



THE ENTHRALLING TALE OF THE FORMATION
AND EVOLUTION OF COMPACT PLANETARY
SYSTEMS

Thesis submitted for the degree of Doctor of Philosophy
at the University of Leicester

by
Thomas Oliver Hands MPhys
Theoretical Astrophysics group
Department of Physics & Astronomy
University of Leicester

June 2016

Abstract

Of the myriad of insights into exoplanetary systems provided by the Kepler mission, one of the most intriguing new discoveries is that of a class of compact planetary systems which include Kepler-11, Kepler-32 and Kepler-90. In such systems, ensembles of several planets are found in very closely packed orbits (often within a few percent of an astronomical unit of one another). These systems present a challenge for traditional formation and migration scenarios, since these planets presumably formed at larger orbital radii before migrating inwards. In particular, it is difficult to understand how some planets in such systems could have migrated across strong mean-motion resonances without becoming trapped, and remaining relatively well-spaced. It is also difficult to explain how such systems remain dynamically cold, as resonant interactions tend to excite orbital eccentricity and lead to close encounters. I present a dynamical study of the formation of these systems, using an N -body method which incorporates a parametrized model of planet migration in a turbulent protoplanetary disc. The study explores a wide parameter space, and finds that under suitable conditions it is possible to form compact, close-packed planetary systems via traditional disc-driven migration, albeit with an over-abundance of mean-motion resonances. I then extend the study to include Jupiter-mass planets exterior to the compact systems, and find that the dynamical effect of these companions can significantly modify the resonant structure of the compact planets. Finally, I extend this work to two-dimensional hydrodynamical simulations in an attempt to model type I migration self-consistently. In particular, I find that clearing of the disc by photoevaporation can halt migration of compact systems, and also discover that planet-disc interactions can - under the right conditions - break mean-motion resonances.

The following chapters from this thesis have been published in the journal *Monthly Notices of the Royal Astronomical Society*:

Chapter 3: *Understanding the assembly of Kepler's compact planetary systems*, **Hands T. O.**, Alexander R. D., Dehnen W., 2014, MNRAS, **445**, 749

Chapter 4: *There might be giants: unseen Jupiter-mass planets as sculptors of tightly packed planetary systems*, **Hands T. O.**, Alexander R. D., 2016, MNRAS, **456**, 4121

Chapter 5: **Hands T. O.**, Alexander R. D., 2016, MNRAS, *in prep*

List of Tables

3.1	Initial conditions for simulations	89
3.2	List of simulation outcomes	90
4.1	K-S test results	112
4.2	List of simulation outcomes	113
4.3	Probabilities of transit for simulated systems	117
5.1	A list of initial conditions used to test PLUTO	130
5.2	Initial conditions for planets in all 4 hydrodynamical models	132

List of Figures

1.1	Schematic diagram showing typical spectra for various classes of young stellar objects.	4
1.2	The spreading ring test for a viscous disc.	11
1.3	The magnetorotational instability	16
1.4	Properties of known exoplanets	21
1.5	How radial velocity works	22
1.6	How the transit method works	24
1.7	Schematic diagram of an elliptical orbit	34
1.8	Schematic diagram of orbital elements	36
1.9	A numerical example of resonant trapping	37
1.10	Schematic diagram showing the setup for deriving the impulse approximation	40
1.11	Locations of Lindblad and Corotation resonances	45
2.1	N-body code test - Figure eight orbit	65
2.2	N-body code test - conservation	66
2.3	N-body code test - elliptical orbit	67
2.4	Smoothed Particle Hydrodynamics	69
2.5	A numerical example of advection schemes	72
2.6	Godunov's method in one dimension	73
2.7	PLUTO scaling test	77
3.1	Tightly-packed planetary systems	82
3.2	Success rate for simulations across parameter space, 2D	102

LIST OF FIGURES

3.3	Success rate for simulations across parameter space, 1D	103
3.4	Semi-major axis distributions from simulations	104
3.5	Occurrence of resonances in simulations	105
4.1	A comparison of systems formed with and without giant planets	108
4.2	The variance of mean-motion resonances between simulations with different giant planets	119
4.3	Radial velocities for planets in my simulations	120
5.1	Demonstration of damping prescription.	126
5.2	PLUTO test run results	131
5.3	Torques on planets, 2planlo and 2planhi	133
5.4	Torques on planets, 3planlo	134
5.5	Torques on planets, 3planhi	135
5.6	Azimuthally averaged surface densities from all models	136
5.7	Surface density maps, 2planhi	137
5.8	Surface density maps, 3planhi	138
5.9	Strip showing development of partial gaps in 2planhi	139
5.10	Evolution of orbital elements, 2planlo	140
5.11	Evolution of orbital elements, 2planhi, 3planlo and 3planhi.	141
5.12	Evolution of resonant argument, 2planlo and 2planhi	142
5.13	Evolution of resonant argument, 3planlo and 3planhi	143
5.14	Migration time-scales from hydrodynamical simulations	144
5.15	Unaveraged torques, 2planhi	146
5.16	Zoomed semi-major axis evolution, 2planhi	147
A.1	A comparison between the thin disc approximation and a full analytical solution	159

*Dedicated to my late grandmother, Hilda May Darby, and my mother and
father.*

Acknowledgements

The journey from complete astrophysics novice to someone capable of writing a coherent thesis on the subject has been a long, perilous, and oftentimes stressful one, and one that has left me with a profound sense of thankfulness to a great many people. First and foremost, I must thank my supervisor Dr. Richard Alexander for his constant guidance and patience, and most of all, friendship. Richard has continuously gone above and beyond the call of duty in his effort to make my Ph.D. experience an enjoyable and fruitful one. He is remarkably relaxed for a man who appears to know almost everything about astrophysics and the English language, and I am incredibly grateful for the opportunity I had to leech as much of his knowledge as I could over the last three and a half years. Secondly, I find myself deeply indebted to Prof. Walter Dehnen, who not only convinced me to apply to the group by guiding me through a fascinating summer project, but continued to act as a supervisor and mentor for years afterwards. Walter has always been there to help me through the more mathematically/numerically challenging parts of this whole experience, and again, I can't thank him enough.

I am naturally indebted to the rest of the Theoretical Astrophysics group at the University of Leicester, particularly the staff - Sergei, Andrew, Mark and Graham - for stimulating discussions and many sorely-needed coffee breaks. Lisa Brant has been incredibly helpful in matters of organisation and travel throughout the whole experience. Additionally I owe many thanks to the students - past and present - of the Batcave. In no particular order: Christian, Lilian, Hasanuddin, Amery, Rob, Hastyar, Pete, Gillian, Alex Dunhill, Clement, Kastytis, Seung-Hoon, Suzan and Martin. Hossam Aly stands out as a man who has always had time for my personal-life and hydrodynamics-related issues. Charlie Field, Alex Chaushev, Marky Mark and Alex Eckersall have kept me sane through many of the harder times, including a weekend trip to Blackpool - easily the hardest of all the times. Stephen Fendyke's insight and knowledge of PLUTO saved me many laborious hours of debugging. Alessia Gualandris has earned a special mention for her tireless patience working with me on N -body simulations of the galactic centre, as has Chris Nixon for his identically tireless patience working with me on MHD simulations.

Of course, I wouldn't be where I am today without my wonderful parents - Martin Hands and Susan Darby - both of whom have been an inspiration and a huge source of support during my entire education. Their support, particularly through the harder times, has been absolutely invaluable, as has their unending belief in my (eventual) success. Many thanks are of course owed to the rest of the Hands/Darby clan, including but not limited to Judy Fletcher, Elaine Cartwright, and my late and oft-missed grandmother Hilda Darby. Another figure who was hugely important to my early education was Sue Rowing of South Bromsgrove High School, without whose guidance I never would've be able to read physics at university.

Finally, I wish to thank the many truly excellent friends from outside of astrophysics who put up with my eccentricities, general moaning and slightly-bizarre work-life balance throughout this whole period. Special mention here has to go to Chris Millar and James Henderson - the former for his relentless dedication to insane road trips and video games, and the latter for putting up with me and my taste in music for nearly 13 years at this point. Patricia König and Marina Mayr deserve thanks for providing a source of constant laughter and a place to sleep when I found myself homeless and often alone in Germany. Theresa Frickel for somehow managing to be monumentally supportive despite living the best part of 1000 miles away. Susi Bauer for her companionship. Katie Dexter for almost-endless cups of coffee and discussions about cellos. Camilla Roome for 13 years of owl-fearing and duck-chasing japes. Lioba (Hähnchen) Franzelin for her support. Katie Herlingshaw for her optimism and excellent accent. MariaGiovanna Garreffa for her pictionary skills. Conor Wildy for many pub trips. Emma Longstaff for sitting through two of the longest bus rides imaginable. Peter Somkuti, Emily Jane Watkinson, Jonathan Youl, Neil Humpage, Gareth John and Andrew Lobban for helping me in my many musical endeavours.

In the (likely) event that I have forgotten someone, I hope my gratitude will be evident regardless.

Contents

1	Introduction	1
1.1	Young Stellar Objects and their discs	2
1.1.1	Spectral Evidence	4
1.1.2	Types of Disc	5
1.1.3	Resolved Observations	7
1.1.4	Theoretical behaviour	8
1.1.5	Sources of viscosity	15
1.2	Observations of Exoplanets	20
1.2.1	Radial Velocity	20
1.2.2	Transits	23
1.2.3	Direct imaging & other methods	26
1.3	Formation of planetary systems	27
1.4	Orbital mechanics	29
1.4.1	The two body problem	30
1.4.2	Resonant trapping	37
1.5	Migration of planets	39
1.5.1	Type I	39
1.5.2	Type II	49
1.5.3	Type III	51
1.6	Summary	52

2	Numerical methods	53
2.1	<i>N</i> -body	53
2.1.1	Time integration	54
2.2	<i>N</i> -body code	61
2.2.1	<i>N</i> -body time-stepping	61
2.2.2	<i>N</i> -body damping	63
2.2.3	<i>N</i> -body code tests	64
2.3	Hydrodynamical	68
2.3.1	PLUTO	75
2.4	Tonight Matthew, I'm going to be...	76
3	Understanding the assembly of Kepler's compact planetary systems	79
3.1	Introduction	79
3.2	Numerical method	83
3.2.1	<i>N</i> -body integrator	83
3.2.2	Migration and eccentricity damping	83
3.2.3	Disc turbulence	85
3.2.4	Simulation set-up	86
3.2.5	Simulations	87
3.3	Results	90
3.3.1	Mean-motion resonances	92
3.3.2	Stability	94
3.4	Discussion	95
3.4.1	Simulation outcomes & preferred parameter values	95
3.4.2	Mean-motion resonances	96
3.4.3	Limitations of the model	97
3.5	Conclusions	101
4	There might be giants	106
4.1	Introduction	106

4.2	Numerical method	107
4.2.1	Runaway gas accretion	109
4.2.2	Linear mass accretion	110
4.2.3	Initial conditions	110
4.2.4	Models	111
4.2.5	Analysis	111
4.3	Results	112
4.4	Discussion	115
4.4.1	Implications	115
4.4.2	Observability	115
4.4.3	Limitations	117
4.5	Summary	118
5	Hydrodynamical simulations of compact systems	121
5.1	Introduction	121
5.2	Method	123
5.2.1	Disc model	124
5.2.2	Resolution and grid setup	125
5.2.3	Boundary conditions & wave damping	126
5.2.4	Gravitational Softening	127
5.2.5	Disc clearing	128
5.2.6	Modifications to PLUTO	130
5.2.7	Coupled code tests	130
5.2.8	Initial conditions	131
5.3	Results	132
5.3.1	Orbital elements	137
5.3.2	Resonances	138
5.4	Discussion	142
5.4.1	Halting migration	142
5.4.2	Disc structure	145

CONTENTS

5.4.3	Resonant behaviour	145
5.4.4	Extending this work	147
5.5	Summary	148
6	Conclusions	149
6.1	The story so far	149
6.2	Future work	151
6.2.1	Planet traps	151
6.2.2	Disc clearing	152
6.2.3	Alternate modelling techniques	153
6.3	Final conclusions	154
6.4	Acknowledgements	155
A	On the validity of the thin disc approximation	156
B	Stochastic forces	160

Chapter 1

Introduction

That, therefore, there are numerous orbital motions belonging to the stars that are being moved across the arch of heaven is evident to those who have even moderately busied themselves with such enquiries.

Aristotle

Metaphysics

Ever since Aristotle - with the help of Eudoxus and Calippus before him - posited that the Universe consists of 55 concentric spheres on which celestial bodies rotate, man has yearned to understand the motion and formation of the planets and stars. Aristotle, of course, famously believed that the Earth was at the centre of this system of spheres, and therefore of the entire universe. This theory was published in his circa 350 B.C. treatise *Metaphysics* (MacMahon, 1896), and variations of this geocentric cosmology continued to be the most widely-accepted view of the Universe through the time of Ptolemy right up to the 16th century. With the publication of his *De revolutionibus orbium coelestium* (*On the Revolutions of the Heavenly Spheres*) in 1543, Nicolaus Copernicus (1473-1543) became the first to popularise Heliocentrism - the idea that the Earth and other planets in our solar system orbit the sun - yet it was not until Johannes Kepler published his first two laws of planetary motion in *Astronomia nova* in 1609 that belief in the Copernican system actually took hold. In this work, Kepler introduced the notion that the orbits of the planets form ellipses with the Sun at one focus. Copernicus had previously assumed that the planets followed circular orbits, which prevented his model from making more accurate predictions than previous geocentric models. Kepler's laws of

planetary motion, in addition to observations of the phases of Venus by Galileo Galilei in 1610, finally allowed heliocentrism to become the dominant system.

Stellar and planetary astronomy has come a long way since the Renaissance, with the discovery of Uranus in 1781, Neptune in 1846 and Pluto in 1932. By 1952, Otto Struve had already suggested that high precision spectroscopy and photometry of stars other than our own might well reveal the presence of exoplanetary bodies (Struve, 1952). Forty years later came the first confirmed detection of planets external to our solar system, orbiting the pulsar PSR1257 + 12 (Wolszczan & Frail, 1992). The next big milestone came only three years later with the detection of a Jupiter-mass companion to a main sequence star - 51 Pegasi - using the exact method proposed by Struve in 1952 (Mayor & Queloz, 1995). Since then there has been a veritable explosion of exoplanet discoveries, which now number in the thousands, and this explosion has constantly provided more questions regarding planet formation theories than it has answered. If we are to begin to understand the formation and evolution of these newly-discovered planets, then it seems prudent to begin by considering the formation and evolution of the stars which they orbit.

1.1 Young Stellar Objects and their discs

Stars form in dense, cold molecular clouds in the interstellar medium (ISM), which collapse in an inside-out fashion, following a theoretical framework established starting with Shu (1977). These clouds consist largely of molecular hydrogen. Once such a cloud is sufficiently large that the gravitational force overcomes the internal energy of the gas, the Jeans instability sets in and the cloud begins to collapse in free-fall. The gravitational energy liberated in this process is radiated in the IR (infra-red) by dust grains, and may excite the molecular hydrogen. The opacity towards the centre of the cloud increases sharply, leading to the formation of a first hydrostatic core, which can collapse no further through this method since the energy cannot be radiated away. Instead, gas external to the first hydrostatic core falls onto it and shocks, the thermal energy from which continues to dissociate the hydrogen molecules and ionise the remaining atoms. These processes absorb the energy of the collapse, allowing it to continue quickly. At the point where most of the hydrogen is ionised, the core is once again in hydrostatic equilibrium and can collapse no further, yielding a protostar. Subsequent

contraction then takes place hydrostatically and thus on a much longer time-scale. The whole collapse process happens in a hierarchical fashion and thus one molecular cloud (which may have a mass of up to $10^5 M_{\odot}$) will form an entire cluster of stars.

This is not the entire story however, as the angular momentum of the initial cloud and first core is much higher than that of the final star, suggesting that a vast quantity of angular momentum must be lost in the transition from cloud to protostar. The solution to this issue involves the formation of a disc of material, from which the core and protostar can accrete (see e.g. [Shu et al., 1987](#)). The evolution of this disc causes matter to fall slowly inwards towards the star, whilst angular momentum is transported out of the system. The formation of the disc can be modelled as follows: if we consider the collapsing cloud to be a rotating sphere, then there is complete symmetry above and below the equatorial plane of this sphere. Thus equal amounts of material fall into the equatorial plane from either side, with equal and opposite momenta. A shock must remove the energy and momentum of the material that is perpendicular to the equatorial plane, leaving only the the velocity component that is in the plane, naturally leading to the formation of a thin disc of gas in this plane ([Hartmann, 1998](#)). The vast majority of the material that falls onto the core and protostar is processed through this disc. By the time it is well formed and the envelope is largely consumed, the disc is truncated at a radius of approximately 0.1 AU by the magnetic field of the star, at which point infalling material is channeled off the disc, along the magnetic field lines and onto the surface of the star. This material actually still has too much angular momentum to be accreted onto the star, and the excess angular momentum is driven away in bipolar outflows. In this way the star consumes the remaining gas in the disc, and eventually reaches the stage at which it begins to burn hydrogen, ending up as a (zero age) main sequence star with little or no remaining disc. The complete evolution of the core, protostar and disc in this phase is covered in substantially more detail in the reviews by [Lada \(2005\)](#), [McKee & Ostriker \(2007\)](#) and [Dunham et al. \(2014\)](#).

It is in the phase where the envelope has developed into an accretion disc and the star has begun to burn deuterium that most of the physics that is interesting for planet formation happens. Such discs provide a rich reservoir of chemical species at various temperatures, hydrodynamical and mechanical processes, all of which can under the right conditions conspire to form planets.

1.1.1 Spectral Evidence

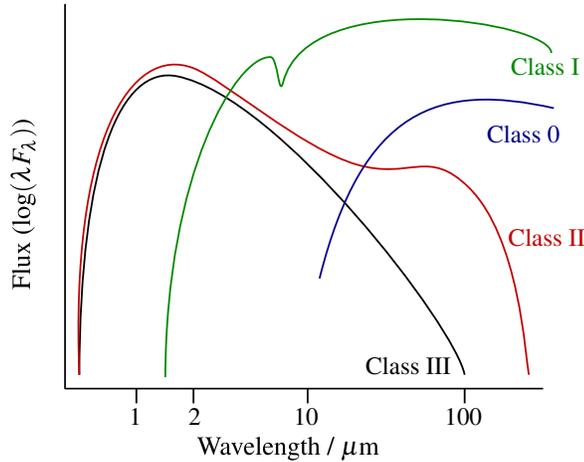


Figure 1.1: Schematic diagram showing typical spectra for various classes of young stellar objects. Class I-III spectra are based on [Lada \(1987\)](#), while the class 0 spectrum is based on [Andre et al. \(1993\)](#). The envelope around class I objects is responsible for the deep absorption feature at around $10 \mu\text{m}$ and is caused by amorphous silicates (see e.g. [Furlan et al., 2008](#)).

Discs around YSOs (Young Stellar Objects) are traditionally inferred from infra-red observations of young stellar objects. Perhaps the most common and useful technique for retrieving information about the disc of a YSO is the use of spectral indices and in particular, the classification system suggested by [Lada \(1987\)](#), which uses the parameter

$$\alpha_{IR} = \frac{d\log(\lambda F_\lambda)}{d\log(\lambda)} \quad (1.1)$$

This spectral index corresponds to the slope of the Spectral Energy Distribution (SED) of an object in the IR, and probes the dusty surface layers of the disc, where dust grains reprocess stellar radiation and re-emit it in the IR. Short and mid-IR wavelengths probe the hot dust that is closer to the star, whilst longer IR wavelengths probe the dust further out. This reprocessing causes the SED of the object to present a distinct excess of IR radiation in comparison to normal stellar photospheres, and allows sources to be grouped into one of four classes, named Classes 0-III. Class 0 sources appear essentially as a low temperature (15-30K) black-body, exhibiting very little emission in the traditional range of a stellar photosphere. For class I sources, $\alpha_{IR} > 0$. Class II sources exhibit a declining slope in the near- and mid-IR, leading to a mildly negative α_{IR} . Class III sources present almost no infra-red excess, meaning that α_{IR} is very negative ($\lesssim -1.5$), and the SED is essentially just that of a stellar photosphere.

Figure 1.1 shows a schematic representation of the spectra in this classification scheme. Note that the boundaries between these classifications are poorly defined, with various authors proposing slightly different schemes (e.g., [Andre et al., 1993](#); [Andre & Montmerle, 1994](#)).

These classifications are made entirely on the basis of the IR-excess in the spectrum of an object, and as such there is no *direct* link to the physical properties of an object. However, one can make some inferences from the physics expected in the star formation process. Class 0 objects are thought to be initial protostellar cores, surrounded by an optically thick envelope. Since they are essentially undergoing adiabatic collapse, they emit very little heat and therefore appear as cold blackbodies. Class I objects are traditionally interpreted as young stars which are no longer totally encompassed in an envelope and have begun to develop a rudimentary accretion disc. Class II objects are well-correlated with so called “classical T Tauri” stars - stars which are pre-main sequence and host significant discs of gas and dust, but no envelope. The large, optically-thick, extended disc reprocesses stellar radiation into all IR wavelengths. Classical T Tauri stars show strong accretion signatures, suggesting that the disc is being accreted onto the star. Class III objects are well-correlated with weak-lined T Tauri stars - stars whose discs have dissipated leaving at most a small debris disc which does minimal reprocessing of stellar radiation. Both types of T Tauri stars are actually classified according to their accretion signatures, with classical T Tauri stars showing strong accretion, while weak-lined T Tauri stars show little or no accretion.

1.1.2 Types of Disc

Protoplanetary Discs

Protoplanetary discs are extended, flattened discs of gas and dust that orbit young stars in the pre-main sequence phase. These objects are generally considered to be in the class II category, and are the direct result of the majority of the circumstellar disc being accreted onto the star in the class 0 and class I phases. Most of the planet formation process is thought to take place in these discs.

Determining the masses of protoplanetary discs is notoriously difficult, since most of the opacity is in the dusty component of the disc, whilst most of the mass (approximately 99%) is in the gaseous component. Nevertheless, spatially resolved, sub-millimeter observations of the optically thin dust in the outer ($> 10\text{AU}$) of discs can be used to derive estimates. For instance [Andrews & Williams](#)

(2007) find that the median radius of such discs is 200AU, whilst the median mass is $0.005 M_{\odot}$, or about 0.9% of the stellar mass. For comparison, they find that the median class I disc mass is 3.8% of the stellar mass, so a vast majority of the disc is accreted in the class I phase. This method of estimating disc masses does ignore larger grains of dust and pebbles in the disc, as these do not emit in the sub-millimetre. The review article by [Williams & Cieza \(2011\)](#) collects together various results for different stellar masses, showing that disc masses in the class II phase are broadly consistent with being 1% of the stellar mass, albeit with a large scatter. This same review provides comprehensive coverage of both observational and theoretical constraints on these discs.

Protoplanetary discs are thought to have a median lifetime of 2-3 Myr, with a maximum of around 10Myr. Throughout the lifetime of the disc, some material accretes onto the star, some is driven away by photoevaporation, and some of the dust agglomerates into larger bodies. The clearing of the vast majority of the disc leads to the transition into a class III object, though the exact interplay of the various processes that lead to this is not currently well understood.

Transitional Discs

Transitional discs exhibit minimal near- and mid-IR excesses - similar to a normal stellar photosphere, but still exhibit a strong excess in the far-IR. The most common interpretation of this is a dust hole in the inner disc, meaning that there is a lack of warm dust emitting in the near- and mid- IR, but there is still emission in the far-IR from the outer disc. In this sense, these discs are the curios of the aforementioned classification scheme, in that they fit somewhere between the class II and class III categories, and thus appear to be stars that are transitioning from having a full, optically-thick protoplanetary disc to an optically-thin debris disc (discussed below). [Espaillat et al. \(2014\)](#) provides a recent review of observations of these objects.

The formation of these discs is not well understood. The fact that there are relatively few of them compared to class II and class III sources suggests that the process that clears the disc is relatively fast, making it difficult to observe discs that are transitioning. The review by [Alexander et al. \(2014\)](#) summarises the processes by which the gas might be cleared quickly from the disc. The primary method by which this happens is photoevaporation, whereby radiation from the host star heats the atmosphere of the disc. Beyond a critical radius in the disc, this heating can cause the thermal energy

of the gas to exceed its gravitational binding energy, allowing it to escape the disc driving an outflow which depletes the disc at this critical radius. Once this happens, the inner disc and outer disc are no longer connected, and the inner disc rapidly drains onto the star, leaving a hole in the middle of the disc. Magnetically-driven winds may also contribute to this depletion. A further option for forming transitional discs is planets, which may carve out a gap in the dust due to gravitational interactions.

The methods above largely concentrate on clearing the gas out of the disc, but note that the lack of dust emission at small radii does not necessarily imply the lack of gas. So while it is possible that the gas drains first, taking small, coupled dust grains with it, transitional discs may also just be lacking in optically thick dust in their innermost regions. This could be due to the fact that the dust grains in the inner disc have agglomerated and grown in size such that only large, solid objects exist, which do not emit in the near- or mid-IR (see e.g., [Testi et al., 2014](#)).

Debris Discs

This final category of discs corresponds to the class III objects that have very little mid-IR emission relative to the Taurus median. These are thought to be objects which have completed the clearing process that formed the transition disc, leaving only a minimal remaining disc of optically thin dust and solid bodies. I shall not discuss these objects in detail here since they are of limited relevance for the work in this thesis, but I direct the interested reader to the recent review by [Matthews et al. \(2014\)](#).

1.1.3 Resolved Observations

An increasingly promising technique of understanding discs is direct imaging of them, which can reveal interesting structures that simply cannot be seen in IR spectra. For instance, polarized light images of the star HD 169142 reveal a gap between 40 and 70 AU in the disc with asymmetric brightness features, possibly suggestive of protoplanets ([Quanz et al., 2013](#)). ALMA observations of discs such as HL Tau have revealed spatially resolved, symmetric gaps in discs (see e.g., [ALMA Partnership et al., 2015](#)). Future results from ALMA will continue to shed light on the finer points of disc structure and evolution.

1.1.4 Theoretical behaviour

Radial profile

The gas within protoplanetary discs behaves in the traditional sense of an astrophysical accretion disc - the majority of the mass moves inwards and is eventually accreted onto the central star, whilst a tiny minority of the mass moves outwards toward infinity, carrying with it the bulk of the angular momentum. This is of course, assuming that some mechanism - for instance, viscosity - exists in the fluid to allow the transport of angular momentum. This viscosity converts the kinetic energy of the gas into heat which is radiated away. The only energy source in the system is the gravitational potential - in this case, of the central star - and thus in essence a viscous accretion disc liberates this potential energy as radiation.

We begin by deriving the mass continuity equation for such a disc following e.g., [Pringle \(1981\)](#). I will be going through this derivation in significantly more detail than the vast majority of sources, in the hope that future budding-astrophysicists might find it useful. To make progress, we assume that the disc is thin (that the vertical coordinate z is much smaller than the radial coordinate R) and therefore treat the disc in a vertically averaged manner, such that the density at any point is completely described by the surface density $\Sigma(R)$. We consider an annulus of width ΔR located at radius R from the centre of our coordinate system. This annulus has a mass that is equal to the surface density of the gas multiplied by the area of the annulus, or

$$m = \int_0^{2\pi} R \Delta R \Sigma(R) d\phi = 2\pi R \Delta R \Sigma(R) \quad (1.2)$$

Since we are concerned with only the mass change within this annulus, we take the partial derivative of this quantity with respect to time, using the product rule and realising that $\frac{\partial R}{\partial t} = 0$:

$$\frac{\partial m}{\partial t} = \frac{\partial}{\partial t} 2\pi R \Delta R \Sigma(R) = 2\pi R \Delta R \frac{\partial \Sigma(R)}{\partial t} \quad (1.3)$$

We now consider the mass flowing into and out of this annulus from the two neighbouring annuli. These are the only sources/sinks of mass for this annulus. Mass flows across the boundary at radius R with velocity $v_R(R)$ (only the radial component of velocity is important) and the boundary at radius $R + \Delta R$ with velocity $v_R(R + \Delta R)$. Therefore the total rate of mass flow into the annulus is

$$\frac{\partial m}{\partial t} = 2\pi R v_R(R, t) \Sigma(R, t) - 2\pi (R + \Delta R) v_R(R + \Delta R, t) \Sigma(R + \Delta R, t) \quad (1.4)$$

Equating equations 1.3 and 1.4, then dividing both sides by $2\pi\Delta R$, we find that

$$R \frac{\partial \Sigma(R)}{\partial t} = \frac{R v_R(R, t) \Sigma(R, t) - (R + \Delta R) v_R(R + \Delta R, t) \Sigma(R + \Delta R, t)}{\Delta R} \quad (1.5)$$

In the limit that $\Delta R \rightarrow 0$, this becomes

$$R \frac{\partial \Sigma(R)}{\partial t} = - \frac{\partial}{\partial R} (R \Sigma v_R) \quad (1.6)$$

and hence we have the equation of mass continuity for a thin disc.

We now turn our attention to the conservation of angular momentum across the disc. Writing the tangential velocity of the gas, v_ϕ as $v_\phi = \Omega R$ where Ω is the angular frequency, the angular momentum of an annulus is

$$L = \int R v_\phi dm = \int_0^{2\pi} \Sigma R^3 \Delta R \Omega d\phi = 2\pi \Omega R^3 \Delta R \Sigma \quad (1.7)$$

We again find the partial differential of this quantity with respect to time:

$$\frac{\partial L}{\partial t} = 2\pi R \Delta R \frac{\partial}{\partial t} (\Sigma \Omega R^2) \quad (1.8)$$

and note that the R^2 could be moved out of the differential but is left in for convenience such that the differential contains the angular momentum per unit area and the constant in front is simply the area. We then consider where this time variability originates. As before there is a flux of angular momentum that comes from the radial drift of material from the two neighbouring annuli, which naturally brings with it its tangential velocity. In addition to this, we have the viscous torques generated by the shear from the neighbouring annuli - a spin-up torque from the annulus at smaller R that is rotating more rapidly, and a spin-down torque from the slower-moving annulus further out. Hence the net change in angular momentum of the annulus is

$$\begin{aligned} \frac{\partial L}{\partial t} = & 2\pi R \Sigma(R, t) R^2 \Omega(R, t) v_R(R, t) \\ & - 2\pi R \Sigma(R + \Delta R, t) (R + \Delta R)^2 \Omega(R + \Delta R, t) v_R(R + \Delta R, t) \\ & - \Gamma_\nu(R, t) + \Gamma_\nu(R + \Delta R, t) \end{aligned} \quad (1.9)$$

where Γ_ν is the viscous torque due to a neighbouring annulus. We equate equations 1.8 and 1.9 and then divide through once again by $2\pi\Delta R$, finding:

$$\begin{aligned} R \frac{\partial}{\partial t} (\Sigma \Omega R^2) = & \frac{R \Sigma(R, t) R^2 \Omega(R, t) v_R(R, t)}{\Delta R} \\ & - \frac{R \Sigma(R + \Delta R, t) (R + \Delta R)^2 \Omega(R + \Delta R, t) v_R(R + \Delta R, t)}{\Delta R} \\ & - \frac{\Gamma_\nu(R, t) - \Gamma_\nu(R + \Delta R, t)}{2\pi \Delta R} \end{aligned} \quad (1.10)$$

In the limit that $\Delta R \rightarrow 0$, this becomes

$$R \frac{\partial}{\partial t} (\Sigma \Omega R^2) + \frac{\partial}{\partial R} (R \Sigma R^2 \Omega v_R) = \frac{1}{2\pi} \frac{\partial}{\partial R} \Gamma_v. \quad (1.11)$$

To consider the viscous torque, we start by defining a z -averaged kinematic viscosity ν :

$$\nu \Sigma = \int_{-\infty}^{\infty} \eta dz, \quad (1.12)$$

where η is the shear viscosity coefficient (Lodato, 2007). The left-hand side of the previous equation multiplied by the rate of shear - $R \cdot d\Omega/dR$ - is hence the viscous force per unit length around the circumference of the annulus. Evidently if there is no shear in the disc then there would be no torque as a result of the viscosity. To obtain the total viscous torque on the annulus, we multiply by the circumference $2\pi R$ and the lever arm of the torque R , finding

$$\Gamma_v = 2\pi R \cdot R \cdot \overline{R \nu \Sigma} \frac{d\Omega}{dR}. \quad (1.13)$$

We can substitute this into equation 1.11, cancelling the 2π and dividing through by R to find

$$\frac{\partial}{\partial t} (\Sigma \Omega R^2) + \frac{1}{R} \frac{\partial}{\partial R} (R \Sigma R^2 \Omega v_R) = \frac{1}{R} \frac{\partial}{\partial R} \left(R^3 \nu \Sigma \frac{d\Omega}{dR} \right) \quad (1.14)$$

The left hand side of this equation can be expanded using the product rule and assuming that $\partial\Omega/\partial t = 0$, we find

$$\Omega R^2 \left(R \frac{\partial \Sigma}{\partial t} + \frac{\partial}{\partial R} (R \Sigma v_R) \right) + R \Sigma v_R \frac{\partial}{\partial R} (\Omega R^2) = \frac{\partial}{\partial R} \left(R^3 \nu \Sigma \frac{d\Omega}{dR} \right). \quad (1.15)$$

The term in brackets is evidently identical to equation 1.6 and therefore 0. If we then divide through by $\frac{\partial}{\partial R} (\Omega R^2)$ and differentiate what remains, we find

$$\frac{\partial}{\partial R} (R \Sigma v_R) = \frac{\partial}{\partial R} \left[\frac{1}{\frac{\partial}{\partial R} (\Omega R^2)} \frac{\partial}{\partial R} \left(R^3 \nu \Sigma \frac{d\Omega}{dR} \right) \right], \quad (1.16)$$

where we can again make use of equation 1.6 to replace the left-hand side, finding

$$\frac{\partial \Sigma}{\partial t} = -\frac{1}{R} \frac{\partial}{\partial R} \left[\frac{1}{\frac{\partial}{\partial R} (\Omega R^2)} \frac{\partial}{\partial R} \left(R^3 \nu \Sigma \frac{d\Omega}{dR} \right) \right]. \quad (1.17)$$

Up to this point, we have made no assumptions about the rotation profile $\Omega(R)$ beyond that it does not vary in time. Hence equation 1.17 is valid for any rotation profile that meets this criterion. If we assume a Keplerian potential and neglect the disc's self-gravity however, we find that $\Omega = (GM/R^3)^{1/2}$, $\partial\Omega/\partial R = -3/2(GM/R^5)^{1/2}$ and $\partial\Omega R^2/\partial R = 1/2(GM/R)^{1/2}$ and thus equation 1.17 becomes

$$\frac{\partial \Sigma}{\partial t} = \frac{3}{R} \frac{\partial}{\partial R} \left[R^{1/2} \frac{\partial}{\partial R} (\nu \Sigma R^{1/2}) \right], \quad (1.18)$$

which is a partial differential equation describing how the value of Σ changes for a given annulus at radius R in the disc. It is a diffusion equation, relating the derivative of time to two spatial derivatives.

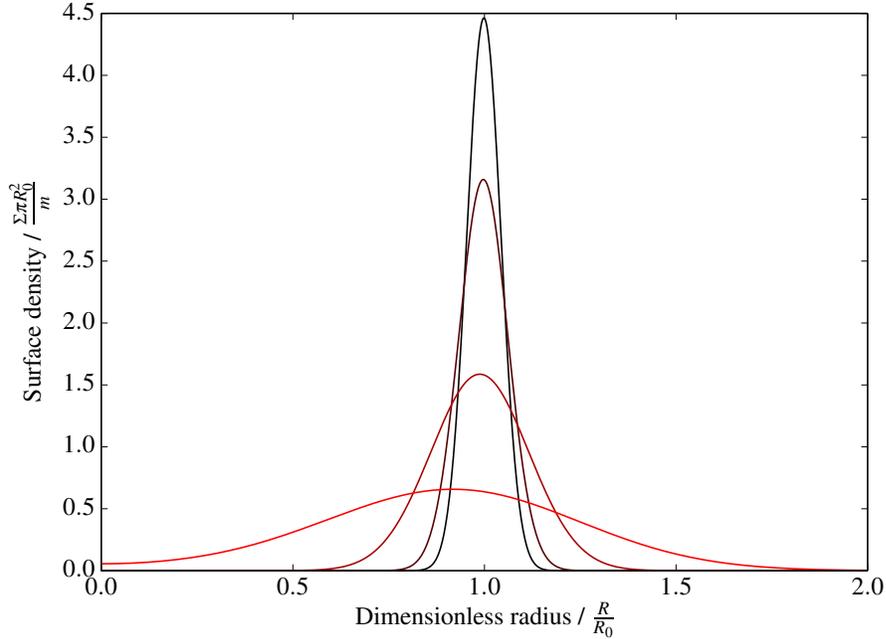


Figure 1.2: The evolution of a spreading ring at different instances of the dimensionless time, $\tau_{ring} = 12vt/R_0^2$. From darkest to lightest (tallest peak to shortest): $\tau_{ring} = 0.004$, $\tau_{ring} = 0.008$, $\tau_{ring} = 0.032$, $\tau_{ring} = 0.200$. The ring spreads out and the peak moves inwards, indicating that the majority of the mass is moving radially inwards. Note that y-axis scaling in this plot differs by a small factor from a similar plot in [Pringle \(1981\)](#), but other authors (e.g. [Dunhill, 2013](#)) have also found this discrepancy.

The form of this equation does not immediately reveal the behaviour of the mass in such a disc to us. Fortunately, there is an analytically tractable solution to the problem that demonstrates the behaviour of equation 1.18 clearly. The initial condition to this problem is a surface density profile of

$$\Sigma(R, t = 0) = \frac{m\delta(R - R_0)}{2\pi R_0}, \quad (1.19)$$

or in other words, an infinitesimally thin ring of matter with total mass m at radius R_0 . The surface density profile then evolves with time in accordance with equation 1.18. The solution to this problem can be expressed at a given dimensionless time $\tau_{ring} = 12vt/R_0^2$ in terms of a modified Bessel function of order 1/4 (see e.g. [Lynden-Bell & Pringle, 1974](#); [Pringle, 1981](#), for a detailed derivation and more information). Figure 1.2 shows plots of this analytical solution as time progresses. It shows that the

majority of mass in the disc moves inwards with time, as a result of the viscosity which allows angular momentum to be conveyed to the outer regions of the domain. This is the general behaviour of a viscous astrophysical disc: the majority of the mass moves inwards, passing its angular momentum to the minority of the mass which moves outwards.

Steady-state solution

The simplest way to obtain an analytical solution describing the surface-density profile, $\Sigma(R)$, of a viscous disc, is to make the assumption that the disc is in a steady-state; that the surface density in the disc does not vary in time. This requires the outer region of the disc be fed by an implicit source. It also requires that the movement of mass through the disc be constant in both space and time. Referring back to our previous annulus-based description of the disc, this means that the mass flowing into the annulus from further out in the disc must be equal to the mass flowing out of the annulus, such that equation 1.4 becomes

$$0 = \frac{\partial m}{\partial t} = 2\pi R v_R(R, t) \Sigma(R, t) - 2\pi(R + \Delta R) v_R(R + \Delta R, t) \Sigma(R + \Delta R, t) \quad (1.20)$$

which implies that the quantity $\dot{M} = -2\pi R v_R(R, t) \Sigma(R, t)$ is constant throughout the disc.

Beginning from equation 1.14, we set the temporal derivative term to zero and multiply through by R , leaving

$$\frac{\partial}{\partial R} (R \Sigma R^2 \Omega v_R) = \frac{\partial}{\partial R} \left(R^3 v \Sigma \frac{d\Omega}{dR} \right). \quad (1.21)$$

We may now rewrite this in terms of our constant \dot{M} , obtaining instead

$$\frac{\partial}{\partial R} (-\dot{M} R^2 \Omega) = 2\pi \frac{\partial}{\partial R} \left(R^3 v \Sigma \frac{d\Omega}{dR} \right). \quad (1.22)$$

We now integrate both sides between $R = R_{in}$ and an arbitrary outer R . We assume that $\frac{d\Omega}{dR} = 0$ at R_{in} , or in other words, that there is no torque at the inner boundary. This leaves us with

$$-\dot{M} R^2 \Omega + \dot{M} R_{in}^2 \Omega_{in} = 2\pi R^3 v \Sigma \frac{d\Omega}{dR}. \quad (1.23)$$

Now we once again use the substitution $\Omega = (GM/R^3)^{1/2}$ and therefore $\partial\Omega/\partial R = -3/2(GM/R^5)^{1/2}$ to obtain

$$-\dot{M} \sqrt{R} + \dot{M} \sqrt{R_{in}} = -3\pi \sqrt{R} v \Sigma. \quad (1.24)$$

Rearranging, we find that

$$\dot{M} \left(1 - \sqrt{\frac{R_{in}}{R}} \right) = 3\pi\nu\Sigma, \quad (1.25)$$

or when R is far from the inner boundary such that $R \gg R_{in}$

$$\dot{M} = 3\pi\nu\Sigma. \quad (1.26)$$

In other words, if we have a steady state disc then the mass accretion rate \dot{M} (which is spatially and temporally constant) is dependent upon the local viscosity and surface density in the disc, and therefore that we must understand the viscosity profile in the disc in order to understand the distribution of mass within it.

Vertical profile

We now consider the z profile of the disc. Hydrostatic equilibrium requires that in order for the disc to be stable in the vertical direction, the pressure must support the fluid against the gravity it feels. If we consider a minute parcel of fluid situated with its lowest point at z , with area A and height Δz , then the force on the top of the parcel due to pressure is

$$F_P(z + \Delta z) = -P(z + \Delta z) \cdot A \quad (1.27)$$

whilst the force experienced by the bottom of the parcel is

$$F_P(z) = P(z) \cdot A \quad (1.28)$$

Since the acceleration due to a gravitational potential Φ is just $\mathbf{a} = -\nabla\Phi$, the gravitational force on this parcel is its mass multiplied by the z component of this acceleration:

$$F_g = -\rho \cdot A \cdot \Delta z \cdot \frac{\partial\Phi}{\partial z} \quad (1.29)$$

Thus, for equilibrium, we require that these forces are in balance, and hence

$$0 = P(z) \cdot A - P(z + \Delta z) \cdot A - \rho \cdot A \cdot \Delta z \cdot \frac{\partial\Phi}{\partial z} \quad (1.30)$$

which we can rearrange to find

$$P(z + \Delta z) \cdot A - P(z) \cdot A = -\rho \cdot A \cdot \Delta z \cdot \frac{\partial\Phi}{\partial z} \quad (1.31)$$

If we now divide through by $A\Delta z$, we find that in the limit $\Delta z \rightarrow 0$

$$\frac{\partial P}{\partial z} = -\rho \frac{\partial \Phi}{\partial z} \quad (1.32)$$

For the meantime we can continue without loss of generality by substituting P for the ideal-gas equation of state $P = \rho \cdot c_s^2/\gamma$ (where c_s is the speed of sound and γ is the adiabatic index) and making no assumptions about the vertical temperature profile. The left hand side of equation 1.32 becomes

$$\frac{\partial P}{\partial z} = \frac{\partial}{\partial z} \frac{\rho c_s^2}{\gamma} = \frac{1}{\gamma} \left(c_s^2 \cdot \frac{\partial \rho}{\partial z} + \rho \cdot \frac{\partial c_s^2}{\partial z} \right) \quad (1.33)$$

which means that equation 1.32 can be rewritten

$$\frac{1}{\rho \gamma} \left(c_s^2 \cdot \frac{\partial \rho}{\partial z} + \rho \cdot \frac{\partial c_s^2}{\partial z} \right) = -\frac{\partial \Phi}{\partial z} \quad (1.34)$$

We now divide through by c_s^2 , and using $\partial \log(f(z))/\partial z = 1/f(z) \cdot \partial f(z)/\partial z$, we find

$$\frac{1}{\gamma} \left(\frac{\partial \log(\rho)}{\partial z} + \frac{\partial \log(c_s^2)}{\partial z} \right) = -\frac{1}{c_s^2} \frac{\partial \Phi}{\partial z} \quad (1.35)$$

which is an equation describing the z profile of an adiabatic disc in a general potential. We may now assume a Keplerian (point-mass) potential, $\Phi = -GM/r$ (where M is the mass of the central object and r is the magnitude of the displacement vector), and equation 1.35 becomes

$$\frac{1}{\gamma} \left(\frac{\partial \log(\rho)}{\partial z} + \frac{\partial \log(c_s^2)}{\partial z} \right) = \frac{1}{c_s^2} \frac{\partial}{\partial z} \left(\frac{GM}{(R^2 + z^2)^{1/2}} \right) \quad (1.36)$$

The potential derivative expands trivially as:

$$\frac{1}{c_s^2} \frac{\partial}{\partial z} \left(\frac{GM}{(R^2 + z^2)^{1/2}} \right) = -\frac{1}{c_s^2} \frac{GM}{(R^2 + z^2)^{3/2}} \cdot z = -\frac{1}{c_s^2} \frac{GM}{R^3} \cdot z \quad (1.37)$$

where we have used the thin disc approximation ($z \ll R$) to neglect the z^2 term in the denominator. I discuss the implications of this assumption further in appendix A. Finally, to make progress, we assume that the disc is isothermal in the z direction and hence that $\partial c_s^2/\partial z = 0$ and $\gamma = 1$. Combining the potential with equation 1.35 then gives

$$\frac{\partial \log(\rho)}{\partial z} = -\frac{1}{c_s^2} \frac{GM}{R^3} \cdot z = -\frac{\Omega_K^2 z}{c_s^2} \quad (1.38)$$

where we have introduced Ω_K , the Keplerian orbital frequency. Finally, integrating both sides from $z = 0$ to some arbitrary z gives

$$\rho = \rho_0 \exp \left(-\frac{\Omega_K^2 z^2}{c_s^2} \right) = \rho_0 \exp \left(-\frac{z^2}{2H^2} \right) \quad (1.39)$$

where $H = c_s/\Omega_k$ is known as the disc-scale height. Evidently the structure of a thin, vertically-isothermal disc is a Gaussian distribution in z , with H as the standard-deviation. This means that 68% of the mass of the disc lies within $\pm H$ of the mid-plane. The above derivation shows that H is itself a function of the sound speed and the radial rotation profile in the disc, and thus the profile of the disc depends upon what assumptions one makes regarding the radial temperature profile.

Angular momentum transport

The final piece in the puzzle is the viscosity of the disc. Evidently equation 1.17 relies upon the viscosity in the disc as a method for transporting angular momentum. The actual source of this transport in discs is unclear (and discussed in more detail later on), hence up to this point I have made no assumptions regarding its nature. However, we can still attempt to understand the long-term behaviour of discs by abstracting away this uncertainty into a dimensionless but physically-justified parameter. Assuming that the viscosity in a disc is turbulent in origin then the eddies generated by the turbulence have some length-scale l and some characteristic velocity v , such that the effective kinematic viscosity can be expressed as

$$\nu \approx lv \tag{1.40}$$

The turbulent motions have a maximum length scale that is the scale height of the disc, and - assuming that they are subsonic - a characteristic velocity that is less than the sound speed. As such, we can use these two quantities and a small constant to parametrise the viscosity:

$$\nu = \alpha c_s H \tag{1.41}$$

where $\alpha \leq 1$. This viscosity prescription was originally suggested by [Shakura & Sunyaev \(1973\)](#), and has remained a popular prescription for viscosity in disc simulations.

1.1.5 Sources of viscosity

The convenient thing about the alpha viscosity prescription presented above is that it allows us to consider the long-term behaviour of accretion discs without actually understanding the source of the viscosity. Protoplanetary discs do have accretion signature (e.g., [Williams & Cieza, 2011](#)), and thus it is clear that gas loses angular momentum in some way such that it falls onto the central star. However,

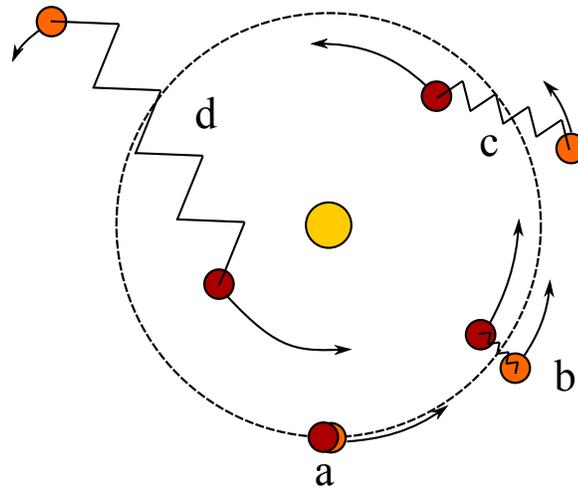


Figure 1.3: Explanation of the magnetorotational instability by way of analogy with a spring. At point a, two fluid elements are stacked vertically and thus orbit the star (yellow circle) at the same radius and with the same velocity, threaded by the same magnetic field line. By point b, a small radial displacement has occurred. The innermost element is now orbiting at higher velocity than the outermost element due to the keplerian nature of the disc. The magnetic field has been stretched and like a spring, attempts to pull the two elements back together. This causes the innermost element to slow down and fall inwards onto a faster orbit, whilst the outermost element speeds up and rises outwards onto a slower orbit. Thus, by point c, the displacement has grown much larger and the relative velocity of the two elements is greater. This runaway effect continues to point d, and it is clear that the elements will continue to be dragged away from one another.

the actual physical mechanism by which this angular momentum is lost from the disc or transported through it is not clear.

Magnetorotational Instability

Water, fire, air and dirt

\$@#!ing magnets, how do they work?

Insane Clown Posse

Miracles

The magnetorotational instability (MRI), sometimes known as the Balbus-Hawley instability, is a powerful shear instability that is in principle present in discs threaded by a weak magnetic field and can in principle produce both turbulence and angular momentum transport. It was first studied in detail in a series of papers starting with [Balbus & Hawley \(1991\)](#). It occurs in discs where the angular velocity decreases with increasing radius. The mechanism is simple: neighbouring fluid elements can

be threaded by the same magnetic field line. If these two elements are slightly displaced from one another radially, then they will be pulled apart by the shear: the innermost element will be moving faster than the outermost element. As they pull apart, the field line threading the two elements exerts a magnetic tension, which attempts to pull the two back together. Thus the magnetic tension robs the inner element of velocity and angular momentum, causing it to fall inward and onto a faster Keplerian orbit. The angular momentum is transferred to the outer element, which proceeds to move outward into a slower orbit. Clearly now the two elements are moving away from one another faster than they were before, yet the magnetic field keeps pulling on each of them, causing a runaway effect. This runaway cycle is the physical explanation of the MRI. Naturally this effect only works in the instance that the magnetic field is sufficiently weak that the magnetic tension does not completely overcome the shear. This leads to the slightly bizarre result that even a vanishingly small magnetic field ($B \rightarrow 0$) leads to an instability (Balbus, 2011).

The traditional MRI derivation concentrates on the ideal MHD regime, in which the protoplanetary disc is completely ionised. In general though, we expect protoplanetary discs to be relatively poorly ionised, due to the limited ability of stellar radiation to penetrate into the inner regions of the disc. In this case, the gas no longer behaves as it would in the ideal MHD regime, but is instead dominated in any given region by one of the following three non-ideal effects (see e.g., Wardle, 1999; Wardle & Salmeron, 2012):

- Ohmic dissipation - caused by collisions between electrons and neutrally charged species. If the density of neutrals is sufficiently high, then the electrons experience collisions often enough that they become decoupled from the magnetic field and the fluid becomes resistive. This effect hence dominates in dense regions, largely the mid-plane of the inner disc.
- Ambipolar diffusion - caused by collisions between positively charged and neutral species. This regime is valid when the magnetic force on the charged particles is dominant over the time average drag force caused by collisions with neutral particles. In this case, both the electrons and ions are well coupled to the magnetic field, and the collisions between ions and neutrals allow the magnetic stresses to be passed from the plasma into the dominant neutral gas. These collisions act to dissipate energy from the magnetic field, and hence damp any waves or turbulence that may arise in the plasma. This effect dominates in sparse regions, for instance the surface of

the disc and the outer disc (Wardle & Ng, 1999).

