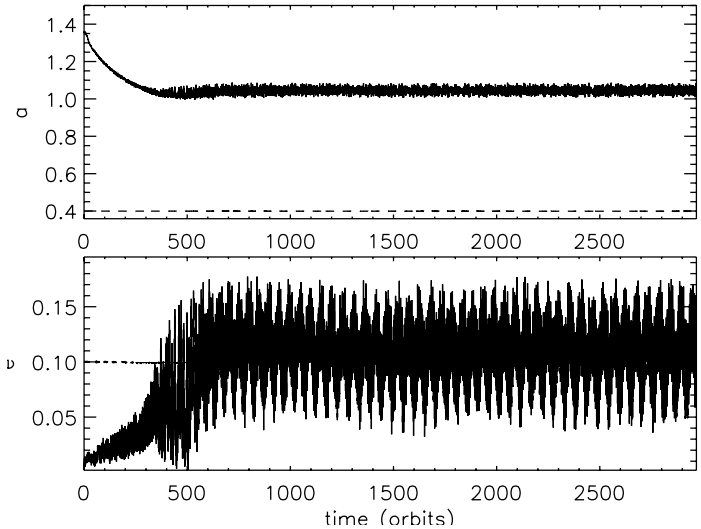

Figure 6. This figure shows the evolution of the semimajor axes and eccentricities of the planet (solid line) and binary system (dashed line) for run C2.

$a_p \simeq 1$. The planet enters the 4:1 resonance and the resonant angle ψ_3 undergoes large amplitude librations, similar to those observed for run I1 in Fig. 5 and in contrast to those observed in Fig. 3 for run D2. e_p also undergoes large amplitude oscillations, with the planet eccentricity reaching $e_p = 0$ just prior to the breaking of the resonance at $t \simeq 500$. Once the planet leaves the resonance, e_p increases to $e_p \simeq 0.15$, and the semimajor axis a_p increases. The planet remains outside the 4:1 resonance for the duration of the simulation, (i.e. for nearly 3000 planetary orbits) with the eccentricity oscillating between values of $0.05 \leq e_p \leq 0.18$. Continuation of this run in the absence of the disc indicates that the planet remains stable for over 6×10^6 orbits. This is in good agreement with the stability criteria obtained by (Holman and Wiegert, 1999).

3.3. Eccentric Disc – Mode 3

A mode of evolution was found in which the planetary migration was halted before the planet could approach the central binary. This only occurred when the central binary had an initial eccentricity of $e_{\rm bin} \geq 0.2$. The migration stalls because the circumbinary disc becomes eccentric. We label runs of this type as 'Mode 3' in table I. Interaction between the protoplanet and the eccentric disc leads to a reduction or even reversal of the time-averaged torque driving the migration. Simulations of this type can be run for many thousands of planetary orbits without any significant net inward migration occurring. Such systems are likely to be stable long after the circumbinary disc has dispersed, and are probably the best candidates for finding stable circumbinary extrasolar planets.

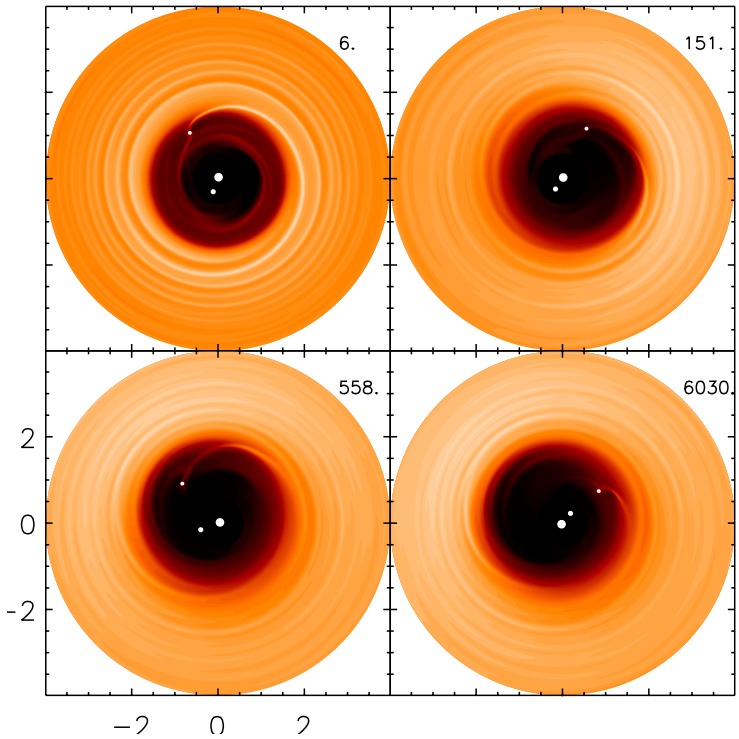

Figure 7. This figure shows the evolution of the disc and planet plus binary system for run J1. Note the formation of the eccentric disc.

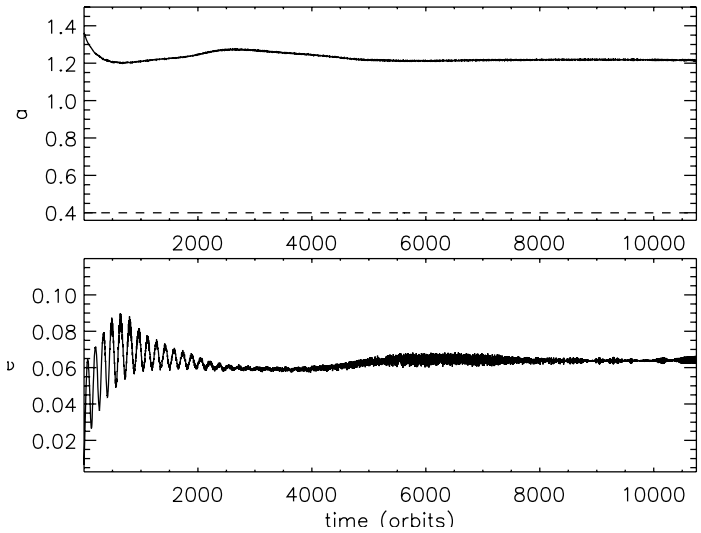

Figure 8. This figure shows the evolution of the semimajor axes and eccentricities of the planet (solid line) and binary system (dashed line) for run J1.

Model J1 is used to illustrate the evolution of the 'Mode 3' class of models. The evolution of the disc, binary, and protoplanet are shown in Fig. 7. Early on the system looks similar to the one shown in Fig. 1. However, the circumbinary disc eventually becomes eccentric. The orbital evolution of the planet and binary is shown in Fig. 8. The upper panel shows the semimajor axes, and the lower panel the eccentricities. Initially the planet undergoes inward migration. After a time of $t \simeq 400$ the migration reverses. This is the time required for the eccentric disc mode to be established. Over longer time scales the planet oscillates between inward and outward migration before settling into a non migratory state. The planet remains orbiting with semimajor axis $a_p \simeq 1.2$, and is unable to migrate in towards the central binary. The resulting system is likely to consist of a central binary with $e_{\rm bin} \leq 0.2$, and a circumbinary planet orbiting with $a_p \simeq 1.2$ and $e_p \simeq 0.06$ after dispersal of the circumbinary disc.

The disc eccentricity in these models is driven by the central binary, and occurs only when the inner binary has significant eccentricity. Simulations performed without a protoplanet included also produced eccentric discs. In these cases, the disc inner cavity was retained, and those only calculations with $e_{\rm bin} \geq 0.2$ gave rise to an eccentric disc. In previous work (Papaloizou, Nelson, and Masset, 2001), an eccentric disc could be excited for companion masses with $q \geq 0.02$, corresponding to a companion of 20 Jupiter masses orbiting a solar mass star. This mode of disc-eccentricity driving occurred even for companions on circular orbits, and arose because of nonlinear mode coupling between an initially small $m = 1$ mode in the disc and the $m = 2$ component of the orbiting companion potential. To operate, this requires significant amounts of gas to be present at the 1:3 resonance location (which is at $r \simeq 2.08 \times a_{\rm bin}$). The initial disc models in this paper have inner cavities that extend much beyond the 1:3 resonance of the binary, so this cannot be the cause of the disc eccentricity observed in runs J1 and K1, and those without protoplanets included. Instead, the disc eccentricity is driven by the $m = 1$ component of the eccentric binary potential which excites a global $m = 1$ mode in the disc.

4. Conclusions

We have examined the evolution of giant protoplanets orbiting in circumbinary discs. For low eccentricity binaries (i.e. $e_{\rm bin} \leq 0.1$), when the planet migrates inwards through interaction with the circumbinary disc the resulting interaction with the binary can lead to the planet being scattered and ejected from the sytem, leading to the formation of a 'free-floating planet'. However, there is also a finite probability that the planet can end up orbiting stably just outside of the 4:1 mean motion resonance. For binaries with significant eccentricity (i.e. $e_{\rm bin} \geq 0.2$), the circumbinary disc becomes eccentric, and this stalls the inward migration of the planet, preventing it from approaching the binary and being ejected. Given that

close binaries often have eccentric orbits, this suggests that circumbinary planets in these systems may be orbiting at safe distances, and could potentially be observable.

Acknowledgements

The computations reported here were performed using the UK Astrophysical Fluids Facility (UKAFF).

References

Duquennoy, A., Mayor, M.: 1991, 'Multiplicity among solar-type stars in the solar neighbourhood. II – Distribution of the orbital elements in an unbiased sample', *A & A* **248**, 485.

Dvorak, R.: 1986, 'Critical orbits in the restricted three-body problem', *A & A* **167**, 379.

Ghez, A. M., Neugebauer, G., Matthews, K.: 1993, 'The multiplicity of T Tauri stars in the star forming regions Taurus-Auriga and Ophiuchus-Scorpius: A 2.2 micron speckle imaging survey', *Ap. J.* **106**, 2005.

Holman, M.J., Wiegert, P.A.: 1999, 'Long-term stability of planets in binary systems', *Ap. J.* **117**, 621.

Kley, W., Burkert, A.: 2000, 'Disks and Planets in Binary Systems', *Discs, Planetesimals, and Planets*, ASP Conference Series, eds. F. Garzon, C. Eiroa, D de Winter, T.J. Mahoney.

Leinert, Ch., Zinnecker, H., Weitzel, N., Christou, J., Ridgway, S. T., Jameson, R., Haas, M., Lenzen, R.: 1993, 'A systematic approach for young binaries in Taurus', *A & A* **278**, 129.

Lin, D.N.C., Papaloizou, J.C.B.: 1986, 'On the tidal interaction between protoplanets and the protoplanetary disk. III – Orbital migration of protoplanets', *Ap. J.* **309**, 846.

Nelson, R.P., Papaloizou, J.C.B., Masset, F., Kley, W.: 2000, 'The migration and growth of protoplanets in protostellar discs', *MNRAS* **318**, 18.

Nelson, R.P.: 2003, 'On the evolution of giant protoplanets forming in circumbinary discs', *MNRAS* **345**, 233.

Papaloizou, J. C. B.; Nelson, R. P.; Masset, F.: 2001, 'Orbital eccentricity growth through disc-companion tidal interaction', *A & A* **366**, 263.

Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P.: 1992, 'Numerical Recipes in FORTRAN', Cambridge Univ. Press, Cambridge, p. 710.

Quintana, E.V., Lissauer, J.J., Chambers, J.E., Duncan, M.J.: 2002, 'Terrestrial Planet Formation in the α Centauri System', *Ap. J.* **576**, 982.

Snellgrove, M. D., Papaloizou, J. C. B., Nelson, R. P.: 2001, 'On disc driven inward migration of resonantly coupled planets with application to the system around GJ876', *A & A* **374**, 1092

Ziegler, U., Yorke, H.: 1997, ' A Nested Grid Refinement Technique for Magnetohydrodynamical Flows', *Comp. Phys. Comm.* **101**, 54.

Induced Eccentricities of Extra Solar Planets by Distant Stellar Companions

T. Mazeh
Wise Observatory, Tel Aviv University, Tel Aviv, Israel

Abstract. This paper explores the assumption that the high eccentricities of the extra solar planets were induced by distant stellar companions. I put some constraints on possible, yet undetected, companions that could generate the observed eccentricities. Distant companions can also induce low eccentricities into the planetary orbits. The results of such an effect can be detected when low but significant eccentricities are measured for short period planets, for which we expect the orbit to have been circularized long ago. The feasibility of such an effect is briefly discussed.

1. Introduction

The relatively high eccentricities of some of the extra solar planets compose one of the surprising features of the newly discovered population (e.g., Schneider, 2003). Naively, one could expect low eccentricity orbits for planets that were formed out of a disk of particles that circle a central star in circular orbits (e.g., Kornet, Bodenheimer and Rozyczka, 2002), as indeed is the case for the planets in our Solar system.

Since the discovery of the planetary high eccentricities, various mechanisms have been proposed to explain their formation and/or dynamical evolution (e.g., Weidenschilling and Marzari, 1996; Chiang, Fischer and Thommes, 2002; Marzari and Weidenschilling 2002). One of these ideas was that the high eccentricities were induced by distant stellar companions. This idea was put forward immediately after Cochran et al. (1996) announced the discovery of a planet orbiting 16 Cyg B. The new planet had two special pronounced features. The first one was its high orbital eccentricity – 0.63 ± 0.08 (Cochran et al., 1997), which was remarkably larger than the eccentricities of all other seven planets known at that time. The other feature was the binarity of the parent star, which has a distant stellar companion – 16 Cyg A, at a separation of about 40 arc-sec (Hoffleit and Jaschek, 1982). Therefore, the suggestion that the two features are interrelated was a natural assumption. Three works (Mazeh, Krymolowski and Rosenfeld, 1997; Holman, Touma and Tremaine, 1997; Innanen et al., 1997) studied this effect for 16 Cyg B and showed that the high eccentricity could have been pumped by the distant companion.

Since the discovery of 16 Cyg Bb a few more planets with highly eccentric orbits were found, the most recent one being HD 80606 with an eccentricity of 0.93 (Naef et al., 2001). Not all those planets are known to have distant stellar companions. It seems therefore that even if we accept this model, we have to assume either that the parent stars of all eccentric planets have distant faint, *yet*

unknown, companions, or that in some cases there is another mechanism behind the high eccentricities of the extra solar planets.

In this article I review briefly the dynamics of the pumping. Following the work of Holman, Touma and Tremaine (1997) and Lin et al. (2000), I put some constraints on the presumed distant companions that could induce the presently known high eccentricities. A similar effect can be observed by detecting small but significant eccentricity for planets with short periods, where we expect the orbit to circularize to a high level. I discuss briefly the feasibility of such an effect.

2. The Eccentricity Modulation

Mazeh and Shaham (1979) have shown that in hierarchical stellar triple systems the third distant star can modulate the inner binary eccentricity. This effect was already pointed out by Kozai (1962), who studied the motion of some asteroids under the gravitation of the Sun and Jupiter. Stellar triple systems are characterized by mass ratios close to unity, whereas the typical mass ratio in the Sun-Jupiter-asteroid system is 1:0.001:0. Nevertheless, the nature of the modulation is the same, as shown by Harrington (1968) and Lidov and Ziglin (1976). A system with two stellar components and a planet has a typical mass ratio of 1:1:0.001, in between the stellar triple systems and the asteroids. As such, its modulation is of the same nature.

Mazeh and Shaham (1979), using a second-order Hamiltonian of the three-body problem (Harrington, 1968, 1969) estimated the modulation period of the eccentricity for a stellar triple system to be

$$P_{\rm mod} \sim P_3 \left(\frac{P_3}{P_{1,2}} \right) \left(\frac{M}{m_3} \right) , \tag{1}$$

where P_3 is the third star orbital period, $P_{1,2}$ is the close binary period, M is the total mass of the binary system, and m_3 is the mass of the distant companion. For our case we get

$$P_{\rm mod} \sim P_{\rm bin} \left(\frac{P_{\rm bin}}{P_{\rm plnt}} \right) \left(\frac{m_{\rm prim}}{m_{\rm comp}} \right) , \tag{2}$$

where $P_{\rm bin}$ is the orbital period of the stellar binary system, $P_{\rm plnt}$ is the orbital period of the planet, and $m_{\rm prim}$ and $m_{\rm comp}$ are the masses of the parent star of the planet and its distant companion, respectively. For 16 Cyg B this formula yields a modulation period of the order of 3×10^8 y, or 100 million inner orbit revolutions. The model assumes that 16 Cyg B is at present in one of its high eccentricity phases that lasts for about 100 millions years.

The long period of the eccentricity modulation is due to the large semi-major axes ratio of the 16 Cyg system, which renders the tidal forces of A exerted on B and its planet very small. One might think that the weakness of the tidal force turns

the amplitude of the eccentricity modulation to be small. This intuition is wrong. It is true that the weakness of the tidal forces makes the eccentricity variation very slow. However, the total increase of the planetary eccentricity, being accumulated for a long period of time, can be substantial. The theory shows that in some cases an initial minute eccentricity can grow by the tidal forces of the distant star up to 0.8 and higher, with a modulation period of the order of 10–100 million years.

3. Two Types of Eccentricity Modulation

From the seminal work of Kozai (1962) we know that the modulation of the planetary eccentricity, e, is associated with two other parameters of the planetary orbit – the longitude of the periastron, g, and the inclination of the planetary orbit relative to the plane of motion of the distant companion, i. Specifically:

$$\frac{de}{dt} \propto e\sqrt{1-e^2}\sin(2g)\sin^2 i \ . \tag{3}$$

Following Kozai we also know that to first order we have two constants of motion for the planetary modulation. The first is

$$\Theta = \cos^2 i(1-e^2) \ , \tag{4}$$

which is proportional to the square of the angular momentum of the planetary motion. The second is the second-order Hamiltonian

$$H_{\mathrm{2nd-order}} = -[1-3\cos^2 i](2+3e^2) + 15[1-\cos^2 i]e^2\cos(2g) \ . \tag{5}$$

These two constants make the motion such that one of the three variables – e, i or g, determines the other two. It is therefore possible to describe the variability of the system by a contour in a parameter space of any two of the three variables. Following Kozai, I present in Figs. 1 and 2 the variability of the planetary orbit in the $\big(e, \cos(2g)\big)$ space. This is especially convenient, as the time derivative of the eccentricity is proportional to $\sin(2g)$, and therefore the sign of $\sin(2g)$ determines whether the eccentricity increases or decreases.

Figure 1 presents such contours for $\Theta = 0.8$. The maximum inclination for this value of Θ is 26.6°, which results in $e = 0$. In this case the second-order derivative of e vanishes, and the orbit does not vary. (Note that the longitude of the periastron is not defined in such a case.) This case is presented in the figure by the upper horizontal boundary. For different values of e, up to about 0.45, we get small eccentricity modulation, as depicted in Fig. 1. The lower boundary of the figure corresponds to an inclination of 0°, where the two planes of motions coincide, and the derivative of the eccentricity vanishes as well.

Kozai (1962) also showed that some systems can show two types of modulation, depending on the value of Θ. For Θ smaller than 0.6, $2g$ can librate, and the e modulation can be substantial. The two types of modulation are depicted in Fig. 2,

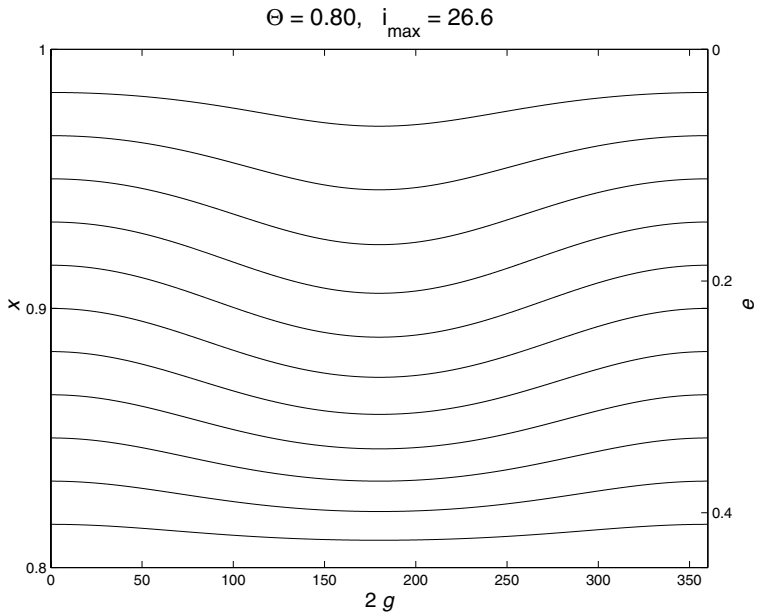

Figure 1. The eccentricity modulation for $\Theta = 0.8$. $x = 1 - e^2$

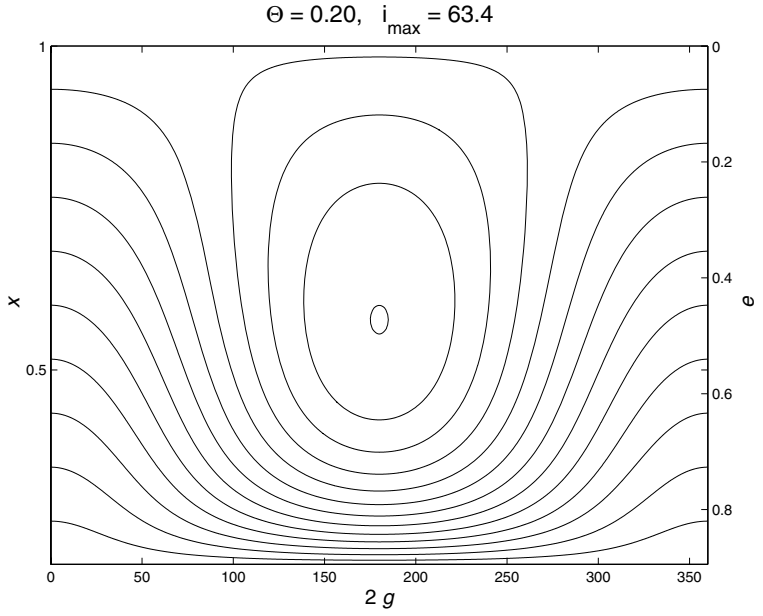

Figure 2. The eccentricity modulation for $\Theta = 0.2$

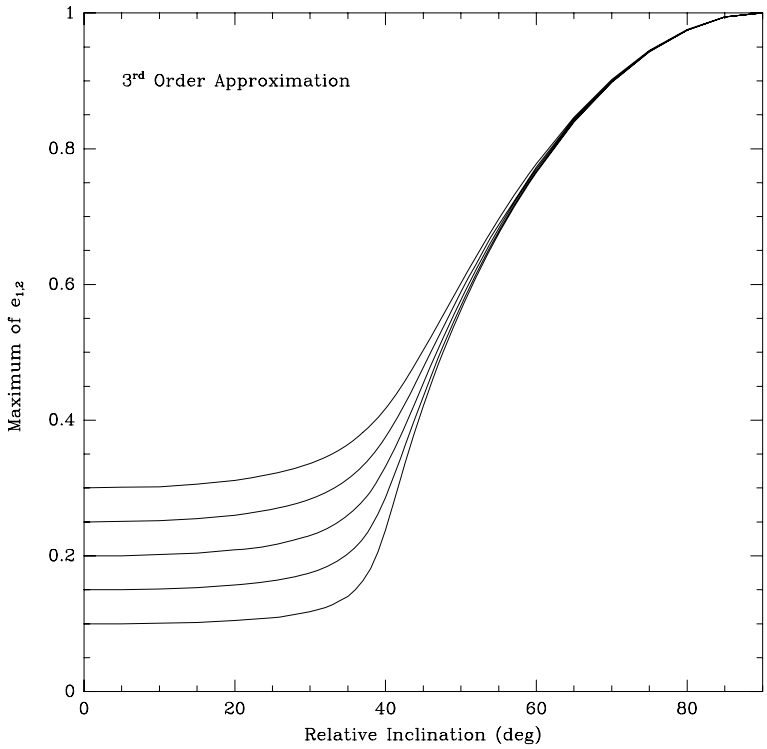

Figure 3. The planetary eccentricity maximum as a function of the initial relative inclinations for different initial planetary eccentricity.

which is plotted for $\Theta = 0.2$. Some of the contours present libration of $2g$, which oscillates around 0° without covering the whole range of angles.

A graph that shows the possible modulations of Cyg 16B is presented in Fig. 3 (Mazeh, Krimolowski and Rosenfeld, 1997). It shows the maximum possible eccentricity of the modulation as a function of the initial inclination, assuming the binary eccentricity is 0.85. To derive the eccentricity variation Mazeh, Krimolowski and Rosenfeld used their third-order expansion of the Hamiltonian, which agrees very well with numerical integration of Newton equations of the problem. The maximum eccentricity starts to grow when the inclination gets larger than $40\,^\circ$. This reflects the fact that for Θ smaller than 0.6 the $\big(e, \cos(2g)\big)$ parameter space is separated into two regions, and some of the curves extend to high eccentricities, as presented in Fig. 2.

4. Suppression of the Eccentricity Modulation

As we have seen, the eccentricity modulation is controlled by the variation of the longitude of the periastron. The slow eccentricity variation can accumulate up to large values only because the longitude of the periastron itself varies on a long time scale. Therefore, any additional drive that pushes the longitude of the periastron into faster precession and shortens the periastron cycle suppresses the accumulation of the eccentricity modulation (Holman, Touma and Tremaine, 1997; Innanen et al., 1997; Lin et al., 2000). One such effect is the famous General Relativity (=GR) precession of the periastron, with a period of

$$P_{\rm GR} = 3.36 \times 10^7 \,{\rm y}(1 - e_{\rm plnt}^2)\left(\frac{P_{\rm plnt}}{1{\rm y}}\right)\left(\frac{a_{\rm plnt}}{1\ {\rm AU}}\right)\left(\frac{m_{\rm prim}}{1M_\odot}\right)^{-1} \tag{6}$$

(Holman, Touma and Tramaine, 1997).

This timescale has to be compared with the period of the eccentricity modulation driven by a stellar companion given above, which will be written now as

$$P_{\rm mod} \sim P_{\rm plnt}\left(\frac{m_{\rm prim}}{m_{\rm comp}}\right)\left(\frac{a_{\rm bin}}{a_{\rm plnt}}\right)^3 (1 - e_{\rm bin}^2)^{3/2}\,, \tag{7}$$

where $e_{\rm bin}$ is the binary eccentricity (Holman, Touma and Tramaine, 1997).

The GR precession period can be derived from the known planetary parameters. Therefore, we can put upper limits on the distance of any possible companion that can strongly modulate the planetary eccentricity of each planet. This is done in Table 1, which lists all the known planets with eccentricity larger than 0.6, together with their $P_{\rm GR}$, which varies between 0.7 and 400 My. I consider two possible distant companions, one with mass of 0.1 $M_\odot$ and eccentricity of 0.4, and another with mass of 0.4 $M_\odot$ and eccentricity of 0.9. The Table lists the corresponding maximum distances of the two possible stellar companions, such that their modulation period would be about equal to $P_{\rm GR}$, so their effect would not be completely suppressed by the GR effect.

The maximum distances of the possible stellar companions range from 25 to 700 AU. Faint companions in this range of distances could have escaped detection. The last column of the Table lists the projected distances of the *known* companions to those planets. One can see that except for Cyg 16 B, the orbits of the known companions are much larger than the largest semi-major axes that could allow for eccentricity pumping. Therefore, the known secondaries could not have induced modulation that accounts for the present high eccentricity.

Table I.

Name	M_1 ($M_\odot$)	$P_{\rm plnt}$ (d)	$a_{\rm plnt}$ (AU)	$e_{\rm plnt}$	$P_{\rm GR}$ (My)	$a_{\rm max,1}$ (AU)	$a_{\rm max,2}$ (AU)	ρ (AU)
HD 74156 b	1.05	51.6	0.276	0.65	0.7	25	80	....
HD 80606	1.1	111.8	0.438	0.93	0.5	25	80	1220
HD 89744	1.4	256.0	0.883	0.70	7.6	90	290	....
Hip 75458	1.05	550.0	1.34	0.71	32	180	620	7980
HD 222582	1.00	577.1	1.36	0.76	31	190	620	4750
16 Cyg B	1.01	798.4	1.69	0.68	66	270	890	830
HD 2039	0.98	1190	2.2	0.69	130	380	1280	....
HD 39091	1.10	2280	3.5	0.63	403	690	2300	....

5. Small Planetary Eccentricity

An eccentricity modulation induced by a distant stellar companion does not necessarily have to be large. The discussion above has shown that if the relative inclination is smaller than 40°, the eccentricity modulation could be a few percent, as depicted in Fig. 1. In fact, we anticipate naively that the relative inclination would be small, specially if the binary and the planet have common origin. The results of such an effect can be observed by detecting small but significant eccentricity for planets with short periods, where we expect the orbit to circularize to a high level.

This effect can also be suppressed by the GR precession. Therefore, for the small eccentricity pumping to work we again need the GR time scale to be long enough. To see if this is plausible we consider a planet with an orbital period of the order of 10 days and a binary with a period of the order of 10 years. The ratio of the two precessions is

$$\frac{P_{\rm GR}}{P_{\rm mod}} \sim 20\left(\frac{P_{\rm plnt}}{10\,{\rm days}}\right)^{8/3}\left(\frac{P_{\rm bin}}{10\,{\rm y}}\right)^{-2}\left(\frac{m_{\rm prim}}{1M_\odot}\right)^{-2/3}\left(\frac{m_{\rm comp}}{m_{\rm prim}}\right)\frac{(1-e_{\rm plnt}^2)}{(1-e_{\rm bin}^2)^{3/2}}\,. \quad (8)$$

The effect can work as long as this ratio is not substantially smaller than unity. This implies that even if the companion mass is only 0.1 or even 0.05 of the primary mass, the effect can still work. Such a small companion is not easy to detect, either by radial velocity or astrometry, and a dedicated observational search for a possible faint companion is probably needed.

6. Discussion

This paper has shown that the GR effect puts strong constraints on the maximum distance of any possible stellar companion that could generate the present high

eccentricity of the eccentric planets. Out of the four known planets that have both high eccentricity and detected stellar companion, only 16 Cyg A could have generated the eccentricity of 16 Cyg Bb. This does not prove that the effect does not work. We probably did not notice many of the faint stellar companions to nearby bright stars, because of the small angular separation and the large visual contrast. We can turn the argument around and use the values of Table 1 as a guide to an observational search for distant companions that caused the high eccentricity.

Another planet in the system can also suppress the eccentricity modulation induced by a distant companion. In such a system, the longitude of the periastron of each of the two planets will be modulated by the tidal forces of both the distant companion and the other planet, suppressing the eccentricity modulation. Furthermore, two planets in some configurations tend to get locked in apsidal resonance, in which the longitude of their periastra vary together (e.g., Lee and Peale, 2003; Beauge, Ferraz-Mello and Michtchenko, 2003). In such systems, the modulation of the distant companion will be minimal.

Small but significant eccentricity of a planetary orbit with short period is an intriguing feature. Such eccentricity was found in a few short period stellar binary systems (Mazeh, 1990; Jha et al. 2000; Mazeh et al., 2001; Tokovinin and Smekhov, 2002). Interestingly enough, a few planets with periods shorter than 4 days show indications for small eccentricities. One example is HD 179949, with a period of 3.09 days and eccentricity of 0.05 ± 0.03 (Tinney et al., 2001). Unfortunately, the finite eccentricity is still not highly significant. When better orbits are available for HD 179949 and similar planets, a search for possible faint companions would be of great interest.

Acknowledgements

I wish to thank O. Tamuz and S. Zucker for illuminating discussions and helpful suggestions.

References

Beauge, C., S., Ferraz-Mello, and T. A. Michtchenko : 2003, 'Extrasolar Planets in Mean-Motion Resonance: Apses Alignment and Asymmetric Stationary Solutions'. *ApJ* **593**, 1124.

Chiang, E. I., D. Fischer, and E. Thommes: 2002, 'Excitation of Orbital Eccentricities of Extrasolar Planets by Repeated Resonance Crossing'. *ApJ* **564**, 105.

Cochran W. D., A. P. Hatzes, R. P. Butler, and G. W. Marcy: 1996, 'Detection of a planetary companion to 16 Cygni B'. *American Astronomical Society, DPS meeting* 28, 12.04.

Cochran W. D., A. P. Hatzes, R. P. Butler, and G. W. Marcy: 1997, 'The Discovery of a Planetary Companion to 16 Cygni B'. *ApJ* **483**, 457.

Harrington R. S.: 1968, 'Dynamical evolution of triple stars'. *AJ* **73**, 190.

Harrington R. S.: 1969, 'The Stellar Three-Body Problem'. *Celes. Mech.* **1**, 200.

Hoffleit D. and C. Jaschek: 1982, *Bright Star Catalogue (4th revised version), New Haven: Yale Univ. Observatory*.

Holman M., J. Touma, and S. Tremaine: 1997, 'Chaotic variations in the eccentricity of the planet orbiting 16 CYG B'. *Nature* **386**, 254.

Innanen K. A., J. Q. Zheng, S. Mikkola, and M.J. Valtonen: 1997, 'The Kozai Mechanism and the Stability of Planetary Orbits in Binary Star Systems'. *AJ* **113**, 1915.

Jha S., G. Torres, R. P. Stefanik, P. Robert, D. W. Latham, and T. Mazeh: 2000, 'Studies of multiple stellar systems - III. Modulation of orbital elements in the triple-lined system HD 109648'. *MNRAS* **317**, 375.

Kornet, K., P. Bodenheimer, and M., Rozyczka: 2002, 'Models of the formation of the planets in the 47 UMa system'. *AandA* **396**, 977.

Kozai Y.: 1962, 'Secular Perturbations of Astroids with High Inclinations and Eccentricity'. *AJ* **67**, 591

Lee, M. H, and S. J. Peale: 2003, 'Secular Evolution of Hierarchical Planetary Systems'. *ApJ* **592**, 1201.

Lidov, M. L., and S.L. Ziglin: 1976, 'Non-restricted double-averaged three body problem in Hill's case'. *Celes. Mech.* **13**, 471

Lin D. N. C., J. C. B. Papaloizou, C. Terquem, G. Bryden and S. Ida: 2000, 'Orbital Evolution and Planet-Star Tidal Interaction'. *Protostars and Planets IV*, ed. V. Mannings, A. P. Boss, S. S. Russell (Tucson: University of Arizona Press), 1111.

Marzari F. and S. J. Weidenschilling: 2002, 'The Jumping Jupiter Model'. *Icarus* **156**, 570.

Mazeh T.: 1990, 'Eccentric Orbits in Samples of Circularized Binary Systems: The Fingerprint of a Third Star'. *AJ* **99**, 675.

Mazeh T., Y. Krymolowski, and G. Rosenfeld: 1997, 'The High Eccentricity of the Planet Orbiting 16 Cygni B'. *ApJL* **L477**, 103.

Mazeh T., D. W. Latham, E. Goldberg, G. Torres, R. P. Stefanik, T. J. Henry, S. Zucker, O. Gnat, and E. Ofek: 2001, 'Studies of multiple stellar systems - IV. The triple-lined spectroscopic system Gliese 644'. *MNRAS* **325**, 343.

Mazeh T., and J. Shaham: 1979, 'The orbital evolution of close triple systems - The binary eccentricity'. *A&A* **77**, 145.

Naef D., D. W. Latham, M. Mayor, T. Mazeh, J. L. Beuzit, G. A. Drukier, C. Perrier-Bellet, D. Queloz, J. P. Sivan, G. Torres, and 2 coauthors: 2001, 'HD 80606 b, a planet on an extremely elongated orbit'. *A&A* **375**, 27.

Schneider, J.: 2003, www.obspm.fr/encycl/encycl.html

Tinney, C. G. et al.: 2001, 'First Results from the Anglo-Australian Planet Search: A Brown Dwarf Candidate and a 51 Peg-like Planet'.2001 *ApJ* **551**, 507.

Tokovinin A. A., and M. G. Smekhov: 2002, 'Statistics of spectroscopic sub-systems in visual multiple stars'. *A&A* **382**, 118.

Weidenschilling S. J., and F. Marzari: 1996, 'Giant Extrasolar Planets? Jumping Jupiters!', *DPS* **28**, 1214.

Is Planetary Migration Inevitable?

Caroline E. J. M. L. J. Terquem
Institut d'Astrophysique de Paris, 98 bis, bd Arago, F-75014 Paris
Université Denis Diderot-Paris VII, 2 Place Jussieu, F-75251 Paris Cedex 5

Abstract. According to current theories, tidal interactions between a disk and an embedded planet may lead to the rapid migration of the protoplanet on a timescale shorter than the disk lifetime or estimated planetary formation timescales. Therefore, planets can form only if there is a mechanism to hold at least some of the cores back on their way in. Once a giant planet has assembled, there also has to be a mechanism to prevent it from migrating down to the disk center. This paper reviews the different mechanisms that have been proposed to stop or slow down migration.

