

GENERATING NON-AXISYMMETRIC
SUBSTRUCTURE IN RINGED
PROTOPLANETARY DISCS

DANIEL PATRICK CUMMINS

A DISSERTATION SUBMITTED TO
IMPERIAL COLLEGE LONDON
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

DEPARTMENT OF PHYSICS
IMPERIAL COLLEGE LONDON

OCTOBER 2022

Copyright Declaration

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-Non Commercial 4.0 International Licence (CC BY-NC). Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose. When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes. Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

Statement of Originality

The work presented in this thesis is original and my own, except where explicitly indicated in the text. Chapter 3 is based on work that has been completed in collaboration with J. E. Owen and R. A. Booth, the results of which are published as:

Cummins, D. P., Owen, J. E. & Booth, R. A. (2022), ‘Extreme Pebble Accretion in Ringed Protoplanetary Discs’, *Monthly Notices of the Royal Astronomical Society* 515(1), 1276–1295.

Chapter 4 is based on work that has been completed in collaboration with J. E. Owen and R. A. Booth. Chapter 5 is based on work that has been completed in collaboration with J. E. Owen.

Acknowledgements

I would like to express my sincere gratitude for my supervisor, James Owen, for his incredible support over these last four years, unhindered by the COVID-19 pandemic. It would not have been anywhere near as enriching and enjoyable an experience without his mentoring, his continuous encouragement to attend and present my research at conferences and summer schools, his facilitation of networking opportunities with visitors to Imperial (both in-person and remotely) as well as at conferences, his efforts in overseeing proposals for computing resources and, most importantly, his proactive attitude and dedication to building and fostering a planet formation group at Imperial, which I am proud to have been a part of.

Abstract

Recent observations of protoplanetary discs have unveiled their wealth of substructure, the origins and implications of which are yet to be fully understood. Of particular interest is whether they promote planet formation, whether they indicate that planet formation is already underway, and what role they play in setting the architecture of planetary systems. In this thesis I present my investigations into physical processes that can give rise to large-scale, non-axisymmetric substructure.

I discuss our current understanding of substructure in protoplanetary discs, and provide an overview of the key processes involved in planet formation. I present the theoretical basis for vortex generation by an accreting planetary embryo. Through hydrodynamics simulations, I show that in massive, axisymmetric dust traps, the growth of a planetary core via pebble accretion produces strong thermal feedback on the disc, making it unstable to vortex formation. I demonstrate how this can further enhance the accretion rate onto the planetary core as well as produce observable non-axisymmetric substructure. I investigate how this process varies with the mass of dust in the ring as well as the local background gas temperature, and show that the thermal feedback always acts to increase the planet's mass. I establish that the required dust masses lie towards the extreme end of those inferred from observations, otherwise the background temperature must be sufficiently low to emphasize the thermal feedback from the embryo's accretion luminosity.

I also investigate how warped and twisted regions of protoplanetary discs cast shadows across the disc, and how this manifests in observations, giving the appearance of non-axisymmetric substructure. I show that this can be used to constrain the geometry of discs in which there is evidence for strong misalignments between their inner and outer regions.

Contents

1	Introduction	10
1.1	Extrasolar Planets	10
1.2	Protoplanetary Discs	11
1.3	Theoretical Framework for Protoplanetary Discs	16
1.3.1	Vertical structure	17
1.3.2	Radial structure	17
1.3.3	Circumstellar dust	21
1.4	Substructure in Protoplanetary Discs	23
1.4.1	Axisymmetric substructure	23
1.4.2	Non-axisymmetric substructure	27
1.4.3	Warped discs	30
1.5	The Formation of Planetary Cores	31
1.5.1	Embryo formation	32
1.6	Thesis Outline	35
2	Numerical Methods	36
2.1	Numerical Hydrodynamics	36
2.1.1	FARGO3D	40
2.1.2	Dust dynamics	42
2.1.3	Pebble accretion	46
2.2	Numerical Radiative Transfer	51
2.2.1	Monte Carlo Radiative Transfer	53
2.2.2	RADMC-3D	56
3	Vortex Formation from an Accreting Planetary Embryo	57
3.1	Introduction	57
3.2	Physical Mechanism	59
3.3	Method	64
3.4	Results	68
3.4.1	Hydrodynamics simulations	68
3.4.2	Radiative transfer simulations	75
3.5	Discussion	78
3.6	Conclusions	85
4	The Impact of Accretion Luminosity on Planet Growth by Pebble Accretion	87
4.1	Introduction	87
4.2	Method	91
4.3	Results	92
4.3.1	Planet growth	92
4.3.2	Disc evolution	97
4.4	Discussion	100
4.4.1	Vortex-assisted pebble accretion	100
4.4.2	Pebble isolation masses in axisymmetric dust traps	106
4.5	Conclusions	108
5	Shadowing in Warped Transition Discs	110
5.1	Introduction	110
5.2	Method	116
5.2.1	Warp and Twist	116
5.2.2	Disc setup	117
5.3	Results	120
5.4	Discussion	123
5.5	Conclusions	128

CONTENTS

6 Summary	129
6.1 Future Directions	130
6.2 Outlook	131
Bibliography	133

List of Figures

1.1	Mass, radius and orbital periods of the confirmed exoplanet population	11
1.2	Transition disc spectral energy distribution	13
1.3	ALMA band 6 observation of HL Tau	15
1.4	Gallery of discs observed in the DSHARP survey	16
2.1	Gas surface density profile with an inner cavity	45
2.2	Dust dynamics numerical scheme validation	46
2.3	Vertically-integrated vs. midplane dust drift velocities	46
2.4	Resolving a planetary embryo's Hill sphere	48
2.5	Grid-based pebble accretion resolution test	49
2.6	Modified FARGO3D flowchart	50
2.7	Micron and millimetre dust grain opacities	52
3.1	Planet mass and pebble surface density evolution in a simple pebble accretion model . . .	61
3.2	Accretion luminosity schematic	62
3.3	Smoothing of the planetary potential	67
3.4	Gas surface density evolution without thermal feedback	69
3.5	Gas surface density evolution with thermal feedback	70
3.6	Vortensity evolution due to thermal feedback	71
3.7	Dust surface density evolution with thermal feedback	72
3.8	Close-up of the vortex formed through accretion luminosity	73
3.9	Planet growth with thermal feedback	74
3.10	Dust mass over time with thermal feedback	75
3.11	Gas surface density with fixed planet mass	76
3.12	Pebble accretion resolution test	76
3.13	Band 7 synthetic observations	78
3.14	Band 6 synthetic observations	79
3.15	Spectral index maps	79
3.16	Comparison of pebble accretion rate to theoretical predictions	81
4.1	Final planet mass reached in dust rings of different masses and temperatures	93
4.2	Planet mass comparison	94
4.3	Pebble accretion rates for $100M_{\oplus}$, $h(a) = 0.035$	95
4.4	Pebble accretion rate comparison for largest scale height disc.	97
4.5	Luminosity radius to Hill radius ratios	98
4.6	Impact of thermal feedback on the dust size distribution	100
4.7	Pebble accretion rates for $h(a) = 0.035$	102
4.8	Pebble accretion rates in $150M_{\oplus}$ rings.	102
4.9	Pebble accretion rate in the $125M_{\oplus}$, $h(a) = 0.035$ ring.	103
4.10	Vortex formation and merging	105
5.1	H -band observation of HD 142527	111
5.2	Submillimetre observation of HD 142527	112
5.3	HD 142527 gas kinematics	114
5.4	Inclination schematic	116
5.5	Dust surface density of a model transition disc	118
5.6	Warp and twist angle parametrization	119
5.7	Model transition disc H -band synthetic observations	120
5.8	Warped and twisted inner disc synthetic observations	122
5.9	Warped and twisted outer disc synthetic observations	124
5.10	Synthetic observation with an alternative warp profile	125
5.11	Warped disc cartoon	127

List of Tables

3.1	Transition disc properties	58
4.1	Ringed disc properties	88
4.2	Gap widths	107

Chapter 1

Introduction

Amongst the most profound questions humanity can ask are those of our origins, from the emergence of life on Earth, to the formation of the Earth itself, and ultimately to the very nature of the Universe. Such questions have been contemplated by the earliest human civilisations, by the minds of some of the greatest natural philosophers, through to modern-day scientists. Indeed, some of the largest unsolved problems in astronomy concern the possibility of life throughout the Universe, and how common planetary systems like our own – the Solar System – are. While questions regarding the possibility or frequency of life can be ambiguous, we are reaching the point where we can begin to answer more well-defined questions, such as how common Earth-like planets in Solar-like planetary systems are.

In order to address questions such as these we must understand the process of planet formation and the conditions under which planets form. Early developments of a theory of planet formation were guided by data gathered within the Solar System. Spectroscopic observations of the Sun, geophysical surveys of Earth, analysis of meteorites found on Earth, missions to both the Moon and Mars, and space probes to study the gas giants and outer Solar System have all provided constraints on the evolution of the Solar System and the formation of its planets. However, the end of the 20th century saw the discovery of planets outside of the Solar System, which immediately challenged the most successful models. The continued discovery of extrasolar planets – “exoplanets” – has only exacerbated this, but by studying the physical properties of these planets, as well as their host stars, we can improve our theory and understanding of planet formation.

1.1 Extrasolar Planets

The discovery of the first planet outside of the Solar System was made by Wolszczan & Frail (1992), who discovered two planets orbiting the pulsar PSR 1257+12, through the deviations they induced in its pulsation period. The first detection of a planet orbiting a main sequence star came soon after: Mayor & Queloz (1995) measured the radial velocity of the G type star 51 Pegasi and detected a regular, sinusoidal perturbation, implying the presence of a planet. The signal in the radial velocity reveals the period of the planet’s orbit as well as an estimate of its mass. Their analysis suggested a planet with mass comparable to that of Jupiter and an orbital period of 4.2 days. This was highly controversial because a giant planet on such a close-in orbit is unlike anything in the Solar System, where the giant planets are found exterior to the terrestrial planets, with periods of tens to hundreds of years. Contemplating the nature of 51 Pegasi b, Guillot et al. (1996) stated “this object is surely the most problematic find in recent memory”.

Over 5000 exoplanets have since been discovered, by a variety of different detection techniques. The two most prolific techniques are the transit and radial velocity methods, finding around 3900 and 1000 respectively, as reported on the NASA Exoplanet Archive¹. The transit method involves measuring the light from a star over some duration. If an orbiting planet passes in front of the star along our line of sight to it, a drop in flux will be observed. The amount of light blocked, and therefore the strength of the transit signal, scales as the square of the ratio of the planet’s radius to the host star’s radius, thus knowledge of the stellar radius yields an estimate of the planet’s radius, and its orbital period can be

¹<https://exoplanetarchive.ipac.caltech.edu>, retrieved 25/10/2022

1.2. Protoplanetary Discs

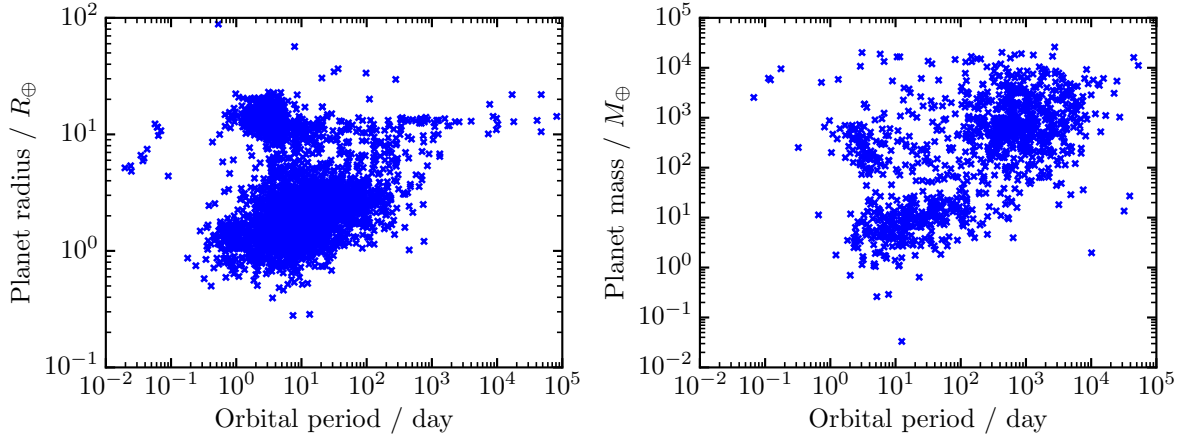


Figure 1.1: The left panel shows the radius–period distribution of the confirmed exoplanet population with radius measurements, presented relative to Earth’s radius. The right panel shows the mass–period distribution, presented relative to Earth’s mass. Uncertainties are not shown. Data obtained from exoplanet.eu on 25/10/2022.

determined by simply measuring how frequently the transit signal occurs. The radial velocity method involves measuring the velocity of a star along our line of sight to it. If the star hosts a planet, it will orbit the barycentre of the system, therefore producing a sinusoidally-varying radial velocity. Kepler’s laws yield the period of the planet, as well as the mass of the planet multiplied by the sine of the inclination of the system with respect to the plane of the sky.

Figure 1.1 shows the radius–period distribution for the sample of the exoplanet population with available radius measurements, as well as the mass–period distribution for the sample with either a direct mass measurement or including the inclination term. Immediately evident is the huge diversity in planet sizes (both masses and radii) and orbital distance from host star, shown here in terms of the orbital period. Many giant planets on close-in orbits, similar to 51 Pegasi b, have been discovered and have come to be known as “Hot Jupiters”. Before concluding that these planets are therefore common, it is worth considering the detection methods. As mentioned above, the vast majority of these planets were detected via either the transit or radial velocity method, both of which are most sensitive to large planets on close-in orbits around small stars. Similarly, only a small fraction of the exoplanet population have masses and radii comparable to Earth, as these primarily lie around the lower limits of detection, thus we cannot simply say that Earth-like planets are rare.

Observational biases aside, the huge diversity in these most fundamental properties of exoplanets is clear, as is the fact that there are many different types of planets simply not found in the Solar System. More detailed analyses reveal the differences in derived properties such as system architectures (e.g. ordering of small and giant planets), eccentricity distributions, period ratios between adjacent planets, to name just a few (Winn & Fabrycky, 2015). This suggests that variations in the initial conditions of planet formation, as well as processes that can occur in planet-forming environments, give rise to a wide variety of planetary system configurations. Understanding these initial conditions, the physical environment of planet formation, as well as the processes involved in planet formation can help us understand these data and develop our theory of planet formation.

1.2 Protoplanetary Discs

One may attempt to gain insight into the processes involved in planet formation by studying the environments in which they form – namely protoplanetary discs. The concept of a protoplanetary disc has

its roots in the observation that each of the Solar System planets orbits the Sun within approximately the same plane, orbiting in the same direction around the Sun, which strongly suggests the formation of planets in a disc. Early ideas may be credited to Swedenborg (1734), though the “nebular hypothesis” is largely due to Kant (1755), and further developed by Marquis de Laplace (1798), which states that the material left over from the collapse of the molecular cloud from which the Sun formed flattened into a disc around the young Sun, eventually forming planets.

The presence of circumstellar material was initially inferred from the spectral energy distributions (SEDs) of young stellar objects. Mendoza Villareal (1966); Mendoza V. (1968) measured the photometry of 33 young stellar objects between $0.9\ \mu\text{m}$ and $5.0\ \mu\text{m}$, finding that all sources showed an excess in infrared (IR) emission above that expected from a blackbody source. They argued that this cannot be due solely to extinction by dust in the interstellar medium (ISM), and that such infrared excesses indicate the presence of dust in the circumstellar environment, which is absorbing stellar radiation and re-radiating it in the infrared.

The sources targeted were primarily T Tauri stars. Named after the prototype T Tau (Joy, 1945), these are young stars (ages typically between 300 kyr to 3 Myr) with strong and broad $\text{H}\alpha$ emission, indicating ongoing accretion, and excess continuum emission in the ultraviolet and infrared. The first direct evidence for disc-like structures around young stellar objects came from near-IR interferometric observations of two such T Tauri stars, HL Tau and R Mon by Beckwith et al. (1984), as well as near-IR imaging of HL Tau by Grasdalen et al. (1984), which showed nebulae similar in spatial extent and mass to the Solar System, and they demonstrated that the nebulae resulted from the light scattered off circumstellar dust grains. Sargent & Beckwith (1987) performed follow-up millimetre-wavelength interferometric observations of HL Tau, detecting continuum emission that was unresolved but coincident with the stellar position, as well as ^{13}CO line emission resembling an edge-on disc, the velocity profile of which revealed Keplerian rotation.

Progress was made in parallel on the theoretical front: Lada & Wilking (1984) surveyed the core of the Ophiuchus molecular cloud and identified three distinct classes of source based on the slope of their SED in the mid-IR. Based on these data, and building on the work of Adams & Shu (1985, 1986), Adams et al. (1987) proposed an evolutionary sequence of a collapsing, rotating core, with these three classes corresponding to phases of the evolution. Disc formation around young stars can therefore be understood as the natural outcome of the spherical collapse of a molecular cloud, while conserving angular momentum.

If we consider a static, uniform, spherical cloud of gas that collapses under its own gravity, the collapse will be spherically symmetric and result in an object supported against gravity by pressure. If the cloud is in uniform rotation, however, the centrifugal force will prevent collapse in the plane perpendicular to the rotation axis. Even if the cloud has non-uniform rotation, angular momentum will be conserved and, unless all angular momenta sum to zero, the collapsing cloud will always have some net rotation. Thus, the general outcome of this process is a pre-main-sequence star accompanied by a circumstellar disc. Following the nebular hypothesis, i.e. that planets are formed from this material, such circumstellar discs are called protoplanetary discs. We can therefore define a protoplanetary disc as a rotationally-supported structure of gas and dust in orbit around a young, pre-main-sequence star, left over from the gravitational collapse of their parent molecular cloud. Protoplanetary discs thus provide the initial conditions for planet formation.