- The Hall effect - caused by the velocity difference between electrons and positively charged species. The velocity difference is the result of ions colliding with neutrals in the gas and hence becoming decoupled from the magnetic field, and instead coupled to the neutral gas. It is hence to the electrons that the magnetic field is frozen, and the electrons drift through the ions and neutrals. Charged grains may also act in the same fashion as ions here. The Hall effect dominates over the other two effects in the middle regions of the disc (Wardle & Salmeron, 2012).

The combination of these effects quenching the development of the MRI can lead to a so called “dead-zone”, first proposed by Gammie (1996). In essence the proposal is as follows: thermal ionisation ensures the disc is well-ionised out to somewhere between 0.1 and 1au. From hereon, the disc is sufficiently cool as to not be ionised, but cosmic rays continue to ionise a narrow surface layer of the disc. This leads to an interior layer of the disc being relatively unionised, dominated by ohmic diffusion and thus not susceptible to the MRI. The idea is then that accretion proceeds around the dead-zone and through the surface layers, but this model did not include ambipolar diffusion which can then go on to also damp the MRI in these surface layers.

Recent progress towards understanding the extent to which these non-ideal effects disrupt the MRI has concentrated on high-resolution, three-dimensional MHD simulations. It has generally been found that ohmic dissipation and ambipolar diffusion always damp the MRI, leading to the existence of magnetic “dead zone” in the disc, where the MRI cannot take hold. The Hall effect, however, can in principle revive these dead zones and allow the MRI to take hold regardless. Lesur et al. (2014a) used such simulations to demonstrate this. They found that at distances of 1-10au from the central star, none of the three non-ideal effects could be neglected, and that the addition of the Hall effect generates a strong azimuthal magnetic field which can aid the growth of the MRI within what would classically be the dead zone, leading to $\alpha \simeq 10^{-2}$ as opposed to a value of $\alpha \simeq 6 \times 10^{-4}$ found without including the Hall term. They found however that in this case the magnetic stresses generated by the MRI were in fact not associated with turbulence in the disc, but rather with laminar stresses that only appeared in the Hall-dominated regime. This effect is actually considered a manifestation of a separate instability - the “Hall–shearing instability” (Kunz, 2008) - rather than a variation of the MRI. It should

also be noted that whether the Hall effect strengthens or suppresses the MRI depends on whether the magnetic field is parallel or anti-parallel to the rotation of the disc. [Wardle & Salmeron \(2012\)](#) found using linear analysis that this effect can increase or decrease the magnetically active column density in a canonical disc by an order of magnitude depending on this orientation.

Further doubt has been cast on the role of the MRI in transporting angular momentum by recent observations. Using ALMA, [Flaherty et al. \(2015\)](#) measured the velocity dispersion in the outer regions ($R > 30AU$) and found that there was an order of magnitude less dispersion than might be expected for MRI induced turbulence. However, if the MRI takes hold in the large-scale, Hall-dominated regime studied by [Lesur et al. \(2014a\)](#) even further out than previously expected, one might not expect such great velocity dispersions anyway.

Other potential sources of viscosity

The MRI has typically been the most studied source of transport due to its potential to generate strong turbulence and high values of α , but it is by no means the only potential source of transport in canonical protoplanetary discs. For a review of possible methods of transport as well as a deeper look at current progress into understanding the MRI, I direct the reader to the recent review by [Turner et al. \(2014\)](#). For the sake of completeness, I shall mention some of these alternatives here.

Discs can be unstable to their own self-gravity in the event that the stabilising influence of pressure and rotation is overcome by the gravitational force exerted by one part of the disc on another. A more formal statement of this is given by the Toomre parameter, which states that the disc is unstable if:

$$Q \equiv \frac{c_s \kappa}{\pi G \Sigma} \lesssim 1 \tag{1.42}$$

where κ is the epicyclic frequency in the disc (Ω for a Keplerian disc). This condition is likely to be met only in the most massive protoplanetary discs, and even then potentially only in the outer regions of the disc (since $c_s \kappa$ drops off faster than Σ for canonical disc models). The instability can lead to the formation of non-axisymmetric structures which exert torques via their mutual gravitational force, thereby producing an effective viscosity. Numerical simulations (e.g., [Lodato & Rice, 2004](#)) suggest that $\alpha \simeq 5 \times 10^{-2}$ is easily achievable for discs up to 25% of the mass of their stellar host, assuming that the disc may cool fast enough. Observations suggest that discs can be particularly massive as objects move from the class I to class II regime, with [Launhardt & Sargent \(2001\)](#) finding a $0.3M_\odot$ disc in the

class I regime, so it is quite probable that gravitational turbulence plays a role in the viscous transport of very young discs, though this phase is likely to be rather short-lived in comparison to the total disc lifetime. Note however that the transport induced by gravitational instability is not necessarily local, since the gravitational force between clumps can couple parts of the disc that are widely separated (Balbus & Papaloizou, 1999).

Pure hydrodynamic instabilities can also cause turbulence which lead to viscosity, though these are less robust than MRI driven turbulence and thus likely to only have a significant impact on the disc in dead zones. Finally, MHD winds driven from the surface of the disc can remove angular momentum from the system, but note that this is not equivalent to a viscosity.

1.2 Observations of Exoplanets

Whilst the likes of Kepler, Copernicus and Galileo barely had the means to study the motion of the planets of our own solar system, advances in astronomical techniques that took place in the four centuries after Kepler's seminal work allowed late 20th century astronomers to begin pondering the existence of planets orbiting other stars in our galaxy. The introductory two paragraphs already briefly covered some of the greater advances in exoplanet detection, from Struve (1952) suggesting that super-Jupiters could be detected via spectroscopy, to Mayor & Queloz (1995) finding the first such exo-Jupiter. Since then, two methods - radial velocity and transit detection - have come to the forefront of exoplanet discovery, with several other methods also being commonly used. Figure 1.4 shows the current zoo of known exoplanets and how they were detected, and below I will summarise each of these techniques. For more information, I refer the interested reader to the recent and comprehensive review by Fischer et al. (2014a).

1.2.1 Radial Velocity

This method relies on the Doppler effect. A planet orbiting a star induces a reflex motion in the star, and both objects orbit their common centre of mass or barycentre. The displacement and velocity of the star from the barycentre is naturally smaller than that of the planet, but the accompanying Doppler shift in the stellar spectrum is detectable if the planet is large enough. Figure 1.5 demonstrates this

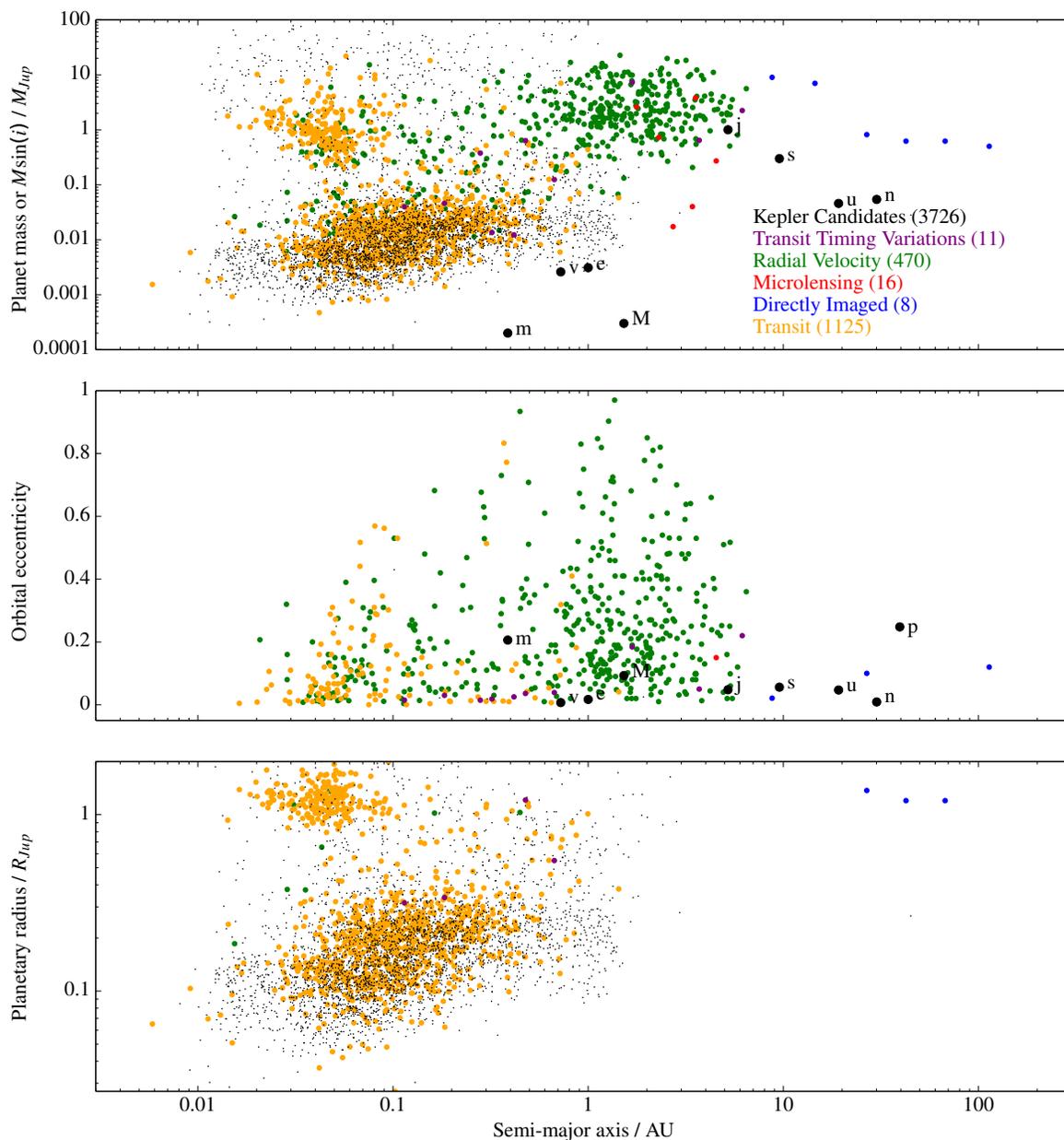


Figure 1.4: Properties of observed exoplanets, coloured by detection method, as of 22/04/2016. Masses for planets that only have measured radii were inferred from the scaling relation $m_p/m_\oplus = (r_p/r_\oplus)^{2.06}$ (Lissauer et al., 2011a) and are provided as a guide only. As such, the radii panel ostensibly gives a better overview of the size distribution of transiting planets, for which radii are generally well constrained. Our own solar system is plotted in the large, black, labelled points on the top two panels for comparison. Note a general trend for closer-in planets to have smaller eccentricities. Data source: exoplanets.org.

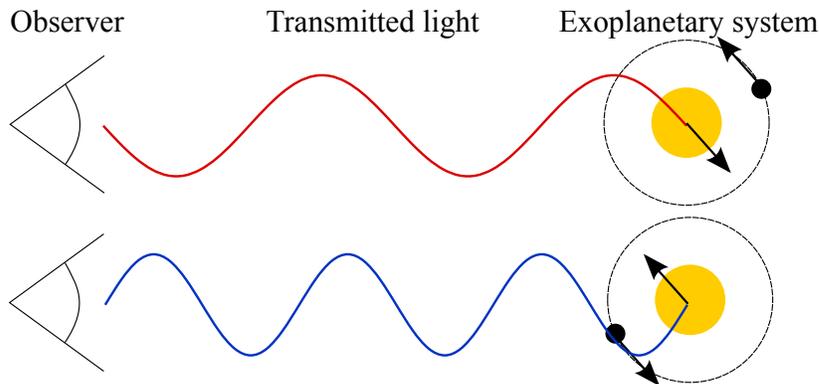


Figure 1.5: Schematic representation of how radial velocity detections work. The planet (black) and star (yellow) both orbit their common barycentre with the same period, though naturally the displacement of the planet from the barycentre is far greater than that of the star. When the planet is moving radially toward the observer, the star is thus moving away from them. This causes the light emitted from the star to appear redshifted. When the planet is on the other side of its orbit, moving radially away from the observer, the star is moving toward the observer. This causes the spectrum of the star to be blueshifted.

effect. More massive planets invoke a larger reflex motion in the star, and thus the mass of the planet can be determined by such measurements, and more massive planets that are nearer to their star are easier to detect. There is, unfortunately, a degeneracy in mass and inclination of planets detected via this method, and the quantity that is actually measured is $M_p \sin(i)/M_*$, which is evidently a lower limit on the mass, and requires precise measurement of the stellar mass. Eccentric orbits may also be detected - for a completely circular orbit, the radial velocity of the star varies perfectly sinusoidally, while an eccentric planet causes a deformed sinusoid which can be fitted to determine the eccentricity. Measuring the orbital period of a planet by this technique requires RV measurements of at least one orbit, which sets an upper limit on periods that can be detected, even with increasingly high precision. The RV method has the notable advantage of giving mass and eccentricity estimates for all of these planets, something which the transit method cannot do.

This technique has a long history. As mentioned before, (Struve, 1952) had already hinted at the possibility, and Griffin (1973) established that the required precision in stellar radial velocity measurements ($\approx 0.01 \text{ km s}^{-1}$) could be obtained by calibrating the spectrometer in question against suitable atmospheric spectral lines. This led to the aforementioned 51 Pegasi b detection. At this point the technique is relatively old and well-developed compared to the other techniques discussed here, with

the oldest RV planet searches having been in operation for almost 30 years (Fischer et al., 2014b).

Notable RV planet searches include Lick (Fischer et al., 2014b), Keck (Cumming et al., 2008), CORALIE (Marmier et al., 2013) and HARPS Mayor et al. (2011). Lick notably discovered the first multiple-planet extrasolar system around a main sequence star, Upsilon Andromedae (Butler et al., 1999). Cumming et al. (2008) use the results of Keck to suggest that between 17% and 19% of stars has a gas giant within 20AU, and that the occurrence rate of gas giants increases by a factor of five beyond periods of 200 days, suggesting a feature in the protoplanetary disc that traps giants at around 1AU (possibly an ice-line). With a 14 year baseline, Marmier et al. (2013) were able to identify 5 gas giants with a period beyond 10 years, finding no significant metallicity excess in all of their host stars. The HARPS survey targeted Neptune and super-Earth mass planets as well as Jupiters, and delivered some staggering statistics with regards to planetary occurrence rate. For instance, including planetary candidates, they find that 54% of stars play host to a super-Earth or Neptune mass planet in a < 100 day orbit, whilst 75% of stars host some mass of planet in a < 10 year orbit. Additionally, 13.9% of stars host some sort of gas giant planet in a less than 10 year orbit, though they note that only 0.5% of stars appear to host a hot Jupiter (< 11 day orbit)- a class of planet once thought to be rather typical of exoplanetary systems, but probably only as a result of biases.

1.2.2 Transits

The most fruitful method of detecting exoplanets to date has been the transit method. This method involves looking the periodic dips in the light curve of the host star, caused by the planet passing through the line of sight between observer and star, obscuring some fraction of the light (see figure 1.6). Detecting exoplanets via transit relies on precise (mmag) photometry of the host star, since the actual fraction of obscuration is relatively low. The dimming fraction f in the stellar light curve caused by a planet of radius R_p transiting in front of a star of radius R_* is easily computed from a simple geometric argument (see e.g. Wright & Gaudi, 2013):

$$f = \frac{R_p^2}{R_*^2}. \quad (1.43)$$

Jupiter is approximately 10% of the radius of the Sun, and thus the dimming fraction in this case is a mere 1%, dropping to 0.01 % for the case of an Earth-sized planet. Evidently, planets with larger radii

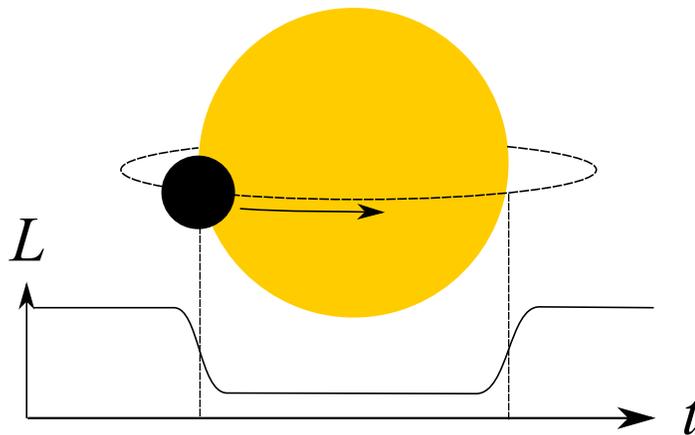


Figure 1.6: Schematic representation of how transiting planets are detected. The planet (black) orbits the star (yellow), following an orbit along the dashed line as it goes. As the planet begins to transit in front of the star, it obscures some of the light being emitted by the star, and the apparent luminosity L of the star begins to drop. As more of the planet obscures the star, the luminosity drops further. The width of this drop is related to the diameter of the planet and its orbital velocity. The obscuration reaches its maximum as the entire planet appears in front of the middle of the star relative to the observer. Note that the overall drop in flux is massively exaggerated for this diagram - drops of order 1% are the maximum in reality.

are easier to detect. To verify the periodicity of the signal, one must also wait to observe 2-3 transits. As a result of this, shorter period planets are easier to detect, since they can be detected with shorter observations. This method can also yield detections of non-transiting planets by way of transit-timing variations (TTVs). These occur when an unseen companion provides a perturbation to the orbit of a transiting planet, causing the transits of this planet to occur in a slightly aperiodic fashion. Dynamical models can then be used to fit these variations, allowing non-transiting companions to be detected with high sensitivity (see e.g., [Nesvorný et al., 2012](#)).

Notable transit surveys include the ground-based WASP (see e.g. [Christian et al., 2006](#)) and the space-based COROT (see e.g. [Deleuil et al., 2011](#)). Of the many transit-based detection missions to date, the Kepler mission has proved to be by far the most successful. Kepler is a space-based telescope in a heliocentric orbit hosting 42 CCDs, originally designed to monitor 170,000 stars over 4 years in the optical waveband over a 105 square degree field of view ([Koch et al., 2010](#)).

The most recent Kepler data release ([Coughlin et al., 2015](#)) contains analysis of the entire 48 month data set of Kepler’s original mission. This catalogue contains 4,293 planetary candidates in total. “Planetary candidates” are inferred from periodic dips in the stellar light curve seen by Kepler

that are characteristic of planets but lacking any further evidence to support their existence. Planets only become “confirmed” or “validated” once follow-up observations or statistical arguments provide further evidence of their existence respectively (see e.g., [Borucki, 2016](#)). 985 Kepler candidates are considered to be “confirmed” planets as a result of follow-up observations, though 9 of these were excluded from the [Coughlin et al. \(2015\)](#) catalogue as a result of issues with their vetting algorithm. The vast majority of the confirmed planets are visible in figure 1.4 as being detected via the transit method, whilst the candidates are plotted as the small grey points.

In the 2014 *Kepler* data release ([Burke et al., 2014](#)) there were 2738 planet candidates around 2017 unique stars. 475 (24%) of these stars host multiple candidates, while 1196 (44%) of the candidates are found in multiple-planet systems. The [Coughlin et al. \(2015\)](#) catalogue contains 4293 candidates around 3355 stars, of which 636 host multiple candidates. 1632 candidates in this catalogue are in multi-candidate systems. This vast increase demonstrates the astounding potential of the transit method to detect large quantities of new planets in a relatively short time period. These figures also demonstrate one very interesting result from the Kepler mission: a huge number of stars host multiple planets.

Tightly-packed Kepler systems and hot super-Earths

Discussion of the Kepler results leads naturally to a class of planets hugely prevalent in this data: super-Earths and sub-Neptunes. These planets are somewhere between 2 and 8 M_{\oplus} , or equivalently between 6×10^{-3} and $2.5 \times 10^{-2}M_{\text{Jup}}$ (figure 1.4)¹. It is readily apparent from this figure that a large number of transiting planets and Kepler candidate planets lie in this regime, and that the vast majority of those that have been detected have far shorter periods than Earth. Whilst this is partially a selection effect, the sheer number of these planets in tight orbits suggests they must be an important and common outcome of planet formation processes.

What is even more startling about this class of planets is that they are often found in multiples. For instance, Kepler-11 ([Lissauer et al., 2011b](#)), Kepler-32 ([Swift et al., 2013](#)) and Kepler-90 ([Cabrera et al., 2014](#); [Schmitt et al., 2014](#)) all contain multiple (> 4) super-Earth planets with periods less than that of Venus. I will discuss the specifics of these systems further in chapters 3 and 4. Thirty years

¹Neptune is $17M_{\oplus}$, so there is some overlap between the terms super-Earth and sub-Neptune.

ago however, the idea that multiple such planets could orbit within the inner confines of our own solar system would have been difficult to justify, but now it seems that this sort of tightly-packed system of planets is commonplace. Furthermore, these systems seem to be remarkably dynamically cold. Kepler-11 has been shown by the transit-timing variation method to have low eccentricities (Lissauer et al., 2013), and the very fact that we observe so many of these systems via the transit method suggests that they are generally both very stable and very coplanar, contrary to what we might expect given the strong gravitational interactions between planets that are so close together.

The need to understand these systems is supported by the previous results I mentioned from the HARPS RV survey (Mayor et al., 2011). In addition to finding that 54% of stars host a super-Earth or a Neptune in a less than 100 day orbit, they also found that 45% of stars host such a planet in a less than 50 day orbit. These percentages were obtained without extrapolating out of the domain of detectable planets, so in all likelihood this figure will only increase with time and sensitivity. Given that Venus orbits with a period of 225 days, the proximity of these super-Earths and Neptunes to their host stars is quite amazing. Results such as these are increasingly showing that our solar system is not representative of an average planetary system.

1.2.3 Direct imaging & other methods

Direct imaging of exoplanets seems in principle straightforward, but is in practice fraught with difficulty. Planets are largely visible due to the optical light they reflect from their host stars, though young planets also emit considerable thermal radiation in the IR. However, the star is always many magnitudes brighter than the planet, making star-planet contrast an issue, and meaning that giant planets that are further away from their hosts are easier to detect. The first direct image of an exoplanet was, for this exact reason, around a brown dwarf in a 55AU orbit, and detected in the IR (Chauvin et al., 2004). It would be another 4 years before an exoplanet was directly imaged around a main sequence star with the discovery of the $8 M_{\text{Jup}}$ β pic b, detected in the near-IR at an orbital radius of 8 AU. Around the same time, a multiple planet system was discovered around HR 8799, consisting of multiple super-Jupiters in wide orbits (Marois et al., 2008). A further super-Jupiter was discovered orbiting internal to these three just two years later (Marois et al., 2010). Direct imaging remains largely sensitive to giant planets with large separations and has therefore been the method of choice for only a handful

of detections, as shown in in figure 1.4. However, [Guyon et al. \(2012\)](#) suggests that the upcoming Extremely Large Telescopes will be able to image rocky planets in the habitable zones of M dwarfs, with later space-based missions achieving the same goal around Sun-like (F-G-K) stars.

The final common method for finding exoplanets is via gravitational microlensing. When one star (a lens) passes in front of another (a source) in the line of sight, the light from the source is distorted by the lens, due to general relativistic effects. A planetary companion to the lens star causes an additional distortion to the light from the source, detectable as a strong and distinctive perturbation in the light curve of the lens. The number of detections by this method so far has been fairly minimal, but dedicated lensing surveys such as OGLE ([Udalski, 2003](#)) are still a promising method of detecting planets that might otherwise be unobservable.

1.3 Formation of planetary systems

Planets form in protoplanetary discs around young stars. This idea dates back as far as [Kant \(1755\)](#), who suggested in his *Allgemeine Naturgeschichte und Theorie des Himmels (Universal Natural History and the Theory of Heaven)* that the solar system formed from material that was initially nebulous. The most widely-accepted model today is based on that of [Safronov \(1972\)](#). In this scenario, solid grains of dust and ice in the protoplanetary disc collide and combine to form planetesimals - solid object large enough to interact via their own gravity. These planetesimals then collide and combine to form terrestrial planets, and some fraction of these terrestrial planets accrete gas and become the cores of gas giants.

The first stage involves the growth of micrometer-sized dust grains into planetesimals that are tens or hundreds of kilometres in radius. The smallest grains stick trivially due to inter-molecular forces. Dust can be concentrated in the mid-plane of the disc by gravity, enhancing this process. Drag causes these larger objects to migrate inwards however, leading to the so called “metre size barrier” where metre-sized objects migrate rapidly inwards and are consumed by the star ([Weidenschilling, 1977](#)). Apart from this, once the particles reach centimetre sizes, they are less likely to stick and more likely to shatter or bounce on collision. Evidently, collisional growth beyond centimetre and metre sizes is challenging, and alternative explanations have been suggested. For instance, an instability known

as the streaming instability can lead to the collapse of self-gravitating clumps of such pebbles into planetesimals (Youdin & Goodman, 2005). The instability is caused by the cm-sized particles experiencing a headwind from the gas, which forces them to clump together in large concentrations. This has proven to be a promising alternative to collisional growth, and suggests that discs of planetesimals might form readily. Indeed, objects in the asteroid belt, Kuiper Belt and Oort Cloud in our own solar system suggest that planetesimal formation is a ubiquitous process, even if we do not yet understand their formation fully. More information on the formation of planetesimals can be found in the recent reviews by Chiang & Youdin (2010) and Johansen et al. (2014).

Assuming that a disc of planetesimals may form, the next stage is that planetesimals collide with one another (see e.g., the review by Raymond et al., 2014). The gravity of two planetesimals is great enough to hold them together once they collide. Larger planetesimals tend to grow faster since the gravitational focusing (increased relative attraction experienced by other bodies) caused by their greater mass means they experience more collisions. This leads them to experience runaway growth up to a mass known as the isolation mass, at which point the protoplanet has consumed all of the planetesimals within its Hill radius and thus has no material left to feed on. Numerical models suggest that only a few planetesimals will enter this runaway phase, the rest remaining as relatively small (see e.g., Greenberg et al., 1978; Wetherill & Stewart, 1989). At this stage, the protoplanets enter an oligarchic growth phase, where smaller protoplanets grow faster than larger ones, but their mass relative to the planetesimals still increases. During this stage, the protoplanets are kept separated from one another by orbital repulsion, being separated by approximately five mutual Hill radii (Kokubo & Ida, 1995, 1998, 2002). Alternative models for the growth of terrestrial-mass protoplanets include pebble accretion (Lambrechts & Johansen, 2012), where the largest planetesimals efficiently sweep pebble-sized objects out of the disc, experiencing rapid growth. Models such as this can help to overcome time-scale issues present in the more conventional planetesimal accretion theory.

Collisions between multiple isolation-mass objects can lead to an increase in size producing terrestrial planets of Earth and super-Earth mass. Alternatively, these objects can go on to become the cores of Jupiter mass planets. This begins with the protoplanets beginning to accrete gas from the disc, undergoing a process known as core accretion. The first self-consistent simulations of this process were carried out by Pollack et al. (1996), but the model was well established long before this.

Perri & Cameron (1974) for instance, suggested that a core might gravitationally attract a large envelope of gas which could collapse onto the core once a critical mass was reached. The theory was further developed by authors such as Bodenheimer & Pollack (1986), Wuchterl (1999) and Hubickyj et al. (2005), largely using progressively more advanced numerical simulations.

Accretion onto the core is slow (hydrostatic) at first - the planet must contract to accrete more mass, and since the rocky core cannot contract, the contraction is slow, limiting the accretion rate. Once the mass of the gaseous envelope around the planet is approximately equal to the mass of the core (this takes of order 1-10 Myr, depending upon dust opacities), the planet can contract much more quickly, and accrete as fast as the disc can feed it. This leads to a runaway growth phase, which is short ($10^4 - 10^5$ yr, compared to disc lifetimes of Myr) but can increase the mass of the planet by an order of magnitude, from something slightly larger than a terrestrial planet to a Jupiter mass giant. The main issue with this method is that reaching the runaway growth phase requires efficient cooling of the gaseous envelope, which is tricky to achieve with canonical grain opacities. Indeed, the time-scale for the hydrostatic growth phase can be as long as the lifetime of a protoplanetary disc. This has led some authors to invoke somewhat-arbitrarily lower opacities to reduce the time-scale of this phase (Hubickyj et al., 2005).

An alternative model for giant planet formation is via gravitational instability, whereby large chunks of the disc collapse under their own self gravity - the same process discussed as a driver of angular momentum transport in section 1.1.5 (see e.g. Boss, 2000). It has been proposed that this method could also form terrestrial planets via stripping of the atmosphere from young, giant planets, and that this model could even reproduce solar system features such as debris discs (Nayakshin, 2010; Fletcher & Nayakshin, 2016). I direct the interested reader to the recent review by Helled et al. (2014) for more information on both models of giant planet formation.

1.4 Orbital mechanics

In a system of N point-mass particles, the potential at the location of particle i is given by the following sum over all other particles

$$\Phi_i = -G \sum_{j=0}^N \frac{m_j}{|\mathbf{r}_{ij}|}. \quad (1.44)$$

where $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$, m_i is the mass of particle i and \mathbf{r}_i is the position vector of particle i . The force that particle i feels as a result of this potential is the negative of the spatial derivative of its potential energy, $\mathbf{F}_i = -m\nabla\Phi_i$, and therefore from Newton's second law:

$$m_i\ddot{\mathbf{r}}_i = -G \sum_{j=0}^N \frac{m_i m_j}{|\mathbf{r}_{ij}|^3} \mathbf{r}_{ij}. \quad (1.45)$$

1.4.1 The two body problem

The simplest way to gain an insight into the motion of a particle in the potential in equation 1.44 is to consider the limiting case that $N = 2$, making the equations analytically tractable. This case is particularly relevant for planetary systems, where the interaction between each planet and its host star generally far outweighs any planet-planet interactions, meaning each planet can be considered as a single body orbiting its host.

We consider two point mass particles, 1 and 2. Looking at equation 1.45, it is clear that m_i can be cancelled from each side and thus

$$\ddot{\mathbf{r}}_1 = -\frac{Gm_2}{|\mathbf{r}_{12}|^3} \mathbf{r}_{12}, \quad \ddot{\mathbf{r}}_2 = \frac{Gm_1}{|\mathbf{r}_{12}|^3} \mathbf{r}_{12} \quad (1.46)$$

Note that here, \mathbf{r}_1 and \mathbf{r}_2 are measured here in some hypothetical inertial (non-accelerating, non-rotating) reference frame. Since $\ddot{\mathbf{r}}_{12} = \ddot{\mathbf{r}}_1 - \ddot{\mathbf{r}}_2$, equations 1.46 can be subtracted from one another, yielding

$$\ddot{\mathbf{r}}_{12} = -\frac{G(m_1 + m_2)}{|\mathbf{r}_{12}|^3} \mathbf{r}_{12}. \quad (1.47)$$

This equation describes the relative motion of the two bodies, under the influence of one another's gravity. Note that at this point the concept of an inertial frame becomes irrelevant to the problem. The equation of relative motion uses only relative vectors, and the length and orientation of these vectors will be the same if measured in a non-rotating but also non-inertial frame whose origin is coincident and moves with the centre of mass of particle 2 (Bate et al., 1971). In the case of planetary dynamics or any other problem where $m_1 \ll m_2$, we see that $G(m_1 + m_2) \approx Gm_2$. We can thus define the standard gravitational parameter, $\mu \equiv Gm_2$ and rewrite equation 1.47 as

$$\ddot{\mathbf{r}} + \frac{\mu}{r^3} \mathbf{r} = 0 \quad (1.48)$$

where we have redefined \mathbf{r}_{12} as \mathbf{r} , and from hereon I shall refer to the vector magnitude $|\mathbf{a}|$ as simply a for convenience. This equation can reveal some surprisingly insightful things about the two body motion. To start, we vector multiply this equation by the radius vector \mathbf{r} .

$$\ddot{\mathbf{r}} \times \mathbf{r} + \frac{\mu}{r^3} \mathbf{r} \times \mathbf{r} = 0. \quad (1.49)$$

The second term on the left hand side is trivially 0 (since $\mathbf{a} \times \mathbf{a} = 0$ for any vector \mathbf{a}). The first term can be integrated, since the product rule shows us that $d/dt(\mathbf{r} \times \dot{\mathbf{r}}) = \dot{\mathbf{r}} \times \dot{\mathbf{r}} + \mathbf{r} \times \ddot{\mathbf{r}}$ and $\dot{\mathbf{r}} \times \dot{\mathbf{r}} = 0$. Thus the integration yields

$$\mathbf{r} \times \dot{\mathbf{r}} = \mathbf{h} \quad (1.50)$$

where \mathbf{h} is simply a constant of integration. That \mathbf{h} is a constant in time implies that both its direction and magnitude do not change in time, and that \mathbf{r} , $\dot{\mathbf{r}}$ and therefore the motion of m_1 around m_2 are confined to a plane. Evidently the vector \mathbf{h} has dimensions of specific orbital angular momentum, and does approximate this quantity well when $m_1 \ll m_2$, showing that angular momentum is conserved. However, to get the true angular momentum of the system, one would have to derive this constant of motion in an inertial frame, with the position and velocity vectors referring to the centre of mass (Murray & Dermott, 1999).

The energy integral

We now derive a second constant of the two body motion using a similar premise. Taking the dot product of equation 1.48 with $\dot{\mathbf{r}}$ yields

$$\dot{\mathbf{r}} \cdot \ddot{\mathbf{r}} + \frac{\mu}{r^3} \dot{\mathbf{r}} \cdot \mathbf{r} = 0 \quad (1.51)$$

Invoking the equality $\dot{\mathbf{r}} \cdot \mathbf{r} = r\dot{r}$ ² and replacing the vector $\dot{\mathbf{r}}$ with \mathbf{v} to avoid confusion of $d|\mathbf{r}|/dt = \dot{r}$ with $|d\mathbf{r}/dt| = v$, we find

$$v\dot{v} + \frac{\mu}{r^3} r\dot{r} = 0 \quad (1.52)$$

²The equality $\dot{\mathbf{r}} \cdot \mathbf{r} = |\mathbf{r}| \cdot d|\mathbf{r}|/dt = r\dot{r}$ comes from expanding the dot product, $\dot{\mathbf{r}} \cdot \mathbf{r} = r_x\dot{r}_x + r_y\dot{r}_y + r_z\dot{r}_z$ and realising that this bears a similarity to the time-derivative of the vector magnitude, $d|\mathbf{r}|/dt = (r_x\dot{r}_x + r_y\dot{r}_y + r_z\dot{r}_z)/|\mathbf{r}|$. An alternate way of thinking about this is that the rate of change of the *magnitude* of a given vector \mathbf{r} is equal to the projection of the vector causing that change $\dot{\mathbf{r}}$ into the direction of the unit vector, $\mathbf{r}/|\mathbf{r}|$.

Inverting the product rule, we realise that $\mathbf{v} \dot{\mathbf{v}} = 1/2 \cdot d v^2 / dt$. Additionally, the second term can be written

$$\frac{\mu}{r^2} \frac{d|\mathbf{r}|}{dt} = \frac{d}{dt} \frac{-\mu}{r}, \quad (1.53)$$

which means that equation 1.52 can be written as

$$\frac{1}{2} \frac{d v^2}{dt} - \frac{d}{dt} \frac{\mu}{r} = 0, \quad (1.54)$$

which we can finally integrate in time to obtain

$$\frac{v^2}{2} - \frac{\mu}{r} = \mathcal{E}, \quad (1.55)$$

where \mathcal{E} is a constant of integration. This constant evidently has units of specific energy and we notice immediately that the first term appears to be the specific kinetic energy of the motion whilst the second term is the specific potential energy of the motion. Indeed, in the limiting case $m_1 \ll m_2$, this constant is the specific orbital energy of m_1 , and proves that this energy is conserved with time. Following a more complicated derivation using the barycentre of the system as an inertial frame, it can be shown that \mathcal{E} is the total energy of the two bodies divided by their reduced mass, $m_1 m_2 / (m_1 + m_2)$ (Murray & Dermott, 1999).

The form of the motion

The final piece of the puzzle is to understand exactly what form the motion in of the smaller body in the plane takes. To do this we must attempt to obtain a form of equation 1.48 which is integrable with respect to time, following Bate et al. (1971). Firstly, we take the cross product of equation 1.48 with respect to \mathbf{h} and rearrange, finding

$$\ddot{\mathbf{r}} \times \mathbf{h} = \frac{\mu}{r^3} (\mathbf{h} \times \mathbf{r}) \quad (1.56)$$

To reduce each side to a simple derivative in time requires a couple of identities. For the left hand side, we can use the product rule to find that

$$\frac{d}{dt} (\mathbf{v} \times \mathbf{h}) = \dot{\mathbf{r}} \times \mathbf{h} + \mathbf{v} \times \dot{\mathbf{h}} = \ddot{\mathbf{r}} \times \mathbf{h}, \quad (1.57)$$

since \mathbf{h} is a constant in time and therefore $\dot{\mathbf{h}} = 0$. The right hand side can be expanded like so:

$$\frac{\mu}{r^3} (\mathbf{h} \times \mathbf{r}) = \frac{\mu}{r^3} (\mathbf{r} \times \mathbf{v}) \times \mathbf{r}, \quad (1.58)$$

where we use the vector triple product ($(\mathbf{b} \times \mathbf{c}) \times \mathbf{a} = \mathbf{c}(\mathbf{a} \cdot \mathbf{b}) - \mathbf{b}(\mathbf{a} \cdot \mathbf{c})$) to obtain

$$\frac{\mu}{r^3}(\mathbf{h} \times \mathbf{r}) = \frac{\mu}{r^3}(\mathbf{v}(\mathbf{r} \cdot \mathbf{r}) - \mathbf{r}(\mathbf{r} \cdot \mathbf{v})), \quad (1.59)$$

which, using $\mathbf{r} \cdot \mathbf{r} = rr$ and $\mathbf{r} \cdot \mathbf{v} = \mathbf{r} \cdot \dot{\mathbf{r}} = r\dot{r}$ becomes

$$\frac{\mu}{r^3}(\mathbf{h} \times \mathbf{r}) = \frac{\mu}{r}\mathbf{v} - \frac{\mu\dot{r}}{r^2}\mathbf{r}. \quad (1.60)$$

Noting that this equation - if divided by μ - has units of inverse time, and that the two terms both bear a resemblance to the unit vector \mathbf{r}/r , we find the time derivative of this unit vector

$$\frac{d}{dt} \frac{\mathbf{r}}{r} = \frac{d\mathbf{r}}{dt} \cdot \frac{1}{r} + \mathbf{r} \frac{d}{dt} \frac{1}{r} = \frac{\mathbf{v}}{r} - \frac{\dot{r}}{r^2}\mathbf{r}, \quad (1.61)$$

and we hence see that equation 1.60 can be rewritten

$$\frac{\mu}{r^3}(\mathbf{h} \times \mathbf{r}) = \mu \frac{d}{dt} \frac{\mathbf{r}}{r}. \quad (1.62)$$

We now use this identity and equation 1.57 to rewrite equation 1.56 as

$$\frac{d}{dt}(\mathbf{v} \times \mathbf{h}) = \mu \frac{d}{dt} \frac{\mathbf{r}}{r}. \quad (1.63)$$

which can be integrated to yield

$$\mathbf{v} \times \mathbf{h} = \mu \frac{\mathbf{r}}{r} + \mathbf{B}, \quad (1.64)$$

where \mathbf{B} is a vector constant of integration. Taking the dot product of \mathbf{r} with this equation yields

$$\mathbf{r} \cdot (\mathbf{v} \times \mathbf{h}) = \mu \frac{\mathbf{r} \cdot \mathbf{r}}{r} + \mathbf{r} \cdot \mathbf{B}. \quad (1.65)$$

Here, we may use the scalar triple product ($\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \mathbf{c} \cdot (\mathbf{a} \times \mathbf{b})$) to rewrite the left-hand side as $\mathbf{h} \cdot (\mathbf{r} \times \mathbf{v}) = \mathbf{h} \cdot \mathbf{h} = h^2$, and the definition of the scalar product to simplify the right-hand side, yielding

$$h^2 = \mu r + rB\cos(\theta), \quad (1.66)$$

where θ is the angle between the radius vector \mathbf{r} and the constant vector \mathbf{B} . This equation can now be trivially rearranged to find r :

$$r = \frac{h^2/\mu}{1 + (B/\mu)\cos(\theta)}. \quad (1.67)$$

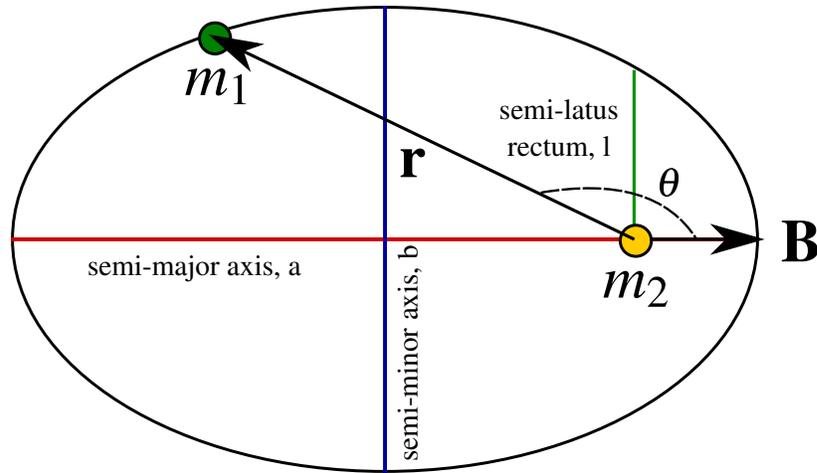


Figure 1.7: Schematic diagram showing an elliptical orbit, as defined by equation 1.67 with $B/\mu = e < 1$. The ellipse has 2 foci, at one of which resides the mass m_2 . The smaller mass m_1 traces an ellipse with m_2 at its origin. The ellipse has two axes across which it is symmetric, the major axis (length $2a$, red) and the minor axes (length $2b$, blue). The radius vector \mathbf{r} subtends an angle θ to a constant vector \mathbf{B} , which points in the direction of closest approach between the two bodies.

This equation is very instructive. It informs us that the magnitude of the radius vector of the orbiting body r changes as the vector itself rotates relative to some stationary vector \mathbf{B} . Moreover, this equation is evidently the equivalent of the equation of a conic section in polar coordinates:

$$r = \frac{l}{1 + e \cos(\theta)}. \quad (1.68)$$

Here l is the semi-latus rectum of the conic section, and e is its eccentricity. The eccentricity e describes the the shape of the conic. Conic sections are divided into 3 classes: ellipses ($0 \leq e < 1$), parabolas ($e = 1$) and hyperbolas ($e > 1$). In the case of a bound planetary orbit, this equation describes an ellipse, and the semi-latus rectum $l = a(1 - e^2)$, where a is the semi-major axis of the ellipse (see figure 1.7). In this case, the eccentricity can be expressed as $e = \sqrt{1 - a^2/b^2}$, where b is the semi-minor axis. Comparing equations 1.67 and 1.68, we see that $h^2/\mu = l$, and therefore we can express the magnitude of the vector \mathbf{h} as

$$h = \sqrt{\mu a(1 - e^2)}. \quad (1.69)$$

Orbital elements & other properties

We see from equation 1.67 that the minimum value of r occurs when $\theta = 0$, i.e. when vectors \mathbf{r} and \mathbf{B} point in the same direction. This is the point of closest approach and is called the periapse of the orbit. Conversely, the maximum value of r occurs when $\theta = \pi$ (\mathbf{r} and \mathbf{B} are anti-parallel). It is the furthest distance that m_1 moves from its larger host m_2 , and is known as the apoapse. This leads us naturally to the idea that \mathbf{B} is a vector that points to the periapse. Again comparing equations 1.67 and 1.68, we see that the magnitude of the vector \mathbf{B} is related to e by $B = \mu e$. This further leads us to intuit that there exists a vector \mathbf{e} known as the eccentricity vector, whose magnitude is equal to the eccentricity of the orbit and whose direction points to the periapse of the orbit. We can recover an expression for this vector by substituting $\mathbf{B} = \mu \mathbf{e}$ into equation 1.64, finding

$$\mathbf{e} = \frac{\mathbf{v} \times \mathbf{h}}{\mu} - \frac{\mathbf{r}}{r}. \quad (1.70)$$

Note that since \mathbf{B} and μ are constant in time, this eccentricity vector and, therefore, the eccentricity of the orbit are also constant in time. We may rewrite this equation using the definition of \mathbf{h} and the vector triple product to obtain an expression for the eccentricity vector that does not require the computation of \mathbf{h} .

The quantities a and e are two of a series of six orbital elements that can be used to completely describe the orbit of body m_1 around m_2 . We already have an expression for h in terms of these elements (equation 1.69), but it would be convenient to express the energy integral (equation 1.55) in this form. Since the integral is constant around the orbit, we consider the simple case of m_1 at its periapse. The radius of the periapse occurs when $\cos(\theta) = 1$ and therefore $r_p = l/(1 + e)$, and we know that $l = a(1 - e^2)$; hence $r_p = a(1 - e)$. Thus, we can replace the r dependence in equation 1.55 with something that depends only on the orbital elements. We eliminate v from equation 1.55 by considering that h at periapse is given by $h = r_p \cdot v_p$. Using the definition of h (equation 1.69), we then find

$$\mathcal{E} = \frac{-\mu}{2a}. \quad (1.71)$$

As a result, we can express the main dynamical properties of an orbit (energy and angular momentum) in terms of its orbital elements.

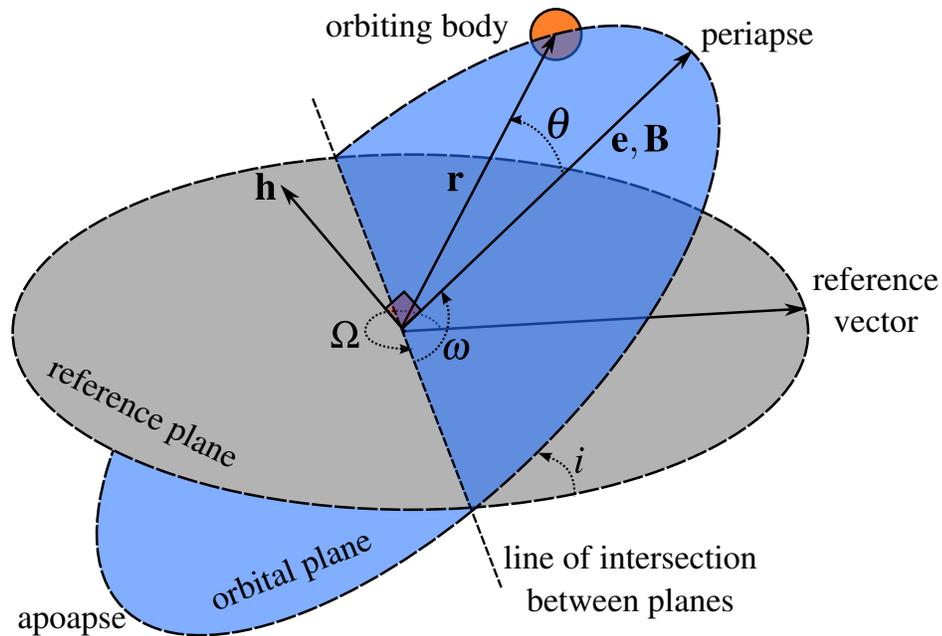


Figure 1.8: Schematic diagram showing an elliptical orbit, inclined to an arbitrary reference plane, with orbital elements and vectors. Solid lines represent vector quantities: the specific orbital angular momentum \mathbf{h} , which is perpendicular to the orbital plane, the radius vector \mathbf{r} from the larger body m_2 to its smaller companion m_1 , the eccentricity vector \mathbf{e} , and our constant vector \mathbf{B} . \mathbf{e} and \mathbf{B} both point directly to the periape of the orbit. Dotted lines represent angle quantities: orbital inclination i , longitude of the ascending node Ω , the argument of periape ω , and the true anomaly θ . See text for more details regarding all quantities.

The other elements we need to define an orbit in space are largely less important to exoplanets, since they define the orientation of the orbit in space (see figure 1.8) in ways that are virtually impossible to measure with current exoplanet missions. We have already seen that m_1 moves in a plane to which \mathbf{h} is perpendicular. Evidently with respect to some reference vector and plane, we can incline this orbital plane - and therefore \mathbf{h} - at an arbitrary angle without changing the physical properties of the orbit. This angle i is known as the inclination. Now that we have allowed the orbit to be inclined with respect to the reference plane, the orbital plane intersects our reference plane. The angle that the line of intersection makes with the reference vector is known as the longitude of the ascending node, Ω . The argument of periape (ω) is the angle that the eccentricity vector makes to the line of intersection. Finally, one must pick a value that defines the position of m_1 along its orbit. The three most common are the eccentric anomaly E , the mean anomaly M and the true anomaly θ - the same angle we have previously used between the vectors \mathbf{e} or \mathbf{B} and \mathbf{r} . Of these, the mean anomaly has the

advantage that it increases linearly with time, but notably does not represent an actual physical angle, and is rather just a measurement of how far along the orbit m_1 has travelled. With these elements we can completely define an arbitrarily-oriented two-body orbit in space.

This simple mathematical treatment of the two body problem has revealed many insightful facts about the orbit of a small body around a large one. When one considers multiple small bodies orbiting a larger one, as in multiple-planet systems, things get rather more complicated, and the resonant interactions between two small bodies can become important.

1.4.2 Resonant trapping

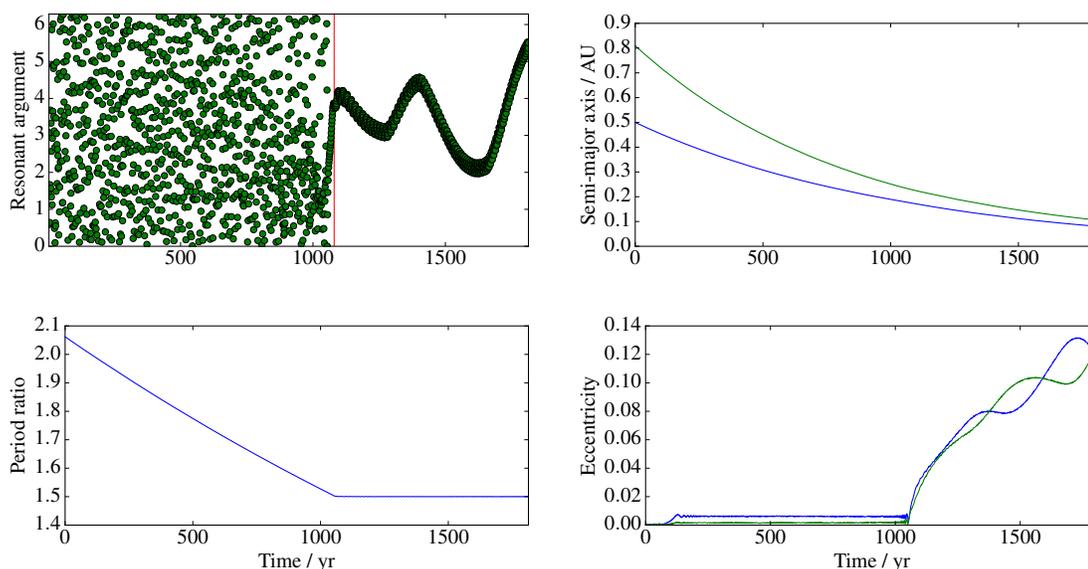


Figure 1.9: An example of resonant trapping. 2 planets are forced to migrate in an N -body simulation (see chapter 2) using the migration forces described in section 2.2.2. The outer planet migrates faster, causing the period ratio to drop from above 2 to 1.5. Once the period ratio is 1.5 at $T = 1080$ yr, the two planets are in the 3:2 resonance. At this point, the resonant argument (equation 1.73) begins to librate, whilst previously it was circulating. It continues to librate for the rest of the simulation, demonstrating that the two planets continue to be in resonance despite the constant migration force applied to them both. Once the planets are in resonance, their future evolution is linked and they migrate inwards together with a constant ratio of periods. The eccentricity of both grows quickly without any external damping.