1. Introduction

Almost 20% of the extrasolar planets detected so far orbit at a distance between 0.038 and 0.1 astronomical unit (au) from their host star.

It is very unlikely that these short-period giant planets, also called 'hot Jupiters', have formed *in situ*. In most of the standard disk models, temperatures at around 0.05 au are larger than 1500 K (Bell et al., 1995; Papaloizou and Terquem, 1999), preventing the condensation of rocky material and therefore the accretion of a solid core there. Even if models with lower temperatures are considered, giant planets may form in close orbits according to the core accretion model only if the disk surface density is rather large and the accretion process very efficient (Bodenheimer et al. 2000; Ikoma et al., 2001). This is because at 0.05 au a solid core of about 40 earth masses has to be assembled before a massive gaseous envelope can be accreted (Papaloizou and Terquem, 1999; Bodenheimer et al., 2000). More likely, the hot Jupiters have formed further away in the protoplanetary nebula and have migrated down to small orbital distances. It is also possible that migration and formation were concurrent (Papaloizou and Terquem, 1999).

So far, three mechanisms have been proposed to explain the location of planets at very short orbital distances. One of them relies on the gravitational interaction between two giant planets, which may lead to orbit crossing and to the ejection of one planet while the other is left in a smaller orbit (Rasio and Ford, 1996; Weidenschilling and Marzari, 1996). However, this mechanism cannot reproduce the orbital characteristics of the extrasolar planets observed so far. Another mechanism is the so-called 'migration instability' (Murray et al., 1998; Malhotra, 1993). It involves resonant interactions between the planet and planetesimals located inside its orbit which lead to the ejection of a fraction of them while simultaneously causing the planet to migrate inward. Such interactions require a very massive disk to move a Jupiter mass planet from several astronomical units down to very small radii, as there has to be at least on the order of a Jupiter mass of planetesimals

inside the orbit of the planet. Such a massive disk is unlikely and furthermore it would be only marginally gravitationally stable. The third and most efficient mechanism involves the tidal interaction between the protoplanet and the gas in the surrounding protoplanetary nebula (Goldreich and Tremaine 1979, 1980; Lin and Papaloizou, 1979, 1993 and references therein; Papaloizou and Lin, 1984; Ward, 1986, 1997a). Here again the protoplanet can move significantly only if there is at least a comparable mass of gas within a radius comparable to that of its orbit. However this is not a problem since this amount of gas is needed anyway in the first place to form the planet.

Note that hot Jupiters can also be produced by the dynamical relaxation of a population of planets on inclined orbits, formed through gravitational instabilities of a circumstellar envelope or a thick disk (Papaloizou and Terquem, 2001). However, if objects as heavy as τ-Boo may be produced *via* fragmentation, it is unlikely that lower mass objects would form that way.

Tidal interaction between a disk and a planet may lead to two different types of migration (Ward, 1997a; Terquem et al., 2000 and references therein). Cores with masses up to about 10 $M_\oplus$ interact linearly with the surrounding nebula and migrate inward *relative* to the gas (type I migration). Planets with masses at least comparable to that of Jupiter interact nonlinearly with the disk and may open up a gap (Goldreich and Tremaine, 1980; Lin and Papaloizou, 1979, 1993 and references therein). The planet is then locked into the angular momentum transport process of the disk, and migrates *with* the gas at a rate controlled by the disk viscous timescale (type II migration). The direction of type II migration is that of the viscous diffusion of the disk. Therefore it is inward except in the outer parts of the disk which diffuse outward.

The drift timescale for a planet of mass $M_{\rm pl}$ undergoing type I migration in a uniform disk is (Ward, 1986, 1997a):

$$\tau_{\rm I}({\rm y}) \sim 10^8 \left(\frac{M_{\rm pl}}{M_\oplus}\right)^{-1} \left(\frac{\Sigma}{{\rm g\ cm^{-2}}}\right)^{-1} \left(\frac{r}{\rm AU}\right)^{-1/2} \times 10^2 \left(\frac{H}{r}\right)^2 \qquad (1)$$

where Σ is the disk surface density, r is the distance to the central star, and H is the disk semithickness. It is assumed here that the torque exerted by the material which corotates with the perturbation can be neglected.

For type II migration, the characteristic orbital decay timescale is the disk viscous timescale:

$$\tau_{\rm II}({\rm yr}) = \frac{1}{3\alpha}\left(\frac{r}{H}\right)^2 \Omega^{-1} = 0.05 \frac{1}{\alpha}\left(\frac{r}{H}\right)^2 \left(\frac{r}{\rm AU}\right)^{3/2} \qquad (2)$$

where α is the standard Shakura and Sunyaev (1973)'s parameter and Ω is the angular velocity at radius r.

It has recently been shown by Masset and Papaloizou (2003) that a planet in the intermediate mass range embedded in a disk massive enough may undergo a runaway migration. This typically happens for Saturn-sized giant planets embedded in disks with a mass several times the minimum mass of the solar nebula. The

timescale for runaway migration can be much shorter than that for type I or type II migration.

For typical disk parameters, the timescales given above are much shorter than the disk lifetime or estimated planetary formation timescales. Therefore, planets can form only if there is a mechanism to hold at least some of the cores back on their way in. Once a giant planet has assembled, there also has to be a mechanism to prevent it from migrating down to the disk center. We now review the different mechanisms that have been proposed to slow down, stop or reverse migration.

2. Stopping Type I Migration

2.1. Stopping Type I Migration at Small Radii

Cores undergoing orbital decay due to type I migration would stop before plunging onto the star if the disk had an inner (magnetospheric) cavity. Merger of incoming cores and subsequent accretion of a massive gaseous atmosphere could then produce a hot Jupiter *in situ* (Ward, 1997b, Papaloizou and Terquem, 1999). However, the extent of magnetospheric cavities is very limited, and therefore this mechanism cannot account for the presence of planets orbiting further away from the central star. For these planets to be assembled, type I migration has to be either avoided or halted. Type I migration would not take place if the interaction between the core and the disk could become nonlinear with a resulting gap formation. However, as discussed by Terquem et al. (2000), this situation is unlikely. We now review the mechanisms that have been proposed to halt or reverse type I migration.

2.2. Migration of Planets on Eccentric Orbits

The migration timescale given by Eq. (1) applies to a planet on a circular orbit. Papaloizou and Larwood (2000) have investigated the case of a planetary core on an eccentric orbit (in an axisymetric disk) with an eccentricity e significantly larger than the disk aspect ratio H/r. They found that the direction of orbital migration reverses for fixed eccentricity $e > 1.1H/r$. This is because the core spends more time near apocenter, where it is accelerated by the surrounding gas, than near pericenter, where it is decelerated. In general, the interaction between the core and the disk leads to eccentricity damping (Goldreich and Tremaine 1980). However, Papaloizou and Larwood (2000) showed that a significant eccentricity could be maintained by gravitational interactions with other cores. Papaloizou (2002) further studied the case of a core embedded in an eccentric disk. He showed that migration may be significantly reduced or even reverse from inward to outward when the eccentricity of the orbit of the core significantly exceeds that of the disk when that is large compared to H/r and the density profile is favorable. In some cases, such a high orbital eccentricity may be an equilibrium solution, and therefore suffers no damping.

2.3. Stopping Migration by a Toroidal Magnetic Field

Terquem (2003) has investigated the effect of a toroidal magnetic field on type I migration for a planet on a circular orbit. When a field is present, in contrast to the nonmagnetic case, there is no singularity at the corotation radius, where the frequency of the perturbation matches the orbital frequency. However, all fluid perturbations are singular at the so-called *magnetic resonances*, where the Doppler shifted frequency of the perturbation matches that of a slow MHD wave propagating along the field line. There are two such resonances, located on each side of the planet's orbit and within the Lindblad resonances. Like in the nonmagnetic case, waves propagate outside the Lindblad resonances. But they also propagate in a restricted region around the magnetic resonances.

The magnetic resonances contribute to a significant global torque which, like the Lindblad torque, is negative (positive) inside (outside) the planet's orbit. Since these resonances are closer to the planet than the Lindblad resonances, they couple more strongly to the tidal potential and the torque they contribute dominates over the Lindblad torque if the magnetic field is large enough. In addition, if $\beta \equiv c^2/v_A^2$, where c is the sound speed and v_A the Alfven velocity, increases fast enough with radius, the outer magnetic resonance becomes less important (it disappears altogether when there is no magnetic field outside the planet's orbit) and the total torque is then negative, dominated by the inner magnetic resonance. This leads to outward migration of the planet.

The amount by which β has to increase outward for the total torque exerted on the disk to be negative depends mainly on the magnitude of β. It was found that, for $\beta = 1$ or 100 at corotation, the torque exerted on the disk is negative when β increases at least as fast as r^2 or r^4, respectively.

The migration timescales that correspond to the torques calculated when a magnetic field is present are rather short. The orbital decay timescale of a planet of mass $M_{\rm pl}$ at radius r_p is $\tau = M_p r_p^2 \Omega_p/|T|$, where T is the torque exerted by the planet on the disk and Ω_p is the angular velocity at radius r_p. This gives:

$$\tau({\rm y}) = 4.3 \times 10^9 \left(\frac{M_{\rm pl}}{M_\oplus}\right)^{-1} \left(\frac{\Sigma_p}{100\ {\rm g\,cm^{-2}}}\right)^{-1} \left(\frac{r_p}{1\ {\rm AU}}\right)^{-1/2} \left(\frac{|T|}{\Sigma_p r_p^4 \Omega_p^2}\right)^{-1} \quad (3)$$

where Σ_p is the disk surface density at radius r_p. In a standard disk model, $\Sigma \sim 100$–10^3 g cm^{-2} at 1 AU (see, for instance, Papaloizou and Terquem 1999). Therefore, $\tau \sim 10^5$–10^6 yr for a one earth mass planet at 1 AU in a nonmagnetic disk, as $|T|/\left(\Sigma_p r_p^4 \Omega_p^2\right) \sim 10^3$ in that case (see Fig. 9 from Terquem 2003). This is in agreement with Ward (1986, 1997a, see Eq. 1 above). In a magnetic disk, $|T|/\left(\Sigma_p r_p^4 \Omega_p^2\right)$ may become larger (see Fig. 9 from Terquem 2003) leading to an even shorter migration timescale. However, it is important to keep in mind that these timescales are *local*. Once the planet migrates outward out of the region where β increases with radius, it may enter a region where β behaves differently and then resume inward migration for instance.

The calculations summarized here indicate that a planet migrating inward through a nonmagnetized region of a disk would stall when reaching a magnetized region. It would then be able to grow to become a terrestrial planet or the core of a giant planet. We are also led to speculate that in a turbulent magnetized disk in which the large scale field structure changes sufficiently slowly a planet may alternate between inward and outward migration, depending on the gradients of the field encountered. Its migration could then become diffusive, or be limited only to small scales.

3. Stopping Type II Migration

3.1. Stopping Type II Migration at Small Radii

Like cores undergoing type I migration, planets subject to type II migration would stop before falling onto the star if they entered a magnetospheric cavity. Tidal interaction with a rapidly rotating star would also halts planet orbital decay at a few stellar radii (where the interaction becomes significant; see Lin et al., 2000 and references therein). Both of these mechanisms have been put forth to account for the present location of the planet around 51 Pegasi (Lin et al., 1996), and to explain the location of hot Jupiters more generally.

A planet overflowing its Roche lobe and losing part of its mass to the central star would also halt at small radii. This is because during the transfer of mass the planet moves outward to conserve the angular momentum of the system (Trilling et al., 1998). The planet stops at the location where its physical radius is equal to its Roche radius. Recent observations of atomic hydrogen absorption in the stellar Lyman α line during three transits of the planet HD209458b suggest that hydrogen atoms are escaping the planetary atmosphere (Vidal-Madjar et al., 2003).

3.2. "The Last of the Mohicans"...

The mechanisms reviewed above cannot account for the presence of giant planets orbiting their parent star at distances larger than about a tenth of an au.

It has been suggested that migration of a giant planet could be stopped at any radius if migration and disk dissipation were concurrent (Trilling et al., 1998, 2002). Note that a massive planet can suffer significant orbital decay only if the mass of gas in its vicinity is comparable to the mass of the planet itself. If the disk is significantly less massive, there is not enough gas around the planet to absorb its angular momentum, and migration is slown down (Ivanov et al., 1999). If the disk dissipates while migration is taking place, then the drift timescale may increase in such a way that the planet stalls at some finite radius. Note however that this requires very fine tuning of the parameters (disk mass, disk lifetime etc.), as for a given disk mass the migration timescale decreases as the orbital radius decreases (see Eq. 2 above). Also, a major problem with this mechanism is to explain how

the disk dissipates. Within this scenario, there is initially enough gas in the disk to push the planet down to some orbital radius. For typical disk parameters, only part of this gas may be accreted by the planet or leak through the gap to be accreted onto the star (Bryden et al., 1999, Kley, 1999). It is therefore not clear how the gas disappears.

A giant planet could survive if after it formed there were not enough material left in the disk for significant migration to occur. It has been suggested that a series of giant planets could actually assemble in the disk and disappear onto the star (Gonzalez, 1997, Laughlin and Adams, 1997). Then at some point the disk mass may be such that one more planet can be formed but not migrate (Lin, 1997). This survivor is sometimes refereed to as the last of the Mohicans.

3.3. Planets Locked in Resonances

Within the context of several planets forming in a disk, another scenario has been suggested to occur when the migration of the innermost planet is stopped by either the star's tidal barrier or a magnetospheric cavity. In that case, a second planet approaching the star would stop when entering a low order resonance with the innermost planet, i.e. when the mean motions of the two planets become commensurate. As shown by Goldreich (1965) in the context of our Solar system, such commensurabilities are stable because angular momentum is secularly transferred between the different objects in just the correct proportion to keep the mean motions commensurate. In the context of the planetary system discussed here, the angular momentum would be transferred from the central star to the innermost planet then to the ring of material trapped between this planet and the next one, then from this ring to the next planet, and so on until the angular momentum is transferred to the disk outer part. The evolution of the central star may cause the whole system to migrate either inward or outward, but the planets remain locked into the resonances (Lin, 1997).

Like with the scenario discussed in the previous subsection, a major problem here is to explain how the disk material trapped in between the different planets eventually disappears.

Resonant trapping of planets can also lead to a reversal of type II migration. This happens when a giant planet (e.g. Jupiter) migrating inward captures into the 2:3 resonance a lighter outer giant planet (e.g. Saturn, Masset and Snellgrove, 2001). The gaps that the two planets open in the disk then overlap, and the imbalance between the torque exerted at Jupiter's inner Lindblad resonance and that exerted at Saturn's outer Lindblad resonance causes the whole system to migrate outward. This outward migration is accompanied by an increased mass flow through the overlapping gaps, as the angular momentum gained by the planets is lost by the disk.

4. Conclusions

We have reviewed the different mechanisms which have been proposed to slow down, halt or reverse inward orbital migration. When the interaction between the disk and the planet(s) is linear, orbital decay can be stopped if a magnetic field is present in the disk or if the planets are on sufficiently eccentric orbits. For a non-linear interaction, no general mechanism has been shown to prevent orbital decay of an isolated planet. Whether type II migration occurs or not may just depend on the mass of gas left in the disk after the planet forms.

References

Bell, K. R., D. N. C. Lin, L. W. Hartmann and S. J. Kenyon: 1995. 'The FU Orionis outburst as a thermal accretion event: Observational constraints for protostellar disk models', *Astroph. J.* **444**, 376–395.

Bodenheimer,P., O. Hubickyj and J. J. Lissauer: 2000, 'Models of the in situ formation of detected extrasolar giant planets', *Icarus* **143** 2–14.

Bryden, G., X. Chen, D. N. C. Lin, R. P. Nelson and J. C. B. Papaloizou: 1999, 'Tidally induced gap formation in protostellar disks: gap clearing and suppression of protoplanetary growth', *Astroph. J.* **514**, 344–367.

Goldreich, P.: 1965, 'An explanation of the frequent occurence of commensurable mean motions in the solar system', *M.N.R.A.S.* **130**, 159–181.

Goldreich, P. and S. Tremaine: 1979, 'The excitation of density waves at the Lindblad and corotation resonances by an external potential', *Astroph. J.* **233**, 857–871.

Goldreich, P. and S. Tremaine: 1980, 'Disk–satellite interactions', *Astroph. J.* **241**, 425–441.

Gonzalez, G.: 1997, 'The stellar metallicity–giant planet connection', *M.N.R.A.S.* **285**, 403–412.

Ikoma, M., H. Emori and K. Nakazawa: 2001, 'Formation of giant planets in dense nebulae: critical core mass revisited', *Astroph. J.* **553**, 999–1005.

Ivanov, P. B., J. C. B. Papaloizou and A. G. Polnarev: 1999, 'The evolution of a supermassive binary caused by an accretion disc', *M.N.R.A.S.* **307**, 791–891.

Kley, W.: 1999, 'Mass flow and accretion through gaps in accretion disks', *M.N.R.A.S.* **303**, 696–710.

Laughlin, G. and F. C. Adams: 1997, 'Possible stellar metallicity enhancements from the accretion of planets', *Astrophys. J.* **491**, L51–L54.

Lin, D. N.C.: 1997, 'Planetary formation in protostellar disks', In *Accretion phenomena and related outflows* , eds. D. T. Wickramasinghe, G. V. Bicknell, and L. Ferrario (ASP Conf. Series, vol. 121), p. 321–330.

Lin, D. N. C., P. Bodenheimer and D. C. Richardson: 1996, 'Orbital migration of the planetary companion of 51 Pegasi to its present location', *Nature* **380**, 606–607.

Lin, D. N. C. and J. C. B. Papaloizou: 1979, 'On the structure of circumbinary accretion disks and the tidal evolution of commensurable satellites', *M.N.R.A.S.* **188**, 191–201.

Lin, D. N. C. and J. C. B. Papaloizou: 1993, 'On the tidal interaction between protostellar disks and companions', In *Protostars and Planets III* , eds. E. H. Levy, and J. I. Lunine (Tucson: University of Arizona Press), p. 749–835.

Lin, D. N. C., J. C. B. Papaloizou, C. Terquem, G. Bryden and S. Ida: 2000, 'Orbital evolution and planet–star tidal interaction', In *Protostars and Planets IV* , eds. V. Mannings, A. P. Boss, and S. S. Russell (Tucson: University of Arizona Press), p. 1111.

Malhotra, R.: 1993, 'The origin of Pluto's peculiar orbit', *Nature* **365**, 819.

Masset, F. and J. C. B. Papaloizou: 2003, 'Runaway migration and the formation of hot Jupiters', *Astrophys. J.* **588**, 494–508.

Masset, F. and M. Snellgrove: 2001, 'Reversing type II migration: resonance trapping of a lighter giant protoplanet', *M.N.R.A.S.* **320**, L55–L59.

Murray, N., Hansen, B., Holman, M. and Tremaine, S.: 1998, 'Migrating Planets', *Science* **279**, 69–72.

Papaloizou, J. C. B.: 2002, 'Global $m = 1$ modes and migration of protoplanetary cores in eccentric protoplanetary discs', *Astron. Astrophys.* **388**, 615–631.

Papaloizou, J. C. B. and J. D. Larwood: 2000, 'On the orbital evolution and growth of protoplanets embedded in a gaseous disc', *M.N.R.A.S.* **315**, 823–833.

Papaloizou, J. C. B. and D. N. C. Lin: 1984, 'On the tidal interaction between protoplanets and the primordial solar nebula. I – Linear calculation of the role of angular momentum exchange', *Astroph. J.* **285**, 818–834.

Papaloizou, J. C. B. and C. Terquem: 1999, 'Critical protoplanetary core masses in protoplanetary disks and the formation of short–period giant planets', *Astroph. J.* **521**, 823–838.

Papaloizou, J. C. B. and C. Terquem: 2001, 'Dynamical relaxation and massive extrasolar planets', *M.N.R.A.S.* **325**, 221–230.

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

Shakura, N. I. and R. A. Sunyaev: 1973, 'Black holes in binary systems: Observational appearance', *Astron. Astroph.* **24**, 337–355.

Terquem, C. E. J. M. L. J.: 2003, 'Stopping inward planetary migration by a toroidal magnetic field', *M.N.R.A.S.* **341**, 1157–1173.

Terquem, C., J. C. B. Papaloizou and R. P. Nelson: 2000, ' Disks, extrasolar planets and migration', In *From Dust to Terrestrial Planets*, eds. W. Benz, R. Kallenbach, G. Lugmair and F. Podosek (ISSI Space Sciences Series, 9, reprinted from Space Science Reviews, 92), p. 323–340.

Trilling, D. E., W. Benz, T. Guillot, J. I. Lunine, W. B. Hubbard and A. Burrows: 1998, 'Orbital evolution and migration of giant planets: modeling extrasolar planets', *Astroph. J.* **500**, 428–439.

Trilling, D. E., J. I. Lunine and W. Benz: 2002, 'Orbital migration and the frequency of giant planet formation', *Astron. Astroph.* **394**, 241–251.

Vidal-Madjar, A., A. Lecavelier des Etangs, J.–M. Désert, G. E. Ballester, R. Ferlet, G. Hébrard and M. Mayor: 2003, 'An extended upper atmosphere around the extrasolar planet HD209458b', *Nature* **422**, 143–146.

Ward, W. R.: 1986, 'Density waves in the solar nebula – Differential Lindblad torque', *Icarus* **67**, 164–180.

Ward, W. R.: 1997a, 'Protoplanet Migration by Nebula Tides', *Icarus* **126**, 261–281.

Ward, W. R.: 1997b, 'Survival of Planetary Systems', *Astroph. J.* **482**, L211–L214.

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

Part V

Missions

The Kepler Mission

A Transit-Photometry Mission to Discover Terrestrial Planets

William J. Borucki and David Koch
NASA Ames Research Center, Moffett Field, CA

Gibor Basri
University of California, Berkeley, CA

Timothy Brown
High Altitude Observatory, UCAR, Boulder, CO

Douglas Caldwell and Edna DeVore
SETI Institute, Mountain View, CA

Edward Dunham
Lowell Observatory, Flagstaff, AZ

Thomas Gautier
Jet Propulsion Laboratory, Pasadena, CA

John Geary
Harvard Smithsonian Center for Astrophysics, Harvard, MA

Ronald Gilliland
Space Telescope Science Institute, Baltimore, MD

Alan Gould
Lawrence Hall of Science, University of California, Berkeley, CA

Steve Howell
WIYN Observatory, Tucson, AZ

Jon Jenkins
SETI Institute, Mountain View, CA

Abstract. The *Kepler* Mission is a NASA Discovery-class mission designed to continuously monitor the brightness of 100,000 main sequence stars to detect the transit of Earth-size and larger planets. It is a wide field of view photometer with a Schmidt-type telescope and an array of 42 CCDs covering the 100 sq. degree field-of-view (FOV). It has a 0.95 m aperture and a 1.4 m primary and is designed to attain a photometric precision of 20 parts per million (ppm) for 12^{th} magnitude solar-like stars for a 6.5-hour transit duration. It will continuously observe 100,000 main sequence stars from 9^{th} to 15^{th} magnitude in the Cygnus constellation for a period of four years with a cadence of 4 measurements per hour. *Kepler* is Discovery Mission #10 and is on schedule for launch in 2007 into heliocentric orbit. A ground-based program to classify all 450,000 stars brighter than 15^{th} magnitude in the FOV and to conduct a detailed examination of a subset of the stars that show planetary companions is also planned. Hundreds of Earth-size planets should be detected if they are common around solar-like stars. Ground-based spectrometric observations of those stars with planetary companions will be made to determine the dependences of the frequency and size of terrestrial planets on stellar characteristics such as type and metallicity. A null result would imply that terrestrial planets are rare.

1. Introduction

Small rocky planets at orbital distances from 0.9 to 1.2 AU are more likely to harbor life than the gas giant planets that are now being discovered with the Doppler-velocity technique. Technology based on transit photometry can find smaller, Earth-like planets that are a factor of several hundred times less massive than Jupiter-like planets. The *Kepler* Mission is designed to discover hundreds of Earth-size planets in and near the habitable zone (HZ) around a wide variety of stars. *Kepler* was selected as NASA Discovery Mission #10 in December 2001. A description of the mission and the expected science results are presented.

2. Scientific Goals

The general scientific goal of the Kepler Mission is to explore the structure and diversity of planetary systems with special emphasis on determining the frequency of Earth-size planets in the HZ of solar-like stars. This is achieved by surveying a large sample of stars to:

- Determine the frequency of 0.8 Earth-radii ($R_{\oplus}$) and larger planets in or near the habitable zone of a wide variety of spectral types of stars;
- Determine the distributions of sizes and orbital semi-major axes of these planets;
- Estimate the frequency of planets orbiting multiple-star systems;
- Determine the distributions of semi-major axis, eccentricity, albedo, size, mass, and density of short period giant planets;
- Identify additional members of each photometrically-discovered planetary system using complementary techniques; and
- Determine the properties of those stars that harbor planetary systems.

3. Photometer and Spacecraft Description

The instrument is a wide FOV differential photometer with a 100 square degree field of view that continuously and simultaneously monitors the brightness of 100,000 main-sequence stars with sufficient precision to detect transits by Earth-size planets orbiting G2 dwarfs. The brightness range of target stars is from visual magnitude 9 through 15. The photometer is based on a modified Schmidt telescope design that includes field flatteners near the focal plane. Figure 1 is a schematic diagram of the photometer. The corrector has a clear aperture of 0.95 m with a 1.4 m diameter $F/1$ primary. This aperture is sufficient to reduce the Poisson noise to the level required to obtain a 4σ detection for a single transit from an Earth-size planet transiting a 12^{th} magnitude G2 dwarf. The focal plane is composed of forty-two 1024×2200 backside-illuminated CCDs with 27 μm pixels.

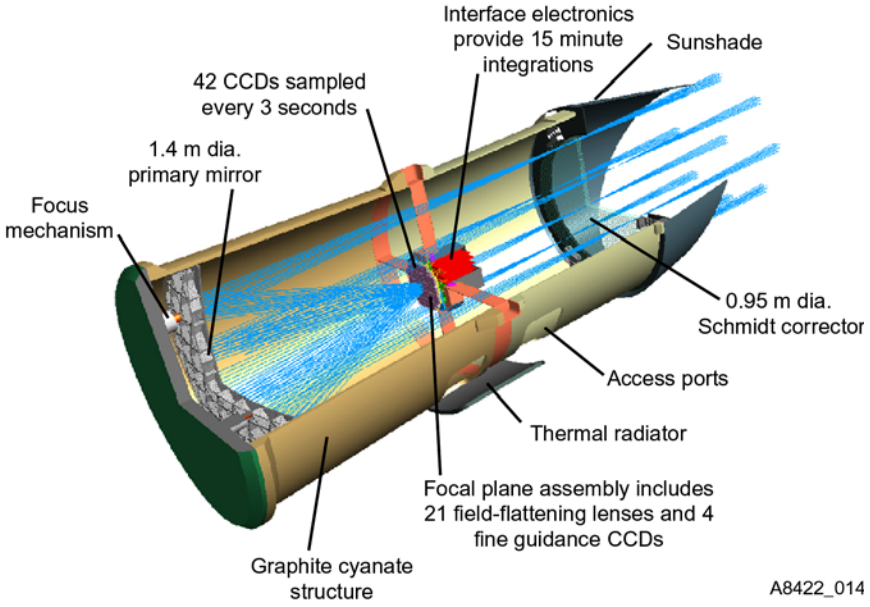

Figure 1. Schematic diagram of Kepler photometer.

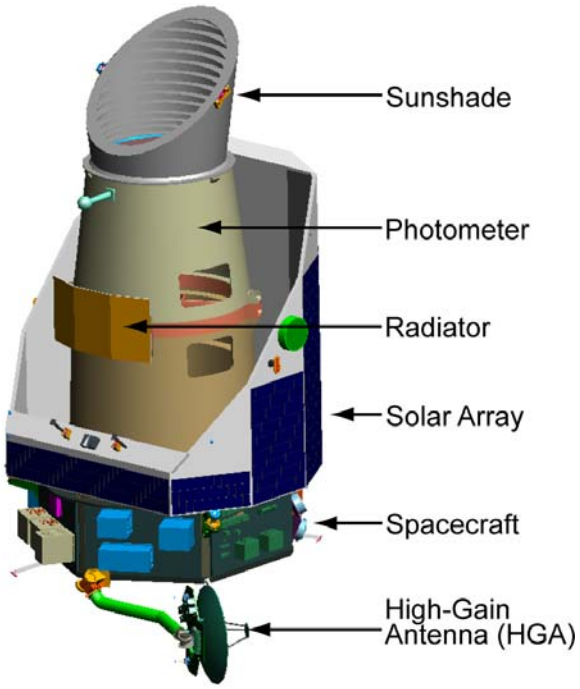

Figure 2. Integrated spacecraft and photometer.

The detector focal plane is at prime focus and is cooled by heat pipes that carry the heat out to a radiator in the shadow of the spacecraft. The low-level electronics are placed immediately behind the focal plane. A four-vane spider supports the focal plane and its electronics and contains the power- and signal-cables and the heat pipes.

The spacecraft bus encloses the base of the photometer, supports the solar panels, and provides attachment for power, communication, and navigation equipment. Several antennas with different frequency coverage and gain patterns are available for uplink commanding and for data downlink. A steerable high-gain antenna operating at Ka band is used for high-speed data transfer to the Deep Space Network (DSN). It is the only articulated component other than the ejectable cover. Approximately 1 GByte/day of data are recorded and then transferred to the ground every few days when contact is made with the DSN. The spacecraft provides very stable pointing using four fine guidance sensors mounted in the photometer focal plane. Small thrusters are used to desaturate the momentum wheels. Sufficient expendables are carried to extend the mission to six years.

Both the instrument and the spacecraft are being built by the Ball Aerospace and Technology Corporation (BATC) in Boulder, Colorado. (See Figure 2.)

4. Scientific Approach

To achieve the required photometric precision to find terrestrial-size planets, the photometer and the data analysis system must be designed to detect the very small changes in stellar flux that accompany the transits. In particular, the variability of the star, on time scales different than that of the transits, is not of interest. Sharp images of the star field are not helpful because the PSF must be over-sampled and because a broad image profile reduces the sensitivity to image motion. This provides the best estimates of image centroids that are used to reduce the systematic error due to image drift. The Kepler Mission approach is best described as "differential relative photometry". In this approach;

- Target stars are always measured relative to the ensemble of similar stars on the same part of the same CCD and read out by the same amplifier.
- Only the time change of the ratio of the target star to the ensemble is of interest. Only decreases from a trend line based on a few times the transit duration are relevant (long-term stability of the trend is not required).
- Target star and ensemble stars are read out every three seconds to avoid drift and saturation
- A broad PSF is used to avoid saturated pixels and to allow image centroids to be tracked and used to reduce systematic errors.
- Correction for systematic errors is critical. See the description in Jenkins et al. (2000).

Photometry is not done on the spacecraft. Instead, all of the pixels associated with each star image are sent to the ground for analysis. This choice allows many different approaches to be used to reduce systematic errors.

The spacecraft is placed in an Earth-trailing heliocentric orbit by a Delta II 2925-10L launch vehicle. The heliocentric orbit provides a benign thermal environment to maintain photometric precision. It also allows continuous viewing of a single FOV for the entire mission without the Sun, Earth or Moon obtruding. Only a single FOV is monitored during the entire mission to avoid missing transits.

A pattern of at least three transits that shows that the orbital period repeats to a precision of at least 10 ppm and that shows at least a 7σ detection is required to validate any discovery. The mission lifetime of four years allows four transits to be observed so that 8σ detections can be obtained from Earth-size planets transiting solar-like stars. A detection threshold of 7σ is required to avoid false positives due to random noise.

Planetary signatures exhibiting mean detection statistics of 7σ will be recognized 50% of the time. The recognition rate for planets exhibiting 8σ detection statistics increases to 84%. This increase means that the false negative event rate falls from 50% to 16%. Signal detection algorithms that whiten the stellar noise, fold the data to superimpose multiple transits, and apply matched filters are employed to search for the transit patterns down to the statistical noise limit (Jenkins, 2002). From measurements of the period, change in brightness and known stellar type, the planetary size, the semi-major axis and the characteristic temperature of the planet can be determined.

Only data from pixels illuminated by preselected target stars (i.e., $\backsim$ 20 pixels/star) are saved for transmission to Earth. Data for each pixel are co- added onboard to produce one brightness measurement per pixel per 15-minute integration.

Data for target stars that are monitored for p-mode analysis are measured at a cadence of once per minute. In the Sun, a series of modes with periods of about 3 minutes and equal in spacing in the frequency domain are excited to a level of about 3 ppm in white light. This level of precision requires the detection of at least 10^{12} photons. *Kepler* provides the necessary photon-electron levels in one month for 3600 dwarf stars brighter than $m_v = 11.4$ in the FOV.

5. Selection of Target Stars and Field of View

Approximately 100,000 target stars must be monitored to get a statistically meaningful estimate of the frequency of terrestrial planets in the HZ of solar-like stars. A FOV centered on a galactic longitude of 70 deg and latitude of + 6 deg satisfies both the constraint of a 55 degree sun-avoidance angle and provides a very rich star field. This FOV falls within the Cygnus constellation and results in looking along the Orion spiral arm (see Figure 3). In the 100 sq degree Kepler FOV, there are

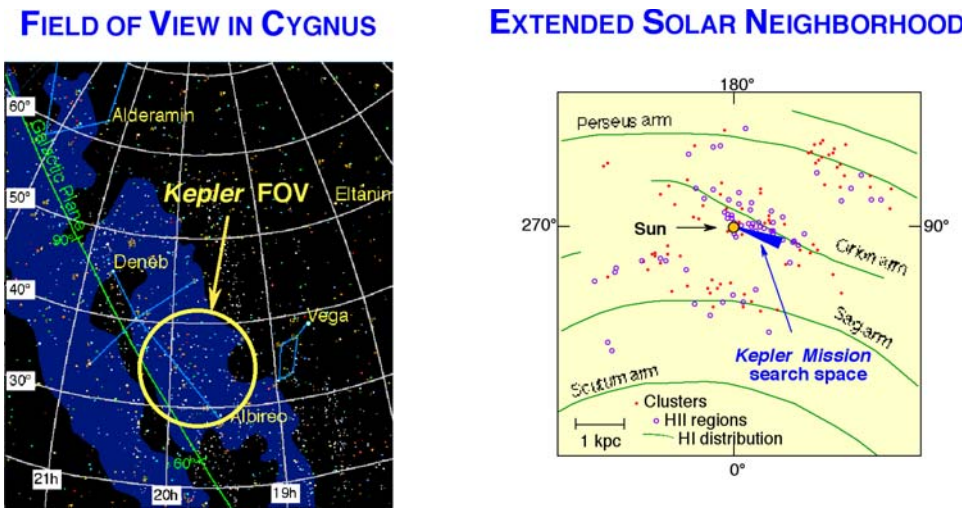

Figure 3. *Kepler* FOV in the Cygnus constellation, looking along the Orion spiral arm.

approximately 450,000 stars brighter than 15th magnitude. Studies are underway to determine the most efficient way of classifying such a large number of stars and choosing the approximately 100,000 late-F, G, and K dwarfs as targets. The current approach is to use wide FOV ground-based telescopes with a new color-filter system to identify both the luminosity class and spectral types and thereby enable the elimination of giants and early spectral types from the target list.

If 100,000 solar-like stars are monitored and if every such star has a single planet with an orbital semi-major axis of 1 AU, then only about 500 planets can be discovered because the geometrical probability that the planets' orbit will be aligned well enough to show transits is only about 0.5%.

6. Expected Results

After the mission ends, the results can be summarized as graphs analogous to those shown in Figure 4. Because both the size and mass of the stars that are found to have planets will be determined by follow-up observations, the size of the planets and their orbital semi-major axes can be determined. If most stars have planets approximately the size of the Earth ($0.9 < R_{\oplus} < 1.2$,) for example, then we should find data points along the curve marked "Earth-size". At distances near 1 AU ($0.8 < a < 1.2$ AU), we expect approximately 25 planets. If most stars have two such planets (like Earth and Venus) in that region, then there will be a point showing that 50 planets were detected there. If stars often have planets 30% larger, then because such planets are more readily detected, points along the curve marked "1.3 $R_{\oplus}$" will be recorded and a data point for 200 planets will be plotted for a semi-major axis near 1 AU. Planets twice the radius of the Earth (i.e.,

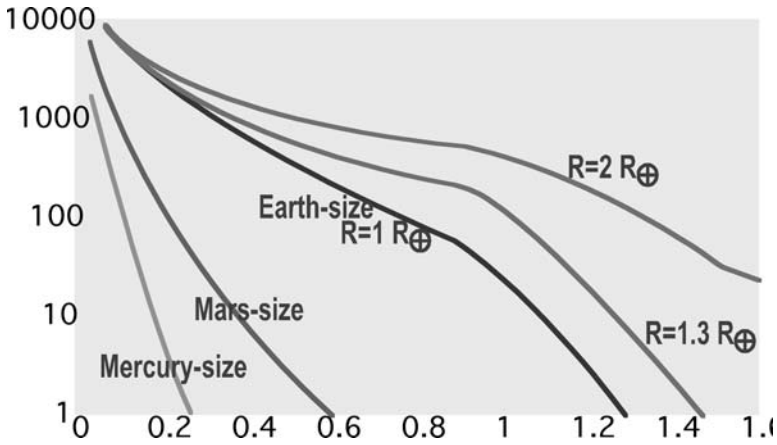

Figure 4. Trace of the number of planets expected to be detected if all stars have one planet at the position indicated. Note the high sensitivity of the results to the planet diameter and the orbital semi-major axis.

approximately 10 times the mass of the Earth) are readily detected even for large or dim stars so about 600 planets should be detected near 1 AU.