With the launch of the Infrared Astronomical Satellite (IRAS, Neugebauer et al., 1984) in 1983, which observed at $12\ \mu\text{m}$, $25\ \mu\text{m}$, $60\ \mu\text{m}$ and $100\ \mu\text{m}$, spectral energy distributions of young stellar objects could be studied more comprehensively. Strom et al. (1989); Skrutskie et al. (1990) combined ground-based photometry, providing wavelength coverage $\lambda < 10\ \mu\text{m}$, with IRAS measurements of 83 sources in the Taurus-Auriga star-forming region. They discovered that some SEDs, despite showing a significant

1.2. Protoplanetary Discs

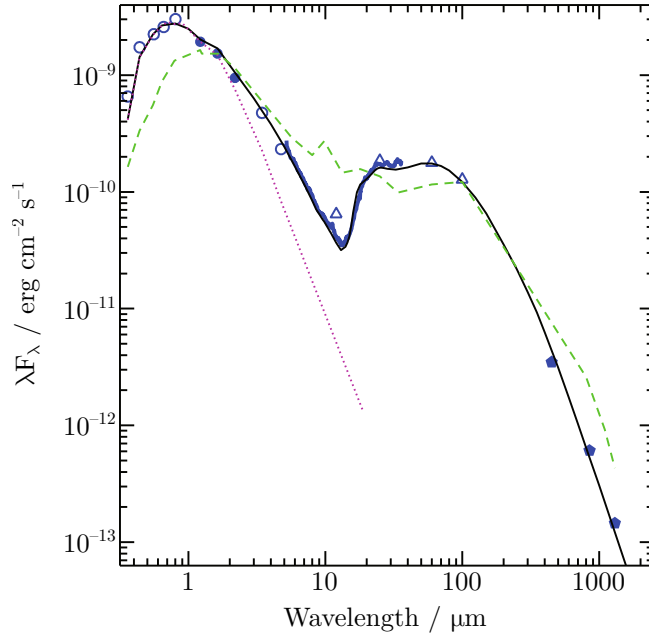


Figure 1.2: SED of UX Tau A, an example of a transition disc. The green dashed line shows the Taurus (T Tauri) median SED, and the magenta dotted line shows a synthetic spectrum for the stellar photosphere. The black solid line is a fit to the UX Tau A photometry: the blue open circles are from ground-based optical L - and M -band measurements, the blue filled circles are 2MASS J -, H - and K -band measurements, the blue solid line is *Spitzer* IRS measurements, open triangles are IRAS measurements and the blue pentagons are millimetre fluxes. Adapted from Espaillat et al. (2007) preprint.

mid-to-far-IR excess, lacked a near-IR excess. The near-IR excess arises from thermal emission by small, warm dust grains and thus originates in the inner disc; the lack of this emission therefore indicates the absence of these grains. This led to the suggestion that these are young stellar objects in transition from disc-hosting to disc-less, which has since led to these becoming known as “transition” discs.

One example of such an SED is that of UX Tau A, shown in figure 1.2 and compared to the median SED of T Tauri stars in Taurus. The measured fluxes closely follow those expected from the stellar photosphere up to around $2\ \mu\text{m}$, beyond which the infrared excess becomes apparent. The median Taurus SED exhibits an excess out to $\approx 100\ \mu\text{m}$ before dropping off following a Rayleigh-Jeans tail. UX Tau A, however, shows a dip in the infrared excess at mid-IR wavelengths (here between approximately $3\ \mu\text{m}$ and $20\ \mu\text{m}$), a feature that is characteristic of these systems.

The lack of a near-IR excess while mid-to-far-IR excess persists implies a disc with a cavity in the dust distribution. Given that most of these sources show signs of accretion, often with rates as high as typical T Tauri stars (e.g. Najita et al., 2007; Cieza et al., 2010; Manara et al., 2014), however, suggests that these cavities cannot be empty, or at least that material must be able to cross them. Indeed, at the time some SEDs showed signs of low levels of dust close to the star (e.g. GM Aur, Calvet et al., 2005).

The first spatially resolved images of protoplanetary discs came with the development of sub-millimetre wavelength interferometers, most notably the Combined Array for Research in Millimeter-wave Astronomy (CARMA), which operated between 2006 and 2015, the Plateau de Bure Interferometer (PdBI), and the Submillimeter Array (SMA). Andrews et al. (2011) observed twelve sources that had been previously identified as transition discs via their SEDs, at $880\ \mu\text{m}$ with the SMA. Achieving a spatial resolution of $0.3''$ (corresponding to ≈ 40 – $75\ \text{AU}$), they resolved dust-depleted cavities with radii between $15\ \text{AU}$ and $73\ \text{AU}$ around the central star in each source. By modelling the emission from each source, they confirmed that the lack of NIR excess was due to an inner cavity.

High spatial resolution imaging of discs has also become possible at near-IR wavelengths, with fa-

cilities such as SPHERE (Spectro-Polarimetric High-contrast Exoplanet REsearch, Beuzit et al., 2008) on the Very Large Telescope (VLT), and GPI (Gemini Planet Imager, Macintosh et al., 2006) on the Gemini South Telescope. These are both adaptive optics and coronagraphic instruments, developed to directly image giant planets orbiting their host stars, both beginning operation in 2014. Observations of protoplanetary discs with this technique have typically focussed on Herbig Ae/Be stars, the more massive (between $2M_{\odot}$ and $8M_{\odot}$) counterpart to T Tauri stars, due to sensitivity limitations. The protoplanetary discs observed at near-IR wavelengths often appear remarkably different to their corresponding (sub-)millimetre images. It is therefore important to consider what is actually being probed at a given observation wavelength. This can be understood by considering the opacity of dust grains of a given size as a function of wavelength. While this is discussed in more detail in chapter 2, it is convenient to provide a simple justification at this point.

The total cross section presented by a dust grain is the sum of its absorption and scattering cross sections. In the limit that the wavelength of radiation (λ) is much smaller than the grain size we may consider the geometric optics limit, wherein the cross section is simply the geometric cross section. At wavelengths longer than the dust grain size, the geometric optics regime is no longer valid and the Rayleigh regime must be considered. In this case the scattering opacity drops off as λ^{-4} , and the absorption opacity as λ^{-1} . This suggests that millimetre-wavelength observations are measuring the light that has been absorbed and then re-emitted by dust grains – i.e. the thermal emission from dust grains – since the scattering opacity at such wavelengths is so low. Similarly, at near-IR wavelengths, scattering dominates, thus imaging of discs at these wavelengths shows the light scattered by dust grains. An additional aspect of this is the size of the dust grains present in the disc: large dust grains reside close to the disc midplane, while smaller grains are distributed vertically through the disc. Scattered light observations therefore image the surface of the disc, while millimetre-wavelength thermal emission probes the midplane. This will be substantiated in the next section.

Millimetre wavelength interferometry truly came into its own with the development of ALMA, the Atacama Large Millimeter/sub-millimeter Array, a collection of 66 antennas operating between 0.32 mm and 3.6 mm. Early observations began in 2011, with baselines (i.e. the separation between any two antennae) up to ≈ 1.5 km. Further developments provided the longest baselines (≈ 15 km), with which ALMA can achieve a maximum spatial resolution of $0.025''$ and $0.017''$ in bands 6 (1.3 mm) and 7 (0.87 mm) respectively, although the actual resolutions for a given observation can vary by $\sim 20\%$ depending on the imaging parameters (ALMA Partnership et al., 2015a). As part of the Science Verification of ALMA Long Baseline Campaign, ALMA targeted the HL Tau region. HL Tau is one of the brightest sources at millimetre wavelengths (e.g. Andrews & Williams, 2005) and has therefore been a well-studied target. Located in Taurus, the nearest star-forming region at approximately 140 pc (Rebull et al., 2004), ALMA can resolve spatial scales corresponding to ~ 5 AU. The highest angular resolution in prior observations of HL Tau was $0.13''$ (~ 20 AU) with CARMA, resolving the disc (Kwon et al., 2011), thus ALMA was expected to produce images with unprecedented detail. Nonetheless, the community was astonished when the ALMA observation, shown in figure 1.3, was published by ALMA Partnership et al. (2015b).

The image reveals a number of bright and dark concentric rings, from orbital separations of ~ 10 AU out to ~ 100 AU, each varying in width and contrast. They suggested that the dark rings are low-density gaps in the disc, caused by the presence of growing planets. The suggestion that planet formation was underway in such a young ($\lesssim 1$ Myr, Briceño et al., 2002) system, with the multiple rings potentially indicating the presence of multiple planets, was indeed a surprise, seemingly at odds with the time-scales for planet formation predicted by models at the time.

More observations at comparable resolutions came, of our nearest (~ 60 pc, van Leeuwen, 2007) young stellar object TW Hya (Andrews et al., 2016) and MWC 758 (Dong et al., 2018a), unveiling the

1.2. Protoplanetary Discs

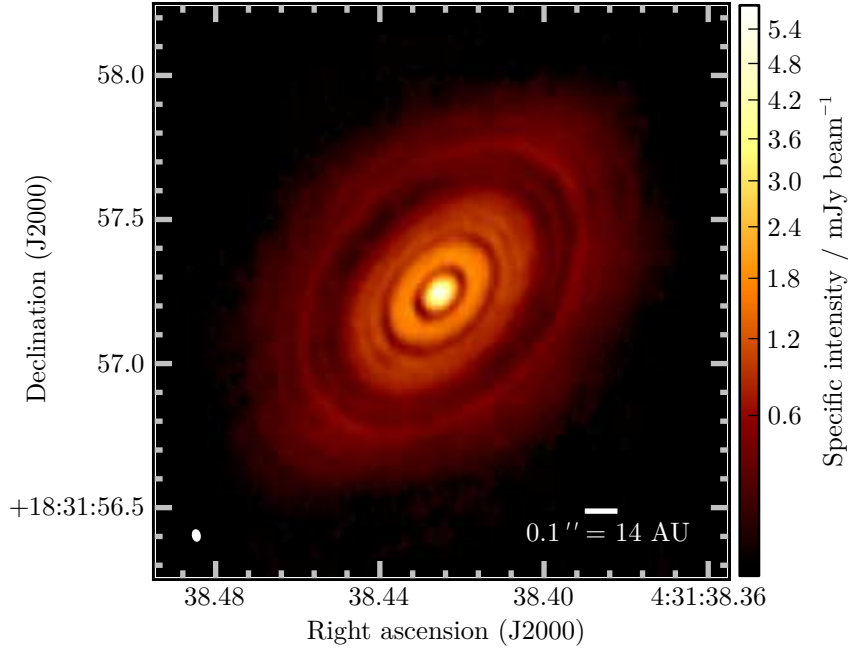


Figure 1.3: The ground-breaking 1.3 mm ALMA observation of HL Tau at an unprecedented spatial resolution of $\approx 0.03''$, revealing multiple concentric bright and dark rings between 10 AU and 100 AU. The beam is represented by the white ellipse in the bottom left corner. Adapted from ALMA Partnership et al. (2015b) preprint.

richness of the substructure in the brightest discs. This provided the motivation and justification for the DSHARP survey (Andrews et al., 2018), in which twenty young stellar objects were observed at $\approx 0.035''$ resolution with ALMA, with the aim of determining the frequency and characteristics of substructure in protoplanetary discs. They targeted some of the brightest nearby discs, but avoided transition discs on the basis that they are already known to contain substructure. The set of 1.25 mm images from the survey is presented in figure 1.4, showing that substructure is ubiquitous amongst this sample of discs. Clearly visible are concentric rings in almost all of the discs, while three show large-scale spiral arms, and two show distinct arc-shaped features. These arc-shaped features, present in HD 143006 and HD 169296, are of particular interest and provide motivation for the investigations presented in chapters 3 and 4.

Subsequent high resolution surveys with ALMA have corroborated that substructure in protoplanetary discs is common (e.g. Long et al., 2018; Cieza et al., 2019), even in discs around low mass stars (Kurtovic et al., 2021). If these are indeed created by planets, or at least indicate that planet formation is underway, they provide constraints on the time-scales of processes required for the formation of planets. While these observations cannot probe the innermost regions of these discs, i.e. within a few astronomical units, where the majority of exoplanets have been detected, the fact that rings are observed out beyond ~ 100 AU in many discs suggests that planets must be forming in the outer regions of these discs. Planets at such large orbital separations evade current transit and radial velocity surveys, but have been detected via direct imaging and microlensing techniques. These suggest that wide separation planetary systems with Jupiter mass planets are not common (e.g. Winn & Fabrycky, 2015), but the empirical constraints on the frequency of lower mass planets remains weak. Furthermore, the presence of planets at large separations, where dynamical time-scales are longer, are difficult to explain in such young systems with current planet formation theory.

Directly comparing the exoplanet population with a hypothetical population of planets inferred from the locations of annular gaps in discs is not a trivial task, however. The confirmed exoplanet population consists almost exclusively of planets with $\sim$ Gyr ages, while protoplanetary discs have $\sim$ Myr ages. Even

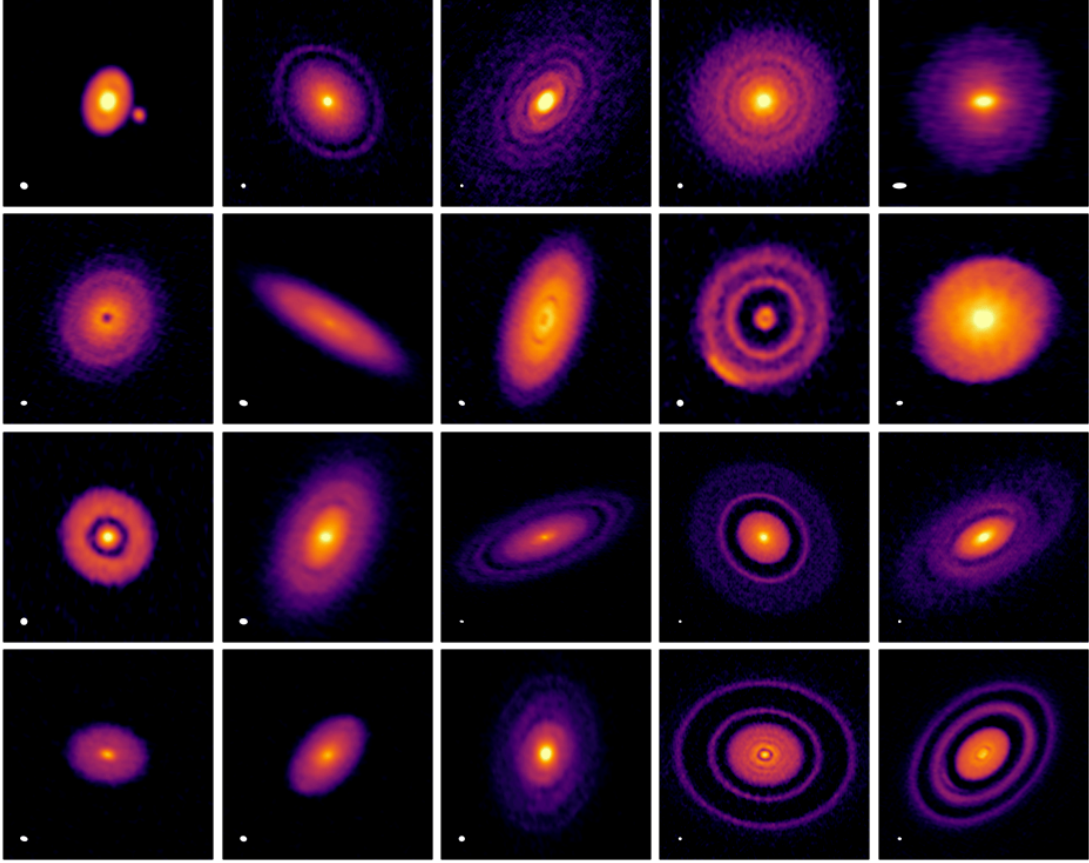


Figure 1.4: Twenty discs observed at a wavelength of 1.25 mm with ALMA in the DSHARP survey. At this wavelength we are measuring the thermal continuum emission from dust grains. The white ellipse in the bottom corner of each image is the beam, typically $0.035''$ in size, which corresponds to ~ 5 AU at the distances to these sources. The colour shows the specific intensity, asinh-scaled to accentuate fainter features, normalised to the maximum flux from each disc (typically of order a few mJy beam^{-1}). A range of substructures can be seen, namely bright and dark concentric rings in most discs, spiral arms (top row, middle column; third row, far right column; bottom row, middle column) and arcs (second row, fourth column; bottom row, far right column). Image from <https://almascience.eso.org/almadata/lp/DSHARP/>.

if planets can be formed out at ~ 100 AU, their interaction with the protoplanetary disc can lead to radial migration (e.g. Lin & Papaloizou, 1986b; Ward, 1997), thus their formation location does not straightforwardly determine where they end up. It is therefore essential to understand the physical processes that occur in discs, that drive their evolution, generate substructure, and how embedded planets interact with them, if we are to make the connection between planet-forming environments and final planetary system configurations.

1.3 Theoretical Framework for Protoplanetary Discs

In this section I provide a mathematical description of protoplanetary discs in order to establish concepts and quantities that will be referred to throughout this thesis, as well as some context for understanding physical processes in protoplanetary discs.