An important dynamical effect when considering the evolution of multi-planet systems is that of Mean-Motion Resonances (MMRs). These resonances occur when the periods of two orbiting

bodies are commensurable, meaning that they can be described by a simple integer relationship. This condition is most simply stated as

$$\frac{n'}{n} = \frac{p}{p+q}, \quad (1.72)$$

where p and q are integers and n and n' are the mean-motions of the bodies. p and q should generally be small since for large p and q even the resonant interactions become essentially negligible. The mean motions are defined by the relation $n = 2\pi/T$, where T is the orbital period. Trivially, one can see that the ratio of mean-motions is the inverse of the ratio of periods. To understand why this behaviour might be important, consider the gravitational forces on a planet. With the planets out of resonance, the force they feel as they reach their closest approach to one another is incoherent with time and hence averages to zero. When in resonance, the interaction is coherent in time - occurring at the same set of points in each orbit - and hence sums up. This interaction not only drives eccentricity growth, but also acts to maintain the resonance. Therefore, planets that are captured in resonances will stay in them until an external perturber breaks them out, the interaction constantly driving eccentricity growth.

We now consider a more robust condition for detecting resonances. First, we define the resonant argument between two particles a and b as

$$\varphi = (p+q)\lambda_b - p\lambda_a - q\varpi_b, \quad (1.73)$$

where $\lambda = M + \varpi$ is the mean longitude, M is the mean anomaly, ϖ is the longitude of pericentre and p and q are integers (Murray & Dermott, 1999). Two particles a and b are considered to be in the $p : p+q$ mean-motion resonance if the resonant argument librates rather than circulates. If the argument is circulating, then over time it takes on every possible value between 0 and 2π . If it is librating, then it takes on only a small range of values and oscillates around some mean angle.

Figure 1.9 demonstrates how this quantity varies for 2 planets in and out of resonance, as well as the other effects that such a resonance has. A single resonance between a single pair of planets is relatively well-behaved, causing only oscillations in the eccentricities and semi-major axes. However, the overlap between two resonances for a single pair of planets can get extremely chaotic (Davies et al., 2014). Mean-motion resonances can then have the effect of destabilising a planetary system. Further

information and analytic treatment of resonances can be found in the textbooks [Armitage \(2010\)](#) and [Murray & Dermott \(1999\)](#), or the recent review article by [Davies et al. \(2014\)](#).

1.5 Migration of planets

Planets embedded in protoplanetary discs interact gravitationally with the surrounding gas, leading to an exchange in angular momentum and energy between the planet and the gas. Gas orbiting exterior to the planet robs the planet of its angular momentum, causing a reduction in the semi-major axis of the planet (see equation 1.69). Gas orbiting interior to the planet imparts angular momentum to the planet, pushing it outwards and increasing the semi-major axis of the orbit. It is thus the total balance of these two effects that determines the net change in angular momentum of the planet, and hence the direction in which it migrates. In general migration is expected to be inwards for canonical protoplanetary disc models, though the specifics depend very much on the temperature and density profiles of the disc, as well as the mass of the protoplanet. Migration has traditionally been split into two regimes - type I and II - based on the mass of the planet, with a regime known as type III only recently coming to light. I aim to provide an overview of the relevant physics in the following discussion, but the interested reader should refer to [Baruteau et al. \(2013\)](#) or [Kley & Nelson \(2012\)](#) for a more detailed review.

1.5.1 Type I

The type I migration regime occurs when the mass of a planet embedded in a disc is sufficiently low that the torques exerted by the planet on the disc do not have a significant effect upon the disc's structure. A first attempt at understanding the scale and nature of the type I migration torque can be made by considering the basic gravitational interaction between a pocket of gas passing a planet. This is referred to as the impulse approximation and comes from the work of [Lin & Papaloizou \(1979\)](#), who originally considered the case of a binary with a small mass ratio.

The impulse approximation

We first consider a parcel of gas of mass m that approaches the planet of mass M with some relative and parallel initial velocity v_{ini} , following the derivation in [Papaloizou & Terquem \(2006\)](#). The gas

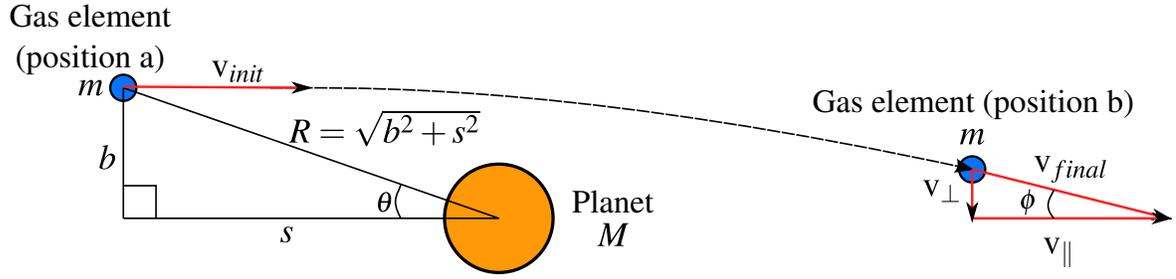


Figure 1.10: Schematic diagram showing variables used in the impulse approximation, in the frame of reference of the planet. Red arrows represent velocity vectors and the dashed line represents the trajectory of the gas as it moves from position a to position b.

approaches with impact parameter b and from initial distance s . The gravitational pull of the planet deflects the gas, giving it a perpendicular velocity component. We make the assumption that the deflection caused by the planet is small, so the final perpendicular velocity of the gas, v_{\perp} , obeys the relation $v_{\perp} \ll v_{\parallel}$, and consequently $v_{\parallel} \approx v_{final} \approx v_{init}$, where v_{final} is the total velocity of the gas parcel after the encounter. Figure 1.10 provides more details regarding the variables in this scenario.

To calculate v_{\perp} , we consider Newton's second law

$$F_{\perp} = m \frac{dv_{\perp}}{dt}, \quad (1.74)$$

or equivalently

$$v_{\perp} = \frac{1}{m} \int_{-\infty}^{\infty} F_{\perp} dt. \quad (1.75)$$

The assumption of a small angle of deflection, ϕ , means we can approximate that the pocket of gas continues travelling in a straight line with velocity v_{init} from $s = -\infty$ to $s = \infty$. This means that we can write

$$v_{init} dt = ds, \quad (1.76)$$

which allows us to rewrite equation 1.75 as

$$v_{\perp} = \frac{1}{m} \int_{-\infty}^{\infty} \frac{F_{\perp}}{v_{init}} ds. \quad (1.77)$$

Referring to figure 1.10, the force exerted on the gas parcel perpendicular to its direction of travel is

$$F_{\perp} = F \sin(\theta) = F \sin\left(\frac{b}{R}\right), \quad (1.78)$$

1.5. MIGRATION OF PLANETS

where F , the gravitational force along the line between the planet and the gas parcel, is given by

$$F = \frac{GMm}{R^2}, \quad (1.79)$$

and so

$$F_{\perp} = \frac{GMm}{R^2} \left(\frac{b}{R} \right). \quad (1.80)$$

Using this expression for the force and $R^2 = \sqrt{b^2 + s^2}$, equation 1.77 becomes

$$v_{\perp} = \int_{-\infty}^{\infty} \frac{GM}{(b^2 + s^2)^{3/2}} \left(\frac{b}{v_{init}} \right) ds, \quad (1.81)$$

which can be further simplified by removing a factor of b^2 from the denominator

$$v_{\perp} = \int_{-\infty}^{\infty} \frac{GM}{\left[b^2 \left(1 + \frac{s^2}{b^2} \right) \right]^{3/2}} \left(\frac{b}{v_{init}} \right) ds = \frac{GM}{b^2 v_{init}} \int_{-\infty}^{\infty} \frac{1}{\left(1 + \frac{s^2}{b^2} \right)^{3/2}} ds. \quad (1.82)$$

We solve this integral using the substitution $\tan(u) = s/b$ and therefore $ds = b \sec^2(u) du$:

$$v_{\perp} = \frac{GM}{b^2 v_{init}} \int_{-\pi/2}^{\pi/2} \frac{b \sec^2(u)}{(1 + \tan^2(u))^{3/2}} du = \frac{GM}{b v_{init}} \int_{-\pi/2}^{\pi/2} \frac{\sec^2(u)}{\sec^3(u)} du. \quad (1.83)$$

This leaves the trivially solvable integral

$$v_{\perp} = \frac{GM}{b v_{init}} \int_{-\pi/2}^{\pi/2} \cos(u) du = \frac{GM}{b v_{init}} [\sin(u)]_{-\pi/2}^{\pi/2} = \frac{2GM}{b v_{init}}, \quad (1.84)$$

and hence we have an expression for the perpendicular velocity of the gas parcel after the interaction.

Evidently, however, this velocity change is not the one that causes a change in the angular momentum.

Rather, it is the small change in parallel velocity that changes the angular momentum. To derive this,

we use the fact that kinetic energy is conserved during the interaction. We may therefore write

$$v_{init}^2 = v_{\perp}^2 + v_{\parallel}^2, \quad (1.85)$$

where $v_{\parallel} = v_{init} - \delta v_{\parallel}$ and δv_{\parallel} is the change in parallel velocity during the interaction. We can hence

write

$$v_{init}^2 = v_{\perp}^2 + (v_{init} - \delta v_{\parallel})^2, \quad (1.86)$$

which expands to

$$v_{init}^2 = v_{\perp}^2 + v_{init}^2 - 2v_{init}\delta v_{\parallel} + \delta v_{\parallel}^2. \quad (1.87)$$

We now rearrange this, neglecting the term δv_{\parallel}^2 since it is small, and substituting our expression for v_{\perp} from equation 1.84. This leaves us with

$$\delta v_{\parallel} = \frac{1}{2v_{init}} \left(\frac{2GM}{bv_{init}} \right)^2 \quad (1.88)$$

which is the change in the parallel velocity of the gas parcel during its interaction with the planet. At this point it is easy to derive the change in the specific angular momentum of the gas parcel. If the planet is on a circular orbit of radius a , then the change in specific angular momentum of the gas is $\delta j = a\delta v_{\parallel}$ or

$$\Delta j = a \frac{2G^2 M^2}{b^2 v_{init}^3}. \quad (1.89)$$

Evidently, by conservation of angular momentum, the planet experiences an equal and opposite change. We can derive the total torque on the planet by integrating equation 1.89 over the entire disc. First however, it is necessary to calculate the time-scale over which these impulses occur. If the angular frequency of the planet is Ω_p and the that of a gas parcel is Ω , then the impulses occur whenever a sufficient time Δt has elapsed that one has completed a full circle relative to the other:

$$|\Omega - \Omega_p| \Delta t = 2\pi. \quad (1.90)$$

We can use the first-order Taylor expansion of Ω_p around a to obtain

$$\Omega(b) = \Omega_p + \frac{d\Omega_p}{da} b, \quad (1.91)$$

which is valid for $b \ll a$. Assuming Keplerian orbits, $\Omega_p = \sqrt{GM_*/a^3}$ and $d\Omega_p/da = -3\Omega_p/(2a)$, so we may write

$$|\Omega - \Omega_p| = \frac{3\Omega_p}{2a} b, \quad (1.92)$$

which allows us to express the time between impulses as

$$\Delta t = \frac{4\pi a}{3\Omega_p b}. \quad (1.93)$$

We now consider these impulses taking place around an annulus of width db , whose mass is given by

$$dm = 2\pi \Sigma a \cdot db, \quad (1.94)$$

as in section 1.1.4. The time-rate of change in angular momentum of the planet due to the interaction with this annulus is the change in total angular momentum of the annulus, $\Delta j \cdot dm$, divided by the time between impulses, Δt . Integrating this over the entire disc gives us

$$\frac{dJ}{dt} = - \int \frac{\Delta j}{\Delta t} dm. \quad (1.95)$$

We now replace dm with equation 1.94, Δt with equation 1.93, and Δj with equation 1.89. Finally, we must replace the initial velocity difference, v_{init} , that appears in the Δj term. The difference in velocities in the case of Keplerian orbits may be written $v_{init} = (a + b)\Omega - a\Omega_p$. Continuing with the assumption of $b \ll a$, we may neglect the term in $b\Omega$, and substitute for $\Omega - \Omega_p$ using equation 1.92, which leaves $v_{init} = (3/2)\Omega_p b$.³ This allows us to integrate across all impact parameters, like so:

$$\frac{dJ}{dt} = - \int_0^\infty \frac{8G^2 M^2 \Sigma a}{9\Omega_p^2 b^4} db. \quad (1.96)$$

Unfortunately this integral diverges due to the assumption of an impact parameter of 0. Instead, we have to make some assumption about a minimum impact parameter, b_{min} . Doing this allows the integral to be solved trivially, leaving

$$\frac{dJ}{dt} = - \frac{8G^2 M^2 \Sigma a}{27\Omega_p^2 b_{min}^3}. \quad (1.97)$$

There are naturally several problems with this estimate - for instance, we have ignored the fact that the planet and gas are in constant rotation, instead assuming that they approach one another on straight lines. It is possible to account for this effect - see e.g., [Goldreich & Tremaine \(1980\)](#) - in which case one obtains a slightly modified estimate but with identical scaling. In spite of this, the impulse approximation reveals several interesting properties of disc-driven migration. For instance, the torque increases with the square of planet mass, such that higher mass planet migrate significantly faster. It also reveals that more massive discs drive faster migration. It is however, still an estimate, and only approximately recovers the actual type I torque calculated by more sophisticated techniques.

³This is the standard way of deriving the impulse approximation, but is not technically an actual first-order approximation to v_{init} . One may instead write the velocity of the planet as $v_p = \sqrt{GM_*/a}$. To expand this to first order for a small displacement b around a gives $v(a + b) = v_p + b(dv_p/da) = v_p - (1/2)\Omega_p b$, such that the velocity difference may be written $v_{init} = (1/2)\Omega_p b$. This only differs by a factor 3 from the other estimate, but causes a factor 27 difference in the final torque estimate, due to the fact that v_{init} is cubed in equation 1.89.

A more detailed description

For a more detailed understanding of the type I migration torque, one must pursue more complicated analyses. In a linearised analysis of the waves excited in a disc by a planet, [Tanaka et al. \(2002\)](#) (see also [Korycansky & Pollack, 1993](#)) showed that the type I migration torque scales as

$$\Gamma_0 = \frac{q^2}{h^2} \Sigma_p a^4 \Omega_p^4, \quad (1.98)$$

where q is the planet-to-star mass ratio, h is the aspect ratio of the disc, a is the semi-major axis of the planet and Ω_p is the local Keplerian orbital frequency at the location of the planet. This expression can be used to calculate an approximate timescale for type I migration

$$\tau_{mig} = \frac{a}{da/dt} = \frac{1}{2} \frac{h^2}{q} \frac{M_*}{\Sigma_p a^2} \Omega_p^{-1}, \quad (1.99)$$

where we see rather trivially from the dependence on q that the migration time-scale reduces in a linear fashion with planet mass. Comparing equation 1.98 to equation 1.97 reveals that the impulse approximation correctly predicted the scaling with planet mass and surface density.

Whilst this expression gives the main scaling relations for type I migration, the exact value of the torque on the planet involves yet more complexities. Indeed, the net torque depends on the interplay between two different types of torques: Lindblad and Corotation resonances, which I will discuss in more detail below.

Lindblad resonances

A potential rotating around the same central mass as a gas disc (or indeed, any other kind of disc, such as a stellar disc) drives oscillations in the disc at locations in the disc where the orbital frequency of the disc Ω_g relative to that of the potential (Ω_p) corresponds to some fraction of the epicyclic frequency in that part of the disc ([Goldreich & Tremaine, 1979, 1980](#)). The epicyclic frequency is defined as the frequency at which an element of gas which experiences a radial displacement will oscillate. Thus, if the element of gas experiences a driving frequency $\Omega_g - \Omega_p$ which is equivalent to some integer fraction $1/m$ of its epicyclic frequency κ_0 , it will oscillate radially around its orbit with increasing amplitude. This condition may be stated mathematically as

$$m(\Omega_g - \Omega_p) = \pm \kappa_0 \quad (1.100)$$

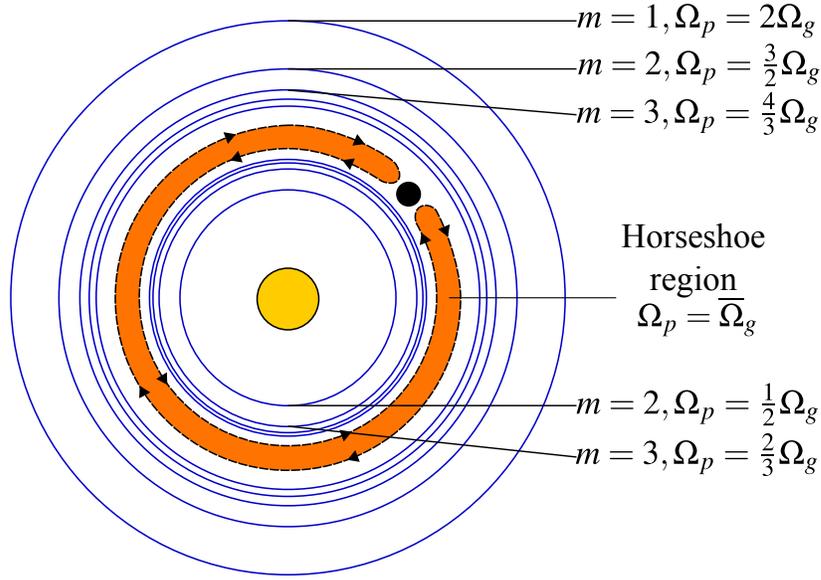


Figure 1.11: Schematic diagram showing the nominal locations of the Lindblad resonances (blue lines) and horseshoe region (orange shaded) relative to a planet (black circle) orbiting a star (yellow circle) in a protoplanetary disc. Material in the horseshoe region traces horseshoe orbits in the disc when viewed from the rotating frame of the planet, as demonstrated by the dashed line with arrows to show the direction of flow at each point.

The difference in signs on the right hand side of this equation corresponds to the 2 different cases; one in which the gas element is external to the planet ($\Omega_g < \Omega_p$, sign is negative) and one in which the gas element is internal to the orbit of the planet ($\Omega_g > \Omega_p$, sign is positive). For a Keplerian disc, the mathematics of the problem is simplified somewhat since $\kappa_0 = \Omega_g$. Figure 1.11 demonstrates the locations of the lowest m Lindblad resonances - note that for high m the comb of resonances piles up with each resonance being increasingly closer to the one before it. In fact, for realistic discs, there is a finite, set distance from this planet at which this pile up occurs. The distance from the planet at which this happens is set by the pressure scale-height of the disc, because for large m , equation 1.100 is no longer strictly valid. Instead, one must consider that the epicyclic frequency is also influenced by the pressure in the disc. The condition for resonance then becomes

$$m(\Omega_g - \Omega_p) = \pm \sqrt{\kappa_0^2(1 + \eta^2)} \quad (1.101)$$

where $\eta \equiv mc_s/(\Omega r)$ (see e.g. Artymowicz, 1993, for a more mathematically rigorous treatment). The Lindblad resonances now pile up at $r = a_p \pm 2H/3$, where a_p is the semi-major axis of the planet. This leads to a cut-off in the torque, which stops it from diverging near to the planet.

It is perhaps easier to elucidate the physical origin of this oscillation by considering the relative motion of the planet and an element of gas. Let us consider the $m = 1$ case - note that the $m = 1$ case only exists in gas exterior to the orbit of the planet. In this case, a given element of gas completes one orbit for every two orbits that the planet completes. So, if we consider that the planet and the gas element experience their first conjunction (point of closest approach) at an angle θ , their second conjunction will occur at the same angle, exactly one orbit later for the pocket of gas. Since the conjunction is where the gas feels the strongest force from the planet, the gas will be displaced radially at this point. Exactly one orbit later, the gas will again be strongly radially displaced. However, since the epicyclic frequency of the gas is equal to its orbital frequency, it will have already completed a full radial oscillation and returned to the radial displacement caused by the first conjunction. Thus at this second conjunction, the gas finds itself displaced from its original radius by a distance that is equal to the displacement from the first conjunction, plus an extra displacement from the second conjunction. We clearly see that this process will continue and the displacement will build until some equilibrium is reached, and that over time the planet will continuously drive a wave from this location. We can perform the same thought experiment in the $m = 2$ case: the planet completes exactly three orbits for every two orbits completed by the gas. A conjunction between the two, therefore, happens every time the gas has completed exactly two orbits. The gas therefore feels a driving force that oscillates with exactly half of its epicyclic frequency, and the radial displacement builds up at successive conjunctions. Note that the preceding description is somewhat simplified for the sake of clarity, but communicates the main aspects of the process.

The waves launched in these locations superpose, leading to the formation of a one-armed spiral density wave. The back-reaction of this spiral-density wave on the planet is what causes the so-called “Lindblad torque”, or “wave torque”. The resulting torque was first calculated analytically in the linear approximation (e.g., [Goldreich & Tremaine, 1979](#)), with later authors such as [Tanaka et al. \(2002\)](#) integrating the wave equations numerically. The latter publication led to a widely-used analytic approximation for the torques that is unfortunately only valid in the isothermal regime, but nevertheless gives a torque formula that relies only on the quantities Γ_0 (equation [1.98](#)), α - the negative of the local surface density power-law exponent in the disc - and β - the negative of the local temperature

power-law exponent in the disc.

$$\Gamma_L/\Gamma_0 = -3.2 - 1.468\alpha. \quad (1.102)$$

[Paardekooper et al. \(2010\)](#) derived a similar expression for the Lindblad torque in the adiabatic case:

$$\gamma\Gamma_L/\Gamma_0 = -2.5 - 1.7\beta + 0.1\alpha, \quad (1.103)$$

where γ is the adiabatic index. In general, we expect both surface density and temperature to decrease with orbital radius in canonical discs, meaning that α and β are both positive. Therefore, it is clear from these formulae that the Lindblad torques generally enact a negative torque and promote inward migration.

Corotation torques

The corotation torque, otherwise known as “horseshoe drag”, occurs as a result of the interaction between the planet and gas that is, on average, orbiting with the same period as the planet. In the frame of reference that rotates with the planet, this material appears to execute closed, horseshoe-shaped orbits through the disc. It is worth noting that the terms “linear corotation torque” and “horseshoe drag” actually refer to two subtly different effects: the former was discovered earlier (e.g., [Goldreich & Tremaine, 1979](#)) and is somewhat simpler, whilst the latter is a non-linear approximation derived by [Ward \(1991\)](#) and directly considers the motion of corotating material in the frame of reference of the planet. [Paardekooper & Papaloizou \(2009\)](#) found that the linear approximation is valid for the case of a low mass planet in a highly-viscous disc, but otherwise the horseshoe drag should be used.

To understand the nature of the horseshoe drag, it is necessary to consider the path that a parcel of corotating gas traces in the disc from the planet’s frame of reference. When the material is orbiting external to the planet, its azimuthal velocity is lower than that of the planet and therefore the planet sees the material coming towards it. When the material is sufficiently close to the planet to be perturbed by the planet’s gravitational potential, it is pulled inwards towards the planet, losing angular momentum and falling into a lower, faster orbit. On this faster orbit the material is travelling faster than the planet and therefore appears to the planet to be moving away, as though it has performed a U-turn. Over the course of many more orbits (since the orbital frequencies of the material and planet are only minutely different), the material loops around and reaches the other side of the planet and the planet sees the

material coming towards it once more. This time the perturbation from the planet's potential draws the material outwards, into a higher angular momentum orbit which is slower. The material then begins to move away from the planet again. We see in this case that the material never actually passes the planet in either direction, but rather draws horseshoe-shapes with the planet as the gap in the horseshoe.

One might naively expect that the 2 interactions - one where the planet loses angular momentum to the gas and another where the planet gains angular momentum from the gas - cancel each other out and therefore over time there is no net torque. However, the important thing to remember is that there is a gradient in both temperature and density in the disc, which leads to the existence of both a vortensity⁴- and an entropy-related horseshoe drag. In the case of an adiabatic disc there is therefore a change in local entropy as the gas moves inward or outward, but the total entropy is conserved. This change in entropy is what results in the net torque on the planet. In the isothermal case, it is the change in local vortensity but conservation of total vortensity that drives the torque (Casoli & Masset, 2009).

The region in which these orbits are possible is known as the horseshoe region, and naturally is rather narrow due to the small sphere of influence of the planet's gravity relative to its stellar host. Thus there is a finite amount of gas and hence a finite amount of angular momentum that may be transferred to or from the planet in this region. The corotation torque may therefore become saturated once the limited angular momentum available has been transferred to the planet. In this case, the corotation torque becomes negligible and only the linear Lindblad torque remains (Paardekooper & Papaloizou, 2009). Naturally the way to avoid this saturation is to have a sufficiently viscous disc that angular momentum is transported into or out of this region fast enough to replenish it. Additionally, numerical simulations by Fendyke & Nelson (2014) suggest that the magnitude of the corotation torque is reduced exponentially with increasing planetary eccentricity.

Paardekooper et al. (2010) derived the following expression for the non-linear horseshoe drag from numerical simulations and theoretical arguments:

$$\gamma\Gamma_{c,HS}/\Gamma_0 = 1.1\left(\frac{3}{2} - \alpha\right) + 7.9\frac{\xi}{\gamma} \quad (1.104)$$

where $\xi = \beta - (\gamma - 1)\alpha$ is the negative of the exponent of the local power law of the specific entropy. This entropy should decrease outwardly in the innermost regions of discs, and thus the horseshoe drag may promote outward migration in these cases.

⁴Vortensity is the curl of the velocity field (vorticity) divided by the local surface density

Total torque

In summary, one can consider the two types of torques to be a result of resonances that are driven by a forcing frequency and act on the two natural frequencies of the gas, κ (Lindblad) and 0 (Corotation) (Binney & Tremaine, 1987). With an understanding of the regimes in which the two torques operate, the natural conclusion is to combine them to obtain a formula for the total torque, scaling with Γ_0 .

D'Angelo & Lubow (2010) attempted to derive such a torque formula in the locally isothermal case using three dimensional discs, and found

$$\Gamma_{tot}/\Gamma_0 = -1.4 - 0.4\beta - 0.6\alpha. \quad (1.105)$$

For the adiabatic case, one can simply combine equations 1.103 and 1.104, but must always be wary of the caveats of such formalisms. The thermal and viscous structure in the disc depends very strongly on which effects are included in the heating of the disc. Bitsch et al. (2013a, 2014b,a) have studied these various effects in great detail using 2D simulations including stellar irradiation, viscous heating and radiative cooling, finding that at various points in the disc lifetime, super-Earth and Neptune mass planets can experience periods of outward migration in discs with sufficiently high metallicity.

In the limit of low eccentricity, the torques causing type-I migration can also damp the eccentricity in inclination of planets. Tanaka & Ward (2004) showed that for planets in an isothermal disc, the eccentricity damping time-scale is shorter than the migration timescale by a factor equal to the square of the disc aspect ratio, or $\tau_{ecc} \approx (H/r)^2 \tau_{mig}$. Cresswell et al. (2007) found using hydrodynamical simulations that eccentricity damping is exponential for small e , but that for $e > 0.1$, $\dot{e} \propto -e^{-2}$. They also found that small non-zero eccentricities tended to increase the migration rate of the planet, whilst larger eccentricities tended to reduce it.

1.5.2 Type II

The assumption underpinning the type I regime was that the protoplanet was not sufficiently massive to significantly alter the structure of its host disc - beyond launching spiral density waves. In the type II regime, the mass of the protoplanet is large enough that the planet can overcome the pressure and viscous forces in the disc and clear material out of the region of its orbit, opening a gap. Crida et al. (2006) used a combination of analytic arguments and numerical results to derive the following gap

opening criterion:

$$P = \frac{3}{4} \frac{H}{r_H} + \frac{50}{q \text{Re}} \lesssim 1 \quad (1.106)$$

where $\text{Re} = r_p^2 \Omega_p / \nu$ is the Reynolds number, $r_H = r_p (q/3)^{1/3}$ is the Hill radius of the planet and q is the planet-to-star mass ratio. If the disc viscosity is based on the aforementioned α prescription (equation 1.41), then we can use $\nu = \alpha c_s H$ and the aforementioned definitions of r_H and Re to rewrite this criterion in terms of the disc aspect ratio $h = H/R$, the planet-to-star mass ratio q , and α . Substituting the expressions for r_H and Re into equation 1.106 gives us

$$P = \frac{3}{4} \frac{H}{r_p} \left(\frac{3}{q} \right)^{1/3} + \frac{50}{q} \frac{\nu}{r_p^2 \Omega_p} \lesssim 1, \quad (1.107)$$

which, using the definition of α , becomes

$$P = \frac{3}{4} \frac{H}{r_p} \left(\frac{3}{q} \right)^{1/3} + \frac{50}{q} \frac{\alpha c_s H}{r_p^2 \Omega_p} \lesssim 1. \quad (1.108)$$

Realising that $c_s = H \Omega_p$, we can eliminate the sound speed from the second term, leaving

$$P = \frac{3}{4} \frac{H}{r_p} \left(\frac{3}{q} \right)^{1/3} + \frac{50}{q} \frac{\alpha H^2}{r_p^2} \lesssim 1. \quad (1.109)$$

Finally, using the disc aspect ratio $h = H/R$, and neglecting the order unity constant $3^{4/3}/4$, the criterion becomes

$$P = \frac{h}{q^{1/3}} + \frac{50 \alpha h^2}{q} \lesssim 1. \quad (1.110)$$

Hydrodynamic simulations by [Duffell & MacFadyen \(2013\)](#) suggest that this criterion ceases to be valid for extremely low values of disc α , with planets of only a few M_\oplus opening gaps for $\alpha \lesssim 10^{-4}$.

With a large gap opened in the disc, the corotation torque becomes insignificant and the Lindblad resonances nearest to the planet are also greatly reduced. Thus, the torque formulae previously discussed are no longer valid, and one finds that the planet actually migrates in a way that is coupled to the viscous evolution of the disc, with a time-scale $\tau_\nu = a_p^2 / \nu$ ([Lin & Papaloizou, 1986](#)), much slower than the typical type I migration time-scale. Numerical simulations such as those by [Nelson et al. \(2000\)](#) have found slow, inward migration as expected, with planets capable of accreting a significant amount of mass during this time. Sufficiently large planets can, however, completely cut off the inner disc from the outer disc, causing the inner disc to drain onto the star and modifying the migration rate

further. Finally, planets which open partial gaps (defined here as $P > 1.5$) may still be surrounded by enough coorbital gas that they experience significant corotation torques, which can lead to a slowing or reversing of type II migration [Crida & Morbidelli \(2007\)](#).

1.5.3 Type III

Type III migration is a special case for planets that open only a partial gap in their host disc, leading to a runaway migration process that takes place over tens of orbits. This runaway process is caused by the fact that the net torque experienced by the planet is proportional to the migration rate, such that the two quantities grow together. The first full analysis of this effect was performed by [Masset & Papaloizou \(2003\)](#). To understand it, we must consider that there are in fact three distinct torques caused by horseshoe material, rather than just one. The one that we already considered is that caused by material that simply corotates with the planet, executing horseshoe turns and then encountering the planet again later on. What we did not consider is that the planet drags this material with it as it migrates inwards. This effect exerts a positive torque on the planet, slowing migration. We also did not consider that there exists material which is on the separatrix between the horseshoe and non-horseshoe regions. This material meets the planet once when it is in the inner disc, performs a u-turn, and is propelled into the outer disc. The planet never encounters it again. Evidently, the planet sees more of this material if it migrates faster, causing it to experience a greater inward torque and encounter yet more such material. However, the planet also drags horseshoe material with it at a faster rate if it migrates faster, and for low mass planets, these two effects cancel out. However, if the planet opens a partial gap, the mass deficit δm of the disc in the region of the planet may satisfy $\delta m \approx m'_p$ where m'_p is the mass of the planet and the material it drags along. In this case, the material passing by the planet once on its journey from inner to outer disc exerts a greater torque than the horseshoe material that the planet drags along, the migration accelerates, the torque from this effect increases and the process runs away. This runaway effect is expected to happen for planets of approximately Saturn mass, where the gap opening criterion $P \approx$ a few.

1.6 Summary

Throughout this chapter, I have provided a review of our current understanding of planet formation and migration in protoplanetary discs, from both a theoretical and observational perspective. In particular, I've highlighted how low-mass planets migrate in the type I regime, and the recent explosion in data surrounding tightly-packed systems of super-Earths. There is a difficulty in explaining how these systems can form and remain dynamically cold, given the interplay of type I migration and resonant trapping. The complex mixture of both classical and fluid dynamics that governs these systems makes them challenging to explore using analytic methods, and therefore we must turn to numerical methods to gain a further insight. In the next chapter I will discuss the numerical methods that I will use throughout the rest of this thesis to explore these problems.

Chapter 2

Numerical methods

During the discussion of planet formation, migration and evolution in the previous chapter, gravitational interactions and fluid mechanics were consistently the most important types of physics. Fortunately, mature numerical techniques to simulate both of these already exist, and can be combined to model type I migration in detail. In this chapter I will summarise the current state-of-the-art methods for modelling both, and describe how I have adapted them for use in the context of type I migration.

2.1 *N*-body

The term *N*-body can refer to many different types of simulations in the astrophysical context, and generally refers to any method which attempts to solve for the mutual interaction of multiple bodies. In general though, the term *N*-body refers to simulations of particles that are interacting gravitationally, and such methods attempt to solve the system of equations given by equation 1.45. Gravitational *N*-body methods are classified in one of two categories; “collisional” simulations, whereby the close interactions of individual bodies matter and are modelled as accurately as possible, and “collisionless” systems, whereby these interactions are considered unimportant and the bodies in the simulation represent the physical system in a statistical sense. For a more technical distinction between these two classifications, we turn to the so-called two-body relaxation time, which defines the time-scale on which the equipartition of kinetic energy via gravitational close encounters occurs. This timescale can

be written approximately as [Dehnen & Read \(2011\)](#)

$$t_{relax} \approx \frac{N}{8 \ln(\Lambda)} t_{dyn}, \quad (2.1)$$

where t_{dyn} is the dynamical time, N is the number of particles and $\Lambda = b_{max}/b_{min}$. b in this case refers to the impact parameter. Collisional simulations deal with phenomena which take place over multiples of the two-body relaxation time-scale of the system, whilst collisionless systems deal with phenomena which take place on less than this time-scale. Collisional N -body dynamics therefore naturally lends itself better to simulations of planetary systems, since N is low, t_{dyn} (in this case, the orbital time-scale) is of the order years and the lifetime of the system is of order gigayears, making it straightforward to see from this equation that the average planetary system will have evolved through many multiples of the two-body relaxation time. Collisionless methods are better suited to simulations of galaxies and clusters thereof, where number densities of particles are sufficiently low and dynamical timescales are sufficiently long that the redistribution of kinetic energy does not happen on reasonable time-scales. Thus for the remainder of this thesis, I shall write exclusively about collisional N -body methods.

2.1.1 Time integration

The natural way to proceed in solving the system of equations generated by equation 1.45 is to integrate them forwards in time, where we know the position $\mathbf{x}(t_0)$ and velocity $\mathbf{v}(t_0) \equiv \dot{\mathbf{r}}(t_0)$ of each particle, and wish to calculate the effect that the force $m\mathbf{a}(t_0) \equiv m\ddot{\mathbf{r}}(t_0)$ has on these variables as a function of time. The quickest way to make any sort of progress in this endeavour is to compute the Taylor expansion of \mathbf{x} and $\dot{\mathbf{x}}$ around t_0 , which will give us an understanding of how these variables vary in time for a small region around t_0 . Thus for an arbitrary time t in this region, we obtain, to second-order:

$$\mathbf{x}(t) = \mathbf{x}(t_0) + \mathbf{v}(t_0)(t - t_0) + \frac{\mathbf{a}(t_0)}{2}(t - t_0)^2 + \mathcal{O}((t - t_0)^3) \quad (2.2)$$

and

$$\mathbf{v}(t) = \mathbf{v}(t_0) + \mathbf{a}(t_0)(t - t_0) + \frac{\dot{\mathbf{a}}(t_0)}{2}(t - t_0)^2 + \mathcal{O}((t - t_0)^3) \quad (2.3)$$

If we now redefine our arbitrary time t to be the original time t_0 plus some small advance called the time-step Δt , this becomes

$$\mathbf{x}(t_0 + \Delta t) = \mathbf{x}(t_0) + \mathbf{v}(t_0)\Delta t + \frac{\mathbf{a}(t_0)}{2}\Delta t^2 + \mathcal{O}(\Delta t^3) \quad (2.4)$$

and

$$\mathbf{v}(t_0 + \Delta t) = \mathbf{v}(t_0) + \mathbf{a}(t_0)\Delta t + \frac{\dot{\mathbf{a}}(t_0)}{2}\Delta t^2 + \mathcal{O}(\Delta t^3) \quad (2.5)$$

It thus becomes clear that we can solve the *N*-body problem approximately and integrate $\mathbf{x}(t_0)$ and $\mathbf{v}(t_0)$ forward in time to i^{th} order through these expansions, though this requires knowledge of increasingly complicated higher-order derivatives of equation 1.45, which naturally become numerically prohibitive to calculate. We thus surmise that the absolute simplest scheme we can devise comes from taking these expansions to first order such that *no* derivatives of the acceleration are required. Doing this, we obtain the Euler scheme:

$$\mathbf{x}(t_0 + \Delta t) = \mathbf{x}(t_0) + \mathbf{v}(t_0)\Delta t + \mathcal{O}(\Delta t^2) \quad (2.6a)$$

$$\mathbf{v}(t_0 + \Delta t) = \mathbf{v}(t_0) + \mathbf{a}(t_0)\Delta t + \mathcal{O}(\Delta t^2) \quad (2.6b)$$

which is evidently accurate to first order, with the errors being proportional to the the square of the step-size, Δt^2 . Of course, errors of the order Δt^2 aren't necessarily all that small, and thus we can intuit that this scheme performs rather poorly unless we choose a very short step size. Reducing the step size means more calculations of the force, which requires the calculation of the magnitude of the vector \mathbf{r}_{ij} , which in turn requires the calculation of an inverse square-root. Inverse square roots are notoriously computationally expensive, and thus we would like to improve the order of our integration and allow larger step sizes without incurring too many more of these operations.

The Leapfrog integrator

The trick here comes from realising that equation 2.5 can be rewritten using the acceleration expanded to second order. Consider this expansion:

$$\mathbf{a}(t_0 + \Delta t) = \mathbf{a}(t_0) + \dot{\mathbf{a}}(t_0)\Delta t + \frac{\ddot{\mathbf{a}}(t_0)}{2}\Delta t^2 + \mathcal{O}(\Delta t^3) \quad (2.7)$$

Rearranging gives us

$$\dot{\mathbf{a}}(t_0)\Delta t = \mathbf{a}(t_0 + \Delta t) - \mathbf{a}(t_0) - \frac{\ddot{\mathbf{a}}(t_0)}{2}\Delta t^2 + \mathcal{O}(\Delta t^3) \quad (2.8)$$

which, substituted into equation 2.5, gives us

$$\mathbf{v}(t_0 + \Delta t) = \mathbf{v}(t_0) + \mathbf{a}(t_0)\Delta t + \frac{\Delta t}{2} \left(\mathbf{a}(t_0 + \Delta t) - \mathbf{a}(t_0) - \frac{\ddot{\mathbf{a}}(t_0)}{2}\Delta t^2 \right) \quad (2.9)$$

where I have momentarily neglected the O terms for the sake of clarity. Expanding out the bracketed term gives

$$\mathbf{v}(t_0 + \Delta t) = \mathbf{v}(t_0) + \frac{\Delta t}{2} (\mathbf{a}(t_0) + \mathbf{a}(t_0 + \Delta t)) - \frac{\ddot{\mathbf{a}}(t_0)}{2} \Delta t^3 \quad (2.10)$$

Here we have a term in Δt^3 which is evidently not the correct 3rd order term for the expansion of the velocities. The correct term would rather be $+\ddot{\mathbf{a}}(t_0)/6 \cdot \Delta t^3$. A simple consideration of the neglected $O(\Delta t^3)$ terms in equation 2.7 shows that if we had also substituted each of these neglected terms into equation 2.5, they would also be incorrect. Thus the equation

$$\mathbf{v}(t_0 + \Delta t) = \mathbf{v}(t_0) + \frac{\Delta t}{2} (\mathbf{a}(t_0) + \mathbf{a}(t_0 + \Delta t)) \quad (2.11)$$

is equivalent to the 2nd order expansion of velocity in equation 2.5, with erroneous terms of order Δt^3 ¹. Thus it becomes clear that we can obtain second-order accurate velocities without calculating the $\dot{\mathbf{a}}(t_0)$ term, but that in order to do this we need to already know the acceleration at the next time-step. Fortunately, equation 2.4 proves that in order to gain a second-order accurate position at the next time-step, we do not require the velocity from the next time-step. Since for the N -body problem the force relies only on the position and not the velocity, this knowledge is sufficient to calculate the term $\mathbf{a}(t_0 + \Delta t)$. Equation 2.4 also confirms that the second-order change to the position relies on $\mathbf{a}(t_0)/2$ in the same way that our new equation for the second-order change to velocity does. We can therefore split equations 2.4 and 2.11 like so:

$$\mathbf{v}' = \mathbf{v}(t_0) + \frac{\Delta t}{2} \mathbf{a}(t_0) \quad (2.12a)$$

$$\mathbf{x}(t_0 + \Delta t) = \mathbf{x}(t_0) + \mathbf{v}' \Delta t \quad (2.12b)$$

$$\mathbf{v}(t_0 + \Delta t) = \mathbf{v}' + \frac{\Delta t}{2} \mathbf{a}(t_0 + \Delta t) \quad (2.12c)$$

where we have defined \mathbf{v}' purely as a temporary velocity. Substituting for \mathbf{v}' confirms that we retrieve equations 2.4 and 2.11, and therefore that executing the three equations 2.12 in the order they are written will solve the N -body problem with second-order accuracy. This scheme is referred to as the kick-drift-kick leapfrog integrator, since the first and third steps change only the velocity of the

¹This can be confirmed by substituting 2.7 back into the equation in as many terms as one cares to deal with. The result is the recovery of equation 2.5.

particles (thus they kick the particle), and the second step causes the particle to “drift” with its present velocity.

The name of the integrator comes from the fact that one can achieve improved numerical efficiency by concatenating the two kick operations between the drift operations, such that the entire scheme is written as

$$\mathbf{x}(t_0 + \Delta t) = \mathbf{x}(t_0) + \mathbf{v}(t_0 - \Delta t/2)\Delta t \quad (2.13a)$$

$$\mathbf{v}(t_0 + \Delta t/2) = \mathbf{v}(t_0 - \Delta t/2) + \mathbf{a}(t_0 + \Delta t)\Delta t \quad (2.13b)$$

The previously intermediate velocity \mathbf{v}' has now become $\mathbf{v}(t_0 + \Delta t/2)$ - the only velocity we have any knowledge of. We calculate this velocity at the halfway point in every timestep, yet we know the positions at the beginning and end of every time-step. Thus the positions and velocities “leapfrog” over one another. It is evident from the form of the above equations that this half-time-step velocity is only accurate to first order, and thus the previous form of the leapfrog from equations 2.12 is preferred for scenarios where accurate output of velocities is required ². This also becomes important for adding velocity dependent forces, where the values of such forces are only accurate to second-order when using the “full-step” velocities from equations 2.12. Section 2.2.2 covers this in more detail.

The integrator derived above is *symplectic*, which means that it provides a canonical map that maps one set of phase-space coordinates directly to another. This process preserves the phase-space volume of a system and has a number of features that make it useful for integrations of planetary systems. The primary advantage is that this integrator exactly solves a Hamiltonian $\tilde{H} = H + H_{err}$, where H is the original Hamiltonian and H_{err} is an “error” Hamiltonian. What this means in practice is that whilst there is some deviation from the initial energy of the system, there is no secular growth in the energy error over time, and instead the energy oscillates a small amount around its initial value. This is evidently useful in planetary dynamics where the vast majority of forces within the system are conservative. Note however that the symplectic nature of the map is not preserved when not using a constant Δt (or rather more obviously, non-conservative forces), though by maintaining time-symmetry, this error can be reduced (Dehnen & Read, 2011). I discuss the adaptive time-stepping

²To maintain second order accuracy with this second version of the leapfrog, one must perform an extra kick using the first-order Taylor expansion of the velocity ($\mathbf{v}(0 + \Delta t/2) = \mathbf{v}(0) + \mathbf{a}(0)\Delta t/2$) to calculate the first intermediate velocity before beginning the scheme.

used for the calculations in this thesis in section 2.2.1. I defer the interested reader to [Yoshida \(1993\)](#) for a more mathematically rigorous derivation of the integrator that proves these properties using Lie algebra. The Leapfrog also implicitly conserves angular momentum, which is again very useful for integrating planetary orbits. As a result of these properties and the numerical efficiency of the leapfrog (equations 2.12 require only that the forces be computed once per time-step to achieve second order accuracy), I use the leapfrog and related integrators for the majority *N*-body integrations within this thesis.

Runge-Kutta methods

For the integrations in this thesis where a symplectic integrator is not directly beneficial, or higher order integration is required, I use one of the family of so-called Runge-Kutta methods (see e.g., [Press et al., 1992](#)). Understanding how these methods work is best done by way of example. For the sake of clarity, we shall move away from *N*-body systems momentarily to a simpler, first-order differential equation

$$\dot{y} = f(t, y). \tag{2.14}$$

If we know the initial value $y(t_0)$, we might hope to find the future values of y using the Euler method (equations 2.6), which for this ODE would be

$$y(t_0 + \Delta t) = y(t_0) + f(t_0, y(t_0))\Delta t. \tag{2.15}$$

We already know that this is accurate to first order and that the error will be of order $O(\Delta t^2)$. Rather than using Taylor expansions to gain an insight on how to improve this (as with the Leapfrog method above), we might instead attempt to use a better estimate for the average time derivative of y across this entire step. We could use the initial gradient of y , which we know from our initial conditions, to estimate \dot{y} halfway through the step, and then use this new value of \dot{y} to advance the system forward a whole time-step. Intuitively, this appears to be better than simply assuming that the gradient of y remains the same across the entire step. Implementing this method gives us the following system of equations:

$$k_1 = f(t, y), \tag{2.16a}$$

$$k_2 = f\left(t + \frac{\Delta t}{2}, y + k_1 \frac{\Delta t}{2}\right), \quad (2.16b)$$

$$y(t_0 + \Delta t) = y(t_0) + k_2 \Delta t. \quad (2.16c)$$

Here, k_1 is simply the initial time-derivative of y that we know from our initial conditions, and k_2 is our estimate of the time-derivative halfway across the time-step. This method is second order accurate, though it does not have the same conservation properties as the Leapfrog.

This scheme may be extended to fourth-order accuracy using two extra estimations of \dot{y} . The first, named k_3 , is another estimate of \dot{y} at the midpoint of the interval, but this time using k_2 rather than k_1 to estimate y in the middle of the interval. Finally, we use k_3 to estimate \dot{y} right at the end of the interval. We then evolve $y(t_0)$ to $y(t_0 + \Delta t)$ using an average of these various gradient estimates, yielding the following system of equations:

$$k_1 = f(t, y), \quad (2.17a)$$

$$k_2 = f\left(t + \frac{\Delta t}{2}, y + k_1 \frac{\Delta t}{2}\right), \quad (2.17b)$$

$$k_3 = f\left(t + \frac{\Delta t}{2}, y + k_2 \frac{\Delta t}{2}\right), \quad (2.17c)$$

$$k_4 = f(t + \Delta t, y + k_3 \Delta t), \quad (2.17d)$$

$$y(t_0 + \Delta t) = y(t_0) + \frac{\Delta t}{6}(k_1 + 2(k_2 + k_3) + k_4). \quad (2.17e)$$

This integrator is the fourth-order Runge-Kutta, as presented in [Press et al. \(1992\)](#), and solves a single first-order ordinary differential equation. Evidently, to use this integrator in N -body context, we need to solve a second-order differential equation

$$\ddot{\mathbf{r}} = \mathbf{a}(t, \mathbf{r}), \quad (2.18)$$

rather than a first-order one. We may, however, split this second-order ODE into two first-order differential equations, these being

$$\dot{\mathbf{r}} = \mathbf{v}, \quad (2.19a)$$

$$\dot{\mathbf{v}} = \mathbf{a}. \quad (2.19b)$$

2.1. N-BODY

We now use the same fourth-order scheme from above to solve these equations simultaneously. The first step remains as before - to use the initial gradient in the position and velocity:

$$\mathbf{s}_1 = \mathbf{v}(t), \tag{2.20a}$$

$$\mathbf{f}_1 = \mathbf{a}(\mathbf{r}), \tag{2.20b}$$

in order to estimate these same gradients at the midpoint of the step:

$$\mathbf{s}_2 = \mathbf{v}(t) + \mathbf{f}_1 \frac{\Delta t}{2}, \tag{2.21a}$$

$$\mathbf{f}_2 = \mathbf{a}(\mathbf{r} + \mathbf{s}_1 \Delta t/2). \tag{2.21b}$$

We then use these new gradients to obtain another estimate of the gradients at the midpoint:

$$\mathbf{s}_3 = \mathbf{v}(t) + \mathbf{f}_2 \frac{\Delta t}{2}, \tag{2.22a}$$

$$\mathbf{f}_3 = \mathbf{a}(\mathbf{r} + \mathbf{s}_2 \Delta t/2), \tag{2.22b}$$

at which point we use these latest gradient estimates to estimate the gradients at the end of the step:

$$\mathbf{s}_4 = \mathbf{v}(t) + \mathbf{f}_3 \Delta t, \tag{2.23a}$$

$$\mathbf{f}_4 = \mathbf{a}(\mathbf{r} + \mathbf{s}_3 \Delta t). \tag{2.23b}$$

Finally, we combine our gradient estimates as in equation 2.17e to obtain our final estimate for the positions and velocities of each particle at the end of the time-step:

$$\mathbf{r}(t_0 + \Delta t) = \mathbf{r}(t_0) + \frac{\Delta t}{6}(\mathbf{s}_1 + 2(\mathbf{s}_2 + \mathbf{s}_3) + \mathbf{s}_4), \tag{2.24a}$$

$$\mathbf{v}(t_0 + \Delta t) = \mathbf{v}(t_0) + \frac{\Delta t}{6}(\mathbf{f}_1 + 2(\mathbf{f}_2 + \mathbf{f}_3) + \mathbf{f}_4). \tag{2.24b}$$

Note that we could also follow a similar series of steps to adapt the second-order Runge-Kutta method (equations 2.16) to second-order ODEs.

This new method is fourth-order accurate, which may seem preferable to the second-order accuracy of the Leapfrog. One must remember, however, that each time we calculate the gradient in velocity, \mathbf{f} , we have to perform a force calculation, which involves the expensive square root operation. So the system of equations presented above requires four force calculations, whilst the Leapfrog

requires only one. We must be careful then, in selecting the optimal method for a given problem. In the instance where our system is conservative and completely Hamiltonian, the symplectic nature of the Leapfrog and the conservation properties it offers may make it preferable, in spite of its larger error.

2.2 *N*-body code

For the purposes of the work presented in the following chapters, I developed my own *N*-body code to integrate systems with small N (< 100). Named PENGUIM (Planets Evolving with *N*-body Gravity Under the Influence of Migration), this code was built in a modular fashion to reduce the complexity of adding new physics to the simulation. The code was developed in C++ and is not yet publically available. It contains implementations of the 2nd order Leapfrog integrator and the 4th order Runge-Kutta integrator, as well as an option for both fixed and adaptive (described below) time-stepping.

2.2.1 *N*-body time-stepping

Time stepping is handled adaptively in PENGUIM. Whilst this violates the symplectic nature of the leapfrog, so do any non-conservative forces applied to the bodies, so preserving the symplectic nature is impossible when using the forces described in section 2.2.2. The use of adaptive time-stepping is important since in a range of randomly-sampled initial conditions, the behaviour of each model is unpredictable and hence a constant, short time-step is not suitable to guarantee both computational efficiency and accurate time integration. I note that longer timesteps would in principle be possible with a higher order integrator, though the computational simplicity of the leapfrog and the ease with which extra, velocity-dependent forces may be added makes it the integrator of choice here.