These values are based upon a realistic distribution of spectral types for a magnitude-limited survey. For a limiting visual magnitude of 14, about 50% of the stars will be early types that are too big to produce a SNR$>$ 7 for an Earth-size planet and/or too massive to produce a pattern with at least three transits during the four-year mission duration. Studies are underway to find somewhat dimmer stars of later spectral type that would have sufficiently large signals to overcome the increased photon shot noise. If there is an efficient way to find such stars, they will replace many of the brighter-but-earlier-spectral-type stars. This change could nearly double the number of small planets found with large values of semi-major axes.

As the semi-major axis decreases, the expected number of discoveries rises very rapidly, even on a log plot. This occurs for two reasons. First, the probability of a detection increases as the inverse of the orbital radius and second, the number of transits that occur during the mission lifetime increases. The detectability of transit patterns rapidly increase with the signal-to-noise ratio (SNR) and the SNR increases with the square root of the number of transits. At the value of the semi-major axis found for the "hot Jupiters" detected by the Doppler velocity technique, i.e., $\backsim$ 0.05 AU, tens of thousands of planets should be found if they are common orbiting solar-like stars. Even if the planets are as small as Mars or Mercury, the number of transits that occur in such tight orbits during a four year mission will be so large ($\backsim$ 400) that small planets should be found in profusion.

Of course, there is no reason to expect that all solar-like stars have planets and therefore the actual data points are likely to fall well below the curves shown in

Figure 4. Nevertheless, if planetary frequencies are as high as 1%, many terrestrial-size planets should be found.

After two years of operation, the *Kepler* Mission should provide a good estimate of the frequency of Earth-size planets with orbital periods as long as six months. Once the mission length reaches four years, a good estimate of that frequency will be obtained for planets in the HZ of solar-like stars. However, because not all stars are expected to have Earth-size planets in the HZ, and because the HZ could extend to the orbit of Mars (Kasting et al., 1993; Franck et al., 2000), it is worth considering the benefit of an extension of the mission from the planned four years to six years. The uppermost curves in Figure 5 show the dependence of the expected number of discoveries of Earth-size planets on the mission duration and semi-major axis. Clearly, the fractional increase in planet discoveries is dramatic for semi-major axes of 1 AU and larger. The increased mission duration raises the number of expected detections by a factor of two at 1 AU and a factor of 3.5 at 2 AU. If Earth-size planets are rare, this increase could be very helpful in providing an estimate of the number of stars that must be observed by the Darwin/TPF mission if it is to fulfill its goal of determining the atmospheric composition of Earth-size planets in the HZ of solar-like stars.

Planets as small as Mars might be habitable if they are well placed in the HZ. Such planets can be detected if the stars are smaller than the Sun or if the planets are in short period orbits. K stars are smaller and cooler than the Sun and have their HZ at distances of 0.2 to 1.0 AU. Hence the orbital periods in the HZ are measured in months and the total SNR of Mars-size planets is sufficient for valid detections. Again note the dramatic increase in the discovery result as the mission duration increases. At a semi-major axis of 0.5 AU, no planets are expected if the mission duration is 3 years. Two are expected for a 4-year mission and six are expected when the mission duration is increased to six years. It is clear from the lower curves, that knowledge of the lower end of the size distribution of rocky planets will be significantly improved by increasing the mission duration to six years.

Giant planets, like 51 Pegasi, with orbits of less than seven days are also detected by the periodic phase modulation of their reflected light without requiring a transit (Borucki et al., 1997; Hatzies, 2003). For the short-period giant planets that do transit, the planetary albedo can be calculated. Information on the scattering properties of the planet's atmosphere can also be derived (Marley et al., 1999; Sudarsky et al., 2000; Seager et al., 2000).

Ground-based Doppler spectroscopy and/or space-based astrometry with SIM can be used to measure the planetary masses, if they are jovian or larger, and to distinguish between a planet and a brown dwarf. These complementary methods can also detect additional massive companions in the systems to better define the structure of each planetary system. The density of any giant planet detected by both photometry and either of the other methods can be calculated.

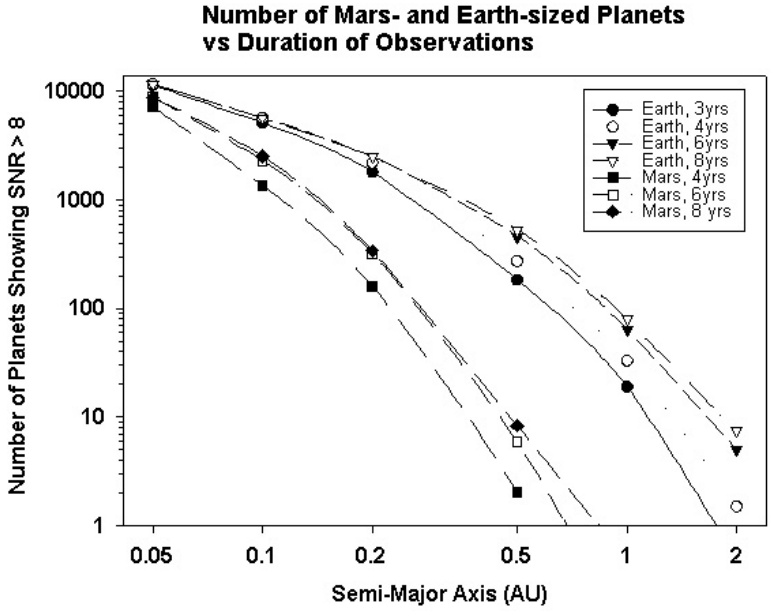

Figure 5. Expected number of planets detected versus mission duration and semi-major axis. The upper curves are for Earth-size planets while the lower curves are for Mars-size planets.

Thus, the results from this mission will allow us to place our solar system within the spectrum of planetary systems in our Galaxy and develop theories based on empirical data by providing a statistically robust census of the sizes and orbital periods of terrestrial and larger planets orbiting a wide variety of spectral types.

White dwarf stars are about the size of the Earth and might be expected to produce a transit signal of similar magnitude. However, because of the gravitational lensing caused by their large-but-compact mass, the transits actually result in an increase in brightness (Sahu and Gilliland, 2003) and are thereby readily distinguished from those of a planet.

7. Influence of Stellar Variability on the Detectability of Transit Signals

Stellar variability sets the limit to the minimum size of planet that can be detected. It reduces the signal detectability in two important ways;

- The variability introduces noise into the detection passband and thereby reduces the signal to noise ratio (SNR) and thus the statistical significance of transits.
- Because the flux of every target star is ratioed to the fluxes of many surrounding stars to reject common-mode instrument noise, variability of the stars used in the normalization introduces noise into the target star signal.

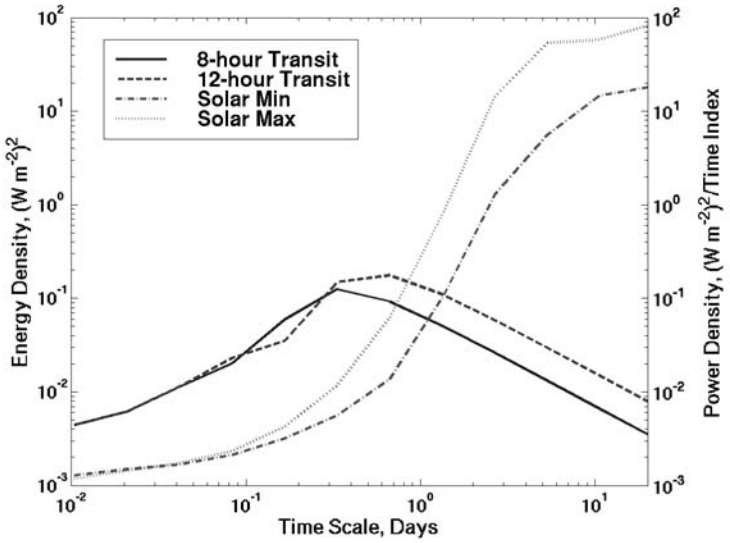

Figure 6. Power spectra of solar variability at Solar Maximum and Minimum and energy spectra of 8-hr and 10-hr transits.

The second concern can be alleviated by measuring the variability of each star relative to an ensemble of others and then iteratively removing the noisiest from the list of comparison standards. To mitigate the effects of the first concern, stars must be chosen that have low variability. Power spectra for the Sun at solar maximum and minimum are shown in Figure 6.

Also shown are the energy spectra for transits with 8- and 10-hour durations. It's clear that most of the solar variability is at periods substantially longer than those associated with planetary transits. In particular, the Sun's variability for samples with duration similar to that for transits is about 10 ppm. For stars rotating more rapidly than the Sun, the power spectrum will increase in amplitude and move to shorter periods thus increasing the noise in the detection passband.

Stellar variability in late-type main sequence stars is usually associated with the interplay of the convective layer and the internal magnetic field. Because the depth of the convective layer is a function of the spectral class of the star and because the activity level is higher when the star is rotating rapidly, the variability of solar-like main sequence stars is related to both their spectral class and rotation rate. Further, because the rotation rate decreases with age, the age of a star is an important variable. Thus we expect that the factors that influence the variability of target stars are age and spectral type.

The age and rotation rate of the Sun are approximately 5 Gyr and 27 days, respectively. The age of the galaxy is about 12 Gyr and about 2/3 of the stars are older than the Sun and are expected to be at least as quiet as the Sun. That extrapolation cannot be verified by examining the actual photometric variability

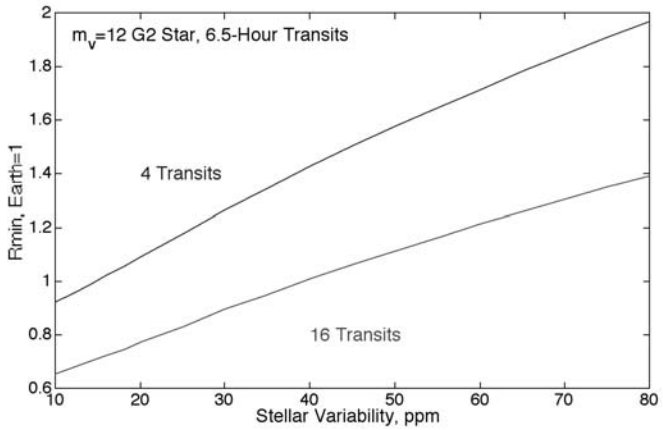

Figure 7. Effect of increased stellar variability on the minimum size planet that can be detected with 8σ.

of solar-like stars because no star other than the Sun has been measured to the requisite precision. However, the $R'_{\rm HK}$ index is believed to be well correlated to stellar variability. It is based on the spectral line profile of the Calcium H and K lines and is readily measured with ground-based telescopes. Figure 7 shows measurements of the $R'_{\rm HK}$ index for a variety of spectral types. Observations show that about 70% of the stars are found to have the index at least as low as that of the Sun which strengthens the argument that most solar-like stars will be at least as quiet as the Sun. Hence we plan to choose approximately 130,000 late type dwarfs to monitor during the first year of observations and then gradually eliminate those that are too variable to find Earth-size planets. This is necessary because the telemetry rate will decrease as the spacecraft recedes from the Earth.

Although most solar-like stars are expected to have low stellar activity levels, planets can still be found around stars with higher activity levels if the size of the planets are somewhat larger than the Earth of if they are found in the habitable zone of later spectral types. Planets in the HZ of K dwarfs have orbital periods of a few months and therefore would show about 16 transits during a four-year mission. Figure 7 shows the minimum size planet required to produce an 8σ detection versus the amplitude of the stellar variability assuming that the frequency distribution of the stellar noise is the same as that of the Sun. The upper curve shows that the amplitude of the stellar noise would need to be at least eight times that of the Sun before it would prevent planets slightly larger than twice the radius of the Earth from being detected. For planets showing 16 transits, planets as small as 1.4 times the radius of the Earth would still be detectable.

Table I. Comparison of Signal and Noise and Minimum Size Planet that Produces an 8σ Signal

Visual magnitude	9	12	14
Stellar signal (photo electrons)	9×10^{10}	6×10^{9}	9×10^{8}
Stellar shot noise (ppm)	3.3	13	35
Instrument noise (ppm)	1.7	7	30
Solar variability (ppm)	10	10	10
Relative signal for Earth transit across the Sun (ppm)	84	84	84
SNR for 4 transits	15	9	3.5
Minimum detectable planet radius (Earth=1) at 8σ	0.5	0.9	2.3

8. Comparative Importance of Various Sources of Noise

As described earlier, it is important to find planets with a size sufficient to produce 8σ detections with three or more transits. Equation 1 shows the relationship (greatly simplified) to the signal to noise ratio (SNR) as a function of ratio of the area of the planet to area of the star (Ap/A*), number of transits (Ntran), and the noise due to stellar variability (v), Poisson noise in the stellar flux (F), and instrument noise (i);

$$\mathrm{SNR} = (N_{\mathrm{tran}})^{1/2}(A_p/A_*)/[(i^2 + v^2) + 1/F]^{1/2}$$

An examination of this equation shows that for very bright stars where $1/F$ is small and when the stellar variability dominates the instrument noise, the SNR is dominated by the stellar variability. The opposite is true for dim stars where $1/F$ is large and the SNR is dominated by the shot noise in the stellar flux. Table I presents calculated noise values for *Kepler* instrument. For all entries in this table, a stellar variability of 10 ppm is assumed. Note that at 14th magnitude, that although the effect of stellar variability is negligible, instrument noise also makes a substantial contribution. Such large values of instrument noise warn that assuming shot-noise-limited performance is unwarranted. It is also clear that for stars 14^{th} magnitude and dimmer, detectable planets must be somewhat larger than the diameter of the Earth if they are to be reliably detected orbiting G2 dwarfs.

9. Validation of Planet Detections

Before a candidate detection can be considered to be a validated planet and the information released to the public, a rigorous validation process must be executed to ensure that it is not due to some other phenomenon (Jenkins et al., 2002). Public release of false positives would ultimately discredit any mission results. There-

fore to be considered a reliable detection the candidate planet detection must meet several requirements.

- The total SNR of the superimposed transits must exceed 7σ. This requirement prevents false positives produced by statistical noise when 8×10^{11} statistical tests are carried out on 10^5 stars for orbital periods from 1 to 700 days.
- At least three transits must be observed that demonstrate a period constant to 10 ppm. This test is independent of the previous test and demonstrates the presence of a highly periodic process. It essentially rules out mistaking stellar phenomenon for transits.
- The duration, depth, and shape of the light curve must be consistent. The duration must be consistent with *Kepler's* laws based on the orbital period. The depth must be consistent over all transits. A weaker requirement is that the shape must be consistent with a "U" shape of a planetary transit rather than a "V" shape of a grazing eclipse of a binary star. Clearly, low-amplitude transits are likely to be too noisy to make this distinction.
- The position of the centroid of the target star determined outside of the transits must be the same as that of the differential transit signal. If there is a significant change in position, the cause of the signal is likely to be an eclipsing star in the background.
- Radial velocity measurements must be conducted to demonstrate that the target star is not a grazing eclipsing binary.
- High precision radial velocity measurements must be made to measure the mass of the companion or provide an upper limit that is consistent with that of a small planet.
- High spatial resolution measurements must be made of the area immediately surrounding the target star to demonstrate that there is no background star in the aperture capable of producing a false positive signal.

It is also possible that future instrumentation on HST and JWST will have sufficient precision to detect the color changes during the transit. A measured color change consistent with the differential limb darkening expected of the target star (Borucki and Summers, 1984) would strengthen the validation. A shape or depth substantially different than expected would point to the possibility of a very close background star that differed in spectral type.

10. Mission Status

The *Kepler* Mission was chosen as Discovery Mission #10 in December of 2001 and funds to proceed were received to start the Phase B work in early 2002. An outline of the mission schedule is shown in Figure 8. Detectors and the optical components have been ordered. In the summer of 2002, a new management team from the Jet Propulsion Lab was chosen to provide overall mission management.

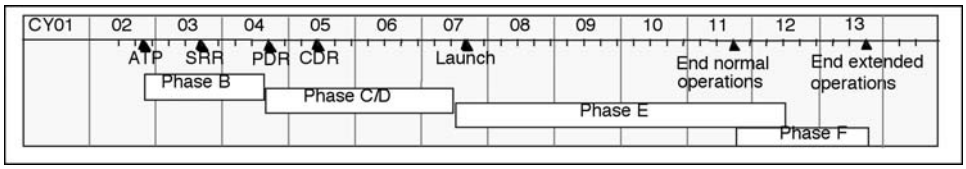

Figure 8. Outline of Kepler Mission schedule.

The JPL team members have been smoothly integrated with those at Ames and BATC. Their addition greatly strengthens the *Kepler* Mission by providing great depth in mission management and engineering.

In summary, the Kepler Mission is on schedule for an October 2007 launch.

References

Borucki, W.J., Summers, A.L.: 1984, The photometric method of detecting other planetary systems, *Icarus* **58**, 121–134.

Borucki, W. J., D. G. Koch, E. W. Dunham, and J. M. Jenkins.: 1997, The *Kepler* Mission: A mission to determine the frequency of inner planets near the habitable zone for a wide range of stars. Planets Beyond the Solar System and the Next Generation of Space Missions, *ASP Conference Series* **Vol. 119**, 153–173.

Franck, S., A. Block, W. von Bloh, C. Bounama, A. J. Schellnhuber, and Y Svirezhev.: 2000, Habitable zone for Earh-like planets in the solar system, *Plan. and Space Sci.* **48**, 1099–1105.

Hatzies, A. P.: 2003, Detecting Extrasolar planets in reflected light using COROT and *Kepler*, Scientific Frontiers in Research on Extrasolar Planets, *ASP Conference Series* **294**, pp. 523–527. D. Deming and S. Seager, eds.

Henry, T. J., D. R. Soderblom, R. A. Donahue, and S. L. Baliunas.: 1996, , *AJ* **111**, 439.

Jenkins, J.M.: 2002, The impact of stellar variability on the detectability of transiting terrestrial planets, *ApJ* **575**, 493–501.

Jenkins, J.M., Caldwell, D.A., and Borucki, W.J.: 2002, Some test to establish confidence in planets discovered by transit photometry, *ApJ* **564**, 495–507.

Kasting, J. F., D. P. Whitmire, and R. T. Reynolds.: 1993, , *Icarus* **101**, 108.

Marley M. S., C. Gelino, D. Stephens, J. I. Lunine and R. Freedman.: 1999, Reflected spectra and albedos of extrasolar giant planets. I. Clear and cloudy atmospheres. *ApJ* **513**, 879–893.

Sahu, K. C. and R. L. Gilliland.: 2003, Near-field microlensing and its effects on stellar transit observations by *Kepler*. *ApJ* **584**, 1042–1052.

Seager, S. , B. A. Whitney, and D. D. Sasselov.: 2000, Photometric light curves and polarization of close-in extrasolar giant planets, *ApJ* **540**, 504–520.

Sudarsky, D., A. Burrows, and P. Pinto.: 2000, Albedo and reflection spectra of extrasolar giant planets, *ApJ* **538**, 885–903.

The Exoplanet Program of the CoRoT Space Mission

C. Moutou and the CoRoT Team
Laboratoire d'Astrophysique de Marseille, Traverse du Siphon, 13376 Marseille cedex, France

Abstract. CoRoT is a space mission devoted to high precision star photometry. To be launched in 2005, CoRoT will use the transit method to discover new planets: likely some highly-coveted terrestrials and, quite certainly, a large number of close-in giants. Our knowledge on these "hot" planets will encounter a significant breakthrough due to a strong sample enlargement and to the possibility to measure the planet mass from ground-based radial velocity follow-up. Here, we present in some details the planet detection capability of CoRoT. Planetary radii will be measured down to Uranus size around solar type stars. Orbital periods will be determined from three detected transits up to 75 days. The possibility to detect more than one planet in a system is also addressed.

1. Introduction

CoRoT is a CNES space mission with 20% European participation. Half of its field-of-view is devoted to the detection of extra-solar planets using the transit method. The instrument is a 27 cm afocal telescope with a wide-field camera, optimized for accurate photometric measurements. Besides planet searches, CoRoT will also detect stellar oscillations (Baglin et al., 2001) and will produce new results on the photometric variability of the main-sequence stars, a knowledge that transit detection will also benefit.

Within a field of view of 3.9 square degrees CoRoT will monitor simultaneously up to 12,000 stars during five time slots of 150 day duration. During the 2.5 year satellite lifetime, a total of 60,000 light curves will be produced at a sampling rate of 8 minutes. In addition, after each long observation run the pointing direction of the satellite will have to be changed enabling 1 or 2 short observation runs of 20 days. During these short runs CoRoT will observe other sky fields within a few degrees of the main ones, so that some 120,000 additional light curves will be also produced. The range of magnitude of the CoRoT's exoplanet targets is 11.5–16. For the brightest stars color information will be also available and will help discriminating single transit events from stellar artifacts.

The photometric accuracy for the exoplanet targets will be of the order of a few 10^{-4} (section 3) in one hour integration time and for a star with $m_v = 14$. The different kinds of planets accessible by CoRoT are presented below.

2. Close-in Giant Planets

A significant breakthrough in our knowledge of this giant planet family is expected from the wide sample CoRoT will bring us back. Indeed, radial-velocity surveys

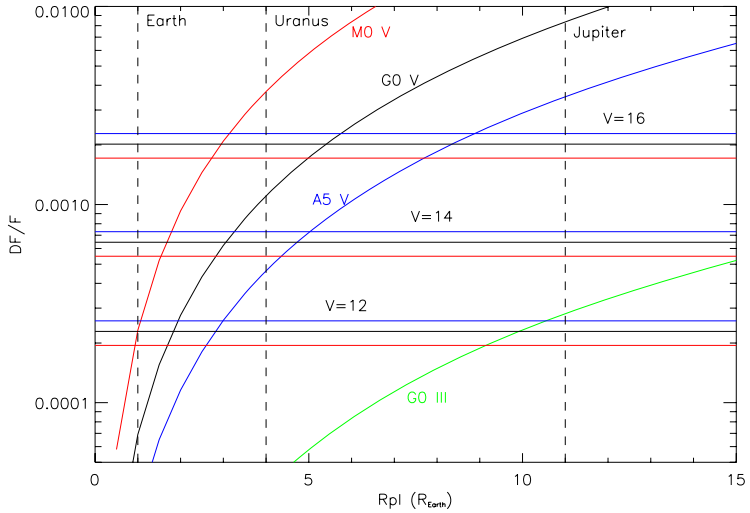

Figure 1. Relative depth of the transit versus planet radius (in Earth radii). The smallest planets correspond to the crossing of horizontal lines (CoRoT detection threshold) and curves, and depends on spectral type and magnitude of the star.

have revealed the existence of planets similar to the famous 51Peg b (Mayor and Queloz, 1995) and established the first statistics on this new kind of objects. Today, these are 17 (/102) planets separated by less than 0.1 AU from their star which are known. These objects are the easiest to detect from space with the transit method and CoRoT will be able to find their radius distribution. Furthermore, a RV follow-up of the stars identified by CoRoT to host a planet is of utmost importance. It will give access to: (i) a determination of the planet mass and density; (ii) a wide field of potential studies (internal structure of close-in giant planets, evolution theories, etc...). When we account for the 180,000 available light curves, and if we drop 50% of the targets due to their radius or binarity, up to 200 such planets (mass $> 0.1 M_{\rm Jup}$ and period < 10 days) are expected, following prediction based on the present RV statistics and the derived mass distribution (Mayor et al., this book).

Another way to study close-in giant planets with CoRoT would be to detect the periodic variations of the star light reflected by the planet and to estimate the planet albedo. Such a possibility has been explored for bright targets in the Seismology channel (Guillot and Vannier, personal comm.).

Although transit detection is easier for dwarf stars than for other spectral types, since transit depth scales as $1/R^2_{\rm star}$, CoRoT will be able to discover hot giant planets around giant stars of type earlier than G5 (Fig. 1) and, thus, will explore the occurrence of planet formation in a wide range of the HR diagram.

Table I. For an A5V, G0V, M0V, G0III star: the transit duration, the maximal star-planet distance (Period=75 days), the probability of correct inclination, the effective temperature range (Albedo=0.2 and 0.8) and the minimal planet radius for a 12 and 15 magnitude star. Jovian planets around G0 giants can be observed around bright stars, up to V=13, and have a large geometrical probability.

Star	Transit duration	Max sma (AU)	Proba. transit	$T_{\rm eff}$ (K)	V=12 R_E	V=15 R_E
A5V	9 hours	0.41	1.9%	600–860	2.4	6.0
G0V	7 hours	0.36	1.4%	370–520	1.7	4.0
M0V	5 hours	0.26	1.1%	160–220	1.0	2.5
G0III	25 hours	0.51	5.5%	680–960	10 ($0.9R_{\rm Jup}$)	–

3. Smaller Planets

As the observation runs will last at maximum 150 days, the longest planetary periods accessible by CoRoT will be around 75 days. Identifying longer periods, from single or dual transit events, cannot be excluded but will require confirmation by complementary observations. The orbital domain accessible by CoRoT is thus mainly inside 0.25 AU or 0.5 AU following the spectral type.

The radius of the smallest detectable planet can be estimated from blind transit detection in simulated light curves (Defaÿ et al., 2001). These light-curves account for instrumental noise, straylight from Earth, zodiacal light and also on estimate of the stellar variability, scaled from space observations of the Sun. This minimum planet radius depends on the spectral type and magnitude of the star, and on the star-planet distance. The performance level of planet detection by CoRoT is summarized in Table 1 and Figure 1. Note that detectable planets can be as small as the Earth if the parent star is a brigth, but very scarce, M0 dwarf. Preliminary spectral classification in the CoRoT fields permits an optimized target selection, storing up dwarf stars and avoiding giants. More generally, planetary radii range from Jupiter size to Earth size, while effective temperature usually remains above 400K. Note that two points are also converging for a strong impact of CoRoT: (i) the planet mass distribution from RV surveys indicates a large proportion of low-mass planet (close to a $1/M$ law); (ii) a RV follow-up of the CoRoT targets is possible with HARPS (Pepe et al., 2000), even for low-mass planets (Bouchy et al., 2003). CoRoT will thus answer the question whether hot terrestrial planets exist and survive.

More details on CoRoT exoplanet detection capabilities can be found in Defaÿ et al. (2001) and in Bordé et al. (2003) which includes a study on the importance of star population and of field crowding.

4. Planetary Systems

Among the planetary systems identified so far, 12 are known to have more than one planet. How CoRoT can contribute to discover multi-planet system? Let us examine various possible cases:

- The star has planets outside the orbit of the detected one. The probability for a transit by another planet drops as $1/a$. Here, note that, in the case of the Solar System, the dispersion in inclination is around 2 degrees (Pluto excluded). This relatively high value weakens the probability for a second detection when a first planet has been found. Of course the radial extent of the system is of importance in the problem. The probability for a second detection is not so bad at short distance from the star (i.e.) the favourite domain of CoRoT. In fact, the possibility to observe another planet outside the orbit of the first one usually relies on other methods: radial velocity for a giant planet up to a few AU, or direct imaging, interferometry or astrometry for giants at larger distances.

- The star has planets inside the orbit of the detected one. In this case, tidal effects probably stir up the orbital inclinations and the chances to get a transit from inner planets are no more very large. Some regions around the star could even be dynamically unstable (Erikson and Rauer, in prep.). Nevertheless, if such planets were present and massive enough, RV follow-up could detect those with the shortest period.

- No planet has been detected by CoRoT. The accurate light curve available with the CoRoT data will be used to estimate some important properties of the star such as: its variability level, its rotational velocity, some limits on the inclination and the masses of the potential planets. Further studies could still lead to the detection of a non-transiting planet!?

The combination of various methods for detecting extrasolar planets will be extremely useful to gather information on planetary systems other than the Solar System and to put further constraints on the planetary formation models.

References

Baglin, A., et al.: 2001, *IAU colloquium* **185**, Leuven.
Bordé, P., Rouan, D., and Léger, A.: 2003, *Astronomy and Astrophysics* **405**, 1137–1144.
Bouchy, F., et al., in *Scientific Frontiers in Research on Extrasolar Planets*.
Defaÿ, C., Deleuil, M., and Barge, P.: 2001, *Astronomy and Astrophysics* **365**, 330–340.
Mayor, M. and Queloz, D.: 1995, *Nature* **378**, 355.
Pepe, F., et al.: 2000, *Proc. SPIE* **4008**, 582–592.

Part VI

Working Group Reports

Report of the Working Group on Detection Methods

T. M. Brown
National Center for Atmospheric Research, 3450 Mitchell Lane, Boulder, CO 80307, USA

J.-R. De Medeiros
Departamento de Física, Universidade Federal do Rio Grande do Norte, 59072-970 Natal, RN., Brazil

Abstract. A group of about fifteen interested scientists spent two long sessions discussing issues related to detecting extrasolar planets, particularly ones of roughly terrestrial size. We arrived at several recommendations, principally: (1) Radial velocity measurement precision should be pushed to its limits, presumably those set by the astrophysics of stars. (2) Data policies for upcoming planet-finding space missions must be written so they do not encourage premature or mistaken discovery announcements. (3) Ground-based transit searches would benefit enormously if some adequately-funded institution would take the lead in fielding a coherent and well-engineered world-wide network of telescopes for this purpose. We should attempt to build international consensus and support for this idea.

1. Introduction

As a planned part of the workshop, about 15 interested participants gathered for two long meetings to discuss issues related to detecting extrasolar planets. We discussed a large range of topics but managed to stay fairly well focused, even while probing widely divergent views. Thus we came to understand a new meaning for the term "self-organized criticality." Discussions of our various sub-topics were necessarily loosely connected with each other, so I present them below as entirely separate summaries, in no particular order.

Group members were invited to contribute their own write-ups to this report. Only one person (Jose-Renan de Medeiros) did so; Section 6 contains his contribution. In the interest of compactness, I have merged our reference lists.

2. Confirmation of Detections (Especially Transits)

Since the beginning of searches for extrasolar planets, the problem of confirmation has been a difficult one. The indirect nature of most methods for detecting distant planets contributes to this uncertainty. Given a signal that may represent an extrasolar planet, how can one assure that it truly results from a planet, and not from Something Else? Several techniques can been applied to answer this question, depending upon the circumstances. A particularly thorny problem that we will soon face concerns transiting extrasolar planets, observed either from the ground

or (with much smaller signal levels) from spacecraft. How are we to confirm the planetary nature of detections from, say, Kepler or Eddington?

One approach is "internal confirmation". This depends on obtaining large quantities of data using a single observing technique, and making a convincing argument that the values obtained could only be produced by an orbiting planet. Sometimes it is possible to apply other consistency checks, such as comparisons against likely confounding variables, or generation of an ephemeris for the planet, followed by prediction of the results of future observations. This approach has been followed with varying degrees of success by planet-hunters old (van de Kamp, 1963) and new (Mayor and Queloz, 1995). In the case of transiting Earth-like planets, it is not a promising avenue because the lifetimes of the missions are only a factor of a few longer than the interesting orbital periods, because the transits occur only rarely (one cannot just obtain a new data point whenever one likes), and because transits carry no evidence for Keplerian (as opposed to merely periodic) behavior.

A better approach, when it is feasible, is "external confirmation". In the weak form of this, one obtains independent observations of the same kind, but with similar or improved accuracy. This approach discriminates against certain kinds of systematic errors, but not all. The strong form of external confirmation is to observe the same planet by different means altogether; this is clearly the desired approach, when it is possible.

In the context of transit searches, the group agreed that the most convincing kind of external confirmation would be through radial velocity measurements. (Astrometry is another valuable possibility, but it is most feasible for planets with large orbital semimajor axes, and hence requires observations spanning very long times.) The challenge in detecting a terrestrial-sized planet via radial velocity methods is formidable: Earth imposes a reflex velocity of only about 8 $\mathrm{cm\,s^{-1}}$ on the Sun, and the signal from Venus is not much bigger. On the other hand, we do not really know the ultimate limits on our ability to measure stellar Doppler shifts (i.e., limits set by processes in the stars, rather than in our instruments). The signal from terrestrial planets may be within reach, at least for some kinds of stars. Moreover, there may be other clues that bear on the existence of a planet and that are not so difficult to observe. For instance, (with our own solar system in mind) detection of an outer Jovian-mass planet would tend to confirm the reality of an inner, transiting, Earth-sized planet. This sort of confirmation is clearly within the grasp of current technology; it would just take time. Another route would be to seek the line-profile distortions (to lowest order, the Doppler shift) caused by the transiting planet passing across the disk of its rotating parent star. For an Earth-sized planet and a stellar equatorial rotation speed comparable to the Sun's, this signal would be about the same size (± 20 $\mathrm{cm\,s^{-1}}$) as the reflex Doppler shift. But its time scale would be hours, not months, so it would be correspondingly easier to separate from instrumental or stellar noise. These considerations led to the first major recommendation of the group:

• It is vital to continue to push the precision limits of radial velocity methods, until they are limited only by astrophysical processes. This should be done on all timescales, from hours to years.

Another concern raised within the group had to do with the way that detection projects (again, particularly transit-search space missions) will decide that a new planet has actually been found. For the credibility of the science and the continued health of the field, it is essential that nobody should announce false planets. This will be particularly true for the first few detections. Obviously, however, powerful forces will tend to subvert this ideal. The space missions, especially, are expensive and high-profile operations, and the pressure to produce results is going to be very large. Moreover, the timing of the US and European efforts suggests that a competition is likely, and only one announcement can be the first. Possibly worst of all, emerging space agency data policies may require that light curves be released to the world at large shortly after they are obtained. This practice would almost guarantee premature announcements, since many people would have much to gain and little to lose by declaring anything resembling a transit as evidence for a new Earth. Accordingly the group has a second recommendation:

• Data and publication policies must not encourage premature or mistaken planet detection announcements.

The community should think carefully about fair and effective ways to implement this recommendation.

Finally, and with a lesser degree of conviction, the group concluded that confirmation of transiting Earth-sized planets would be less problematic if Kepler and Eddington observe the same field in the sky, at least for part of their respective mission lifetimes.

3. Planets around Post-Main-Sequence Stars

We had a relatively short discussion about the prospects for locating planets circling post-main-sequence stars. We concluded that it is both possible and desirable to take some steps in this general direction.

David Latham presented some very interesting results concerning the rapid rotation rate of some giant stars. Their large angular momentum could be explained if, as they expanded following core hydrogen exhaustion, they had ingested a Jupiter-sized planet. In this case, one has evidence for a planet that is no longer there, which is not entirely satisfactory. The idea suggests connected lines of research, however. For instance, one might examine stars with large rotational velocities for radial velocity signals from surviving giant planets in larger orbits, and detailed chemical studies of the rapidly-rotating giants should help determine whether the excess angular momentum might have been dredged up from their cores during

their evolutionary process. A written contribution to this part of the discussion, by J. De Medeiros, follows this summary.

We discussed several possible strategies for identifying planets around white dwarfs. Astrometry is promising, since there is a good population of relatively nearby white dwarfs, and since their masses are rather small. White dwarf radial velocities are difficult to measure well, since their spectral lines are few and heavily pressure-broadened. We did, however, discuss the possibility that white dwarf pulsations (which are excellent clocks) might be used to measure the time-varying distance between the Earth and the star, in a fashion similar measuring Doppler shifts. The relatively long periods of white dwarf pulsations (on the order of 100 s) mean that measurement of their phases would have to be carried out with extraordinary accuracy, however. A third approach would be to seek the recombination spectrum resulting from ionization of a planetary atmosphere by ultraviolet radiation from a hot white dwarf (e.g., Chu et al., 2001).

Finally, we considered planets circling pulsars. Such planets were, after all, the first ones known outside our solar system (Wolszczan and Frail, 1992). That no other detections have followed the first one suggest that rather special circumstances are needed in order to produce planets around pulsars. The fact remains, however, that a systematic search for such planets has not been mounted, and it should be.

4. Other Detection and Characterization Methods

We spent some time considering several "blue-sky" techniques for detecting or characterizing extrasolar planets.