It is convenient to first consider the static vertical structure of protoplanetary discs, determined to first order by the balance between gravity and gas pressure, before addressing their radial structure,

1.3. Theoretical Framework for Protoplanetary Discs

where we should expect time-evolution on the basis that protoplanetary discs are accretion discs. Their geometry makes the cylindrical polar coordinate system a natural choice. I consider an axisymmetric disc, such that gas pressure P and density ρ_g are functions of cylindrical radial distance r and vertical distance z only, in orbit around a star of mass M_* . I will work in the limit of $M_{\text{disc}} \ll M_*$ such that the disc is non-self-gravitating. This is observationally justified: dust masses obtained from (sub-)millimetre fluxes are $\lesssim 1 \times 10^{-3} M_*$ (e.g. Andrews & Williams, 2005, 2007; Ansdell et al., 2016), and assuming a dust-to-gas mass ratio of 0.01 yields disc masses of $\sim 0.01 - 0.1 M_*$.

1.3.1 Vertical structure

Assuming that, vertically, protoplanetary discs are in hydrostatic equilibrium, we may write

$$\frac{1}{\rho_g} \frac{\partial P}{\partial z} = - \frac{GM_*}{r^2 + z^2} \frac{z}{\sqrt{r^2 + z^2}}. \quad (1.1)$$

Using the observational fact that discs are thin, we can make the approximation $z^2 \ll r^2$, which yields

$$\frac{1}{\rho_g} \frac{\partial P}{\partial z} = -\Omega_K^2 z, \quad (1.2)$$

where I have introduced the Keplerian orbital frequency $\Omega_K = \sqrt{GM_*/r^3}$. Making a further assumption that the circumstellar gas obeys the ideal gas law, we can substitute for the pressure using $c_s^2 = \partial P / \partial \rho_g$, to obtain

$$\frac{1}{\rho_g} \frac{\partial \rho_g}{\partial z} = - \frac{\Omega_K^2}{c_s^2} z. \quad (1.3)$$

It is convenient to define the vertical pressure scale height of the gas, $H_g(r) = c_s(r) / \Omega_K(r)$; using this we obtain a differential equation for the gas density as a function of height:

$$\frac{1}{\rho_g} \frac{\partial \rho_g}{\partial z} = - \frac{1}{H_g^2} z. \quad (1.4)$$

Integrating over the vertical dimension yields a Gaussian profile for the gas density, whose width is characterised by $H_g(r)$:

$$\rho_g(r, z) = \rho_{\text{mid}}(r) \exp\left(-\frac{z^2}{2H_g^2}\right), \quad (1.5)$$

where ρ_{mid} is the gas density at the disc midplane ($z = 0$).

Since discs are thin it is often convenient to consider vertically integrated quantities such as surface density,

$$\Sigma_g(r) = \int_{-\infty}^{\infty} \rho_g(r, z) dz, \quad (1.6)$$

and pressure, $\mathcal{P} = c_s^2 \Sigma_g$. Note also that a thin disc implies that orbital speeds are supersonic:

$$\frac{H_g}{r} \equiv h = \frac{c_s}{v_K}. \quad (1.7)$$

h is referred to as the aspect ratio, and is typically $\lesssim 0.1$ (e.g. Andrews et al., 2011; Dullemond et al., 2018), so the orbital speed is highly supersonic.

1.3.2 Radial structure

The early theory of the radial structure and dynamical evolution of astrophysical discs is due to Lynden-Bell & Pringle (1974). They considered the evolution of a Keplerian disc wherein neighbouring annuli

exchange angular momentum through viscosity. In this picture, the viscous stress between material orbiting at slightly different speeds creates a torque, through which angular momentum can be exchanged.

Considering the viscous force between two neighbouring annuli, the torque exerted on the inner annulus by the outer one is

$$G = 2\pi\nu\Sigma_g r^3 \frac{\partial\Omega_K}{\partial r}, \quad (1.8)$$

where ν is the kinematic viscosity. Since Ω_K decreases with radius, the torque on the inner annulus is negative, thus the equal and opposite torque on the outer annulus is positive, i.e. angular momentum is transported radially outwards. The conservation of mass for an axisymmetric thin disc can be expressed as

$$r \frac{\partial\Sigma_g}{\partial t} + \frac{\partial}{\partial r}(r\Sigma_g v_{r,g}) = 0, \quad (1.9)$$

and the conservation of angular momentum ($\Sigma_g r^2 \Omega_K$) can be expressed similarly:

$$r \frac{\partial}{\partial t}(\Sigma_g r^2 \Omega_K) + \frac{\partial}{\partial r}(r^3 \Sigma_g \Omega_K v_{r,g}) = \frac{1}{2\pi} \frac{\partial G}{\partial r}. \quad (1.10)$$

Combining these two equations and making use of the viscous torque from equation (1.8) yields the time evolution of the gas surface density:

$$\frac{\partial\Sigma_g}{\partial t} = \frac{3}{r} \frac{\partial}{\partial r} \left[r^{1/2} \frac{\partial}{\partial r} (\nu \Sigma_g r^{1/2}) \right]. \quad (1.11)$$

Being second-order in the spatial derivative and first-order in time, this is a diffusion equation for the surface density, implying that material diffuses radially, driven by viscosity within the disc. Being non-linear, only numerical solutions are possible, however, insight into this process can be gained by making the straight-forward simplification of spatially constant viscosity, which makes this equation linear. With a sensible choice of initial condition for the radial surface density distribution (e.g a ring), one can show (numerically, or analytically with Green's functions, see e.g. Frank et al. (2002)) that beyond some characteristic radius, a small amount of material carries angular momentum outwards, and interior to that radius mass moves inwards. Furthermore, this characteristic radius itself moves outwards over time. The diffusion may therefore be interpreted as the viscous “spreading” of material. Moreover, by comparing equations (1.9) and (1.11), the velocity at which gas moves radially inwards can be identified as

$$v_{r,g} = -\frac{3}{\Sigma_g r^{1/2}} \frac{\partial}{\partial r} (\nu \Sigma_g r^{1/2}), \quad (1.12)$$

which is referred to as the gas viscous speed.

The physical nature of the kinematic viscosity, ν , is as yet unknown. Lynden-Bell & Pringle (1974) argued that discs should be turbulent due to their high Reynolds number, given estimates of molecular viscosity, and therefore that turbulence may be a source of angular momentum transport. One may consider (e.g. Pringle, 1981) the viscosity that would arise from turbulent motions, whose characteristic length scales are of order the gas scale height, with characteristic speeds of order the sound speed. This is the essence of the “alpha viscosity prescription”, originally proposed by Shakura & Sunyaev (1973):

$$\nu = \alpha c_s H_g. \quad (1.13)$$

While this appears to merely shift the unknown from the physical process into a new, dimensionless parameter α , it enables a physical interpretation of the viscosity in terms of tangible disc properties, namely the sound speed and pressure scale height.

We can gain some insight into how turbulence provides viscosity, and therefore transports angular

1.3. Theoretical Framework for Protoplanetary Discs

momentum, by considering the stress tensor in the Navier-Stokes equations, and cast the dominant terms as an effective α parameter (e.g. Balbus & Hawley, 1998). This leads to an expression of the strength of angular momentum transport in terms of fluctuations in radial and azimuthal velocities, δv_r and δv_ϕ , but also magnetic and gravitational terms:

$$\alpha = \left\langle \frac{\delta v_r \delta v_\phi}{c_s^2} - \frac{B_r B_\phi}{4\pi c_s^2} - \frac{g_r g_\phi}{4\pi G \rho_g c_s^2} \right\rangle. \quad (1.14)$$

This is a density-weighted, time-average of the various stresses that may arise in the flow. The first term represents the hydrodynamic stress, B_r and B_ϕ are the radial and azimuthal components of the magnetic field, contributing to the magnetic stress, and g_r and g_ϕ are the radial and azimuthal components of the gravitational field, giving rise to a gravitational stress. The latter terms are typically weak as discs tend to be non-self-gravitating, as discussed above. The magnetic term may be important in ionised region of discs, and indeed, one of the prevailing views of angular momentum transport in discs is via the magnetorotational instability (MRI, Balbus & Hawley, 1991). In this process, small perturbations in the magnetic field cause small deviations in the flow of ionised fluid elements. The intrinsic shear flow of the Keplerian disc thus amplifies these deviations, and the magnetic tension between neighbouring fluid elements transfers angular momentum. Such magnetic stresses drive turbulent transport of angular momentum with efficiencies of $\alpha \sim 10^{-3}$ – 10^{-2} . As ionisation levels drop off however, the MRI no longer operates and a “dead zone” forms (Gammie, 1996). This process therefore likely only occurs in the hot, inner regions of discs where ionisation levels are high. Alternative processes, such as the removal of angular momentum by magnetised disc winds have since gained favour (e.g. Bai, 2016); suffice it to say there remains significant debate as to the nature of angular momentum transport in discs, and therefore the value of α .

Purely hydrodynamical turbulent processes, such as the vertical shear instability (VSI, Nelson et al., 2013) transport angular momentum with $\alpha \sim 10^{-4}$ (Stoll & Kley, 2014). From an observational perspective, estimates of α can be obtained by considering the measured accretion rates onto host stars. Evolutionary lifetimes of discs (i.e. $M_{\text{disc}}/\dot{M}_*$) yield values of $\alpha \sim 10^{-2}$ (e.g. Hartmann et al., 1998), while parametric models of sub-millimetre observations of discs imply values of $\alpha \sim 10^{-4}$ (e.g. Andrews et al., 2009). Measurements of the velocity dispersions in the TW Hya disc through the broadening of emission lines suggest $\alpha \sim 10^{-3}$ – 10^{-4} (Teague et al., 2016; Flaherty et al., 2018). Pinte et al. (2016) performed detailed modelling of the HL Tau disc, requiring $\alpha \sim 10^{-4}$. However, velocity dispersions can also indicate the presence of gravitational instability within protoplanetary discs, and thus may provide some constraints on the gravitational contribution to α : Hall et al. (2020) identified a characteristic deviation of gas velocities from Keplerian rotation, termed the “GI wiggle”, of which evidence was found in recent ALMA observations of the disc around Elias 2-27 (Paneque-Carreño et al., 2021), and from which it may be possible to derive a value of α (e.g. Terry et al., 2022). Just as the physical mechanism is uncertain, the inferred values of α range substantially in order of magnitude.

A crucial ingredient in analyses of disc evolution is the surface density, and therefore the disc mass. To first order, we should expect the composition of protoplanetary discs to reflect that of the ISM from which they formed. The ISM is inferred to have a dust-to-gas mass ratio of $\sim 1\%$ (Bohlin et al., 1978), i.e. their mass is dominated by that of the gas. Gas masses in protoplanetary discs are notoriously difficult to measure, however: the most abundant species is molecular hydrogen which, being a symmetric molecule, has no permanent dipole moment. The most abundant molecule that does is CO, but this is generally optically thick and therefore traces temperature rather than density. One may look for CO isotopologues, such as ^{13}CO or C^{18}O , which are less likely to be optically thick due to their lower abundances, but then suffer from low signal-to-noise limitations. And even if such species are optically thin, conversion to a meaningful total gas density requires intimate knowledge of the ongoing chemical reactions in the disc,

the pressure and temperature structure of the disc, as well as an understanding of processes such as the freeze-out of gaseous species onto the surfaces of dust grains.

The gas surface density of a viscous disc can be obtained by evolving equation (1.11) over time, given an initial condition. Alternatively, if the viscosity can be expressed as $\nu \propto r^\gamma$ (where γ is some index), equation (1.11) has a self-similar solution (Lynden-Bell & Pringle, 1974), which yields a power-law decline in surface density up to some critical radius, beyond which the surface density exponentially decays:

$$\Sigma_g = \Sigma_c \left(\frac{r}{R_c} \right)^{-\gamma} \exp \left[- \left(\frac{r}{R_c} \right)^{2-\gamma} \right], \quad (1.15)$$

where Σ_c is the gas surface density at the critical radius R_c . One interesting prediction is from the “minimum-mass Solar nebula” (MMSN, Weidenschilling, 1977a; Hayashi, 1981), a model of the Sun’s circumstellar disc from which the Solar System planets formed, which suggests $\Sigma \propto r^{-3/2}$, based on the masses and orbital locations of the planets as observed today. Sub-millimetre imaging surveys of discs suggest $\Sigma_d \propto r^{-1}$ on larger scales (e.g. Andrews et al., 2009, 2010, 2011) but, as will be shown in §1.4.1, the distribution of large grains can differ significantly from that of the gas. For simplicity, many models assume a power law, typically $\Sigma \propto r^{-1}$, i.e. $\gamma = 1$.

Along with the gas surface density, the radial temperature profile sets the pressure structure of the disc, and therefore the vertical scale height of the disc. Kenyon & Hartmann (1987) performed fits to SEDs of discs, suggesting that discs are flared, wherein the aspect ratio increases with distance from star, according to $H_g \propto r^{9/8}$. D’Alessio et al. (2001) presented models of irradiated discs, solving self-consistently for their vertical structure, finding $H_g \propto r^{5/4}$. This was justified by Chiang & Goldreich (1997), who analytically considered the temperature of a thin, passively heated disc, whose temperature is set solely by flux from the central star, arriving at $T(r) \propto r^{-3/7}$, which corresponds to $H_g \propto r^{9/7}$. Thus, the pressure scale height is typically assumed to take the form

$$H_g = h_0 R_0 \left(\frac{r}{R_0} \right)^q, \quad (1.16)$$

where h_0 is the aspect ratio of the disc at some reference radius R_0 . For simplicity, many models adopt $T \propto r^{-1/2}$ and $H_g \propto r^{5/4}$. The aspect ratios inferred from both models and observations are typically a few percent ($\sim 0.05 - 0.1$, e.g. Kenyon & Hartmann, 1987; D’Alessio et al., 2001; Andrews et al., 2011; Dullemond et al., 2018; Law et al., 2021b).

In the above calculations I assumed that the circumstellar gas orbits at the Keplerian angular frequency. While this serves as a good approximation, a simple consideration of the radial force balance shows that the force due to the thermal pressure gradient causes the gas to orbit at a slightly different frequency:

$$r\Omega_g^2 - \frac{GM_*r}{(r^2 + z^2)^{3/2}} - \frac{1}{\rho_g} \frac{\partial P}{\partial r} = 0. \quad (1.17)$$

Considering, for simplicity, the mid-plane, i.e. $z = 0$, and substituting $\Omega_K = \sqrt{GM_*/r^3}$:

$$r\Omega_g^2 - r\Omega_K^2 - \frac{1}{\rho_g} \frac{\partial P}{\partial r} = 0. \quad (1.18)$$

Solving now for Ω_g :

$$\Omega_g^2 = \Omega_K^2 \left(1 + \frac{1}{r\rho_g\Omega_K^2} \frac{\partial P}{\partial r} \right). \quad (1.19)$$

This is usually written as:

$$\Omega_g^2 = \Omega_K^2 (1 - \eta), \quad (1.20)$$

1.3. Theoretical Framework for Protoplanetary Discs

where

$$\eta = -\frac{1}{r\rho_g\Omega_K^2}\frac{\partial P}{\partial r} = -\frac{c_s^2}{v_K^2}\frac{\partial \log P}{\partial \log r}. \quad (1.21)$$

η is the ratio of the force due to the thermal pressure gradient to the gravitational force on a gas fluid element due to the central star (Adachi et al., 1976; Weidenschilling, 1977b). This therefore shows that gas does not orbit at the Keplerian speed, but instead at a speed modified by the radial pressure gradient. In the typical case where the gas pressure decreases with distance from the star (as both temperature and density decrease with distance from the star), the gas orbits at a sub-Keplerian speed.

It can be illuminating to re-write η in terms of some more tangible disc properties. Substituting $P = c_s^2\rho_g$ into equation (1.21) and expanding the derivative using the product rule,

$$\eta = -\frac{1}{r\rho_g\Omega_K^2}\left(c_s^2\frac{\partial \rho_g}{\partial r} + \rho_g\frac{\partial c_s^2}{\partial r}\right). \quad (1.22)$$

Assuming that the disc temperature follows a power law of index q , this can be written as

$$\eta = -\frac{1}{r\rho_g\Omega_K^2}\left(c_s^2\frac{\partial \rho_g}{\partial r} + qc_s^2\frac{\rho_g}{r}\right). \quad (1.23)$$

Substituting for the gas scale height, $H_g = c_s/\Omega_K$, yields

$$\eta = -\left(\frac{H_g}{r}\right)^2\left(\frac{\partial \log \rho_g}{\partial \log r} + q\right). \quad (1.24)$$

Since $c_s^2 \propto r^q$ and $\Omega_K \propto r^{-3/2}$, we can write

$$H_g \propto \frac{r^{q/2}}{r^{-3/2}} \propto r^{(q+3)/2} \propto r^k, \quad (1.25)$$

where $k = (q + 3)/2$. Thus, in terms of k ,

$$\eta = -\left(\frac{H_g}{r}\right)^2\left(\frac{\partial \log \rho_g}{\partial \log r} + 2k - 3\right). \quad (1.26)$$

This shows that the gas orbital speed deviates from the Keplerian speed by an amount which depends on the disc aspect ratio, the flaring index and the radial density gradient. In the absence of a pressure gradient, the gas will orbit at the Keplerian speed – this may happen at, for example, a pressure bump in the disc, resulting in a localised enhancement in gas density. Note that a simple enhancement in density is insufficient – the enhancement must be sufficiently strong to minimise the pressure gradient, which depends on the temperature profile. We can also see that if the density gradient instead *increases* with distance from the star, as would be the case on the inner side of the pressure bump, the gas can orbit at a super-Keplerian speed. This is of particular importance in the interpretation of substructure in (sub-)millimetre wavelength observations of discs, as will be shown in the next section.