The adopted time-stepping scheme is the so called “block-step” scheme, whereby the time-step of the *N*-body code is required to be some integer factor of a longer block-step. This system has advantages where it is important for the integrator to work and update on a certain time-scale, which is crucial for interactions with other parts of our numerical method which will be described later. The *N*-body integrator is hence allowed to select from a range of timesteps in accordance with the relation

$$\tau_{grav} = \frac{1}{2^n} \tau_{block} \quad (2.25)$$

where the integer factor n is referred to as the 'rung' of the timestepping scheme. This rung defines the length of one time-step and is selected in order to maximise the accuracy and minimise the computational requirements of the simulation. The choice of rung is made according to the criterion

$$\tau_{optimal} = \min\left(\epsilon \frac{\sqrt{|\Phi_i|}}{|\mathbf{a}_i|}\right). \quad (2.26)$$

Here, ϵ is a small parameter (of order 10^{-3} or less), Φ is the gravitational potential at the location of the planet i , and \mathbf{a}_i is the planet's acceleration. At the beginning of every time step, this value is determined for each body in the system. The scheme then may then move up or down one rung - or stay on the same rung - depending on how far the optimal time step is from the current timestep. The timestep may always move up one rung (to a shorter timestep), but will not be allowed to move down a rung (to a longer timestep) if this could potentially cause the N -body integrator to end up out of block step. The main advantage of this scheme is that it allows the N -body code to stay exactly in sync with another code such as PLUTO, but also allows for exact output times. This scheme is largely similar to that suggested in [Dehnen & Read \(2011\)](#), albeit with each body on the same time step rather than on separate ones. Again one could in principle achieve a speed increase in the code by allowing each particle to occupy a separate rung, but for low N simulations such as those presented here, this speed increase would be minimal.

We can obtain some insight into the behaviour of the time-stepping criterion in equation 2.26 by considering the simplifying $N = 2$ case. Using equations 1.44 and 1.45, we then find an optimal time-step of

$$\tau_{optimal} = \epsilon \sqrt{\frac{R^3}{GM}}. \quad (2.27)$$

For a circular orbit of radius R and with $\epsilon = 2\pi$, this time-stepping criterion becomes the orbital period. Therefore one can trivially ensure that such an orbit is resolved into m discrete time-steps by selecting a value $\epsilon = 2\pi/m$. This criterion therefore lends itself trivially to problems involving planetary systems around single stars, since it allows one to ensure that the orbit is accurately integrated by selecting the appropriate value of m . The default value of m in PENGUIM is a very conservative 1000, selected to ensure that Keplerian orbits are accurately integrated.

The natural issue with this time-stepping criterion occurs when one considers systems where $N > 2$. In these systems, one can envisage the existence of points in the simulation domain where $|\mathbf{a}_i|$ for a given particle becomes very close to 0, meaning that equation 2.26 would give rise to a very large time-step. The most obvious such configuration lies on the line connecting two bodies. For any two given masses, there will be a point on this line at which the gravitational forces exerted on a test particle cancels out exactly. Evidently such configurations will also occur in PENGUIM simulations of planetary systems, leading to a large optimum time-step for any body that is near to this point. The effect that this issue might have in PENGUIM is minimised by the use of a global time-step; whilst one particle may lie near to such a point, the other particles in the simulation will still require a shorter time-step, forcing the code to use a shorter overall time-step.

2.2.2 *N-body damping*

Simulating the interaction between a planet and its parent disc that leads to migration is a numerically challenging problem, with full hydrodynamical simulations (discussed later) requiring vast amounts of computing time. In order to emulate the effects of planet migration without incurring such a high computational cost, I use toy forces which enact a change in both semi-major axis and eccentricity to each planet.

The way in which these damping forces are implemented in PENGUIM is as follows. At each time step I damp the velocity of each planet such that the orbital elements a and e are reduced. The forces are velocity-dependent and are applied as part of a modified Leapfrog integrator, using the “generalised Leapfrog method”, such that the extra accelerations are integrated to second order and pseudo-symplectically. The forces are, for migration and eccentricity damping respectively

$$\ddot{\mathbf{r}}_a = -\frac{\dot{\mathbf{r}}}{2} \frac{1}{\tau_a} \quad (2.28)$$

$$\ddot{\mathbf{r}}_e = \frac{2}{3} \left(\frac{\mu}{h} \hat{\mathbf{h}} \times \hat{\mathbf{r}} - \dot{\mathbf{r}} \right) \frac{1}{\tau_e} \quad (2.29)$$

We can write the standard second-order kick-drift-kick Leapfrog as

$$K\left(\frac{\tau}{2}\right) D(\tau) K\left(\frac{\tau}{2}\right), \quad (2.30)$$

where τ is the simulation timestep, K is the kick operator (which evolves the velocities exactly for $\dot{\mathbf{r}} = 0$) and D is the drift operator (which evolves the positions exactly for $\dot{\mathbf{r}} = 0$). The above forces

can be applied only at a time when the exact positions and velocities of the particles are known at the same time - either in the middle of a step of the above algorithm or at the end. I hence introduce a new operator $E(\tau)$ which performs the damping operations and write the new leapfrog as follows:

$$E\left(\frac{\tau}{2}\right)K\left(\frac{\tau}{2}\right)D(\tau)K\left(\frac{\tau}{2}\right)E\left(\frac{\tau}{2}\right). \quad (2.31)$$

The effect of these forces on the velocities at each time-step can be integrated in the limit of $\dot{\mathbf{r}} = 0$ (as required by the generalised Leapfrog) exactly by writing them in terms of radial and azimuthal components as

$$v_R(t + \tau) = v_R(t)\exp\left(-\frac{\tau}{2\tau_a} - \frac{2\tau}{3\tau_e}\right) \quad (2.32)$$

$$v_\phi(t + \tau) = v_e^2 + (v_\phi^2(t) - v_e^2)\exp\left(-\frac{\tau}{2\tau_a} - \frac{4\tau}{3\tau_e}\right) \quad (2.33)$$

where

$$v_e^2 = \frac{\mu}{r} \left(1 + \frac{3\tau_e}{4\tau_a}\right)^{-1}. \quad (2.34)$$

These forces damp the *unperturbed* orbital elements in an orbit-averaged sense, as can be seen from considering the change in orbital angular momentum and energy affected by the forces. The osculating orbital elements hence do not vary exactly as in equations 3.2 and 3.4, but I consider this method to be the most satisfactory prescription for physically realistic damping. I note that this method can provide considerable efficiency improvements compared to alternatives (e.g, [Lee & Peale, 2002](#)) since the orbital elements do not need to be calculated at each time-step.

2.2.3 *N*-body code tests

Before using my *N*-body code PENGUIM to run production simulations, it was necessary to test the performance of the integration methods and time-stepping criterion discussed above. An interesting first test of the code is the three-body solution discovered by [Moore \(1993\)](#), in which three equal-mass bodies orbit around one another drawing closed, figure-eight shaped tracks as they go. This solution was shown to be stable by [Simó \(2002\)](#) and as such, should continue to trace figure eights ad infinitum if numerical integration is sufficiently accurate. As initial conditions for this test I use those calculated by Carles Simó in [Chenciner & Montgomery \(2000\)](#). The solution is periodic and revisits its original

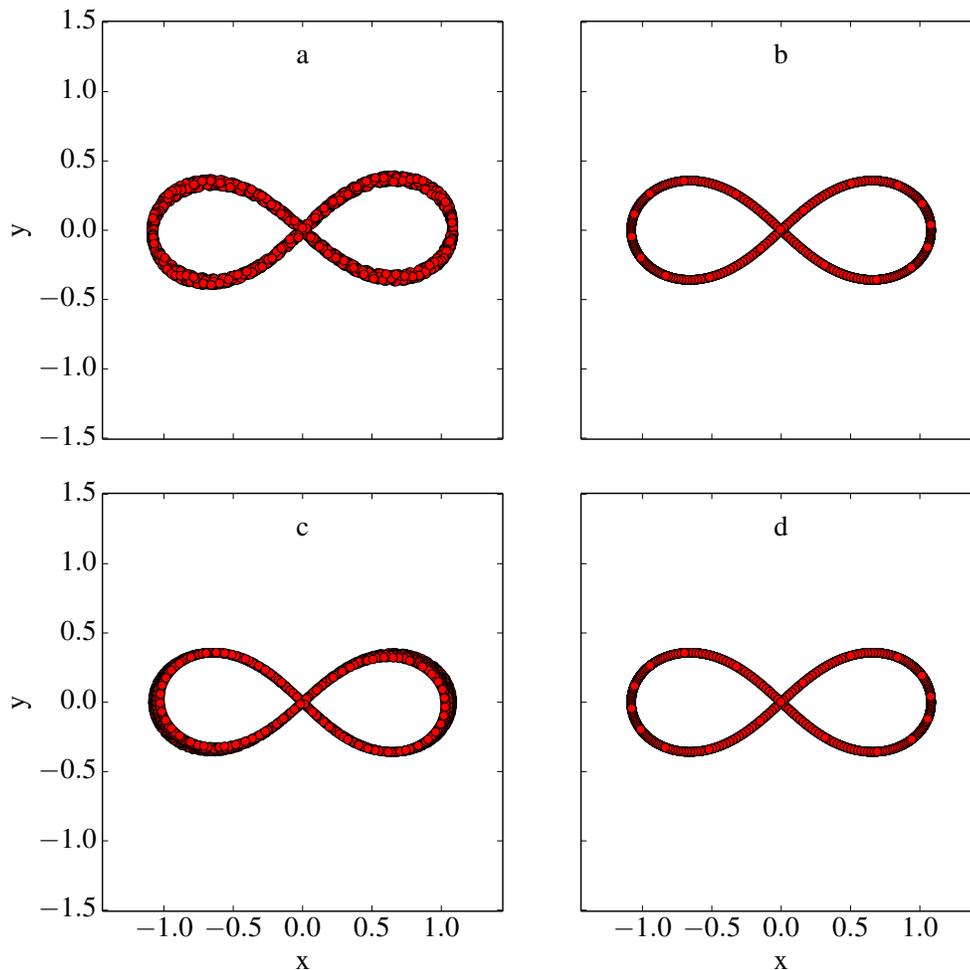


Figure 2.1: Figure eight orbit used to test N -body integration methods and time-stepping. The initial configuration of 3 bodies is evolved for 1000 time units (around 158 periods). The track of one of the three bodies is plotted throughout the entire simulation. Panel a: Leapfrog integrator, constant time-step $\Delta t = 0.1$. Panel b: Leapfrog integrator, adaptive time-step. Panel c: Runge-Kutta 4th order, $\Delta t = 0.1$. Panel d: Runge-Kutta 4th order, adaptive time-step.

configuration once every 6.32591398 time-units. I use this problem to test the Leapfrog and Runge-Kutta 4th order integrators with both adaptive and fixed time-stepping. Figure 2.1 shows the results of these tests.

Evidently, PENGUIM is capable of integrating this problem accurately for a prolonged period, with the bodies continuing to trace figure eights. This test also shows the clear advantage of using adaptive time-stepping. The thickness of the tracks in the constant time-step simulations suggests the the

2.2. N-BODY CODE

motion is not integrated sufficiently accurately and that the bodies wander slightly away from the initial figure eight orbit as a result of these errors. In the adaptive time-stepping cases, the bodies continue to trace the exact same figure eight throughout the entire simulation. This could also be achieved using a much shorter, constant time-step, but a well-tuned adaptive time-step is more computationally efficient.

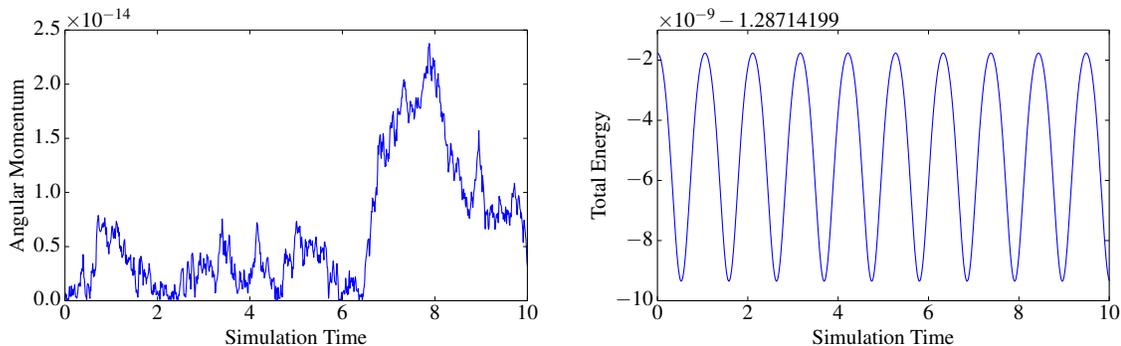


Figure 2.2: Energy and angular momentum conservation using the Leapfrog integrator on the figure eight problem, $\Delta t = 0.0001$.

As an addition to this first test, I ran the same initial conditions in PENGUIM using the Leapfrog integrator, a constant time-step of $\Delta t = 0.0001$, and a very high output frequency. The aim of this was to establish if the Leapfrog implementation in PENGUIM behaves as it should; that is to say, there is no secular evolution of the energy error and the initial angular momentum of the system (which in this case, is 0) is conserved. Figure 2.2 shows the result of this test. As expected, there is no secular growth in energy error, rather the total energy oscillates around its initial value. The angular momentum evolves along what appears to be a random walk, undergoing only small changes that can be attributed to round-off error - there is clearly no secular evolution. Evidently, the Leapfrog is behaving as expected.

For a second test and so that PENGUIM may be compared to another code, I integrate the orbit of a highly eccentric ($e = 0.9$) Jupiter-mass body around a Solar-mass body for 2616 orbital periods. I run this same test in both PENGUIM and long-standing N -body code MERCURY (Chambers & Migliorini, 1997). Figure 2.3 shows the resulting orbital tracks.

Here the issue with constant time-stepping becomes abundantly clear - the orbit precesses sig-

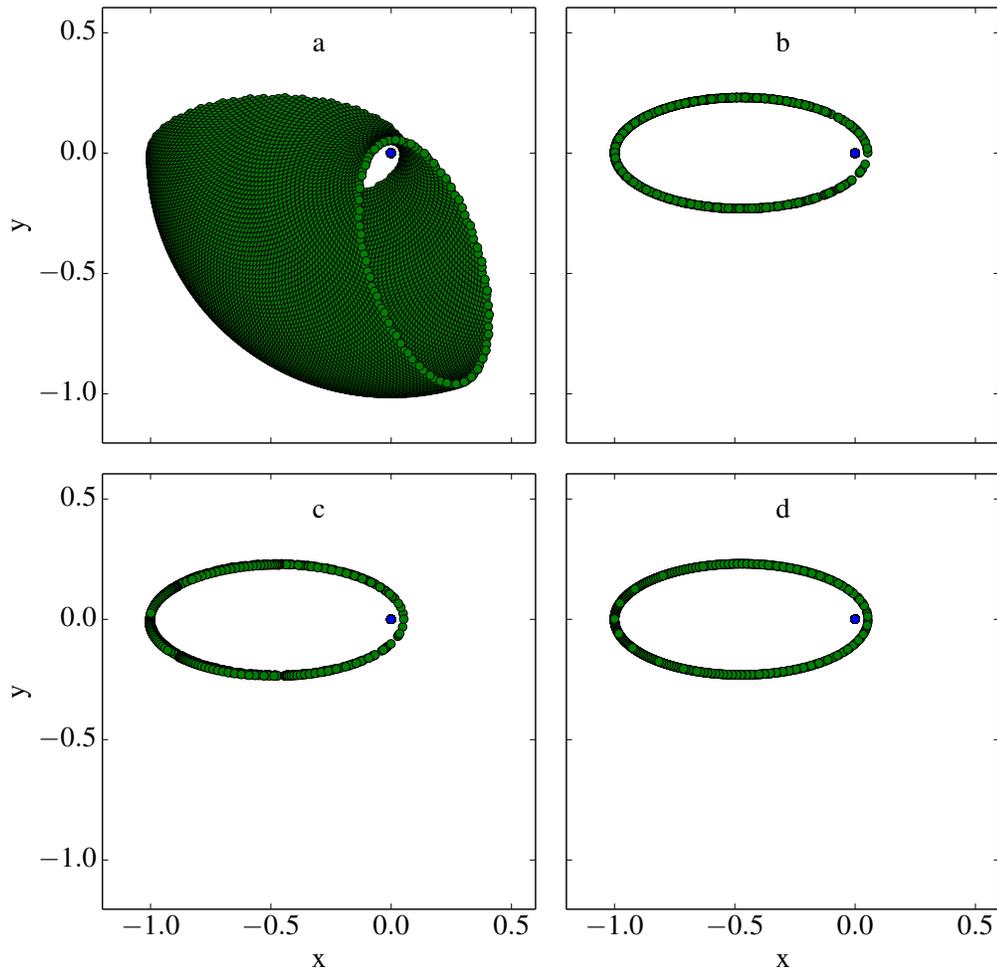


Figure 2.3: Highly elliptical orbit used to test N -body integration methods and time-stepping. The initial configuration of 3 bodies is evolved for 1000 time units (around 2616 orbital periods). The track of the two bodies is plotted throughout the entire simulation. Panel a: Leapfrog integrator, constant time-step $\Delta t = 0.0001$. Panel b: Runge-Kutta 4th order, constant time-step $\Delta t = 0.0001$. Panel c: Leapfrog integrator, adaptive time-step. Panel d: Mercury, multi-variable symplectic integrator, $\Delta t = 0.0001$.

nificantly with when a constant time-step is used with the Leapfrog integrator. I note that this is a known phenomenon with the Leapfrog (see e.g., [Dehnen & Read, 2011](#)). The Runge-Kutta performs much better with an identical constant time-step, though due to the extra force calculations is nearly four times more computationally expensive. The Leapfrog with adaptive time-stepping, which is the scheme I use for most integrations in the following chapters, performs admirably well compared to MERCURY, with minimal precession and the planet continually retracing the same track. These tests

suggest that the implementations of the Leapfrog, Runge-Kutta and adaptive time-stepping algorithm within PENGUIM function correctly, and that the code may be used to integrate planetary systems with confidence.

2.3 Hydrodynamical

The earlier discussion of thin accretion discs demonstrated the importance of fluid mechanics in astrophysics. In that particular case the problem was analytically tractable due to its axisymmetry, however, with the discussion of type-I migration above, it is evident that we also wish to understand non-axisymmetric, complicated behaviours in the fluid. Thus we must consider a more general case.

The Eulerian equations of hydrodynamics for compressible, three-dimensional flow are given by

$$\frac{D\rho}{Dt} + \rho \nabla \cdot \mathbf{v} = 0 \quad (2.35a)$$

$$\rho \frac{D\mathbf{v}}{Dt} = -\nabla p - \rho \nabla \Phi \quad (2.35b)$$

$$\rho \frac{D}{Dt} \left(\frac{e}{\rho} \right) = -p \nabla \cdot \mathbf{v}. \quad (2.35c)$$

Here, ρ is the mass density of the fluid, \mathbf{v} is the velocity vector of the fluid, Φ is the gravitational potential, e is the internal energy density and p is the pressure. D/Dt is known as the comoving, Lagrangian, or material derivative

$$\frac{D}{Dt} \equiv \frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla. \quad (2.36)$$

These equations are a system of linked hyperbolic partial differential equations, which cannot be solved analytically for the general case. We hence turn once again to numerical methods to approximate a solution.

Broadly speaking, the simulation of hydrodynamics in an astrophysical contexts is carried out using one of two approaches; either a *Eulerian* approach, whereby the fluid quantities are tracked and advected across a grid of fixed points, or a *Lagrangian* approach, whereby the fluid quantities are tracked by particles or other interpolation points which move through space.

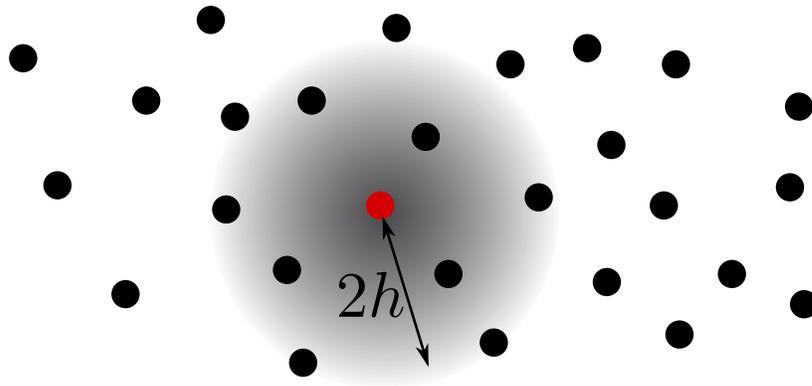


Figure 2.4: Schematic diagram showing how SPH works. A smoothed average for the density of the red particle is calculated using a smoothing kernel over a radius $2h$, where h is the so-called smoothing length.

Smoothed Particle Hydrodynamics (SPH)

SPH works by discretizing the *mass* contained in a fluid, and thus belongs to the Lagrangian class of codes. The fluid is therefore described by a series of particles each carrying (usually) a set amount of mass, which move in space and trace the flow of the fluid. The problem is then to estimate the density of the fluid at any point in space in order to solve equations 2.35. One can conceive of many different approaches to doing this: assigning the particles to a grid and calculating the density within each cell, considering a finite sphere of volume around the particle, or partitioning the space up using tessellation techniques to separate each individual particle into a characteristic volume (see e.g., [Springel, 2010](#)). The main issue in selecting a method is that the discreteness of the mass distribution leads to unwanted noise in the density estimation, and we ideally wish to minimise this noise. SPH approaches this problem by calculating the density at each point using a *smoothing kernel*, such that particles near to the point in question contribute more to the density estimate than those further away ([Gingold & Monaghan, 1977](#); [Lucy, 1977](#); [Monaghan, 1992](#)). Figure 2.4 provides a schematic of this approach. I refer the interested reader to the excellent review of [Price \(2012a\)](#) for more information. SPH-based codes have a number of advantages and disadvantages with respect to other codes, and the literature comparing various approaches has turned into something of a quagmire (see e.g., [Agertz et al., 2007](#); [Price, 2008, 2012b](#)) that I shall not be tackling here. The important thing is naturally to select the approach which has the best properties for the problem at hand. With regards to simulations

of type I migration, the limiting factor is that high resolution is required in the horseshoe region around the planet to model the migration correctly, and SPH generally incurs a higher computational cost than other fluid dynamics codes for similar spatial resolution. As a result, I stick with the Eulerian techniques discussed below to minimise computational time.

Fixed-grid and Adaptive Mesh Refinement (AMR)

These codes solve the Eulerian equations of hydrodynamics by discretising *space-time* on a grid. Fixed-grid codes keep the discretization of space constant in time, whilst AMR codes adapt the discretization to increase resolution in areas of interest. To begin in the discussion, I follow [Roe \(1986\)](#), and we consider a general system of one-dimensional conservation laws that can be written in the following way

$$\mathbf{q}_t + \mathbf{F}_x = 0 \tag{2.37}$$

where \mathbf{q} is a vector of state variables and \mathbf{F} is a vector of fluxes. For simplicity I have neglected source terms here. An application of Gauss's divergence theorem leads to the integral form of the conservation laws

$$\oint (\mathbf{q}dx + \mathbf{F}dt) = 0 \tag{2.38}$$

for a closed contour in (x, t) . This second form is more general since it holds even in the event that the vectors are not differentiable. These two ways of casting the conservation laws lead to two different methods of solving the underlying problem: finite-difference and finite-volume schemes. Finite-difference schemes take Taylor series expansions of the hyperbolic differential equations in question, hence solving these to some order of terms.

The simplest case of equation [2.37](#) occurs when there is only one state variable with a constant flux velocity. In this case we can write the differential form as

$$\frac{\partial q(x, t)}{\partial t} + u \frac{\partial q(x, t)}{\partial x} = 0. \tag{2.39}$$

This equation describes the advection of a function $q(x, t)$ in a one-dimensional space, with u being a constant. For this simple case, the equation can be solved trivially using the method of characteristics. Using this method, we seek to find a new coordinate s in which we may express $x(s)$ and $t(s)$, and in

2.3. HYDRODYNAMICAL

which the advection equation is an ordinary differential equation rather than a partial one. This means we attempt to find the function $F(q, x(s), t(s))$ such that $dq/ds = F$. The curve defined by $(x(s), t(s))$ is known as a characteristic (Holmes, 2006). The total derivative of q may be expressed with respect to s as

$$\frac{dq}{ds} = \frac{\partial q}{\partial s} \frac{dt}{ds} + \frac{\partial q}{\partial x} \frac{dx}{ds} \quad (2.40)$$

and we note that this becomes equation 2.39 for $dt/ds = 1$ and $dx/ds = u$, which allows us to write $dq/ds = F(q, x(s), t(s)) = 0$. Integrating the first 2 ODEs here with the initial condition $x(0) = x_0$, $t(0) = 0$, we find that $s = t$ and $x = x_0 + us = x_0 + ut$. Interestingly, this has also proven that $dx/dt = u$ without any a-priori knowledge about u . We see that equation 2.39 simply expresses that $dq/dt = dq/ds = 0$, meaning that if we follow an individual pocket of q on its journey through space-time, the amount of q inside it is constant. This means that $q(x, t) = f(x_0) = f(x - ut)$, and therefore that q is constant along lines in the (x, t) plane defined by $x = x_0 + ut$. These lines are the characteristics of the advection equation, and give describe how the position x of some information changes in the scalar field q .

We now consider ways in which we might solve equation 2.39 numerically. To solve this equation using the finite-difference approach, we attempt to discretise the partial differentials such that the can be computed from the value of q on a grid (Dullemond, 2011)³. In the following, subscript i denotes the cell on the spatial grid whilst superscript n denotes the position on the temporal grid. We know the function $q(x, t^n)$ for each i and wish to compute the function $q(x, t^{n+1})$ for each i . One possible discretization that achieves this is

$$\frac{q_i^{n+1} - q_i^n}{t_{n+1} - t_n} + u \frac{q_{i+1}^n - q_{i-1}^n}{x_{i+1} - x_{i-1}} = 0. \quad (2.41)$$

This is known as the centred differencing scheme, since the gradient of q at x_i is computer using only the values of q to either side of it. We can rearrange trivially for q_i^{n+1} . Numerical integration using this scheme leads to some issues, however. The results of a simple integration are shown in in figure 2.5. This integration began with $q = 1$ for $x < 0$ and $q = 0$ for $x = 1$. The advection velocity is $u = 1$ and the integration ran for 3000 time-steps where $\Delta t = 0.0001$. The initial jump in q has propagated to

³The lecture notes provided by C. Dullemond (http://www.ita.uni-heidelberg.de/~dullemond/lectures/num_fluid_2011/) have proven invaluable to me for understanding the various approaches to computational fluid mechanics.

2.3. HYDRODYNAMICAL

$x = 0.3$ as we would expect analytically, but the jump has been smeared out and there are unphysical oscillations upstream (against the direction of flow) of the jump. Applying physical intuition to this scheme reveals one of the issues: we have approximated the derivative in x using one cell that is upstream of the cell in question and one cell that is downstream. Information in physical systems should clearly move from upstream to downstream, and not be allowed to propagate upstream. If we fix $u > 0$, then the upstream direction is defined by $x < x_i$, and we can estimate the differential using the following method

$$\frac{q_i^{n+1} - q_i^n}{t_{n+1} - t_n} + u \frac{q_i^n - q_{i-1}^n}{x_i - x_{i-1}} = 0, \quad (2.42)$$

which only takes information from upstream of cell i . Using the same initial value problem as before, I have performed the same integration with this new method. The results are shown in in figure 2.5. Fixing the flow of information to be from upstream to downstream has evidently solved the previous issue of noise, but the jump remains smeared. Note that this smearing could also be reduced by increasing the spatial or temporal resolution, but that the assumption of differentiability will also lead to *some* smearing.

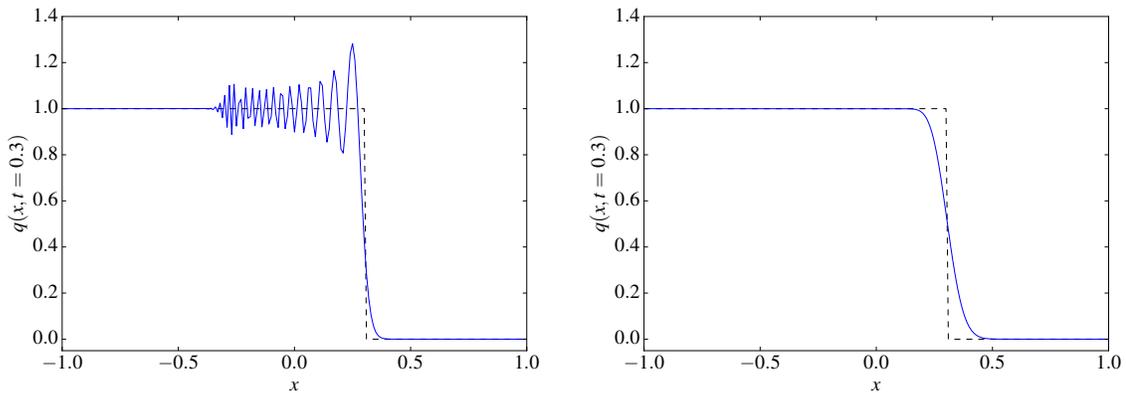


Figure 2.5: Numerical integrations of the differential linear advection equation (equation 2.39) in solid blue, with the analytic solution plotting as black, dashed lines. The analytic solution is that computed from the method of characteristics in the text. Left: centred-differencing scheme. Right: Upwind-differencing scheme. Note that in both cases the step function is smeared out due to the assumption that the function $q(x, t)$ is differentiable.

Modern finite-difference schemes use markedly more advanced techniques for estimating the differentials, and the above discussion is not intended to be a comprehensive one, but rather one that

highlights the salient point of such schemes, and I refer the reader to the ZEUS code paper (Stone & Norman, 1992) for a more in-depth look at how a modern finite-difference scheme functions.

Most of the finite-volume schemes in use today take a very different approach, and rely on the integral form of the conservation equations, calculating the fluxes passing through the closed surface of a finite-volume element in space-time. The majority of these schemes have in fact evolved based on a completely different underlying physical interpretation of the discretization problem to finite-difference schemes, pioneered by Godunov (1959). The finite-difference approach assumes that the solution across the discretized space is smooth, meaning that sharp discontinuities (and therefore shocks) are implicitly handled poorly by the scheme (Woodward & Colella, 1984). Godunov's alternative was to consider the grid and therefore the fluid as a whole to consist of a series of discontinuous solutions and solve the equations at the boundaries between these discontinuities. Naturally if two neighbouring cells have very similar properties, then the discontinuity is small and Godunov's method approximates the smooth solution excellently. This method then also has the advantage that it solves more accurately for shocks in the fluid. Schemes based on this principle have progressed massively and been used extensively over the last 30 years.

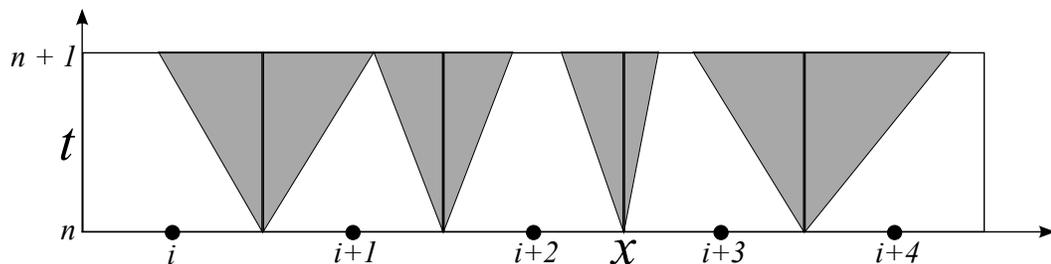


Figure 2.6: Schematic diagram showing how Godunov's method works. The 1 dimensional space is discretized into a grid, with cell-averaged state variables being stored at the centre point. To evolve the system from time n to $n + 1$, a self-similar Riemann problem is solved at the interface between each pair of cells. Waves (grey shaded regions) or characteristics propagate out from the interfaces as a result of this solution. The waves traversing cell $i + 1$ are about to interact at the end of the time-step, meaning that the time-step could not be any longer.

Godunov's idea (see figure 2.6) requires the solution of a so-called Riemann problem at the interface between each cell. This is a non-linear problem whereby for a given set of hyperbolic equations, the initial states either side of the boundary can each be prescribed as a semi-infinite state. Note that the initial condition for $q(x, t)$ that I integrated numerically in figure 2.5 is actually a Riemann prob-

lem, since the state either side of $x = 0$ is constant. When solving the Euler equations via this method, one finds that three features develop: a contact discontinuity in the centre with a wave (rarefaction or shock) travelling off to each side. As time evolves, these waves propagate further into the 2 cells to either side of the boundary. In the case of linear conservation laws, these waves are simply the characteristic lines of the conservation laws, which in turn are lines that represent the propagation of a signal through space-time in an advective system. It is important to ensure that the waves driven by neighbouring boundaries do not overlap and thus this sets an upper-limit on the time-step. Once these solutions have been calculated for the entire domain, the volume-averaged quantities can be updated in a conservative fashion and the process repeated again for the new values (Roe, 1986).

The volume averaged quantities can be calculated in one of two ways. One can consider the new solution defined by the propagating waves as a piecewise uniform distribution and perform the required averaging across each cell. The preferred method however, involves invoking the integral form of the conservation laws described by equation 2.38. Referring back to figure 2.6, the region in space-time occupied by the boundaries of one cell as it progresses from time n to $n + 1$ forms a complete contour in (x, t) . At time $n + 1$ this region contains an exact solution to the series of hyperbolic equations, albeit an exact solution to an approximate problem. Thus we invoke the integral form, finding that

$$(\mathbf{q}_i^{n+1} - \mathbf{q}_i^n)\Delta x + (\mathbf{F}_{i+1/2}^{n+1/2} - \mathbf{F}_{i-1/2}^{n+1/2})\Delta t = 0. \quad (2.43)$$

Here, superscripts denote the time at which a variable is measured and subscripts denote the position. Hence the values \mathbf{F} are calculated at the interfaces between i and its neighbouring cells, and are time averages of the the fluxes. Note however that since the Riemann problem that is solved is self-similar, the flux is constant throughout the time-step and thus no actual averaging is required. We see then that we can easily and conservatively solve for the state vector of an individual cell \mathbf{q}_i^{n+1} given the original state \mathbf{q}_i^n and the solution to the Riemann problem at each boundary. This approach has the advantage of explicitly using the integral form of the conservation laws, meaning that non-differentiable features such as contact discontinuities and shocks can be advected cleanly and with little smoothing.

Godunov's original description of his method included an exact Riemann solver for the Euler equations. Such a solver is however rather numerically expensive and requires iterative methods to work. Thus the majority of modern codes use approximate Riemann solvers, often using methods such as linearising the Euler equations in order to easily calculate an approximate solution that conserves

most of the non-linear behaviour.

2.3.1 PLUTO

The PLUTO code (Mignone et al., 2007) is designed to integrate a system of general hyperbolic conservation laws that can be written as

$$\frac{\partial \mathbf{q}}{\partial t} = -\nabla \cdot \mathbf{T}(\mathbf{q}) + \mathbf{S}(\mathbf{q}) \quad (2.44)$$

Here \mathbf{q} remains a vector of state variables, $\mathbf{T}(\mathbf{q})$ is a rank-2 tensor describing the fluxes of \mathbf{q} , and $\mathbf{S}(\mathbf{q})$ describes additional source terms. Any system of hyperbolic equations, including equations 2.35 can be recast in this form.

The code integrates such equations using the Riemann problem based formalism described above, and uses the method described by equation 2.43 to update volume-averaged state vectors. It additionally performs an interpolation step to establish cell-interface values of \mathbf{q} starting from their volume-averages. It provides a selection of different Riemann solvers that approximate the problem in various different ways, and allows for easy set-up of various physical problems. As a result of this, I use PLUTO for all the hydrodynamical simulations described in chapter 5 of this thesis.

Modifications to PLUTO

By default, PLUTO has no apparatus with which to deal with N -body dynamics. As a result of this, I coupled my N -body code to the standard distribution of PLUTO to solve for planet-planet interactions.

I use the standard Runge-Kutta 4th order integrator to integrate planetary orbits. This is in contrast to the Leapfrog integrator used throughout the rest of this thesis. The change from the leapfrog was motivated by the complexity of keeping the fluid and N -body codes synchronised, since in the Leapfrog, the force calculation is performed half way through a time-step. Time-stepping is handled adaptively using the formalism described in section 2.2.1, which allows the N -body code to stay in block-step with the PLUTO in the (rare) event that PENGUIM wishes to use a time-step shorter than that of PLUTO.

The softened gravitational potential exerted by the N -body particles on the fluid is calculated at each time-step by a call to the PENGUIM API, performed as part of PLUTO's standard potential-calculation routine. The corresponding acceleration is calculated in a separate part of the PLUTO

main loop later on, and added to the accelerations that PENGUIM calculates on its own. I note that an efficiency increase could be obtained by performing these operations simultaneously, but that this would complicate the modifications to PLUTO significantly.

PLUTO scaling test

In order to optimise the runtime of simulations performed with PLUTO, it was necessary to determine how well the modified version of PLUTO scales over multiple processors. To this end, a scaling test was completed. I ran PLUTO for one hour using a test setup on varying numbers of processors, in order to discover how many timesteps the code completed within an hour on a given number of cores. The results of this test can be seen in figure 2.7. This test was run on 16, 32, 48, 64, 80, 96, 128, 160, 192, 256, 320, 384, 512, 768 and 1024 cores. The test setup included one planet, and used a grid resolution of 500 cells in R (from $R = 0.3$ to $R = 6$) and 1200 in azimuth. The simulation is essentially identical to those described in section 5.2.7 in all but runtime. Naturally it was vital to perform this test at production resolution in order to determine the optimal number of cores for the production runs. As a result of this test I elected to use 512 cores for the production runs, as the scaling of the code gets noticeably worse beyond this number.

2.4 Tonight Matthew, I'm going to be...

In this introductory section, I have described much of the current progress in understanding the formation of planetary systems in discs, and their subsequent dynamical evolution. Many unanswered questions remain in this field, however. The existence of dynamically cold, tightly-packed systems of super-Earths remains difficult to explain given the physics I have discussed here. Type I migration naturally leads to convergent migration if larger planets are exterior to smaller ones, since the migration rate scales with planet mass. This in turn, naturally leads to resonant trapping which excites eccentricity. How then, do systems such as Kepler-11 and Kepler-32 form, if the traditional scenario of them forming further out in the disc and migrating inwards doesn't lead to stable systems? Can other effects explain the breakdown of these resonances? What processes differentiate this class of compact planetary systems from our own?

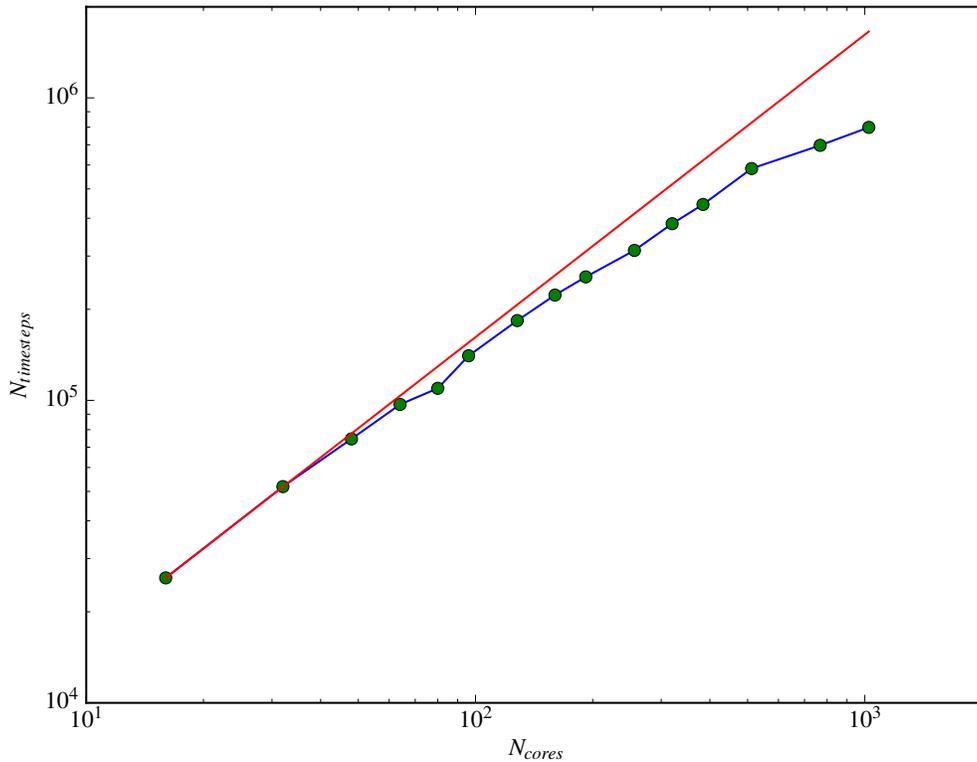


Figure 2.7: Results of the scaling test for the modified version of PLUTO. Points show actual data, blue line connects the points as a guide for the eye. The red line shows linear scaling, again as a guide for the eye. The code scales well up to 512 cores, after this there is a drop off.

For the remainder of this thesis, I will use the numerical methods described in chapter 2 to explore the physics of planetary systems undergoing type I migration. As we have seen in this introduction, the process of type I is intrinsically very complicated, and resolving the region in which the horseshoe torque is generated in particular requires very high resolution hydrodynamical simulations which are very computationally expensive. As a result of this, I explore the problem first using an N -body method with toy forces to describe the disc interaction, rather than performing expensive hydrodynamical simulations. This allows the exploration of a greater parameter space. To begin with in chapter 2, I will apply this method to compact systems of super-Earths, then in chapter 3 I consider the additional effect that a perturbing Jupiter mass planet exterior to a compact system would have. Finally in chapter 4, I move on to full hydrodynamical simulations using PLUTO, to model the interaction of

2.4. TONIGHT MATTHEW, I'M GOING TO BE...

multiple migrating super-Earths in a realistic disc, in an attempt to understand how migration of such systems can come to an end without the planets migrating onto the host star, and without being locked in resonance with their neighbouring planets.

Chapter 3

Understanding the assembly of Kepler's compact planetary systems

If thou, dear reader, art bored with this wearisome method of calculation, take pity on me, who had to go through with at least seventy repetitions of it, at a very great loss of time.

Johannes Kepler
Astronomia nova

3.1 Introduction

Of the myriad of recent advances in the study of extra-solar planets, perhaps the most interesting is the discovery of large numbers of systems of multiple planets. The first multiple-planet system around a main sequence star, Upsilon Andromedae, was reported by [Butler et al. \(1999\)](#), and we now know that many, if not most, planets form in multiple systems. The *Kepler* mission has provided astronomers with a novel insight into this class of planetary system.

Kepler detections are made by the transit method, discussed in section [1.2.2](#). The telescope observes temporal variation in the brightness of a star, watching for a characteristic, periodic dip in the intensity, which is caused by a planet transiting in front of it. This method allows the characterisation of the period of the planet, as well as the radius of the planet relative to the radius of its stellar host. Retrieving the actual radius of the planet from this data requires accurate characterisation of

the radius of the host star, whilst measuring masses requires combining the transit method with other methods such as radial velocity or dynamical models. As a result of these difficulties, even the bulk compositions of *Kepler* planets are generally not well known. Chapter 1 already covered the vast and increasing multiplicity observed in the *Kepler* sample. Statistical considerations suggest that the false positive rate in *Kepler*'s multiple-planet systems is low (Lissauer et al., 2012). When one considers non-detections, it seems likely that the majority of planets form in multi-planet systems.

Among this avalanche of new data, arguably the most surprising discovery was a new class of compact, tightly-packed systems. These systems, for which *Kepler-11* is the prototype (Lissauer et al., 2011b, 2013), consist of several Neptune- or super-Earth-size planets, typically within a few tenths of an AU of their host star, and often within a few hundredths of an AU of one another. The planets are well ordered in size, with larger planets typically orbiting at larger radii (see Fig. 3.1). The proximity of these planets to their stellar host causes problems for theories trying to explain how they might have formed. In general there are two schools of thought on how these systems might form, both of which build upon the classical theory that planets are formed in dust- and gas-rich discs around young stars. The first theory posits that they formed *in situ* (e.g., Hansen & Murray, 2012; Chiang & Laughlin, 2013), essentially forming in their current observed locations with minimal relocation within the disc. The second theory states that the planets themselves formed much further out in the disc, but a combination of various torques from gas in the disc caused inward migration of the planet. These planets are of sufficiently low mass that they would be expected to migrate in the type-I (low-mass) regime (Kley & Nelson, 2012), where migration rate scales with mass. Both theories have their advantages and disadvantages.

The dynamical properties of these systems make the case even more complicated. The planets are invariably dynamically cold, with low eccentricities and low mutual inclinations. However, the proximity of the planets to one another and their stellar host suggests that violent dynamical interactions leading to growth of both properties might not be uncommon. Indeed, many of these compact systems appear to be close to dynamical instability (Deck et al., 2012). Mean-motion resonances between adjacent planets are seen in some cases, but not all (e.g., Lissauer et al., 2013; Swift et al., 2013). At this point I refer the reader back to section 1.4.2. If the planets in these systems had once been trapped in resonance, we would expect them to still be trapped and to have forced eccentricities. This presents

issues for the migration theory since the scaling of migration rate with planetary mass combined with the mass ordering of these systems suggests that planets would converge on one another, which has been shown to cause resonant trapping (see e.g., [Baruteau & Papaloizou, 2013](#)). Overall then, compact planetary systems represent a striking contradiction: they apparently require delicate assembly in order to avoid being destroyed by dynamical instabilities, but their prevalence suggests a robust formation mechanism.

Returning to the previous two models, the lack of MMRs might indicate that the *in situ* scenario is more plausible. However, [Swift et al. \(2013\)](#) show that the dust sublimation radius for Kepler-32 is outside the innermost planet’s orbit for the length of all but the most unrealistic disc lifetimes, meaning that one would not expect enough solid material to exist in the disc to form at least some of these planets at their present locations. Moreover, conventional protoplanetary disc models contain too little mass in all forms at small radii to form systems such as Kepler-32 *in situ* ([Swift et al., 2013](#)). [Hansen & Murray \(2013\)](#) used Monte Carlo simulations of *in situ* formation to show that this mechanism can qualitatively reproduce the distribution of tightly-packed systems – albeit with a slight shift towards longer periods – while [Hansen & Murray \(2014\)](#) suggested that tidal dissipation may be responsible for bringing planets inside the expected dust sublimation radius to shorter-period orbits. More recently [Raymond & Cossou \(2014\)](#) also found that the range of disc models required for *in situ* formation of Kepler’s sample of tightly-packed super-Earths is unrealistic.

The natural alternative to *in situ* formation is that hot super-Earths formed at larger radii and migrated inwards. [Raymond et al. \(2008a\)](#) made predictions of observational signatures that might help to distinguish between *in situ* and migration models, and find that mean-motion resonances would be expected if such systems formed by migration. Several authors have conducted simulations exploring the scenario in which hot super-Earths form from inwardly migrating planetary embryos. [Terquem & Papaloizou \(2007\)](#) performed calculations which can form 2–5 super-Earths in very tight orbits since their migration is halted at the inner edge of the disc, whilst [McNeil & Nelson \(2010\)](#) found that super-Earths of up to $3\text{--}4M_{\oplus}$ form readily in a sufficiently massive disc. [Cossou et al. \(2014\)](#) explored the problem with a more recent model of type-I migration and found that super-Earth sized objects can form from embryos migrating from 1–20AU and either pile up at the inner edge of the disc they are embedded in, or become giant enough to migrate outwards and become the core of a giant planet.

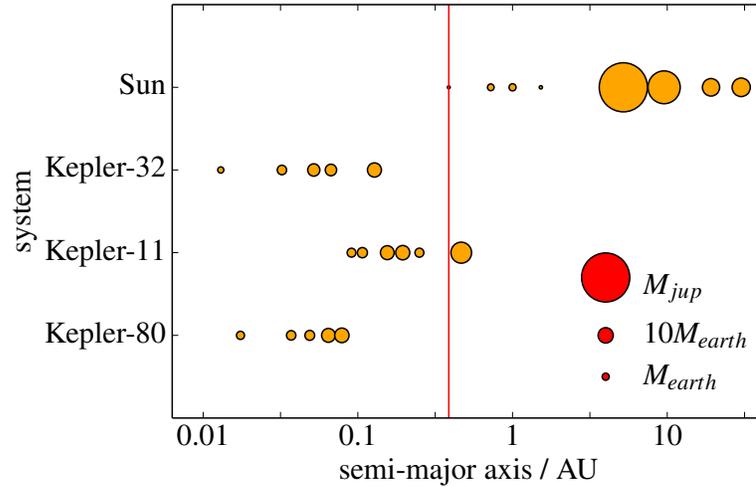


Figure 3.1: Comparison between tightly-packed planetary systems and our own. The orbit of Mercury is shown by the red line.

Note that the line between *in situ* formation and formation further out in the disc can be blurred once the embryos have migrated sufficiently far in, since they may continue to accrete material once the disc has dissipated via collisions with other bodies (Cossou et al., 2014).

In this chapter I investigate assembling compact, tightly-packed planetary systems by traditional, disc-driven migration. I adopt an N -body approach, using a variety of parametrised forces designed to mimic those that each planet would experience in a real protoplanetary disc. I consider the influence of each of the main elements of the Type I regime: planetary migration, eccentricity damping and disc turbulence. This results in a large parameter space with many inherent uncertainties, so I use a statistical approach, running large numbers of models of individual systems in order to understand what balance of parameters is most conducive to building these systems. This approach allows us to capture the essential physics of the simultaneous migration of multiple planets, while avoiding the computational cost of full hydrodynamic calculations. In essence my calculations are a proof-of-concept, designed simply to understand whether simultaneous migration of multiple planets is a viable model for the assembly of tightly-packed planetary systems. I find that it is possible to assemble such systems via disc-driven migration across a wide range of parameters.

3.2 Numerical method

Following the simultaneous migration of many planets in a hydrodynamic calculation is computationally expensive, and the large parameter space makes this approach infeasible here. I instead model the assembly of compact multi-planet systems using an N -body approach, employing parametrized forces to mimic the effects of disc-driven migration, eccentricity damping and turbulent forcing.

3.2.1 N -body integrator

I compute gravitational interactions (both star-planet and planet-planet) using a direct summation N -body code. I adopt a modified 2nd-order kick-drift-kick leapfrog method (see section 2.2.2) for time integration, with adaptive time-stepping to ensure numerical accuracy whilst minimising the computational cost¹. The star and planets are modelled as point masses, and no gravitational softening is used. I assign physical radii to all the particles, but these are used only to identify physical collisions between particles.

3.2.2 Migration and eccentricity damping

Gravitational torques from a protoplanetary disc drive both planet migration and eccentricity damping. The low planet masses considered here are expected to migrate in the Type I regime (Kley & Nelson, 2012) where this torque is given by (e.g., Tanaka et al., 2002; Paardekooper & Papaloizou, 2009)

$$\Gamma = -C \frac{q^2}{h^2} \Sigma_p a^4 \Omega_p^4. \quad (3.1)$$

Here a is the orbital semi-major axis of the planet, q is the planet:star mass ratio, Ω_p is the angular frequency of the planet and h is the disc aspect ratio, and C is an order-unity constant. The torque scales with the square of the planet mass, so the Type I migration rate da/dt increases linearly with planet mass M_p . I follow other authors (e.g., Lee & Peale, 2002; Rein, 2012), and parametrize the migration rate as

$$\frac{da}{dt} = -\frac{a}{\tau(M_p)}. \quad (3.2)$$

¹Strictly, the use of adaptive time-stepping violates the symplectic nature of the integrator but in practice this is unimportant, as the system is in any case non-Hamiltonian due to my use of non-conservative damping forces.