Some information about extrasolar planet atmospheres (for instance, something about cloud particle size distributions) might be learned by measuring the polarization signal from light reflected from a planet. This polarization signal would be extremely weak, however. Since the reflected light even from a close-in Jupiter-sized planet is only a few parts in 10^5 of the direct light from its star, and since the polarization imposed on the reflected light is at most a few percent, one would need precision of parts in 10^7 in order to obtain useful diagnostics. This is unreachable by present instruments, but the idea should be remembered for future reference.

In a similar vein, we wondered about the detectability of the extrasolar analogue of radio emission from Jupiter, caused by its interaction with the solar wind. For close-in planets the radiated intensity would presumably be larger than from Jupiter, because of the higher ram pressure exerted by the stellar wind at small orbital distances. None of us, however, had the expertise to estimate the resulting radio flux or its detectability. We therefore left this topic also for another day.

As a last idea, I presented the outline of a small space mission, the purpose of which would be to perform both reflected-light and transmission spectroscopy on known close-in extrasolar planets. The concept involved an imaging telescope with

about 0.5 m aperture, equipped with suitable detectors and grisms to give $R \approx 50$ spectral resolution on stellar targets, over a wavelength range of $0.35\ \mu\text{m} \leq \lambda \leq 2.5\ \mu\text{m}$. Such a system could measure albedos, phase functions, and the strengths of molecular features on all of the known planets with orbital semimajor axes smaller than 0.1 AU, and could perform useful transmission spectroscopy on (mostly as yet undiscovered) transiting planets down to about $V = 10$.

5. Organization

We talked briefly about the state of ground-based transit searches. These are not now as productive as they should be, because they tend to be poorly-funded and part-time projects, often using equipment scavenged from other programs. These characteristics impede efforts to maintain high duty cycles and efficient data analysis, and these failings in turn have a large effect on the number of promising transit candidates discovered. This discussion led to the group's third recommendation:

• **Some adequately-funded institution should take the lead in in fielding a coherent and well-engineered world-wide network of telescopes to identify extrasolar planet transits. We should attempt to build international consensus and support for this idea.**

Based on the example of the GONG network, which involves systems of similar size but considerably greater complication than we envisioned for a transit network, we guessed that the total cost to install such a network would be in the neighborhood of $10M.

6. The Lithium-Rich and the High-Rotator Giant Stars

Two puzzling events in observational stellar astrophysics in recent years have been the discovery of the lithium-rich single giant stars, some of them possessing surface lithium abundance approaching the cosmic value of the interstellar medium and young main-sequence stars (e.g. Wallerstein and Sneden, 1982; Brown et al., 1989; Reddy et al., 2002), and the single giant stars with enhanced rotation (De Medeiros and Mayor, 1999). The abnormally large lithium contents in cool giants are difficult to understand,since lithium is destroyed in main-sequence stars with deep subphotospheric convective envelope and is further depleted on the ascent of the red giant branch, where a deepening convective zone dilutes remaining lithium. Different scenarios have been proposed to explain such abnormal behavior in lithium: either fresh lithium synthesized in the so-called beryllium transport mechanism (following the mechanism proposed by Cameron and Fowler, 1971), or a preserved intrinsic lithium content (Brown et al., 1989); pollution from debris of nova-explosion of a white dwarf companion; engulfing of a brown dwarf into the

atmosphere of an evolving giant stars; and lithium production during helium-shell flashes (e.g. Gratton and D'Antona, 1989). The enrichment of the lithium envelope of the giant star by surrounding planetary material is also proposed by different authors (e.g. Brown et al., 1989; Livio and Soker, 2002). Single giant stars with enhanced rotation, essentially late-G and K type stars, present $V \sin i$ from about 10 to 50 km/s, which is substantially higher than the mean $V \sin i$ for these spectral regions (De Medeiros and Mayor, 1999; De Medeiros et al., 1996). One of the most exciting scenarios to explain the root-cause for such abnormal rotation is based on the engulfment of planets by giant stars on the evolution of late-type stars (Siess and Livio, 1999). Livio and Soker (2002) have estimated that at least 3.5% of the evolved solar-type stars will be significantly affected by the presence of planetary companions. In the broad sense, the study of lithium-rich giant and rapidly-rotating giants might represent a key problem to understand the evolution of giant stars hosting planets. In this context, among the new stellar properties to be analyzed it seems important to measure different abundances, carbon isotope ratios as well as additional multi-wavelength diagnostics to establish the nature of lithium-rich and high-rotator giant stars.

References

Brown, T., et al.: 1989, *ApJS* **71**, 293.
Cameron, A. G. W., and Fowler, W. A.: 1971, *ApJ* **164**, 111.
Chu, Y.-H., Dunne, B.C., Gruendel, R.A., and Bradner, W.: 2001, *ApJ* **546**, L61.
De Medeiros, J. R., and Mayor, M.: 1999, *A&AS* **139**, 433.
De Medeiros, J. R., Rocha, C., and Mayor, M.: 1996, *A&A* **314**, 499.
Gratton, R. G., and D'Antona, F.: 1989, *A&A* **215**, 66.
Livio, M., and Soker, N.: 2002, *ApJ* **571**, L161.
Mayor, M., and Queloz, D.: 1995, *Nature* **378**, 355.
Reddy, B. E., Lambert, D. L., Hrivnak, B. J., and Bakker, E. J.: 2002, *Astron. J.* **123**, 1993.
Siess, L., and Livio, M.: 1999, *MNRAS* **308**, 1133.
van de Kamp, P.: 1963, *AJ* **68**, 515.
Wallerstein, G., and Sneden, C.: 1982, *ApJ* **255**, 577.
Wolszczan, A. and Frail, D. A.: 1992, *Nature* **355**, 145.

Dynamical Interactions Among Extrasolar Planets and their Observability in Radial Velocity Data Sets

Report of Working Group IV

Gregory Laughlin
UCO/Lick Observatory, University of California, Santa Cruz, CA, 95064 USA

John Chambers
NASA Ames Research Center, Planetary Systems Branch, Mountain View, CA, 95032 USA

Abstract. For certain multiple-planet systems such as GJ 876 and 55 Cancri, which have been observed for a large number of orbital periods, and which have strong planet-planet gravitational interactions, the approximation that the planets are orbiting on independent Keplerian ellipses is inadequate. This phenomena is of immediate importance to the interpretation of the increasing number of known multiple-planet systems, and it can potentially be used to remove the $\sin(i)$ degeneracy. Multiple-planet fitting sparked a great deal of discussion during the working group session on planet-planet interactions. In this report, we provide a self-contained summary of the techniques and the inherent potential of self-consistent dynamical fits to interacting planetary systems.

1. Introduction

The steady accumulation of Doppler radial velocity (RV) observations of nearby stars has made it clear that many systems have more than one planet. The first extrasolar multiple-planet system was discovered around υ Andromedae (Butler et al., 1999), and in the past two years, pairs or triplets of planets have been detected around a number of additional stars, including GJ876 (Marcy et al., 2001), 47 UMa (Fischer et al., 2002), and 55 Cancri (Marcy et al., 2002). Within the current planet catalog, there are 11 multiple-planet systems containing 26 planets (Schneider, 2002). Additional planetary companions may ultimately be found around more than half of the stars with one detected planet (Fischer et al., 2001).

The presence of two or more interacting planets in a system dramatically increases our potential ability to constrain and understand the processes of planetary formation and evolution. Planet-planet interactions can be reasonably subdivided into three categories: (i) interactions during the planet formation and nebular phases, (ii) ongoing secular or resonant interactions which manifest themselves on observable timescales of decades or less, and (iii) long-term dynamical interactions that can sculpt a planetary system over timescales comparable to the lifetime of the parent star. Short-term dynamical interactions (category ii) are of particular interest because they have immediately observable consequences, and indeed, several multiple-planet systems are showing clear signs of ongoing interaction. In this working group report, we summarize the current situation regarding planet-planet interactions that are *directly observable* within RV data sets.

2. Non-Keplerian Interactions

When confronted with a radial velocity (RV) data set that contains the signature of more than one planet, one is faced with a rather tricky question: How are the individual motions of the planets to be separately pulled out of the composite curve? Isaac Newton, in his analysis of observations of our own solar system, was the the first to address the question of dynamical interactions among multiple planets. In Book I, Section II, proposition 69, of the Principia (Newton, 1687), he writes:

> "And hence, if several lesser bodies revolve about a greatest one, it can be found that the orbits described will approach closer to elliptical orbits, and the description of the areas will become more uniform [...] if the focus of each orbit is located in the common center of gravity of all the inner bodies."

In modern terms, one says that to first order, orbits in a multiple-planet system are Keplerian when written in terms of Jacobi coordinates.

One can visually evaluate the validity of the Keplerian approximation in Jacobi coordinates by integrating the full equations of motion for the planets in a given multiple-planet system. The results of four such integrations (for Upsilon Andromedae, HD 168443, 55 Cancri, and GJ 876) are shown in the separate panels of Figure 1. In each case, the integration is started at the epoch of the first published RV observation, and the trajectories of the planets are plotted from that moment through to the present (November 01, 2002). A glance at Figure 1 makes it clear that in the Upsilon Andromedae system (and to a lesser extent in the HD 168443 system) the planetary orbits are dynamically separated to the extent that direct planet-planet perturbations are not very apparent over the timescale of the RV observations. In these cases, as with most of the other known multiple-planet systems, Newton's approximation that the planets trace fixed elliptical orbits about the center of mass of all interior bodies works very well. For systems such as GJ 876, and 55 Cancri, however, the direct perturbations between the planets give the motions a distinctly non-Keplerian flavor. The planetary orbits are aperiodic, and over the time covered by the RV campaigns, their excursions explore thick elliptical annuli surrounding the star. These annuli arise largely from precession of the orbits. Therefore, in order to determine whether a self-consistent N-body fit to the motion is required, one can simply integrate the nominal initial orbital parameters of the system, and look to see whether the resulting motion consists of nearly fixed ellipses.

3. Multiple Planet Fitting

When planets display significant non-Keplerian motion, self-consistent fitting is necessary in order to characterize the system. One must specify the initial masses, positions, and velocities of the planets relative to the star at the epoch of the first

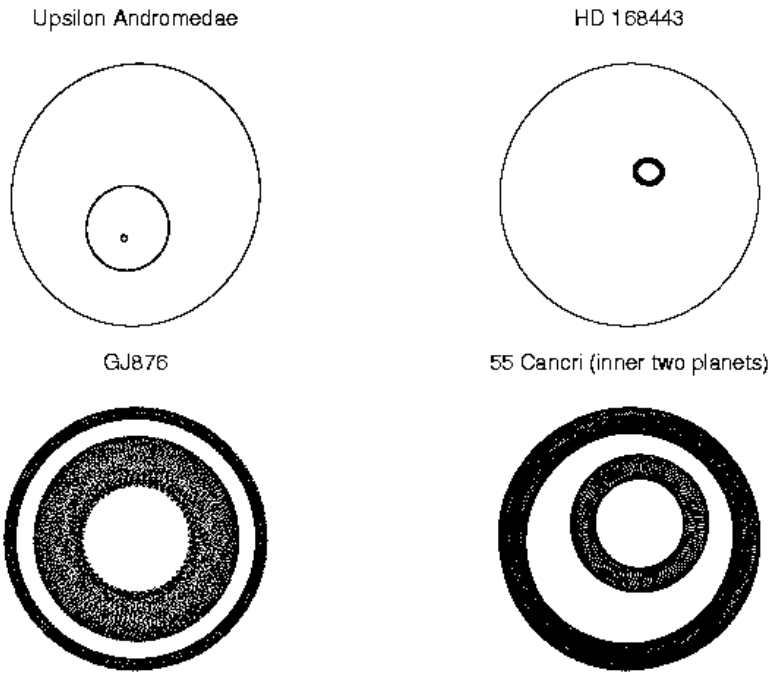

Figure 1. Orbital motion in the center-of-mass frame for four known multiple-planet systems. In each diagram, the orbital motion has been integrated over the actual time frame of the RV observations.

observation, and then examine how the integrated stellar reflex motion, $v_{\rm model}$, compares with the observed RV, $v_{\rm obs}$, at each epoch for which an RV observation exists. Every set of initial conditions for the planets thus corresponds to a particular value of

$$\chi^2 = \frac{1}{N-M} \sum_{i=1}^{N} \left(\frac{v_{\rm obs}(t_i) - v_{\rm model}(t_i)}{\sigma_i} \right)^2$$

where N is the number of RV observations, M is the number of parameters (osculating orbital elements), and σ_i are the individual RV errors for the N observations. The minimum value for χ^2 corresponds to the best system model.

The basic idea is then to vary the initial conditions for the planets so that a minimum value for χ^2 is obtained. A value $\chi^2 \sim 1$ indicates that an adequate model for the system has been found. Because a particular initial condition leads to deterministic motion, a given orbital fit also constitutes a prediction of where future RV measurements will fall. We note that due to a historical quirk, orbital fits to RV data are usually reported in terms of the statistic $\sqrt{\chi^2}$, rather than in terms of χ^2.

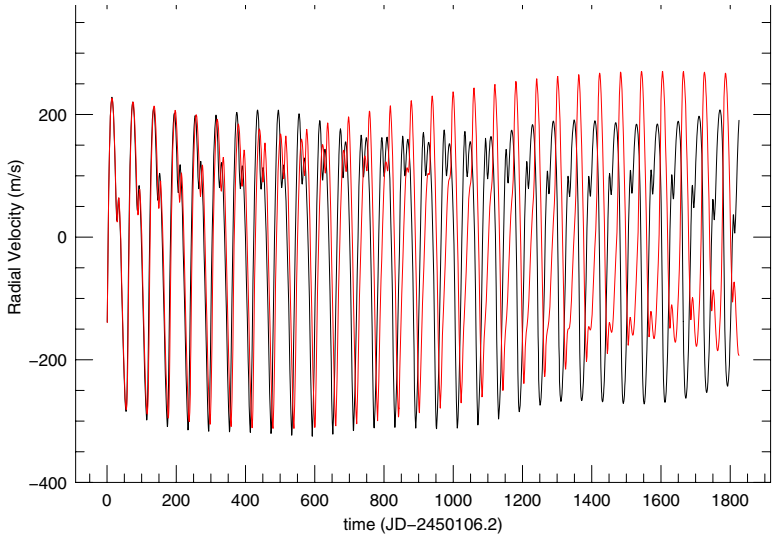

Figure 2. Synthetic RV variations for GJ 876, assuming a super-position of two fixed Keplerian motions for the planets (*red line*) and an N-body integration using the same elements (*black line*).

3.1. GJ 876: A SUCCESSFUL TEST CASE

To date, the GJ 876 system has provided by far the best and most straightforward opportunity for self-consistent fitting. As reported by Marcy et al. (2001), two-Keplerian fits to the RV data from the Keck telescope and to a combined data set from Keck and Lick, suggest that the star (a 0.32 $M_\odot$ M dwarf) is accompanied by two planets on $\sim$ 30 day and $\sim$ 60 day orbits. The RV amplitude variations suggest a minimum mass of order 0.6 $M_{\rm JUP}$ for the inner planet, and 1.9 $M_{\rm JUP}$ for the outer planet. This combination of a low-order (2:1) commensurability, relatively large planetary masses in comparison to the 0.32 $M_\odot$ primary, and observations spanning more than thirty orbits of the outer companion, combine to make self-consistent fits necessary to accurately model the system. To see this, consider Figure 2, which compares synthetic RV curves arising from orbital elements given in Marcy et al. (2001). The red line shows the RV curve that results from the super-position of the two Keplerian reflex motions. The black line shows the RV curve resulting from the full three-body integration. After three orbits of the outer planet, the motion begins to deviate noticeably from the dual-Keplerian approximation. After several years, the motions have diverged completely!

As reported in Laughlin and Chambers (2001), and as found independently by Rivera and Lissauer (2001), when planet-planet interactions are strong, one can obtain self-consistent orbital fits through the use of a calculus-based Levenberg-Marquardt minimization algorithm (e.g., Press et al., 1992) driving N-body integrations of prospective planetary configurations. If the Levenberg-Marquardt algorithm is started with osculating planetary orbital elements corresponding to the best

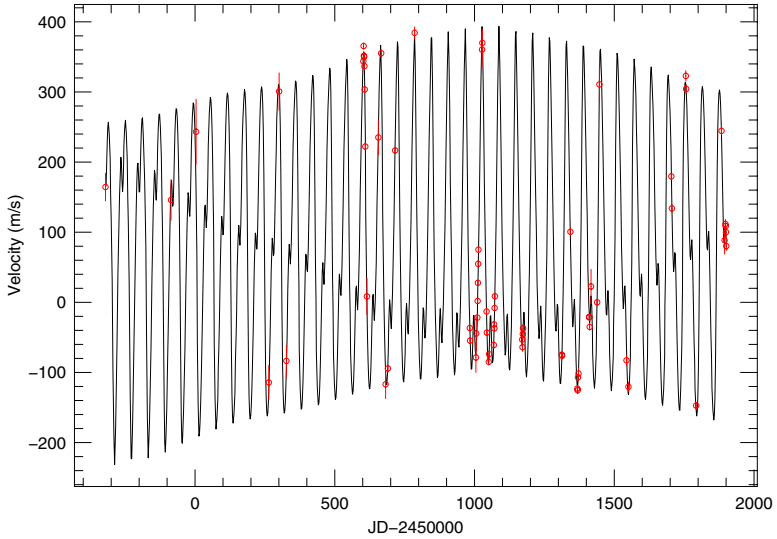

Figure 3. Synthetic RV variations for the best fit generated by a Levenberg-Marquardt scheme applied to the GJ 876 Keck+Lick radial velocities reported by Marcy et al. (2001).

dual-Keplerian fit to the radial velocities, then a significant improvement to the fit is rapidly obtained. Our self-consistent model for the combined Keck and Lick RV data from Marcy et al. (2001) has a $\sqrt{\chi_\nu^2}$ value of 1.46 and an rms scatter of 13.95 $\mathrm{m\,s^{-1}}$. This fit is shown in Figure 3, and represents a significant improvement to the best dual-Keplerian fit which has $\sqrt{\chi_\nu^2} = 1.88$.

Fits which take planet-planet interactions into account can conceivably remove the $\sin(i)$ degeneracy and reveal the true masses of the planets. In the case of GJ 876, the combined RV data favors a co-planar inclination of $\sin i = 0.775$ for the system. There is, however, a broad minimum around this best-fit value. At $\sin i = 0.55$, the $\sqrt{\chi_\nu{}^2} = 1.47$, and the rms scatter is 14.69 $\mathrm{ms^{-1}}$. Only for $\sin i < 0.50$ does the fit begin to worsen significantly. This result is in substantial agreement with the result found by Rivera and Lissauer (2001). Future and ongoing RV measurements will allow the system model to be gradually refined.

Nauenberg (2002) performed a similar analysis using the Keck data alone, and argues that the Keck RV subset is best fit by a co-planar system having $\sin(i) \sim 1$. His his rms scatter and χ^2 values are higher than those found by Laughlin and Chambers (2001) and Rivera and Lissauer (2001), however.

It is interesting to remark that if the GJ876 planets are indeed co-planar, and if the orbits are viewed edge-on, then planet-planet interactions will strongly affect the predicted transit epochs. For example, if one uses the best dual-Kepler fit, one finds epochs of JD=2452599.27, JD=2452629.39, and JD=2452659.51 for the next three transits of the inner planet. Using a co-planar $i = 90°$ self-consistent fit to exactly the same RV data, one obtains completely different predictions for the next three transit times; JD-2452607.85, JD=2452638.11, and JD=2452668.28. Further-

Table I. **Assumed and Fitted Elements for a Synthetic System**

Parameter	Planet 1 (Assumed Elements)	Planet 2 (Assumed Elements)	Planet 1 (Fitted Elements)	Planet 2 (Fitted Elements)
Period (day)	30.000	30.000	29.919	30.037
K ($\mathrm{m\,s^{-1}}$)	50.000	100.00	49.233	99.694
Eccentricity	0.0500	0.0500	0.0418	0.0401
ω (deg)	180.00	180.00	195.02	174.29
Periastron Time (JD)	10.000	20.000	11.253	19.512

more, the individual transit epochs for GJ 876c (the less massive inner planet) are not equally spaced. Libration generated by the resonant interaction causes delays and early arrivals of the transit center by up to 4 hours on either side of the mean period. A 4-hour delay or advance is highly significant when compared to observation intervals spaced by several minutes. Even more remarkably, if the planets have a mutual inclination relative to one another, it is possible that the node of one of the planets can precess into an inclination where transits begin to occur. This possibility is examined in an upcoming paper.

4. Genetic Algorithms as Applied to Multiple-Planet Fitting

Self-consistent fits to the GJ 876 data set show that the GJ 876 system lies deep within the 2:1 mean motion and secular resonances, with libration angles as small as 5°. Indeed, Lee (2002) have shown that these highly damped librations require that strong dissipative processes were present as the planetary orbits evolved into their present configuration.

It is precisely because the GJ 876 system lies so deeply within the 2:1 resonance that a dual-Keplerian fit provides an initial guess adequate for Levenberg-Marquardt minimization to successfully refine the model of the system. For other planetary configurations, however, especially those which lie near the separatrix surrounding a resonance, small variations in the osculating orbital elements of the planets at the initial epoch lead to very strong sensitivity of $\sqrt{\chi^2}$, coupled with an overall $\sqrt{\chi^2}$ landscape that is topologically rugged on larger scales.

Multiple-planet fitting techniques for RV data sets can be tested by computing a stellar reflex motion arising from a specified configuration of planets, and then sampling this motion at cadences and with uncertainties appropriate to the current RV data sets. For instance, as a representative example, one can adopt the updated Keck Telescope RV data set for GJ 876 (G. W. Marcy 2001, private communication) which includes 63 RV measurements spanning 1532 days. Each individual velocity has an associated measurement uncertainty that is estimated from the in-

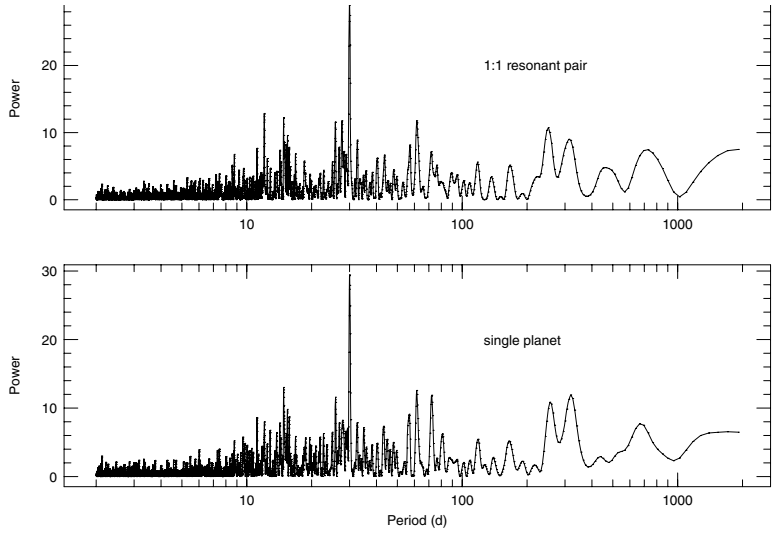

Figure 4. Upper panel: Power spectrum corresponding to a synthetic radial velocity data set produced using the hypothetical planetary system shown in Table I, sampled at the epochs and velocity precision of the GJ 876 Keck radial velocities. Lower Panel: Power spectrum corresponding to a single planet on a circular orbit with $K = 75$m/s, and $P = 30$d, also sampled at the epochs and velocity precision of the GJ 876 data set.

ternal cross-correlation of the lines within the spectra. These uncertainties range from 2.8 to 8.3 m s^{-1}, with an rms value of 5.14 m s^{-1}.

In Table I, we specify osculating orbital elements of a hypothetical 1:1 resonant pair of planets orbiting a solar mass star. This synthetic system exhibits large (tadpole-like) librations about the equilateral co-orbital resonant configuration, and is thus not well-suited to either a Keplerian fit or to the Levenberg-Marquardt minimization procedure that worked for GJ 876.

Using these orbital elements at the epoch of the first radial velocity point as initial conditions, we integrated the hypothetical system forward in time for the duration spanned by the actual RV observations of GJ 876. We then sampled the RV of the parent star in response to the hypothetical 1:1 resonant pair at the 63 observational epochs. For each point, we then added RV noise drawn from a Gaussian distribution of half-width given by the actual quoted velocity errors. This yields a synthetic RV data set from which we can attempt to reconstruct the orbital parameters of the planets.

The top panel of Figure 4 shows the power spectrum of the synthetic RV data set. There is a single significant peak at the fundamental 30 day period, and the spectrum gives little indication of the presence of two planets in the data. In particular, the accordion-like modulation of the RV curve due to the 300 day libration frequency between the planets is not immediately apparent in the power spectrum, as can be seen by comparing with the power spectrum produced by a single planet

of $K = 75 \mathrm{m\,s^{-1}}$ on a circular 30 day orbit (bottom panel of Figure 4). In both cases, the power spectrum shows only a single significant peak.

In order to fit the synthetic 1:1 resonant system, we have employed a two-stage method. In the first stage, we use a genetic algorithm (Goldberg, 1989) as implemented in FORTRAN by Carroll (1999)[1] for public domain use. The genetic algorithm starts with an aggregate of osculating orbital elements, each referenced to the epoch of the first RV observation ($T_0 = 2450602.0931$). Each set of elements (genomes) describes a unique three-body integration and an associated radial velocity curve for the central star. The fitness of a particular genome is measured by the χ^2 value of its fit to the RV data set. At each generation, the genetic algorithm evaluates the χ^2 fit resulting from each parameter set, and cross breeds the best members of the population to produce a new generation.

Because the fundamental 30-day period of the system is clearly visible in the power spectrum, the genetic algorithm is constrained to search for 1:1 resonant configurations in which the initial period ranges for the planets are $29\,\mathrm{d} < P_1 < 31\,\mathrm{d}$, and $29\,\mathrm{d} < P_2 < 31\,\mathrm{d}$. The initial arguments of periapse and mean anomalies of the two planets are allowed to vary within the allowed 2π range. The RV half-amplitudes of the planets are required to fall in the range $0\,\mathrm{m\,s^{-1}} < K_1 < 150\,\mathrm{m\,s^{-1}}$ and $0\,\mathrm{m\,s^{-1}} < K_2 < 150\,\mathrm{m\,s^{-1}}$. The planetary eccentricities are allowed to vary within the full $0 < e < 1$ range.

With these very liberal constraints on the space of osculating initial orbital elements, the Genetic Algorithm rapidly identifies a set of parameters having $\chi = 1.27$. This tentative fit is then further improved by use of Levenberg-Marquardt minimization to produce a fit to the data having $\chi = 0.97$. This fit is shown in Figure 5, and in the third and fourth columns of Table I. The excellent agreement between the input system and the fitted solution shows that a resonant pair of this type is readily identifiable if it exists within a RV data set comparable to the GJ 876 velocities. As the RV surveys continue, an increasing number of target stars will have velocity data sets of this quality.

5. 55 Cancri – A Challenge For Multiple-Planet Fitting

The $P = 14.65$d, $M\sin(i) = 0.84 M_{\mathrm{JUP}}$ object orbiting the nearby super-metal-rich $0.95 M_{\odot}$ star 55 Cnc was the fourth RV planet to be discovered (Butler et al., 1997). Soon after the discovery, residuals in the radial velocities indicated the likely presence of additional planet(s) in the system. After five further years of monitoring, the residuals can be modeled as arising from two additional planets, one with a period $P = 44.3\,\mathrm{d}$ and $M\sin(i) = 0.21 M_{\mathrm{JUP}}$, and a second with period $P \sim 5400\,\mathrm{d}$ and $M\sin(i) = 4.00 M_{\mathrm{JUP}}$ (Marcy et al., 2002). The period ratio of the inner two planets in the Marcy et al. (2002) model is $P_c/P_b = 3.02$, which is close

[1] See `http://cuaerospace.com/carroll/ga.html`

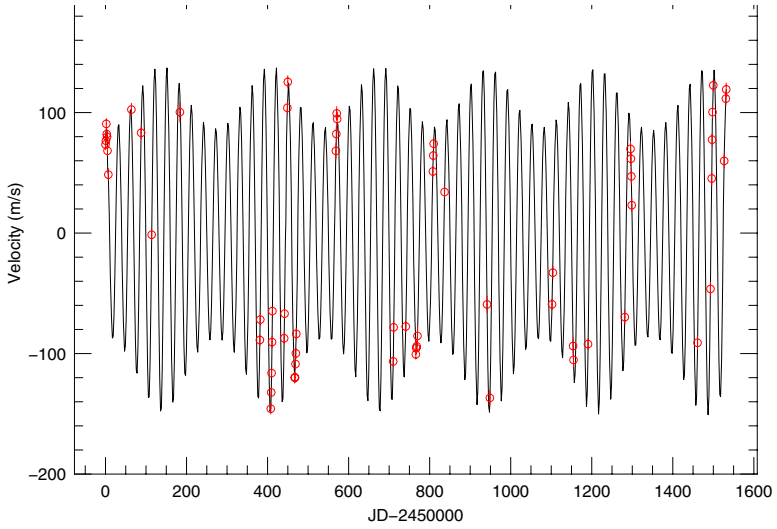

Figure 5. Synthetic RV variations for the planetary system shown in Table I (solid line). The circular points with small vertical lines corresponding to error bars represent a sample of the system with the properties of the GJ876 data set. A superimposed dotted line line shows a fit to the sampled data obtained with a combined Genetic Algorithm and Levenberg-Marquardt procedure. It is hard to distinguish the difference between these two curves at the resolution of the plot.

to the 3:1 commensurability. This means that over the ∼5000 day timespan that the star has been observed, the planet-planet interactions between the inner and middle planets will tend to add in a constructive way. This is illustrated in Figure 6, which shows the running difference between the RV curve of the star under the influence of summed Kepler motions, and under full four-body motion. The discrepancy between the two versions of the stellar motion grows to $\Delta V > 100\,\mathrm{m\,s^{-1}}$ after 10 years of observation, which indicates that self-consistent fitting should be required in order to correctly describe a system in this configuration.

Inspired by the successful attempts to fit the GJ 876 RV data set, we first used the Levenberg-Marquardt minimization routine driving a four-body integrator to produce a self-consistent fit. As with GJ 876, the best summed Keplerian fit was used as an initial guess.

One soon finds, however, that the 3-Keplerian fit to the 55 Cancri data does not provide a similarly propitious point of departure for improvement of the orbital fit. When the summed Keplerian fit is used as an initial guess, the code converges to a self-consistent solution with $\sqrt{\chi^2} = 1.85$. This fit is shown in the first column of Table II. Experimentation with the starting conditions shows that there is a very strong sensitivity of χ^2 to small variations in the initial conditions, coupled with a χ^2 landscape that is topologically rugged on large scales. One must therefore resort to alternate methods to locate the true configuration of the system.

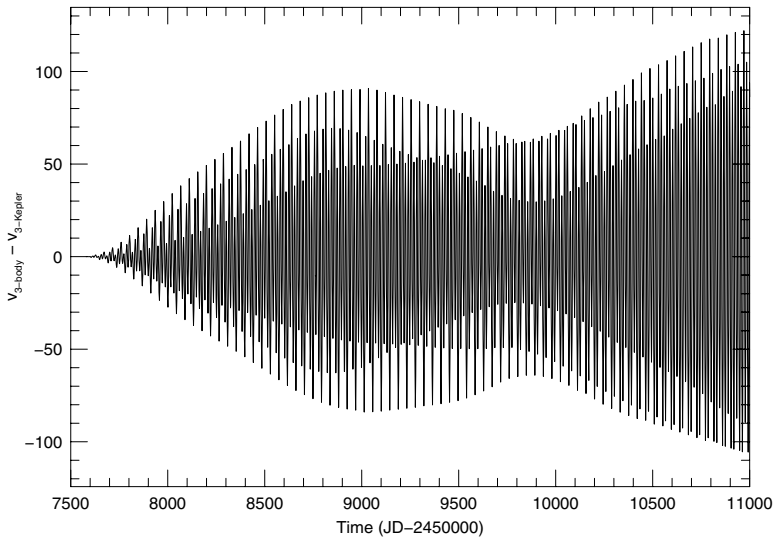

Figure 6. Running difference between the 55 Cnc RV reflex motion under the influence of summed Keplerian motion, and under the full four-body Newtonian motion. Here, three planets are assumed with semi-major axes of 0.11, 0.24, and 5.5 AU. After 10 years, the two simulations of stellar motion differ by $100\,\mathrm{m\,s^{-1}}$, indicating that a self-consistent Newtonian fit is required to correctly fit the motion (see also Figure 1).

We first used a scheme which turns on the planet-planet perturbations in a gradual way, and which was successfully adopted for GJ876 by Rivera and Lissauer (2001). We decreased both the masses of the planets and the magnitudes of the stellar reflex velocities by a factor of 10^6. This allows the Levenberg-Marquardt N-body code to recover the 3-Keplerian fit listed in the first column of Table II. We then gradually increased both the masses of the planets and the radial velocities in a series of discrete increments. After each increment, we allowed the Levenberg-Marquardt minimization to converge to a self-consistent fit. When the radial velocities have grown to their full observed values, the code produces a self-consistent $\sqrt{\chi^2_\nu} = 1.82$ fit, which we list in the second column of Table II. (All of the fits in Table II correspond to Epoch JD 2447578.730, the time at which the first RV observation of the star was made).

With the exception of the eccentricity of the middle planet, which has dropped from its large value of $e = 0.34$, the osculating orbital elements for the self-consistent fits are quite similar to the summed Kepler fit. In contrast to the situation with GJ 876, the imposition of planet-planet interactions has not improved the χ^2 statistic. Furthermore, examination of the 3:1 resonance arguments, $\theta_1 = 3\lambda_c - \lambda_b - 2\varpi_b$, $\theta_2 = 3\lambda_c - \lambda_b - \varpi_c - \varpi_b$, and, $\theta_3 = 3\lambda_c - \lambda_b - 2\varpi_c$, show that while the model systems are close to resonance, none of the resonance arguments are librating for any of the fits. This is illustrated in Figure 7, where the time variation of the resonant argument θ_1 is plotted for both the 3-Keplerian fit and for the self-

Table II. **Three-Planet Models for the 55 Cancri System**

	3-Kepler Fit	LM Fit	GA Fit
P_b (days)	14.653	14.645	14.652
P_c (days)	44.255	44.517	44.301
P_d (days)	4780.58	4840.46	5064.362
MA_b (deg)	271.161	253.309	302.693
MA_c (deg)	235.764	316.890	38.849
MA_d (deg)	3.8528	349.341	21.033
e_b	0.028	0.038	0.020
e_c	0.250	0.217	0.000
e_d	0.145	0.158	0.116
ϖ_b (deg)	67.92	69.92	38.35
ϖ_c (deg)	19.84	350.74	234.15
ϖ_d (deg)	184.35	203.06	180.58
Mass b ($M_{\rm JUP}$)	0.838	0.832	0.835
Mass c ($M_{\rm JUP}$)	0.204	0.186	0.184
Mass d ($M_{\rm JUP}$)	3.571	3.794	3.692
epoch (JD)	2457578.730	2457578.730	2457578.730
$v_{\rm epoch}$ offset ($\rm m\,s^{-1}$)	6.4541	1.824	16.679

consistent fits. It seems clear that the 55 Cancri system will likely turn out to be very interesting dynamically. Even if it is not in resonance today, it likely was in the resonance in the past, and the fact that it is currently not indicates an intriguing history, possibly including tidal dissipation.

The failure of the Levenberg-Marquardt routine to significantly improve the χ^2 statistic suggests that a true global minimum in the three-planet parameter space was simply not located by the *locally* convergent Levenberg-Marquardt algorithm, suggesting the use of the globally convergent genetic algorithm, as described in the previous section. Unfortunately, however Extensive use of the genetic algorithm also fails to find a significantly improved χ^2 value over that provided by the summed Keplerian fit. This is in contrast to the algorithm's excellent performance on the above-described test problem. The best fit which we evolved is listed in the third column of Table II. This fit has $\sqrt{\chi^2}$=1.82, and has osculating orbital elements which are very similar to the fit obtained by slowly increasing the strength of the planet-planet interactions (second column of Table II).