1.3.3 Circumstellar dust

Thus far I have only considered the dynamics and distribution of the gas. While the solid component of protoplanetary discs only contributes 1% of their mass, it is these dust grains that planetary cores are formed from, thus understanding their properties and dynamics is essential, especially considering the $\gtrsim 10$ order of magnitude difference in size between interstellar dust grains and planetary cores. Mathis et al. (1977) performed a fit to the observed interstellar extinction between wavelengths of 0.11 μm and 1 μm . They determined that the number of dust grains, n , with a given size, s , follows a power law

distribution:

$$\frac{dn}{ds} \propto s^{-7/2}, \quad (1.27)$$

between grain sizes of $0.025 \mu\text{m}$ and $0.25 \mu\text{m}$. $n(s)ds$ is the number of grains with sizes in the interval $(s, s+ds)$ – this is the “MRN distribution” of dust grain sizes. We might naively expect this size distribution to hold within protoplanetary discs; how, then, are planetary cores formed from such grains? For spherical grains with mass $m \propto s^3$ this implies a mass distribution of

$$\frac{dn}{dm} \propto m^{-11/6}. \quad (1.28)$$

This is consistent with the result obtained by Dohnanyi (1969), who considered a collisional model of asteroids and their debris, suggesting that a collisional cascade produces a distribution $dn/dm \propto m^{-1.837}$, an index remarkably close to $-11/6$. An MRN distribution of sizes is therefore not limited to sub-micron-sized grains.

Early planet formation theory suggests that dust grains collide and grow through sticking (e.g. Safronov, 1972; Goldreich & Ward, 1973). Turbulence in the gas drives collisions between dust grains which, for sufficiently low velocities, can result in sticking due to van der Waals interactions or electrostatic forces (e.g. Weidenschilling & Cuzzi, 1993). This growth through collisions is referred to as “coagulation”. The relative velocities between dust grains, driven by turbulence, increases with grain size, thus coagulation cannot proceed uninhibited, as larger grains eventually fragment upon collision (e.g. Dominik & Tielens, 1997; Blum & Wurm, 2008).

Birnstiel et al. (2011) developed an analytical model for a population of spherical dust grains in coagulation–fragmentation equilibrium, wherein primitive dust grains grow through collisions, eventually reaching sizes at which the collision velocities result in fragmentation, replenishing the smaller grains. They performed numerical simulations of the evolution of such a population to find a steady-state size distribution. Their results support an MRN size distribution, with the maximum grain size given by the fragmentation limit:

$$s_{\text{frag}} \approx \frac{2\Sigma_{\text{g}}}{\pi\alpha\rho_{\bullet}} \frac{v_{\text{frag}}^2}{c_{\text{s}}^2} \quad (1.29)$$

where v_{frag} is the fragmentation velocity and $\rho_{\bullet}$ is the material density of dust grains, typically considered to be of order 1 g cm^{-3} , depending on the assumed composition (e.g. Birnstiel et al., 2018). The fragmentation velocity depends on the composition and structure of the dust grains, estimated to be $\sim 1 \text{ m s}^{-1}$ for silicate grains and 10 m s^{-1} for icy grains (Gundlach & Blum, 2015).

Commonly adopted compositions of dust grains in protoplanetary discs are magnesium-iron silicates, which may host icy mantles in the cold, outer regions of protoplanetary discs (Natta et al., 2007). Dust grains are typically assumed to be spherical, although this is likely a gross simplification. Both laboratory experiments and numerical simulations show that coagulation tends to produce non-spherical, porous aggregates, the bulk density of which decreases with size (e.g. Meakin & Donn, 1988; Dominik & Tielens, 1997; Blum & Wurm, 2008; Blum, 2018). However, collisions and drag forces from the gas tend to cause erosion and compaction of the aggregates (e.g. Miyake & Nakagawa, 1993; Dominik & Tielens, 1997).

Evidence of growth from sub-micron-sized grains up to larger sizes can be obtained from observations of protoplanetary discs. If the thermal emission from dust grains with opacity κ_{ν} at a temperature T_{d} is optically thin, the flux measured at a distance D is (e.g. Hildebrand, 1983)

$$F_{\nu} = \frac{M_{\text{d}}\kappa_{\nu}B_{\nu}(T_{\text{d}})}{D^2}, \quad (1.30)$$

where M_{d} is the dust mass and B_{ν} is the Planck function which, in the Rayleigh-Jeans (i.e. long-

1.4. Substructure in Protoplanetary Discs

wavelength) regime, is well-approximated by $B_\nu \propto \nu^{-2}$. If the opacity follows a power-law in frequency, i.e. $\kappa_\nu \propto \nu^\beta$, the gradient of the SED is the spectral index:

$$\alpha = \frac{d \log F_\nu}{d \log \nu}, \quad (1.31)$$

where $\alpha = 2 + \beta$ (Beckwith et al., 1990). By performing multi-wavelength observations, the spectral index between two wavelengths can be measured, ultimately yielding a value of β . By modelling the size-dependent opacities of dust grains, one can thereby determine the sizes of the emitting grains. Better still, by performing spatially resolved imaging, for example at ALMA bands 6 and 7 (1.3 mm and 0.88 mm), the spectral index between these two wavelengths can be calculated over the disc. This technique typically yields millimetre to centimetre sizes (e.g. Pérez et al., 2012, 2015b; Tazzari et al., 2016, 2021a,b).

An alternative method is from measuring the polarisation emission at millimetre wavelengths since, if the dust grains have sizes comparable to the observation wavelength, their scattering opacities should be large (e.g. Kataoka et al., 2015, 2017). This method typically yields grain sizes of few hundred microns (e.g. Ohashi & Kataoka, 2019; Ohashi et al., 2020; Mori & Kataoka, 2021) and is thus the subject of a growing debate in the community. Polarised emission can also arise from the alignment of dust aggregates with the magnetic field (e.g. Lazarian, 2007; Andersson et al., 2015). Furthermore, the contribution to the polarised emission from scattering depends on the viewing angle of the disc as well as the location in the disc. The significant uncertainties in dust grain composition, internal structure and density, as well as optical properties, make this a challenging problem to solve. Regardless, there is overwhelmingly strong evidence for grain growth in protoplanetary discs beyond primitive, interstellar dust grains.

1.4 Substructure in Protoplanetary Discs

With a description of the density and temperature structure of protoplanetary discs, as well as an understanding of the gas dynamics and the size distribution of dust grains, it is now possible to attempt to interpret observations of protoplanetary discs, particularly with respect to the substructure suggested from SED measurements and infrared images, and understand the physical processes that may be responsible for creating them. Understanding the origin of these substructures, in particular if and how they are related to planet formation, is key to understanding the formation and architecture of planetary systems, as well as reconciling our understanding of planet formation with the observed exoplanet population.

1.4.1 Axisymmetric substructure

Initially proposed to explain the lack of near-infrared excess in the spectral energy distributions of “transition discs”, the existence of substructure in the form of inner cavities was confirmed with early (sub-)millimetre observations (Brown et al., 2009; Andrews et al., 2009; Hughes et al., 2009; Isella et al., 2010; Andrews et al., 2011; Mathews et al., 2012; Tang et al., 2012; Fukagawa et al., 2013). As discussed in §1.2, high spatial resolution observations of young stellar objects with ALMA have since revealed the wealth of substructure present in their protoplanetary discs. Many discs observed at millimetre wavelengths with ALMA show concentric rings of emission from millimetre-sized dust grains (e.g. ALMA Partnership et al., 2015b; Andrews et al., 2016; Fedele et al., 2017; Andrews et al., 2018; Long et al., 2018). However, one of the primary confounding factors in interpreting these observations is that they generally probe the dust component, which does not necessarily directly trace the gas that drives the dynamics.

Whipple (1972) demonstrated how the aerodynamic interaction between the gas and dust can concentrate dust at a local axisymmetric pressure maximum. Equation (1.20) states that gas does not orbit the central star at the Keplerian frequency, but at a speed modified by a factor $\sqrt{1-\eta}$, due to pressure support. The acceleration of a dust grain due to this force is negligible (it is of order the mass ratio of a dust grain to a hydrogen molecule lower than the acceleration of the gas), therefore its orbital speed is unaffected directly by gas pressure. However, this difference in speeds results in a drag force, such that the relative velocity $\Delta \mathbf{v}$ decays on a time-scale t_s , namely the stopping time:

$$t_s = \frac{m_{\text{dust}} |\Delta \mathbf{v}|}{|\mathbf{F}_D|}. \quad (1.32)$$

Assuming spherical dust grains of radius s , the full expression for the drag force is

$$\mathbf{F}_D = -\frac{C_D}{2} \pi s^2 \rho_g \Delta \mathbf{v}^2. \quad (1.33)$$

The drag coefficient, C_D , depends on the Reynolds number and viscosity of the flow. For dust grains in protoplanetary discs, where the mean free path is much larger than the size of dust grains and the relative velocity is small compared to the mean kinetic molecular speed, one can make use of the Epstein approximation (Epstein, 1924; Whipple, 1972) for the drag coefficient, such that the drag force can be written as

$$\mathbf{F}_D \approx -\frac{4\pi}{3} s^2 \rho_g v_{\text{th}} \Delta \mathbf{v}, \quad (1.34)$$

where v_{th} is the mean thermal speed of the gas:

$$v_{\text{th}} = \sqrt{\frac{8k_B T}{\pi \mu m_H}}, \quad (1.35)$$

with μ the mean molecular weight. In this case, the stopping time can be expressed as

$$t_s = \sqrt{\frac{\pi}{8}} \frac{s \rho_{\bullet}}{\rho_g c_s}. \quad (1.36)$$

The stopping time frequently appears alongside the Keplerian frequency in expressions, i.e. as a ratio to the local dynamical time-scale. The product of these two gives the Stokes number, $\text{St} = t_s \Omega_K$, a dimensionless quantity which represents the strength of the coupling between the dust and gas. Thus, for spherical grains of radius s and density $\rho_{\bullet}$, the Stokes number at the disc midplane is

$$\text{St} = \frac{\pi s \rho_{\bullet}}{2 \Sigma_g}. \quad (1.37)$$

Dust grains with $\text{St} \ll 1$ have their velocity relative to the gas rapidly suppressed through drag forces, and are thus strongly coupled to the gas flow. At the other extreme, objects with $\text{St} \gg 1$ are decoupled from the gas flow and essentially remain on Keplerian orbits – this typically corresponds to large (e.g. kilometre-sized) bodies for typical protoplanetary disc conditions.

This drag force necessarily results in an exchange of angular momentum between the gas and dust. If the gas orbits at a sub-Keplerian speed, the dust grains lose angular momentum and drift inwards towards the star, a process termed “radial drift”. Note that this is characterised by the η parameter, the sign of which depends on the gas pressure gradient (equation (1.21)). If there exists a positive pressure gradient in the disc, the gas will orbit at a super-Keplerian speed, giving angular momentum to the dust grains and therefore moving them to larger orbital radii. When $\eta = 0$, gas orbits at the Keplerian velocity so there is no drag force on the dust, thus no dust drift occurs. A local pressure maximum will

1.4. Substructure in Protoplanetary Discs

therefore trap dust grains.

Takeuchi & Lin (2002) calculated the terminal radial drift speed of dust grains as a function of their Stokes number:

$$v_{r,d} = \frac{v_{r,g}/St - \eta v_K}{St + 1/St}. \quad (1.38)$$

This shows a dependence on η , which is minimised at gas pressure maxima, at which point dust drift speed is of order the viscous speed of the gas (for $St \lesssim 1$; larger bodies remain on Keplerian orbits). This drift speed is maximised for $St = 1$ which, under typical protoplanetary disc conditions at $\sim$ tens of astronomical units, corresponds to dust grains with sizes

$$s \sim 2 \text{ cm} \left(\frac{\Sigma_g}{10 \text{ g cm}^{-2}} \right) \left(\frac{\rho_\bullet}{3 \text{ g cm}^{-3}} \right)^{-1} St. \quad (1.39)$$

Dust grains with a Stokes number of order unity will therefore be trapped most efficiently.

Note that this interaction between the gas and dust grains also results in dust grains settling towards the disc midplane: while the gas is pressure-supported against the vertical component of stellar gravity, dust grains are not, and thus accelerate towards the midplane until reaching a terminal velocity due to gas drag. Turbulence in the gas also stirs dust grains in the vertical direction, causing them to settle into a layer of finite thickness, dependent on their Stokes number (e.g. Dubrulle et al., 1995):

$$\frac{H_d}{H_g} \approx \sqrt{\frac{\alpha}{St}}. \quad (1.40)$$

This approximation is valid in the limit that the dust-to-gas ratio is small, which clearly breaks down as larger grains settle in a thinner layer. However, recent ALMA observations of edge-on discs show that millimetre-sized grains occupy a very thin layer (e.g. Villenave et al., 2020, 2022), while NIR observations reveal the large vertical extent of micron-sized grains (e.g. Avenhaus et al., 2018).

Just as turbulent diffusion resists the settling of dust grains at the disc midplane, it also acts to resist trapping at radial pressure maxima, where $\eta = 0$. Clarke & Pringle (1988) calculated the diffusive flux as

$$\mathbf{v}_{\text{diff}} = -\frac{\nu}{Sc} \frac{\Sigma_g}{\Sigma_d} \nabla \left(\frac{\Sigma_d}{\Sigma_g} \right), \quad (1.41)$$

where Sc is the Schmidt number, which is the ratio of momentum diffusivity to mass diffusivity, typically assumed to be of order unity (e.g. Johansen & Klahr, 2005; Youdin & Lithwick, 2007).

The concentric bright and dark rings observed in protoplanetary discs have therefore been interpreted as being due to the presence of dust traps, with the traps forming due to local maxima in gas pressure (e.g. Pinilla et al., 2012a; Dullemond et al., 2018; Rosotti et al., 2020b). Observational confirmation of this remains elusive though, as the corresponding gas substructures are yet to be identified, given the difficulties in directly measuring gas densities described above. It is therefore unknown whether gas pressure maxima capable of trapping dust at the observed locations exist. Fortunately, the MAPS programme (Öberg et al., 2021) is beginning to explore this observationally (Law et al., 2021a), and techniques are being developed to make inferences from gas kinematics (e.g. Teague et al., 2019).

A variety of plausible mechanisms have been shown to form axisymmetric pressure maxima, e.g. planet-carved gaps (Lin & Papaloizou, 1979; Rice et al., 2006; Zhu et al., 2012), photoevaporation (Clarke et al., 2001; Alexander et al., 2006; Alexander & Armitage, 2007; Owen & Kollmeier, 2019), magnetohydrodynamical effects (Johansen et al., 2009; Uribe et al., 2011; Bai & Stone, 2014), processes which reduce the local viscosity (Johansen et al., 2011; Dullemond & Penzlin, 2018) and processes that can naturally increase the local dust density, such as the freeze-out of gaseous species at their respective icelines (Cuzzi & Zahnle, 2004; Kretke & Lin, 2007; Okuzumi et al., 2016; Owen, 2020). These can

explain the origin of the observed rings, each with varying degrees of success.

Perhaps the most prevalent explanation is that of gap-opening by planets, which results due to the gravitational perturbation from the planet on the disc. Goldreich & Tremaine (1979) showed that the planetary potential exerts torques on the disc at specific orbital resonances, exciting waves which transport angular momentum away from the planet. Material interior to the planet loses angular momentum and moves towards the central star, while material exterior to the planet gains angular momentum, thus moving outwards. This opens an annular gap at the planet's semimajor axis (e.g. Papaloizou & Lin, 1984).

Opening a gap requires a sufficiently massive planet, depending on the physical properties of the disc. Specifically, the rate of angular momentum transfer from the tidal torque must exceed the viscous angular momentum transfer rate, and the extent of the planet's Roche lobe must be comparable to the scale height of the disc. Lin & Papaloizou (1986a) showed that once the planet's Roche lobe becomes comparable to the scale height, the waves excited by the planet reach the sound speed, resulting in shock formation and thus angular momentum exchange. The width of the gap is therefore expected to be of order the scale height of the disc. This reasoning leads to the concept of the thermal mass, i.e. the mass at which the planet's Roche lobe is comparable in size to the gas scale height. The size of the Roche lobe in this limit is typically approximated by the Hill radius which, for a planet of mass M_P orbiting a star of mass M_* with a semimajor axis a , is

$$R_H = a \left(\frac{M_P}{3M_*} \right)^{1/3}, \quad (1.42)$$

thus the thermal mass is often taken to be the mass at which the planet's Hill radius equals the scale height, leading to the following expression:

$$M_{\text{th}} = 3M_* h^3. \quad (1.43)$$

If the disc viscosity is sufficiently high, however, the viscous transport of material would exceed the rate at which the planet repels it, thus reaching the thermal mass is not a sufficient condition for gap opening. Bryden et al. (1999) considered gap-opening in a viscous disc, proposing

$$M_{\text{gap}} \approx 40M_* \alpha h^2. \quad (1.44)$$

Crida et al. (2006) built upon this, accounting for the planet mass and disc aspect ratio, arriving at the widely-adopted gap-opening criterion:

$$\frac{3H_g}{4R_H} + \frac{50M_* \alpha h^2}{M_P} \lesssim 1. \quad (1.45)$$

It has been demonstrated theoretically (e.g. Goodman & Rafikov, 2001; Rafikov, 2002) and via numerical simulations (e.g. Dong et al., 2011; Duffell & MacFadyen, 2012) that even low mass planets can open gaps, however. Furthermore, Duffell & MacFadyen (2013) showed that equation (1.45) is not valid for low viscosities and low planet masses. Complex treatments of gap-opening remain a challenge, thus empirical scalings of gap widths as a function of planet mass, aspect ratio and viscosity (e.g. Duffell & MacFadyen, 2013; Fung et al., 2014; Duffell, 2015; Kanagawa et al., 2015) are often invoked.