3.2. NUMERICAL METHOD

In contrast to previous work, I define the migration time-scale to be a function of planetary mass

$$\tau(M_p) = \tau M_{scale}/M_p. \quad (3.3)$$

Here τ and M_{scale} are reference values: I choose $M_{scale} = 3 \times 10^{-5} M_\odot$ (approximately $10 M_\oplus$), and treat the migration time-scale τ as an input parameter. I implement migration (and eccentricity damping; see below) in my code following a method described in the section 2.2.2. This method differs from that of Lee & Peale (2002) in that the damping is implemented directly as forces at each time-step, and is similar to the method implemented by Rein & Liu (2012). This approach captures the most important features of Type I migration (the migration rate and scaling with planet mass) while remaining computationally inexpensive. In adopting this method, I assume that all planets migrate inwards only. Several authors have recently shown that outward migration is possible for super-Earth sized planets, and that this behaviour can be important in preventing the loss of planets on to the star (Bitsch et al., 2013a, 2014b; Cossou et al., 2014). I note however that the only the most massive planets (those with $M_p \gtrsim 5 M_\oplus$) considered here are large enough to potentially migrate outwards in the disc models considered by these studies, and even then only in specific regions of the disc (e.g., Bitsch et al., 2014a). I consider a range of migration time-scales $\tau = 10^{3.5} - 10^{5.5}$ yr, except in the case of Kepler-11. It was clear from early work with this system that a longer migration time-scale was preferred, and hence for Kepler-11 I consider $\tau = 10^4 - 10^6$ yr. I do not base my choice of time-scale on a specific or evolving disc model; instead I explore a broad range of the relevant physical parameters, without reference to the disc properties. It is however possible to relate the migration time-scales presented here to canonical disc models using simple arguments. For instance, for a $1 M_\odot$ star and typical parameters ($\Sigma = 1000 \text{ g cm}^{-2}$, $a = 1 \text{ AU}$, $h = 0.05$, $M_p = M_{scale}$), equation 3.1 gives a migration time-scale $\tau = 6.3 \times 10^4$ yr. My adopted range of migration time-scales therefore corresponds approximately to a range of disc surface densities $\Sigma \simeq 10^2 - 10^4 \text{ g cm}^{-2}$ at 1 AU.

For low eccentricities ($e \lesssim 2h$) in the Type I regime, both analytic arguments and numerical arguments find that planet-disc interactions lead to exponential damping of the planet's eccentricity (Tanaka & Ward, 2004; Cresswell et al., 2007; Bitsch & Kley, 2010). I apply eccentricity damping in a similar fashion to that described above, with the damping rate given by

$$\frac{de}{dt} = -\frac{Ke}{\tau(M_p)}. \quad (3.4)$$

I therefore assume that the eccentricity damping rate de/dt is proportional to the migration rate da/dt ; K is the constant of proportionality. This assumption is motivated by several theoretical studies. [Tanaka & Ward \(2004\)](#) used a linear analysis for low-mass planets to show that $\tau_{ecc} \approx (H/r)^2 \tau_{mig}$, suggesting a $K \sim 10^2$ for canonical disc models. Similarly, [Lee & Peale \(2002\)](#) found that a value in the range $K = 10\text{--}100$ is required to form the GJ 876 system (although the planets in GJ 876 are notably more massive than those considered here). Hydrodynamical simulations performed by [Cresswell et al. \(2007\)](#) and [Bitsch & Kley \(2010\)](#) both show excellent agreement the results of [Tanaka & Ward \(2004\)](#) in the low e regime, finding eccentricity damping time scales to be a few tens of orbits. I treat K as an input parameter which sets the strength of the eccentricity damping, and consider values in the range $K = 10^{0.5}\text{--}10^{2.5}$. Note that no inclination damping is included in these simulations, since all planets are initialised on co-planar orbits.

3.2.3 Disc turbulence

Protoplanetary discs are turbulent, and in addition to transporting angular momentum and driving accretion, this turbulence results in stochastic fluctuations in the local gas density. These local density variations lead in turn to stochastic variations in the planet-disc torque, which cause the planet’s orbit to undergo a random walk ([Nelson & Papaloizou, 2004](#); [Nelson, 2005](#); [Oishi et al., 2007](#)). Full magnetohydrodynamic simulations of planet-disc interactions in a turbulent disc remain extremely computationally expensive, with some authors choosing to apply a stochastic potential to the disc to perform such simulations in an efficient manner (e.g., [Laughlin et al., 2004](#); [Baruteau & Lin, 2010](#); [Pierens et al., 2012](#)). Previous work has shown that this process can also be well-approximated in N-body calculations by applying stochastic forcing to the planets. Following the method of [Rein & Papaloizou \(2009\)](#), I use a modified discrete time Markov process to generate stochastic forces in both the ϕ and r direction at every point along the planet’s orbit, hence adding noise to the acceleration of each planet and sending its orbital elements on a random walk (around the smooth net migration rate described in Section 2.1). The standard Markov process is zero-mean and Gaussian, and is defined by two parameters: the root-mean-square (RMS) force-per-unit-mass $\sqrt{\langle F^2 \rangle}$ and the auto-correlation time τ_c . The RMS force is a free parameter in my model, whilst the auto-correlation time is set to Ω^{-1} for each planet. I note that this is a high estimate of τ_c and that numerical simulations show that it can

3.2. NUMERICAL METHOD

be as low as $0.5\Omega^{-1}$ (Oishi et al., 2007). I characterise the strength of the stochastic forcing in terms of the dimensionless parameter $\beta = \sqrt{\langle F^2 \rangle} / F_g$, where F_g is the gravitational acceleration due to the star's gravity. This means that the absolute magnitude of the stochastic forces grows as planets move inwards, so strictly my distribution of forces is no longer Gaussian (though it remains zero-mean). I vary β between 10^{-5} and 10^{-8} in all simulations presented here.

My parametrisation of the turbulence can be compared to previous works for scaling. Paardekooper et al. (2013) parametrize their noise according to the scale $F_0 = \pi G \Sigma / 2$, with $\Sigma = 1.5 \times 10^4 \text{g cm}^{-2}$, varying $\sqrt{\langle F^2 \rangle} / F_0$ between 0.01 and 0.1. The corresponding range for my β parameter (required to reproduce the absolute magnitude of the forces) is $\beta = 2.5 \times 10^{-7}$ to 2.5×10^{-6} at 0.1AU.

3.2.4 Simulation set-up

My simulations have three free parameters: the migration time-scale τ , the eccentricity damping constant K , and β , which characterises the strength of the disc turbulence. Predicted values of all three of these parameters span at least 2–3 orders of magnitude, so in order to span this vast parameter space I perform 10,000 simulations of each system, each with randomly chosen values of τ , K & β . The values of each parameter are sampled uniformly in log-space in order to understand how order of magnitude changes in their values affect the evolution of each system.

I apply my model to three well-known systems (Kepler-11, -32 and -80), the parameters of which are given below. I initially place all planets on circular, co-planar orbits with randomly chosen orbital phases. All planets are present at the start of the simulation and migrate together – the consequences of which I discuss briefly in Section 3.4.3. I expect that planets form beyond the snow-line, so I initially place the inner-most planet in each system at $a = 1\text{AU}$ (a typical location for the snow-line in protoplanetary discs; e.g., Garaud & Lin 2004). Beyond this I place the remaining planets in their observed order, with separations motivated by an oligarchic spacing argument (e.g., Kokubo & Ida, 1998). In this scenario protoplanets form with separations of $5\text{--}10r_H$, where r_H is the mutual Hill radius of the two planets

$$r_H = \left(\frac{M_1 + M_2}{3M_\star} \right)^{1/3} \frac{a_1 + a_2}{2}. \quad (3.5)$$

I note, however, that for the masses considered here this typically places planet pairs inside the 2:1 mean-motion resonance. Observations suggest that many planet pairs are captured into this resonance,

and as I expect most of my planets to undergo convergent migration I instead require a somewhat wider initial spacing. I assign initial separations by randomly sampling a Gaussian distribution, with mean $30r_H$ and standard deviation $5r_H$. This results in most, but not all, planet pairs starting off outside the 2:1 resonance. This setup is essentially the simplest set of realistic initial conditions that could plausibly reproduce the observed planetary systems. I recognise, however, that my choice of initial conditions is (necessarily) rather arbitrary, and discuss the consequences of these choices in Section 3.4.3. Note also that because the separations, initial phases and turbulent forcing are all randomly chosen, simulations with the same free parameters can have very different outcomes. Consequently a large number of runs is required to characterise the parameter space adequately.

I halt my simulations when one of my criteria are met: i) a planet is ejected from the system; ii) two planets physically collide; iii) the simulation time exceeds 15τ ; or iv) a planet attains a semi-major axis that is less than that of the innermost observed planet in a system. The first two criteria represent failed models (i.e., runs which do not reproduce the observed system architectures). Due to the tight packing and high multiplicity of these systems, typically a large number (approx. 70%) of simulations are ended for this reason. The third criterion (a time-limit of 15τ) accounts for the fact that a small number of runs evolve into configurations where little or no migration occurs (usually where resonant torques dominate over migration); systems which satisfy this criterion have essentially stopped evolving. The final stopping criterion is deemed to be a “successful” outcome, inasmuch as the innermost planet has migrated to its observed position before any ejections or collisions occur in the system. This stopping criterion is somewhat arbitrary, but there is good reason to believe that the migration of super-Earths might be stopped at small orbital radii in real discs (e.g., [Masset et al., 2006a](#)). I discuss the possible physical origins of such a stopping mechanism in further detail in Section 3.4.3.

3.2.5 Simulations

I use the method presented above to model the assembly of 3 systems: Kepler-11, Kepler-32 and Kepler-80. The parameters I adopt for each of these systems are given in Table 3.1. I summarise the properties of each individual system below.

Kepler-11

Kepler-11 is the prototype tightly-packed planetary system, and consists of six super-Earth- to Neptune-size planets orbiting an approximately solar-mass star (Lissauer et al., 2011b). All six (known) planets orbit within 0.5 AU of the host star. The innermost 5 planets are within the orbit of Mercury, having periods between 10 and 47 days, and eccentricities confirmed by dynamical studies of less than 0.02 (Lissauer et al., 2013). These planets also appear to exhibit no mean-motion resonances, although the innermost pair of planets is near to the 5 : 4 resonance. The outermost planet is somewhat anomalous, due to its large separation from the others, and its properties are less well constrained.

I adopt the parameters for this system from Lissauer et al. (2013), who used dynamical fits to the observed transit timing variations (TTVs) to determine the planetary masses. The inner five planets are mass-ordered, with more massive planets orbiting at larger distances from the star, while the fifth planet is less massive. The mass of the outermost planet (Kepler-11g) is not well constrained by the TTVs, and the dynamical analysis of Lissauer et al. (2013) yields only a weak upper limit to its mass. In my models this large value leads to unrealistically rapid migration of the outer planet, so I instead adopt a mass of $8M_{\oplus}$ for planet g, as assumed in the dynamical fitting models in Lissauer et al. (2013).

Kepler-32

Kepler-32 is one of the most compact multi-planet systems discovered to date, with five planets in orbital periods ranging from 0.7–22.8d (Swift et al., 2013). It also exhibits what appears to be an interlocking mean-motion resonance; planets e and b are apparently in (or very close to) a 1:2 resonance whilst planets b and c are similarly close to a 2:3 resonance. The planets are well ordered in planetary radius, with the planets becoming progressively larger further away from the star. I adopt the stellar parameters, orbits and planetary radii from table 3 of Swift et al. (2013). However, in this case the masses of the planets have not been measured directly. I therefore infer masses from the measured planetary radii, using a simple power-law scaling

$$\frac{M_p}{M_{\oplus}} = \left(\frac{r_p}{r_{\oplus}}\right)^{2.06}, \quad (3.6)$$

derived from a fit to the Earth and Saturn (Lissauer et al., 2011a). There are likely to be large inaccuracies in these mass estimates: Lissauer et al. (2013) note that even after reducing their size estimates,

Object	m (M_{\odot}/M_{\oplus})	a (AU)	r (R_{\odot}/R_{\oplus})
Kepler-11 *	0.961	—	1.053
Kepler-11 b	1.9	0.091	1.8
Kepler-11 c	2.9	0.107	2.87
Kepler-11 d	7.3	0.155	3.11
Kepler-11 e	8.0	0.195	4.18
Kepler-11 f	2.0	0.25	2.48
Kepler-11 g	8.09	0.466	3.33
Kepler-32 *	0.57	—	0.53
Kepler-32 f	0.65	0.013	0.81
Kepler-32 e	2.31	0.0323	1.50
Kepler-32 b	5.07	0.0519	2.20
Kepler-32 c	4.17	0.067	2.00
Kepler-32 d	7.74	0.128	2.70
Kepler-80 *	0.72	—	0.637
Kepler-80 f	1.46	0.017	1.2
Kepler-80 d	2.31	0.037	1.5
Kepler-80 e	2.63	0.049	1.6
Kepler-80 b	7.16	0.065	2.6
Kepler-80 c	8.34	0.079	2.8

Table 3.1: Simulation parameters for each of my 3 systems. Units relative to the sun are used for stellar components of systems, units relative to earth are used for their planetary companions. See section 3.2.4 for a detailed description of the origin of these values.

the planets in the Kepler-11 system are all significantly less massive than one would estimate from this relation.

Kepler-80

Kepler-80 (also known as KOI-500) is an extreme example of a tightly-packed system, with five planets in orbital periods that range from just 3.0–9.5d. As with Kepler-32 the planets are well ordered in size, and there is a possible 4-body interlocking resonance of 4 : 6 : 9 : 12 between the outer five planets (Xie, 2013). Parameters for this system are drawn from the NASA Exoplanet Archive (<http://exoplanetarchive.ipac.caltech.edu>), apart from planet f, whose period and radius are taken from Xie (2013). Only radii and orbital periods are available for the planets of Kepler-80, so I derive semi-major axes from the orbital periods (using the provided stellar mass), and again use the scaling relation (equation 3.6) to obtain mass estimates for the planets.

System	S	S/O	S/U	T	C	E
Kepler-11	3270	2089	1181	1	5753	977
Kepler-32	1964	1916	48	153	7140	896
Kepler-80	2106	2031	75	40	6907	987

Table 3.2: The number of runs ending in a particular outcome for each set of 10,000 runs. S: Runs that finished without a collision or ejection event. S/O: Subset of S that finished with the planets correctly ordered. S/U: As S/O but for incorrectly ordered planets. T: Total number of runs from S that were stopped due to running for longer than 15τ . C: Total number of runs ending in a collision of 2 bodies. E: Total number of runs ending in the ejection of a body.

3.3 Results

As described in Section 3.2.5, I ran 10,000 models for each of the three systems. Given the chaotic nature of the models, and my arbitrary choice of initial conditions, I initially take a very simple view of the results. I define a model as “successful” if the innermost planet migrates to its observed position or the simulation time reaches 15τ without any ejections or collisions occurring in the system. I further split this group of successful runs into two, depending on whether or not the order of the planets was preserved or not. The numbers of successful/unsuccessful runs for each system are given in Table 3.1, while Figures 3.2 & 3.3 show how the fraction of successful runs for each system varies as a function of the input parameters.

While the overall fraction of successful models is rather low, Figure 3.2 shows that this is essentially an artefact of the large parameter space I consider; in each case there is always a region of parameter space where most of the models are successful. It is immediately clear from Figures 3.2 and 3.3 that my models show a clear preference for stronger eccentricity damping and lower stochastic forcing. The reason for this is clear: the close packing of the planets means that even small perturbations can lead to orbit crossing and subsequent collisions or ejections. A lower level of stochastic forcing reduces the probability of orbit crossing, as does more efficient eccentricity damping. All three sets of models require $\beta \lesssim 10^{-6.5}$, suggesting that modest or low levels of turbulence are required in discs that form compact systems. Similarly, I require the eccentricity damping parameter $K \gtrsim 10^{1.5}$ in order to form tightly packed systems effectively. This is towards the upper end of the range of values found in previous numerical calculations, as discussed above, but is certainly plausible for planets in the Type I migration regime.

3.3. RESULTS

When I consider the migration time-scale τ , however, there is a distinct difference in the results for Kepler-11 compared to the other two systems. My models of Kepler-32 and -80 models show no strong variations with τ : the fraction of successful runs is essentially constant across the range of migration time-scales I consider, and shows only a weak decline for very short migration time-scales ($\tau < 10^4$ yr). By contrast, for Kepler-11 I find a strong preference for $\tau \gtrsim 10^5$ yr, and see essentially no successful models with $\tau \lesssim 10^{4.5}$ yr.

On closer inspection, however, I find that the preference for longer migration time-scales in Kepler-11 is almost entirely due to the ordering of the outer two planets (f & g). Kepler-11 is the only system I consider where the planets are not ordered in increasing size. If I relax my criteria to include models where planets f & g are allowed to switch positions (see the dotted histogram in Figure 3.3), then I again find that the fraction of surviving systems is approximately constant with τ , as for Kepler-32 and -80. This behaviour is again readily understood. In my models planet g is approximately fmy times more massive than planet f, so the planets’ orbits converge rapidly as they migrate. If migration is relatively slow then the probability of capture into a mean-motion resonance is high (see e.g., [Mustill & Wyatt, 2011](#)), but resonant capture becomes progressively less likely for faster migration rates, making it much more likely that planet g “overtakes” planet f. The mass of planet g is poorly constrained by current observations ([Lissauer et al., 2013](#)), however, so it is not clear whether this result is significant. Otherwise, I find that migration time-scales $\tau \approx 10^4$ – 10^6 yr readily lead to successful assembly of these compact systems. This range of migration time-scales is broadly consistent with the predictions of Type I migration models, but in general my calculations do not set strong constraints on the required migration rate.

Figure 3.4 shows the distribution of final semi-major axes for each planet in each system at the end of my successful runs. Also shown are some “best-fit” models, in which the final positions of the planets are close to their observed locations. Generally the systems produced by my models are indeed very tightly packed, with the majority of systems in the Kepler-11 case containing all 6 planets within 1AU. In most cases my models end with the outer 3 planets in the Kepler-11 system exterior to their true positions, but even in these cases the tightly-packed nature of the system is clearly maintained. By contrast, my models of Kepler-32 and Kepler-80 tend to produce systems that are in fact more tightly-packed than the real systems, particularly with regards to the spacing of the first

and second planets. Thus, although the precise orbital configurations of the observed systems are somewhat unusual outcomes, my models show that it is clearly possible to assemble systems of five or more tightly-packed at sub-AU radii through simultaneous, disc-driven migration

3.3.1 Mean-motion resonances

Further insight into my results can be obtained by considering the occurrence of mean-motion resonances in the successful runs. Calculating the resonant argument for multiple planet pairs in thousands of individual simulations is computationally expensive, so I instead perform a simple analysis on the period ratios in the final orbital configuration of each run to establish if planets are in resonance. I take the ratio of periods between each pair of adjacent planets, find the closest integer ratio, and consider it a possible resonance if the actual ratio is within 0.5% of this integer value. I acknowledge that resonant behaviour can be observed even between planets that are $\sim 5\%$ away from the exact commensurability (see e.g., [Raymond et al., 2008b](#)). However, I found that for my low-eccentricity systems the 0.5% tolerance was sufficient, since resonant pairs tend to remain very close to exact commensurability, while non-resonant pairs are typically very far from the nearest first-order resonance. This analysis is performed in each of the final 10 output snapshots from each run (i.e., over 100 yr), and the planets are considered to be in resonance if the nearest integer ratio is the same across all 10 snapshots, and always falls within the 0.5% tolerance.

I find that vast majority of runs end with each pair of adjacent planets in resonance, as shown in figure 3.5. However, non-resonant configurations are not uncommon: for each pair of adjacent planets in Kepler-11, a few percent of runs end with no resonance. In my Kepler-11 models the outer two planets are almost invariably trapped in one of several different resonances at the end of the simulations. This is a result of the strongly convergent migration described above, and partly explains the difficulty my models have in keeping these planets as well spaced as they are in reality. Moreover, this same mechanism explains the difficulty in spacing planets e and f correctly: planet f is completely dominated by its resonant interaction with g, allowing planet e to migrate inwards unperturbed. This in turn accounts for the dearth of resonances between planets e and f seen in figure 3.5 - these being the only two planets that are outside of a resonance in the majority of my runs. However, given the large uncertainty in the mass of planet g (discussed above), the significance of this behaviour is unclear.

3.3. RESULTS

In general, my models of Kepler-32 and -80 show similar behaviour to Kepler-11, with the caveat that in reality a number of (possible) mean-motion resonances are observed in these systems. In both of these systems the period ratio of the innermost pair of planets at the end of my simulations is significantly smaller than in reality (see Fig.3.4), but otherwise my models have some success in reproducing the observed MMRs in these systems. In Kepler-32 34% of successful runs end with planets e and b in a 2 : 1 resonance, and 33% with planets b and c in a 3 : 2 resonance; the 1:2:3 resonance in Kepler-32 occurs 52 times in 1964 successful runs (i.e., 2.6% of the time). In Kepler-80, the interlocking resonance between planets d through c occurs only once in my successful runs, though the individual resonances occur separately many times. The 3:2 resonances between planets d and e and planets e and b occur 49% and 40% of the time respectively, whilst the 4:3 resonance between planets b and c occurs 20% of the time.

The prevalence and population of resonances depends weakly on the model parameters. For Kepler-11 the innermost pair of planets is twice as likely to be found out of resonance for values of β above the median (10.8% of successful runs) than below (5.7%). [Rein \(2012\)](#) similarly found that increased stochastic forcing could reduce the number of observed resonances in systems. This effect is not as pronounced in my results as in those of [Rein \(2012\)](#) however, with the majority of my planet pairs still found in first order resonances even in simulations with higher-than-median values of β . When I consider the dependence on the migration time-scale τ , I see similar behaviour to that found by [Paardekooper et al. \(2013\)](#): faster migration leads to planets crossing first-order resonances such as 2 : 1 and 3 : 2, and instead becoming trapped in small-period ratio first-order resonances such as 4 : 3 and 5 : 4. For instance, in Kepler-32, the innermost pair of planets are found in the 3 : 2 resonance at the end of 22% of successful simulations for values of τ above the median. For values below the median, just 4% appear to be in this resonance, with the corresponding percentage for the 4 : 3 resonance jumping from 7% to 15%. A similar trend is seen in the other two systems. Additionally, I find that τ can affect the percentage of planet pairs that finish the simulation in no resonance. In Kepler-32 simulations with τ larger than the median, 1% of all planetary pairs are in no perceivable resonance, this figure jumping to 5.2% for values of τ below the median. However, it is still clear that my models in general over-predict the occurrence of MMRs, and I discuss this issue further in Section [3.4.2](#).

3.3. RESULTS

Interestingly, whilst the overall prevalence of resonances is similar across all three systems, figure 3.5 shows that there are marked variations in which resonances are preferred. For instance, over 25% of adjacent pairs end trapped in the 2:1 resonance in Kepler-32, with less than half this number ending in the same resonance in Kepler-80. This is particularly curious given the striking similarities between these two systems. My models of both Kepler-32 and Kepler-80 generally favour the 3 : 2 resonance, while Kepler-11 favours instead the 2:1 resonance, with the 3:2 resonance occurring much less often. It is clear that the minor differences between these systems (e.g., in planet mass or initial spacing) can make a large difference to the preferred configurations of resonances in my models.

3.3.2 Stability

A key additional consideration is the orbital stability of the configurations – it is entirely possible that many of my systems will become unstable once the damping and forcing due to the disc is no longer present. Simulations of super-Earth embryos conducted by [Cossou et al. \(2014\)](#) suggest that this indeed may be the case, with the removal of disc-driven damping leading to instability and collisions in their calculations. I performed basic stability tests on the successful Kepler-11 models by taking the final results of each and evolving them using the 4th order integrator described by [Yoshida \(1990\)](#) for 5Myr². I consider models to be unstable if they undergo an ejection or collision during this time, or if they show signs of Lagrange instability. Following [Deck et al. \(2013\)](#), I define Lagrange instability to be a change in the semi-major axis of any planet by more than 5% from its initial value over the course of the 5Myr. Of the 2089 successful Kepler-11 models, 5.03% underwent an ejection or collision event within the 5Myr integration. A further 0.24% were found to be Lagrange unstable. Generally I find that the unstable models are those with little eccentricity damping, since they have higher eccentricities at the beginning of the stability test which can quickly lead to orbit crossing. A small percentage (1.82%) of the models required time-steps which were prohibitively short, and hence were not completed. The remaining 92.92% of the models were stable for the full 5Myr. This rudimentary analysis shows that the majority of my models are stable on \gtrsim Myr time-scales, and consequently that the stability of my models does not affect my results significantly.

²I restrict this analysis to Kepler-11 since I expect the stability of Kepler-32 and Kepler-80 to be dominated by the uncertainty in their planetary masses.

3.4 Discussion

3.4.1 Simulation outcomes & preferred parameter values

Table 3.2 shows the total number of runs that ended in each type of outcome for each system. The tiny fraction of runs which are ended due to the simulation time limit implies that this criterion has little effect on my results, and a manual inspection of the final states of these models reveals that they are generally very close to reaching the normal, position-based stopping criterion anyway. In general, runs which end without a collision or ejection event end with the planets in the same order. Kepler-11 is the notable exception to this trend, with 20.89% of runs ending successfully and a further 11.50% ending without any catastrophic event but with planets f and g having swapped positions. The apparent ease with which these planets switch positions suggests that these systems may not have formed in the orders that they are currently observed, particularly given that planet f in Kepler-11 does not follow the otherwise prevalent trend of mass ordering. Instead the planets could have formed in a different order, with the combination of strong damping, convergent migration and disc turbulence facilitating their rearrangement without violent interactions.

As previously stated, my results set no specific constraints on the migration time-scale required for the assembly of these systems. This suggests that limits on disc parameters are instead set by my allowed values of K and β , and also means that any realistic Type I migration time-scale allows the assembly of these systems via disc-driven migration. This is convenient since it allows a number of potential migration stopping mechanisms to be invoked in ending the migration of the planets (see Section 3.4.3). The eccentricity damping parameter K is much more strongly constrained, with values $K < 10^{1.5}$ almost always ending in a collision or ejection. Stronger damping naturally allows for easier assembly, since the combination of tight-packing and resonant configurations in my results can lead to orbit crossing if left unchecked. In spite of this, the preferred range of K values is broadly consistent with those found in previous studies (see Section 1.5) and thus compatible with typical Type I migration models.

The preferred range of β values in my simulations is $\beta \lesssim 10^{-6.5}$, with larger values usually leading to collisions, ejections or planet reordering. To obtain some insight into the meaning of this value, I calibrate my β parameter against previous magnetohydrodynamical studies of migration in turbulent

discs. For a disc with [Shakura & Sunyaev \(1973\)](#) viscosity parameter $\alpha = 0.007$ and aspect ratio $h = 0.07$, [Nelson & Papaloizou \(2004\)](#) found the RMS of the fluctuating specific torque to be $\sigma \simeq 2 \times 10^{-5}$. In my units this corresponds to $\beta \simeq 6 \times 10^{-5}$. However, the disc aspect ratio adopted by [Nelson & Papaloizou \(2004\)](#) is significantly larger than expected at the sub-AU radii I consider here: for a $1M_{\odot}$ star, a gas temperature of 1000K yields $h \simeq 0.02$ at $r = 0.1\text{AU}$. Since the strength of turbulent fluctuations scales approximately as h^2 , a realistic value of β at these orbital radii is likely to be at least an order of magnitude lower: I estimate that the torque fluctuations in the calculations of [Nelson & Papaloizou \(2004\)](#) correspond to $\beta \approx 10^{-6}$ at $r \approx 0.1\text{AU}$. This is around the upper limit of suitable values in my simulations, and suggests that modest levels of disc turbulence ($\alpha \approx 10^{-3}$ – 10^{-2}) are compatible with the assembly of compact planetary systems by simultaneous migration. I note, however, that in real discs the [Shakura & Sunyaev \(1973\)](#) α -parameter depends on fluctuations in the velocity and magnetic fields (e.g., [Balbus, 2011](#)), while the migration torque varies due to fluctuations in the gas density. Establishing a relationship between the turbulent stresses which drive angular momentum transport and the density fluctuations that give rise to stochastic migration (i.e., between α and β) therefore requires detailed magnetohydrodynamic calculations, and is beyond the scope of this work.

3.4.2 Mean-motion resonances

One of the principle arguments against forming this class of systems (in particular Kepler-11) via convergent migration is that many of the adjacent pairs of planets within them do not appear to be in MMRs. As discussed above, my models have significant difficulty replicating this dearth of resonances. It is clear from figure 3.2 that stronger turbulence does not reduce the incidence of resonances significantly, but instead simply reduces the survival rate of systems. An alternative explanation for the paucity of resonances, especially in Kepler-11, is hence required. Mechanisms that could lead to the breakdown of resonances have already been studied in great detail, largely in an attempt to explain an observed pile-up of planetary pairs that have period ratios slightly larger than the closest MMR ([Lissauer et al., 2011a](#); [Fabrycky et al., 2014a](#)).

[Goldreich & Schlichting \(2014\)](#) have suggested that the paucity of resonances observed in exoplanetary systems may be a side-effect of disc-driven eccentricity damping. However, the presence of

this effect in my simulations appears to have had little effect on the long-term maintenance of resonances. Given that Kepler-11 is estimated to be 8.1Gyr old (Lissauer et al., 2013) and protoplanetary disc lifetimes are of order several Myr, long term dynamical evolution of these systems could play an important role in finalising their architectures. For instance, the excitation of eccentricity in the post-disc phase can cause orbit crossing and close-encounters, leading to the breakdown of resonances (Ida & Lin, 2010). A cursory analysis of MMRs present at the end of my Kepler-11 stability analysis suggests that this effect is not significant after 5Myr for my models.

Several authors have investigated the possibility that the paucity of resonances can be attributed to tidal interactions. Lee et al. (2013) find that tidal dissipation can drive some planets out of resonances, but that the effect is generally not strong enough to account for all resonance breaking. Lithwick & Wu (2012) show that tidal dissipation alone can only explain the complete distribution of period ratios if the damping is unexpectedly strong. They propose instead that planets can be repulsed from a resonance by a dissipative process such as tidal damping acting on the forced eccentricity which is driven by the resonance. Baruteau & Papaloizou (2013) show that planets with periods larger than 10 days are unlikely to be repulsed from resonance by tidal interactions. Instead they propose that the wake generated by a companion planet while the two are migrating in resonance can reverse convergent migration. Another promising possibility was suggested by Chatterjee & Ford (2014), who found that interactions with a disc of planetesimals can lead to migration of pairs of resonant planets. This can in turn disrupt the MMR and leads to the final period ratio between the two planets being slightly larger than that of the initial MMR. Whatever the mechanism, there is clearly a fine balance to be struck between breaking resonances in systems such as Kepler-11, and maintaining them in systems such as Kepler-80.

3.4.3 Limitations of the model

First, I note here that caution should be used when considering the outcome of models with low τ and high K values (i.e., those with very rapid eccentricity damping). The simulations of Cresswell et al. (2007) suggest that eccentricity damping occurs on several tens of orbital time-scales, while my formalism assumes only that all effects happen on time-scales greater than the orbital period (see section 2.2.2). For models in which the outermost planets are initially placed far from their host star,

extreme values of low τ and high K can result in an initial eccentricity damping time-scale of less than one orbital period. However, this occurs only in a minority of my models, and even my strongest eccentricity damping acts on times-scales no smaller than $\approx 10\text{yr}$ – far longer than the orbital time-scale at the final orbital radii. Hence I still consider these models to be valid on the basis that the most important period of migration (when the planets are most tightly spaced) is unaffected.

Naturally, for the sake of this simplified calculation, I have made several assumptions which are not valid for actual planetary systems. For instance, the parametrization used here is evidently not self-consistent; previous studies such as those discussed in Section 1.5 have shown that eccentricity damping and migration time-scales are not independent. Given the uncertainties in these values however, it is not possible to perform a more realistic calculation at this time without reverting to full hydrodynamical models. In spite of this, it is encouraging that the range of τ and K values for which my models are successful is broadly consistent with the results of these more sophisticated calculations.

The initial conditions in my models have been chosen on the basis of simple physical arguments, but the planets considered here most likely did not form concurrently with their current masses. According to the standard core-accretion model of planet formation, these planets would have spent considerable time as lower mass planetary embryos, accreting material from a protoplanetary disc (Raymond et al., 2013) and even merging with other embryos. In particular, several planets in Kepler-11 have bulk densities which are consistent with a gaseous atmosphere component, while the planets themselves are sub-Jupiter mass. This suggests that they underwent a phase of relatively slow accretion from the gas disk (Lissauer et al., 2013). If the planets formed at different times, then one would expect that the innermost planets would have formed before their further-out counterparts and hence begun migrating first. This sort of sequential migration could vastly alter the final spacing of the planets, and therefore the landscape of resonances seen in Figure 3.5. Allowing for different formation times and periods of slower migration with lower planetary masses may alleviate some of the difficulties with spacing the planets in my models. Furthermore, I have seen in my results for Kepler-11 that the order of planets in such systems is not necessarily fixed, and allowing the planets to begin in different orders may alter the preferred region of parameter space or the resulting distributions of period ratios and semi-major axes.

Stopping

The largest simplification in the calculations presented here is that planet migration is halted arbitrarily. In reality it is not clear how Type I migration comes to an end, or indeed why it should end so much closer to the host star for these planets than in other planetary systems. There are, however, two mechanisms that are commonly suggested to halt Type I migration, and the large range of possible migration time-scales allowed in my models suggests that either of them may be viable.

The first of these mechanisms is disc clearing (see, e.g., [Alexander et al., 2014](#), and references therein). Protoplanetary disc dispersal is driven by processes (disc accretion and mass-loss due to winds) which are largely independent of planet formation, and removal of the disc gas inevitably halts disc-driven planet migration. The role of disc clearing in halting Type II migration has been studied in detail by various different authors (e.g., [Armitage et al., 2002](#); [Alexander & Pascucci, 2012](#)), but the implications for the Type I regime are largely unexplored. Typical protoplanetary disc lifetimes are 1–10Myr, but final disc clearing occurs rather more rapidly (in $\sim 10^5$ yr). Once disc clearing begins, material at sub-AU radii is simply accreted on to the star, resulting in an exponential decline in the disc surface density (on time-scales $\approx 10^5$ yr) which simply “strands” migrating planets at their current positions. This is therefore a plausible mechanism for halting relatively slow Type I migration (i.e., $\tau \gtrsim 10^4$ – 10^5 yr), but is difficult to reconcile with shorter migration time-scales. In addition, disc dispersal at sub-AU radii is essentially scale-free, so (to first order) I do not expect disc clearing to alter the architectures of compact planetary systems dramatically.

The alternative to disc clearing is the presence of “traps” in the disc: sharp radial changes in the disc structure which result in locations at which an embedded planet experiences no net torque. Such traps may occur at the edges of of a dead zone in an partially ionized disc (e.g., [Gammie, 1996](#)), where the inner disc is truncated by the stellar magnetic field (typically at a few stellar radii; e.g., [Hartmann et al., 1994](#); [Bouvier et al., 2007](#)) or at the dust sublimation radius (which occurs at $r \approx 0.2$ AU; e.g., [Eisner et al., 2005](#)). This mechanism is more plausible for the shorter migration time-scales that I consider ($\tau \lesssim 10^4$ yr), when the migration phase is less likely to overlap with the end of the disc lifetime. Moreover, as traps occur at specific locations in the disc, they are likely to alter the architectures of migrating planetary systems significantly. [Masset et al. \(2006a\)](#) show via hydrodynamical simulations that this mechanism is effective for planets in the super-Earth to sub-

Neptune regime out to 5AU in an MMSN-style disc. The aforementioned migration simulations by [Terquem & Papaloizou \(2007\)](#) invoked magnetospheric truncation of the disc as a method of halting migration. [Ida & Lin \(2010\)](#) conduct population synthesis models of super-Earth systems and find that a magnetospheric cavity in the disc can halt the migration of planetary embryos as they approach the disc edge at around 0.1AU, leading to multiple super-Earths in short-period orbits. Planet-planet interactions may then break any resonances that have formed. This magnetospheric truncation trap has the advantage of stopping the innermost planet in each system approximately at their observed locations, but it is not clear why only some systems contain such short-period planets. Alternate traps further out in the disc and their evolution with time have been considered in some detail by [Hasegawa & Pudritz \(2011\)](#), showing that different traps move in distinct fashions as the disc evolves. Once a planet is caught in one of these traps, it moves in lock-step with the trap as the disc evolves. This typically occurs on a time-scale much longer time-scale than that of Type I migration, and could allow the planets to survive in the disc until the end of its lifetime (e.g. [Lyra et al., 2010](#); [Bitsch et al., 2013b](#)). As mentioned earlier, the larger planets in my simulations may also experience periods of outward migration, depending on their masses and locations at various points during the disc's evolution. [Cossou et al. \(2014\)](#) show that embryos that are large enough to end up in these zones of outward migration generally end up further out as the cores of giant planets, which suggests that the planets in the systems I consider here spent the vast majority of their lifetimes migrating inwards. Overall it is clear that halting the migration of super-Earths and their progenitors is indeed possible, but further work is required to understand these processes in detail.

Tides

The proximity of the planets in my models to their stellar hosts suggests that tidal interactions could be a significant contributor to their long-term evolution. Tides can not only circularise planetary orbits, but also shrink them (e.g., [Ogilvie, 2014](#)). As a result, it may not be necessary for the innermost planet in each of these systems to have reached its observed position as a result of disc-driven migration. Instead, the planets may have migrated part-way towards their present positions in a disc, with the rest of the migration being a result of tidal dissipation on Gyr time-scales ([Hansen & Murray, 2014](#)). Tidal forces would damp eccentricity in this long phase of evolution, which could in turn affect the results

of my stability analysis. As mentioned above, many studies have suggested that tidal forces may also play a role in shaping the distributions of resonances in exoplanetary systems. However, given the uncertainties in the composition of these planets, it is difficult to conduct a more comprehensive study of tidal effects on them at this time.

3.5 Conclusions

I have performed N -body simulations of the migration of the planetary systems Kepler-11, Kepler-32, and Kepler-80, using parametrised forces to investigate the feasibility of assembling these systems through traditional, disc-driven migration. I find that forming such tightly-packed systems via this method is possible under the right circumstances for realistic disc parameters. Interestingly, the slight difference in architecture between Kepler-11 and the other two systems means that this system favours longer migration time-scales, though the significance of this difference is impossible to ascertain without better constraints on the masses of all of these planets. My results suggest that disc turbulence cannot explain the lack of resonances in compact systems, since even moderate levels of turbulence result in catastrophic disturbances to the systems. Further work is required to understand the lack of resonances in these systems, and also to investigate the impact of more realistic formation and migration models.

3.5. CONCLUSIONS

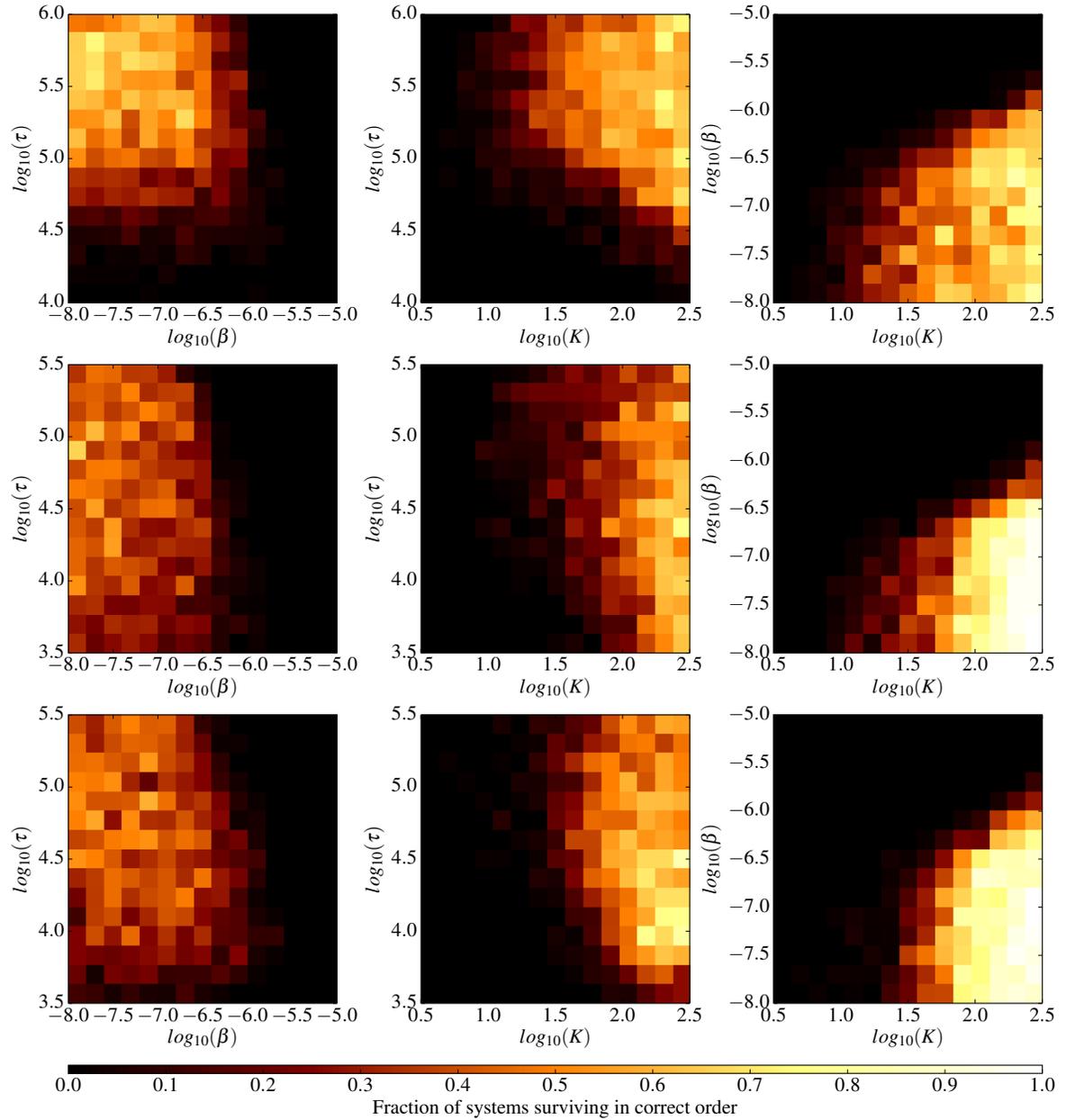


Figure 3.2: Histograms showing success rate for each region of flattened parameter space across all simulations. Top row: Kepler-11, middle-row: Kepler-32, bottom-row: Kepler-80.

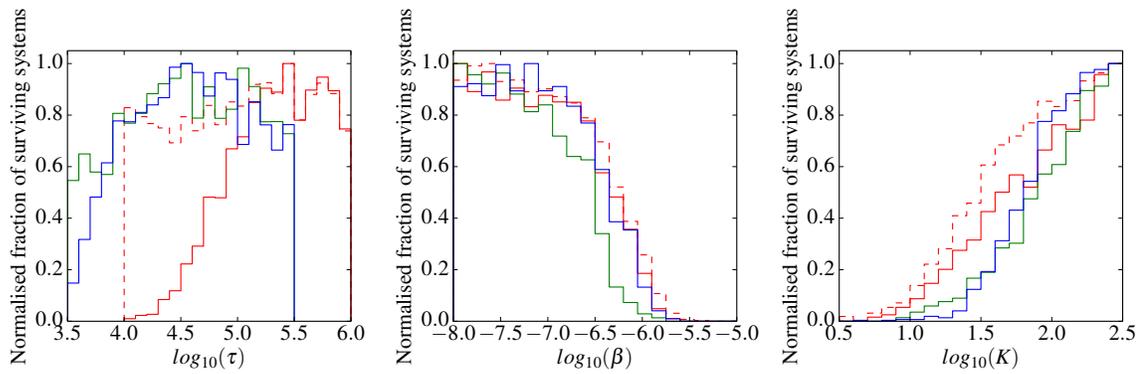


Figure 3.3: Histograms showing success rate for each region of flattened parameter space across all simulations. Red: Kepler-11, dashed-red: as with red line but allowing planets 5 and 6 to maintain order or switch positions with one another, green: Kepler-32, blue: Kepler-80. The Y axis shows the fraction of systems run in each bin that fulfilled the success criteria, normalised to a maximum of 1 for easy comparison. Note that the shapes of the distributions are very close.

3.5. CONCLUSIONS

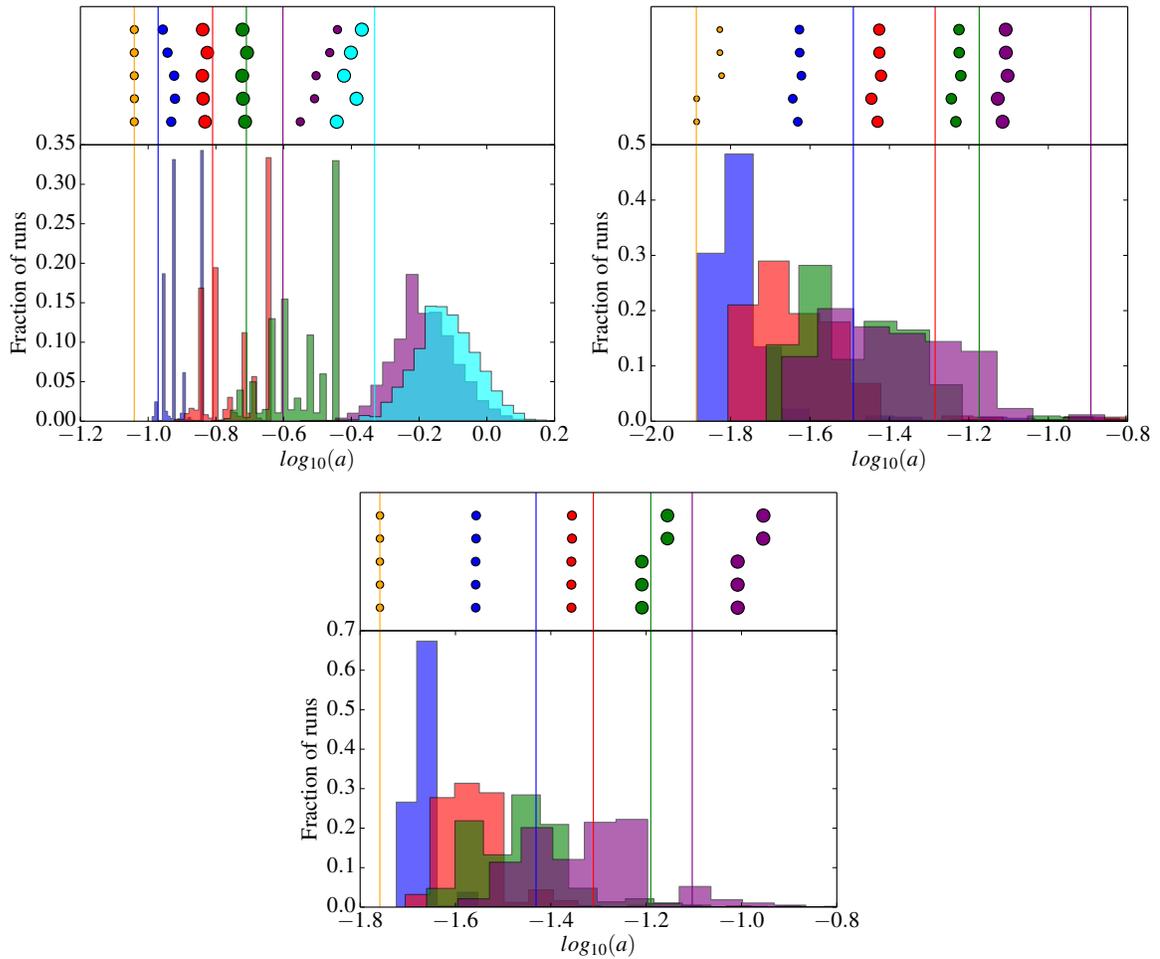


Figure 3.4: Histograms showing distribution of each planet in each system across all successful runs. Each colour represents a specific planet in a specific system. Solid lines show the actual positions of the planets in each system, whilst points plotted above show planetary positions in some representative “best-fit” models, fitted by semi-major axis. Top left: Kepler-11, Top right: Kepler-32, Bottom: Kepler-80.

3.5. CONCLUSIONS

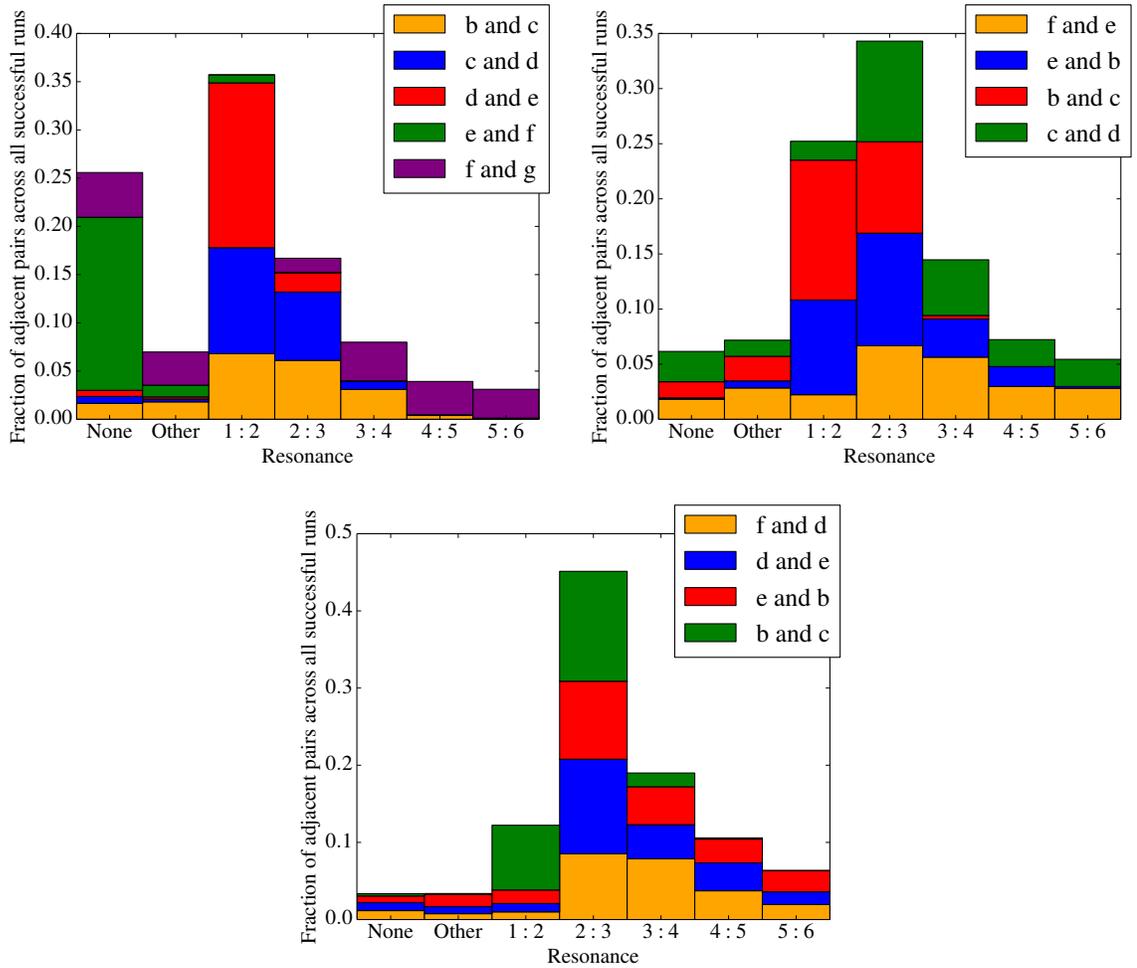


Figure 3.5: Fractions of adjacent pairs of planets that are in each resonance at the end of my successful runs for Kepler-11 (top-left), Kepler-32 (top-right) and Kepler-80 (bottom). Pairs are numbered in order of increasing semi-major axis. The “Other” bar is a sum over all resonances that contain less than 3% of the total number of pairs.