In summary, we have investigated three numerical strategies for producing self-consistent three-planet fits to the 55 Cancri RV data set. All three methods provide

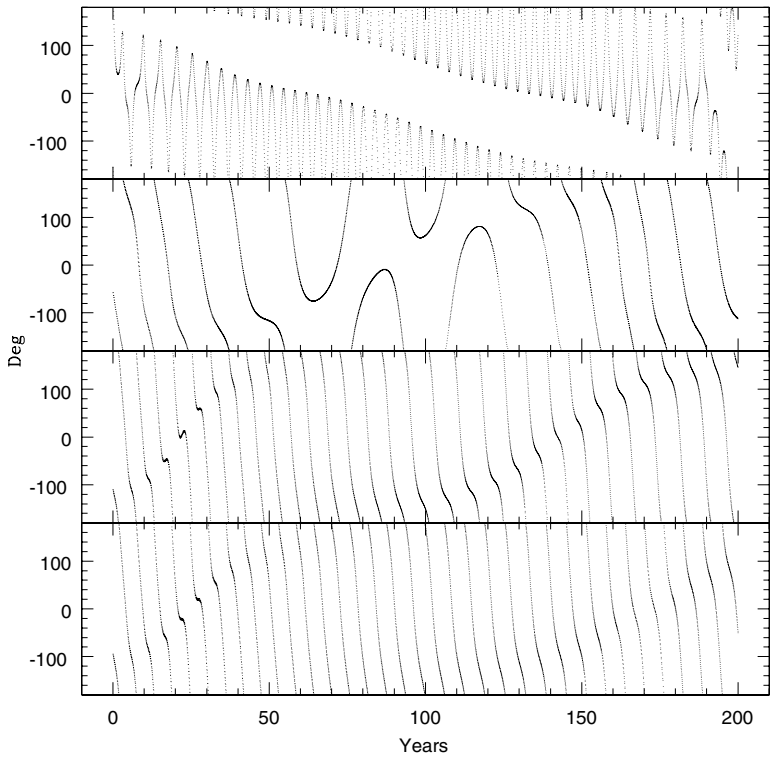

Figure 7. Time behavior of the 3:1 resonant argument $\theta_1 = 3\lambda_c - \lambda_b - 2\varpi_b$, for *Top Panel* summed Keplerian fit, *Second Panel* self-consistent 4-body fit #1 (see Table II), *Third Panel* self-consistent 4-body fit #2, *Bottom Panel* self-consistent 4-body fit #3.

fits with χ^2 statistics that are essentially equivalent of the best 3-Kepler fit, and all three fits are in quite good agreement, suggesting that the system lies just outside of the 3:1 resonance. We stress, however, that it is not yet completely clear whether the system is indeed in the 3:1 resonance. Further dynamical fitting and further observation of the star will be required to definitively identify the dynamical relationship between the inner and the middle planet.

6. Conclusions

The most important long-term benefit of self-consistent dynamical fitting techniques for multiple-planet systems lies in their ability to break the $\sin(i)$ degeneracy and thus determine the true masses and mutual inclinations of extrasolar planets. In a few systems, the true masses can be found when a planet transits the parent star, but these cases will be relatively rare. For stars brighter than V=10, for which 3-5 m/s RV precision can be readily obtained, we estimate that there are $\approx$ 10 transiting planets with periods less than a week, and $\approx$ 10 transiting

planets with periods in the range $7\,\mathrm{d} < P < 200\,\mathrm{d}$. Self-consistent fitting, on the other hand, can be applied to any system containing more than one planet, given a sufficient baseline of observation. Self-consistent fitting may also help with uncovering the presence of additional planets in a system, or demonstrating that an additional planet (beyond the first two, say) is unnecessary to get an adequate fit.

References

Butler, R. P., Marcy, G. W., Williams, E., Hauser, H., and Shirts, P. 1997, 'Three New "51 Pegasi-Type" Planets'. *Astrophys. J.* **474**, L115-L118.

Butler, R. P., Marcy, G. W., Fischer, D. A., Brown, T. M., Contos, A. R., Korzennik, S. G., Nisenson, P., Noyes, R. W. 1999, 'Evidence for Multiple Companions to υ Andromedae'. *Astrophys. J.* **526**, 916–927.

Fischer, D. A., Marcy, G. W., Butler, R. P., Vogt, S. S. Frink, S., and Apps, K. 2001, 'Planetary Companions to HD 12661, HD 92788, HD 38529 and Variations in Keplerian Residuals of Extrasolar Planets'. *Astrophys. J.* **551**, 1107–1118.

Fischer, D. A., Marcy, G. W., Butler, R. P., Laughlin, G., Vogt, S. S. 2002, 'A Second Planet Orbiting 47 Ursae Majoris', *Astrophys. J.* **564**, 1028–1034.

Goldberg, D. E. 1989 'Genetic Algorithms in Search, Optimization and Machine Learning' 1989. Addison-Wesley: Reading, MA.

Laughlin, G., and Chambers, J. E. 2001, 'Short-Term Dynamical Interactions among Extrasolar Planets', *Astrophys. J. Letters* **551**, L109–L113.

Lee, M. H. and Peale, S. J. 2002, 'Dynamics and ORigin of the 2:1 Orbital Resonances of the GJ 876 Planets', *Astrophys. J.* **567**, 596–609.

Marcy, G. W., Butler, R. P., Fischer, D. A., Vogt, S. S., Lissauer, J. J., Rivera, E. J. 2001, 'A Pair of Resonant Planets Orbiting GJ 876'. *Astrophys. J.* **556**, 296–301.

Marcy, G. W., Butler, R. P., Fischer, D. A., Laughlin, G., Vogt, S. S., Henry, G. W. and Pourbaix, D. 2002, 'A Planet at 5 AU Around 55 Cancri'. *Astrophys. J.* In Press.

Nauenberg, M. 2002, 'Determination of Mases and Other Properties of Extrasolar Planetary Systems with More than One Planet', *Astrophys. J.* **568**, 369–376.

Newton, I.: 1687, 'The Principia, Mathematical Principles of Natural Philosophy'. I. B. Cohen and A. Whitman (trans.) 1999. Berkeley: University of California Press.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery B. P. 'Numerical Recipes in FORTRAN, 2nd Edition' 1992. Cambridge: Cambridge University Press.

Rivera, E. J., and Lissauer, J. J. 2001, 'Dynamical Models of the Resonant Pair of Planets Orbiting GJ 876', *Astrophys. J.* **558**, 392–402.

Schneider, J. 2002, 'Extrasolar Planets Encylopdedia'. `http://cfa-www.harvard.edu/planets/catalog.html`.

Part VII

Future and Conclusions

Theories of Planet Formation: Future Prospects

Jack J. Lissauer
Space Science Division 245-3, NASA Ames Research Center, Moffett Field, CA 94035, USA
*(*jlissauer@ringside.arc.nasa.gov*)*

Abstract. Several major questions in theories of planet formation have remained a concern for over a decade: How do planetesimals form (especially the growth from mm–km size)? What core mass (if any!) is needed for a planet to accumulate substantial quantities of hydrogen and helium? What is the distribution of sizes and orbits of terrestrial planets? How common are habitable planets? Many of these questions are likely to be answered in the next decade, some by purely theoretical advances, others by a combination of additional observations or experiments and theoretical research.

1. Introduction

Questions concerning the existence and characteristics of other worlds been pondered by Natural Philosophers for at least the past few millennia. Aristotle believed that earth, the densest of the four elements, fell towards the center of the Universe, so no other worlds could possibly exist; in contrast, Democritus and other early atomists surmised that the ubiquity of physical laws implies innumerable Earth-like planets must exist in the heavens (Lewis, 1998). Some theologians have considered Earth and humanity to be God's special place, unique in the entire universe, whereas others see no reason why God would have bothered creating stars other than the Sun and not also have surrounded them with planets teeming with life.

Some aspects of the question of Earth's uniqueness remain ill-constrained, but others have yielded to scientific investigation. Copernicus, Kepler, Galileo and Newton convincingly demonstrated that our home planet is not the center of the Universe, and that other worlds qualitatively similar to the Earth orbit the Sun. Telescopic observations, and more recently interplanetary spacecraft, have told us a great deal about these neighboring worlds. This conference and book honor the 60^{th} birthday of Michel Mayor, who, together with collaborators and competitors from around the world, has greatly extended our horizons by discovering $\sim$100 planets in orbit about stars other than our Sun.

Questions of origins are generally quite difficult to answer: It is almost always easier to figure out how something works than to understand how that something got to be the way it is. We can now accurately predict the locations of the planets within our Solar System tens of millions of years into the future. In contrast, first-order questions about the origin of our Solar System remain unanswered, and quite possibly some them have not yet even been asked.

The complexities of the physics and chemistry involved and the uncertainties in initial conditions have prevented theorists from developing models of planetary formation that are complete and predictive. Rather, models have been developed to fit

the observations of star forming regions, our Solar System and extrasolar planetary systems. We have data on a limited number of planetary systems, and furthermore these data are strongly biased — only extrasolar planets massive enough and located sufficiently close to their stars to produce detectable variations in the star's radial velocity have been identified, and we would not be enquiring about these issues if our Solar System didn't have characteristics amenable to the development of life forms advanced enough to ponder them (Wetherill, 1994).

Most models of planetary formation endeavor to simulate the growth of our Solar System, and some models try to produce planets observed about other stars; other configurations come from simulations that tried but failed to reproduce observed systems. Thus, theory shares many of the biases of observations. However, the sensitivity of model outcomes to small changes in initial conditions, as well as to the observed variations of circumstellar disk properties, imply that a great deal of diversity can be expected among extrasolar planetary systems (Lissauer, 1995).

Despite all of the challenges involved, the intellectual and technological advances of the past century have greatly improved our understanding of planetary cosmogony and leave us poised to investigate the problem in far greater detail. This article focuses on some of the important questions that theoretical research on planetary cosmogony will address in the coming years. For a summary of current models of planet formation, see the reviews by Lissauer (1993, 1999) and Lissauer and Lin (2000).

2. Recent History

Before endeavoring to address the future, let us consider some changes in our perception of planet formation within the past couple of decades.

In the 1970's, the prevailing models viewed planet formation to be an orderly process of growth (Safronov, 1969; Wetherill, 1980). The acceptance of the giant impact theory of lunar formation by most planetary cosmogonists (Hartmann *et al.*, 1986) marked a shift to a view in which stochastic events played a central role in the formation of our Solar System. Indeed, the outcome of current models of terrestrial planet growth is extraordinarily sensitive to initial conditions, e.g., moving just one of over 100 lunar-sized planetesimals along its orbit by 1 meter can change the ultimate outcome from a system with 5 terrestrial planets to one in which only 2 terrestrial planets survive (Chambers *et al.*, 2002).

Prior to the discovery of the first extrasolar planets in the 1990's, most astronomers either believed that planets are rare or that our Solar System is typical. Giant planets were expected to travel on low eccentricity orbits beyond the ice condensation line, and some models suggested that planets could not be much more massive than Jupiter.

Observations of extrasolar planets have provided most of the impetus for theoretical advances in planetary cosmogony over the past decade. The first extrasolar

planets were discovered in quite unexpected locations: Two planets with masses a few times that of the Earth were observed around a millisecond pulsar (Wolszczan and Frail, 1992), a planet with $M \sin i$ just under half of Jupiter's mass was identified orbiting only 0.05 AU from the star 51 Pegasi (Mayor and Queloz, 1995), and a several Jupiter mass companion was detected in a quite eccentric ($e = 0.4$) orbit about 70 Virginis (Marcy and Butler, 1996). Known extrasolar planets have a broad distribution of masses peaking near the mass of Jupiter, but this sample is strongly biased towards higher mass planets. About 10% of stars have giant planets with orbital periods of less than 3 years (Marcy *et al.*, 2004), and, apart from very close-in planets whose orbits have presumably been tidally circularized, most of these planets have orbital eccentricities well above those of the giant planets within our Solar System.

Theoretical advances are more difficult to predict than are observational ones, as they are less technologically based and a larger fraction of the progress is made through rare brilliant advances. While I can confidently predict that numerical studies will accelerate with improvements in computer algorithms and hardware, and if history is any guide, both analytic and numerical modeling will continue to have a synergistic relationship with observations, my crystal ball becomes cloudy when I try to view specific future theoretical advances.

I attempted to peer through such clouds in the mid-1990's, when I published a paper using theoretical models of planetary growth to speculate on the diversity of plausible planetary systems (Lissauer, 1995) only a few months prior to the announcement of the first planet known to orbit a normal star other than the Sun (Mayor and Queloz, 1995). While past performance is not a guarantee of future results (a fact painfully apparent to just about everyone who bought tech stocks in 1999 or 2000), examining how these predictions have fared may be useful to assessing the discussion in the following sections.

In the introduction of my 1995 article, I stated "Extrapolating from the one planetary system known to orbit a main sequence star to a model of the variety of planetary systems which may be present throughout the galaxy is a daunting challenge surely fraught with pitfalls. Detailed predictions are almost certain to be erroneous." This turned out to be one of my most successful predictions. I nonetheless continued forward, and as you shall see, I stumbled into a few of those pits.

A large piece of the article concerns dynamical stability "The spacings of planets are limited by dynamical stability considerations...", and my basic thesis was that our planetary system and the satellite systems of the giant planets are about as full as possible for them to be stable for the observed lifetime of the system. The prediction that other planetary systems are nearly dynamically full remains untested. However, the statement "...the major planets in most, but not all, planetary systems probably have low orbital inclinations and eccentricities" has not received support from observations (although, since planets have thus far been detected around 10% of the stars surveyed, it has not yet been definitively refuted).

Planetary migration has been a hot research topic ever since Mayor and Queloz's (1995) seminal paper on 51 Peg (Lin *et al.*, 1996). I did note that "Planetary orbits may be able to migrate substantially as a result of such gravitational torques (Goldreich and Tremaine, 1980). Thus, even fully formed planets are not safe from being consumed by the central star. Entire planetary systems or portions thereof could in principle be lost...", but neither I, nor to my knowledge anyone else, predicted that many planets would migrate inwards only to stop just prior to falling into a firey grave. Because the prevailing theory of giant planet formation was then and is now the core instability model (Pollack *et al.*, 1996; Wuchterl *et al.*, 2000; Lissauer, 2004a), which predicts giant planet growth times comparable to the observed lifetimes of protoplanetary disks, I suggested that "...some planetary systems consist of only terrestrial type planets ... gas giants are more likely to exist and to be more massive in high surface mass density, long-lived protoplanetary disks." Today, I would probably also note that planetary migration could eliminate some planetary systems entirely, and replace "exist" in the above statement by "form", to account for the possibility that the most massive disks might lead to the demise of most planets formed within them.

Most successful of my predictions were those which emphasized our ignorance, e.g., "we do not have a good understanding of the factors that determine the final masses of giant planets" and paid homage to the creativity of Nature, "The variety of planets and planetary systems in our galaxy must be immense, and even more difficult to imagine and predict than was the diversity of the outer planet satellites prior to the *Voyager* mission."

Given the difficulty of making theoretical predictions of planetary system properties based on models of planetary growth, what is the value of doing theoretical simulations? The answer is manifold. Fundamentally, understanding "how" is as important as understanding "what". More appealing to observers, perhaps, is that theory is already being used to interpret observations: Multiple planet systems in which no mutual perturbations are observed are assumed to have all planets orbiting in the same direction based on formation scenarios. Limits on masses and inclinations are provided by stability considerations. Stability of test particles constrains where Earths may orbit.

Note that, powerful as current observing techniques are, they still only provide us with a limited view. Consider the response of the Sun to perturbations by the planets, as viewed by radial velocity observers (Figure 1) and astrometrists (Figure 2) located at optimal positions for their various techniques. While Jupiter and Saturn analogs should soon be detectable, other planets in our Solar System induce much smaller signals, or, in the case of astrometric perturbations by Uranus and Neptune, signals that are difficult to interpret without many decades of data.

Thus, theoretical modeling is useful, but that doesn't mean that predictions about future theoretical models are likely to fare well. In an effort not to make the same mistakes twice (it's more fun to make different kinds of mistakes), my

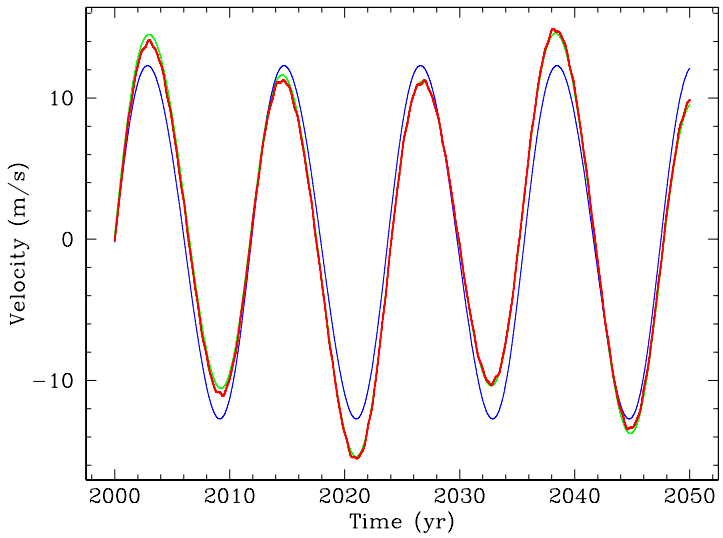

Figure 1. Model of radial variations in the velocity of the Sun as viewed from the plane of Jupiter's orbit. The thin curve shows variations caused by Jupiter, the medium curve also includes Saturn's gravitational forcing and the thick curve includes perturbations by all nine planets. Note that the medium curve is quite close to the thick curve at all times, indicating that radial velocity variations of the Sun are dominated by the perturbations of Jupiter and Saturn. Courtesy Elisa Quintana.

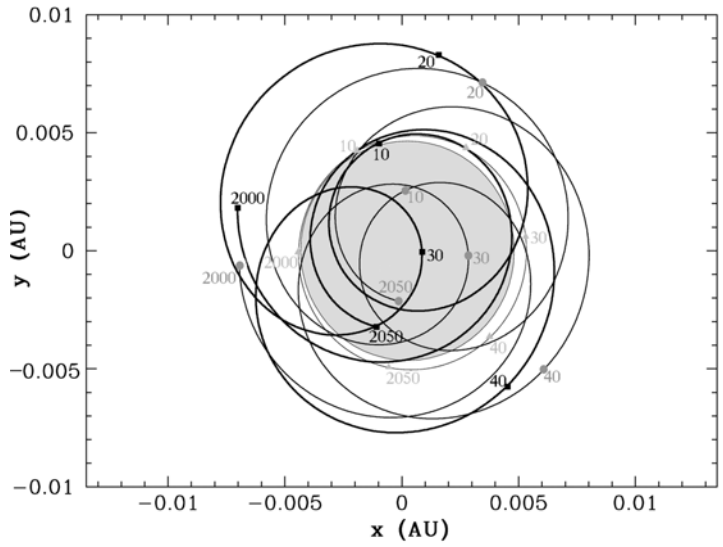

Figure 2. Variations in the Sun's position resulting from perturbations of Jupiter alone (thin curve with dates marked by triangles), Jupiter and Saturn (medium curve with dates marked by circles) and all nine planets (thick curve with dates marked by squares) as they would be observed along a line from the Solar System perpendicular to the plane of the Jupiter's orbit. The axes give the Sun's displacement from the center of mass of the system in AU. If the observer were located 10 parsecs away, the amplitude of the apparent motion of the Sun due to perturbations by Jupiter would be 5×10^{-4} arcseconds. The shaded circle represents the size of the solar disk. Courtesy Elisa Quintana.

focus in the remaining sections of this article is to emphasize the questions that theoreticians will be addressing, rather than to suggest answers.

3. Questions From Our Solar System

The overwhelming majority of all elements heavier than boron were produced by nucleosynthetic processes within stars, and to a good approximation primitive material in the Solar System is isotopically and (apart from volatiles) elementally uniform. However, small variations in isotopic abundances can tell us a great deal about the processes involved in planetary formation, and the timescales over which these processes took place (cf. Halliday, 2004). Much of the value of these chronometers depends on understanding the origin of the radioactive isotopes whose decay has produced the most useful signatures. Most of these nuclei are presumed to have been formed in stars, subsequently been incorporated within the solar nebula, and then condensed into small bodies prior to their radioactive decay. However, the presence of small pockets of nearly pure ^{22}Ne, presumably from the decay of ^{22}Na ($t_{1/2} = 2.6$ years) clearly indicates that some pre-Solar System decay products were incorporated into primitive bodies without being mixed. Some of the isotopes with half-lives of $10^5 - 10^6$ years, and thus the potential to provide excellent information on the timescales of various aspects of the planetary formation process, are light nuclei that might be produced within the forming Solar System as a result of irradiation by the young, active Sun (Gounelle *et al.*, 2001). Determining the origins of these isotopes could help us pin down timescales for various processes within the protoplanetary disk, and possibly even within its precursor molecular cloud core.

Primitive meteorites are planetary building blocks that were not incorporated into a planet until their recent landing on Earth. Studies of such meteorites reveal clues to the Earth's bulk composition that cannot be obtained from direct observations. Chondrules are small (typically a few mm) igneous inclusions that are quite abundant in most primitive meteorites. Their composition and morphology imply that they were formed within the Solar System, prior to being incorporated into large bodies. They were heated to temperatures of $\sim$1600 K, but only remained close to their peak temperatures for a very short time, and rapidly cooled by at least several hundred degrees. CAIs are very refractory inclusions found in some primitive meteorites; CAIs appear to have formed a few million years prior to chondrules (Amelin *et al.* , 2002). Chondrules and CAIs imply very energetic processing, localized in time and/or space, occurred within the solar nebula. Various heating mechanisms have been proposed, including lightning (Desch and Cuzzi, 2000), passage through shock waves (Ciesla and Hood, 2002; Desch and Connolly, 2002) and flash heating near the Sun (Shu *et al.*, 1997), but no consensus yet exists. Thus, processes that modified much of the solid material that formed planets are poorly understood.

The gas within a planet's atmosphere is prevented from collapsing onto the planet's surface by pressure gradients. In contrast, satellites remain in orbit because centrifugal forces balance gravity. Models of protoplanetary disks imply that they are primarily centrifugally supported, but that pressure gradients in the gas can provide of order 1% of the force needed to balance gravity (Adachi *et al.*, 1976; Weidenschilling, 1977). Thus, the gas in a protoplanetary disk orbits about the star at about 1% less than the keplerian speed in order to maintain a balance between the star's inward gravitational pull and the combined outwardly directed centrifugal force and pressure gradient. Small dust in the disk travels along with the gas, but large bodies orbit at essentially the keplerian rate. These large bodies face a headwind from the slower-moving gas, but because the gas density is so low, they are not greatly affected. The problem occurs for medium-sized bodies, which collide with an amount of gas roughly equal to their own mass every orbit. These bodies, ranging in size from centimeters to tens of meters, orbit substantially faster than the gas, but have small enough mass/cross-sectional area ratios that the headwind removes a considerable fraction of their orbital angular momentum. The orbits of meter-sized bodies could spiral inwards from 1 AU until they reach the star's surface (or evaporate) in as little as 100 years (Weidenschilling, 1977). How do bodies grow fast enough not to be lost to the stellar inferno?

Once bodies grow to several kilometers in size, they are safe from gas drag for the timescale at which the gas is believed to be present (a few million years), but many theoretical questions remain about later stages of solid-body growth. In the outer fringes of the Solar System, characteristic timescales (orbital, collisional, accretional) are quite long, and it is difficult to understand how the larger bodies in the Kuiper belt could have grown to their current sizes before perturbations by the giant planets excited their relative motions to a point at which accretion would have tailed off or stopped. One possibility is that the Kuiper belt was once many orders of magnitude more massive than it is today (Luu and Jewitt, 2002), but then the removal of all of this material must be explained.

The problems with meteorites are not limited to the primitive objects. Most differentiated meteorites come from the crusts of asteroidal bodies, and crusts don't represent the bulk of asteroids. Crusts are, however, the most easily accessible part of an asteroid. But why then are there a considerable number of iron meteorites from the cores of differentiated bodies, but hardly any dunite meteorites representing their mantles? Tidal stripping of mantles and crusts of asteroids during the epoch of planetary growth by large bodies that have subsequently been ejected from the asteroid belt may well help to account for the origins of metal meteorites and asteroids (Asphaug *et al.*, 2002). However, the deficit of dunite meteorites (relative to expected fraction of asteroid volume) is so severe that much uncertainty remains.

Dynamical models suggest that comets now in the Oort cloud formed in the region of the giant planets, and were subsequently delivered to their current orbits by a combination of planetary encounters, which sent them outwards on nearly parabolic trajectories, and stellar perturbations and/or the galactic tidal field, which

raised their perihelia above the reach of the planets. However, using estimates of the current population of the Oort cloud from the flux of new comets, and accounting for losses due to overzealous planetary perturbations and subsequent encounters with passing stars and giant molecular clouds, there are believed to been so many planetesimals destined to be sent outwards to become either Oort cloud members or interstellar vagabonds that they would block each other's escape routes from the planetary region (Stern and Weissman, 2001).

A collision problem closer to home is presented by the Late Heavy Bombardment. The impact rate on the Moon (and Earth) around 3.9–3.8 billion years ago was hundreds of times as large as during subsequent times, and probably also substantially larger than in the prior few hundred million years, although this latter point remains somewhat controversial (Hartmann *et al.*, 2000; Morbidelli *et al.*, 2001). What could have caused such a sharp spike in cratering two-thirds of a billion years after the Solar System formed? A big asteroid which was collisionally disrupted near a resonance allowing rapid delivery of the fragments to Earth? A somewhat smaller asteroid tidally disrupted during a close encounter with Venus or Earth? A runaway planet in the asteroid belt (Chambers and Lissauer, 2002)? Late rearrangement of Uranus and Neptune (Thommes *et al.*, 2002)? We don't yet know, but the answer may hold important clues to the late stages of the planet formation process.

The formation of the giant planets presents another host of problems to theoreticians. Probably the best known of these is that the timescale of giant planet growth predicted by the favored core-instability model (several million years, Pollack *et al.*, 1996) is comparable to or longer than the lifetimes of many protoplanetary disks (Strom *et al.*, 1993). Lower opacity in the outer atmospheres of proto-giant planets may lead to more rapid accretion rates (Pollack *et al.*, 1996; Ikoma *et al.*, 2000), but the models require additional work.

In-situ growth of Uranus and Neptune presents even more serious problems, because the escape velocities of these planets are so much greater than orbital speeds at their distance from the Sun that even at half their current masses Uranus and Neptune could excite planetesimals to such high eccentricities that ejection from the Solar System becomes much more common than accretion (Lissauer *et al.*, 1995). Although it sounds quite radical, the model of forming Uranus and Neptune between Jupiter and Saturn and subsequently scattering them outwards (Thommes *et al.*, 1999; 2002) may offer the best hope of surmounting these difficulties.

Several minor constituents in Jupiter's atmosphere, C, N, S, Ar, Kr and Xe, were found to have abundances relative to H approximately three times as large as in the Sun's atmosphere (Young, 2003). Presumably, the super-solar abundances of these gases entered Jupiter in solids. However, the condensation temperatures of these elements (or the compounds likely to contain these elements within a solar mixture of gases), or more reasonably the temperatures at which they are efficiently incorporated into water ice, vary substantially. Only at temperatures less than 30 K will they all go into ice (Bar-Nun *et al.*, 1988), and thus only at temperatures less

than 30 K would we expect that they would have comparable enhancements relative to solar. But these temperatures are far lower than the expected temperatures where Jupiter was forming. The processes responsible for these uniform supersolar abundances are not known, although some scenarios have been proposed (Owen, 2004).

Even after the giant planets form there may be problems keeping them around. As mentioned above, torques from the protoplanetary disk may drive planets into the star (cf. Ward, 2004). If giant planets are on orbits that are too closely spaced, some of them may be ejected from the system (Levison *et al.*, 1998; Ford *et al.*, 2004).

4. Questions From Extrasolar Planets

A host of new theoretical questions has arisen from the discoveries of extrasolar planets in recent years. The first of these to arise concern pulsar planets. The low eccentricity orbits and close spacing of the planets orbiting PSR 1257+12 suggest that these objects were formed subsequent to the birth of the pulsar, and indeed that the late stages of their growth probably resembled the late stages of terrestrial planet accumulation within our Solar System. However, the formation and characteristics of their protoplanetary disk may have been quite different (Podsiadlowski, 1993). The various models for circumpulsar protoplanetary disks predict different frequencies and properties of the pulsar planets that they produce, and thus observations of additional pulsar planets, or of indications that pulsar planets are rare, will guide theory here.

As mentioned above, the observed orbital distribution of giant exoplanets came as quite a surprise. Although it is possible that the majority of giant planets will be found to travel on low eccentricity orbits beyond the ice condensation line, clearly at the very least a significantly minority are not so "well-behaved". Gravitational interactions between planets and their surrounding natal disks are probably responsible for a substantial amount of inward migration of many giant planets (Ward, 2004). Disk-planet interactions may also play a role in exciting planetary eccentricities. Gravitational scatterings among giant planets are almost certainly responsible for producing the high eccentricities of some extrasolar planets. Tides raised by the star on nearby planets clearly have circularized the orbits of planets with orbital periods of less than a week, and tides on the star raised by planets may play a role in the pile-up of planets with periods of 3–5 days. Theoretical work is needed to refine (and possibly in some cases refute) these zeroth-order models.

Giant planets are being found more frequently around stars with higher metallicity (Gonzalez *et al.*, 2001). This correlation supports the core-instability model of giant planet formation, as higher metallicity presumably leads to more solids and more rapid core growth within protoplanetary disks. But how strong this correlation is, if it even exists for planetary systems resembling our Solar System, and

whether or not it is as simple as more solids equals more planets, are outstanding questions. For instance, planets that become massive while a substantial amount of gas remains in the disk may migrate into the star as a consequence of their gravitational interactions with the disk (Goldreich and Tremaine, 1980; Ward, 1986). How planetary properties depend on stellar mass and binarity is related to many important issues that have not yet been substantially addressed either observationally or theoretically, although some studies are underway (e.g., Quintana *et al.*, 2002).

The one exoplanet for which we have both size and mass information, the transiting companion to HD 209458, is larger than predicted by models developed prior to the measurements. Possible causes for the planet's inflated size are tides and the dynamic transport of radiative energy received from the star downwards to depths of tens of bars (Guillot and Showman, 2002).

5. Addressing The Problems

New data will clearly be extremely important for providing inspiration to and constraints on models of planetary formation. Observations of exoplanets are likely to be paramount in this regard (Lissauer, 2004b), but observations of protoplanetary disks, star-forming clouds and within our Solar System may also be vital.

Advances in basic physics will improve theories of planetary cosmogony. Experiments to determine the equation of state of hydrogen and helium at high pressures will improve our understanding of the interior structure of giant planets (Wuchterl *et al.*, 2000), and resolve questions about the masses of their heavy-element cores, if any. But enough of calls for help to observers and experimentalists — what can theorists do to improve our understanding of planetary cosmogony? Theoretical work can also better elucidate the equation of state of hydrogen/helium mixtures at high pressures. A better understanding of turbulence, both in disks and in atmospheres, would also be extremely useful.

Improvements in numerical models resulting from better hardware and algorithms should help clarify many issues. The formation of long-lived clumps of masses down to a few times that of Jupiter has been found for some initial conditions (Boss, 2004), but these disks evolve so rapidly that their starting point is clearly unreasonable. Only calculations that begin with a stable disk, which then becomes unstable through the accumulation of additional mass or cooling, will determine whether or not the giant gaseous protoplanet model is viable. State-of the art simulations of the dynamics of the late stages of terrestrial planet formation typically produce planets with more eccentric orbits than those of the terrestrial planets within our Solar System (Chambers, 2001). Are the higher eccentricities the result of neglecting gas drag, small planetesimals and collisional fragmentation, or is our planetary system simply at one extreme of the distribution? Numerical studies with faster machines and algorithms will allow for the incorporation of more bodies and physics into the simulations and provide the answers. More de-

tailed fluid mechanical models of off-center giant impacts and their aftermath will help us better understand the origin of Earth's Moon, and the likelihood of finding similar bodies elsewhere.

Computational improvements will also advance our understanding of disk-planet interactions and of the core instability model of giant planet growth. However, these topics might well receive comparable or greater boosts from new ideas rather than simply better computer models.

Thorny issues like the formation of chondrules and the accumulation of chondrules and matrix into kilometer-sized planetesimals probably require new ideas, either to support one of the existing scenarios or to come up with a new one. Likewise, new ideas are probably required to explain the abundances of volatile heavy elements in Jupiter.

6. Conclusions

Prognostications typically overestimate the changes that occur over time scales of decades (consider predictions of flying cars and other wonders made at the 1939 World's Fair in New York). In contrast, forecasts centuries into the future tend to be overly conservative. In part, this reflects deficiencies in the human imagination. However, another extremely important factor is that the average rate of change greatly exceeds the median rate. Consider the implication of joining the galactic club of civilizations a billion years more advanced than our own (assuming such a community exists!). Such a revelation would lead to changes far more fundamental than the invention of movable type, the industrial revolution and the information revolution have brought to us within the past millennium. All of our inquiries about how planets form would easily be answered by such advanced beings. Unless/until this happens, we have many questions to address whose answers may lead us towards the goal of understand the formation of planets in general and our Solar System in particular.

References

Adachi, I., Hayashi, C., and Nakazawa, K.: 1976, 'The gas drag effect on the elliptic motion of a solid body in the primordial solar nebula', *Prog. Theor. Phys.* **56**, 1756–1771.

Amelin, Y., Krot, A. N., Hutcheon, I. D., and Ulyanov, A. A.: 2002, 'Lead isotope ages of chondrules and calcium-aluminum-rich inclusions', *Science* **297**, 1678–1683.

Asphaug, E., Petit, J.-M., and Rivkin, A. S.: 2002, 'Removing Mantles from Cores: Tidal Disruption of Ancient Asteroids', *Lunar Planetary Inst. Conf. Abstracts* **33**, 2066.

Bar-Nun, A., Kleinfeld, I., and Kochavi, E.: 1988, 'Trapping of gas mixtures by amorphous water ice', *Phys. Rev. B* **38**, 7749–7754.

Boss, A. P.: 2004, 'Outlook: Testing Planet Formation Theories', this volume.

Chambers, J. E.: 2001, 'Making More Terrestrial Planets', *Icarus* **152**, 205–224.

Chambers, J. E., and Lissauer, J. J.: 2002, 'A New Dynamical Model for the Lunar Late Heavy Bombardment', *Lunar and Planetary Inst. Conf. Abstracts* **33**, 1093.
Chambers, J. E., Quintana, E. V., Duncan, M. J., and Lissauer, J. J.: 2002, 'Symplectic Algorithms for Accretion in Binary Star Systems', *Astron. J.* **123**, 2884–2894.
Ciesla, F. J., and Hood, L. L.: 2002, 'The nebular shock wave model for chondrule formation: Shock processing in a particle-gas suspension', *Icarus* **158**, 281–293.
Desch, S. J., and Connolly, H. C.: 2002, 'A model of the thermal processing of particles in solar nebula shocks: Application to the cooling rates of chondrules', *Meteoritics Planetary Sci.* **37**, 183–207.
Desch, S. J., and Cuzzi, J. N.: 2000, 'The generation of lightning in the solar nebula', *Icarus* **143**, 87–105.
Ford, E. B., Rasio, F. A., Yu, K.: 2004, 'Chaotic interactions in multiple planet systems', this volume.
Goldreich, P., and Tremaine, S.: 1980, 'Disk-satellite interactions', *Astrophys. J.* **241**, 425–441.
Gonzalez, G., Laws, C., Tyagi, S., and Reddy, B. E.: 2001, 'Parent Stars of Extrasolar Planets. VI. Abundance Analyses of 20 New Systems', *Astron. J.* **121**, 432–452.
Gounelle, M., Russell, S. S., Weiss, D., Mullane, E., Mason, T. F. D., and Coles, B. J.: 2001, 'Vanadium Isotopes as a Record of Early Solar System Irradiation: a Progress Report', *Meteoritics Planetary Sci.* **36**, A71.
Guillot, T., and Showman, A. P.: 2002, 'Evolution of "51 Pegasus b-like" planets', *Astron. Astrophys.* **385**, 156–165.
Halliday, A.: 2004, 'Time scales in the solar system', this volume.
Hartmann, W. K., Phillips, R. J., Taylor, G. J. (eds.): 1986, *Origin of the Moon*, Lunar and Planetary Institute, Houston.
Hartmann, W. K., Ryder, G., Dones, L., and Grinspoon, D.: 2000, 'The time-dependent intense bombardment of the Earth/Moon system', *in* R. M. Canup and K. Richter (eds.), *Origin of the Earth and Moon*, Univ. Arizona Press, Tucson, pp. 493–512.
Ikoma, M., Nakazawa, K., and Emori, H.: 2000, 'Formation of Giant Planets: Dependences on Core Accretion Rate and Grain Opacity', *Astrophys. J.* **537**, 1013–1025.
Levison, H. F., Lissauer, J. J., and Duncan, M. J.: 1998, 'Modeling the Diversity of Outer Planetary Systems', *Astron. J.* **116**, 1998–2014.
Lewis, J. S.: 1998, *Worlds Without End: The Exploration of Planets Known and Unknown*, Helix, Reading, pp. 7–28.
Lin, D. N. C., Bodenheimer, P., Richardson, D. C.: 1996, 'Orbital migration of the planetary companion of 51 Pegasi to its present location', *Nature* **380**, 606–607.
Lissauer, J. J.: 1993, 'Planet formation', *Annu. Rev. Astron. Astrophys.* **31**, 129–174.
Lissauer, J. J.: 1995, 'Urey Prize Lecture: On the Diversity of Plausible Planetary Systems', *Icarus* **114**, 217–236.
Lissauer, J. J.: 1999, 'How Common are Habitable Planets?' *Nature* **402**, C11–C14.
Lissauer, J. J.: 2004a, 'Formation of Giant Planets and Brown Dwarfs', *in* J.-P. Beaulieu, A. Lecavelier and C. Terquem, (eds.), *Extrasolar Planets, Today and Tomorrow*, ASP Conference Series 321, Astronomical Society of the Pacific, San Francisco, p. 271.
Lissauer, J. J.: 2004b, 'Concluding Remarks: Onwards Towards Extrasolar Earths!' *in* J.-P. Beaulieu, A. Lecavelier and C. Terquem, (eds.), *Extrasolar Planets, Today and Tomorrow*, ASP Conference Series 321, Astronomical Society of the Pacific, San Francisco, p. 421.
Lissauer, J. J., and Lin, D. N. C.: 2000, 'Diversity of Planetary Systems: Formation Scenarios and Unsolved Problems', *in* J. Bergeron and A. Renzini (eds.), *From Extrasolar Planets to Cosmology: The VLT Opening Symposium*, Springer, Berlin, pp. 377–390.
Lissauer, J. J., Pollack, J. B., Wetherill, G. W., and Stevenson, D. J.: 1995, 'Formation of the Neptune System', *in* D. P. Cruikshank (ed.), *Neptune and Triton*, Univ. Arizona Press, Tucson, pp. 37–108.
Luu, J. X., and Jewitt, D. C.: 2002, 'Kuiper Belt Objects: Relics from the Accretion Disk of the Sun', *Annu. Rev. Astron. Astrophys.* **40**, 63–101.