Given the degeneracies between planet mass, viscosity and scale height, it is difficult to infer a planet mass from an observed gap depth and width without additional constraints. Establishing whether the bright and dark rings observed in discs are indeed due to gaps opened by planets is therefore difficult. Zhang et al. (2018) attempted to do this for the discs observed in the DSHARP survey (see figure 1.4),

1.4. Substructure in Protoplanetary Discs

inferring a population of planets showing little-to-no agreement with the observed exoplanet population in the planet mass vs. semimajor axis plane. Of course, exoplanets at these orbital separations are typically beyond our current detection methods, thus the discrepancy is not fatal for this hypothesis.

1.4.2 Non-axisymmetric substructure

While rings appear to be ubiquitous in high-resolution images of protoplanetary discs, a small fraction show departures from axisymmetry, both at near-infrared and (sub-)millimetre wavelengths, specifically in the form of spiral arms and crescent or horseshoe features, the interpretations of which I now discuss.

Crescent features

A handful of discs show crescent-shaped features in their (sub-)millimetre emission which, following the hypothesis that substructure at these wavelengths is due to dust trapping, suggests that the dust in these discs is concentrated azimuthally. These asymmetries can be small in scale and exist alongside axisymmetric features (e.g. Andrews et al. 2018: HD 143006, HD 163296 and Hashimoto et al. 2021: DM Tau), while in the most extreme cases all the millimetre emission comes from one large-scale crescent (e.g. van der Marel et al. 2013: IRS 48; Ohashi 2008; Casassus et al. 2013: HD 142527; Pérez et al. 2014: SAO 206462, SR 21).

A number of mechanisms have been proposed to explain the origin of the observed non-axisymmetric substructures. One interpretation of these observations is that as-yet undetected stellar/substellar companions are present in these systems, which drive eccentricity in the discs, and the observed asymmetry arises simply due to the reduced orbital velocity of material at apocentre (Ataiee et al., 2013). However, the azimuthal density contrast achieved in this process is inconsistent with observations. Companions on eccentric orbits can also generate similar signatures as they carve eccentric cavities in the circumbinary disc, creating a non-axisymmetric density profile at the cavity edge (e.g. Ragusa et al., 2017; Poblete et al., 2019; Ragusa et al., 2020), though it is difficult to understand how this mechanism can explain the more localised trapping of larger solids seen at longer wavelength observations, e.g. IRS 48 (van der Marel et al., 2013, 2015). An alternative explanation proposed by Mittal & Chiang (2015) is that, since any inherent large-scale azimuthal asymmetry in the disc shifts the barycentre of the system, in sufficiently massive discs this can give rise to an $m = 1$ azimuthal mode, causing the gas to form a horseshoe-shaped annulus that concentrates dust grains via drag (Zhu & Baruteau, 2016; Baruteau & Zhu, 2016). This mechanism requires particularly massive, self-gravitating discs, however. While many of the discs which show such asymmetries *are* massive, they have been determined to be gravitationally stable. Furthermore, self-gravity acts to alter the trapping locations of dust grains as a function of their size, which is in contention with the multi-wavelength observations of HD 142527 and IRS 48.

The most prevalent interpretation, however, is that an anticyclonic vortex forms in the gas, trapping the dust azimuthally (e.g. Lyra & Lin, 2013). An anticyclonic vortex is befitting as it represents a pressure maximum and is therefore capable of trapping dust in a similar fashion to the axisymmetric pressure maximum described above (e.g. Barge & Sommeria, 1995; Tanga et al., 1996; Klahr & Henning, 1997; Godon & Livio, 1999, 2000). Moreover, anticyclonic vortices are stable to the intrinsic shear of the Keplerian flow and therefore have the potential to be long-lived. However, it is uncertain whether they are stable in real, three-dimensional discs, where they are subject to dissipation by viscosity (e.g. Meheut et al., 2012), in addition to destruction via elliptical instabilities (Kerswell, 2002; Lesur & Papaloizou, 2009; Lyra & Klahr, 2011).

It is therefore worth considering how vorticity can evolve in a protoplanetary disc to see how it might

be generated. Starting with the momentum equation for the gas, neglecting viscosity,

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla) \mathbf{v} = -\frac{1}{\rho} \nabla P - \nabla \Phi, \quad (1.46)$$

where Φ is the gravitational potential, the vorticity, $\boldsymbol{\omega} = \nabla \times \mathbf{v}$, can be exposed by rewriting the advection term:

$$(\mathbf{v} \cdot \nabla) \mathbf{v} = \frac{1}{2} \nabla \mathbf{v} \cdot \mathbf{v} - \mathbf{v} \times \nabla \times \mathbf{v}. \quad (1.47)$$

Taking the curl of the momentum equation thus yields

$$\nabla \times \frac{\partial \mathbf{v}}{\partial t} + \nabla \times \frac{1}{2} \nabla \mathbf{v} \cdot \mathbf{v} - \nabla \times \mathbf{v} \times \boldsymbol{\omega} = -\nabla \times \frac{1}{\rho} \nabla P - \nabla \times \nabla \Phi. \quad (1.48)$$

The kinetic energy and gravitational potential terms vanish since the curl of a gradient is zero. The third term can be rewritten as

$$\nabla \times \mathbf{v} \times \boldsymbol{\omega} = -(\mathbf{v} \cdot \nabla) \boldsymbol{\omega} + (\boldsymbol{\omega} \cdot \nabla) \mathbf{v} - \boldsymbol{\omega} (\nabla \cdot \mathbf{v}) - \mathbf{v} (\nabla \cdot \boldsymbol{\omega}). \quad (1.49)$$

The last term on the RHS is zero since the divergence of the curl of a quantity is zero. We thus have

$$\frac{\partial}{\partial t} (\nabla \times \mathbf{v}) + (\mathbf{v} \cdot \nabla) \boldsymbol{\omega} - (\boldsymbol{\omega} \cdot \nabla) \mathbf{v} + \boldsymbol{\omega} (\nabla \cdot \mathbf{v}) = -\nabla \times \frac{1}{\rho} \nabla P \quad (1.50)$$

Adopting the notation for the advective derivative,

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + (\mathbf{v} \cdot \nabla), \quad (1.51)$$

we can write

$$\frac{D \boldsymbol{\omega}}{Dt} - (\boldsymbol{\omega} \cdot \nabla) \mathbf{v} + \boldsymbol{\omega} (\nabla \cdot \mathbf{v}) = -\nabla \times \frac{1}{\rho} \nabla P. \quad (1.52)$$

In order to account for mass conservation, we can make use of the continuity equation which, in its advective form, is

$$\frac{D \rho}{Dt} = -\rho \nabla \cdot \mathbf{v}. \quad (1.53)$$

Thus, substituting for $\nabla \cdot \mathbf{v}$ in equation (1.52), we obtain

$$\frac{D \boldsymbol{\omega}}{Dt} - (\boldsymbol{\omega} \cdot \nabla) \mathbf{v} - \frac{\boldsymbol{\omega}}{\rho} \frac{D \rho}{Dt} = -\nabla \times \frac{1}{\rho} \nabla P. \quad (1.54)$$

Combining the time derivatives, we obtain an expression for the rate of change of vortensity, $\boldsymbol{\omega}/\rho$:

$$\frac{D}{Dt} \left(\frac{\boldsymbol{\omega}}{\rho} \right) = \left(\frac{\boldsymbol{\omega}}{\rho} \cdot \nabla \right) \mathbf{v} - \frac{1}{\rho} \nabla \times \frac{1}{\rho} \nabla P. \quad (1.55)$$

Application of the quotient rule to last term yields

$$\frac{D}{Dt} \left(\frac{\boldsymbol{\omega}}{\rho} \right) = \left(\frac{\boldsymbol{\omega}}{\rho} \cdot \nabla \right) \mathbf{v} + \frac{\nabla \rho \times \nabla P}{\rho^3}. \quad (1.56)$$

This is the equation for the conservation of vortensity. The first term on the right-hand side represents the change in vortensity due to vortex tilting or stretching, which vanishes in protoplanetary discs with no flow in the vertical direction. The second term is the baroclinic term; in barotropic flows, where density and pressure gradients are aligned, vortensity is advectively conserved. In baroclinic flows, however, vortensity is not conserved and vortices can be generated.

1.4. Substructure in Protoplanetary Discs

Axisymmetric Keplerian discs are, in general, stable to the formation of vortices. In the absence of viscosity, vortensity is a conserved quantity. Vortices are therefore of interest because, once generated, the conservation of vortensity implies that they can be long-lived. Viscosity acts to dissipate vortensity, however, the viscous time-scales in protoplanetary discs are generally orders of magnitude longer than the dynamical time-scales, allowing vortices to survive for hundreds of orbits.

If vortices are sufficiently large, their velocity perturbation becomes supersonic, thus creating a shock which causes them to lose energy. This can be seen by considering a patch of the disc containing the vortex, small enough that the local approximation (the “shearing sheet”, Goldreich & Lynden-Bell, 1965) can be employed. Within this patch the orbital motion around the central star can be treated as uniform (rigid body) rotation plus a linear shear flow. If we consider a Cartesian coordinate system (x, y) centred on and rotating with the core of the vortex, such that $\hat{x}$ points in the radial direction away from the star and $\hat{y}$ points in the azimuthal direction, i.e. along the global sense of rotation around the star, then

$$\mathbf{v}_y = -S_0 x \hat{\mathbf{y}}, \quad (1.57)$$

where $S_0 = S(x=0)$ is the rate of orbital shear, given by:

$$S = -r \frac{d\Omega}{dr}. \quad (1.58)$$

For a Keplerian disc, $S_0 = \frac{3}{2} \Omega_{K,0}$, thus

$$\mathbf{v}_y = -\frac{3}{2} \Omega_{K,0} x \hat{\mathbf{y}}. \quad (1.59)$$

It is then straight-forward to show that the distance from the centre of the vortex at which the velocity reaches the sound speed is $x = \frac{2}{3} H_g$. This sets a maximum semiminor axis of an anticyclonic vortex in a Keplerian disc since, if they exceed this, the sound waves radiated deplete their energy.

Interest in vortices was instigated when Barge & Sommeria (1995), considered the trajectories of dust grains in the vicinity of a vortex. They demonstrated that vortices trap dust grains with an efficiency that is maximised for dust grains of Stokes number ~ 1 , due to their pressure structure. If a vortex can concentrate solids to the point that the local dust density can exceed the Roche density, gravitational collapse and thus planetesimal formation may ensue (e.g. Lyra et al., 2008, 2009).

Large-scale crescent-shaped features have most commonly been observed in transition discs. It is therefore interesting to consider whether the mechanism responsible for creating the transition disc signature can also be related to the formation of azimuthal asymmetries. Vortex formation in protoplanetary discs has been shown to proceed through various mechanisms. Klahr & Bodenheimer (2003) demonstrated that in discs with a radial entropy gradient, azimuthal perturbations are baroclinically unstable and can lead to the formation of vortices. The most commonly invoked mechanism to form a large-scale vortex in a protoplanetary disc is the Rossby wave instability (RWI). The RWI can be triggered at steep radial density gradients (Lovelace et al., 1999; Li et al., 2000, 2001); such a gradient may be generated at the viscosity transition between regions in which the magnetorotational instability can operate (e.g. Varnière & Tagger, 2006; Lyra et al., 2009; Lyra & Mac Low, 2012). Steep radial density gradients are also found at the outer edge of a gap carved in the disc by a giant planet. Indeed, hydrodynamics simulations of protoplanetary discs containing a giant planet have successfully produced large-scale vortices (e.g. Masset et al., 2006; de Val-Borro et al., 2007; Meheut et al., 2010; Lin & Papaloizou, 2011; Lin, 2012). The subject of vortex formation and its connection to planet formation will be explored further in chapters 3 and 4.

Spiral arms

A number of discs when imaged at near-IR wavelengths, tracing micron-sized grains at the disc surface, show spiral arms (e.g. SAO 206462: Muto et al. 2012; MWC 758: Grady et al. 2013; HD 100453: Wagner et al. 2015; V1247 Ori, Ohta et al. 2016; LkH α 330: Akiyama et al. 2016; Uyama et al. 2018; CQ Tau: Uyama et al. 2020; WaOph 6: Brown-Sevilla et al. 2021; SR 21: Follette et al. 2013; Muro-Arena et al. 2020). These have primarily been interpreted in terms of the spiral density wave theory of Lin & Shu (1964). As described above in the case of gap-opening, wherein the planet launches waves as it exchanges angular momentum with the disc, the planet exerts a periodic force on fluid elements, which resonates with the epicyclic frequency at Lindblad resonances (Goldreich & Tremaine, 1979). The Keplerian velocity profile shears this wave into a spiral pattern.

Early work in this area focussed on the interaction between giant planets and the disc, with hydrodynamics simulations confirming theoretical predictions (e.g. Bryden et al., 1999; Kley, 1999), though spiral arm patterns had also been seen in simulations of discs containing a terrestrial mass planet. Ogilvie & Lubow (2002) considered the linear waves generated by a low mass planet, and showed that the Fourier modes interfere constructively to produce a one-armed spiral pattern.

Although the substructures discussed thus far have related to the distribution of dust, spiral arms are generated in the gas, thus it is important to understand how this affects the dust distribution that is probed in both near-IR and (sub-)millimetre wavelength observations. Indeed, given the eagerness to propose that annular substructures are created by planets, it is perhaps surprising that the spiral arms they are expected to generate are not observed.

Juhász et al. (2015) showed that the spiral arms generated by planets in 2D simulations have a density contrast almost an order of magnitude lower than those in observations. Although their simulations were gas-only, the small grains probed in near-IR observations are expected to follow the gas due to their low Stokes numbers. Zhu et al. (2015) performed a similar analysis using 3D simulations, demonstrating the importance of the vertical dimension in this problem, and achieving brightness contrasts more compatible with observations.

The situation is complicated, however, by the fact that corresponding spiral arms are not seen in (sub-)millimetre observations of the same discs, and that in the discs that do show spiral arms, the spiral arms do not appear to be associated with annular gaps as one would expect (e.g. Brown-Sevilla et al., 2021). There is therefore some debate as to the nature of spiral arms in protoplanetary discs (e.g. Dong et al., 2018b), and the conditions under which they can be observed (e.g. Speedie et al., 2022).

1.4.3 Warped discs

In addition to the substructure observed in protoplanetary discs, some systems show evidence of a warp. A disc is warped if the angular momentum vector, and thus the plane of its orbital motion, varies with radial distance from the central object. If the orbital plane of an internal companion is misaligned with respect to disc, the torques it exerts on the disc can drive a warp. We might also expect warped discs to be twisted, such that the angular momentum vector varies in azimuth as well as radius (e.g. Bardeen & Petterson, 1975). Papaloizou & Pringle (1983) identified two regimes of warp propagation, which depend on the thickness of the disc compared to the viscosity. If $H/r > \alpha$, viscosity is slow to damp perturbations and pressure forces dominate, causing the warp to propagate as a wave. If $H/r \lesssim \alpha$, the warp propagates diffusively.

Hughes et al. (2009) found evidence for a warp in the transition disc around GM Aur, identifying an 11° rotation between the orientation of the disc in sub-millimetre continuum emission and the CO line emission. Rosenfeld et al. (2012) proposed a warp to explain the kinematics of TW Hya, which had previously been suggested to contain a warp due to features observed in scattered light imaging (Roberge

1.5. The Formation of Planetary Cores

et al., 2005).

The transition discs around Herbig Ae/Be stars HD 142527 (Casassus et al., 2013) and HD 100453 (Wagner et al., 2015) show evidence for strong misalignments ($\sim 70^\circ$) between their inner and outer discs (e.g. Marino et al., 2015). Such extreme misalignments would require a companion on an orbit strongly misaligned with the disc in order to warp the inner disc to that same plane – this is in contention with the expectation that a companion should share the same orbital plane as the disc from which it formed.

Owen & Lai (2017) considered the dynamical process of “secular precession resonance”, which had previously been proposed as a means of generating misalignments between a protoplanetary disc and an external binary companion by Batygin & Adams (2013). In this case, however, the configuration of interest is a companion orbiting between the initially aligned inner and outer regions of a transition disc, separated by a cavity. Owen & Lai (2017) showed that an approximately flat configuration (a small mutual misalignment of $\approx 5^\circ$), can become excited into a strongly misaligned ($\gtrsim 60^\circ$) state due to a resonance between the nodal precession of the inner disc and the precession of the companion.

They considered the torque on the inner disc by the companion, the torque on the companion by the outer disc, and the torque on the inner disc by the outer disc. This gives rise to six individual precession frequencies (i.e. two between each pair of the three interacting components), the values of which depend on the masses and orbital separations of each component. As the system evolves, due either to accretion onto the companion or as the companion migrates, the precession frequencies of the inner disc and the companion cross, resulting in a resonant response and a strong misalignment. In this calculation, the inner and outer discs are treated as rigid discs, however, the torque from the companion would likely warp their edges, i.e. the outer edge of the inner disc, and the inner edge of the outer disc.

Most intriguing is the fact that these two systems, HD 142527 and HD 100453, both show spiral arms in their outer disc when observed at near-IR wavelengths, raising the question of whether these are related in some way to the warp and twist in the disc. The observational signatures of warps in protoplanetary discs will be explored further in chapter 5.

1.5 The Formation of Planetary Cores

Having described the environments in which planets are thought to form, it is timely to address the subject of planet formation itself. In short, there are two leading theories to explain the formation of planets within protoplanetary discs: via core accretion (Pollack et al., 1996) and through gravitational instability (GI, Boss, 1997). In the latter, over-dense regions of the protoplanetary disc undergo gravitational collapse to form a planetary core surrounded by a gaseous envelope. Toomre (1964) considered the stability of a uniformly rotating sheet, showing that the condition for gravitational collapse requires $Q < 1$, where

$$Q = \frac{c_s \kappa}{\pi G \Sigma}, \quad (1.60)$$

with κ being the epicyclic frequency. For Keplerian discs, $\kappa = \Omega_K$. Such a condition is only likely to be met in the cold, outer regions of particularly massive discs.