Chapter 4

There might be giants

4.1 Introduction

The recent explosion of results in the field of extra-solar planet detection has revealed several new and distinct populations of planets that are of great interest from a dynamical perspective. In particular, the *Kepler* mission has discovered a multitude of compact systems, each consisting of 5 or 6 planets in the super-Earth to Neptune mass regime and all orbiting within 1AU of their host star (see e.g., [Lissauer et al., 2011b](#); [Swift et al., 2013](#); [Quintana et al., 2014](#)). This class of planets appears to be very common. Both radial velocity surveys and the data from the *Kepler* mission agree in suggesting that $\geq 50\%$ of stars host at least one close-in super-Earth ([Chiang & Laughlin, 2013](#)). Multiplicity is also very common. The sixth *Kepler* data release shows that 1640 (39.3%) of a total of 4175 Kepler Objects of Interest are in multiple-planet systems, with 656 (20.6%) of a total of 3191 candidate systems containing multiple planets ([Mullally et al., 2015](#)). These figures are an increase from the 38.4% and 19.9% reported respectively for these statistics in the previous data release ([Burke et al., 2014](#)), and include perhaps the most extreme example of a compact system yet: Kepler-90. This system contains 7 planets, two of which are gas giants, the outermost of which orbits its host star at 1.01AU and has a radius roughly equivalent to that of Jupiter ([Cabrera et al., 2014](#); [Schmitt et al., 2014](#)).

Radial velocity surveys have revealed a large population of Jupiter-mass objects orbiting their stellar hosts exterior to 1AU (see e.g., [Marmier et al., 2013](#)). The *Keck* survey suggests that between

17–20% of Sun-like stars could host gas giant planets within 20AU (Cumming et al., 2008), whilst *HARPS* finds that 14% of such stars host a gas giant in an orbit of 10 years or shorter (Mayor et al., 2011). In spite of the relatively high incidence of both close-in super-Earths and gas giants, Kepler-90 represents the only known example of a system which contains both of these types of planets. This is likely to be a result of the limited sensitivity of each detection method, with *Kepler* having limited completeness exterior to 1AU and radial velocity surveys being both insensitive to lower-mass planets and unable to perform follow up on dim, distant *Kepler* targets (see Fischer et al., 2014a, for a recent review). It is unlikely that the two populations are mutually exclusive, and future observational campaigns may shed light on the overlap between them.

In my previous work in chapter 3, I considered the possibility of assembling compact systems of super-Earths via disc-driven (Type I) migration. I found that this method can reliably produce systems analogous to Kepler-11 or Kepler-32, albeit with a greater occurrence of mean-motion resonances than in the observed systems. In this chapter I will explore the possibility that some compact systems may contain additional, undetected Jupiter-mass companions orbiting exterior to the known planets. I run a suite of numerical simulations of the assembly of such systems. The premise is that these planets form much further out in the disc and then migrate inwards in the Type I regime as a result of their gravitational interaction with the gas disc. In each simulation I allow one of the outer embryos to undergo runaway growth as in the core accretion model of planet formation (discussed in chapter 1), rapidly accreting gas from the disc and growing exponentially in mass to become a gas giant. I follow the evolution of these systems numerically, and investigate the effects of outer giant planets on the orbital architectures of the observable (inner) planets.

4.2 Numerical method

I follow the method from chapter 3, simulating the migration of multiple super-Earth mass planets using an N -body integrator with imposed migration forces. This method is based on that of Rein & Papaloizou (2009), and imposes an exponential decay of both semi-major axis and eccentricity on each planet, while simultaneously adding a stochastic forcing component in the r and θ directions to simulate the effect of disc turbulence. These forces are controlled by 3 free parameters: τ , which

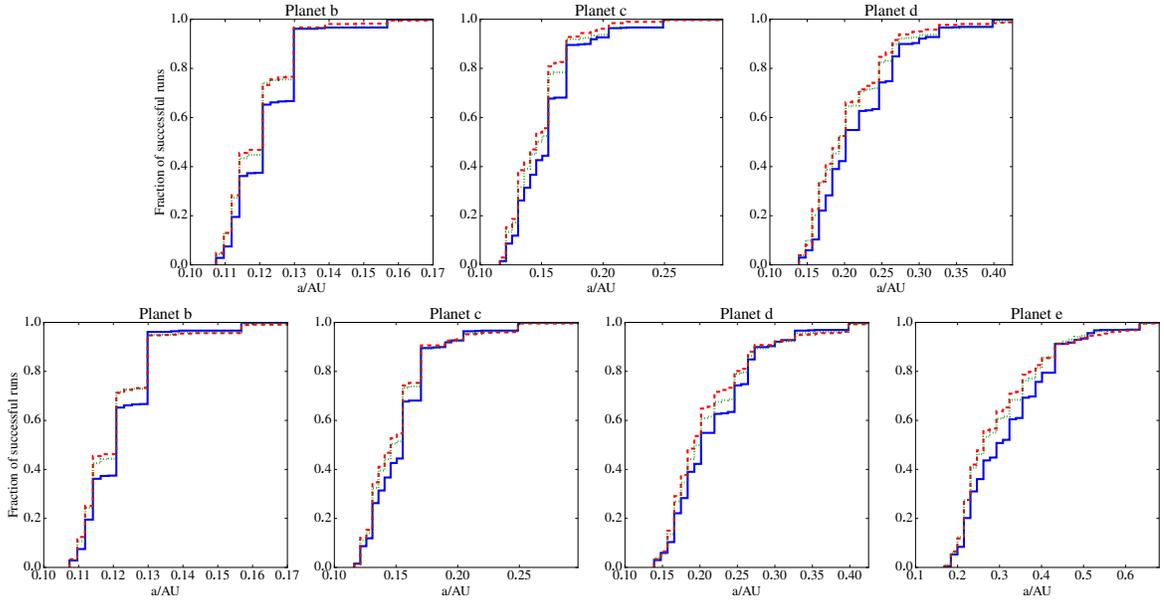


Figure 4.1: Comparison between distributions of final semi-major axes of interior, super-Earth mass planets in the control case with no giant (blue/solid), the case with planet e (top row) or f (bottom row) planet becoming a $1M_{\text{jup}}$ giant at 1AU (green/dotted) and the case with planet e/f becoming a $3M_{\text{jup}}$ giant at 1AU (red/dashed). The linear growth case is not plotted. The interaction with the giant generally allows the lower-mass planets to occupy tighter orbits by breaking widely-spaced resonances between them.

sets the migration time-scale of a planet; K which sets the ratio between the eccentricity-damping time-scale and the migration time-scale; and β , which sets the RMS strength of the stochastic forces relative to the gravitational force exerted on each planet by the host star. Note that τ also scales inversely proportional to the mass of each planet, meaning that higher-mass planets migrate faster, as is expected of the Type I regime. For the purposes of this study I vary τ between $10^{3.5}$ and $10^{5.5}$ yr, K between $10^{1.5}$ and $10^{2.5}$, and β between 10^{-6} and 10^{-8} . A discussion of the physical significance of these parameter values can be found in chapter 3. Note that the overall range of the parameter space is reduced in this work, since in my previous work I showed that relatively high levels of stochastic forcing and relatively low values of eccentricity damping led to almost all compact systems being disrupted during their assembly.

4.2.1 Runaway gas accretion

In addition to parametrized disc forces, I include a simple prescription for the runaway gas accretion, to simulate the rapid growth of a protoplanet from a large super-Earth up to a Jupiter-mass giant (see chapter 1 for more details). This process is an important part of the core accretion theory of planet formation (see e.g. [Helled et al., 2014](#), for a recent review), and begins after a protoplanet has been accreting gas slowly on to its envelope for some time (typically \sim Myr). Once the envelope mass becomes comparable to the core mass the envelope contracts rapidly and matter is accreted on to the planet as fast as local disc conditions permit, until the planet becomes massive enough to open a gap in the disc ([Pollack et al., 1996](#)).

In my model, the planet initially maintains a constant mass M_0 as it migrates inwards. When it reaches some arbitrary distance from the star, the planet mass grows as

$$M = M_0 \exp\left(\frac{t}{\tau_g}\right), \quad (4.1)$$

where M is the mass of the planet at time t after the growth process begins, and τ_g is a characteristic growth time-scale. This time-scale naturally varies with the disc parameters in the vicinity of the planet, with [D'Angelo & Lubow \(2008\)](#) finding from two- and three-dimensional hydrodynamical simulations that this process takes several hundreds of orbits. I hence set τ_g to 10^3 yr. The migration and eccentricity damping forces and stochastic forces for the planet in question are switched off at the point when runaway growth begins. While in reality such a planet would likely continue to migrate inwards in the gap-opening Type II regime, the rate of such migration is sufficiently slow compared to the Type I regime that the movement would be negligible compared to the other super-Earths in the system (e.g., [Baruteau et al., 2014](#)).

Since the migration force is turned off once runaway growth begins, the point at which this process is started essentially sets the radius at which all giant planets in my models will orbit. I select 1 AU as this distance, giving the giants a period of about a year. This radius was selected due to it being the minimum distance at which I would expect *Kepler* to not see the majority of planets, since at this radius even a small inclination will lead to a planet not transiting. This means that all of the giant planets in my results orbit in a small annulus around 1AU, with minor deviations being caused by dynamical interactions with the other planets during the growth process.

4.2.2 Linear mass accretion

As an alternative to the model in which one planet grows exponentially, I test a scenario in which the outermost 4 planets grow from low initial masses - approximately equivalent to the masses of the innermost 2 planets - to their larger, super-Earth masses as they migrate inwards. This ensures that any behaviour seen in my simulations with exponential gas accretion are truly down to the dynamical effect of the giant, and not simply an artefact of my growth prescription. In this case, the mass of each planet increases linearly at a rate

$$\frac{dM}{dt} = \frac{M_{final} - M_{initial}}{5\tau}, \quad (4.2)$$

where τ is the global migration timescale for that simulation and is itself unrelated to the mass of the planet.

4.2.3 Initial conditions

For the purposes of these simulations, I use a model planetary system designed to be analogous to compact systems such as Kepler-11 or Kepler-32. This system consists of 6 planets that are initially in the super-Earth to Neptune mass regime. The masses are based loosely on those of Kepler-11, and assigned such that the planets are mass-ordered from lowest- to highest-mass with increasing distance from the star, hence emulating the configuration of other compact systems. From inner- to outer-most, the planets in this system have initial masses of 1.9, 2.9, 7.3, 8.0, 8.7 and 9.3 M_{\oplus} respectively, and I label them a–f in order of increasing initial mass/semi-major axis. The jump in mass between the second and third planets is similar to that observed in Kepler-11 (Lissauer et al., 2013). Radii are then assigned according to the masses of the planets, using the power-law $M_p/M_{\oplus} = (r_p/r_{\oplus})^{2.06}$ (Lissauer et al., 2011a), although planetary radii are used only for detecting collisions.

The initial positions of the planets are picked in a similar way to in chapter 3, with the innermost planet being placed just exterior to the expected snow-line (at 1.5AU) and each subsequent planet being separated from its neighbours by an oligarchic spacing argument (e.g., Kokubo & Ida, 1998). Pairs of adjacent planets are separated by a random number of mutual Hill radii (r_H), picked from a normal distribution. Here I use a mean of $28r_H$ with a standard-deviation of $5r_H$, selected such that

adjacent planets are in general initially situated exterior to the 2:1 resonance. Initial phases for each planet are selected at random from a uniform distribution, and all simulations are co-planar.

4.2.4 Models

I run 6 distinct sets of models in total. In the first 4 I vary which of the outer two planets undergoes exponential growth into a giant (planet e or f), and the final mass of the giant (1 or 3 M_{jup}). In the fifth set of models, I grow the outermost 4 planets from 2.93, 3.00, 3.06 and 3.12 M_{\oplus} respectively to their final masses (7.3, 8.0, 8.7 and 9.3 M_{\oplus} respectively) using the linear growth model. The final set of models is a control, in which none of the planets grow into a giant. Each set consists of 1000 individual models, with initial positions and phases of the planets being varied randomly between each. Values for each of the three forcing parameters are also picked randomly and uniformly in log space from the prescribed range.

Each model runs until the innermost planet (planet a) in the system has reached 0.1 AU, chosen for its similarity to the semi-major axis of the innermost planet in Kepler-11. Models in which planets collide or a planet is ejected from the system are discarded, as are the small minority of models in which two planets switch positions.

4.2.5 Analysis

For each of the sets of models in which a giant planet was formed, I compare the distribution of the final position of each individual planet to the control case using a Kolmogorov–Smirnov (K-S) test. For each planet in each set of models this yields the probability that the presence of the giant significantly affects the final location of the planet. The results from this analysis are shown in table 4.1.

I also apply a resonance detecting algorithm to each simulation in order to establish the distribution of mean-motion resonances between the remaining super-Earth mass planets in the final systems. This algorithm looks to see if the condition stated in equation 1.73 and discussed in section 1.4.2 is satisfied for 2 planets. I look for evidence of resonant behaviour in the last 20,000 yr (200 snapshots) of each simulation. At each of these snapshots I calculate the period ratio between each pair of adjacent planets, and then find the nearest rational number to this ratio in the form $p/(p + q)$. I impose the

4.3. RESULTS

limit that $p, p + q < 9$ ¹. Equation 1.73 is then used to calculate the resonant argument across the last 200 snapshots with these values of p and q . The algorithm looks for evidence of circulation in the sequence of resonant arguments. If the current sequence has a) a mean between $3\pi/4$ and $5\pi/4$, b) a range larger than 5.75 and c) a standard deviation larger than 1.25, it is considered to be circulating. Once circulation is detected, then the algorithm begins building a new sequence from the point at which the last one ended, again looking for circulation. If no circulation is detected within at least the final 5000 yr of the simulation, the resonant argument is considered to be librating at the end of the simulation and thus the planets are in resonance. The definition of libration and circulation used by the algorithm is necessarily arbitrary and values for all the limits have been tweaked by hand to avoid false positives (due to the low sampling frequency of my output snapshots). In all cases the sampling frequency of the resonant argument is much lower than the orbital frequency of either planet, but examination of the evolution of several hundred resonant arguments by eye suggests that this algorithm returns very few false positives.

4.3 Results

	Planet e, 1M _{jup}	Planet e, 3M _{jup}	Planet f, 1M _{jup}	Planet f, 3M _{jup}	Linear growth
Planet b	2.16×10^{-4}	1.21×10^{-4}	5.85×10^{-4}	1.69×10^{-4}	1.80×10^{-8}
Planet c	3.03×10^{-5}	8.60×10^{-8}	8.26×10^{-3}	3.12×10^{-4}	1.20×10^{-8}
Planet d	2.79×10^{-8}	3.00×10^{-6}	3.74×10^{-3}	2.27×10^{-5}	6.53×10^{-5}
Planet e	N/A	N/A	4.88×10^{-4}	1.96×10^{-6}	8.49×10^{-9}
Planet f	N/A	N/A	N/A	N/A	3.66×10^{-32}

Table 4.1: Probabilities from the K-S test for each case with a giant planet in comparison to the control case. Each value represents the probability that the distribution of the position of each planet across all successful runs was picked from the same underlying distribution.

The results of the K-S tests comparing the distribution of final planet positions between the control case and the cases with a giant are shown in table 4.1. Figure 4.1 shows exactly how these distributions compare to the control in cases where planet e or f becomes a giant, while figure 4.2 illustrates how the presence of a giant affects the final distribution of mean-motion resonances among the super-Earths. It is clear from the K-S test results that the presence of a giant planet during the evolution of compact

¹This limit on the magnitude of p and $p + q$ is imposed to prevent spurious detection of very weak resonances.

4.3. RESULTS

System	S	S/O	S/U	T	C	E
No giant (control)	801	774	27	1	195	4
Planet e, $1M_{\text{jup}}$	760	726	34	0	234	6
Planet e, $3M_{\text{jup}}$	599	568	31	0	360	41
Planet f, $1M_{\text{jup}}$	772	731	41	0	223	5
Planet f, $3M_{\text{jup}}$	705	660	45	1	279	16
Linear growth	912	912	0	1	87	1

Table 4.2: Outcome types for each set of 1000 runs. S: Runs that finished without a collision or ejection event. S/O: Subset of S that finished with no planets having swapped positions. S/U: Subset of S in which some planets swapped positions. T: Number of runs from S that were stopped once simulation time exceeded 15τ . C: Number of runs that ended in a collision between two planets. E: Number of runs that ended in a planetary ejection.

systems can have a significant effect on their final structure. Figure 4.1 suggests that the effect of a giant planet generally tends to push the interior super-Earths into more tightly packed orbits. This figure also illustrates how making the giant planet more massive amplifies this effect, with planets b and c being found on tighter orbits in the $3M_{\text{jup}}$ case than the $1M_{\text{jup}}$ case. The reason for this effect is evident from figure 4.2. Compared to the control case, the cases with giant planets show a lower incidence of widely-spaced resonances (such as 2:1) with a correspondingly higher incidence of tighter resonances (such as 3:4). This suggests that the dynamical effect of the giant is to break the interior super-Earths out of wide resonances, allowing convergent migration to push them into tighter ones, which naturally results in more tightly-packed orbits. The linear growth model on the other hand appears to cause a dramatic increase in the occupancy of the 2:3 resonance, at the expense of both more- and less-tightly packed resonances. The dynamical reasoning for this is clear: having the outer planets spend most of the simulation time at lower masses means they perturb the inner planets to a lesser extent, meaning that once the interior planets are in the 2:3 resonance, they are unlikely to break out.

It is worth considering how the incidence of resonances in my simulations compares to observations. With only one potential example of an observed system similar to those I form, this is naturally a difficult prospect. Nevertheless, [Cabrera et al. \(2014\)](#) note that the three super-Earths in Kepler-90 appear to be close to a 2:3:4 Laplace resonance. A cursory examination of my results shows that the incidence of this particular resonance chain more than doubles between the control case and the case with planet f becoming a $1M_{\text{jup}}$ planet, being present in 3.5% of the successful runs in the former case

4.3. RESULTS

and 8.2% of runs in the latter case. A similar increase in the occurrence of this resonance chain is seen in the case where planet e becomes a $1M_{\text{Jup}}$ planet, with 5.9% of runs exhibiting this behaviour. This suggests that outer giant companions preferentially lead to the formation of resonant chains between super-Earths.

The distribution of the planet neighbouring the giant is always significantly different than in the control case, but interior planets are affected to differing extents. For instance, it is clear from figure 4.1 that the orbit of planet c is significantly altered by planet e becoming a giant, but in the case where planet f becomes a $1M_{\text{Jup}}$ or $3M_{\text{Jup}}$ giant, the distribution of planet c is altered to a lesser extent, with the difference to the control distribution being more exaggerated in the former case. The giants are the same mass and at the same location in all of these cases, the only difference being the addition of an extra super-Earth (planet e) when planet f becomes a giant. This suggests that having an extra super-Earth between a planet and the giant can act to shield the planet from the dynamical effect of the giant. The size of this effect depends upon the final mass of the giant and the position of the perturbed planet relative to the giant, with larger giants having a more significant effect.

The breakdown of simulation outcomes in table 4.2 reveals more about the effect the presence of a giant has on the evolution of the super-Earths. Larger giants naturally result in more collisions and ejections, suggesting that there may be a lower incidence of compact systems with very high mass outer planets, or at the very least a trend for fewer super-Earth mass planets in such systems. I note that the 4 ejections in the “no-giant” scenario are all caused by the innermost planet being scattered by its nearest neighbour into the path of planet c, which is almost 4 times larger than planet a. The encounter between planet a and planet c is then sufficient to push planet a onto a marginally hyperbolic orbit ($e \approx 1.01$). I also note that in a very small minority of cases, the giant planet does not complete its growth before the simulation ends. However, this is only a tiny fraction of my runs ($\approx 2\%$) and the planet is still generally many times larger than the super-Earths, so I do not count this as a separate outcome.

4.4 Discussion

4.4.1 Implications

my results suggest that giant companions could affect compact systems in a similar way to disc turbulence; causing the breakdown of widely-spaced mean-motion resonances, and allowing the formation of tighter ones. This effect was explored in the context of disc turbulence by Rein (2012). I thus suggest that compact systems with tighter mean-motion resonances provide better candidates in searches for giant companions, since the tighter resonances may indicate that a giant has allowed the super-Earths to migrate through more widely-spaced ones. This mechanism may help to explain the formation of systems such as Kepler-36, the two (known) planets of which are near to the 7 : 6 resonance (Carter et al., 2012). Paardekooper et al. (2013) suggested that this may be the result of turbulence in the disc breaking wide resonances and thus allowing convergent migration to push the 2 planets into closer orbits. I propose that an exterior giant companion could provide an alternate formation channel for such systems.

Similarly, it seems that the formation of the 2:3:4 Laplace resonance is amplified by the presence of a giant planet. This leads us to believe that systems of known super-Earths exhibiting such a resonant configuration would also be good candidates in follow-up searches for giant companions.

I also note that a more realistic approach may be some combination of my linear and runaway growth models, allowing all of the super-Earths to grow linearly to their final masses before allowing one of them to undergo runaway growth. Whilst the linear growth scenario cannot explain the formation of extremely tightly-packed systems such as Kepler-36, it does push a significant number of planetary pairs interior to the 2:1 resonance relative to the control case. The combination of this effect with the later perturbation caused by a giant could lead to even higher occupation of very tight resonances.

4.4.2 Observability

As I hypothesise that the giant planets in my simulations would not be seen by *Kepler*, I now estimate what fraction of these planets would actually be detectable. Since my simulations are all co-planar (by construction), I have to make some assumption regarding the inclination of the giant relative to

4.4. DISCUSSION

the rest of the system in order to say whether or not the giant will transit. [Fabrycky et al. \(2014b\)](#) found by comparing transit impact parameters of adjacent planets in compact systems that the mutual inclinations were in the range 1-2.2°. I assume that this distribution extends to my giants also, and therefore assign random inclinations to my giants from a Gaussian distribution with a mean of 1.6° and a standard deviation of 0.6°, such that the entire range suggested by [Fabrycky et al. \(2014b\)](#) is included within 1σ of the mean. The distribution is truncated at 0° and 3.2°. I assume that the giant in each simulation transits if

$$\sin i \leq \frac{R_p + R_*}{a} \quad (4.3)$$

where i is the randomly assigned inclination, R_p is the planetary radius, R_* is the stellar radius (set to R_\odot) and a is the semi-major axis of the planetary orbit. This condition ensures that the giant will be transiting regardless of the longitude of periapsis. Note that since I halt migration once the runaway growth phase begins, all of the giants in my models are at almost exactly 1AU, apart from small deviations caused by dynamical interactions with the other planets. Nevertheless, for the sake of self-consistency, I take the semi-major axes of the giants straight from my models, assign a random inclination, and then determine if they will transit using the above criterion. The results of this test are contained in table 4.3. It is clear that the vast majority ($\approx 99\%$) of these giant companions would be undetectable via transit even with this rather modest inclination distribution. Note however that two factors could change the frequency of transits predicted by this model. Firstly, some of the planets excluded by this simple model as non-transiting would in fact be visible as transits assuming that their longitude of periapsis was such that they passed between the star and the line of sight. Secondly, the radius R_p assigned to the giants using the power-law from section 4.2.3 my $1M_{jup}$ models is approximately 1.5 times larger than the actual radius of Jupiter. Hence a small minority of planets that would nominally appear as grazing transits would not in fact be visible. In spite of these factors, I feel that these figures represent a good estimate of what percentage of giant companions would be visible to transit studies, and suggest that there could be a not-insignificant number of *Kepler* systems harbouring unseen giants.

Using the same inclination distribution, I can also establish if such giants would be detectable by radial velocity (RV) surveys. I use equation 1 from [Fischer et al. \(2014a\)](#) to calculate radial velocities for all simulated planets. An example distribution can be seen in figure 4.3: the low range of

Planet e, 1M_{jup}	Planet e, 3M_{jup}	Planet f, 1M_{jup}	Planet f, 3M_{jup}
0.96%	0.88%	1.23%	0.91%

Table 4.3: Percentage of giants found to be detectable by transit using my simplified transit model.

inclinations gives only a small spread in reflex velocities, and typical signals are $\gtrsim 25\text{m/s}$. For bright stars such a signal is easily detectable via RV observations (see e.g., [Mayor et al., 2011](#)). However, the majority of *Kepler* host stars are much too faint for RV follow-up, and detection would be further hindered by the relatively long time-scales ($\gtrsim 1\text{yr}$) on which the RV signal oscillates. With a dedicated campaign it may be possible to detect “hidden giants” around the brightest *Kepler* host stars, but otherwise they are likely to remain undetected by the current generation of planet-hunting facilities.

4.4.3 Limitations

There are a number of necessary limitations imposed upon these models in order to reduce the vast parameter space to something computationally-viable. The majority of these, such as the arbitrary stopping criterion and an over-prediction of the abundance of MMRs, are discussed previously in chapter 3. However, the introduction of the runaway growth prescription brings about several more parameters worth consideration. The growth time-scale τ_g and final mass of a planet undergoing runaway growth depend sensitively upon the structure and composition of the disc and will naturally vary from case-to-case in reality. A full exploration of the effect that changing these parameters has is beyond the scope of this study, and I believe 10^3yr to be a reasonable estimate of the time-scale at the small orbital radii considered in this proof-of-concept study. The choice of radius at which runaway growth begins (1AU) is also arbitrary, and in this case was chosen to be the minimum radius at which a giant planet might be expected to exist without being detected by *Kepler* as a transiting planet. Thus the perturbations caused by the giants in my models are the maximum effect that one might expect to see in a tightly-packed *Kepler* system, and any real giants might have a less significant impact. Future work could concentrate on how the degree of perturbation changes as the spacing between the giant and the star is changed.

4.5 Summary

In this paper I have investigated the dynamical impact of a giant companion on the formation of tightly-packed planetary systems. A giant planet can break widely-spaced mean-motion resonances and push compact systems into tighter ones, leading to more tightly-packed orbits and to the formation of tight, Laplace-resonant chains. The magnitude of this effect is dependent upon which of the planets becomes a giant, with planets that are nearer to the giant being more strongly perturbed, and also increases for more massive giant planets. I suggest that this could provide an alternate channel for the assembly of *Kepler* systems that are close to tight resonances, and that in turn these systems may prove to be promising candidates in searches for far-out giant companions.

4.5. SUMMARY

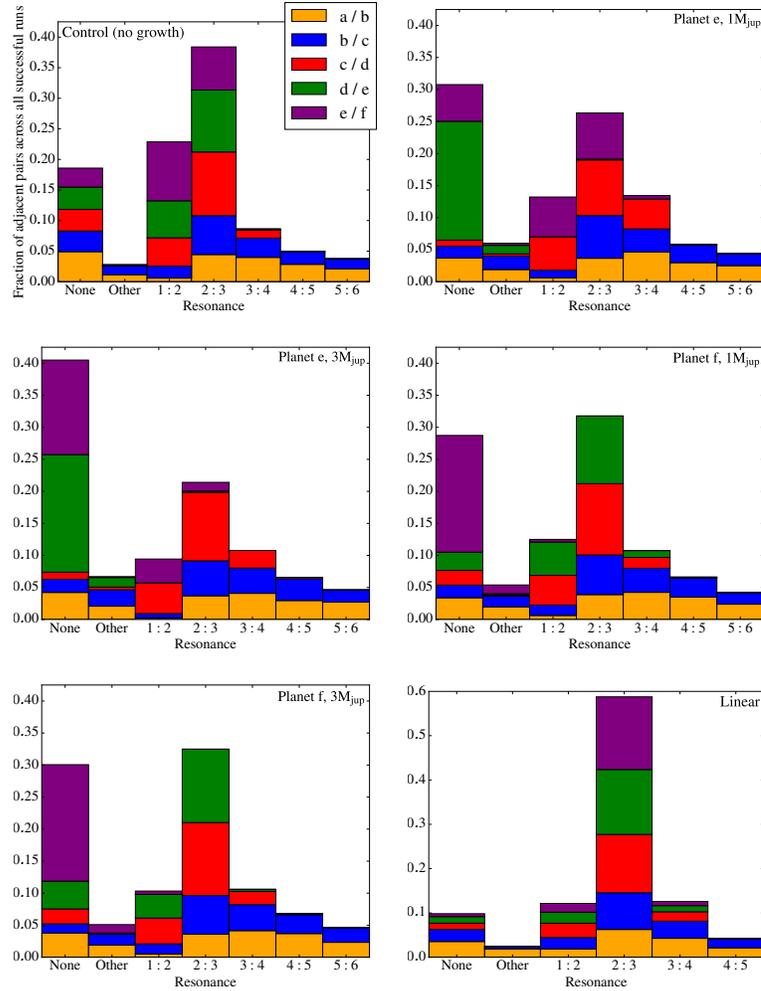


Figure 4.2: Fractions of adjacent pairs of planets that are in each mean-motion resonance at the end of my successful runs. Planets are labelled a–f in order of increasing semi-major axis. The “Other” bar is a sum over all resonances that contain less than 3% of the total number of pairs. The numbers in the key show which two adjacent planets are represented by each colour. The simulations represented in each plot are (clockwise from top-left) the control, planet e becoming a $1M_{\text{jup}}$ giant, planet e becoming a $3M_{\text{jup}}$ giant, the linear growth model, planet f becoming a $3M_{\text{jup}}$ giant and planet f becoming a $1M_{\text{jup}}$ giant. In comparison to the control case, the cases with a giant companion exhibit fewer widely-spaced mean motion resonances, and more tight ones. For instance, note the relative increase in the occurrence of the 3 : 4 resonance compared to the 2 : 3 resonance between planets a and b between the control case and the other cases. Note also that the scale on the y-axis of the bottom-right plot is different to the others.

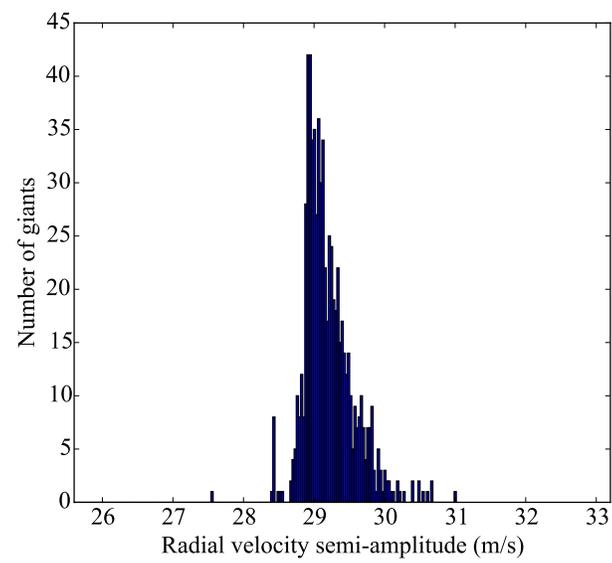


Figure 4.3: Example radial velocity distribution for my models where planet e becomes a $1M_{\text{jup}}$ giant.

Chapter 5

Hydrodynamical simulations of compact systems

5.1 Introduction

My work on type I migration in the previous two chapters has focussed on the use of parametrised forces to replicate the properties of the regime discussed in [1.5.1](#), in particular emulating the mass-dependence in equation [1.98](#). This technique has allowed me to explore a broad parameter space of migration times-scales, eccentricity damping time-scales and turbulent density fluctuations, covering a wide range of potential disc parameters (discussed in section [3.2.2](#)). Covering this parameter space is important since there is great uncertainty in the composition, parameters and even physics in protoplanetary discs, discussed in section [1.1](#). Taking this simplified approach does, however, neglect some of the physics of the problem, and it is worth considering the effect that this physics would have on my previous results.

One of the largest uncertainties in the work presented in the previous two chapters is the lack of a self-consistent way to halt type I migration. I previously suggested that magnetospheric truncation of the inner disc might provide a trap in which the innermost planet might get stuck, halting its migration. Perhaps the most natural way of halting migration, however, is the clearing of the disc. In section [1.1.2](#), I briefly discussed the way in which protoplanetary discs might be cleared by photoevaporation. In this scenario (see e.g., [Alexander et al., 2014](#)), there is a critical radius between 1 and $5a_{\text{au}}$ at which

high energy radiation from the star heats the disc atmosphere sufficiently that a wind is driven out from the atmosphere. This depletes the disc at this critical radius, eventually cutting off the inner disc from the outer disc. At this point the inner disc drains completely and rapidly - in approximately 10^4 - 10^5 yr - onto the star, since it is no longer being replenished by the outer disc. This drainage is caused by the viscosity of the inner disc.

The impact of this effect on type I migration is immediately clear - equation 1.98 suggests that the migration torque scales linearly with disc surface density, Σ . Thus, we expect the type I torque to drop off linearly as the density does. The important question then, is if this clearing process can cause the torque to drop off quickly enough. To obtain some idea, we may compare the time-scale for type I migration at some characteristic radius (equation 1.99) to the viscous time - the time taken for viscosity to change the disc surface density by order unity. The viscous timescale can be approximated as

$$\tau_v = \frac{R^2}{\nu}, \quad (5.1)$$

where R is the orbital radius. In the case of an α viscosity prescription, this becomes

$$\tau_v = \frac{R^2}{\alpha c_s H} = \frac{R^2}{\alpha \Omega_p H^2}. \quad (5.2)$$

If we assume a modest $\alpha = 10^{-3}$ and a disc scale-height of $H = 0.05$, then the viscous time-scale at 1au is $\approx 6 \times 10^4$ yr. Now, assuming a $10 M_\oplus$ planet is at 1AU when the clearing process begins, and using this value $\tau_{\text{clear}} = \tau_v = 6 \times 10^4$ yr, then $\tau_{\text{mig}} > \tau_{\text{clear}}$ if $\Sigma_p < 1000\text{g/cm}^2$ in equation 1.99. This critical surface density is considered reasonable for canonical discs, and will evidently scale inversely and linearly with planet mass. Should we assume a stronger viscosity of $\alpha = 10^{-2}$, the clearing time becomes $\tau_{\text{clear}} = \tau_v = 6 \times 10^3$ yr, and clearing will proceed faster than migration if $\Sigma_p < 10,000\text{g/cm}^2$. Of course, this comparison is only an approximation, but it shows that the time-scales for clearing and migration can be comparable for physical disc models.

When moving to full hydrodynamical simulations, many factors become important that I did not have to consider in the previous chapters. For instance, one must select an equation of state for the gas in the disc, and appreciate the effect that this might have on any torques. In section 1.5.1, I discussed at some length the generation of a corotation torque in discs by material that is - on average - coorbital with the planet. In the case of a sufficiently viscous disc, the torque collapses to a linear

approximation (Goldreich & Tremaine, 1979). In the non-linear, inviscid or semi-inviscid case, the torque is generated by material that executes u-turns and therefore horseshoe orbits in the vicinity of the planet. In adiabatic discs, entropy is conserved during these u-turns, leading to a torque in the case of an entropy gradient (Masset & Casoli, 2009). A similar argument applies to the vortensity in such discs. In globally isothermal discs however, only the torque generated by the vortensity gradient is present.

In locally isothermal discs, the source of the torque is rather less clear (Paardekooper & Papaloizou, 2009). Casoli & Masset (2009) studied the difference between the globally and locally isothermal cases, finding that a supplementary torque exists in the locally isothermal case, caused by vortensity that is generated near to the planet. This torque can be up to one third of the normal, vortensity-generated torque in magnitude, but is difficult to derive analytically. Paardekooper (2014) attacked this issue from an analytical and numerical perspective, whilst D’Angelo & Lubow (2010) derived the fitting formula 1.105 for the total torque in a locally isothermal disc, including both the corotation torque and the Lindblad torques. It should be noted that the models to which this formula was fit used a sufficiently high viscosity that the corotation torque may be the linear variant as opposed to the horseshoe variant.

In this chapter, I will perform hydrodynamical models of locally-isothermal discs with a basic model of disc clearing. The primary purpose of these simulations is to understand how disc clearing affects the the migration of planets in the type I regime, and if it can be invoked as a method to halt migration.

5.2 Method

Hydrodynamical simulations are performed using the hydrodynamical code PLUTO (Mignone et al., 2007), the methodology of which is discussed at length in section 2.3. In order to minimise the computational expense of these simulations, I perform them in two dimensions, using a cylindrical r, ϕ grid. The justification for this two-dimensional set-up and some corrections made to match three-dimensional simulations are discussed below.

5.2.1 Disc model

I initialise the disc in an axisymmetric fashion, with the surface density Σ varying as

$$\Sigma(R) = \Sigma_{1\text{AU}} \left(\frac{R}{1\text{AU}} \right)^{-1}, \quad (5.3)$$

where $\Sigma_{1\text{AU}}$ is the reference surface density at 1AU. I use $\Sigma_{1\text{AU}} = 1000\text{g/cm}^2$ as the fiducial value for all simulations here. The pressure scale height of the disc is similarly set using a power-law:

$$H(R) = \frac{c_s(R)}{\Omega_k} = H_{1\text{AU}} \cdot \left(\frac{R}{1\text{AU}} \right)^q \quad (5.4)$$

where I assume that $q = 5/4$, leading to a disc profile that flares moderately with increasing R . Again, $H_{1\text{AU}}$ is the reference value at 1AU, set to $H_{1\text{AU}} = 0.05\text{AU}$. The disc is locally isothermal in that the profile $c_s(R)$ is constant in time, thus implicitly assuming that the disc radiates and cools very efficiently. The temperature power-law is related to that in c_s by $T(R) \propto c_s(R)^2$. Thus we can obtain the resulting temperature power-law by squaring equation 5.4 and then multiplying through by Ω_k^2 , finding

$$c_s(R)^2 = H_{1\text{AU}}^2 \cdot R^{2q} \cdot \Omega_k^2 = H_{1\text{AU}}^2 \cdot R^{2q} \cdot \frac{GM}{R^3} \quad (5.5)$$

and therefore that

$$T(R) \propto R^{2q-3} \quad (5.6)$$

which for $q = 5/4$ implies $T \propto R^{-1/2}$. In the nomenclature of section 1.5.1, the negative of the surface density power law is $\alpha = 1$ and the negative of the temperature power law is $\beta = 1/2$. Substituting this into equation 1.105 - the D'Angelo & Lubow (2010) scaling for type I torques - we obtain $\Gamma_{\text{tot}} = -2.2\Gamma_0$, where Γ_0 is the scaling torque given by equation 1.98. The negative sign here implies that we would expect each planet to lose angular momentum at every point in the disc, and thus that we expect inward migration with a time-scale that varies depending upon local disc properties.

I note that in chapter 3, I found that migration time-scales between $10^{3.5}$ and $10^{5.5}$ years could easily produce compact systems, and that in section 3.2.2, I found that a time-scale of $\tau_{\text{mig}} \approx 10^{4.5}\text{yr}$ corresponds to a disc of $\Sigma_{1\text{AU}} = 1000\text{g/cm}^2$ with $H_{1\text{AU}} = 0.05$. As such, the disc model here represents a sort of ‘‘median’’ from the parametrised models in chapters 3 and 4, and is thus a disc in which we would expect the assembly of compact systems to be possible. Apart from that, 1000g/cm^2 was the

5.2. METHOD

critical surface density I found using the time-scale comparison in section 5.1, making it the ideal surface density with which to test disc clearing as a way to halt migration.

Prescribing a surface density power-law in the R direction naturally leads to a pressure gradient and therefore a force in the radial direction. This force can be compensated for by introducing a correction to the Keplerian orbital velocity. Lodato (2007) shows that for a disc with power law density $\rho \propto R^{-\alpha}$, the tangential velocity required to compensate both for the pressure gradient and the gravitational pull of the host star is

$$v_\phi = v_k \cdot \left[1 - \alpha \left(\frac{c_s}{v_k} \right)^2 \right]^{1/2}. \quad (5.7)$$

where $v_k = \sqrt{GM/R}$ is the Keplerian orbital velocity in the mid-plane. I apply this correction to the initial conditions in all of my simulations.

Finally, I include no viscosity in the disc, other than the small viscosity which is inherent to the numerical scheme. Whilst it is anticipated that some form of viscosity will be present in protoplanetary discs, the source (discussed in section 1.1.4) and therefore magnitude is not clear. As discussed in section 1.5.2, Duffell & MacFadyen (2013) showed that using minimal viscosity reduces the planet mass required to clear material out of the co-orbital region. As such, I anticipate that the co-orbital torques in my results will be less than they might be in the case of a viscous disc.

5.2.2 Resolution and grid setup

In order to avoid unwanted numerical dissipation, we require that the grid cells in the simulation are approximately square. That is to say, $\Delta r = r\Delta\phi$ for each and every cell. This can be ensured that requiring that the relation

$$\log\left(\frac{r_R}{r_L}\right) = N_r \log\left(\frac{2 + \Delta\phi}{2 - \Delta\phi}\right) \quad (5.8)$$

is satisfied in the R direction, or that the cells are logarithmically spaced in R .

Preliminary tests and previous work (e.g., Paardekooper et al., 2013) suggested that of order 10 cells are required per disc scale height H in order to ensure that the disc dynamics causing type I migration are properly resolved. For the four models presented here, I use a grid that spans from $R = 0.25\text{au}$ to $R = 3.1\text{au}$, and a full 2π in ϕ . The R grid contains 504 logarithmically spaced cells, whilst the ϕ grid contains 1256 uniformly spaced cells. This leads to resolutions of ~ 7.4 cells per disc

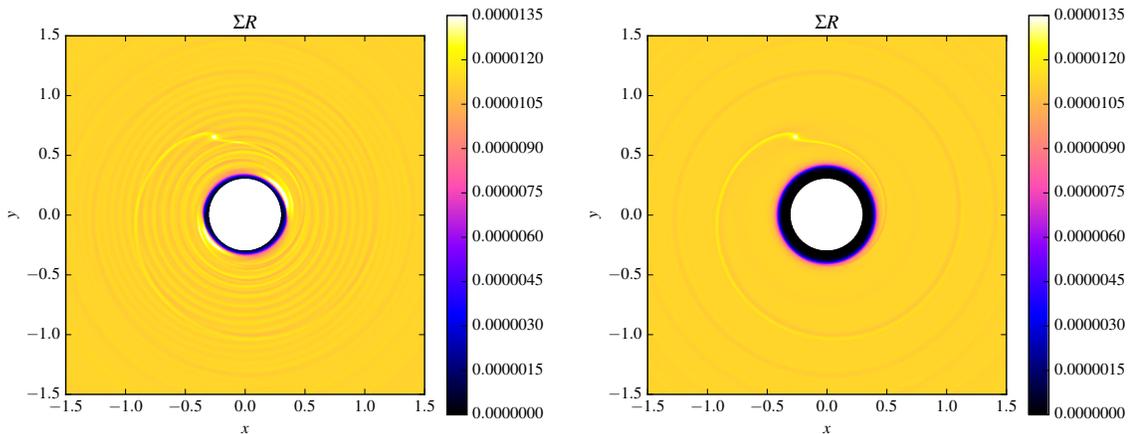


Figure 5.1: . Radius-normalised surface density plots from a simulation without (left) and with the [de Val-Borro et al. \(2006\)](#) wave-damping prescription. The damping suppresses unphysical artefacts in the disc caused by the boundaries very effectively, leaving only the spiral density waves driven by the planet. The trade-off is that the inner part of the disc is completely drained of material.

scale-height at 0.3 AU, ~ 10 cells per disc scale-height at 1 AU and ~ 12.6 cells per disc scale-height at 2.5 AU.

5.2.3 Boundary conditions & wave damping

I chose to use zero-torque boundary conditions for the two radial edges of the computational domain, with the boundary in the ϕ direction being periodic. In principle, this means that mass at the radial edges of the domain can fall out of the computational domain. However, due to the effects discussed below, the majority of this mass is removed before it reaches the exact edge of the domain.

Preliminary runs showed that the unaveraged torques over each orbit became extremely noisy after a few tens of orbits, due to the wake generated by the planet reflecting off the inner and outer boundaries of the simulation and producing spurious torques. Additionally, the lack of equilibrium in the disc at the computational boundaries causes ripples to propagate through the disc, which also have some effect on the initial torques. In order to combat this, I adopt the wave-damping prescription suggested by [de Val-Borro et al. \(2006\)](#). In this prescription, the following equation is integrated at each time-step in each cell belonging to two boundary regions:

$$\frac{dx}{dt} = -\frac{x - x_0}{\tau} X(R) \quad (5.9)$$

where x represents the surface density Σ and the two velocity components in each cell, v_ϕ and v_R . $X(R)$ is a parabolic function that is 1 at the domain boundary and 0 at the boundary of the wave-damping zones. This ensures that waves are damped both smoothly and more strongly as they approach the simulation boundaries, without having too much affect on the rest of the computational domain. I used a simple quadratic in r with the correct scaling for this function. The values of Σ_0 , $v_{\phi 0}$ and v_{R0} provide the “default” values to which the damping attempts to restore each variable in each cell, which I chose to be 0, v_k and 0 respectively. The combination of the quadratic term in equation 5.9 and $\Sigma_0 = 0$ means that the disc settles into a state where the surface density tapers smoothly off to 0 in the wave-damping regions, which serves to shield the rest of the computational domain from the zero-torque boundary condition. As mentioned by [de Val-Borro et al. \(2006\)](#), this approach clearly does not conserve mass or angular momentum, though in practice I found the disruption from this effect to be minimal compared to the spurious waves reflecting off the disc boundaries. An example of the efficacy of this prescription is provided in figure 5.1.

For my four main models, I choose that the inner damping region extends from the inner boundary of the grid at $R = 0.25\text{au}$ to $R = 0.28\text{au}$. The outer damping region extends from $R = 2.8\text{au}$ out to the outer boundary $R = 3.1\text{au}$. The final free parameter here is the wave-damping time-scale τ . I set this time-scale to

$$\tau = \frac{kR}{c_s(R)}, \quad (5.10)$$

which is a measure of the time taken for sound waves to propagate out to radius R in the disc. k is a constant, the exact value of which does not greatly affect the results. For the simulation set-up described here, $k = 5$ was found to minimise noise from the boundaries without removing too much mass from the simulation.

5.2.4 Gravitational Softening

It has been shown that in order to reproduce migration rates from 3D simulations in 2D, the gravitational potential of the embedded planet must be softened. This involves modifying the Newtonian potential (equation 1.44) and acceleration (equation 1.46) such as to avoid the singularity at $r = 0$, and is done here following the standard Plummer softening technique, in which each point-mass particle

is represented as a Plummer sphere (see e.g., [Dehnen & Read, 2011](#)). The potential and acceleration of a point mass then become

$$\Phi_i = -G \sum_{j=0}^N \frac{m_j}{\sqrt{|\mathbf{r}_{ij}|^2 + \epsilon^2}}, \quad (5.11)$$

and

$$m_i \ddot{\mathbf{r}}_i = -G \sum_{j=0}^N \frac{m_i m_j}{(|\mathbf{r}_{ij}|^2 + \epsilon^2)^{3/2}} \mathbf{r}_{ij}, \quad (5.12)$$

where ϵ is a small parameter known as the softening length.

This modification is physically motivated - a real planet is not a point mass and thus gas can get no closer than the planet's surface. More importantly in this case however, the assumption of vertical averaging that allows 2D simulations means that gas that would otherwise be vertically separated from the planet by some portion of the disc scale height H is instead co-planar with the planet, nearer to it and therefore able to exert a stronger force than it otherwise might in 3D. It thus seems prudent to soften the potential of the planet by some factor of H in attempt to emulate the vertical separation from the 3D case.

Selecting the correct value of ϵ must be done in a way that causes the results to closely match those of 3D simulations. Different authors have used different, albeit similar values to achieve this. For instance, [de Val-Borro et al. \(2006\)](#) and [Paardekooper et al. \(2013\)](#) use $\epsilon = 0.6H_p$, where H_p is the scale-height of the disc at the location of the planet. [Fendyke & Nelson \(2014\)](#) use $\epsilon = 0.4H_p$. Smaller smoothing lengths generally lead to stronger torques and faster migration. For calibration, I have compared the results of my models to equation 1.105, derived by [D'Angelo & Lubow \(2010\)](#) for 3D, locally isothermal discs. This is the closest set of 3D simulations to my 2D ones, although their simulations include viscosity and mine do not. As a result of calibrating against this formula, I opted to use $\epsilon = 0.4H_p$. My simulations still tend to produce systematically slower type I migration than [D'Angelo & Lubow \(2010\)](#) found, though this is minimal and is likely as much to do with the lack of viscosity in my simulations as the smoothing.

5.2.5 Disc clearing

Introducing a full model of photoevaporative disc clearing to the code would introduce a host of new parameters and computational complexity that are prohibitive for a proof-of-concept study. Instead

we make the simplifying assumption that we consider only the inner disc - the fraction of the disc that is within the critical clearing radius - and only after the time at which photoevaporation has cut it off from the outer disc.

To model the clearing in the inner disc, we must first understand the way in which viscosity clears the disc. Turning first to the steady-state solution to equation 1.18 (derived earlier in section 1.1.4), we find that the accretion rate is described by

$$\dot{M} = 3\pi\nu\Sigma, \quad (5.13)$$

which we may again rewrite for an α viscosity prescription, finding

$$\dot{M} = 3\pi\alpha c_s H \Sigma. \quad (5.14)$$

The only temporally non-constant variable on the right-hand side of this equation is Σ , and thus we see that the mass change in an annulus is dependent only upon the mass already contained within it. If we were to integrate this then, we would naturally obtain an exponential whereby the mass clears at first quickly, and asymptotes toward 0.

Invoking a source term as the outer boundary is rather unphysical given that the assumption underpinning this study is that the inner disc is cut off from the outer disc. A possibly more relevant solution is the [Lynden-Bell & Pringle \(1974\)](#) self-similar solution. This solution shows that the surface density decays as a power-law in time. Again, the surface density decreases rapidly at first before asymptoting towards 0.

In both of these cases, the surface density declines in a self-similar fashion. For the sake of simplicity, I therefore chose to model the decay of the inner disc using an exponential with the e-folding time as a free parameter:

$$\frac{d\Sigma(r, \phi)}{dt} = -\frac{\Sigma(r, \phi)}{\tau_{\text{clear}}}, \quad (5.15)$$

where I set $\tau_{\text{clear}} = 5000\text{yr}$. The clearing time-scale chosen here of 5000 yr means that 86% of the disc material will be lost within the 10^4 years expected of photoevaporation, this being broadly consistent with the time-scales discussed in the introduction to this chapter, [Alexander et al. \(2014\)](#) and associated literature.

Table 5.1: Parameters for my test runs using my modified version of PLUTO to calculate the torques on one planet starting at 1au.

	$\Sigma_{1\text{AU}} / \text{g/cm}^2$	$H_{1\text{AU}} / \text{AU}$	$M_{\text{planet}} / M_{\oplus}$
control	1000	0.05	3
lowmass	1000	0.05	1.5
highmass	1000	0.05	6
lowsigma	500	0.05	3
highsigma	3000	0.05	3
thickdisc	1000	0.1	3

5.2.6 Modifications to PLUTO

The wave-damping formalism described above was implemented in PLUTO as an “internal” boundary condition - PLUTO’s name for changes to the computational domain made during the enforcement of boundary conditions. All other aspects of the simulation are implemented using standard methods in the stock build of PLUTO. Other than this, I used the modifications described in section 2.3.1 to model N -body dynamics within PLUTO.

5.2.7 Coupled code tests

In order to validate that my modified version of PLUTO reproduces type I migration as described in the literature, I ran a series of test simulations to ensure that the torque scaled as expected in equation 1.98. Using the same numerical setup described above, I altered H , Σ , and the mass of the planet, and compared the torques on this planet over at least 300 orbits to the prediction of [D’Angelo & Lubow \(2010\)](#).