Marcy, G. W., Fischer, D. A., Butler, R. P., Vogt, S. S.: 2004, 'Systems of Multiple Planets', this volume.
Marcy, G., and Butler, P.: 1996, 'A Planetary Companion to 70 Virginis', *Astrophys. J. Letters* **464**, L147–L151.
Mayor, M., and Queloz, D.: 1995, 'A Jupiter-mass companion to a solar-type star', *Nature* **378**, 355–359.
Morbidelli, A., Petit, J. M., Gladman, B., and Chambers, J.: 2001, 'A plausible cause of the late heavy bombardment', *Meteoritics Planet. Sci.* **36**, 371–380
Owen, T.: 2004, 'A new constraint on the formation of Jupiter', this volume.
Podsiadlowski, P.: 1993, 'Planet formation scenarios', *in* J.A. Phillips, S.E. Thorsett and S.R. Kulkarni (eds.), *Planets Around Pulsars*, ASP Conference Series 36, Astronomical Society of the Pacific, San Francisco, pp. 149–165.
Pollack, J. B., Hubickyj, O., Bodenheimer, P., Lissauer, J. J., Podolak, M., and Greenzweig, Y.: 1996, 'Formation of the Giant Planets by Concurrent Accretion of Solids and Gas', *Icarus* **124**, 62–85.
Quintana, E. V., Lissauer, J. J., Chambers, J. E., and Duncan, M. J.: 2002, 'Terrestrial Planet Formation in the α Centauri System', *Astrophys. J.* **576**, 982–996.
Safronov, V. S.: 1969, *Evolution of the Protoplanetary Cloud and Formation of the Earth and Planets*, Nauka Press, Moscow (in Russian), English translation: NASA TTF-677, 1972.
Shu, F. H., Shang, H., Glassgold, A. E., and Lee, T.: 1997, 'X-rays and fluctuating X-winds from protostars', *Science* **277**, 1475–1479.
Stern, S. A., and Weissman, P. R.: 2001, 'Rapid collisional evolution of comets during the formation of the Oort cloud', *Nature* **409**, 589–591.
Strom, S. E., Edwards, S., and Skrutskie, M. F.: 1993, 'Evolutionary timescales for circumstellar disks associated with intermediate- and solar-type stars', *in* E.H. Levy and J.I. Lunine (eds.), *Protostars and Planets III*, Univ. Arizona Press, Tucson, pp. 837–866.
Thommes, E. W., Duncan, M. J., and Levison, H. F.: 1999, 'The formation of Uranus and Neptune in the Jupiter-Saturn region of the Solar System', *Nature* **402**, 635–638.
Thommes, E. W., Duncan, M. J., and Levison, H. F.: 2002, 'The Formation of Uranus and Neptune among Jupiter and Saturn', *Astron. J.* **123**, 2862–2883.
Ward, W. R.: 1986, 'Density waves in the solar nebula: Differential Lindblad torque', *Icarus* **67**, 164–180.
Ward, W. R.: 2004, 'Angular Momentum Constraints on Type II Planet Migration', this volume.
Weidenschilling, S. J.: 1977, 'Aerodynamics of solid bodies in the solar nebula', *Mon. Not. R. Astr. Soc.* **180**, 57–70.
Wetherill, G. W.: 1980, 'Formation of the terrestrial planets', *Ann. Rev. Astron. Astrophy.* **18**, 77–113.
Wetherill, G. W.: 1994, 'Possible consequence of absence of "Jupiters" in planetary systems', *Astrophys. Space Sci.* **212**, 23–32.
Wolszczan, A., and Frail, D.: 1992, 'A Planetary System around the Millisecond Pulsar PSR1257+12', *Nature* **255**, 145–147.
Wuchterl, G., Guillot, T., and Lissauer, J. J.: 2000, 'Giant Planet Formation', *in* V. Mannings, A. P. Boss and S. S. Russell (eds.), *Protostars and Planets IV*, Univ. Arizona Press, Tucson, pp. 1081–1109.
Young, R. E.: 2003, 'The Galileo Probe: How it has changed our understanding of Jupiter', *New Astron. Revs.*, **47**, 1–51.

Direct Detection Of Exoplanets: A Dream or a Near-Future Reality?

D. Rouan
LESIA - CNRS UMR 8109, Observatoire de Paris, 92195 Meudon cedex, France

Abstract. We focus here on the question of the direct imaging of exoplanets. Beyond its mediatic impact, this technique will bring several new constraints with respect to other methods and will open the door to spectroscopy and thus to the physics of the planets, ultimately to the search of life signatures. It is generally assumed to be in excess of 23 magnitudes, however, this contrast depends largely on wavelength and other parameters and we first examine how it varies in different situations of planet/star couples, using as far as possible realistic models of planetary evolution. The conclusion is that several very favorable cases do exist: they are the most promising for a direct detection in the near future. There is an amazing richness in concepts and genuine ideas to reach this goal. Of course, the huge difficulty lies in the contrast between the star and the nearby planet. Two main families of instruments are today considered for actually succeeding in planet imaging: a high performance coronagraph on a single telescope and a space nulling interferometer at thermal wavelengths. In the case of coronagraphy, we review briefly several experimental setups, with some illustration based on recent proposals of optimized ground-based and space experiments dedicated to that aim, using the interferometric four quadrant coronagraph that our group proposed. On the ground, even with an extremely powerful adaptive optics system (Strehl ration of 90 %), it is clear that the speckle noise will be the main limitation: contrast in magnitude as large as Δm $= 15$ are however possible in the K band, provided that one uses a combination of a coronagraph and differential imaging or polarimetry at two wavelengths: this is likely the most promising concept for direct planet detection from the ground. *Young* planets which are bright in the IR will likely be the first detected bodies in a not so far future. On the other hand, we show that with a dedicated coronagraph on a large space telescope, a classical Jupiter is indeed detectable from space at 20 μm for a star closer than 10 pc, while the more favorable cases of a young (hot) giant planet allows detection at 6 μm for a star belonging to the closest star forming region at 50 pc. For instance, the mid-IR instrument MIRI onboard the JWST (ex NGST), equipped with a coronagraph, will have this capability. The second avenue, nulling interferometry from space, is obviously a difficult one, but it is also a very promising one since it has the potential of detecting life tracers thanks to spectroscopy in the thermal infrared. It is a major space program, with years of technological research on difficult issues and it requires probably a joint effort between agencies to become one day a reality. Based on today's know-how, it is however possible to predict that projects like TPF or Darwin will have the capability to detect earth-like planets and to make their IR spectrum on one to two dozens of cases, after a monitoring of typically one hundred of nearby stars.

1. Introduction

By direct imaging, we mean the detection of genuine photons from the planet, unambiguously separated from those of the parent star. They could be either photons reflected by the planetary surface or those emitted through thermal radiation. The first direct imaging of extrasolar planets will certainly have an important media impact, but, more important, this capability will bring new pieces of information

on the physical parameters of these objects: albedo, orbital elements, combined information on temperature and composition of the atmosphere – even seasonal variations could be monitored – and ultimately detection of life tracers. It is known that the powerful technique of radial velocity measurements do select massive planets of rather short period, so that any other technique which does not have this bias is a complementary one; direct detection is one of them: it favours also big planets, but rather on distant orbits because of the contrast with respect to the star. On the other hand, combining several detection techniques has powerful advantages to understand the nature of the planets: for instance, in the case of HD209458B, the detection of the transit (Charbonneau et al., 1999; Henry et al., 1999), that came after the identification of the planet through radial velocity measurement, brought the confirmation that it was indeed a gas-giant planet with a mass and a size correctly predicted by models. Another important consequence of direct detection will be to obtain statistics on many systems. Finally, such a breakthrough will open the door to spectroscopy which is the way to actually tackle the physics of those objects (atmosphere composition, thermal structure) and possibly detect the signature of life: either life prerequisite such as H_2O and CO_2 or the signature of life itself, for instance through bands of chlorophyl analogs or O_2 and O_3 lines (Léger et al., 1996).

2. The Contrast Issue

The detection of faint extended sources or point-like companions near a bright astrophysical object is not an easy task and requires both a high angular resolution and a high dynamic range. Let's for instance consider a standard Jupiter at 10 pc around a G2V star: it is 23 magnitudes (i.e. 1.3×10^9 times) fainter, in the visible, than the star it orbits around and separated from it by only $0.5''$. This is typical of a tentative detection of the light reflected by the planet: the situation will improve with a larger planet radius, a medium orbit radius (in principle, the halo of scattered light decreases as r^{-3}, while the planet luminosity as r^{-2}), a large albedo. On the other hand, it could be the thermal emission of the planet that is looked for. In the Rayleigh-Jeans domain, then the contrast is roughly $5 \times 10^{-6}(R_p/R_*)^2(T_p/T_*)$: it will improve if the planet is more massive, younger (thus still hot) and on a wider orbit. Such differences in conditions explain why, depending on the wavelength of observation and the actual configuration (size, age, orbit radius, parent star, etc.) there is a large range of conditions that could make the life much easier than the 23 magnitudes contrast we first quoted. For instance, a young massive ($10M_{\rm Jup}$) planet, still on the contraction phase (Burrows et al., 1997) – and thus at a rather high temperature –, orbiting a cold M star and observed at 20 μm is now only 4.8 magnitude fainter than its star. Between those two extreme cases, there is a variety of conditions that could be potentially frequent and may represent the right door to enter the era of the direct detection of extrasolar planets. This is illustrated in

Table I. Flux ratio between a star and an orbiting giant planet at various wavelengths in the infrared. The ratio is expressed in magnitudes.

Star	λ (μm)	Young Jupiter 50 Myr 50 pc, 5 AU			Jupiter 5 pc, 1 AU	Old Jupiter 5 Gyr 10 pc, 5 AU		
		$1M_j$ 515 K	$5M_j$ 1094 K	$10M_j$ 1539 K	$1M_j$ 315 K	$1M_j$ 170 K	$5M_j$ 285 K	$10M_j$ 460 K
G2V	6	11.56	8.25	7.28	16.02	22.22	16.90	12.80
	10	9.35	7.46	6.79	12.01	16.25	12.30	10.25
	20	8.41	7.11	6.58	10.25	12.47	10.30	9.17
M2V	6	9.67	6.38	5.39	16.87	21.76	15.03	10.92
	10	7.54	5.66	4.99	11.63	14.45	10.50	8.44
	20	6.64	5.36	**4.81**	9.25	10.70	8.35	7.39

Table I which gives the contrast in magnitudes for different situations of a Jupiter-like planet at different ages, masses and distances, orbiting a solar type star or a colder M star.

2.1. Is Coronagraphy Needed: Do We Have to Hide the Star?

On the ground, if photon noise were the limit in the visible, then the detection of one earth requires one night of integration on a 30m telescope while imaging one Jupiter needs one night of integration on a 8m telescope. Indeed the speckle noise is the actual limit in any situation and makes the situation much more difficult, since a fraction of the starlight is scattered by apertures, dust, phase defects due to atmospheric turbulence, optics, etc. One shows that the enhancement in brightness produced by one earth is equivalent to one speckle produced by a phase defect of ... 2 picometer! This is awfully small. Atmospheric speckles, in fact, smooth out quickly by averaging and the dominant effect should come from uncorrected systematic phase errors: this is why it is believed that the dominant noise is likely quasi-static speckle noise that will suffer low frequency drifts (thermal evolution, for instance). Rather short exposures could then be required to sense the evolution so that one can cancel the phase distorsion by some servo system. Once the speckle noise is removed, at small angular distances the photon noise can dominate again and long integrations can lead to success. However, because of the short exposures, read-out noise would dominate at larger angular distances, while, in any cases, saturation on the star occurs with long or short exposures, because of the contrast which is looked for. This is the reason why we do believe that coronagraphy is mandatory to cancel out as far as possible the starlight, especially if we consider that with an interferometric coronagraph as proposed recently (Gay and Rabbia,

1996; Roddier and Roddier, 1997; Rouan et al., 2000), it is indeed possible to probe angular distances which are smaller than λ/D. Is detection of an Earth even possible from the ground? The challenge is incredibly difficult but not out of reach, R. Angel said at the last SPIE conference in Kona (Angel et al., 2003), provided that several conditions are met: a) we must use very large apertures ($D > 30$ m); b) since we are interested in few spatial frequencies only, adequate spatial filtering must be done; c) it is mandatory to control in real-time the phase by measuring the complex amplitude on the final imaging array (e.g. using interferometric techniques), the only way to probes the phase defects on the complete imaging path; d) using photon counting (e.g. Marconi L3CCD), then detecting 6 photons is enough for characterization of a given speckle; e) given the extremely small value of the phase variation one is looking for, typically a frame frequency of 4 kHz is needed; f) finally, using post-processing for optimum time filtering will improve the detection, but this will requires to store Tbytes of data. Can we share the optimism of Roger Angel? That is the question. Given this set of difficulties, one can guess that the first planet that will be directly detected will be a giant one and that a space experiment will almost certainly be required for the detection of an earth-like planet.

2.2. Two Main Avenues: Nulling Interferometry and Large Single Dish Coronagraphy

As most of the studies (e.g. TPF and Darwin) conducted today conclude, there are likely only two main avenues to explore: on one side nulling interferometry in the mid-infrared where the gain is on the better contrast in the infrared, while the large basis provides the angular resolution, on the other side coronagraphy or high dynamic imaging with a single dish telescope where the gain is in simplification of the instrumental setup. In the last case, this could be in the visible or near-IR in space or from ground-based large telescopes, provided high performance adaptive optics is used, or in the thermal IR, using cold telescopes in space. In all cases, however, a solution must be found for controlling the residual diffracted light from the bright source, either by cancelling it as far as possible, or by limiting its effect to a limited portion of the image, leaving a significant portion of the star environment unpolluted.

3. Coronagraphy on a Single Dish Telescope

In the recent years, there has been a wealth of ideas about coronagraphic solutions. One must mention of course the classical Lyot coronagraph where an opaque disk is used to block light, but it can be shown that it is unable to fulfil the requirements set by planet detection. The main reason is the closest angular distance reachable which cannot be smaller than the opaque disk radius, i.e. a few λ/D, while the

residual light remains still important. Among the different other approach, one is to make use of non-circular apertures combined with apodization. This is the case of the square pupil coupled to apodization (Nisenson and Papaliolios, 2001): in their ASA (Apodized Square Aperture) concept, an off-axis square telescope produces an intermediate pupil where a mask with a smoothly modulated transimission performs an apodization of the final PSF. To the same family belongs the Gaussian pupil shape (Jacquinot and Roizen-Dossier, 1964; Spergel, 2001) where the pupil features a strange shape delimited by combination of Gaussian profiles. In both cases, the idea is both to suppress the Airy rings and to concentrate the diffracted light in peculiar zones that leave a large fraction of the field free from any diffracted light and thus appropriate to detection of a faint object. Those approaches are interesting for a space experiment, and the ASA is one of the most serious candidate for the TPF mission. Laboratory tests are very encouraging, however they have not been totally proven against the effect of atmospheric turbulence or of phase defects due to pointing errors and imperfect optics. The manufacturing of a large pupil which must feature an excellent definition of the edges in another difficulty in the case of the Spergel solution. The question of the saturation of the detector at the star location may be a serious problem and, as in any apodization, there is a price to pay: first a loss of equivalent transmission that increases with the rejection factor and a reduction in actual resolution in terms of the PSF full width at half maximum which becomes typically 4–5 λ/D instead of λ/D. A second approach is to use a mask in the focal plane that modify the phase, rather than the amplitude as does the Lyot mask. The first concept of a coronagraph based on phase modification was proposed a few years ago (Gay and Rabbia, 1996). The principle is to produce a π phase shift on one of the arms of a Michelson interferometer. Shortly after (Roddier and Roddier, 1997), was proposed the idea of a stellar coronagraph with a phase mask, but now in the focal plane. The principle of this coronagraph is to put at the focal plane a small transparent disk of diameter $\approx 0.56\ \lambda/D$ (D is the telescope diameter) that produces a phase shift of π. In the four quadrants coronagraph (4QC) proposed more recently (Rouan et al., 2000; Riaud *et al.*(2003); Boccaletti et al., 2002,), the mask geometry is different but the basics principle is the same: half of the amplitude of the star PSF is changed of sign so that a destructive interference is produced – on the star only – while preserving the photons of the companion. Excellent performances are predicted and even total nulling can be obtained in the ideal case (circular pupil with no central obscuration and no phase distorsion) with the 4QC. Finally the solution of combining both techniques (phase mask and apodization) has been explored recently by Soummer and Aime (2002); they showed that prolate-spheroïdal functions could provide an excellent rejection and even a total one, when coupled to either a classical Lyot mask or better to a Roddier and Roddier phase mask. Those functions can also apply to either a circular pupil or to a square-shaped pupil. For all such solutions the magnitude of the rejection depends strongly on the quality of the wavefront: $\lambda/1000$ is a typical wavefront quality which is aimed at. This is an essential specification, that can hardly be found

at the required level in the real world, because of defects on the optics, atmospheric turbulence and pointing errors. A genuine proposal that does not belong to any of the families listed above is the use of an occulting mask at large distance from the telescope: this is the Free Flying Occulter proposed in the TPF study conducted by the TRW consortium. To avoid shadowing effects, the distances involved require a space experiment with two satellites separated by typically 10^5 km. The proposal by TRW consists in a NGST-like 8m telescope and a 70 m square occulter, both on an heliocentric orbit following the earth. The advantage of this solution is its relative simplicity, but obviously the operation of such a dual space system is not trivial and manufacturing of a large occulter that is also apodized is not the least of the problems.

3.1. An Example of Coronagraphic Device: The Four Quadrant Phase Mask

In this peculiar focal plane coronagraph, the transparent mask is made of four quadrants: two quadrants on one diagonal are without phase-shift and the two other experience a π phase-shift. The light from the bright source, when exactly centered on the mask, cancels out. This is possible because the amplitude has been split into two equal parts, one of them suffering a change of sign (due to the π phase shift): a destructive interference occurs then inside the geometric pupil area, provided that an appropriate Lyot stop is put at the pupil image to block most of the diffracted starlight. Any companion or structure (disk, for instance) which is not centered on the mask does not suffer any significant nulling and is transmitted with practically no attenuation. Fig. 1 illustrates the principle of the 4QC and what happens in the different planes (focal or pupil or conjugated of one or the other).

The four quadrants phase mask coronagraph (4QC hereafter) is a rather novel solution that appears very efficient since it is less sensitive to jitter than other phase masks and benefits from a peculiar property: the symetrization of the residual speckles. To improve very significantly (up to 6 magnitudes) the rejection of the remaining stellar light, a large fraction of the residual speckle pattern which is essentially centro-symmetric – in case of a very good correction – can be removed by subtracting the same image rotated of 180deg. This feature is analyzed in greater details elsewhere (Boccaletti et al., 2002). The 4QC is a compact device, easy to install in any imaging instrument. We'll describe in the next section two cases of instruments where a 4QC could bring the capability of planet detection in a rather near-future: the MIRI instrument aboard the JWST (ex-NGST) and the Planet Finder instrument that ESO plans to install on one VLT.

Achromaticity of the π phase shift may be a concern when a broad band measurement, using the 4QC, is looked for. Solutions have been proposed (Abe et al., 1998; Mawet et al., 2003), however, in all practical cases the limits in nulling performances imposed by other causes than chromaticity are largely dominant, so that

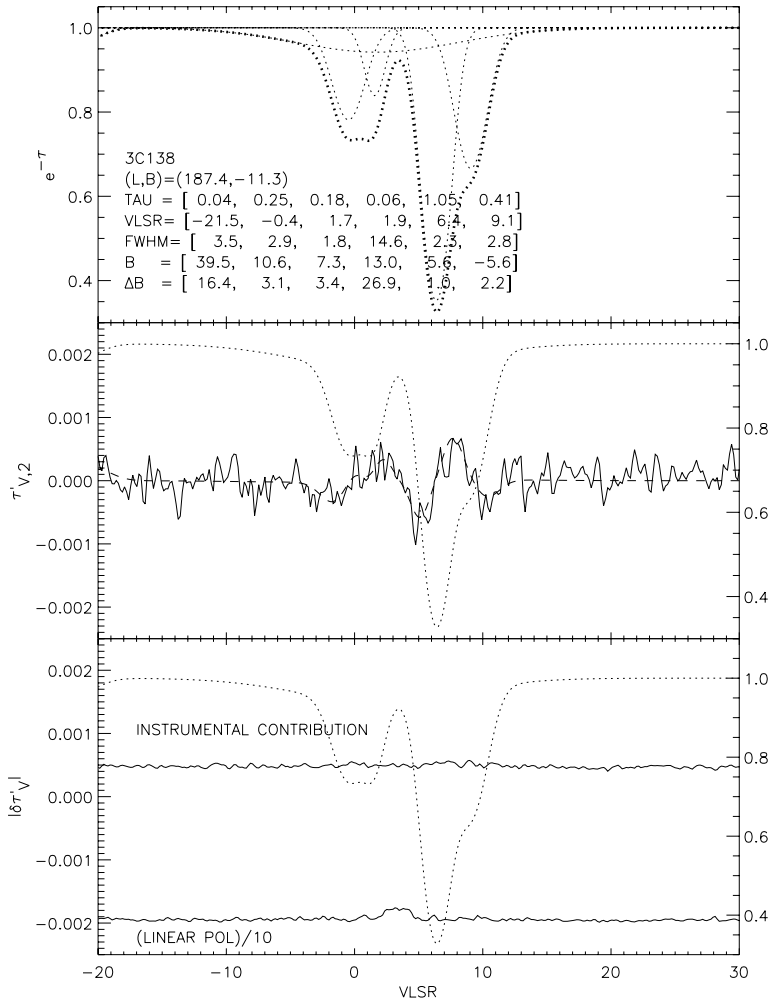

Figure 1. Principle of the 4QC. The optical scheme is shown, as well as images of amplitude or the intensity at different planes (focal or pupil planes).

a simple plate with a step in thickness of $\lambda/[2(n-1)]$ between adjacent quadrants is fully satisfactory, even with the spectral resolution of classical broad-band filters.

A demonstration experiment was developed at the Paris-Meudon Observatory in the visible (Riaud *et al.*(2003),). To manufacture the 4QC, we used a high quality plate of silica where a layer of the same material was deposited on two quadrants to form a step of thickness e, such that $(n-1)e = \lambda_0/2$, with $\lambda_0 = 650$ nm. The device was integrated in a bench where the different functions depicted on Fig. 1 are fulfilled with rather standard but high quality optics. In order to assess the detection of a faint companion, a secondary source, 300 times fainter, was created

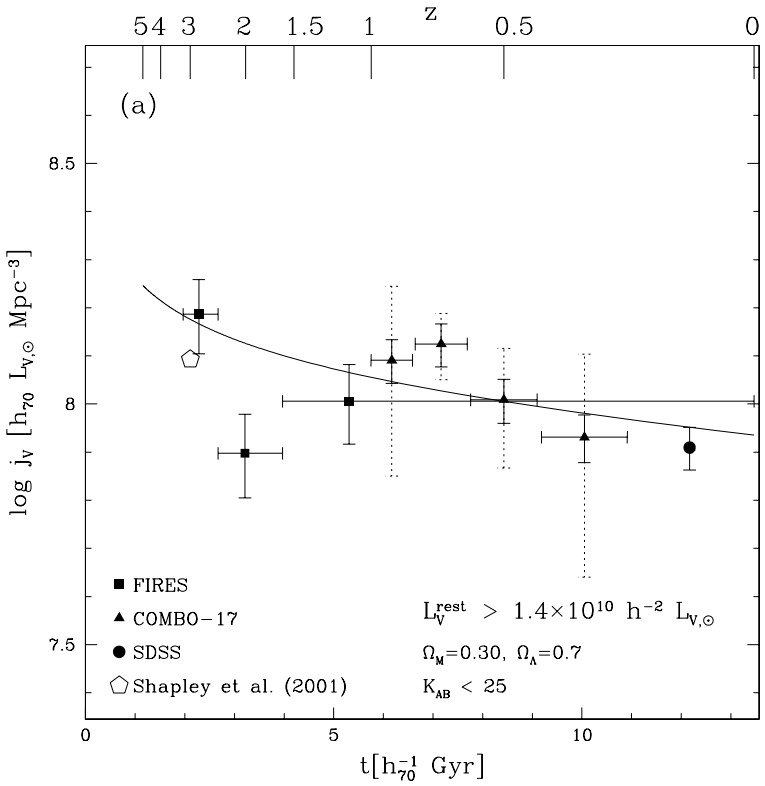

Figure 2. Example of result obtained in the lab with the experimental setup. A companion 300 times fainter than the artificial star appears 40 times brighter than the starlight residual when the star is exactly centered on the 4QC. The rejection in flux is 12500 and 44000 in peak brightness.

by a reflection on a plate. The image on Fig. 2 shows an example where the *planet* appears as a source 40 times brighter than the residual of the starlight: an extinction of 44000 was obtained at the location of the peak. This appears promising.

With a rejection factor as high as 12500, one could already pretend to try detecting hot extrasolar planets orbiting nearby stars. However, in the real life there are several limitation that prevent to actually reach such a high rejection factor. To begin with, actual telescopes have a central obscuration. For a ground-based instrument, the contribution of turbulent atmospheric tip-tilt dominates: a very efficient adaptive optics system is mandatory, but still this is the true limit. Indeed, a test of a 4QC on telescope was recently done (Puga et al., 2003): it confirmed that without a very low residual tip-tilt, the rejection factor can be fairly poor. For a space telescope, practically all sources of nulling degradation are at about a same level: the residual tip-tilt, the pupil shape (petals of the primary mirror on the JWST for example), and the surface roughness. We explore in the following two simulations of a ground-based and a space experiment, where those different factors are taken into account.

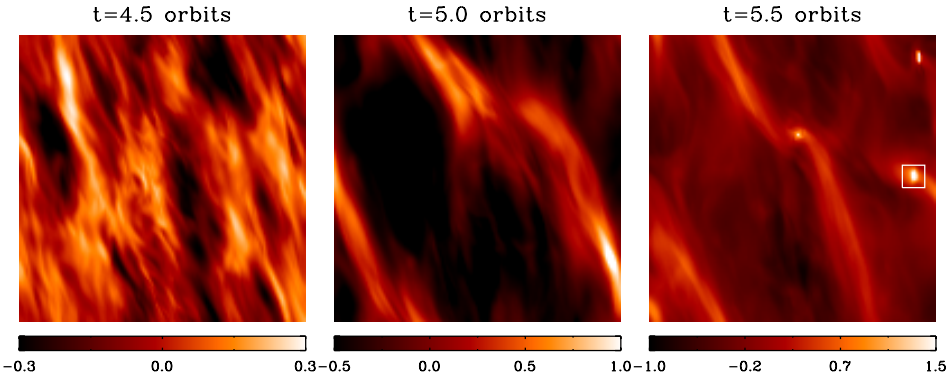

Figure 3. Example of a simulation using a 4QC done for the ESO Planet Finder study. Residuals of the brightness stellar light is plotted as a function of angular distance to the star, after substraction of the centro-symmetric pattern. The conditions of the simulation are: RON = 5e-, $\lambda = 2.2$ μm, $\Delta\lambda = 0.33$ μm, sky brightness: mK = 16 (arcsec)$^{-2}$, Strehl ratio = 90% , $r_0 = 1$ m, $\tau_0 = 10$ ms, overall efficiency = 0.4, integration time = 10 s. Vertical dashed lines are labeled in units of λ/D. The case of an optimized Lyot coronagraph (thick line) is overplotted

3.2. A Possible Concept for the ESO Planet Finder Project

ESO, that envisions second generation VLT instrument, selected two consortia to study the capability of an instrument – the Planet Finder – that would be dedicated to high contrast imaging, with peculiar emphasis on planet detection. We present here some preliminary results of the study conducted by one of the consortia leaded by a group of institutes in France (Lagrange et al., 2003). The basic concept is to have a top-level adaptive optics system associated in a first step to a coronagraph. Simulations have been performed concerning the capability of a 4QC coronagraph on such a dedicated instrument. We present on Fig. 3 a sample result of this simulation. A realistic Strehl ratio of 92 % was assumed: this is the maximum figure that one can expect to reach when everything is pushed to the limit of the current technology in adaptive optics (with e.g. 1000 actuators on the deformable mirror). The residuals, after 10 s of observation, are essentially dominated by speckle noise. This is one of the important conclusions of the study: residual speckle noise is always the main limitation in performances, when star brighter than 8–10 are observed. This means that reaching contrast larger than Δm = 17, the driving specification for exoplanet detection, requires additional techniques than simple coronagraphy, even at a high Strehl ratio. Differential imaging combined with coronagraphy is certainly one very promising way (Racine et al., 1999). This technique supposes to acquire simultaneously images in several narrow-band filters, with at least one centered on a spectral feature characteristic of the planet or of the star. The methane absorption near 1.6 μm is one of the most noticeable band to apply this technique.

Other differential techniques such as spectro-astrometry, and differential polarimetry are also promising ways to improve the rejection of residual speckles:

they should be explored. The second consortium that is working on the VLT PF concept (Feldt et al., 2003) is indeed exploring by simulation some of those techniques.

3.3. A CORONAGRAPH ON THE INSTRUMENT MIRI ABOARD THE JWST (EX NGST)

A team is presently studying the imager of MIRI, the mid-IR instrument of the JWST (ex-NGST), within the framework of the phase A study sponsored by ESA and NASA (Dubreuil et al., 2003). Several 4QC are envisioned to produce the coronagraphic functionality which is part of the high-level specifications of the imager in MIRI. This choice was made because of the poor capability of a classical Lyot disk in term of contrast (up to 10^4) and angular distance to the bright source that can be reached ($0.8\ \lambda/D$ vs. $3\ \lambda/D$). The status of the current phase A study concerning the 4QC is discussed in Boccaletti et al. (2003). Three 4QC at 10.5, 12.5 and 17 μm are envisioned in MIRI. A complete model of the nulling efficiency was developed with the aim of assessing in an exhaustive way different sources of nulling degradation: the effect of departure of the pupil from a disk (JWST primary is made of petals) and the phase aberrations, including pointing errors. Using realistic figures on this phase corrugation, a nulling performance of $\tau = 110$ is found when all phase errors are considered. We showed on the other hand that a 4QC made of a simple plate with steps in thickness of $\lambda_0/[2(n-1)]$ for two of the quadrants (in other words a perfect 4QC at λ_0), features an intrinsic nulling which is largely better (typically $\tau = 200 - 250$), whatever the concept for the telescope is (e.g. Lookheed Martin or TRW design): this simple design has thus been retained for MIRI.

The results of the simulation are very encouraging: we show on Fig. 4 the plot of the achievable contrast with a 4QC on MIRI at 6 and 20 microns. The brightness of the residual is plotted versus the angular distance to the star, on a magnitude scale giving directly the difference in magnitude between the star and the planet. In this figure, a M2V star is considered, but other cases have been studied. Symbols correspond to different conditions of distance of the star and temperature of the planet, according to its age, using models (Burrows et al., 1997). The main conclusion is that in several cases, the direct detection at 20 μm of a planet analog to a classical (old) Jupiter is possible. At this wavelength, despite the loss in angular resolution due to the longer wavelength, the detection of analog of our Jupiter is feasible, thanks to the much better contrast in flux. At 6 μm, it is rather the detection of young hot Jupiters which becomes optimum at this wavelength; in that case, the detection could be possible only around young stars, this is why we only considered a distance of 50 pc corresponding to the closest star forming region. The advantage of the 4QC of reaching very small angular distances is essential in this case.

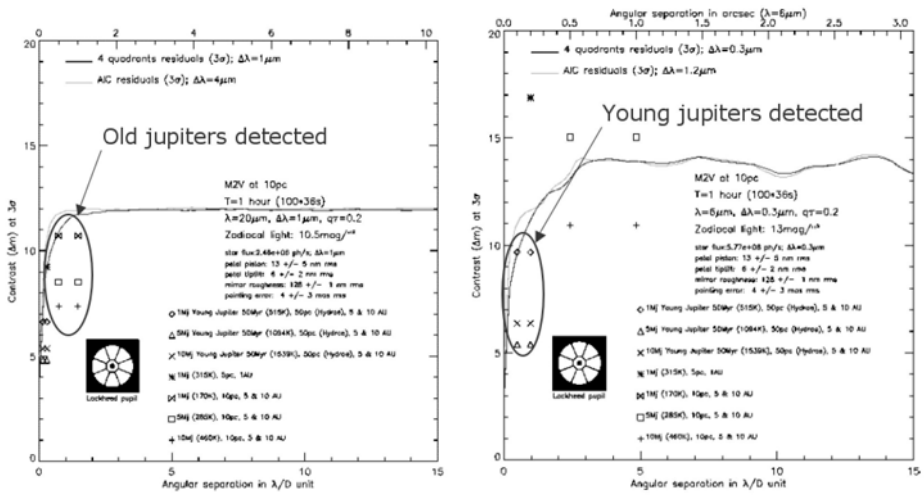

Figure 4. Left: plot of the achievable contrast with a 4QC on MIRI (JWST), to detect at 20 μm a planet of Jupiter's size, versus the angular orbital radius to its star (M2V); the symbols correspond to different cases of distance of the star and temperature of the planet, according to its age (models from Burrows et al. have been used). At this wavelength, despite the loss in angular resolution due to the long wavelength, the detection of analog of our Jupiter appears feasible, thanks to the much better contrast in flux. Right: same as previously, but at 6 μm: it is rather the detection of young hot (young) Jupiters which is optimum at this wavelength thanks to the angular resolution

4. The Second Avenue: Nulling Interferometry

Clearly, the solutions based on interferometry from space are more complex with more critical technological issues, but they are directed towards a more challenging goal: detecting and characterizing earth analogs (TPF, Darwin). The need for canceling out most of the stellar light is even more critical when an Earth has to be found, and the idea of a nulling interferometer, first proposed by Bracewell (1978), has been explored in a thorough way by various groups e.g. (Léger et al., 1996; Angel and Woolf, 1997; Guyon and Roddier, 2002; Riaud et al., 2002). The principle of the nulling interferometer is that it is possible to produce a dark central fringe when combining the light from two telescopes, by simply introducing a π phase shift on one of the optical path. Note that the recombination is done at a pupil level, in such a way that the fringe pattern is projected onto the plane of the sky than rather modulating the PSF as in a Fizeau interferometer. If now the star image corresponds to this dark fringe, while the planet location is rather on a bright fringe, then the contrast between the two objects can be improved by a huge factor. In fact, using only two telescopes is not sufficient because of the finite diameter of the star: there are leaks of stellar light for the external points of the disk which are not perfectly cancelled out because they are no longer at the center of the dark fringe. Using more telescopes in the recombination, placed on a well chosen pattern, allows to have a deeper nulling function that varies rather

as θ^4 than as the θ^2 law of the Bracewell original 2-telescopes concept: this is then sufficient to remove practically all stellar light. The rejection factor required to detect an earth around a solar type star in the thermal infrared is only 10^5, so that one other important source of noise is the speckle noise introduced by phase defect in the optics. A solution proposed by Mariotti et al. (1996) is to apply some spatial filtering using either a hole or fiber optics at the focal plane of each telescope before recombination: the high frequencies introduced by the phase corrugation at a small scale are efficiently filtered. Finally, introducing some temporal modulation is required in order to distinguish a putative planet from the exo-zodiacal light that must be likely present, as it is the case in our solar system. Continuous rotation of the whole interferometer is one solution: modulation of the signal by a planet passing through the fringe pattern projected onto the sky would be at a different frequency and significantly different shape from the modulation of the extended zodiacal light. Another solution, adopted for the present Darwin concept, is to have an internal modulation of the optical path difference. As a guideline, let's briefly examine the Darwin project, to have an idea of the technological issues and of the capabilities, as they are estimated today, after several years of studies in research institutes and in industry. The concept for the telescopes pattern is now the Laurance interferometer made of 6 elementary free-flier telescopes at the summits of an hexagon, with a central station for recombination. One important point that precursor missions will have to assess is the demonstration that precision formation flying is indeed possible. The background noise, which includes photon noise of the solar zodiacal light as well as fluctuations of the mean value, should be defeated by mean of an internal modulation of the OPD at a frequency ν_0, significantly higher than the characteristic break in the $1/f$ noise spectrum. One common problem to any system that intends to use a phase shift for the nulling of the stellar light is the achromaticity: indeed, producing an optical path difference which is proportional to the wavelength is not an easy task. Several solutions have been proposed, the most efficient appearing the use of the property of a π phase shift at a reflection on a surface or at the crossing of a focus, however polarization properties must be taken into account and rather sophisticated optical schemes have to be used, for instance with roof-top mirrors (Wallace et al., 2000). In the case of Darwin, as well of TPF, where the spectrum between 6 and 18 μm has to be measured, this question of achromaticity put severe constraints that require a vigorous R&D program. If all those specifications are satisfied, which apparently is not out of reach, then the expected number of detections can be evaluated assuming a mission duration of few years and a typical integration time – for detection only – of one day per target star. Fig. 5 summarizes those numbers in two cases: when the frequency of finding an earth is as high as 40% or when it is as low a 10 %; the strategy of sharing the time of the mission between detection and spectroscopy of detected planets is obviously different depending on which case will be found. The predicted number of earth-like planets that will be detected and spectroscopically studied is between 8 and 24 in this simulation.