The theory of core accretion details the assembly of a planetary core through the collisional growth of dust grains into successively larger objects, which contracts a gaseous envelope once sufficiently massive (e.g. Safronov, 1972; Mizuno, 1980; Stevenson, 1982; Pollack et al., 1996). Once a gravitationally bound object – a planetesimal – is formed through the coagulation of dust grains, its accretion cross section becomes enhanced due to gravitational focussing, thus the gravitational attraction of smaller bodies begins to be significant. This allows a planetesimal to grow further by accreting dust grains as well as larger bodies, including other planetesimals, ultimately forming a planetary core – this is the process of planetesimal accretion. Runaway gas accretion begins once the envelope accreted by the core becomes

self-gravitating (e.g. Rafikov, 2006), and terminates once the planet opens a deep gap in disc (e.g. Bate et al., 2003). This requires the solid cores of giant planets to be formed on a time-scale $\lesssim 10^6$ yr, i.e. before disc dispersal (e.g. Strom et al., 1989; Haisch et al., 2001; Mamajek, 2009), so that a gaseous envelope can be accreted. This is difficult to achieve by planetesimal accretion based on predicted time-scales (e.g. Lissauer & Stevenson, 2007), especially for the outermost planets.

An additional complication is the formation of planetesimals themselves. The fragmentation barrier prevents growth beyond a few centimetres. There are further barriers to growth, such as the radial drift barrier: as shown above, once grains reach sizes corresponding to $St \sim 1$, they experience rapid radial drift towards the star. For a typical disc with $\Sigma_g \propto r^{-1}$, $T \propto r^{-1/2}$ and $h/r = 0.05$ around a $1M_\odot$ star, drift timescales can be as short as 1000 years at distances of a few AU. There is therefore a huge jump in the sizes of solids achieved through coagulation and the sizes required for embryo formation.

1.5.1 Embryo formation

A mechanism proposed by Goldreich & Ward (1973) to form planetesimals proposes that dust settling towards the disc midplane results in the formation of a thin, gravitationally unstable disc of solids. This would subsequently fragment and form planetesimals via gravitational collapse. This has been shown to fail since, once the local dust-to-gas mass ratio in the midplane exceeds unity, shear instabilities occur (Weidenschilling, 1980; Cuzzi et al., 1993; Weidenschilling, 1995), in particular the Kelvin-Helmholtz instability (Johansen et al., 2006; Barranco, 2009), preventing gravitational collapse.

Let us return briefly to the dynamics of dust grains to proceed. The treatment of the radial drift of dust in §1.4.1 did not account for the back-reaction on the gas: dust drifts radially due to the drag force from the gas, thereby exchanging angular momentum. There is necessarily an equal amount of angular momentum transferred to the gas, but since we typically assume a dust-to-gas mass ratio of order 1%, the effect on the gas is negligible and this back-reaction is typically ignored. However, if the dust-to-gas ratio is significantly enhanced, this effect must be accounted for.

The radial drift velocity of the gas, equation (1.38), can be generalised to account for the angular momentum gained by the gas – this is the Nakagawa-Sekiya-Hayashi (NSH) equilibrium (Nakagawa et al., 1986). Youdin & Goodman (2005) showed that such an aerodynamically coupled system of gas and solids, described by the NSH equilibrium, is linearly unstable to the growth of perturbations. This “streaming instability” can be understood as follows: if we consider a population of dust grains drifting inwards from the outer disc, if any initial clumping occurs (e.g. from random motions), the back-reaction on the gas becomes stronger, which in turn reduces the drift of dust, which leads to even more clumping as drifting dust catches up. Clumps whose density exceeds the Roche density, i.e. the limit at which self-gravity exceeds the tidal forces due to Keplerian shear, can undergo gravitational collapse into planetesimals. This typically forms objects following a power-law mass function, the upper limit of which corresponds to Ceres-sized objects (Johansen et al., 2015; Simon et al., 2016).

Pebble accretion

Perhaps the most promising route to form planetary cores is not through the collisional, oligarchic growth of planetesimals, but instead through the accretion of smaller solids onto a planetary embryo. It is well established from observations that protoplanetary discs contain large dust grains, and in substantial amounts. This motivated Johansen & Lacerda (2010) to consider the accretion of such “pebbles and rocks” (corresponding to Stokes numbers between 0.01 and 1 for their choice of disc properties) onto a ~ 100 km planetary seed. Through shearing sheet hydrodynamics simulations they found that the largest grains, with $St \sim 1$, rapidly spiral in towards the planet. Ormel & Klahr (2010) performed detailed calculations of the trajectories of such pebbles and derived their accretion rates onto the embryo, finding

1.5. The Formation of Planetary Cores

a maximum for $St \sim 1$. Due to their optimal coupling to the gas, they were found to accrete from distances comparable to the embryo’s Hill radius – significantly larger than the embryo’s geometric cross section – for embryos larger than $\sim 10^3$ km.

Lambrechts & Johansen (2012) shed further light on the pebble accretion process, showing how the combination of the embryo’s gravitational potential and gas drag enhance the capture of pebbles by the embryo. Assuming that the pebbles form a thin layer at the midplane due to vertical settling, pebble accretion can be considered 2D, where the accretion rate onto the embryo is

$$\dot{M}_{\text{peb}} = 2v_{\infty} b_{\text{PA}} \Sigma_{\text{peb}}, \quad (1.61)$$

where v_{∞} is the approach velocity of the pebble relative to the embryo, b_{PA} is the accretion cross section for pebble accretion (PA), and Σ_{peb} is the surface density of pebbles.

Two important quantities to establish are the planetary embryo’s Hill radius and Bondi radius. The Hill radius, defined in equation (1.42), quantifies the distance at which the planet’s gravity dominates over that from the host star. The Bondi radius is the distance from the planet at which its escape speed is the sound speed:

$$R_{\text{B}} = \frac{GM_{\text{P}}}{c_{\text{s}}^2}. \quad (1.62)$$

The approach velocity $v_{\infty} \sim v_{\phi, \text{d}}(a + b_{\text{PA}}) \approx (3/2)\Omega_{\text{K}}(a)b_{\text{PA}} + v_{\text{hw}}$, where v_{hw} is the headwind velocity felt by the pebble due to the deviation of the gas flow from Keplerian. They hence identified two regimes of pebble accretion, where either the headwind or shear velocities dominate the approach velocity. This can be expressed in terms of the relative sizes of the embryo’s Bondi and Hill radii, R_{B} and R_{H} respectively: for low embryo masses, where $R_{\text{B}} > R_{\text{H}}$, the approach velocity $v_{\infty} \approx v_{\text{hw}}$ (the “headwind” or “Bondi” regime), while for $R_{\text{H}} > R_{\text{B}}$, the approach velocity $v_{\infty} \approx (3/2)\Omega_{\text{K}}(a)b_{\text{PA}}$ (the “shear” or “Hill” regime).

They showed that in the latter case, the most optimally coupled pebbles, i.e. with Stokes numbers between 0.1 and 1, will all be accreted if they enter the Hill sphere. For those less optimally coupled, i.e. with $St \leq 0.1$, however, they argued for a Stokes number dependence:

$$b_{\text{PA}} = \begin{cases} St^{1/3} R_{\text{H}} & St \leq 0.1 \\ R_{\text{H}} & 0.1 < St \leq 1 \end{cases} \quad (1.63)$$

This therefore suggests a pebble accretion rate of $\dot{M}_{\text{peb}} \sim \Omega_{\text{K}}(a) R_{\text{H}}^2 \Sigma_{\text{peb}}$. In fact, multiple formulations of the ensuing pebble accretion rate exist (e.g. Lambrechts et al., 2014; Guillot et al., 2014; Morbidelli et al., 2015; Ida et al., 2016); Ormel (2017) proposes

$$\dot{M}_{\text{peb}} \approx 2\Omega_{\text{K}}(a) St^{2/3} R_{\text{H}}^2 \Sigma_{\text{peb}}(a), \quad (1.64)$$

with which such formulations are consistent to within a factor of unity. Most importantly, they all show that the pebble accretion rate is maximised for pebbles with $St \sim 0.1$ –1.

Lambrechts & Johansen (2012) showed that such accretion rates reduce planetary core growth time-scales by factors up to $\sim 10^2$ – 10^4 , depending on their orbital separation. Such has been the appeal of this mechanism that it is more-or-less simply invoked as a means of forming planetary cores in many planet formation models. A crucial step in its onset, however, is the requirement that the pebble layer can indeed be considered as 2D. The requirement is thus that the embryo’s Hill radius exceeds the thickness of the pebble layer (Lambrechts & Johansen, 2012; Lambrechts et al., 2014). This condition for $R_{\text{H}} > H_{\text{peb}}$ can be expressed in terms of a transition mass. By estimating the pebble scale height

from equation (1.40), the embryo mass must exceed

$$M_P \gtrsim 3M_* \left(\frac{H_g}{a} \right)^3 \left(\frac{\alpha}{\text{St}} \right)^{3/2}. \quad (1.65)$$

Adopting typical values of $M_* = M_\odot$, $H_g/a = 0.05$, $\alpha = 10^{-3}$ and $\text{St} = 1$, this requires an embryo mass $\sim 10^{-3} M_\oplus$. This is higher than the most massive objects produced through the streaming instability, thus further growth is required from planetesimals to reach this transition mass, beyond which rapid pebble accretion rates can be achieved.

Of course, pebble accretion can take place in the Bondi regime, just at a slower rate. Lorek & Johansen (2022) studied the growth of embryos by planetesimal accretion up to this transition mass. They considered a large planetesimal within a population of smaller planetesimals, formed in a filament through the streaming instability. They found that the embryo accretes all available material in the filament within a typical disc lifetime, but only in the inner disc (within 5–10 AU). The streaming instability forms multiple filaments, however, thus further growth could be achieved through the accretion of other planetesimals, as well as be supplemented by the accretion of pebbles, albeit at a reduced rate below the transition mass, which was not accounted for in their work.

Pebble accretion has been shown to become halted once the core is large enough to carve a gap in the disc, exterior to which drifting pebbles become trapped, unable to reach the planet (Morbidelli & Nesvorný, 2012; Lambrechts et al., 2014; Bitsch et al., 2018). More specifically, the drift of pebbles towards the planet is prevented once the planet forms a pressure maximum exterior to its orbit. The mass at which this occurs is the so-called “pebble isolation mass”. Lambrechts et al. (2014) performed hydrodynamics simulations to determine the mass at which the velocity exterior to its orbit becomes super-Keplerian, thus forming a local axisymmetric pressure maximum, finding

$$M_{\text{iso}} \approx 20 \left(\frac{H_g/r}{0.05} \right)^3 M_\oplus. \quad (1.66)$$

They also found a weak dependence on the disc viscosity, but this was not substantiated.

Ataiee et al. (2018) performed 2D dust and gas hydrodynamics simulations to further investigate the pebble isolation mass. They corroborated the $M_{\text{iso}} \propto h^3$ scaling, but found that diffusion of dust within the consequent pressure maximum broadens the width over which pebbles are trapped outside the planet’s orbit, which can increase the pebble isolation mass by up to an order of magnitude larger in high viscosity discs.

Bitsch et al. (2018) argued that equation (1.66), being empirically established, is an expression of the mass at which the wake driven by the planet shocks, leading to gap-opening, and that this is not necessarily the same process as that which reverses the pressure gradient exterior to the planet’s orbit, halting pebble drift. They therefore sought to build upon equation (1.66) by performing a parameter study with 3D hydrodynamics simulations. They also found a dependence on the cube of the aspect ratio, as well as a dependence on viscosity and the original, unperturbed pressure gradient, assuming a power-law disc:

$$M_{\text{iso}} = 25M_\oplus \left(\frac{h}{0.05} \right)^3 \left[0.34 \left(\frac{-3}{\log_{10} \alpha} \right)^4 + 0.66 \right] \left[1 - \frac{1}{6} \left(\frac{\partial \log P}{\partial \log r} + 2.5 \right) \right]. \quad (1.67)$$

Once pebble accretion has ceased, the planetary core can contract a gaseous envelope, beginning a phase of gas accretion. If the core is able to accrete enough gas that this envelope becomes self-gravitating, runaway gas accretion can proceed. Given the estimates of the pebble isolation mass in equations (1.66) and (1.67), pebble accretion may well be able to produce the cores of gas giants. Given the relatively

1.6. Thesis Outline

short accretion time-scales afforded by pebble accretion, it is plausible that planet formation is underway in the young systems observed. It is therefore important to consider the link between planet formation and the observed substructure in discs.

1.6 Thesis Outline

In this thesis I detail my investigations into non-axisymmetric disc substructure, with a focus on ringed discs. My investigations have been carried out primarily through numerical simulations, thus chapter 2 provides extensive details of the methods and codes used to this end. Chapters 3 and 4 present results of simulations of pebble accretion onto planetary embryos forming in axisymmetric dust traps, showing how this can lead to the formation of vortices, and thus observable non-axisymmetric features. In chapter 5 I consider static models of warped and twisted transition discs, and generate synthetic observations to determine the extent of shadowing from the disc structure. In chapter 6 I provide a summary of my results and discuss future directions for this work.

Chapter 2

Numerical Methods

Understanding the nature of protoplanetary discs and the physical processes that take place within them involves determining how they evolve in time, through the dynamics of both the gas and dust, interactions between large bodies (such as forming or formed planets) and the disc, and how they are affected by radiation from both the central star and nearby stars. Solving the equations of motion for the gas and dust on a global scale, i.e. over the entire disc, generally involves solving the Navier-Stokes equations in multiple dimensions, although certain symmetries (i.e. axisymmetry, or symmetry about the midplane) often allow the number of dimensions to be reduced. By including the potential of large bodies in the equations of hydrodynamics, their interactions with and impact on the disc structure and evolution can also be calculated. The outcome of such planet-disc interactions can therefore be achieved by performing hydrodynamics simulations. This is key to linking protoplanetary discs to the observed exoplanet population, since the material present at a given location in the disc determines the local potential for planet formation, and planet-disc interactions drive migration which can significantly change a planet's semimajor axis.

While hydrodynamics simulations calculate how material moves through the disc, internal energy can be transported independently of this flow. Heat can be generated within the disc through processes such as turbulent heating, shocks, collisions between large bodies and accretion onto planetary embryos. In addition, the radiative heating and cooling of dust grains – the dominant source of opacity in protoplanetary discs – determines the disc's equilibrium temperature and therefore the composition gradients and abundances of molecular species, thus the radiative transfer of energy through the disc is also a crucial element in understanding the nature of protoplanetary discs.

In order to test theory, simulations must ultimately be compared to observations. Since the only information about protoplanetary discs comes in the form of electromagnetic radiation, it is crucial to understand not only how observed signals are generated, but also how they propagate through the disc and any processes that affect them. Radiative transfer can therefore be used not only to determine how radiation affects the object, but also as a tool for aiding in the interpretation of observations. Radiative transfer simulations can be used to convert the outcome of hydrodynamics simulations (e.g. densities and velocities) into observable quantities such as intensities and wavelengths – hence they can be used to perform synthetic observations.

As their names suggest, these simulations are performed by solving the equations of hydrodynamics and radiative transfer respectively. However, for all but the simplest cases these equations cannot be solved analytically, thus numerical methods must be employed. There are a multitude of ways to perform both hydrodynamics and radiative transfer simulations, using techniques which may be better suited to certain situations, thus in this chapter the details of the numerical methods used to perform the simulations presented in this thesis are provided.

2.1 Numerical Hydrodynamics

In order to calculate the evolution of a protoplanetary disc over many dynamical timescales, numerical methods must be employed to solve the equations of motion of the gas and dust components throughout the disc. Calculating the motion of each and every gas molecule and dust grain is utterly infeasible:

2.1. Numerical Hydrodynamics

considering 1 g of molecular hydrogen gas alone, which constitutes 6×10^{23} molecules, is already far beyond the capabilities of even the most powerful and optimised hardware, let alone the $\gtrsim 10^{30}$ g expected in protoplanetary discs. However, provided that the molecular mean free path is much smaller than the characteristic length scales of the dominant physical processes in protoplanetary discs, both the dust and gas components can be treated as fluids, allowing the use of a statistical description involving conserved, macroscopic quantities. Within this treatment, a fluid is composed of fluid elements which must be sufficiently large to contain enough particles to define macroscopic properties, but sufficiently small compared to the length scales of interest.

A general fluid flow as a function of time, t , can be described by considering the mass per unit volume, i.e. the fluid density, ρ , at each point in space, $\mathbf{r}$. Conservation of mass dictates

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0, \quad (2.1)$$

where $\mathbf{v}$ is the velocity of the flow at each point in space. The fluid velocity can be determined by applying Newton's second law to each fluid element:

$$\rho \left(\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} \cdot \nabla \mathbf{v} \right) = -\nabla P + \nabla \cdot \mathbf{T} + \mathbf{F}_{\text{ext}} - [2\boldsymbol{\Omega} \times \mathbf{v} + \boldsymbol{\Omega} \times (\boldsymbol{\Omega} \times \mathbf{r}) + \boldsymbol{\Omega} \times \mathbf{r}] \rho, \quad (2.2)$$

where the forces on the fluid are the gradient of the thermal pressure, P ; viscosity, captured through the viscous stress tensor, $\mathbf{T}$; as well as any external forces, $\mathbf{F}_{\text{ext}}$. In the case of a protoplanetary disc, this includes the gravitational force from the host star, $-\rho \nabla \Phi$ (as protoplanetary discs are typically thought to have masses $M_{\text{disc}} \lesssim 0.1 M_*$, self-gravity can usually be ignored). As the disc is rotating around the star with angular velocity $\boldsymbol{\Omega}(\mathbf{r})$, it is also necessary to include the Coriolis and centrifugal forces. These two equations can be solved by relating the fluid pressure and density through the ideal gas law:

$$P = c_s^2 \rho, \quad (2.3)$$

where c_s is the fluid sound speed, given by

$$c_s = \sqrt{\frac{k_B T}{\mu m_H}}, \quad (2.4)$$

where k_B is Boltzmann's constant, T is the fluid temperature, μ is the mean molecular weight of the fluid and m_H is the mass of a hydrogen atom.