The results of this comparison are shown in figure 5.2. The torque scales largely as expected, with the simulations always being within less than a factor of 2 of the prediction, no matter which parameter is varied. There is a general trend for simulations with higher mass planets to agree more closely with the prediction, which I believe is explained simply in terms of resolution. [Paardekooper et al. \(2010\)](#) showed that the width of the horseshoe region varies with the square root of the planet-to-star mass ratio, and the inverse of the disc scale height. As such given that I use the same resolution in each of these test, the horseshoe region and therefore corotation torque are better resolved in cases with higher-mass planets and thinner discs. So it is perhaps not surprising given that my simulations are in two-dimensions and this resolution factor, that there is a difference between my results and the

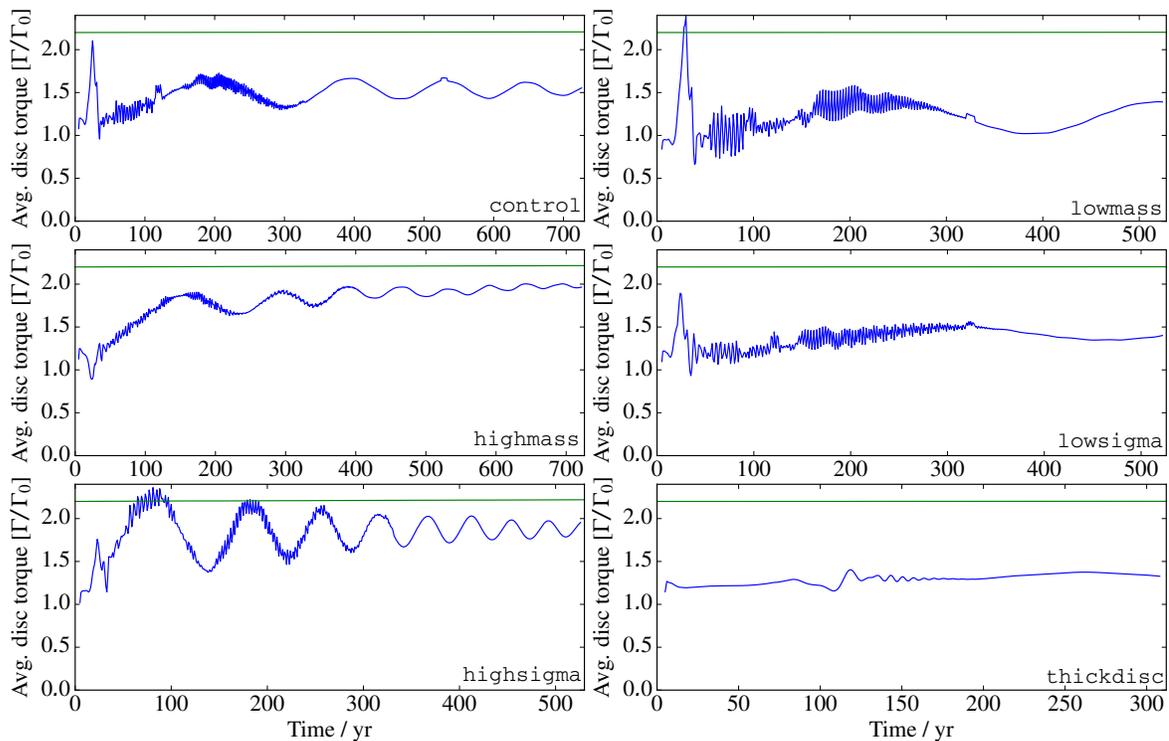


Figure 5.2: Results of tests of my modified version of PLUTO vs. expectations from previous literature. The torque seems to scale as expected with margins of error at the tens of per-cent level. Blue lines denote torques measured from simulations, green lines denote torques calculated from equation 1.105. Parameters are defined in table 5.1. Simulation names are in the bottom right of each panel.

D’Angelo & Lubow (2010) fitting formula. Furthermore, there is a possibility that in this inviscid disc, the corotation torque is become saturated, reducing the torque further.

In summary, the results from these test runs demonstrate that the modified version of PLUTO reproduces the migration time-scales for the type I regime with good accuracy. Further evidence of this is provided by how closely the torques in the simulations described later in this chapter match with the predictions.

5.2.8 Initial conditions

I run four high-resolution simulations of multiple planets migrating inwards together. The models are split into 2 and 3 planet models, and high and low mass planets. The initial conditions for all four models are displayed in table 5.2. The masses in the high-mass cases were selected such that the gap opening criterion P (equation 1.110) is of order a few, and thus the planets are marginally gap opening.

Table 5.2: Initial conditions for planets in all 4 models

	Planets	Planetary masses / M_{\oplus}	Initial semi-major axes / AU
2planlo	a,b	2, 4	1.23, 2.0
2planhi	a,b	5, 10	1.23, 2.0
3planlo	a,b,c	2, 3, 4	0.768, 1.23, 2.0
3planhi	a,b,c	5, 8, 10	0.768, 1.23, 2.0

The masses in the low mass case were selected such that no noticeable gap would be expected. The initial semi-major axes were selected such that each adjacent pair of planets begins the simulation just exterior to the 2 : 1 mean-motion resonance, but also such that the planets begin the simulation within the expected critical radius at which photoevaporation cuts off the inner disc.

5.3 Results

I ran each of the 4 models for as long as computational constraints would allow. For most the models this was between 8000 and 10,000 yr, but due to computational issues, 2planlo ran for only 5000 years. The code makes one dump of the simulation state every 1yr to reduce disc space.

Naturally one of the most important quantities in type I migration is the total torque exerted by the disc on each planet. I use two different methods to calculate the torque. The first uses the instantaneous acceleration from the disc on each planet at each output. This has the advantage of excluding torque contributions from the other planets, but the instantaneous torque is rather noisy and thus requires some averaging. The second simply calculates the change in angular momentum of each planet between two outputs, and this represents the average total torque on the planet between the outputs, from both the disc and the other planets. Having these two different torque values makes it possible to compare the relative significance of planet-planet and planet-disc interactions. Due to the low output frequency of the dumps, some smoothing is required to reduce noise and make the torques comprehensible. All torques are thus plotted as moving averages over 50 outputs, which I note may wipe out some of the finer structure. Figures 5.3, 5.4 and 5.5 show both torque calculations for each planet in each simulation.

To normalise the torques in these plots, I use the quantity Γ_0 (equation 1.98) with values for Σ_p , a_p and H_p being taken at each planet’s location at the beginning of each simulation. For comparison I

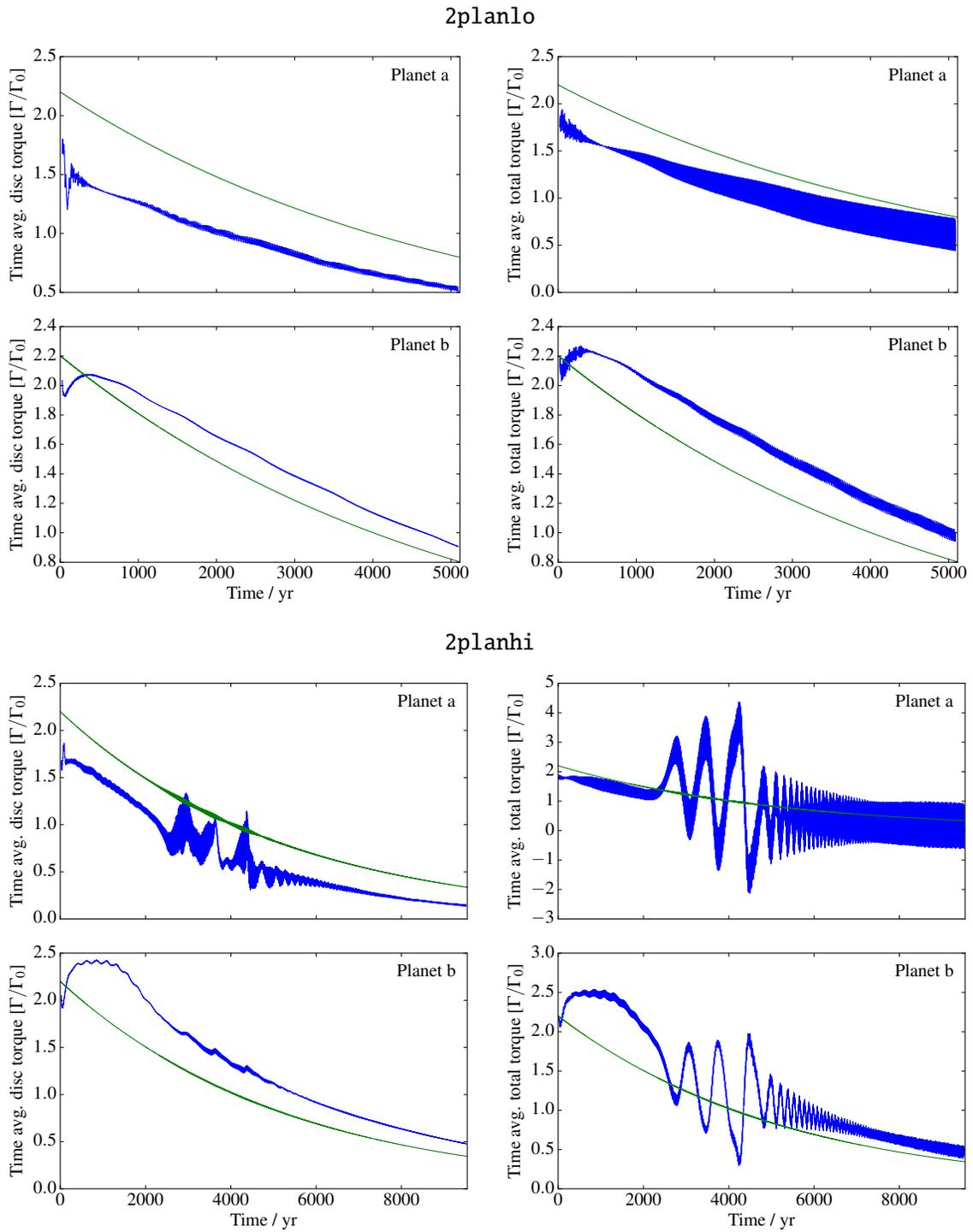


Figure 5.3: Torques acting on each of the planets in 2planlo and 2planhi. Left: torque exerted by the disc on the planet, right: total torque from disc and other planets. The torques are a moving average over 50 yr of output files. Blue lines denote torques measured from simulations, green lines denote torques calculated from equation 1.105. .

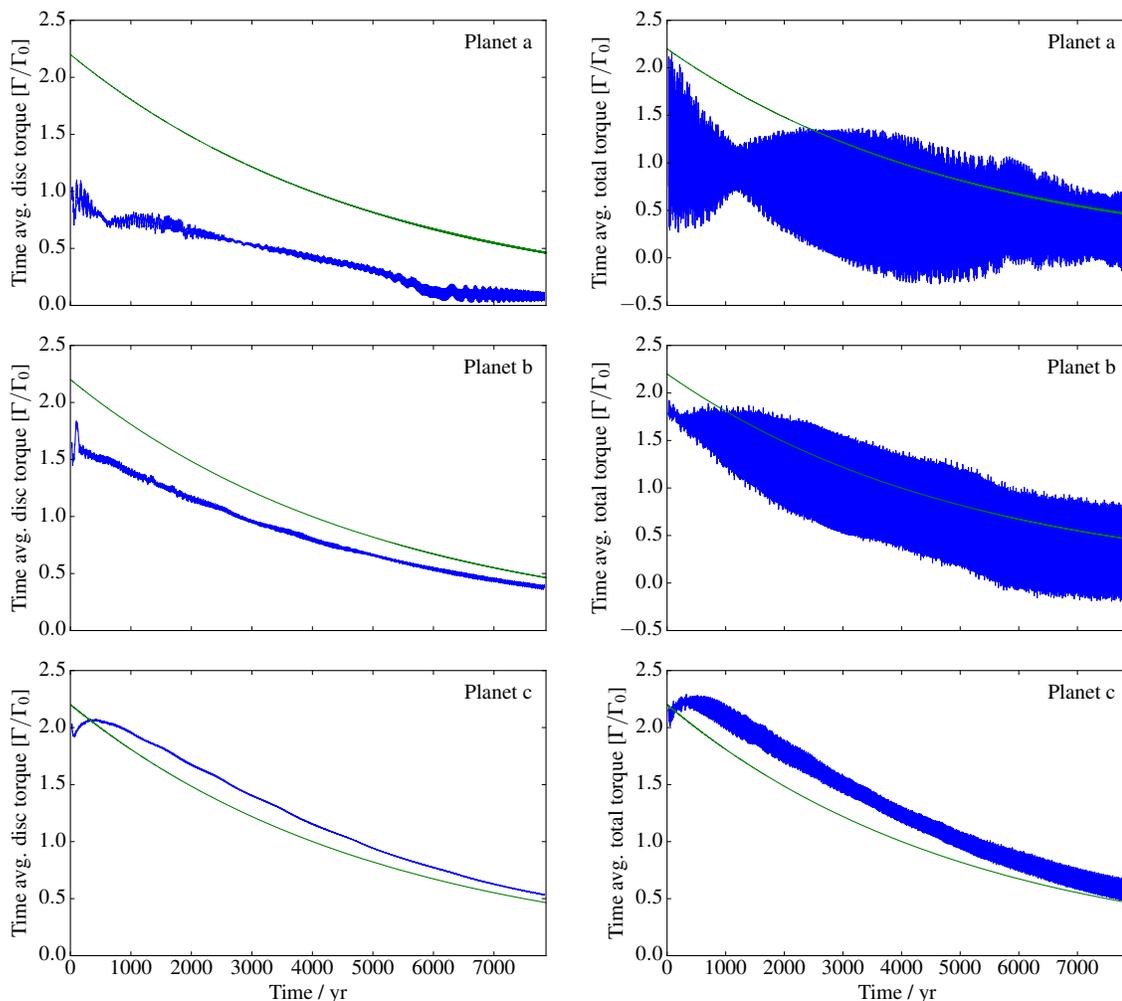


Figure 5.4: As figure 5.3 but for model 3plan1o.

also calculate the torques that would be expected from the [D’Angelo & Lubow \(2010\)](#) scaling formula for locally isothermal discs (equation 1.105).

The second property that is interesting from the perspective of planet-disc interactions is the actual structure of the disc during the simulation, and how the planet observes this structure. To this end, I make plots of the instantaneous, azimuthally-averaged surface density (figures ?? and 5.6) to highlight axisymmetric features in the discs. Note that since the initial surface density profile scales as R^{-1} and the disc clearing is self-similar, this relation holds throughout the simulation, and therefore dividing the instantaneous surface density by R can help to highlight features in the disc.

Maps of the surface density can help to pick out non-axisymmetric features. However, in the case

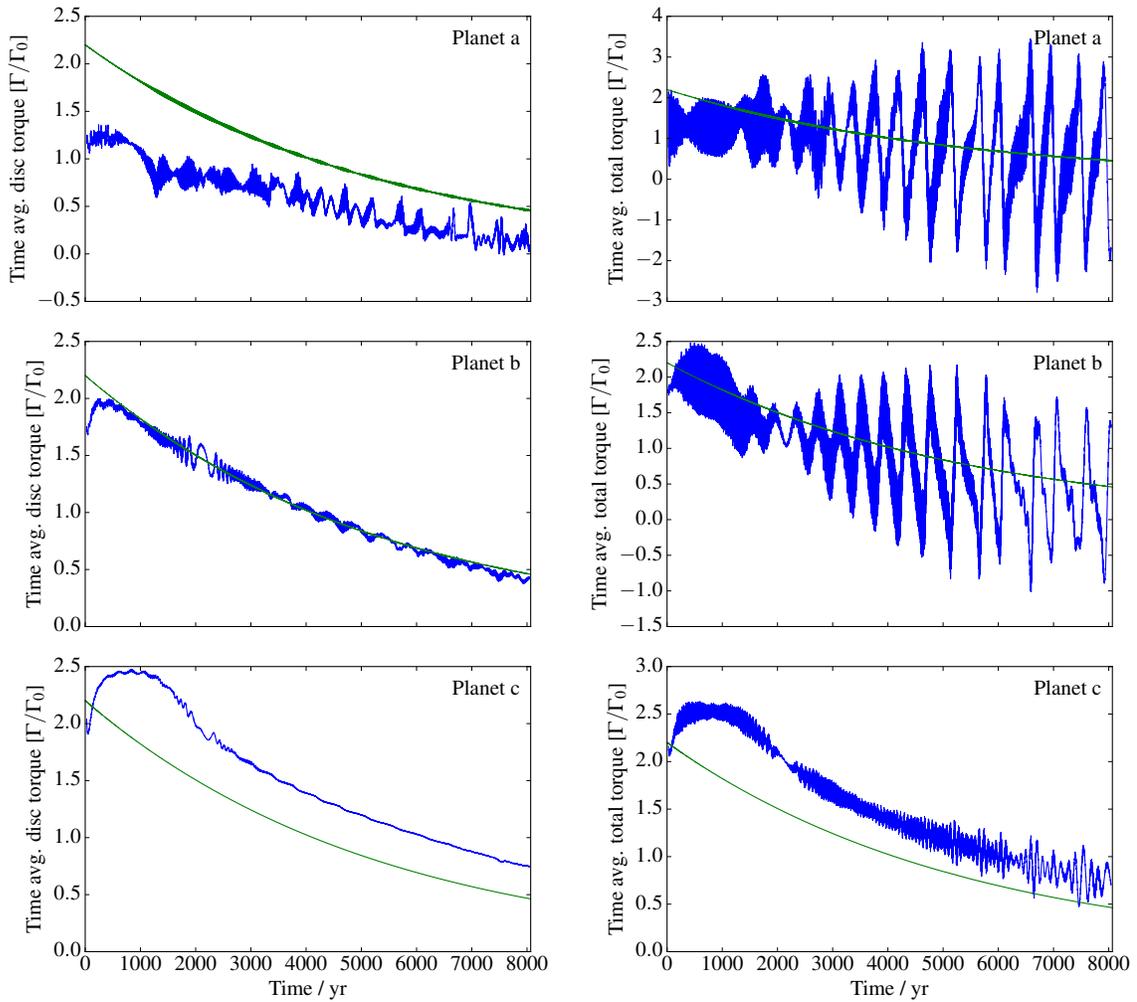


Figure 5.5: As figure 5.3 but for model 3planhi.

of an orbiting planet, features like this that do not orbit in resonance with the planet are likely to be dynamically unimportant. It is thus necessary to understand how non-axisymmetric features move in the frame of reference of each planet. To achieve this, I average the surface density maps over 10 adjacent output files, in a frame of reference that rotates with the planet. In this way, features that are consistent within the planet's frame of reference stand out, whilst others are washed out. These maps, as well as instantaneous surface density maps, can be seen in figures 5.7 and 5.8 for the two high mass cases.

The stand-out feature in the azimuthally-averaged density plots is the partial gaps carved out by the higher mass planets. As anticipated in the set-up, planets of a few P (equation 1.110) clear out

5.3. RESULTS

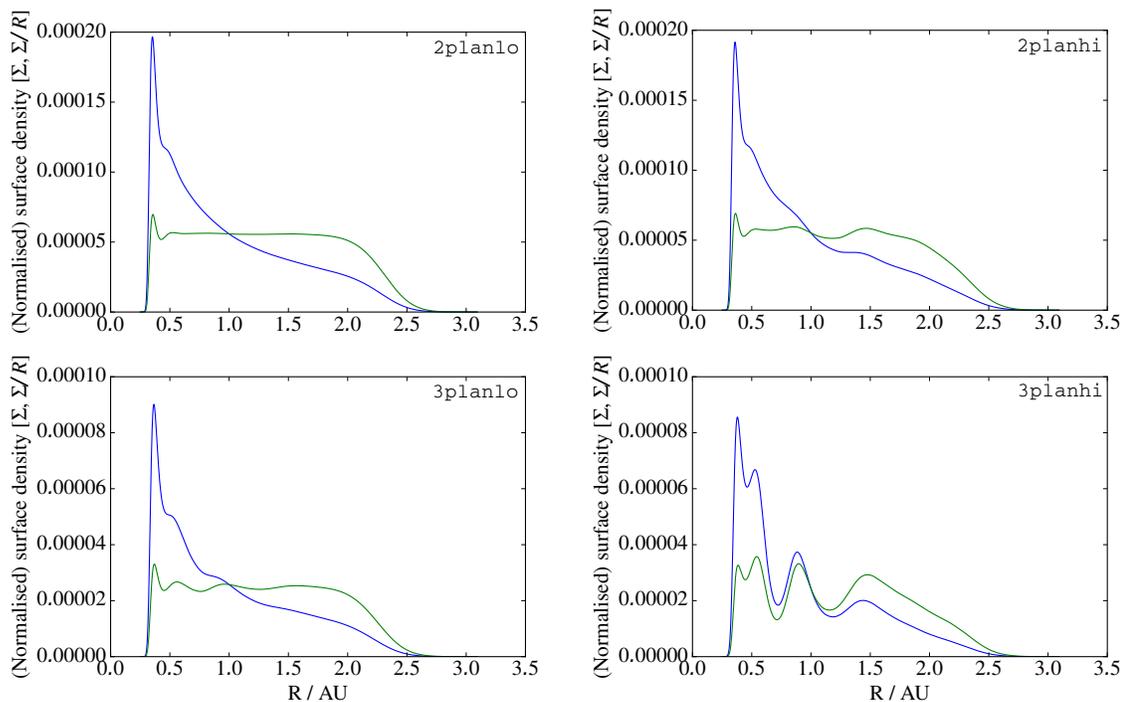


Figure 5.6: Azimuthally-averaged, instantaneous surface densities from 2planlo (top-left) and 2planhi (top-right) at $T = 3500$ yr, and 3planlo (bottom-left) and 3planhi (bottom-right) at $T = 7500$ yr. Blue lines show surface density Σ (M_{\odot}/AU^2), green lines show Σ/R . The planets open partial gaps in their host disc in the high mass cases, but barely make any imprint at all in the low mass cases.

gaps of up to 50% in surface density. Planets in the innermost regions of the disc where H/R is smaller actually carve deeper gaps, in spite of their lower mass. A time strip is plotted in figure 5.9 to demonstrate how the gaps grow deeper as the simulation continues.

In the temporally-averaged, rotating-frame density maps, it is clear that in an average sense, planets only care about their own wakes. In the instantaneous maps, one can see wakes from multiple planets overlapping and interacting, and might naively assume that this modifies the type I migration. However, in the rotating frame maps, the wakes from the other planets are smeared out completely, leaving only the wave generated by the planet in whose frame the map rotates. So, in spite of the interesting resonant behaviour in 2planhi (discussed below), the rotating maps do not suggest that any axisymmetric structures or interactions between the individual planet wakes contribute significantly to such behaviour.

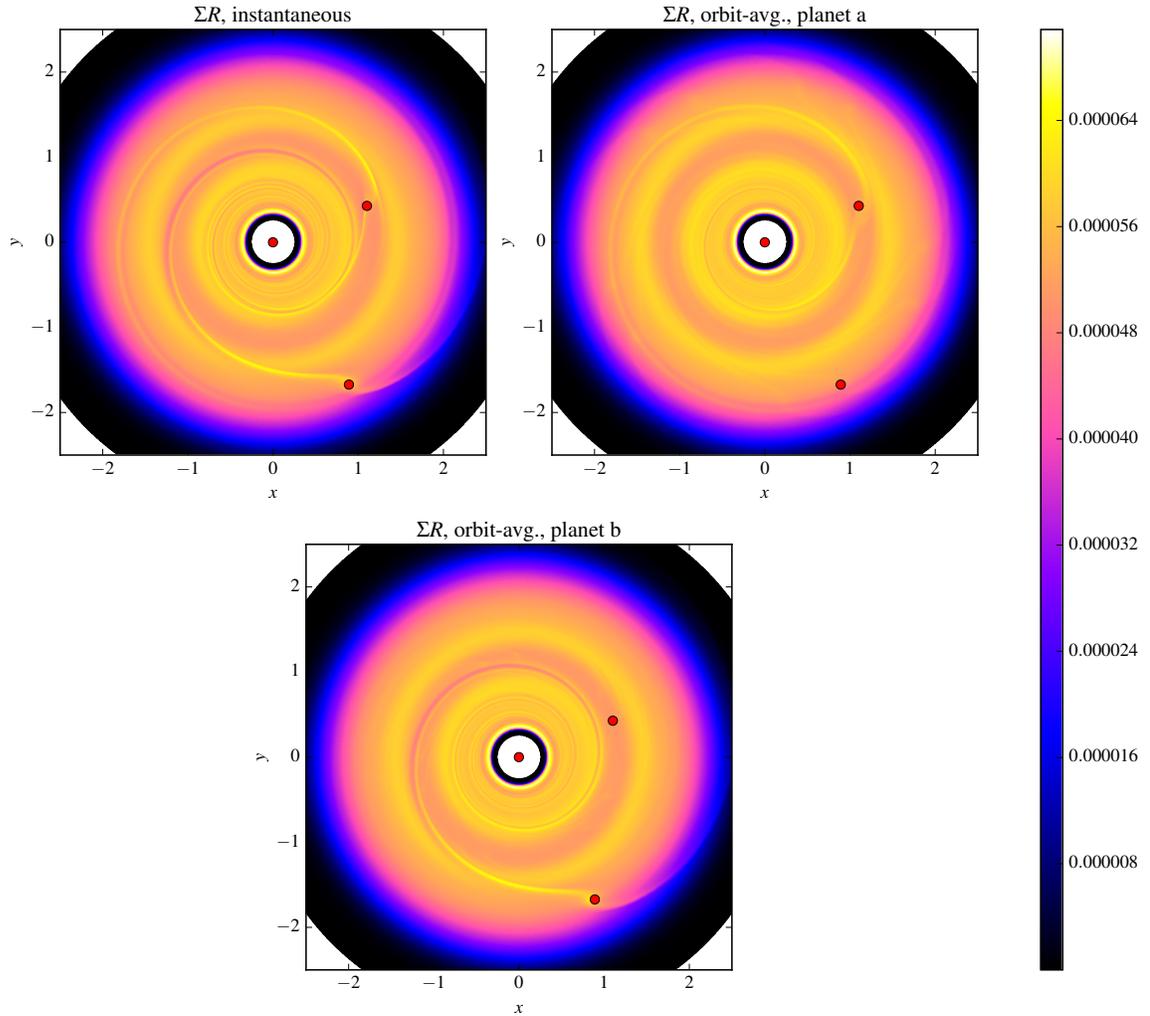


Figure 5.7: Instantaneous and orbit-averaged (per planet) surface density maps from $T = 3500\text{yr}$ in my 2planhi simulation.

5.3.1 Orbital elements

The evolution of the eccentricity e and semi-major axis a (see section 1.4.1) can provide an insight into what exactly the planets are experiencing. The evolution of these elements is plotted in figures 5.10 and 5.11. All planets migrate smoothly inwards throughout the simulation, as expected for the given disc setup. Small eccentricities are damped away exponentially by the disc. However, there are periods of eccentricity growth in three of the models. The cause of this growth is mean-motion resonances.

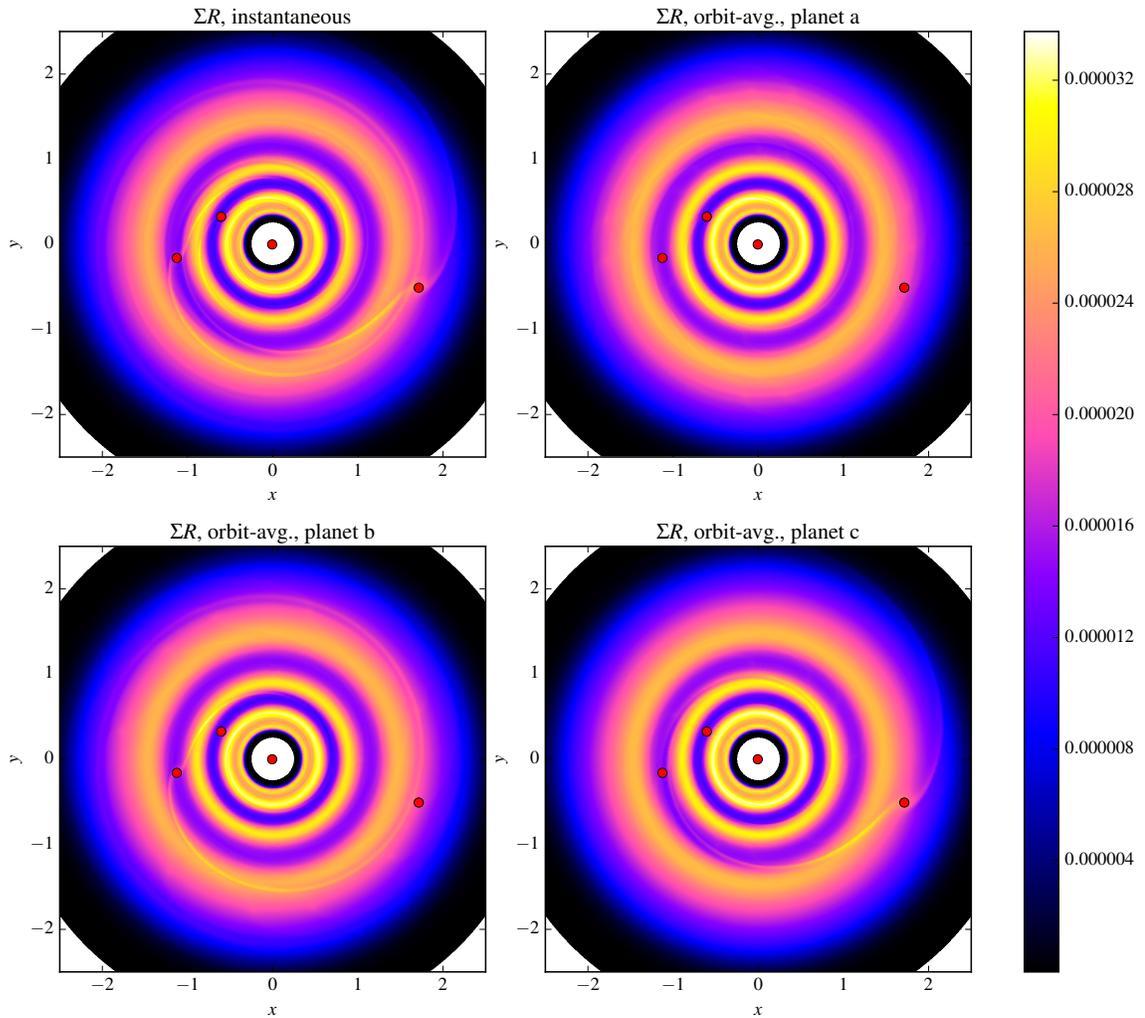


Figure 5.8: Instantaneous and orbit-averaged (per planet) surface density maps from $T = 8000\text{yr}$ in my 3planhi simulation.

5.3.2 Resonances

I search for resonant behaviour in the simulations using the same method outlined in section 4.2.5, and plot the evolution of the resonant argument (equation 1.73). For each pair of adjacent planets, I have plotted the 2 : 1 resonant argument and associated period ratio in figures 5.12 and 5.13. Again, the key to interpreting these plots is that the planets are in resonance if the resonant argument librates, rather than circulates. Given that all pairs of planets were set up on orbits that were close to the 2 : 1 resonance, one might reasonably expect some of them capture in resonance.

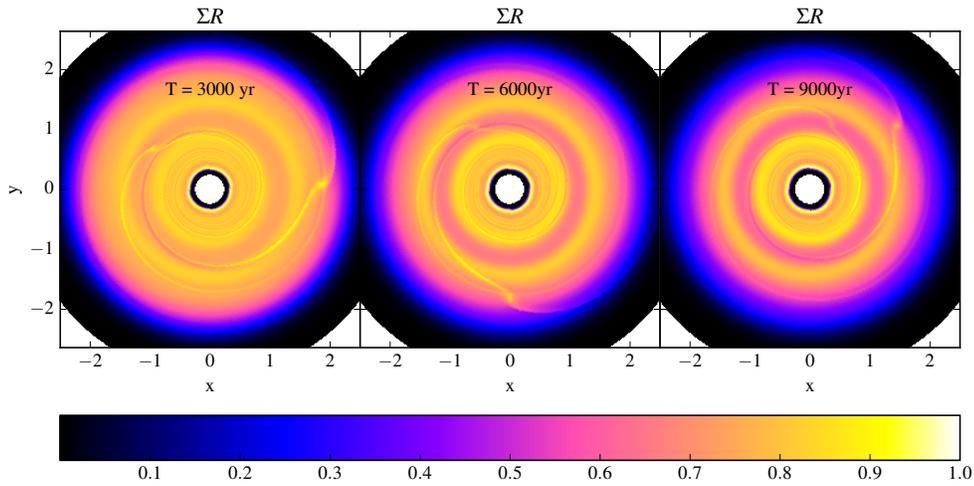


Figure 5.9: Plots of the disc in `2planhi` at three different times. The quantity plotted is ΣR , normalised to a maximum of 1.0. This shows the development of the partial gaps in the disc as the simulation proceeds.

In the `2planlo` case, displayed in figure 5.12, the resonant argument clearly circulates for the entirety of the simulation. The most interesting of the results by far is `2planhi` (again, displayed in figure 5.12), in which the argument circulates for the first 2000yr, then librates until 5000yr, at which point it begins to circulate once more. This 3000yr period is marked by the period ratio oscillating slightly above 2 : 1. The oscillations in the ratio grow in amplitude, and are matched in their own period by growing eccentricity oscillations (shown in figure 5.11). The meaning of this behaviour is clear: at 2000yr, the two planets capture into the 2 : 1 resonance. The resonant interaction pumps causes the eccentricity growth. At 5000yr, some mechanism breaks the resonance between the two planets, and the planet-disc interaction quickly damps the eccentricity.

The total torques displayed in figure 5.3 present a very nice demonstration of the relative importance of mean-motion resonances here. Up until the capture of the 2 planets in resonance, the total torque and disc torque are very closely linked. Once the planets are in resonance, the total torque displays oscillations that far surpass anything coming from the disc alone, indicating that the planet-planet interaction has become as significant in magnitude, if not more significant, than the planet-disc interaction.

5.3. RESULTS

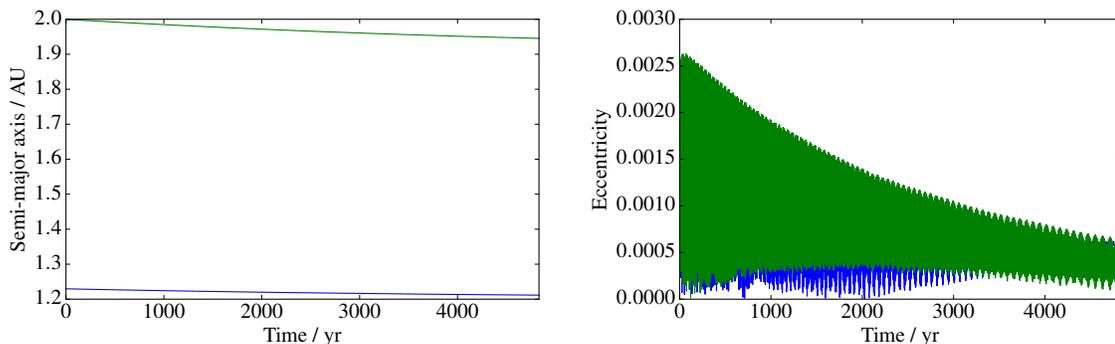


Figure 5.10: Evolution of orbital elements in my `2planlo` simulation. Blue: planet a. Green: planet b. . Since these planets do not capture into resonance, their eccentricity is much more well-behaved than in `2planhi` (below).

Perhaps most interesting is that the torques from the disc (the planet-disc interactions) on the innermost planet are also rather well behaved up until the two planets become resonant, but then begin to oscillate quite wildly as the eccentricity changes. One would expect the total torque to oscillate in this period as the contributions from the resonant partner become non-negligible, but the change in the disc torque is curious. The surface density maps for the resonant period (figure 5.7) do not appear to show any obvious structures forming in the disc that might alter the torque, which strongly suggests that the breaking of the resonance is due to the interplay between the eccentricity pumping and the standard type I torque.

The `3planlo` case also demonstrates resonant behaviour, with the resonant arguments in figure 5.13 demonstrating circulation of the resonant argument for planets a and b up until $T = 4500\text{yr}$, and libration thereafter. This is notably not matched by the period ratio oscillations seen in the `2planhi` case and indeed, it appears as if this resonance is stable in the disc. The eccentricity does grow (figure 5.11), but not by the same magnitude as in `2planhi`, which is to be expected as the planets are lower mass. The eccentricity also does not oscillate and appears to be reaching a steady state in both cases at the end of the simulation - another indicator that the resonance is stable.

Finally, the `3planhi` case demonstrates yet more interesting behaviour, that is qualitatively similar to `2planhi`. In this case, the resonant argument between the innermost two planets, a and b, circulates for the first 1500yr of the simulation. At this point, the two capture into the 2 : 1 resonance, and the period ratio begins to oscillate as in the `2planhi` case. Unlike in `2planhi` however, there does not

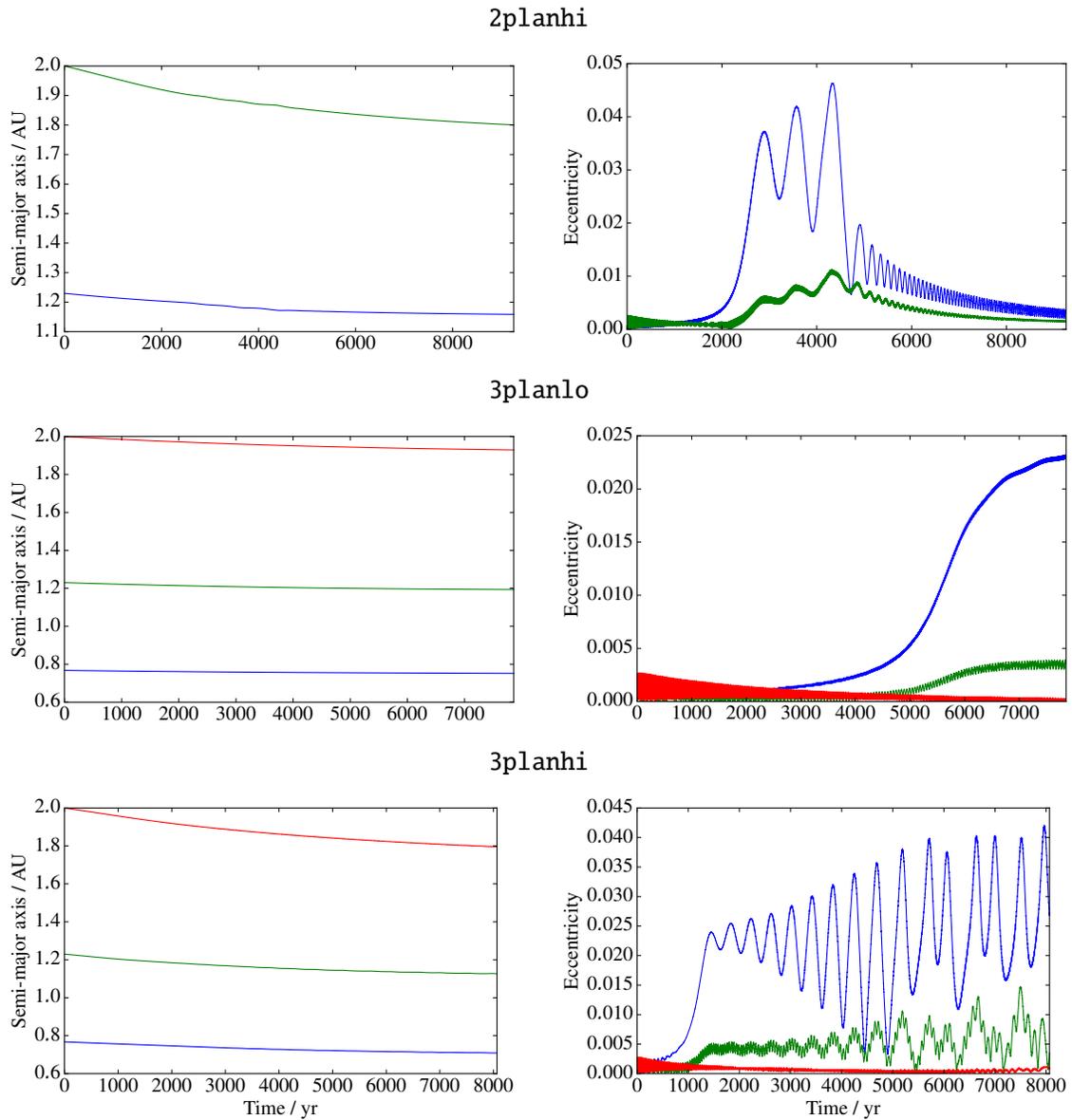


Figure 5.11: Evolution of orbital elements in my 2planhi, 3planlo and 3planhi simulations. Blue: planet a. Green: planet b. Red: planet C. Note in 2planhi the coupled pumping of eccentricity between the two planets between $T = 2000$ and $T = 5000$ yr, while they are in the 2 : 1 resonance. This may be partially responsible for the eventual breaking of the resonance. In 3planlo, planets a and b enter the 2 : 1 resonance at $T = 4500$ yr, at which point their eccentricity grows. However, they do not exhibit the eccentricity oscillations seen in the two high mass cases. In 3planhi however, the capture into resonance at $T = 1500$ yr of the two innermost planets drives eccentricity growth and oscillations similar to the 2planhi case.

appear to be a point at which the resonance is completely broken, although the resonant angle gets rather messy as the oscillations grow in amplitude. The orbital elements in figure 5.11 tell a similar

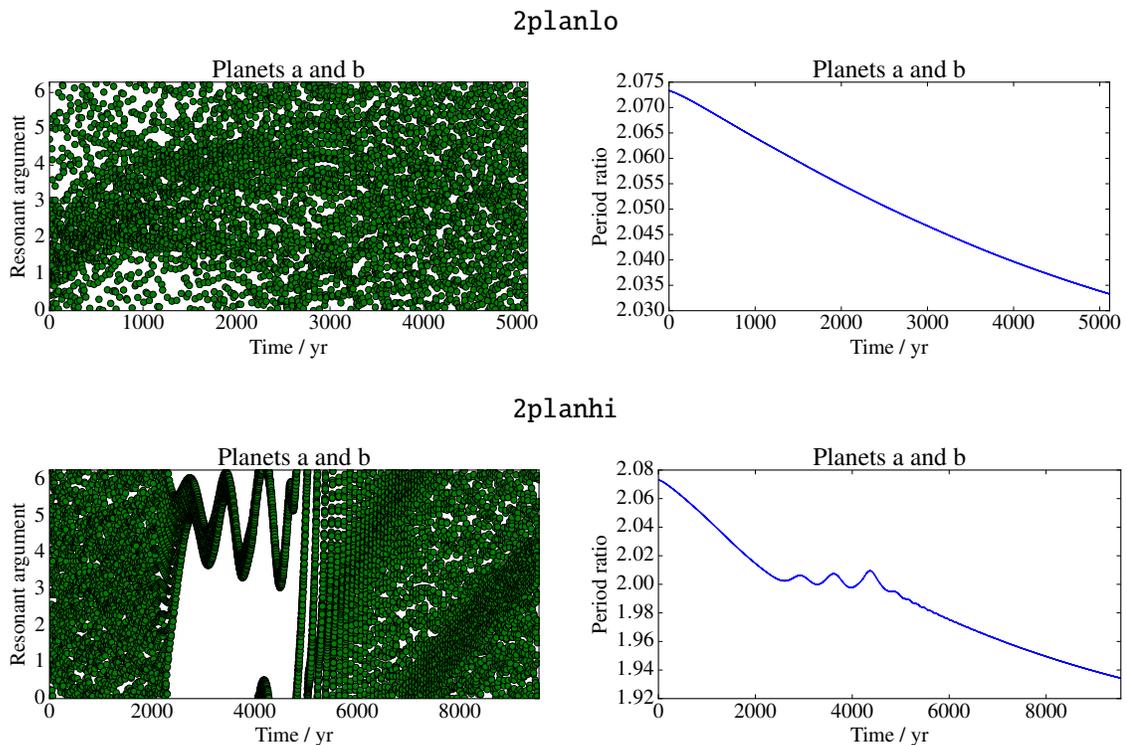


Figure 5.12: Evolution of 2:1 resonant argument (left) and period ratio (right) for the planets in 2planlo (top) and 2planhi (bottom). No resonant behaviour is observed in 2planlo. In 2planhi, the 2 planets appear to be trapped in a resonance from $T = 2000$ yr until they are broken out of the resonance at $T = 5000$ yr.

story, with growing eccentricity oscillations and again, oscillations in both the disc and total torques (figure 5.5).

5.4 Discussion

5.4.1 Halting migration

From the figures of torque presented above, it seems clear that migration is coming to a halt rather rapidly. In spite of the rather high initial surface density and therefore fast migration in the disc setup, the torques seem to be asymptoting. The plots of semi-major axis equally show that the rate of migration is reducing rapidly. To aid visualisation of this, I plot migration time-scales averaged over 500yr periods in figure 5.14, showing a clear tendency for the time-scale to grow exponentially. This figure also quite helpfully demonstrates the noise in migration rate that is generated by resonant

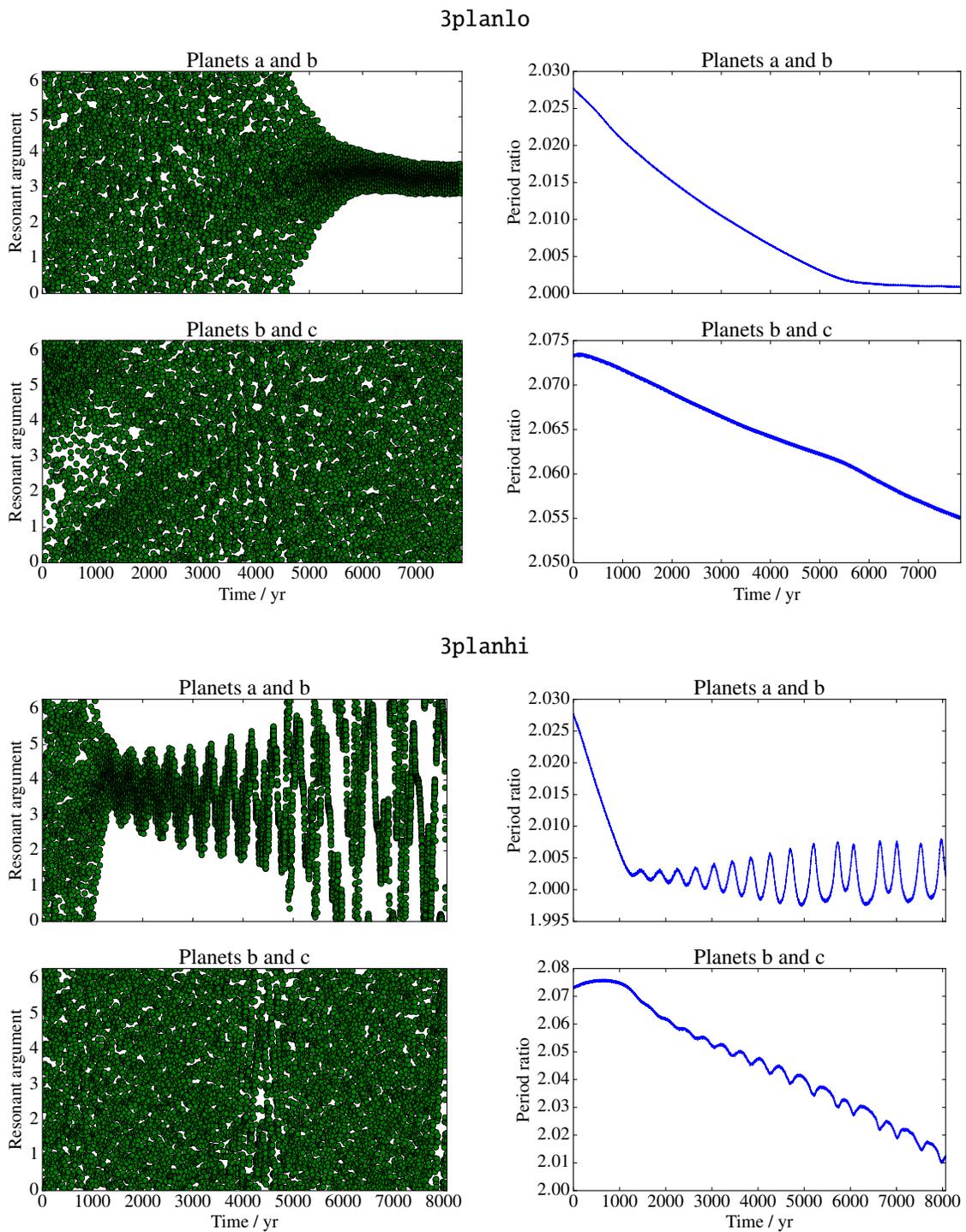


Figure 5.13: Evolution of 2:1 resonant argument (left) and period ratio (right) for pairs of adjacent planets in 3planlo (top) and 3planhi (bottom). The innermost 2 planets in 3planlo appear to be trapped in a resonance from $T = 4500$ yr until the end of the simulation. In 3planhi, the innermost 2 planets appear to be trapped in a resonance from $T = 1500$ yr until the end of the simulation.

5.4. DISCUSSION

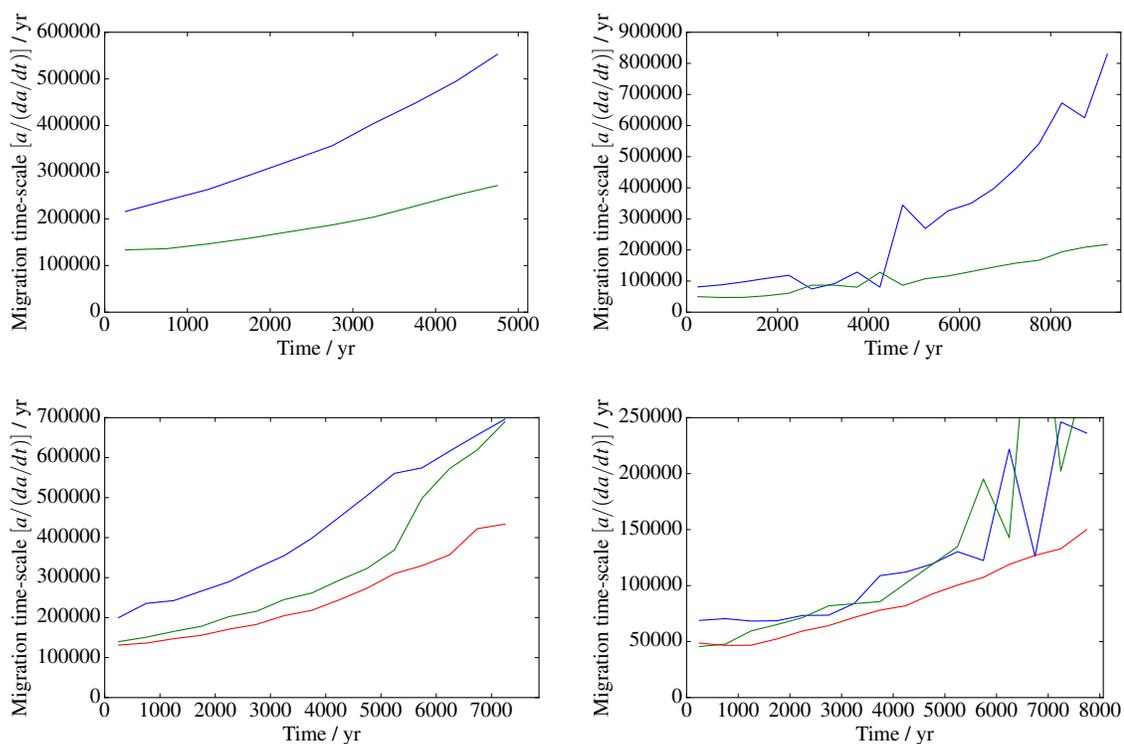


Figure 5.14: Migration time-scales for each planet as a function of simulation time for 2planlo (top-left), 2planhi (top-right), 3planlo (bottom-left) and 3planhi (bottom-left). Blue: planet a. Green: planet b. Red: planet c. The time-scales all appear to be asymptoting towards infinity, but are already an order of magnitude longer than the clearing time-scale at the beginning of the simulation.

interactions - significant even when averaged over 500 yr.

It is clear that for the regime we are in here, there is no doubt that the inner disc will drain before the planets migrate even close to their stellar host, since the migration time-scale is longer than the clearing time-scale by factor of a few even at the beginning of the simulation. As a result of this, it appears that the time-scale argument presented in the introduction of this chapter holds. Based on these plots, one could conceivably multiply the disc surface density in these simulations by a factor of a few without inducing migration onto the host. Obviously there is merit to pursuing more varied disc and clearing models to understand the effect that these factors have on the ability of clearing to halt migration, but it does indeed appear to be possible to stop type I migration in this way.