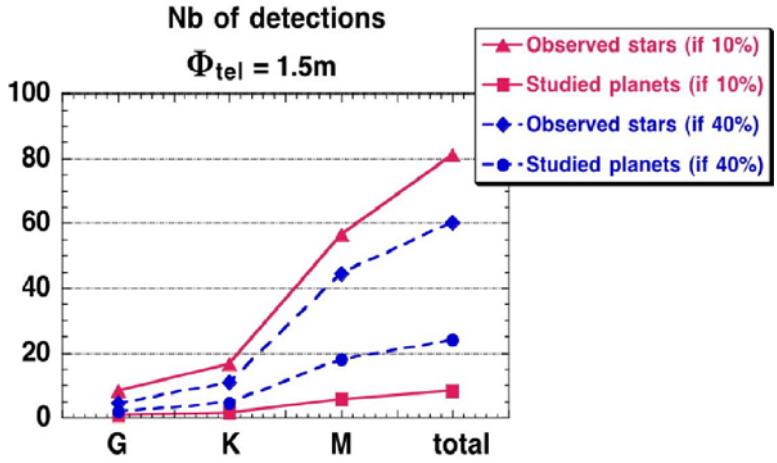

Figure 5. plot of the predicted number of detections that Darwin could achieve with 6 telescopes of 1.5 m diameter. Solid (red) and dotted (blue) curves are for the case where respectively 10 % and 40 % of stars have an earth-like planets. The lower curves correspond to the actual detection of a planet and the upper curves to the number of stars actually monitored (if the success of detection is larger than 40 %, then fewer stars will be looked at, so that the time devoted to spectroscopy on detected ones will be optimized).

5. Conclusion

The first direct detection of an extrasolar planet will certainly be a major mediatic event, but more important a breakthrough in the field because it will be the first step allowing further spectral analysis and thus detailed physics of planets, including the search for biomarkers. It will also open another window in the discovery space of exoplanets, essentially populated today by the massive planets discovered by the powerful radial velocity technique. The problem of the very strong contrast between the star and the planet is of course a difficult one from the observational point of view, but a point which is essential to realize is that there is an immense range of circumstances, differing by orders of magnitude in the difficulty to reach a direct detection: from the case of an Earth in the habitable zone detected in the visible from ground – an outstandingly challenging situation –, to the case of a *young* Jupiter on a wide orbits around an M star, detected from space in the thermal infrared – within reach of existing technology. The two instrumental concepts that are today the most seriously considered for direct detection, are an optical telescope with a high performance coronagraphic device on one side, and a nulling interferometer in the thermal infrared on the other. Several clever solutions have been proposed for the first family. The simulations we present on one of them, do show that from the ground a coronagraph associated to a top-level adaptive optics provide a high gain in dynamic (typically $\Delta m = 15$ magnitudes), but cannot be sufficient for planet detection because of the speckle noise which is largely dominant: it must be associated to another technique. A very promising one is

differential imaging between two wavelengths where the star and the planet differ largely (absorption features or polarization for instance): residual speckles should substract efficiently between two wavelengths. From space, it is not possible today to predict which solution will be the most successful (interferometry or coronagraph), however, realistic simulations prove that the JWST (ex-NGST) equipped with a four-quadrant coronagraph will likely achieve the detection of Jupiter-like planets in the 10–20 μm domain at the end of this decade. For having the chance to see an earth-type planet discovered and may be the proof that it hosts life, we will have to wait that all technical issues raised by the nulling interferometers in space be solved, i.e. probably not before the end of the next decade. However today there are no obvious technical stoppers as shown by the studies conducted for TPF or Darwin. We can thus be reasonably confident that the goal will be reached, especially if, for a major mission of this kind, a coordinated and/or joint efforts by NASA, ESA and other agencies can be put in place, keeping in mind that expensive and challenging programs in R&D, as well as precursors are mandatory.

Acknowledgements

I wish to thank the organizing committee for having invited me. Some Figures and results of simulations have been kindly communicated by colleagues: A. Boccaletti, P. Riaud, A. Léger; I wish to thank them warmly.

References

Abe, L., Vakili, F., and Boccaletti, A.: 1998, *A&A* **374**, 1161.
Angel, R.: 2003, SPIE meeting, Aug. 2002, Waikoloa, Hawaii.
Angel, R. and Woolf, N.: 1997, *ApJ* **457**, 373.
Boccaletti, A., Riaud, P., and Rouan, D.: 2002, *PASP* **114**, 132.
Boccaletti, A., Riaud, P., Rouan, D., and Baudrand J.: 2003, SPIE meeting, Aug. 2002, Waikoloa, Hawaii.
Bracewell, R. N.: 1978, *Nature* **274**, 780.
Burrows, A., et al.: 1997, *ApJ* **491**, 856.
Charbonneau, C., et al.: 1999, *ApJ* **529**, 45.
Dubreuil, D., et al.: 2003, SPIE meeting, Aug. 2002, Waikoloa, Hawaii.
Feldt, R., et al.: 2003, SPIE meeting, Aug. 2002, Waikoloa, Hawaii.
Henry, G., Marcy, G., Butler, P., and Vogt, S.: 1999, *ApJ* **529**, L41.
Gay, J. and Rabbia, Y.: 1996, *CR. Acad. Sci. Paris*, **332**, Serie II b, p. 265–271.
Guyon, O. and Roddier, F.: 2002, *A&A* **391**, 379.
Jacquinot, P. and Rozien-Dossier, B.: 1964, *Progress in Optics*, Vol. **3**, Amsterdam: North-Holland.
Lagrange, A.-M., et al.: 2003, SPIE meeting, Aug. 2002, Waikoloa, Hawaii.
Léger, A., et al.: 1996, *Icarus* **123**, 249.
Mawet, D., et al.: 2003, SPIE meeting, Aug. 2002, Waikoloa, Hawaii.
Mariotti, J.-M., et al.: 1996, *A&AS* **116**, 381.
Nisenson, P. and Papaliolios, C.: 2001, *ApJ* **548**, 201.

Puga, E., Feldt, M., Henning, Th., et al.: 2003, SPIE Meeting, Aug. 2002, Waikoloa, Hawaii.
Racine, R., Nadeau, D., and Doyon, R.: 1999, *PASP* **111**, 587.
Riaud, P., Boccaletti, A., Rouan, D., Lemarquis, F., and Labeyrie A.: 2001, *PASP* **118**, 1145.
Riaud, P., Boccaletti, A., Baudrand, J., and Rouan, D.: 2003, *PASP* **115**, 712.
Riaud, P., et al.: 2002, *A&A* **396**, 345.
Roddier, F. and Roddier, C.: 1997, *PASP* **109**, 815.
Rouan, D., Riaud, P., Boccaletti, A., Clénet, Y., and Labeyrie, A.: 2000, *PASP* **112**, 1479.
Rousset, G., Lacombe, F., Puget, P., Gendron, E., Arsenault, R., et al.: 2000, *Proc. SPIE*, Vol **4007**, 72.
Spergel, D. and Kasdin, J.: 2001, *AAS Meeting* **199**, #86.03.
Soummer, R. and Aime, C.: 2002, *A&A* **389**, 334.
Wallace, K., Hardy, G., and Serabyn, E.: 2000, *Nature* **406**, 700.

Planets With Detectable Life

Tobias Owen
Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, HI 96822, USA

1. Introduction

This workshop honors Michel Mayor for his watershed detection of the first extrasolar planetary system in 1995. It is worth remembering that prior to this discovery, there was an intense debate among well-qualified scientists as to whether or not other planetary systems existed. Now we have come to know over 100 planets circling other stars, with more discoveries announced almost every month.

The search for life outside the solar system is currently stuck in the same debate that once surrounded the quest for other planetary systems. Excellent scientists support a spectrum of views ranging from a strong belief that life is such an improbable state of matter that life on Earth is probably all the life there is, through grudging admission that there may be other planets inhabited by colonies of bacteria, to a sense that millions of technically advanced civilizations are contemplating this same question throughout the galaxy (Goldsmith and Owen, 2002).

Unfortunately, theories for the origin of life are in much worse shape than theories for the origin of planets. Thus the predictive abilities of science are even more severely limited in this case. We need to "do the experiment", we have to go out and look for signs of life.

2. Life As We Know It

What are those signs? Life on Earth may not be the only life in the universe, but it is the only life we know. Hence we must start with a consideration of life as we know it, attempting to generalize from its most basic characteristics to find constraints that environments must satisfy to be habitable and to identify the observable effects of living systems on those environments. As we are in no position to travel to extrasolar planets, the effects we seek are essentially confined to changes in atmospheric composition that we can detect spectroscopically from our remote vantage point. For example, the fact that our atmosphere is 21 percent oxygen would be a sure indication of life on Earth to an alien spectroscopist on the other side of the galaxy. Without the green plants to keep producing it, our present supply of oxygen would disappear in several million years, just 0.1% of geological time.

The first thing to keep in mind is that despite its dizzying diversity to the naked eye, at the molecular level all life on Earth is fundamentally the same. It all uses

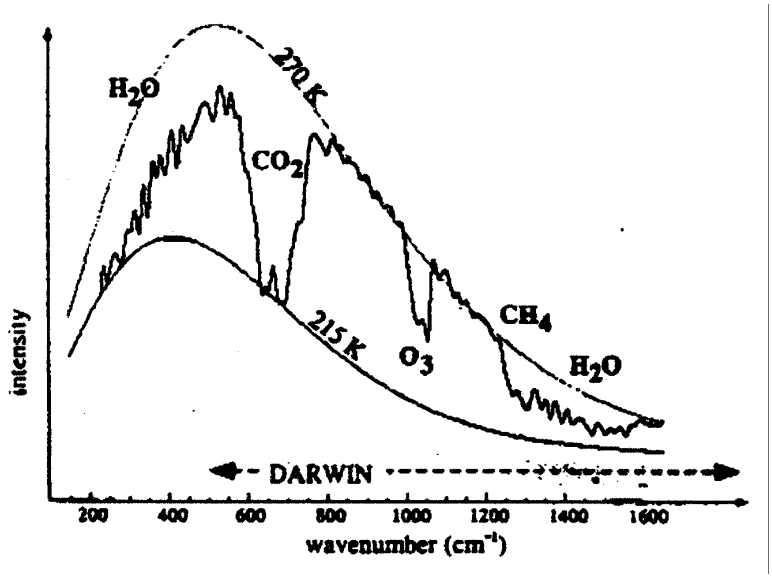

Figure 1. The infrared spectrum of the Earth, recorded by the TES instrument on the Mars Global Surveyor spacecraft (Christensen and Pearl, 1997; Selsis, 2000). Note the strong bands of CO_2, O_3 and CH_4 at concentrations of 345, 0.05 and 2 ppm, respectively. Blackbody curves for 270 K and 215 K are shown for comparison. The projected spectral coverage of the Darwin interferometer is shown above the x-axis.

the same 20 amino acids to build proteins and it all uses the same genetic code, carried by DNA and RNA. At the elemental level, it is even simpler. Just four elements, carbon, nitrogen, oxygen and hydrogen make up more than 95% (by mass) of living matter. This is certainly a positive sign for those who believe in the prevalence of life on other worlds, as these are the four most abundant chemically active elements in the universe. If life depended on holmium and hafnium, one would not be optimistic about its cosmic abundance.

Looking into the fundamental chemistry of life, we are struck by its dependence on carbon as the main compound-forming element and water as the essential solvent. Everywhere on Earth we find water there is life, the only places that are sterile are marked by the absence of this essential substance. A quick examination of the properties of carbon and water suggests that life's dependence on them is no coincidence.

In addition to its cosmic abundance, carbon has the essential ability of being able to form the complex molecules needed for information storage and transfer, as well as others that become structural elements and catalysts, and still others that store and transfer energy. These molecules are easily formed, they are stable but not too stable - the downfall of silicon which is often proposed in science fiction as an alternative to carbon. To illustrate carbon's gregariousness, it is enough to compare the two stable compounds that nitrogen can form with hydrogen (NH_3

and N_2 H_4) with the enormous number (not rigorously known) that carbon can achieve: CH_4 C2 H2 .... C_{90} H_{154} (enneaphyllin) .... etc. And this is just C and H!

Water is also replete with vital virtues. After H_2 , it must be the most abundant molecule in the universe. Its unusually high dielectric constant makes it an excellent solvent, it is liquid over a large range of temperature and it's the right range for the reactions that drive the complex chemistry of living systems. Water vapor is an excellent greenhouse gas, its high latent heat of vaporization helps moderate the climate and regulate the internal temperatures of the organisms that use it. Ice floats, so life in water can survive cold weather and planetary ice ages.

We conclude that life on other worlds is highly likely to use carbon and water, while nitrogen's cosmic abundance suggests it will also be a component of other living systems just as it is on Earth. We are not being overly restrictive with this conclusion because we don't expect life elsewhere to use the same *compounds* of C, H, N, O as we find on Earth. We don't expect the same amino acids, proteins, DNA, RNA, etc. It's not going to look like us and could well be inedible and smell terrible, but life elsewhere will probably rely on carbon as its main structural element, and water as its solvent, and incorporate nitrogen in its compounds. So where will we find it?

3. "Had We But World Enough And Time..."

Andrew Marvell was talking about love when he wrote that line, but it applies as well to life. We need a planet to provide the habitat for life and we need the global climate on that planet to be reasonably stable for at least 4.5 billion years if we hope for intelligent life. We also need a long-lasting source of energy-thermal, chemical, or best of all, star light.

These requirements and desiderata translate to the need for an Earth-like planet in a nearly circular orbit at the appropriate distance from a sun-like star. We need a star like the sun because we want it to have a main sequence lifetime of at least 4.5 billion years, which requires stars of spectral type later than F5. M dwarfs have immensely long life times but their habitable zones are tiny and close to the stars. Hence we prefer stars of types F5 – K5, which still leaves us with some 10's of billions of stars in our galaxy alone. The distance from the star has some latitude. The deeper requirement in this case is to allow liquid water to exist on the planet's surface, which translates to a range of distances that define an annulus around the star (actually the space between two concentric spheres) known as the star's *habitable zone.* Stars slowly get hotter with time, so the habitable zone gradually moves outward. Furthermore, the temperature of a planet will also depend on the composition of its atmosphere: how much of a greenhouse effect it can produce. Thus the outer boundary of the habitable zone is not sharply defined.

In our own solar system, Venus (at 0.72 AU) is outside the habitable zone. It is too close to the sun for water to be stable on its surface. Instead, a runaway

greenhouse effect caused any original oceans to boil, and the resulting overheated atmosphere allowed water vapor to reach altitudes where it was easily dissociated by ultraviolet photons from the sun. The hydrogen escaped, leading to a 150 times enrichment of D/H on the planet, which has been observed (Donahue et al., 1982; de Bergh et al., 1991).

On the other hand Mars (at 1.5 AU) is still within the zone. The problem for Mars is not that it's too far from the sun, it is too small to sustain the thick atmosphere that would provide the necessary greenhouse effect to keep it warm. An Earth-size planet in the orbit of Mars could be habitable.

Even the present Mars may have harbored the origin of life early in its history during an episode when the atmosphere was thicker and liquid water ran across the surface and pooled in its impact craters (Owen, 1997). Just how warm and wet Mars was in that early time and how long the periods of temperate climate lasted are still hotly debated (Squyres and Kasting, 1994; Forget and Pierrehumbert, 1997; Owen and Bar-Nun, 2000). As this stage in our ignorance we can even hold out the hope that life evolved to forms that survived in warm, wet regions underground, just as life has done on Earth (Onstott et al., 1999). The discovery that liquid water may still erupt from time to time onto the Martian surface (Malin and Edgett, 2000) adds to this hope. With such heady prospects, Mars remains an exciting target in the search for life in the solar system.

A less likely but still intriguing target is offered by Jupiter's icy satellite Europa. This moon is warmed from the inside by the dissipation of tidal energy from Jupiter, the same engine that drives the astonishing volcanic activity on the inner satellite Io. Here the thought is that beneath the icy crust of Europa there may be a warm ocean of water, and in that ocean there could be life (Gaidos et al., 1999).

This idea is driven by an analogy with the submarine vents on Earth, where a profusion of life is found at depths that are well beyond the reach of the sun. There are two problems with this idea, however. The first is that our submarine vents have lifetimes of only a hundred years or so. It seems doubtful that life could *originate* in such a timescale, particularly when the chemicals required would be so diluted by the surrounding surfeit of water. The second point is that evidence suggesting that the most ancient ancestor of life on Earth originated under high temperature conditions such as those at a submarine vent has been severely challenged by studies of the cross linking bases in RNA. Using computer models, Galtier el a. (1999) found that cytosine and guanine linkages were less prevalent in ancient life forms than they are in heat loving organisms today . Those C-G links are more heat-resistant than adenine-uracil links. Finally, it is worth noting that Ganymede and Callisto, the outer two of the giant Galilean moons do not have atmospheres of nitrogen and methane, unlike Titan, Saturn's largest satellite. So it may well be that the moons in Jupiter's retinue are deficient in the carbon and nitrogen life seems to need.

Nevertheless, the recognition that tidal energy from a giant planet could make one of its satellites habitable is a new idea, and broadens our definition of the habitable zone.

4. Finding Life On Brave New Worlds

So now we have convinced ourselves that the stuff of life is plentiful in the galaxy and we have defined the constraints that a life-bearing planet must satisfy. In our solar system, we have found only one planet that has life on it, Earth, and another that is at least in the habitable zone of our star, but may (or may not!) be too small to sustain a viable ecosystem. What about all those other worlds out there whose existence we now know?

The sad fact is that all of these super-giant planets are themselves totally unsuited to be adobes of life. Not only that, most of them have migrated through the habitable zones of their systems to their present positions, thereby wiping out any Earth-like planets that might have existed. The other giants occupy eccentric orbits that will again cause them to wreak havoc on the types of planets we seek, or prevent them from forming in the first place. As of 1 January 2003, we are still lacking a system with a giant planet in a nearly circular orbit at a distance of about 5 AU from its star, with no other giants between it and the star. The good news is that we could only hope to detect such a system the last year or so, when the accumulated observations would have covered enough of the giant planet's orbit to make identification certain. Thus we may hope for the first discoveries of systems whose basic architecture is like our own in the very near future.

This will be reassuring, but we will still not have the certainty that these systems have Earth-size planets in their habitable zones. This information will come from the Kepler (see Borucki, this volume) and Eddington missions, which will start returning data on transits of terrestrial planets in 2007. At that point, we will be able to make sound, statistical arguments for the number of Earth-like planets in the galaxy. The next step will be the use of interferometers in space to separate these planets form the glare of their stars, allowing us to do spectroscopy to see what gases their atmospheres contain.

What gases do we seek? Once again we revert to the single example of life that we know. We have seen that plentiful oxygen is a sure sign of life. Trees and grass are not required to produce it. The blue-green bacteria produced the first global abundance of oxygen on Earth some 2.5 billion years ago, and their descendants are still releasing oxygen today. Similar organisms could be thriving on Earth-like planets throughout the galaxy. Other bacteria on our own planet also give themselves away by the gases they produce, as the 2 ppm of methane in our atmosphere testifies. This methane is produced by a certain type of bacteria known (obviously) as methanogens, that live primarily in swamps and in the guts of grass eating animals. Other bacteria generate hydrogen sulfide, while still others produce

other gases that would not survive in the atmosphere of a planet in a habitable zone unless they had a continuous source.

This perspective defines our task: we are looking for gases that would not exist in planetary atmospheres if living organisms were not present to produce them.

The current protocol for the two space-born interferometers under consideration, the Terrestrial Planet Finder (TPF) and Darwin , is to focus on ozone as a surrogate for O_2 (Leger, 2000). The great advantage of this approach is that the strongest O_3 absorption falls at 9.6 μm, a region of the spectrum where the contrast between the brightness of the planet (from its thermally emitted IR radiation) and the brightness of the star is at a minimum. Methane has its strongest band nearby at 7.6 μm , so both gases can be searched for along with the prominent CO_2 band at 15 μm (Fig. 1). Other potentially interesting disequilibrium gases such as N_2O , NH_3 , and SO_2 can also be detected in this same spectral region (Selsis, 2000). Should it prove possible to obtain spectra of candidate planets at 0.76 nanometers, direct detection of O_2 would become feasible at rather low spectral resolution (Owen, 1980).

5. False Positives

How can we be absolutely certain that the discovery of a disequilibrium gas in a planetary atomsphere is really a sign of life? In fact, there are other ways these gases could be generated, but we can eliminate these false positives with proper care.

Consider oxygen. Huge amounts of this gas will be produced by the runaway greehouse phenomenon we discussed when considering the habitability of Venus. This will be a temporary effect, however, as the oxygen liberated by the photolytic destruction of water vapor in a planet's upper atmosphere will soon combine with crustal rocks. There are several ways to eliminate the potential confusion. The first step might be to use the planet's distance from its star and the spectral energy distribution and surface temperature of the star to calculate the planet's surface temperature. This would tell us if a runaway were possible. We could look for the presence of large amounts of CO_2 together with the oxygen in the planet's atmosphere, which would signal that a runaway was indeed underway. Most convincingly, if we could measure the planet's surface temperature, there would be no doubt about the source of its atmospheric oxygen.

In the case of methane and other reduced gases, the key parameters will be the size of the planet, its distance from its star, and the age of the star. We are still uncertain about the composition of the Earth's earliest atmosphere, but if it ever was in a reducing state, it would have been during the first billion years of the planet's history. In this early epoch, bombardment by volatile-rich planetesimals might have been able to generate reducing conditions if the planet were sufficiently cold to minimize the amount of water vapor in its atmosphere. Once the planet warms up,

water vapor gets into the atmosphere in sufficient amounts to generate OH though photolysis and this molecular fragment will oxidize any reduced compounds such as CH_4 .

Thus any inner planet in any system throughout the galaxy will ineluctably convert to a CO_2 -dominated atmosphere if life fails to develop. We see this clearly with Mars and Venus in our own system. So if we find a significant amount of methane in the atmosphere of a distant Earth-sized planet that is in the habitable zone of a star over one billion years old, we can be quite certain that we are looking at an inhabited world. The inhabitants may simply be bacteria, but they would nevertheless demonstrate that the transformation of non-living to living matter was not a unique event in the galaxy.

6. The Future

All of these possibilities and surely others as well, will receive detailed scrutiny and elaboration before we actually undertake the first direct searches for evidence of life on other worlds like ours. We can identify some of the steps along the way to this goal, with an approximate time table as follows:

2003–2011 Continuing investigations of Mars by orbiters, landers and rovers, leading ultimately to the return of samples from Mars.

2007 Launch of the Kepler and Eddington missions. First specific searches for Earth-like planets around sun-like stars.

2015 Launch of a Space-borne interferometer, modeled on TPF + Darwin: We begin a systematic search for life on other worlds.

This is the formal outline of what the agencies are planning. But if past history is any kind of guide, individual astronomers using ever more sophisticated and capable ground-based instruments may actually "jump the queue" and give us an answer much sooner. And of course if that long-awaited message from a technologically advanced civilization somewhere else in the galaxy finally makes its way to our anxiously searching radio telescopes, we could have our answer tomorrow!

The point of all this is that we have indeed reached the stage in the development of our own civilization where experiment is taking the place of speculation in our deeply rooted quest for life elsewhere in the universe. Is living matter a miracle or a commonplace phenomenon? By the year 2020 we should surely know.

References

Christensen, P. R. and Pearl, J. C.: 1997, 'Initial Data from the Mars Global Surveyor thermal emission spectrometer experiment: Observations of the Earth'. *J. Geophys. Res.* **102**, 10875–10880.

de Bergh, C., Bezard, B., Owen, T., Crisp, D., Maillard, J.-P., and Lutz, B. L.: 1991, 'Deuterium in Venus: Observations from Earth.' *Science* **251**, 547–549.

Donahue, T. M., Hoffman, J. H., Hodges, R. R. Jr., and Watson, A. J.: 1982, 'Venus was Wet: A Measurement of the Ratio of D/H' *Science* **216**, 630–633.

Forget, F., and Pierrehumbert, R. T.: 1997, 'Warming Early Mars with CO_2 Clouds that Scatter Infrared Radiation', *Science* **278**, 1273–1276.

Gaidos, E. J., Nealson, K. H., and Kirshvink, J. L.: 1999, 'Life in Ice-covered Oceans' *Science* **284**, 1631–1633.

Galtier, N., Tourasse, N. and Gouy, M.: 1999, 'A Nonhyperthermophilic Common Ancestor to Extant Life Forms' *Science* **283**, 220–223.

Goldsmith, D. and Owen, T.: 2002*, The Search for Life in the Universe, 3e,* University Science Books, Sausalito, 573 pp.

Léger, A.: 199, 'Jean-Marie Mariott and Robin Laurance, Pioneers of the Darwin Mission' in *Darwin and Astronomy, The Infrared Space Interferometer,* ESA SP-451, pp. 5–10.

Malin, M. C. and Edgett, K. S.: 2000, ' Evidence for Recent Groundwater Seepage and Surface Runoff on Mars', *Science* **288**, 2330–2335.

Onstott, T. C., Phelps, T. J., Kieft, T., Colwell, F. S., Balkwill, D. L., Fredrickson, J. k., and Brockmann, F. J.: 1999, in J. Seckbach, (ed.) *Enigmatic Microorganisms and Life in Extreme Environments*, Kluwer, Dordrecht pp 487–500.

Owen, T.: 1997, 'Mars: Was there an ancient Eden?' in C. Cosmovici, S. Bowyer, D. Wertheimer, (eds.) *Astronomical and Biochemical Origins and the Search for Life in the Universe,* pp. 203–218.

Owen, T.: 1980, 'The Search for Early Forms of Life in Other Planetary Systems: Future Possibilities Afforded by Spectroscopic Techniques', in M. Papaglannis (ed.), *Strategies for the Search for Life in the Universe*, D. Reidel, Dordrecht, pp. 177–185.

Owen, T. and Bar-Nun, A.: 2000, 'Volatile Contributions from Icy Planetesimals' in R. M. Cancup and K. Righter, (eds.) *Origin of the Earth and Moon, University* of Arizona press, Tucson, pp. 459–471.

Selsis, F.: 2000, 'Physics of Planets I: Darwin and the Atmospheres of Terrestrial Planets' in *Darwin and Astronomy, The Infrared Space Interferometer*, ESA SP-451, pp. 133–140.

Squyres, S. W., and Kasting, J. F.: 1994, 'Early Mars: how warm and how wet?' *Science* **265**, 744–749.

Outlook: Testing Planet Formation Theories

A. P. Boss
Department of Terrestrial Magnetism, Carnegie Institution, 5241 Broad Branch Road, N.W., Washington, D.C. 20015-1305 U.S.A.

Abstract. The discovery of the first planetary companion to a solar-type star by Mayor and Queloz (1995) launched the extrasolar planetary systems era. Observational and theoretical progress in this area has been made at a breathtaking pace since 1995, as evidenced by this workshop. We now have a large and growing sample of extrasolar gas giant planets with which to test our theories of their formation and evolution. The two competing theories for the formation of gas giant planets, core accretion and disk instability, appear to have testable predictions: (i) Core accretion seems to require exceptionally long-lived disks, implying that gas giants should be somewhat rare, while disk instability can occur in even the shortest-lived disk, implying that gas giants should be abundant. The ongoing census of gas giants by the spectroscopic search programs will determine the frequency of gas giants on Jupiter-like orbits within the next decade. (ii) Core accretion takes millions of years to form gas giants, while disk instability forms gaseous protoplanets in thousands of years. Determining the epoch of gas giant planet formation by searching for astrometric wobbles indicative of gas giant companions around young stars with a range of ages ($\sim$ 0.1 Myr to $\sim$ 10 Myr) should be possible with the Space Interferometry Mission (SIM). (iii) Core accretion would seem to be bolstered by a higher ratio of dust to gas, whereas disk instability occurs equally well for a range of dust opacities. Determining whether a high primordial metallicity is necessary for gas giant planet formation can be accomplished by spectroscopic and astrometric searches for gas giants around metal-poor stars. Eventually, ice giant planets will be detectable as well. If ice giants are found to be much more frequent that gas giants, this may imply that core accretion occurs, but usually fails to form a gas giant. Terrestrial planets will be detected through photometry by Kepler and Eddington, astrometry by SIM, and imaging by Terrestrial Planet Finder and Darwin. Ultimately these detections will clarify the process of Earth formation by collisional accumulation, the only contending theory.

1. Introduction

Given the wealth of knowledge about our own Solar System, nearly all work on the theory of planetary system formation has been focused on our own system. Little attention was paid to how the planet formation process might operate under other circumstances, with the noteworthy exception of Wetherill (1996), who studied the formation of terrestrial planets in the case of central stars and protoplanetary disks with varied masses, and with widely varying assumptions about the location of any gas giant planets in the system. This single-mindedness has led to a fairly mature, widely-accepted theory of terrestrial planet formation by the collisional accumulation of progressively larger solid bodies (Wetherill, 1990). Even after decades of research, however, there are two very different hypotheses about how the gas giant (Jupiter and Saturn) and ice giant (Uranus and Neptune) planets formed.

Beginning with the first discovery of an uncontested extrasolar gas giant planet by Mayor and Queloz (1995), the attention of many planetary formation theorists

have shifted toward trying to understand the origin of the often unexpected properties of extrasolar planetary systems. As a direct result of entering this new era, theories of planetary system formation and evolution are beginning to evolve in order to encompass these new planetary systems. The desire to create a unified theory of planetary system formation, applicable both to extrasolar systems and to our Solar System, will undoubtedly lead to creative tensions that will test our most basic ideas of how these processes occurred.

In this workshop summary, we will review the current status of theoretical work on gas and ice giant planet formation by core accretion and disk instability, the two competing mechanisms. Given the difficulty of deciding between these two competing mechanisms purely on the basis of theoretical arguments, we note that observations of extrasolar planets and star-forming regions must play the central role in deciding between these two mechanisms. We thus conclude by pointing out a number of observational tests that should be applied to help settle the issue of planetary origins. This summary is based in part on a recent review paper (Boss, 2002b).

2. First Census of Neighboring Planetary Systems

The first confirmed discovery of an extrasolar planet orbiting a solar-type star was that of 51 Pegasi's $\approx 0.5 M_J$ (M_J = Jupiter's mass) companion (Mayor and Queloz, 1995). Subsequent discoveries have come so rapidly that review articles are quickly outdated – Marcy and Butler's (1998) review listed only 8 extrasolar planet candidates, whereas roughly 100 planet candidates exist as of September, 2002 (Figure 1). Most of the latest discoveries are either comparable to or greater than Jupiter in mass, or have semi-major axes greater than several AU, implying that most of the higher mass, shorter period planets have already been found, at least for those stars which have been studied for several years or more.

Prior to 1995, the only example of a planetary system was our own, and so it was natural for theorists to concentrate on forming the Solar System's planets. As is evident from Figure 1, there exist a number of planetary systems which do not even remotely resemble our own. The very first Jupiter-mass planet discovered, 51 Pegasi's planet (Mayor and Queloz, 1995), has a semimajor axis of ≈ 0.05 AU, 100 times smaller than that of Jupiter, and a surface temperature ~ 10 times higher, because of its proximity to its star. Several more such "hot Jupiters", and even "hot Saturns", have now been found. Extrasolar planets seem to be more or less evenly distributed throughout the range of semimajor axes from about 0.04 AU to 3 AU (at least when viewed in terms of the log of the separation), with many of those orbiting outside 0.1 AU having orbits which are significantly more eccentric than that of Jupiter. Assuming a random distribution of planetary orbital inclinations, the median true masses of these objects are likely to be about 30% greater than the

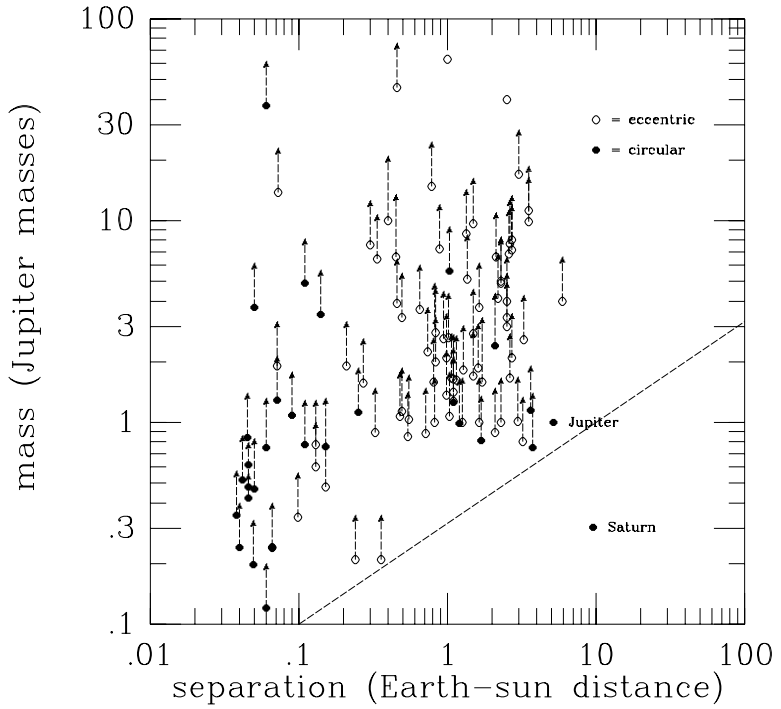

Figure 1. Discovery space for extrasolar planets and brown dwarfs as of September, 2002. The oblique dashed line illustrates the current sensitivity limit for spectroscopic detections, showing that current limits are sufficient to detect true Jupiter-analogues. Filled circles represent roughly circular orbits, while open circles represent eccentric orbits. All of these objects were found by the spectroscopic method (Mayor and Queloz, 1995; Marcy and Butler, 1998) which yields only a lower limit on the companion's mass. Objects with masses above $\sim 13 M_J$ can burn deuterium and hence may be best classified as brown dwarf stars, rather than as gas giant planets. Nearly all of these objects are in orbit around solar-mass main sequence stars.

minimum masses found by Doppler spectroscopy (Figure 1), so that many of these planets are considerably more massive than Jupiter, by up to a factor of 10 or so.