In order to solve equations (2.1) and (2.2) for the state of the protoplanetary disc over time, the Eulerian approach is to define a grid of fixed points spanning the disc and calculate the fluid properties (velocity, pressure and density) at each grid point, such that they are constant throughout each grid cell. By providing the initial disc conditions, i.e. values of each of these fluid properties at all locations in disc, these equations can be solved to calculate their values at these points at a future time. This is in contrast to the Lagrangian approach, wherein the trajectory of individual fluid parcels are computed.

The most popular Lagrangian approach is the method of smoothed-particle hydrodynamics (SPH), developed independently by Lucy (1977) and Gingold & Monaghan (1977), wherein the fluid is discretised into packets of mass – the particles – whose physical properties are spread over some smoothing length, and their trajectories are individually calculated. A stand-out feature of this method is that the resolution correlates with density, since higher density regions will necessarily contain more particles, while in grid-based methods the resolution is manually chosen when defining the grid; this can be either advantageous or disadvantageous, depending on the specifics of the problem. SPH codes also require an artificial viscosity to prevent interparticle penetration, which can make these codes somewhat diffusive. This

places lower limits on the viscosities and therefore the regimes of fluid flow that can be accurately simulated – this will be relevant to some of the discussion in chapter 5. The simulations used throughout this thesis are all performed using fixed-grid, Eulerian codes.

Finite differencing

By defining the fluid properties on a grid, the differentials in equations (2.1) and (2.2) can be approximated as finite differences of quantities between adjacent points on the grid, e.g. $\frac{\partial \rho}{\partial x} \rightarrow \frac{\Delta \rho}{\Delta x}$. The time derivative of a quantity such as the fluid density, ρ , in the i^{th} grid cell at the n^{th} timestep can therefore be expressed as

$$\left. \frac{\partial \rho}{\partial t} \right|_{i,n} = \frac{\rho_i^{n+1} - \rho_i^n}{\Delta t} + \mathcal{O}(\Delta t), \quad (2.5)$$

where $\mathcal{O}(\Delta t)$ indicates that terms of order $(\Delta t)^2$ and higher have been dropped; this approximation is therefore accurate to first order in Δt . This provides an explicit method for calculating the density at a later time $t + \Delta t$, using values at the current time t . Clearly, the accuracy of this finite difference increases for smaller timesteps, but short timesteps mean that long run-times are required. This same argument can be applied to spatial derivatives, leading to the conclusion that smaller grid cells yield greater accuracy, at the cost of run-time. A feature of spatial derivatives that can be exploited, however, is that second-order accuracy can be obtained by using centred differences:

$$\left. \frac{\partial \rho}{\partial x} \right|_{i,n} = \frac{\rho_{i+1}^n - \rho_{i-1}^n}{2\Delta x} + \mathcal{O}(\Delta x)^2. \quad (2.6)$$

Solving the finite difference equations at the grid boundaries requires knowledge of the fluid properties beyond the locations at which they have been defined. The computational domain is therefore surrounded by “ghost cells” within which the fluid properties are defined according to some boundary conditions, chosen for the specific problem. The equations are therefore not solved on these ghost cells, but merely facilitate the solution at the very edges of the computational domain.

While the accuracy of a finite difference scheme is important, a more essential requirement is its stability. The amount of time, Δt , by which one may advance the state of the disc in a single timestep must be limited to prevent spurious fluctuations from growing. Fluctuations arise from the fact that finite differencing schemes use information from neighbouring points to calculate quantities at a given point at a later time, but causality demands that information may only be obtained from those neighbouring points at a speed no greater than the speed of fluid flow. The Courant-Friedrichs-Lewy (CFL) condition (Courant et al., 1928) therefore states that the timestep cannot be greater than the time it takes for the fluid to traverse one grid cell:

$$\frac{v_x \Delta t}{\Delta x} \leq 1. \quad (2.7)$$

For the solution of linear PDEs this is a necessary and sufficient condition for stability, and must be satisfied in each spatial dimension. For non-linear PDEs such as equation (2.2), however, it is not sufficient, and a safety factor is typically used such that $v_x \Delta t / \Delta x \leq C$, where $C < 1$ depends on the specific problem. Developing a numerical scheme is therefore rarely as simple as replacing differentials with finite differences. In fact, naively adopting equations (2.5) and (2.6) yields the “forward-time, centred-space” scheme, which is unconditionally unstable (e.g. Press et al., 1992). Over the best part of the last century, numerous modifications have been developed to address various issues in finite differencing schemes for the solution of PDEs, and the most appropriate method is generally problem-dependent.

Finite difference methods are robust, efficient and straight-forward to implement and build upon.

2.1. Numerical Hydrodynamics

ZEUS-2D (Stone & Norman, 1992) was developed as a general purpose code for modelling astrophysical fluids. Being freely distributed, it has been highly influential and led to the development of many so-called “ZEUS-style” codes. These solve the equations of hydrodynamics using time-explicit finite difference schemes on an Eulerian grid, as well as a number of techniques such as operator splitting, the use of a staggered grid, and upwind methods to improve accuracy and stability in computing the advection terms, each of which will be substantiated hereafter. Grid staggering refers to the concept of storing scalar quantities at cell centres while vector quantities are stored at cell edges. Since many vectors are computed from the spatial derivatives of scalar quantities, centred differences of cell-centred quantities yield second-order accurate quantities at cell edges.

Operator splitting

The method of operator splitting breaks the solution of equations (2.1) and (2.2) into steps, motivated by the fact that the terms can be grouped generically into source terms (forces) and advection terms, each of which lend themselves to different solution methods. In general, a linear operator can be expressed as the sum of other linear operators; the operation is commutative, thus each operation can be performed in succession, with the result from previous steps being used in subsequent steps.

In their conservative forms, the equations of hydrodynamics can both be expressed in the form

$$\frac{\partial q}{\partial t} + \nabla \cdot (q\mathbf{v}) = \mathcal{S}(q, \mathbf{v}, t), \quad (2.8)$$

where q represents a fluid property such as density, and $\mathcal{S}$ is an operator representing the source terms – pressure, gravity etc. The method of operator splitting decomposes equation (2.8) into

$$\frac{\partial q}{\partial t} - \mathcal{S}(q, \mathbf{v}, t) = 0 \quad (2.9)$$

and

$$\frac{\partial q}{\partial t} + \nabla \cdot (q\mathbf{v}) = 0 \quad (2.10)$$

such that a complete time step is composed of a partial update from $q(t) = q^n \rightarrow q^{n+1/2}$ in the source step, followed by the partial update $q^{n+1/2} \rightarrow q^{n+1} = q(t + \Delta t)$ in the transport step.

The equations are also dimensionally split, i.e. they are solved component-wise. This therefore raises the issue of the order in which to perform these steps. One may envisage a scenario wherein calculating the same component first every time leads to a non-physical drift along one direction; numerical schemes have therefore been developed which alternate the order each timestep, or even perform the source step along one component followed by the full source and transport steps along the other, followed by the transport step for the first component. Regardless, for non-linear operators, operator splitting is only ever an approximation to the true solution. In developing their schemes for solving astrophysical hydrodynamics problems, Hawley et al. (1984) and Norman & Winkler (1986) showed that there is little difference between solutions generated via different orderings, and that this operator-split method still provides a more accurate solution than a single step based on old data.

Specialising to 2D polar coordinates (r, ϕ) to reflect the geometry of a protoplanetary disc, the equations solved in each step are as follows. The source step considers the terms describing forces on the fluid:

$$\frac{\partial v_r}{\partial t} = -\frac{1}{\rho} \frac{\partial P}{\partial r} - \frac{\partial \Phi}{\partial r} + \frac{(v_\phi + r\Omega_f)^2}{r}, \quad (2.11)$$

$$\frac{\partial v_\phi}{\partial t} = -\frac{1}{r\rho} \frac{\partial P}{\partial \phi} - \frac{1}{r} \frac{\partial \Phi}{\partial \phi}. \quad (2.12)$$

Ω_f is the angular velocity of the reference frame. Note that the continuity equation is not solved in this step as it contains no source terms.

The nonlinear term $v \frac{\partial v}{\partial x}$ in equation (2.2) can result in wave steepening and ultimately shock formation. This is commonly treated using the addition of an “artificial viscosity”, as developed by Von Neumann & Richtmyer (1950), which acts to smooth out shocks on spatial scales equivalent to a few cells. This artificial viscosity is also applied as a source term, $\nabla \cdot \mathbf{Q}$, where $\mathbf{Q}$ is a stress tensor, defined so as to only become significant if there is a sharp velocity gradient at a cell boundary. Specifically, it is identically zero unless the velocity gradient represents contraction, and is vanishingly small unless the velocity gradient is sufficiently large.

The velocities calculated in the source step are then used to advect the fluid density and momentum. Since this advection step involves the transport of fluid, it is important to ensure that mass and momentum are conserved. Thus, in this transport step, the integral form of equations (2.1) and (2.2) are used so that a conservative differencing scheme can be employed. The integral form of equation (2.10) can be obtained by applying the divergence theorem, which in 2D yields

$$\frac{\partial}{\partial t} \int_A q \, dA + \int_C q \mathbf{v} \cdot \hat{\mathbf{n}} \, dl = 0, \quad (2.13)$$

where l is the boundary of a fluid element with area A , and $\hat{\mathbf{n}}$ is a vector normal to the boundary. This therefore states that the rate of change of a quantity q in a given grid cell is equal to the flux of that quantity through the cell edges. The definition of velocities at cell edges is therefore naturally advantageous here, however, it is necessary to interpolate scalar quantities such as density to the cell edges in order to accurately determine their fluxes. A further complication is that, since the continuity equation contains no source terms, the density will still be defined at time t while the transport step takes place at time $t + \Delta t/2$; a method is therefore required to extrapolate the density forward in time. This can be achieved by using the velocity at the cell edge to integrate back in time over $\Delta t/2$ to yield the location at which to interpolate the quantity q .

The simplest method to approximate q at the interpolation location would be to use the value at the centre of the neighbouring cell, but ensuring that the cell “upwind” is chosen; since information flows in the direction of the fluid velocity, the cell in the direction opposite to the velocity must be chosen for stability (Courant et al., 1952). A second-order accurate method was proposed by van Leer (1977), which uses linear interpolation between the two cell centres either side of the edge of interest. This can give rise to spurious local extrema however, and potentially become unstable, thus van Leer (1977) included a “monotonisation” procedure which ensures that the density variation is monotonic: if the density variation across the cell is *not* monotonic, the method reduces to the first-order upwind scheme, thus sacrificing accuracy for stability. Another method, which is third-order accurate, is to use piecewise-parabolic advection (PPA, Colella & Woodward, 1984), in which the density distribution is instead approximated using parabolae. This method performs exceptionally well at maintaining sharp features, but comes at a high cost in computing time (Stone & Norman, 1992). Regardless of the interpolation scheme, two possible values to use at the cell edge arise from the interpolation either side of the cell edge; again this is resolved by using the upwind criterion.

2.1.1 FARGO3D

One such ZEUS-style code is FARGO3D (Benítez-Llambay & Masset, 2016), which has been developed with a specific focus on protoplanetary discs. It solves the equations of hydrodynamics using time-explicit, finite difference methods on a fixed Eulerian grid and employs operator splitting and upwind interpolation as described above. Its name derives from an algorithm it utilises: Fast Advection in Rotating Gaseous

2.1. Numerical Hydrodynamics

Objects (FARGO), which allows for large timesteps through a modification to the azimuthal transport of material (Masset, 2000). Van Leer’s method is used to evaluate fluxes during the transport step, except during the FARGO algorithm, when PPA is used.

In order to satisfy the CFL criterion, **FARGO3D** implements the following timestep constraints:

- Sound speed $\Delta t_1 = \Delta x_i / c_s$,
- Gas velocity $\Delta t_2 = \Delta x_i / |v_{x_i}|$,
- Artificial viscosity $\Delta t_3 = 4\sqrt{2}|\Delta x_i / \Delta v_{x_i}|$,
- Fluid viscosity $\Delta t_4 = \Delta v_{x_i}^2 / (4\nu)$,

where x_i refers generally to any spatial coordinate, i.e. r, ϕ, θ , thus Δx_i refers to any cell width in the grid, e.g. Δr , $r\Delta\phi$, and Δv_{x_i} refers to a difference in velocity along that direction. The maximum allowed timestep is then chosen according to

$$\Delta t = 0.44 \min \left[\left(\sum_n \Delta t_n^{-2} \right)^{-1/2} \right], \quad (2.14)$$

where Δt_n are the individual timestep constraints listed above, such that Δt will always be smaller than any individual Δt_n , ensuring stability. $\min$ indicates that the minimum of this value over the entire disc is chosen. 0.44 is chosen to be the safety factor (cf. equation (2.7)).

In order to reduce artefacts that may arise from the use of a finite domain, **FARGO3D** implements damping at the boundaries. Damping also serves to reduce the perturbations that arise from the fact that it is not possible to precisely satisfy hydrostatic equilibrium on a staggered grid. The damping takes place over damping zones at the inner and outer boundaries, which extend from r_{bound} to r_{damp} , which are calculated according to

$$r_{\text{damp}}^{\pm} = r_{\text{bound}}^{\pm} a_{\text{damp}}^{\mp 2/3}, \quad (2.15)$$

as implemented in Benítez-Llambay et al. (2016). The $+$ and $-$ label the outer and inner boundaries respectively. Perturbations are damped according to

$$\frac{dq}{dt} = -\frac{q - q_0}{\tau} R(r), \quad (2.16)$$

where q represents either the surface density or velocity, q_0 is its value at $t = 0$, τ is the damping time timescale and $R(r)$ is a parabolic function whose value is unity at the edge of the domain and zero at the interior edge of the damping zone (de Val-Borro et al., 2006).

Since the primary intended use of **FARGO3D** is the simulation of protoplanetary discs, the dynamics of any embedded planets must also be calculated. This is achieved using fifth order Runge-Kutta integration with Cash-Karp coefficients (Cash & Karp, 1990). The timestep for integration is that determined by the CFL condition, which the authors claim is suitable for planetary systems that do not undergo close encounters. This therefore makes **FARGO3D** less suitable for tightly-packed, multi-planet systems, where close encounters are likely.

Extensive testing of the code and algorithms are presented in Benítez-Llambay & Masset (2016), including standard tests such as the one-dimensional shock tube (Sod, 1978) and the viscous spreading of a gaseous ring (cf. Lynden-Bell & Pringle, 1974). **FARGO3D** has also been compared (albeit in its original form, **FARGO**) to other codes that use various alternative methods (both grid-based and SPH) to solve the equations of hydrodynamics in a protoplanetary disc containing an embedded planet (de Val-Borro et al., 2006). **FARGO** produced results which showed a good agreement with the other codes

as well as superior performance over SPH codes, which found weaker spiral waves driven by planets (cf. §1.4.2) due to their inherent diffusivity, and low resolution in planet-opened gaps.

The FARGO Algorithm

In a Keplerian disc, where the orbital velocity decreases with distance from the central object, the high orbital velocity of material at the inner boundary of a simulation domain imposes a significant limit on the timestep that can be achieved due to the CFL condition. Similarly, the geometry of protoplanetary discs makes the use of polar coordinates a natural choice; cells at the inner boundary will therefore be narrow, further limiting the timestep – equation (2.14) shows that smallest timestep over entire disc is taken. A sensibly chosen simulation domain will keep the region of interest far from the boundaries so as to avoid artefacts that may arise due to the finite domain, thus the necessity of using a shorter timestep than is required in the region of interest significantly increases computational cost.

The FARGO algorithm (Masset, 2000) is a modification to the standard transport algorithm used in explicit schemes used to solve the equations of hydrodynamics on a polar grid, capable of achieving an order-of-magnitude increase in timestep. The continuity equation takes the form

$$\frac{\partial \rho}{\partial t} + \frac{1}{r} \frac{\partial (r \rho v_r)}{\partial r} + \frac{1}{r} \frac{\partial (\rho v_\phi)}{\partial \phi} = 0, \quad (2.17)$$

where the second and third terms correspond to radial and azimuthal transport respectively. Without loss of generality, this can be re-written as

$$\frac{\partial \rho}{\partial t} + \frac{1}{r} \frac{\partial [r \rho (v_\phi - \bar{v}_\phi)]}{\partial r} + \frac{\bar{v}_\phi}{r} \frac{\partial \rho}{\partial \phi} + \frac{1}{r} \frac{\partial (\rho v_\phi)}{\partial \phi} = 0, \quad (2.18)$$

where $\bar{v}_\phi$ is the azimuthally averaged velocity. Thus, by using the technique of operator splitting, the azimuthal transport step can be decomposed into transport at velocity $v_\phi - \bar{v}_\phi$, followed by an additional step to solve

$$\frac{\partial \rho}{\partial t} + \frac{\bar{v}_\phi}{r} \frac{\partial \rho}{\partial \phi} = 0. \quad (2.19)$$

By subtracting the average azimuthal velocity at each annulus in the grid before the transport step, the timestep constraint on the azimuthal velocity becomes a constraint on the residual velocity, $v_\phi - \bar{v}_\phi$. Since this is not guaranteed to be maximised at the inner boundary, as is the case for v_ϕ , the timestep is not limited by the inner boundary where, by design, nothing of interest takes place.