5.4.2 Disc structure

The planets in the high mass simulations are large enough that they clear partial gaps in their host disc. These gaps take some time to develop, and are at most a factor 2 in surface density. The opening of these partial gaps does not, however, appear to have much affect on the overall migration rate, with the torques remaining consistently close to the [D'Angelo & Lubow \(2010\)](#) prediction even as the gaps develop. This is not surprising - a partial gap affects the corotation torque on the planet, which is already weaker relative to the Lindblad torques in a locally isothermal disc than it would be in a disc with an entropy gradient. Quite apart from that, with an inviscid disc there is always the threat of the corotation torque becoming saturated, meaning that the depletion of this region would matter even less. Whilst it would be interesting to run the same simulations again with a moderate viscosity so understanding exactly how the partial gaps and corotation torques change, it is not immediately obvious that this would have any affect on the results.

5.4.3 Resonant behaviour

The most surprising result is that of `2planhi`, in which two planets spend 3000yr trapped in resonance before being broken out. The reason for this resonance breaking is not immediately clear, but there is evidently a change in the type I migration torque during the period in which the planets are in resonance. These variations in the torques are matched in period by the eccentricity oscillations driven by the resonance. [Papaloizou & Larwood \(2000\)](#) found that the Lindblad torques could actually reverse in direction, promoting outward migration, when the $e > 1.1(H/R)$. [Fendyke & Nelson \(2014\)](#) confirmed this, additionally showing that the corotation torque can be reduced by low eccentricities. The latter study also showed that the Lindblad torques decrease in magnitude before the sign flips.

We can gain some insight into what is going on here by simply considering the unsmoothed disc torques. In the torque figures above, I smoothed the torques over many orbits, since the instantaneous torque on a planet varies over its epicycle. The simulations saved the instantaneous torques every 1yr. At the time when the resonance breaks, the period of planet a in `2planhi` is approximately 1.31yr. Thus, if we average over enough orbits, as in the torque figures above, we might expect to obtain an average torque across an orbit or epicycle. This naturally smooths out any variations in the torques that happen on only a few orbital periods. [Figure 5.15](#) shows the unsmoothed, instantaneous torques

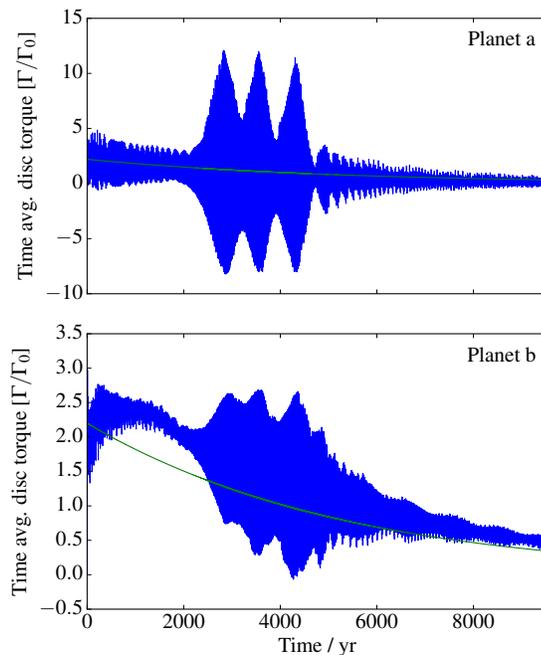


Figure 5.15: Torques acting on each of the planets in 2planhi. Left: torque exerted by the disc on the planet, right: total torque from disc and other planets. The torques are unaveraged, and completely instantaneous. Blue lines denote torques measured from simulations, green lines denote torques calculated from equation 1.105.

for this same simulation. During the resonant period, the instantaneous torque on planet a not only oscillates wildly, but also becomes strongly negative for at least part of the planet’s orbital cycle. Before and after the resonant period, the torque is almost always positive. These strong variations in the torque might easily overcome the resonant torque and break the resonance. This seems particularly plausible given that the period ratio is in a trough before the resonance breaks. The trough suggests that the two planets have just been pushed closer together, and a period of negative torque on planet a would explain this. A zoomed-in view of planet a’s migration confirms this picture of events: figure 5.16 shows that the semi-major axis of the planet actually grows immediately before the resonance is broken.

I therefore propose that the mechanism of resonance breaking is then that the eccentricity growth causes the torques to oscillate wildly. Whilst they remain positive in a time-averaged sense, there are short periods over which planet a experiences negative torques, which push it closer to planet b. The torque oscillations grow with the eccentricity, until at some point a period of negative torquing

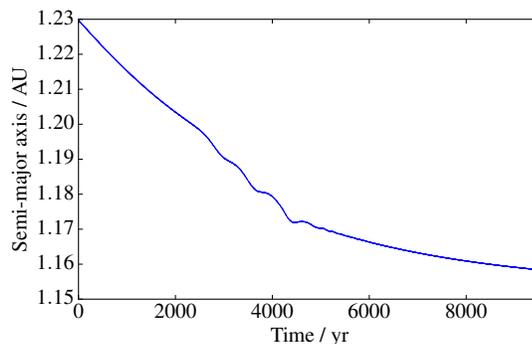


Figure 5.16: Zoomed in view of the migration of planet a in 2planhi. During 2000 and 5000 yr, whilst the planet is in resonance with planet b, there are oscillations in the semi-major axis.

is enough to push planet a inside of the 2 : 1 resonance, breaking the resonance. I stress, however, that the aim of this calculation was to study stopping rather than resonance breaking, and as such, the data from these simulations is not enough to say for certain what the mechanism is. Further work is required with more physical discs and more frequent outputs to establish exactly what effect the disc torque has on two planets in resonance, and vice-versa.

Other studies regarding resonance breaking by discs are rare. [Paardekooper et al. \(2013\)](#) found that the addition of stochastic forcing (similar to that used in the previous two chapters) to hydrodynamical simulations could help planets to move through resonances. They did not however, observe any behaviour intrinsic to a laminar disc that might break resonances. If this mechanism is shown to function in physical discs, then it could go a long way to explaining the dearth of mean-motion resonances in compact systems.

5.4.4 Extending this work

The natural extension of this work is to include more realistic physics, starting with a small α viscosity (see section 1.1.4). This would bring my results more in line with the fitting formula from [D’Angelo & Lubow \(2010\)](#) by reducing the likelihood of corotation torque saturation, as well as reducing the depth of the partial gaps in the disc. In addition to this, it may be worth performing similar simulations with an adiabatic equation of state. This would enable one to gauge the effect of the entropy-related corotation torque, which is stronger than the vortensity-related torque and therefore may reduce the

range of surface densities at which disc clearing is a viable method of halting migration.

One of the challenges with a simulation involving three planets of different masses migrating in unison is the sheer dynamic range required in the simulation domain. The width of the horseshoe region depends upon the mass of the planet, as I discussed in section 5.2.7, meaning that at different parts of the disc, there is a different resolution requirement to resolve the physics in this region. This may explain why for the lower mass planets in my simulations, the torque is less than predicted based on 3D simulations by (D'Angelo & Lubow, 2010). Therefore, I stress that future studies should consider the effect of differing horseshoe region resolutions in different parts of the disc, particularly when considering the resonance-breaking mechanism highlighted above.

5.5 Summary

Type I migration of tightly-packed planets can be stopped by the clearing of the inner disc as a result of photoevaporation. This can be shown to order-of-magnitude using a simple time-scale comparison, but the numerical simulations presented herein have shown that planets that are within the critical clearing radius R_{crit} when the clearing process begins can become stranded as the inner disc clears. Moreover, I have identified a previously unseen resonance breaking mechanism that appears to be caused by the interplay between resonance-pumped eccentricity growth and the subsequent change in type I torques. Future work is required to understand the exact mechanism by which this happens, and in what parameter space of discs it applies.

Chapter 6

Conclusions

6.1 The story so far

Throughout this thesis, I have built a case for how compact planetary systems, such as Kepler-11, Kepler-32 and Kepler-90, might form. This scenario begins with formation from small dust and solid particles at radii beyond the snow-line in a circumstellar, protoplanetary disc. Torques from the disc then rob the planets of the angular momentum causing them to move closer to the star, possibly within the snow-line. The challenge is then to explain why this migration does not lead to the trapping of adjacent planets in mean-motion resonances, and what might cause the migration of the planets to stop.

In chapter 1, I reviewed much of the basic physics surrounding our knowledge of exoplanetary systems today. Beginning with observations of young stars and their discs, I summarised the various methods of detecting planets, and covered the theoretical dynamical behaviour of both discs and planets analytically. I gave a detailed discussion of type I migration and mean-motion resonances, and explained why the combination of these two effects leads to resonant trapping, which in turn leads to the excitement of eccentricity. I introduced the concept of compact planetary systems, and explained the basics behind the numerical techniques used throughout the rest of this thesis.

Chapter 3 presented my early attempts at modelling the assembly of compact planetary systems. I began by discussing the two approaches to forming such systems: *in situ* formation and migration. I argued that *in situ* formation is unlikely and thus explored the idea of forming super-Earths further out

in the disc and migrating them into a compact configuration. I modelled this using an N -body code with parametrized forces representing the various aspects of type I migration. Using an N -body code allowed me to explore a wide variety of parameters, and I considered how these parameters related to physical discs. I found that this migration-based approach can produce sufficiently compact systems that are stable for Myr time-scales in many physically-motivated disc configurations, but produces an over-abundance of planets in mean-motion resonances compared to what is observed.

Chapter 4 began with a simple premise: perhaps some of these mean-motion resonances could be broken or altered by dynamically-important, Jupiter-mass companions on the periphery of compact systems. These planets would not necessarily be detectable by transit-based searches due to their wide orbits and possible inclinations. I found that the addition of such giants to the models from chapter 3 could break widely-spaced, first-order mean-motion resonances such as the 2 : 1, allowing super-Earths to continue convergent migration. This mechanism could go some way to explaining the paucity of widely-spaced mean-motion resonances in tightly-packed Kepler systems, where Kepler’s incompleteness beyond periods of ≈ 1.5 yr could hide giants. Conversely, this led me to suggest that transiting planetary systems with tighter or even 3-body Laplace resonances might be better candidates for follow-up studies searching for Jupiters.

Finally, in chapter 5, I presented results of high-resolution, two-dimensional hydrodynamical simulations of super-Earth and sub-Neptune mass planets undergoing type I migration. The aim of this chapter was to overcome some of the physical limitations of the work in chapters 3 and 4 by modelling type I migration consistently, and attempting to understand how disc clearing might alter the migration rate and change the evolution of these systems. I focused on a disc model that represented the “canonical” case from chapter 3 - that is, a disc that would promote inward migration in a way that was already known to be conducive to the assembly of tightly packed systems. These simulations yielded several interesting results that were not immediately clear from the N -body models. To begin with, I witnessed resonant breaking as a result of what appears to be nothing more than planet-disc interactions. This mechanism for resonance breaking appears to be poorly explored in the literature and warrants further analysis, with increasingly realistic disc models. Secondly, I saw that even the addition of a very simple disc clearing model was enough to cause migration rates to asymptote towards 0. Thus it appears that disc clearing applied at the correct time during the early evolution of

planetary systems could halt migration.

6.2 Future work

One of the greatest issues faced in the models presented in this thesis is determining how type I migration comes to a stop. In chapter 5 I attempted to understand this better using an ad-hoc prescription for disc clearing. This approach is however, rather simple, and neglects some of the more recent work regarding both disc clearing and disc structure. These additional pieces of physics might also help to reduce the rate of resonances in tightly-packed systems even further, bringing the models closer to the observations.

6.2.1 Planet traps

Recently there has been a great deal of interest in planet traps: so-called “zero-torque” regions of the disc where the planet experiences no net torque from the disc and therefore can remain stationary. Such traps can occur at discontinuities in the disc, for instance at the inner-edge of the MRI (magnetorotational instability) dead-zone, where there is a sudden jump in disc viscosity, or at the aforementioned magnetospheric truncation radius of the disc. Many studies have considered the trapping of single planets (e.g., [Masset et al., 2006b](#)), but there has been little interest in how this extends to multiple planets. What happens, for instance, when one of a resonant pair becomes trapped? [Migaszewski \(2015\)](#) explored this process using analytical prescriptions for the evolution of the disc and migration torques, and found that planets in resonance could potentially push their partners through a trap or become stuck, causing such a difference in migration time-scale that the resonance would be broken. Either outcome could potentially have a huge effect on the structure of planetary systems. This is an area in which future missions such as CHEOPS, PLATO and associated RV data could make a huge difference to models; both the resonant trapping and migration processes are dependent upon the mass of the planets, and thus understanding how the period ratios of real planets vary according to their mass ratio could shed a lot of light on how planets get trapped and broken out of resonance.

In the short-term, I would like to investigate how imposed traps in physically-realistic locations alter the migration of multiple planets in hydrodynamical simulations. [Bitsch et al. \(2014c\)](#) performed

hydrodynamical simulations in this vein, prescribing different viscosities to active and dead parts of the disc based on assumptions about at which points the MRI would be active. However, these simulations were only used to produce torque maps for a single planet. Extending these simulations to include multiple, live planets is an important step and would provide a huge insight into the assembly of planetary systems.

The long-term goal is naturally to investigate self-consistently modelled traps. This poses a much greater challenge since the generation of traps depends on a number of factors that are normally excluded from simulations due to computational expense. For instance, resolving the growth of the MRI in 3D requires resolution that is prohibitively high for simulations of migration (see e.g., [Lesur et al., 2014b](#)). Nevertheless, for sufficiently low-mass planets (such that they do not affect the structure of the disc), it is possible to run a suite of such high-resolution simulations without embedded planets and use the resulting information about disc evolution to influence the set-up of lower-resolution models with prescribed traps.

6.2.2 Disc clearing

In chapter 5, I have presented simulations using a simplified disc clearing model as a way to halt migration. The natural extension to this work is to have the disc clear in a self-consistent manner. In section 1.1.2, I talked somewhat about the process of photoevaporation that may lead to disc clearing. The effect of this process on the migration of giant planets in the type II regime by authors such as [Alexander & Armitage \(2009\)](#), [Moeckel & Armitage \(2012\)](#) and [Rosotti et al. \(2015\)](#). This process leads to a gap forming in the disc at a radius on the order of 1AU, isolating the inner disc and causing it to accrete onto the star whilst the out disc is evaporated. This clearing process takes approximately 10^5 yr (e.g., [Alexander et al., 2006](#)), and can quite naturally have some sort of effect on planet migration, since planets near to the critical radius end up experiencing less torque from the inner disc. For sufficiently massive planets, the gap opened by the planet can also have some effect on the disc clearing.

However, to date, no such study exists in the type I regime. The implications of this mechanism for the formation of tightly-packed systems could be vast. For instance, if the planets form sufficiently close to the end of the disc lifetime, disc clearing could halt the migration of all of the planets in a disc

almost simultaneously. This is especially promising given that I found that 10^5 yr is a feasible migration time-scale for the assembly of tightly-packed systems. Additionally, given that the inner disc drains first leaving a gap at AU radii, planets closer to the star will be directly embedded in the disc for longer. This could in principle lead to an interesting modification to the resonant configurations observed in my previous, completely-simultaneous migration models. Therefore, implementing a simplified model of photoevaporative clearing in PLUTO would enable me to gain a much deeper insight into the final stages of assembly of tightly-packed systems. This could naturally lead to simulations including both the effects of planet traps and disc clearing.

6.2.3 Alternate modelling techniques

From the results and discussion surround the two techniques presented within this thesis to model the assembly of compact systems, it is clear that each has advantages and disadvantages. The parametrized N -body models enabled me to quickly explore a large parameter space, but missed physical effects such as resonance breaking that could prove to be crucial. The hydrodynamical models captured this physics, but unfortunately are sufficiently computationally expensive that only a few permutations of disc parameters may be explored. Given the delicate balance between physical accuracy and computational expense, it is tempting to try to find a method of simulating compact systems that combines some of the advantages of both of these techniques.

The natural step would be to switch to a one dimensional disc model based on the equations derived in 1.1.4, slashing the computational costs associated with full two-dimensional simulations. [Alexander & Armitage \(2007\)](#); [Alexander & Armitage \(2009\)](#) already added a robust model for photoevaporative winds to one dimensional simulations, and the torque formulae provided in section 1.5.1 appear to agree well with my two-dimensional simulations, suggesting that they could be use to emulate type I migration in such a set-up. [Coleman & Nelson \(2014\)](#) have in fact, already approached this problem with some success in the context of compact planetary systems. Performing further simulations in this vein and calibrating them against two-dimensional simulations such as those presented in chapter 5 could help to understand the usefulness of this one-dimensional approach. Whilst I have shown in chapter 5 that wake interactions between neighbouring planets are not important, it is not immediately clear if the resonance-breaking behaviour that occurred in my simulations can be captured in 1D. If

this mechanism is the result of type I torques varying with eccentricity as in [Papaloizou & Larwood \(2000\)](#); [Fendyke & Nelson \(2014\)](#), then it is quite possible that 1D simulations could model this behaviour with additional parametrized torques.

This is only one of a multitude of possibilities. One could envisage improving the simple N -body method from chapters 3 and 4 in many different ways. For instance, imposing a disc model and using the prescriptions in section 1.5.1 without directly modelling the disc, or simply adding the torque scaling with eccentricity derived by [Fendyke & Nelson \(2014\)](#). These same parametrized models could be applied to new systems in the future, particularly as new compact systems are discovered and characterised, potentially with the giant companions hypothesised in chapter 4. Ever-increasing accuracy in the determination of mass and radius of these planets will help to constrain the initial conditions for such models. Increasingly well-resolved observations of discs as planet-forming systems (see e.g., [Quanz et al., 2013](#)) will also help in this regard, by giving observationally-motivated initial locations and orbital spacings for super-Earths. Such a constraint was missing from the models in this thesis, which used a simple oligarchic spacing argument.

6.3 Final conclusions

This thesis has approached the problem of forming compact planetary systems from the perspective that the planets in these systems form beyond the snow-line in a disc, and then migrate convergently toward their host star. I have shown that this method is a plausible way of forming these systems for a range of different disc parameters and using two different simulation methods. Furthermore, I have investigated the over-abundance of mean-motion resonances generated by these models, and found that both planet-planet and planet-disc interactions are capable of reducing the frequency of MMRs. In particular, I have highlighted a potential new mechanism for breaking resonances as a result of the change in type I migration torques for eccentric planets. Further work should concentrate on increasing the physical accuracy of these models, and collecting more information regarding this mechanism.

6.4 Acknowledgements

This research has made use of the Exoplanet Orbit Database and the Exoplanet Data Explorer at exoplanets.org. I gratefully acknowledge support from an Science & Technology Facilities Council (STFC) PhD studentship. Astrophysical research at the University of Leicester is supported by an STFC Consolidated Grant (ST/K001000/1). This research used the ALICE High Performance Computing Facility at the University of Leicester. Some resources on ALICE form part of the DiRAC Facility jointly funded by STFC and the Large Facilities Capital Fund of BIS. This work also used the DiRAC *Complexity* system, operated by the University of Leicester IT Services, which forms part of the STFC DiRAC HPC Facility (<http://www.dirac.ac.uk>). This equipment is funded by BIS National E-Infrastructure capital grant ST/K000373/1 and STFC DiRAC Operations grant ST/K0003259/1. DiRAC is part of the UK National E-Infrastructure.

Appendix A

On the validity of the thin disc approximation

In section 1.1.4, I derived equation 1.35 from Newton's second law and considerations of the pressure and gravitational potential. For the sake of clarity, I shall restate the equation here:

$$\frac{1}{\gamma} \left(\frac{\partial \log(\rho)}{\partial z} + \frac{\partial \log(c_s^2)}{\partial z} \right) = -\frac{1}{c_s^2} \frac{\partial \Phi}{\partial z} \quad (\text{A.1})$$

Here, $\Phi = \Phi(R, z)$ is an arbitrary potential and $c_s = c_s(z)$ is an arbitrary sound-speed profile that may vary in the R and z direction. In continuing to simplify this expression for a Keplerian potential, I made the assumption that $z \ll R$; that is to say that the disc is thin. This assumption is however not strictly necessary to obtain an analytic density profile. To discuss this in more detail, I will begin by eliminating the potential derivative from the equation. To begin, we rearrange to find

$$\frac{1}{\gamma} \frac{\partial \log(\rho)}{\partial z} = -\frac{1}{c_s^2} \frac{\partial \Phi}{\partial z} - \frac{1}{\gamma} \frac{\partial \log(c_s^2)}{\partial z} \quad (\text{A.2})$$

which can be trivially integrated from 0 to arbitrary z , yielding

$$\frac{1}{\gamma} [\log(\rho(z'))]_0^z = - \int_0^z \frac{1}{c_s(z')^2} \frac{\partial \Phi}{\partial z'} dz' - \frac{1}{\gamma} [\log(c_s(z')^2)]_0^z \quad (\text{A.3})$$

Applying integration by parts ($\int u dv = uv - \int v du$) eliminates the potential derivative entirely, at the expense of gaining a derivative of the sound-speed profile:

$$\frac{1}{\gamma} [\log(\rho(z'))]_0^z = - \left[\frac{\Phi}{c_s(z')^2} \right]_0^z + \int_0^z \Phi \frac{\partial}{\partial z'} \frac{1}{c_s(z')^2} dz' - \frac{1}{\gamma} [\log(c_s(z')^2)]_0^z \quad (\text{A.4})$$

which naturally makes the ensuing algebra easier for sufficiently complex potentials and sufficiently simple sound-speed profiles. We now again apply the assumption that the disc is isothermal in z , giving $\gamma = 1$ and eliminating the second and third terms which have a dependence on the gradient of sound-speed profile. This leaves only the first term, into which we substitute the familiar Keplerian potential $\Phi = -GM/r$ with $r = \sqrt{R^2 + z^2}$, finding

$$\log \frac{\rho}{\rho_0} = -\frac{GM}{c_s^2} \left(\frac{1}{R} - \frac{1}{\sqrt{R^2 + z^2}} \right) \quad (\text{A.5})$$

or equivalently

$$\rho = \rho_0 \cdot \exp \left[-\frac{GM}{c_s^2} \left(\frac{1}{R} - \frac{1}{\sqrt{R^2 + z^2}} \right) \right]. \quad (\text{A.6})$$

Finally, we remove a factor of $1/R$ from the inner bracket, obtaining

$$\rho = \rho_0 \cdot \exp \left[-\frac{GM}{Rc_s^2} \left(1 - \frac{1}{\sqrt{1 + \frac{z^2}{R^2}}} \right) \right] = \rho_0 \cdot \exp \left[-\frac{R^2}{H^2} \left(1 - \frac{1}{\sqrt{1 + \frac{z^2}{R^2}}} \right) \right] \quad (\text{A.7})$$

where we have used the standard definition $H = c_s/\Omega_K$ and $\Omega_K^2 R^2 = GM/R$ to recast the equation in terms of the disc scale height. It is readily apparent that this equation is not as instructive as [1.39](#), since it is not as easy to picture, but it does retain the information regarding the scale height H . It was, however, derived without the assumption of a thin disc and as such is useful for setting up hydrodynamical simulations such that they are closer to equilibrium far from the midplane than they would be using equation [1.39](#). It should be noted that abandoning the thin-disc approximation means one should also consider the fact that the circular velocity is not strictly Keplerian out of the midplane, and that the pressure gradient in R can be corrected more exactly. For instance, [Takeuchi & Lin \(2002\)](#) show that gas out of the midplane rotates at a sub-Keplerian velocity. The thin disc approximation may be trivially recovered from equation [A.7](#) by expanding the term $1/\sqrt{1 + \frac{z^2}{R^2}}$ as a Taylor series in the limit $(z/R)^2 \rightarrow 0$. The zeroth-order term is simply 1, and the first-order term reduces to $-1/2 \cdot z^2/R^2$, thus making equation [A.7](#) equal to equation [1.39](#).

Here I only wish to use this equation [A.7](#) to verify the validity of the thin-disc approximation in canonical protoplanetary discs. [Figure A.1](#) shows a comparison between equations [1.39](#) and [A.7](#) for different radii and disc-thicknesses. It is clear that the two profiles diverge severely only when the aspect ratio of the disc is larger than about 0.1 and thus for canonical protoplanetary discs ($h < 0.1$)

the approximation is valid. It should be noted however that at large radii even for small h , the flaring disc profile causes some noticeable difference between the thin-disc approximation and equation [A.7](#).

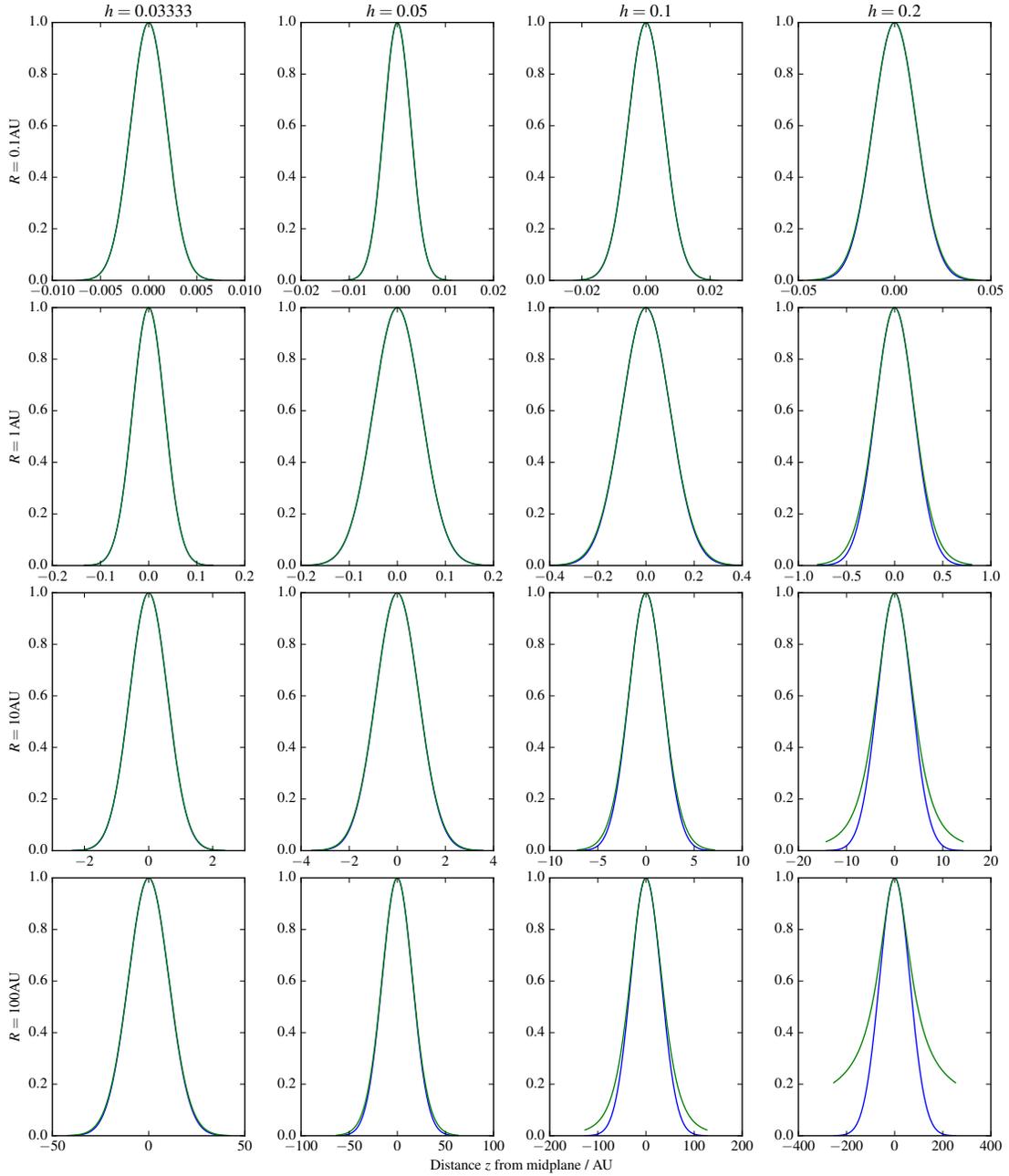


Figure A.1: A validation of the thin disc approximation. Green line represents the exact z profile of the disc, blue line the thin-disc approximation. The value h provided is the disc aspect ratio H/R at $R = 1 \text{ AU}$. The disc flares such that $H \propto R^{5/4}$. The y-axis in each plot is the volume density relative to that at the midplane.

Appendix B

Stochastic forces

In addition to the “smooth” torques discussed in section 2.2.2, stochastic forces are applied to the planets with varying magnitudes in order to simulate the effect of disc turbulence on the planets. I follow the method of [Rein & Papaloizou \(2009\)](#), using a discrete time Markov process to generate stochastic forces in both the θ and r direction at every point along the planet’s orbit, hence adding noise to the acceleration of each planet and sending its orbital elements on a random walk. This process is zero-mean and Gaussian, and is defined by two parameters: the root-mean-square force $\sqrt{\langle F^2 \rangle}$ and the auto-correlation time τ_c . The auto-correlation function of this process is given as

$$R(\tau) = A \cdot \exp(-|\tau|/\tau_c). \quad (\text{B.1})$$

In discrete time, the value x of the process is advanced as

$$x_{k+1} = e^{-\delta t/\tau_c} x_k + w_k, \quad (\text{B.2})$$

where w_k is the result of an Independent and Identically Distributed random variables (IID) process, for which we sample a normal distribution with variance

$$Q_d = \frac{Q}{2\tau_c} (1 - e^{-\delta t/\tau}) \quad (\text{B.3})$$

where Q is the variance of the resulting Markov process. The meaning of these equation is simply that the next value of x depends only on its current value and the result of a random process, hence the process is said to be memory free ([Kasdin, 1995](#)).

If we define the variance of the process as the mean-square force per unit mass, $\langle F^2 \rangle$, we find $Q = 2\langle F^2 \rangle \tau_c$. Since we are dealing with orbits, the natural correlation time for the process is the orbital period of the planet. The RMS force $\langle F^2 \rangle$ is chosen at each time-step as

$$\beta = \frac{\sqrt{\langle F^2 \rangle}}{a_i},$$

where β is a constant and a_i is the acceleration of particle i . This approach maintains a constant magnitude for the stochastic forces relative to the gravitational forces on the planet.

Bibliography

ALMA Partnership et al., 2015, [ApJ](#), 808, L3

Agertz O., et al., 2007, [MNRAS](#), 380, 963

Alexander R., Armitage P., 2007, *Monthly Notices of the Royal ...*, 13, 1

Alexander R. D., Armitage P. J., 2009, [ApJ](#), 704, 989

Alexander R. D., Pascucci I., 2012, [MNRAS](#), 422, L82

Alexander R. D., Clarke C. J., Pringle J. E., 2006, [MNRAS](#), 369, 229

Alexander R., Pascucci I., Andrews S., Armitage P., Cieza L., 2014, [Protostars and Planets VI](#), pp 475–496

Andre P., Montmerle T., 1994, [ApJ](#), 420, 837

Andre P., Ward-Thompson D., Barsony M., 1993, [ApJ](#), 406, 122

Andrews S. M., Williams J. P., 2007, [ApJ](#), 659, 705

Armitage P., 2010, *Astrophysics of Planet Formation*. Cambridge University Press, <http://books.google.co.uk/books?id=zjLZQnFfMs8C>

Armitage P. J., Livio M., Lubow S. H., Pringle J. E., 2002, [MNRAS](#), 334, 248

Artymowicz P., 1993, [ApJ](#), 419, 155

Balbus S. A., 2011, in Garcia P. J. V., ed., *Physical Processes in Circumstellar Disks around Young Stars*. pp 237–282

BIBLIOGRAPHY

- Balbus S. A., Hawley J. F., 1991, *ApJ*, 376, 214
- Balbus S. A., Papaloizou J. C. B., 1999, *ApJ*, 521, 650
- Baruteau C., Lin D. N. C., 2010, *ApJ*, 709, 759
- Baruteau C., Papaloizou J. C. B., 2013, *ApJ*, 778, 7
- Baruteau C., et al., 2013, ArXiv:1312.4293,
- Baruteau C., et al., 2014, *Protostars and Planets VI*, pp 667–689
- Bate R., Mueller D., White J., 1971, *Fundamentals of Astrodynamics*. Dover Books on Aeronautical Engineering Series, Dover Publications, <https://books.google.co.uk/books?id=UtJK8cetqGkC>
- Binney J., Tremaine S., 1987, *Galactic dynamics*
- Bitsch B., Kley W., 2010, *A&A*, 523, A30
- Bitsch B., Crida A., Morbidelli A., Kley W., Dobbs-Dixon I., 2013a, *A&A*, 549, A124
- Bitsch B., Boley A., Kley W., 2013b, *A&A*, 550, A52
- Bitsch B., Morbidelli A., Lega E., Kretke K., Crida A., 2014a, ArXiv:1408.1016,
- Bitsch B., Morbidelli A., Lega E., Crida A., 2014b, *A&A*, 564, A135
- Bitsch B., Morbidelli A., Lega E., Kretke K., Crida A., 2014c, *A&A*, 570, A75
- Bodenheimer P., Pollack J. B., 1986, *Icarus*, 67, 391
- Borucki W. J., 2016, *Reports on Progress in Physics*, 79, 036901
- Boss A. P., 2000, *ApJ*, 536, L101
- Bouvier J., Alencar S. H. P., Harries T. J., Johns-Krull C. M., Romanova M. M., 2007, *Protostars and Planets V*, pp 479–494
- Burke C. J., et al., 2014, *ApJS*, 210, 19

BIBLIOGRAPHY

- Butler R. P., Marcy G. W., Fischer D. A., Brown T. M., Contos A. R., Korzennik S. G., Nisenson P., Noyes R. W., 1999, *ApJ*, 526, 916
- Cabrera J., et al., 2014, *ApJ*, 781, 18
- Carter J. A., et al., 2012, *Science*, 337, 556
- Casoli J., Masset F. S., 2009, *ApJ*, 703, 845
- Chambers J. E., Migliorini F., 1997, in AAS/Division for Planetary Sciences Meeting Abstracts #29. p. 1024
- Chatterjee S., Ford E. B., 2014, ArXiv:1406.0521,
- Chauvin G., Lagrange A.-M., Dumas C., Zuckerman B., Mouillet D., Song I., Beuzit J.-L., Lowrance P., 2004, *A&A*, 425, L29
- Chenciner A., Montgomery R., 2000, *Annals of Mathematics*, 152, 881
- Chiang E., Laughlin G., 2013, *MNRAS*, 431, 3444
- Chiang E., Youdin A. N., 2010, *Annual Review of Earth and Planetary Sciences*, 38, 493
- Christian D. J., et al., 2006, *MNRAS*, 372, 1117
- Coleman G. A. L., Nelson R. P., 2014, *MNRAS*, 445, 479
- Cossou C., Raymond S. N., Hersant F., Pierens A., 2014, ArXiv:1407.6011,
- Coughlin J. L., et al., 2015, preprint, ([arXiv:1512.06149](https://arxiv.org/abs/1512.06149))
- Cresswell P., Dirksen G., Kley W., Nelson R. P., 2007, *A&A*, 473, 329
- Crida A., Morbidelli A., 2007, *MNRAS*, 377, 1324
- Crida A., Morbidelli A., Masset F., 2006, *Icarus*, 181, 587
- Cumming A., Butler R. P., Marcy G. W., Vogt S. S., Wright J. T., Fischer D. A., 2008, *PASP*, 120, 531
- D'Angelo G., Lubow S. H., 2008, *ApJ*, 685, 560

BIBLIOGRAPHY

- D'Angelo G., Lubow S. H., 2010, *ApJ*, 724, 730
- Davies M. B., Adams F. C., Armitage P., Chambers J., Ford E., Morbidelli A., Raymond S. N., Veras D., 2014, *Protostars and Planets VI*, pp 787–808
- Deck K. M., Holman M. J., Agol E., Carter J. A., Lissauer J. J., Ragozzine D., Winn J. N., 2012, *ApJ*, 755, L21
- Deck K. M., Payne M., Holman M. J., 2013, *ApJ*, 774, 129
- Dehnen W., Read J. I., 2011, *The European Physical Journal Plus*, 126, 55
- Deleuil M., Moutou C., Bordé P., 2011, in *European Physical Journal Web of Conferences*. p. 01001 ([arXiv:1105.1887](https://arxiv.org/abs/1105.1887)), [doi:10.1051/epjconf/20101101001](https://doi.org/10.1051/epjconf/20101101001)
- Duffell P. C., MacFadyen A. I., 2013, *ApJ*, 769, 41
- Dullemond C., 2011, Chapter 3 Advection algorithms I. The basics, http://www.ita.uni-heidelberg.de/~dullemond/lectures/num_fluid_2011/Chapter_3.pdf
- Dunham M. M., et al., 2014, *Protostars and Planets VI*, pp 195–218
- Dunhill A. C., 2013, PhD thesis, University of Leicester, <http://hdl.handle.net/2381/28373>
- Eisner J. A., Hillenbrand L. A., White R. J., Akeson R. L., Sargent A. I., 2005, *ApJ*, 623, 952
- Espaillet C., et al., 2014, *Protostars and Planets VI*, pp 497–520
- Fabrycky D. C., et al., 2014a, *ApJ*, 790, 146
- Fabrycky D. C., et al., 2014b, *ApJ*, 790, 146
- Fendyke S. M., Nelson R. P., 2014, *MNRAS*, 437, 96
- Fischer D. A., Howard A. W., Laughlin G. P., Macintosh B., Mahadevan S., Sahlmann J., Yee J. C., 2014a, *Protostars and Planets VI*, pp 715–737
- Fischer D. A., Marcy G. W., Spronck J. F. P., 2014b, *ApJS*, 210, 5

BIBLIOGRAPHY

- Flaherty K. M., Hughes A. M., Rosenfeld K. A., Andrews S. M., Chiang E., Simon J. B., Kerzner S., Wilner D. J., 2015, *ApJ*, **813**, 99
- Fletcher M., Nayakshin S., 2016, preprint, ([arXiv:1605.08329](https://arxiv.org/abs/1605.08329))
- Furlan E., et al., 2008, *ApJS*, **176**, 184
- Gammie C. F., 1996, *ApJ*, **457**, 355
- Garaud P., Lin D. N. C., 2004, *ApJ*, **608**, 1050
- Gingold R. A., Monaghan J. J., 1977, *MNRAS*, **181**, 375
- Godunov S. K., 1959, *Mat. Sb. (N.S.)*, **47** (89), 271
- Goldreich P., Schlichting H. E., 2014, *AJ*, **147**, 32
- Goldreich P., Tremaine S., 1979, *ApJ*, **233**, 857
- Goldreich P., Tremaine S., 1980, *ApJ*, **241**, 425
- Greenberg R., Hartmann W. K., Chapman C. R., Wacker J. F., 1978, *Icarus*, **35**, 1
- Griffin R., 1973, *MNRAS*, **162**, 243
- Guyon O., Martinache F., Cady E. J., Belikov R., Balasubramanian K., Wilson D., Clergeon C. S., Mateen M., 2012, in *Adaptive Optics Systems III*. p. 84471X, [doi:10.1117/12.927181](https://doi.org/10.1117/12.927181)
- Hansen B. M. S., Murray N., 2012, *ApJ*, **751**, 158
- Hansen B. M. S., Murray N., 2013, *ApJ*, **775**, 53
- Hansen B. M. S., Murray N., 2014, *ArXiv:1405.2342*,
- Hartmann L., 1998, *Accretion Processes in Star Formation*
- Hartmann L., Hewett R., Calvet N., 1994, *ApJ*, **426**, 669
- Hasegawa Y., Pudritz R. E., 2011, *MNRAS*, **417**, 1236
- Helled R., et al., 2014, *Protostars and Planets VI*, pp 643–665

BIBLIOGRAPHY

- Holmes M., 2006, Introduction to Numerical Methods in Differential Equations. Texts in Applied Mathematics, Springer, New York, <https://books.google.co.uk/books?id=Bymo5PKzVJQC>
- Hubickyj O., Bodenheimer P., Lissauer J. J., 2005, *Icarus*, 179, 415
- Ida S., Lin D. N. C., 2010, *ApJ*, 719, 810
- Johansen A., Blum J., Tanaka H., Ormel C., Bizzarro M., Rickman H., 2014, *Protostars and Planets VI*, pp 547–570
- Kant I., 1755, Allgemeine Naturgeschichte und Theorie des Himmels oder Versuch von der Verfassung und dem mechanischen Ursprunge des ganzen Weltgebäudes nach Newtonischen Grundsätzen abgehandelt. Petersen, Königsberg and Leipzig, http://www.deutschestextarchiv.de/book/show/kant_naturgeschichte_1755
- Kasdin N., 1995, Proceedings of the IEEE, 83
- Kley W., Nelson R. P., 2012, *ARA&A*, 50, 211
- Koch D. G., et al., 2010, *ApJ*, 713, L79
- Kokubo E., Ida S., 1995, *Icarus*, 114, 247
- Kokubo E., Ida S., 1998, *Icarus*, 131, 171
- Kokubo E., Ida S., 2002, *ApJ*, 581, 666
- Korycansky D. G., Pollack J. B., 1993, *Icarus*, 102, 150
- Kunz M. W., 2008, *MNRAS*, 385, 1494
- Lada C., 1987, Star forming regions; Proceedings of the Symposium
- Lada C. J., 2005, *Progress of Theoretical Physics Supplement*, 158, 1
- Lambrechts M., Johansen A., 2012, *A&A*, 544, A32
- Laughlin G., Steinacker A., Adams F. C., 2004, *ApJ*, 608, 489
- Launhardt R., Sargent A. I., 2001, *ApJ*, 562, L173

BIBLIOGRAPHY

- Lee M. H., Peale S. J., 2002, *ApJ*, 567, 596
- Lee M. H., Fabrycky D., Lin D. N. C., 2013, *ApJ*, 774, 52
- Lesur G., Kunz M. W., Fromang S., 2014a, *A&A*, 566, A56
- Lesur G., Kunz M. W., Fromang S., 2014b, *A&A*, 566, A56
- Lin D. N. C., Papaloizou J., 1979, *MNRAS*, 186, 799
- Lin D., Papaloizou J., 1986, *The Astrophysical Journal*
- Lissauer J. J., et al., 2011a, *ApJS*, 197, 8
- Lissauer J. J., et al., 2011b, *Nature*, 470, 53
- Lissauer J. J., Marcy G. W., Rowe J. F., Bryson S. T., Adams E., Buchhave L. A., Ciardi 2012, *ApJ*, 750, 112
- Lissauer J. J., et al., 2013, *ApJ*, 770, 131
- Lithwick Y., Wu Y., 2012, *ApJ*, 756, L11
- Lodato G., 2007, *Nuovo Cimento Rivista Serie*, 30
- Lodato G., Rice W. K. M., 2004, *MNRAS*, 351, 630
- Lucy L. B., 1977, *AJ*, 82, 1013
- Lynden-Bell D., Pringle J. E., 1974, *MNRAS*, 168, 603
- Lyra W., Paardekooper S.-J., Mac Low M.-M., 2010, *ApJ*, 715, L68
- MacMahon J., 1896, *The Metaphysics of Aristotle*. Bohn's classical library, G. Bell and sons, <https://books.google.co.uk/books?id=vXfLMuNttTUC>
- Marmier M., et al., 2013, *A&A*, 551, A90
- Marois C., Macintosh B., Barman T., Zuckerman B., Song I., Patience J., Lafrenière D., Doyon R., 2008, *Science*, 322, 1348

BIBLIOGRAPHY

- Marois C., Zuckerman B., Konopacky Q. M., Macintosh B., Barman T., 2010, [Nature](#), 468, 1080
- Masset F. S., Casoli J., 2009, [ApJ](#), 703, 857
- Masset F. S., Papaloizou J. C. B., 2003, [ApJ](#), 588, 494
- Masset F. S., Morbidelli A., Crida A., Ferreira J., 2006a, [ApJ](#), 642, 478
- Masset F. S., Morbidelli A., Crida A., Ferreira J., 2006b, [ApJ](#), 642, 478
- Matthews B. C., Krivov A. V., Wyatt M. C., Bryden G., Eiroa C., 2014, [Protostars and Planets VI](#), pp 521–544
- Mayor M., Queloz D., 1995, [Nature](#), 378, 355
- Mayor M., et al., 2011, preprint, ([arXiv:1109.2497](#))
- McKee C. F., Ostriker E. C., 2007, [ARA&A](#), 45, 565
- McNeil D. S., Nelson R. P., 2010, [MNRAS](#), 401, 1691
- Migaszewski C., 2015, [MNRAS](#), 453, 1632
- Mignone A., Bodo G., Massaglia S., Matsakos T., Tesileanu O., Zanni C., Ferrari A., 2007, [ApJS](#), 170, 228
- Moeckel N., Armitage P. J., 2012, [MNRAS](#), 419, 366
- Monaghan J. J., 1992, [ARA&A](#), 30, 543
- Moore C., 1993, [Physical Review Letters](#), 70, 3675
- Mullally F., et al., 2015, [ApJS](#), 217, 31
- Murray C. D., Dermott S. F., 1999, Solar system dynamics
- Mustill A. J., Wyatt M. C., 2011, [MNRAS](#), 413, 554
- Nayakshin S., 2010, [MNRAS](#), 408, L36
- Nelson R. P., 2005, [A&A](#), 443, 1067

BIBLIOGRAPHY

- Nelson R. P., Papaloizou J. C. B., 2004, [MNRAS](#), **350**, 849
- Nelson R. P., Papaloizou J. C. B., Masset F., Kley W., 2000, [MNRAS](#), **318**, 18
- Nesvorný D., Kipping D. M., Buchhave L. A., Bakos G. Á., Hartman J., Schmitt A. R., 2012, [Science](#), **336**, 1133
- Ogilvie G. I., 2014, [ArXiv:1406.2207](#),
- Oishi J. S., Mac Low M.-M., Menou K., 2007, [ApJ](#), **670**, 805
- Paardekooper S.-J., 2014, [MNRAS](#), **444**, 2031
- Paardekooper S.-J., Papaloizou J. C. B., 2009, [MNRAS](#), **394**, 2283
- Paardekooper S.-J., Baruteau C., Crida A., Kley W., 2010, [MNRAS](#), **401**, 1950
- Paardekooper S.-J., Rein H., Kley W., 2013, [MNRAS](#), **434**, 3018
- Papaloizou J. C. B., Larwood J. D., 2000, [MNRAS](#), **315**, 823
- Papaloizou J. C. B., Terquem C., 2006, [Reports on Progress in Physics](#), **69**, 119
- Perri F., Cameron A. G. W., 1974, [Icarus](#), **22**, 416
- Pierens A., Baruteau C., Hersant F., 2012, [MNRAS](#), **427**, 1562
- Pollack J. B., Hubickyj O., Bodenheimer P., Lissauer J. J., Podolak M., Greenzweig Y., 1996, [Icarus](#), **124**, 62
- Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P., 1992, Numerical recipes in C. The art of scientific computing
- Price D. J., 2008, [Journal of Computational Physics](#), **227**, 10040
- Price D. J., 2012a, [Journal of Computational Physics](#), **231**, 759
- Price D. J., 2012b, in Capuzzo-Dolcetta R., Limongi M., Tornambè A., eds, Astronomical Society of the Pacific Conference Series Vol. 453, Advances in Computational Astrophysics: Methods, Tools, and Outcome. p. 249 ([arXiv:1111.1259](#))

BIBLIOGRAPHY

- Pringle J. E., 1981, [Annual Review of Astronomy and Astrophysics](#), 19, 137
- Quanz S. P., Avenhaus H., Buenzli E., Garufi A., Schmid H. M., Wolf S., 2013, [ApJ](#), 766, L2
- Quintana E. V., et al., 2014, [Science](#), 344, 277
- Raymond S. N., Cossou C., 2014, [MNRAS](#), 440, L11
- Raymond S. N., Barnes R., Mandell A. M., 2008a, [MNRAS](#), 384, 663
- Raymond S. N., Barnes R., Armitage P. J., Gorelick N., 2008b, [ApJ](#), 687, L107
- Raymond S. N., Kokubo E., Morbidelli A., Morishima R., Walsh K. J., 2013, [ArXiv:1312.1689](#),
- Raymond S. N., Kokubo E., Morbidelli A., Morishima R., Walsh K. J., 2014, [Protostars and Planets VI](#), pp 595–618
- Rein H., 2012, [MNRAS](#), 427, L21
- Rein H., Liu S.-F., 2012, [A&A](#), 537, A128
- Rein H., Papaloizou J. C. B., 2009, [A&A](#), 497, 595
- Roe P. L., 1986, [Annual Review of Fluid Mechanics](#), 18, 337
- Rosotti G. P., Ercolano B., Owen J. E., 2015, [MNRAS](#), 454, 2173
- Safronov V. S., 1972, Evolution of the protoplanetary cloud and formation of the earth and planets.
- Schmitt J. R., et al., 2014, [AJ](#), 148, 28
- Shakura N., Sunyaev R., 1973, [Astronomy & Astrophysics](#), 24, 337
- Shu F. H., 1977, [ApJ](#), 214, 488
- Shu F. H., Adams F. C., Lizano S., 1987, [ARA&A](#), 25, 23
- Simó C., 2002, in Chenciner A., Cushman R., Robinson C., Xia Z. J., eds, [Celestial Mechanics](#), Dedicated to Donald Saari for his 60th Birthday. p. 209
- Springel V., 2010, [MNRAS](#), 401, 791

BIBLIOGRAPHY

- Stone J. M., Norman M. L., 1992, *ApJS*, 80, 753
- Struve O., 1952, *The Observatory*, 72, 199
- Swift J. J., Johnson J. A., Morton T. D., Crepp J. R., Montet B. T., Fabrycky D. C., Muirhead P. S., 2013, *ApJ*, 764, 105
- Takeuchi T., Lin D. N. C., 2002, *ApJ*, 581, 1344
- Tanaka H., Ward W. R., 2004, *ApJ*, 602, 388
- Tanaka H., Takeuchi T., Ward W. R., 2002, *ApJ*, 565, 1257
- Terquem C., Papaloizou J. C. B., 2007, *ApJ*, 654, 1110
- Testi L., et al., 2014, *Protostars and Planets VI*, pp 339–361
- Turner N. J., Fromang S., Gammie C., Klahr H., Lesur G., Wardle M., Bai X.-N., 2014, *Protostars and Planets VI*, pp 411–432
- Udalski A., 2003, *ApJ*, 590, 284
- Ward W. R., 1991, in *Lunar and Planetary Science Conference*.
- Wardle M., 1999, *MNRAS*, 307, 849
- Wardle M., Ng C., 1999, *MNRAS*, 303, 239
- Wardle M., Salmeron R., 2012, *MNRAS*, 422, 2737
- Weidenschilling S. J., 1977, *MNRAS*, 180, 57
- Wetherill G. W., Stewart G. R., 1989, *Icarus*, 77, 330
- Williams J. P., Cieza L. A., 2011, *ARA&A*, 49, 67
- Wolszczan A., Frail D. A., 1992, *Nature*, 355, 145
- Woodward P., Colella P., 1984, *Journal of Computational Physics*, 54, 115

BIBLIOGRAPHY

Wright J. T., Gaudi B. S., 2013, *Exoplanet Detection Methods*. p. 489, doi:[10.1007/978-94-007-5606-9_10](https://doi.org/10.1007/978-94-007-5606-9_10)

Wuchterl G., 1999, in Gimenez A., Guinan E. F., Montesinos B., eds, *Astronomical Society of the Pacific Conference Series Vol. 173, Stellar Structure: Theory and Test of Connective Energy Transport*. p. 185

Xie J.-W., 2013, *ApJS*, **208**, 22

Yoshida H., 1990, *Physics Letters A*, **150**, 262

Yoshida H., 1993, *Celestial Mechanics and Dynamical Astronomy*, **56**, 27

Youdin A. N., Goodman J., 2005, *ApJ*, **620**, 459

de Val-Borro M., et al., 2006, *MNRAS*, **370**, 529