However, the discoveries to date have also been reassuring in several ways. Evidence for a dozen or so systems containing two or more very low mass companions have been found so far, implying a planetary-like configuration of several smaller bodies orbiting a central star, which is totally unlike the hierarchical configuration of multiple star systems. The discovery of the first (and so far only) transiting planet, around the star HD 209458 (Charbonneau et al., 2000; Henry et al., 2000), has provided the best evidence yet that many if not most of these objects are indeed gas giant planets: the planet's mass is $\approx 0.7 M_J$, with a radius and a density (Mazeh et al., 2000) roughly equal to that expected for a hot Jupiter. In addition, sodium has been detected in the atmosphere of HD 209458's planet (Charbonneau et al., 2002), exactly as predicted for a hot Jupiter (Seager and Sasselov, 2000). These ongoing discoveries have conclusively shown us that the Solar System is not the

only outcome of the planet formation process. The bulk of these new objects are likely to be gas giant planets, and their formation process and characteristics should be explainable by any general theory of planetary system formation.

3. Gas Giant Planet Formation

There are two logical extremes for forming gas giant planets, namely from the "bottom up" (core accretion), or from the "top down" (disk instability). Historically speaking, by far most of the efforts to understand giant planet formation have been performed in the context of the core accretion mechanism. As a result, the strengths and weaknesses of the core accretion mechanism are much better known than those of the competing disk instability mechanism, the latter of which has only recently been revived and subjected to serious theoretical investigation.

3.1. Core Accretion

The terrestrial planets are nearly universally believed to have formed in the inner solar nebula through the collisional accumulation of successively larger, solid bodies – sub-micron-sized dust grains, kilometer-sized planetesimals, lunar-sized planetary embryos, and finally Earth-size planets (Wetherill, 1990). The core accretion mechanism envisions the same basic process as having occurred in the outer solar nebula as well, leading to the formation of $\sim 10\ M_\oplus$ solid cores, on roughly circular orbits initially, which then accrete massive gaseous envelopes from the disk gas (Mizuno, 1980). The $\sim 10\ M_\oplus$ solid cores are expected to form through runaway accretion (e.g., Lissauer, 1987), where the largest bodies grow the fastest because their self-gravity increases their collisional cross-sections. An atmosphere forms at an early phase, by accretion of solar nebula gas, and as the protoplanet continues to grow by accreting both gas and planetesimals, this atmosphere eventually can no longer be supported in hydrostatic equilibrium and contracts. This contraction leads to a short period of atmospheric collapse, during which the protoplanet quickly gains the bulk of its final mass (Pollack et al., 1996). At that point, the accretion of solar nebula gas is assumed to be terminated.

The time scale for core accretion to proceed depends strongly on the initial surface density of solids. The surface density in the giant planet region is often assumed to be 5 to 10 times that of the "minimum mass solar nebula" in models of core accretion (Lissauer, 1987). Models which calculate both the accretion of gas as well as of planetesimals (Pollack et al., 1996) show that with a surface density of solids $\sigma_s = 10$ g cm^{-2} at 5.2 AU, core accretion requires $\approx 8 \times 10^6$ years to form Jupiter. When σ_s is decreased by just 25%, the time required increases by a factor of 6. Time scales of 8 million years or more exceed current estimates of disk life times for typical solar-type young stars of a few million years in regions of low mass star formation (Briceño et al., 2001), and of less than a million years

in regions of high mass star formation (Bally et al., 1998). Speeding up the core accretion process by increasing the assumed surface density at 5.2 AU does not appear to work for Jupiter: when σ_s is increased to 15 g cm^{-2}, the formation time drops to $\sim 2 \times 10^6$ years (Pollack et al., 1996), but produces a central core with a mass exceeding the possible range for Jupiter (Guillot et al., 1997), even given the great uncertainties in models of the Jovian interior. Models of the structure of HD 209458's hot Jupiter seem to require that this extrasolar planet have no core at all, in order to match its observed radius (Hubbard et al., 2002; Guillot, this volume).

Things get even worse out at Saturn's orbital radius, as core accretion proceeds more slowly as the orbital radius increases and orbital periods increase. All of these calculations (Pollack et al., 1996) already assume optimum conditions for the growth of cores: infinite reservoirs of accretable solids and gas, maximum possible gravitational cross-sections for collisions, no collisional fragmentation, and the absence of competition from other nearby, runaway protoplanets. Some recent core accretion models have shown that the time scale for envelope growth depends strongly on the assumed core mass and somewhat on the assumed dust grain opacity (Ikoma et al., 2000). Ikoma et al. (2000) claim that a nebula lifetime of more than 100 million years is needed to form Jupiter and Saturn, or else migration of protoplanets may have to be considered. The recent paper by Kokubo and Ida (2002) assumes that disks last for 100 million years, sufficiently long for several Jovian planets to form from massive disks by core accretion. However, the core accretion models by Kornet et al. (2002) have found that the two giant planets in the 47 UMa system might have formed in about 3 million years, assuming that the protoplanetary disk had a mass of 0.164 $M_\odot$, a local enhancement of the dust-to-gas ratio between 1 AU and 4 AU, no competing protoplanets, no loss of planetesimals by gas drag, and no migration of the protoplanet. While work continues on resolving the time scale problem, core accretion would seem to be competitive only in relatively long-lived protoplanetary disks (Lissauer and Lin, 2000).

Core accretion models of the *in situ* formation of hot Jupiters have also been attempted. The critical core mass needed for gaseous envelope collapse onto the core could be as low as 2-3 $M_\oplus$ at 0.1 AU, according to Ikoma et al. (2001). However, other calculations found that core accretion could only proceed very close to the star if solids were transported inward at a high rate in order to feed the growing core (Bodenheimer et al., 2000). The latter models suffer from the same time scale problem that occurs at more traditional distances.

The results of Bodenheimer et al. (2000) suggest that hot Jupiters formed farther out than their current distances, and then experienced inward orbital migration to their present parking orbits (Lin et al., 1996; Bodenheimer et al., 2000). The necessary orbital migration is likely to have resulted from gravitational interactions between the planet and the gas disk, though planet-planet gravitational interactions have also been studied (Ford, this volume). However, prior to the planet becoming large enough to open a gap in the disk, the time scale for inward migration (Type

I migration) is only about 10^4 years for a 10 $M_\oplus$ body. Hence it may be difficult for a growing core to survive long enough to accrete a gaseous envelope and then open a disk gap before the core is lost by inward migration onto the star (Ward, 1997; Papaloizou and Larwood, 2000; Miyoshi et al., 1999). Once a protoplanet grows large enough to remove the gas from its immediate vicinity by opening a disk gap, the protoplanet thereafter migrates (inward, or outward, depending on location in the disk) along with the disk (Type II migration). Opening a disk gap may slow but not stop subsequent growth by accretion of gas (Bryden et al., 1999). The observational fact that extrasolar planets found so far have a wide range of orbital distances from their stars may imply that their Type II orbital migration times were comparable to the life times of their inner disks ($\sim 10^6$ years), a life time that is roughly consistent with estimated disk viscosities based on theoretical estimates of likely sources of turbulent viscosity.

A number of arguments for and against core accretion may be noted. Forming Jupiter by core accretion is consistent with the absence of planets in the asteroid belt (Wetherill, 1996). Core accretion also is consistent with large core masses and non-solar bulk compositions. However, estimated core masses for the gas giant planets have dropped dramatically since 1995, and are now thought to be in the range from 0 to 10 $M_\oplus$ for Jupiter, and from 6 to 17 $M_\oplus$ for Saturn (Guillot et al., 1997; Guillot, 1999a). Obviously a core mass which is too small to initiate gas accretion would seem to rule out core accretion, unless the core somehow dissolved after the planet formed. HD 209458's planet appears to be hydrogen-rich (Hubbard et al., 2002) and need not even have a core. The non-solar atmospheric compositions of Jupiter and Saturn (Guillot, 1999b) are likely to be at least in part the result of several billion years of cometary impacts (e.g., Comet Shoemaker-Levy 9's spectacular demise in Jupiter's atmosphere). In the core accretion scenario, it is unclear what process could have limited Saturn's mass to its present value, roughly 1/3 that of Jupiter, given its apparently larger core mass. Considering these concerns about core accretion, it seems worthwhile to examine the prospects of the other possibility for gas giant planet formation.

3.2. Disk Instability

The only known alternative to core accretion is disk instability, where gas giant protoplanets form rapidly through a gravitational instability of the gaseous portion of the disk (Cameron, 1978; Boss, 1997). The disk instability mechanism had been neglected for years largely because it could not easily account for the initial estimates of the core masses of Jupiter and Saturn (in the range of 20 to 30 Earth masses), or for their non-solar bulk abundances. However, it seems likely that a significant solid core could form in a giant gaseous protoplanet by the process of sedimentation of dust grains to the center of the protoplanet, prior to contraction of the protoplanet to planetary densities and temperatures high enough to dissolve or melt the solids (Boss, 1997, 1998a). This is the same process that is thought to

lead to the formation of solid planetesimals in the core accretion mechanism, as dust grains sediment downward toward the midplane of the disk, only now in the spherical geometry of a gaseous protoplanet. A disk instability leading to gaseous protoplanet formation and the process of core formation by sedimentation of dust grains would occur essentially simultaneously, within $\sim 10^3$ years. Contraction of the protoplanet to planetary densities requires another $\sim 10^6$ years or so. For a Jupiter-mass protoplanet with a solar abundance of metals, sedimentation of all of the metals could lead to the formation of at most a $\sim 6 M_\oplus$ core, a core mass more or less in the middle of the currently estimated range (Guillot, 1999a). The impacts of the fragments of Comet Shoemaker-Levy 9 with Jupiter, combined with recent gas giant interior models (Guillot, 1999b), suggest that the present atmospheres of the gas giant planets reflect their accretion history more than their primordial compositions. A Jupiter formed by disk instability may then have experienced much the same accretion history as a Jupiter formed by core accretion, leading to similar envelope enrichments.

The standard core accretion model (Pollack et al., 1996) seems to requires a surface density at 5.2 AU which implies at least a marginally gravitationally unstable nebula, because midplane temperatures in the solar nebula drop quickly beyond the asteroid belt (Boss, 1998c) and are constrained to values below about 50 K in the outer nebula by the presence of low temperature molecular species seen in comets. Detailed three dimensional hydrodynamical models have shown that such a marginally unstable disk will become strongly non-axisymmetric and form trailing spiral arms within just a few rotation periods (Boss, 1998a). When followed with a sufficiently high spatial resolution calculation, these spiral arms break-up into high-density clumps containing enough mass to be self-gravitating and tidally stable (Boss, 2000). Hydrodynamical models with a full thermodynamical treatment, including three dimensional radiative transfer in the diffusion approximation (Boss, 2001), have shown that a disk instability proceeds in a similar manner as in the previous calculations (Boss, 2000), which employed thermodynamical assumptions which were more favorable for the growth of self-gravitating clumps. This similarity results because the time scale for cooling from the disk surface is comparable to the dynamical (orbital) time scale, so that clump formation is slowed, but not prevented, by compressional heating. The energy produced by compressional heating at the disk midplane is transported to the disk surface by convective cells, with convective velocities at 10 AU being large enough to transport heat to the disk surface on the orbital time scale (Boss, 2002c).

Disk instability can produce self-gravitating protoplanets with cores in $\sim 10^3$ years, so there is no problem with forming gas giant planets in even the shortest-lived protoplanetary disks. Even if core accretion can form gas giant planets in about 3 million years (Kornet et al., 2002), disk instability evidently will outrace core accretion, if it can occur in the first place. Disk instability is enhanced in increasingly massive disks, and so it should be able to form planets at least as massive as Jupiter, given that Jupiter-mass clumps form even in disks with masses

of $\sim 0.1 M_{\odot}$. Disk instability sidesteps any problem with Type I orbital migration, and with gap-limited mass accretion, because the clumps form directly from the gas without requiring the prior existence of a solid core subject to Type I drift that can disappear before opening a gap. Once they are formed, the clumps quickly open a disk gap, preventing Type I motion with respect to the gas, but only after most of the protoplanet's mass has already been captured. Thereafter the protoplanet migrates with the disk; in the case of the Solar System, little orbital migration appears to be necessary, implying a short life time for the solar nebula (see below). Rapid Jupiter formation by disk instability appears to be compatible with terrestrial planet formation by collisional accumulation and may help limit the growth of bodies in the asteroid belt (Kortenkamp and Wetherill, 2000; Kortenkamp et al., 2001; Thébault and Brahic, 1999). In fact, disk instability could even help to speed the growth of the Earth and other terrestrial planets (Kortenkamp et al., 2001). As in the core accretion mechanism, the ongoing accretion of comets is needed to explain the non-solar compositions of the envelopes of the gas giant planets.

Disk instability is not without its own problems, however, though potential solutions may exist in some cases. It is unclear how massive a core formed by sedimentation of dust grains would be, particularly if the interior temperature should become too high for water ice to remain solid. Whether or not a marginally gravitationally-unstable disk will evolve in such a way as to produce self-gravitating clumps is also unclear, as gravitational torques may simply redistribute mass and angular momentum instead. The pioneering work on this question had insufficient numerical resolution (25,000 particles, versus over 1 million grid cells in current disk instability models, e.g., Boss, 2003) to allow self-gravitating clumps to form (Laughlin and Bodenheimer, 1994). As a result, disk instabilities may require some sort of trigger to produce clumps, such as the accumulation of gas in a magnetically-dead zone of the disk, episodic accretion of infalling gas onto the disk (Boss, 1997), or perhaps a close encounter with another star. The conditions under which newly-formed clumps survive to become gas giant planets also remain to be understood, though calculations by Mayer et al. (2002) suggest that the clumps can survive as they orbit in the disk. Disk instability would also have trouble forming sub-Jupiter-mass planets, unless one invokes tidal stripping of the protoplanet's envelope during a phase of rapid inward orbital migration, or photoevaporation of the protoplanet's envelope (see below).

4. Ice Giant Planet Formation

The same two mechanisms that have been advanced for explaining the formation of the gas giant planets are also possibilities for explaining the formation of the ice giant planets.

4.1. CORE ACCRETION

Because of their relatively modest gaseous envelopes, the ice giant planets are somewhat less constrained than the gas giant planets by the observed life times of protoplanetary disks; e.g., it has been suggested that they formed as recently as 3.9 billion years ago (Levison et al., 2001). Nevertheless, there remains a severe time scale problem for forming ice giants by core accretion. Time scales for collisions to occur increase with increasing orbital period. As a result, collisional accumulation is even slower in the outer solar nebula than in the gas giant planet region. Furthermore, lower surface densities of solids at greater heliocentric distances slow collisional growth even more. Perhaps most importantly, because the escape velocity from the Solar System at 20 AU to 30 AU is ~ 8 km s^{-1}, comparable to orbital velocities and to the relative velocities between growing planetary embryos, the effect of mutual encounters of embryos is to excite orbital eccentricities so much that the embryos soon cross the orbits of Saturn or Jupiter. As a direct result, embryos can be ejected on hyperbolic orbits, lost by impact with the gas giant planets, or perturbed onto comet-like orbits (Lissauer et al., 1995). In fact, it has been asserted that ice giant planets cannot form in the standard core accretion model (Levison and Stewart, 2001). Theorists have artificially increased collisional cross sections of the growing planetary embryos by several orders of magnitude in order to try to gain some understanding of what might have happened in the outer Solar System (Levison et al., 1998). Even in this unphysical case, the bodies that do grow have eccentric orbits that inhibit further growth and do not resemble the more nearly circular orbits of the outer planets.

One approach to solving the problem of the ice giant planets has been to invoke a hypothetical drag force that would damp orbital eccentricities in the outer Solar System (Levison and Stewart, 2001), due perhaps to extended planetary envelopes, or to interactions with the remaining gas and planetesimals in the nebula. One could also imagine the runaway accretion of a single embryo all the way to Uranus-size (Bryden et al., 2000), instead of assuming the oligarchic growth of multiple embryos (Kokubo and Ida, 1998). Perhaps the leading suggestion for forming the ice giant planets by core accretion is the idea (Thommes et al., 1999) that the ice giants were formed between Jupiter and Saturn (i.e., that core accretion was able to form *four* cores between 5 and 10 AU), and then were gravitationally scattered outward to their current orbits, following the rapid growth of the gaseous envelopes (and masses) of one or two of the other cores that were destined to become Jupiter or Saturn. This scenario might then be only a little more difficult than that of forming just the gas giant planets by core accretion.

4.2. DISK INSTABILITY

Recently it has been proposed that disk instability might be capable of forming the ice giant planets, provided that the Solar System formed in a region of high mass star formation, similar to the Orion nebula cluster (Boss et al., 2002). The

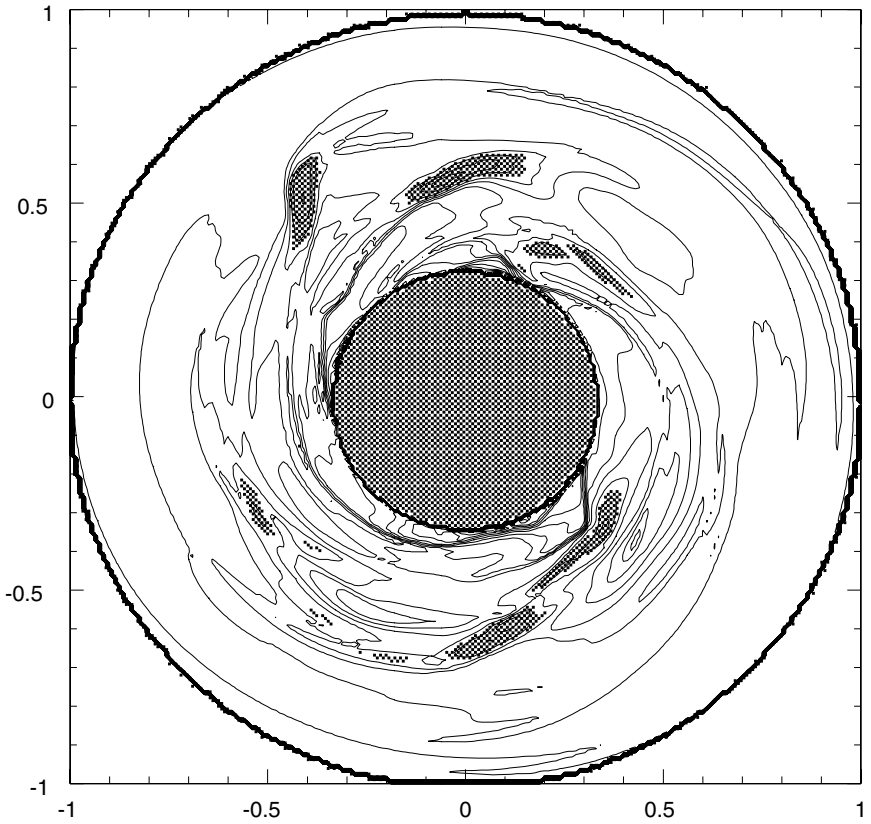

Figure 2. Equatorial density contours for a three-dimensional radiative hydrodynamics calculation of a disk instability possibly leading to the formation of the giant planets. The region shown is 30 AU in radius, with an inner region of radius 10 AU excised. Four well-defined clumps (cross-hatched regions) with masses about twice that of Jupiter have formed after 340 years of evolution.

conventional view of Solar System formation is that the presolar cloud collapsed in a region of low-mass star formation, similar to Taurus-Auriga. In such a quiescent setting, the background UV flux is likely to be low and limited largely to the flux from the protosun, once it forms.

However, in regions similar to the Orion Trapezium environment, extreme ultraviolet radiation (EUV) from the massive stars would photoevaporate the disk gas outside a radius of about 10 AU in about $\sim 10^5$ years (Bally et al., 1998). Recent observations of young stars in the Eta Carina nebula have revealed the presence of protoplanetary disks being exposed to a level of EUV radiation roughly 100 times higher than that in Orion (Smith et al., 2002), leading to proportionately shorter outer disk life times. Once the disk gas is removed, the outermost protoplanets would then be exposed to EUV radiation, and provided that they do not contract to planetary densities in a time much less than $\sim 10^4$ to $\sim 10^6$ years (DeCampli and Cameron, 1979), their gaseous envelopes will also be photoevaporated by the

incoming EUV flux, leaving behind the cores previously formed by sedimentation and coagulation of their dust grains. Given the formation by disk instability of four gas giant protoplanets with masses of order 1 to 2 M_J, close to the orbits of Jupiter, Saturn, Uranus, and Neptune (Boss, 2003; Figure 2), EUV photoevaporation of the gaseous envelopes of the outermost three protoplanets would leave behind cores with partial gaseous envelopes, producing planets similar in composition to Saturn, Uranus, and Neptune (Boss et al., 2002). The innermost gas giant, destined to become Jupiter, is inside the critical radius where the disk gas cannot be removed by EUV radiation because of the protosun's gravitational attraction, and so does not lose any envelope gas.

Much remains to be studied in this unconventional mechanism for ice giant planet formation, but this mechanism does have the major advantage of being applicable in general, as most stars are believed to form in regions similar to Orion. If correct, this would mean that the Solar System need not have formed under somewhat special circumstances, and so need not be the exception, but could instead be the rule among planetary systems. While roughly 5% of nearby sun-like stars are circled by planetary systems quite unlike our own (Figure 1), the remaining sun-like stars might still shelter planetary systems similar to the Solar System.

5. Conclusions

Thanks to the pathbreaking work of Michel Mayor, Didier Queloz, and their colleagues, we are now embarked on a grand journey to explore the possibilities for other habitable planets in our corner of the Milky Way galaxy. As we proceed on this journey, we are likely to learn enough about the characteristics of planetary systems to be able to differentiate between the two competing theories of the origin of gas and ice giant planets. We conclude by highlighting a few key observational tests that should aid in this process.

5.1. Observational Tests for Gas Giant Planet Formation

If core accretion requires an exceptionally long-lived disk in order to have sufficient time for the accretion of a massive gaseous envelope, then gas giant planets might be no more frequent than long-lived disks. If disk instability is able to occur, on the other hand, it should lead to the formation of gas giants in even the shortest-lived protoplanetary disk. Hence, one basic test is to determine the frequency of gas giant planets: are they rare or commonplace? The results to date (Figure 1) would suggest that gas giants are abundant, but only continued efforts by the ground-based radial velocity surveys will reveal the true gas giant planetary census, particularly for the longer period planets more closely resembling Jupiter.

Because core accretion is slow with respect to disk instability, if core accretion dominates, young stars should not show evidence of gas giant companions until

they reach ages of several million years or more. If disk instability dominates, however, even the youngest stellar objects may have gas giant planets. Dating the epoch of gas giant planet formation is thus another means to differentiate between the two contenders (Boss, 1998b). Several techniques could be employed. NASA's Space Interferometry Mission (SIM) will search for the astrometric wobbles of young stars caused by gas giant companions, beginning perhaps as early as 2009. Nearby low-mass star-forming regions such as Taurus and Ophiuchus will be the primary hunting grounds. The Atacama Large Millimeter Array (ALMA) will be able to form mm-wave images of protoplanetary disks with ~ 1 AU resolution and search for disk gaps created by massive protoplanets. The radial velocity technique is hampered in dating the epoch of gas giant planet formation because of the rapid rotation rates (and hence broad spectral lines), chromospheric activity, and variability of young stars.

Disk instability appears to be relatively insensitive to the opacity of the disk, which is dominated by dust grains, and thus to the metallicity of the host star and disk (Boss, 2002a). It is unclear if core accretion is helped or hindered by higher metallicity, but the overall effect should be to raise the surface density of solids and thus to speed the growth of cores. Another test is thus to see if a high primordial metallicity is necessary for gas giant planet formation. This will require either spectroscopic or astrometric ground-based searches for gas giants around metal-poor stars, the latter using the Keck Interferometer (KI) or the Very Large Telescope Interferometer (VLTI).

Finally, the mass of Jupiter's core remains as an important clue to its origin. If Jupiter's core is much more massive than $\sim 6 M_{\oplus}$, then it probably could not have formed by disk instability, unless it had lost part of its gaseous envelope by EUV stripping and then migrated inward to 5.2 AU. A large Jupiter core would seem to support formation by core accretion. A Jupiter polar orbiter mission to probe the planet's gravitational field might be needed to better constrain the Jovian interior.

5.2. Observational Tests for Ice Giant Planet Formation

If core accretion can occur, but seldom occurs in a disk long-lived enough for a gas giant planet to form, then the typical result may be a system of failed cores, i.e., a system of ice giant planets, unaccompanied by gas giants. If disk instability dominates, inner gas giants should be the rule, accompanied by outer ice giant planets in systems which formed in Orion-like regions and experienced EUV envelope stripping. In Taurus-like regions, disk instability should produce only gas giants, unaccompanied by outer ice giants. Ground-based radial velocity surveys for "hot Neptunes" will shed light on the frequency of short period ice giants. The Corot, Kepler, and Eddington space missions will use photometry to detect "hot" and "warm Neptunes" by the transit method. "Cold Neptunes" could be detected astrometrically by the KI, VLTI, or SIM.

The question of planetary system architectures is another potential factor: where are the ice giants orbiting with respect to the gas giants? If extrasolar planetary systems are typically as well-ordered as the Solar System (inner terrestrial planets, intermediate gas giants, outer ice giants), such an architecture would imply formation *in situ*, or else an orderly, disk-driven inward orbital migration from more distant regions. However, if extrasolar planetary systems are more often disordered, this would imply that gravitational interactions between the protoplanets led to a phase of chaotic evolution where information about the primordial planetary orbits has been lost. In the latter case, it may be hard to place constraints on the formation mechanisms involved.

5.3. Detection of Extrasolar Earths

The ultimate goal of the search for extrasolar planetary systems is to find terrestrial-like planets orbiting in the habitable zones of their stars, planets that might well be analogous to the Earth, though perhaps at a much different phase of planetary evolution. While this goal still seems distant, it is much closer than it was in 1995, before Mayor and Queloz (1995) made their epochal discovery. Both NASA and ESA are planning to fly space missions that will first estimate the frequency of Earth-like planets by transit photometry (Kepler and Eddington, respectively), and then detect and characterize Earth-like planets by direct imaging, either with an optical coronagraph or with an infrared interferometer (Terrestrial Planet Finder and Darwin, respectively). The detection of extrasolar Earths will tell us much about the outcome of the collisional accumulation process in different stellar environments, and will begin to answer the question of the existence of life elsewhere in the universe. History will record that this journey began in 1995 with the work of two Swiss astronomers.

Acknowledgements

I thank the meeting organizers for their support of my attendance at this grand celebration of Michel Mayor's life and achievements. This work has also been partially supported by the U.S. National Aeronautics and Space Administration through grant NAG 5-10201.

References

Bally, J., L. Testi, A. Sargent, and J. Carlstrom: 1998, 'Disk Mass Limits and Lifetimes of Externally Irradiated Young Stellar Objects Embedded in the Orion Nebula'. *Astron. J.* **116**, 854–859.

Bodenheimer, P., O. Hubickyj, and J. J. Lissauer: 2000, 'Models of the *in Situ* Formation of Detected Extrasolar Giant Planets'. *Icarus* **143**, 2–14.

Boss, A. P.: 1997, 'Giant Planet Formation by Gravitational Instability'. *Science* **276**, 1836–1839.

Boss, A. P.: 1998a, 'Evolution of the Solar Nebula. IV. Giant Gaseous Protoplanet Formation'. *Astrophys. J.* **503**, 923–937.
Boss, A. P.: 1998b, 'Astrometric signatures of giant–planet formation'. *Nature* **393**, 141–143.
Boss, A. P.: 1998c, 'Temperatures in protoplanetary disks'. *Ann. Rev. Earth Planet. Sci.* **26**, 53–80.
Boss, A. P.: 2000, 'Possible Rapid Gas Giant Planet Formation in the Solar Nebula and Other Protoplanetary Disks'. *Astrophys. J.* **536**, L101–L104.
Boss, A. P.: 2001, 'Gas Giant Planet Formation: Disk Instability Models with Thermodynamics and Radiative Transfer'. *Astrophys. J.* **563**, 367–373.
Boss, A. P.: 2002a, 'Stellar Metallicity and the Formation of Extrasolar Gas Giant Planets'. *Astrophys. J.* **567**, L149–L153.
Boss, A. P.: 2002b, 'Formation of gas and ice giant planets'. *Earth Planet. Sci. Lett.* **202**, 513–523.
Boss, A. P.: 2002c, 'Evolution of the Solar Nebula. V. Disk Instabilities with Varied Thermodynamics'. *Astrophys. J.* **576**, 462–472.
Boss, A. P.: 2003. 'Rapid Formation of Ice Giant Planets by Disk Instability'. *Astrophys. J.*, submitted.
Boss, A. P., G. W. Wetherill, and N. Haghighipour: 2002, 'Rapid Formation of Ice Giant Planets'. *Icarus* **156**, 291–295.
Briceño, C., A. K. Vivas, N. Calvet, L. Hartmann, R. Pacheco, D. Herrera, L. Romero, P. Berlind, G. Sánchez, J. A. Snyder, and P. Andrews: 2001, 'The CIDA–QUEST Large–Scale Survey of Orion OB1: Evidence for Rapid Disk Dissipation in a Dispersed Stellar Population'. *Science* **291**, 93–96.
Bryden, G., X. Chen, D. N. C. Lin, R. P. Nelson, and J. C. B. Papaloizou: 1999, 'Tidally Induced Gap Formation in Protostellar Disks: Gap Clearing and Suppression of Protoplanetary Growth'. *Astrophys. J.* **514**, 344–367.
Bryden, G., D. N. C. Lin, and S. Ida: 2000, 'Protoplanetary Formation. I. Neptune'. *Astrophys. J.* **544**, 481–495.
Cameron, A. G. W.: 1978, 'Physics of the primitive solar accretion disk'. *Moon Planets* **18**, 5–40.
Charbonneau, D., T. M. Brown, D. W. Latham, and M. Mayor: 2000, 'Detection of Planetary Transits Across a Sun-Like Star'. *Astrophys. J.* **529**, L45–L48.
Charbonneau, D., T. M. Brown, R. W. Noyes, and R. L. Gilliland: 2002, 'Detection of an Extrasolar Planet Atmosphere'. *Astrophys. J.*, **568**, 377–384.
DeCampli, W. M., and A. G. W. Cameron: 1979, 'Structure and Evolution of Isolated Giant Gaseous Protoplanets'. *Icarus* **98**, 367–391.
Guillot, T.: 1999a, 'A comparison of the interiors of Jupiter and Saturn'. *Planet. Space Science* **47**, 1183–1200.
Guillot, T.: 1999b, 'Interiors of Giant Planets Inside and Outside the Solar System'. *Science* **286**, 72–77.
Guillot, T., D. Gautier, and W. B. Hubbard: 1997, 'New Constraints on the Composition of Jupiter from Galileo Measurements and Interior Models'. *Icarus* **130**, 534–539.
Henry, G. W., G. W. Marcy, R. P. Butler, and S. S. Vogt: 2000, 'A Transiting "51 Peg-Like" Planet'. *Astrophys. J.* **529**, L41–L44.
Hubbard, W. B., A. Burrows, and J. I. Lunine: 2002, 'Theory of Giant Planets'. *Ann. Rev. Astron. Astrophys.* **40**, 103–136.
Ikoma, M., H. Emori, and K. Nakazawa: 2001, 'Formation of Giant Planets in Dense Nebulae: Critical Core Mass Revisited'. *Astrophys. J.* **553**, 999–1005.
Ikoma, M., K. Nakazawa, and H. Emori: 2000, 'Formation of Giant Planets: Dependences on Core Accretion Rate and Grain Opacity'. *Astrophys. J.* **537**, 1013–1025.
Kokubo, E., and S. Ida: 1998, 'Oligarchic Growth of Protoplanets'. *Icarus* **131**, 171–178.
Kokubo, E., and S. Ida: 2002, 'Formation of Protoplanet Systems and Diversity of Planetary Systems'. *Astrophys. J.* **581**, 666–680.

Kornet, K., Bodenheimer, P., and Różyczka, M.: 2002, 'Models of the formation of the planets in the 47 UMa system'. *Astron. Astrophys.* **396**, 977–986.

Kortenkamp, S. J., and G. W. Wetherill: 2000, 'Terrestrial Planet and Asteroid Formation in the Presence of Giant Planets. I. Relative Velocities of Planetesimals Subject to Jupiter and Saturn Perturbations'. *Icarus* **143**, 60–73.

Kortenkamp, S. J., G. W. Wetherill, and S. Inaba: 2001, 'Runaway Growth of Planetary Embryos Facilitated by Massive Bodies in a Protoplanetary Disk'. *Science* **293**, 1127–1129.

Laughlin, G., and P. Bodenheimer: 1994, 'Nonaxisymmetric Evolution in Protostellar Disks'. *Astrophys. J.* **436**, 335–354.

Levison, H. F., L. Dones, C. R. Chapman, S. A. Stern, M. J. Duncan and K. Zahnle: 2001, 'Could the Lunar "Late Heavy Bombardment" Have Been Triggered by the Formation of Uranus and Neptune?'. *Icarus* **151**, 286–306.

Levison, H. F., J. J. Lissauer, and M. J. Duncan: 1998, 'Modeling the Diversity of Outer Planetary Systems'. *Astrophys. J.* **116**, 1998–2014.

Levison, H. F., and G. R. Stewart: 2001, 'Remarks on Modeling the Formation of Uranus and Neptune'. *Icarus* **153**, 224–228.

Lin, D. N. C., P. Bodenheimer, and D. C. Richardson: 1996, 'Orbital migration of the planetary companion of 51 Pegasi to its present location'. *Nature* **380**, 606–607.

Lissauer, J. J.: 1987, 'Timescales for planetary accretion and the structure of the protoplanetary disk'. *Icarus* **69**, 249–265.

Lissauer, J. J., J. B. Pollack, G. W. Wetherill, and D. J. Stevenson: 1995, 'Formation of the Neptune System'. In: *Neptune and Triton*, D. P. Cruikshank (Ed.), University of Arizona Press, Tucson, AZ, pp. 37–107.

Lissauer, J. J., and D. N. C. Lin: 2000, 'Diversity of Planetary Systems: Formation Scenarios and Unsolved Problems'. In: *From extrasolar planets to cosmology: the VLT opening symposium*, A. Renzini (Ed.), Springer, Berlin, pp. 377–390.

Marcy, G. W. and R. P. Butler: 1998, 'Detection of Extrasolar Giant Planets'. *Ann. Rev. Astron. Astrophys.* **36**, 57–97.

Mayer, L., T. Quinn, J. Wadsley, and J. Stadel.: 2002, 'Formation of Giant Planets by Fragmentation of Protoplanetary Disks'. *Science* **298**, 1756–1759.

Mayor, M. and D. Queloz: 1995, 'A Jupiter-mass companion to a solar-type star'. *Nature* **378**, 355–359.

Mazeh, T., D. Naef, G. Torres, D. W. Latham, M. Mayor, J. L. Beuzit, T. M. Brown, L. Buchhave, M. Burnet, B. W. Carney, D. Charbonneau, J. B. Laird, F. Pepe, C. Perrier, and D. Queloz: 2000, 'The Spectroscopic Orbit of the Planetary Companion Transiting HD 209458'. *Astrophys. J.* **523**, L55–L58.

Miyoshi, K., T. Takeuchi, H. Tanaka, and S. Ida: 1999, 'Gravitational Interaction Between a Protoplanet and a Protoplanetary Disk. I. Local Three–Dimensional Simulations'. *Astrophys. J.* **516**, 451–464.

Mizuno, H: 1980, 'Formation of the giant planets'. *Prog. Theor. Phys.* **64**, 544–557.

Papaloizou, J. C. B., and J. D. Larwood: 2000, 'On the orbital evolution and growth of protoplanets embedded in a gaseous disc'. *Mon. Not. Roy. Astron. Soc.* **315**, 823–833.

Pollack, J. B., O. Hubickyj, P. Bodenheimer, J. J. Lissauer, M. Podolak, and Y. Greenzweig: 1996, 'Formation of the Giant Planets by Concurrent Accretion of Solids and Gas'. *Icarus* **124**, 62–85.

Seager, S., and D. D. Sasselov: 2000, 'Theoretical Transmission Spectra During Extrasolar Giant Planet Transits'. *Astrophys. J.* **537**, 916–921.

Smith, N., J. Bally, J. Thiel, and J. A. Morse: 2002, 'Discovery of Proplyd-Like Objects in the Extreme Environment of the Carina Nebula'. *Bull. Amer. Astron. Soc.* **34**, 1260.

Thébault, P., and A. Brahic: 1999, 'Dynamical influence of a proto–Jupiter on a disk of colliding planetesimals'. *Planet. Space Science* **47**, 233–243.

Thommes, E. W., M. J. Duncan, and H. F. Levison: 1999, 'The formation of Uranus and Neptune in the Jupiter–Saturn region of the Solar System'. *Nature* **402**, 635–637.
Ward, W. R.: 1997, 'Protoplanet Migration by Nebula Tides'. *Icarus* **126**, 261–281.
Wetherill, G. W.: 1990, 'Formation of the Earth'. *Ann. Rev. Earth Planet. Science* **18**, 205–256.
Wetherill, G. W.: 1996, 'The Formation and Habitability of Extra-Solar Planets'. *Icarus* **119**, 219–238.