This algorithm also reduces the truncation error in the transport step. If, for example, the fluid is to be transported a distance equivalent to 4.7 cells, the FARGO algorithm decomposes 4.7 into $-0.3 + 5$, the material is transported -0.3 cells and then the results are simply shifted by 5 cells, which introduces no numerical error, and the transport by -0.3 cells introduces a lower numerical error than 4.7 otherwise would have.

During this modified azimuthal transport step, PPA is used rather than van Leer, as the former performs better at capturing shocks, which is relevant for the highly supersonic azimuthal velocities in protoplanetary discs. **FARGO3D** includes the option to use the standard advection algorithm, enabling a comparison and validation of the FARGO algorithm.

2.1.2 Dust dynamics

Treatment of dust in protoplanetary discs requires special consideration as the dust is directly affected by the gas, as described in §1.4.1, and this interaction must be accounted for. In protoplanetary discs, dust may be treated as a pressure-less fluid, as the acceleration of an individual dust grain from the gas

2.1. Numerical Hydrodynamics

pressure gradient is negligible. Dust grains therefore do not have their orbital speeds directly affected by the pressure gradient as is the case for gas, and this difference in orbital speeds gives rise to a drag force. In the limit that $\Sigma_g \gg \Sigma_d$, the gas is unaffected by this drag force and its acceleration is therefore constant during the interaction.

The equation of motion for a dust grain coupled to the gas via drag forces in the Epstein regime, such that the drag force is a linear function of velocity (equation (1.34)), is

$$\frac{\partial \mathbf{v}_d}{\partial t} + \mathbf{v}_d \cdot \nabla \mathbf{v}_d = -\frac{1}{t_s}(\mathbf{v}_d - \mathbf{v}_g) + \mathbf{a}_d, \quad (2.20)$$

where $\mathbf{v}_g$ is the gas velocity and $\mathbf{a}_d$ is the acceleration of the dust grain due to forces other than drag, such as stellar gravity. t_s is the dust grain's stopping time, defined in equation (1.36). As was the case for the gas dynamics, the solution of this equation can be split into source and transport steps. In the source step, since $\mathbf{v}_d$, $\mathbf{v}_g$ and $\mathbf{a}_d$ at time t are known, an explicit calculation of $\mathbf{v}_d(t + \Delta t)$ is possible. However, in order to make use of the timestep calculated for the gas according to the CFL condition, a semi-implicit method can be used, wherein the gas velocity at time $t + \Delta t$ is used. Considering just the source terms, and re-writing equation (2.20) with $k = t_s^{-1}$ to improve clarity,

$$\frac{d\mathbf{v}_d}{dt} + k\mathbf{v}_d = k\mathbf{v}_g + \mathbf{a}_d, \quad (2.21)$$

the equation of motion is now in a form that can be solved using the integrating factor $\exp(kt)$:

$$\frac{d}{dt}[\exp(kt)\mathbf{v}_d] = k\exp(kt)\mathbf{v}_g + \exp(kt)\mathbf{a}_d. \quad (2.22)$$

Integrating over the timestep from $t' = t$ to $t' = t + \Delta t$

$$\int_t^{t+\Delta t} \frac{d}{dt'}[\exp(kt')\mathbf{v}_d] dt' = \int_t^{t+\Delta t} [k\exp(kt')\mathbf{v}_g + \exp(kt')\mathbf{a}_d] dt', \quad (2.23)$$

the LHS can be evaluated using the fundamental theorem of calculus:

$$\int_t^{t+\Delta t} \frac{d}{dt'}[\exp(kt')\mathbf{v}_d] dt' = \exp[k(t + \Delta t)]\mathbf{v}_d(t + \Delta t) - \exp(kt)\mathbf{v}_d(t). \quad (2.24)$$

Integrating the RHS by parts,

$$\int_t^{t+\Delta t} k\exp(kt')\mathbf{v}_g dt' = \left[\exp(kt')\mathbf{v}_g \right]_t^{t+\Delta t} - \int_t^{t+\Delta t} \exp(kt') \frac{d\mathbf{v}_g}{dt'} dt' + \int_t^{t+\Delta t} \exp(kt')\mathbf{a}_d dt', \quad (2.25)$$

we now make the assumption that both the gas and dust accelerations are constant from $t' = t$ to $t' = t + \Delta t$ and can therefore be taken out of the respective integrands:

$$\int_t^{t+\Delta t} k\exp(kt')\mathbf{v}_g dt' = \left[\exp(kt')\mathbf{v}_g \right]_t^{t+\Delta t} - \left(\frac{d\mathbf{v}_g}{dt'} - \mathbf{a}_d \right) \int_t^{t+\Delta t} \exp(kt') dt'. \quad (2.26)$$

This integral can now be evaluated:

$$\begin{aligned} \int_t^{t+\Delta t} k\exp(kt')\mathbf{v}_g dt' &= \left[\exp(kt')\mathbf{v}_g \right]_t^{t+\Delta t} - \frac{\mathbf{a}_g - \mathbf{a}_d}{k} \left[\exp(kt') \right]_t^{t+\Delta t} \\ &= \exp[k(t + \Delta t)]\mathbf{v}_g(t + \Delta t) - \mathbf{v}_g(t)\exp(kt) \\ &\quad - \frac{\mathbf{a}_g - \mathbf{a}_d}{k} \exp[k(t + \Delta t)] + \frac{\mathbf{a}_g - \mathbf{a}_d}{k} \exp(kt). \end{aligned} \quad (2.27)$$

Note that this expression contains the gas velocity $\mathbf{v}_g(t + \Delta t)$, which is known provided the gas source step is performed first. Equating equations (2.24) and (2.27) and dividing through by $\exp[k(t + \Delta t)]$ yields

$$\mathbf{v}_d(t + \Delta t) = \mathbf{v}_d(t) \exp(-k\Delta t) + \mathbf{v}_g(t + \Delta t) - \mathbf{v}_g(t) \exp(-k\Delta t) + \frac{\mathbf{a}_d - \mathbf{a}_g}{k} [1 - \exp(-k\Delta t)]. \quad (2.28)$$

Approximating the gas acceleration as the change in gas velocity over the source step, we can use $\mathbf{v}_g(t + \Delta t) = \mathbf{v}_g(t) + \mathbf{a}_g \Delta t$ to obtain the expression for the updated dust velocity:

$$\mathbf{v}_d(t + \Delta t) = \mathbf{v}_d(t) \exp(-k\Delta t) + \mathbf{a}_g \Delta t + \left[\mathbf{v}_g(t) + \frac{\mathbf{a}_d - \mathbf{a}_g}{k} \right] [1 - \exp(-k\Delta t)]. \quad (2.29)$$

This expression was first presented in Rosotti et al. (2016), where it was used to facilitate the addition of dust to FARGO3D. While it is presented here in terms of the stopping time $t_s = k^{-1}$, it can be expressed in terms of the dust grain size through equation (1.36), so that dust species can either represent a population of grains of either a fixed stopping time, Stokes number or grain size. To account for dust diffusion, an effective diffusive velocity, as defined in equation (1.41), is added to the velocity at which the dust is advected during the transport step. An additional timestep constraint is implemented to ensure stability of the dust, identical in form to that of the gas:

$$\Delta t = \Delta x_i / |v_{x_i,d}|. \quad (2.30)$$

In order to verify that this produces the correct dust velocity, I performed a simulation in which the dust was left to evolve from an axisymmetric initial condition with $\Sigma_d / \Sigma_g = 0.01$ everywhere in the disc to a steady state, and compared the resulting dust surface density and radial velocity to the analytic solution of the steady-state advection-diffusion equation. The continuity equation for the dust surface density in 1D (i.e. for an axisymmetric disc) is

$$\frac{\partial \Sigma_d}{\partial t} + \frac{1}{r} \frac{\partial}{\partial r} (r F_d) = 0, \quad (2.31)$$

where F_d is the radial dust flux. For an axisymmetric disc this is given by

$$F_d = \Sigma_d v_{r,d} - \nu \Sigma_g \frac{\partial}{\partial r} \left(\frac{\Sigma_d}{\Sigma_g} \right), \quad (2.32)$$

where the first term on the right-hand side is the advective term and the second captures dust diffusion. In steady state, where $F_d = 0$, this can be expressed as an ordinary differential equation for the dust-to-gas mass ratio:

$$\frac{d}{dr} \left(\frac{\Sigma_d}{\Sigma_g} \right) = \frac{v_{r,d}}{\nu} \frac{\Sigma_d}{\Sigma_g}. \quad (2.33)$$

In order to obtain a steady-state solution, a gas surface density profile containing a local pressure maximum capable of trapping dust was used to prevent all the dust from simply drifting to the inner edge. The gas surface density used was

$$\Sigma_g = \frac{\Sigma_0}{2} \left(\frac{r}{R_0} \right)^{-1} \left[1 + \operatorname{erf} \left(\frac{r - R_1}{\sqrt{2}\sigma} \right) + \epsilon \right], \quad (2.34)$$

where $\Sigma_0 = 10^{-3} M_* R_0^{-2}$, $R_1 = 1.2 R_0$. $\sigma = 0.3 R_0$ controls the width of the pressure maximum, and $\epsilon = 10^{-3}$ prevents the surface density from vanishing at small radii in the initial condition. This imprints an inner cavity on a standard power-law surface density profile with index $\gamma = 1$ (cf. §1.3.2), creating a

2.1. Numerical Hydrodynamics

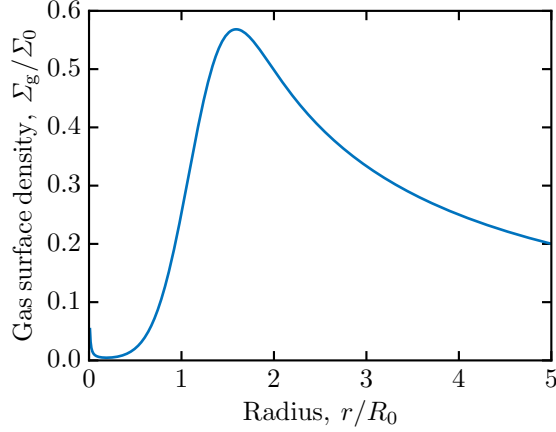


Figure 2.1: Gas surface density profile used for testing the dust dynamics numerical scheme. The inner disc is significantly depleted of gas, which creates a pressure maximum capable of trapping dust.

pressure maximum and therefore a dust trap at $r \approx 1.5R_0$. This functional form is depicted in figure 2.1.

The disc had a flaring index of $k = 5/4$ and an aspect ratio $h_0 = 0.05$. The disc viscosity was prescribed via the alpha viscosity prescription, with $\alpha = 10^{-4}$, and the central star’s mass was $M_* = 2M_\odot$. The grid contained 768 cells in azimuth, uniformly spaced between 0 and 2π , while 512 logarithmically spaced cells were used to span the disc from $0.1R_0$ to $5R_0$ in radius.

I obtained the analytic solution by numerically integrating equation (2.33) using a fourth-order Runge-Kutta method, with $v_{r,d}$ given by equation (1.38) and ν by equation (1.13). As the FARGO3D simulation requires time for the dust to reach steady-state, and during that time the gas surface density will evolve due to viscous accretion, I used the gas surface density and velocity at the time at which the dust had reached a steady state as the gas surface density and velocity in calculating the analytical solution. However, I increased their resolution from 512 cells in radius to 192000 via cubic spline interpolation, so that stable integration could be achieved.

I performed this test for dust with a fixed Stokes number, $St = 0.01, 0.1$ and 1 , and also for fixed grain size dust, with $s = 1$ mm. In each case a dust density of $\rho_\bullet = 3 \text{ g cm}^{-3}$ was used. A comparison of the resulting dust surface densities and radial velocities is shown in figure 2.2 for the case of fixed $St = 0.1$. This dust velocity implementation in FARGO3D is in excellent agreement with the analytical solution, and thus produces the correct location of the dust trap and the mass of dust within. While the dust radial velocity is zero at $r = 1.5$, as expected for a dust trap, it also reaches zero at $r \approx 0.2$. The latter is an unstable equilibrium, where η is maximised, thus diffusion prevents dust from becoming trapped there.

It is also worth noting a subtlety in this comparison: dust drift velocity depends on drag, which depends on gas orbital velocity. The expression for the dust radial velocity contains η , the derivation (in §1.3.2) of which was in terms of the gas pressure P and density ρ_g . An equivalent derivation using vertically integrated pressure $\mathcal{P}$ and surface density Σ_g can be performed, which yields the expression

$$\eta_{v1} = -\left(\frac{H}{r}\right)^2 \left(\frac{d \log \Sigma_g}{d \log r} + 2k - 3\right). \quad (2.35)$$

However, substituting $\Sigma_g = \sqrt{2\pi}H\rho_{g,\text{mid}}$ into equation (1.26) (cf. Alexander & Armitage, 2007) gives

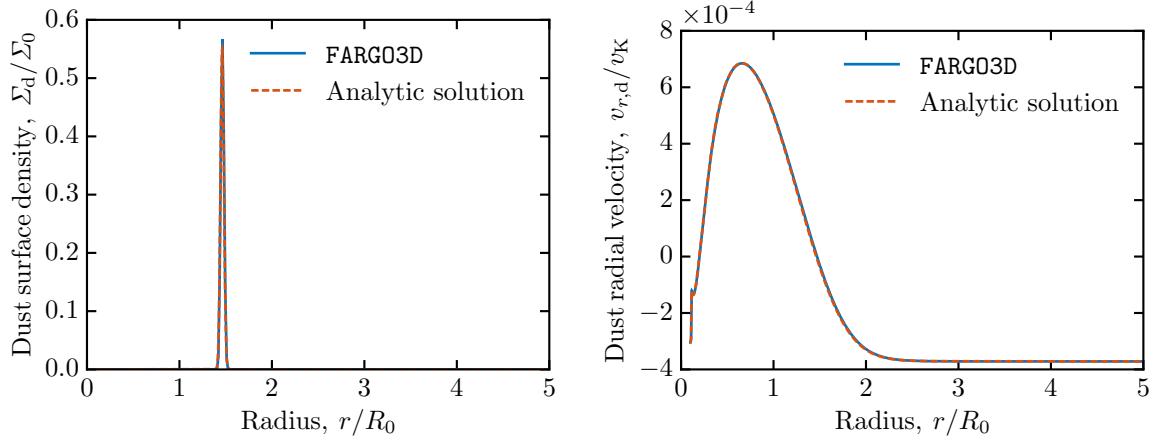


Figure 2.2: Comparison of the surface density (left) and radial velocity (right) of $St = 0.1$ dust resulting from the semi-implicit scheme implemented in FARGO3D with the analytic solutions, showing excellent agreement. The dust drifts to the pressure maximum at $r \approx 1.5R_0$ (where the radial velocity is zero) and forms a peak, the width of which is set by viscosity.

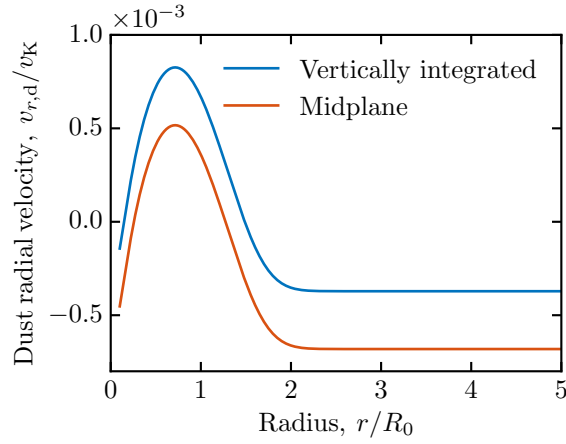


Figure 2.3: Comparison of the dust radial velocity obtained using vertically integrated and midplane quantities η . In addition to producing a different magnitude of velocity, the location of the dust trap is shifted slightly in radius.

the deviation from Keplerian rotation at the disc midplane as

$$\eta_{\text{mid}} = -\left(\frac{H}{r}\right)^2 \left(\frac{d \log \Sigma_g}{d \log r} + k - 3\right). \quad (2.36)$$

The difference that this makes to the dust radial velocity is shown in figure 2.3. The dust velocity is zero at slightly different radial positions – specifically at $0.15R_0$ and $1.27R_0$, compared with $0.26R_0$ and $1.49R_0$ in the vertically integrated case, thus the location of the dust trap varies by $\sim 10\%$. However, the disc properties do not change significantly over such distances, thus the properties of the dust trap itself would remain largely unchanged. In order to make a direct comparison to FARGO3D, which uses vertically integrated pressure and density in 2D simulations, this test made use of equation (2.35).

2.1.3 Pebble accretion

One further modification that I made to FARGO3D is the addition of a pebble accretion routine. Given that hydrodynamics simulations deal with fluid elements which comprise large numbers of particles, the accretion of individual dust grains onto the surface of a planet is not possible – especially with a

2.1. Numerical Hydrodynamics

point-mass representation of the planet. The aim of this pebble accretion routine is therefore to have the *physics* of pebble accretion resolved. Specifically, the planet’s Hill sphere should be supplied with dust at the appropriate rate set by the physical properties of both the planet and the disc. However, since the accretion clearly cannot be resolved directly, dust must be removed from the planet’s vicinity to prevent unphysically large dust surface densities inside the Hill sphere. Thus, by removing dust from the planet’s Hill sphere at a rate appropriate for the physics of pebble accretion, the surface density in the Hill sphere will reach a steady balance between the physically derived supply rate from the simulation and the sensibly chosen removal rate. Provided that the removal rate is sensibly chosen (and dependent on the dust surface density in the Hill sphere), if the supply rate is resolved then the removal rate, and hence accretion rate onto the planet, will then be independent of the actual form of the removal rate; all that will happen is the surface density in the Hill sphere will moderate itself to match the supply rate provided by the physics of the simulation.

The pebble accretion rate onto a planet at semimajor axis a is calculated numerically by summing the masses in each cell within the planet’s Hill sphere; by calculating $\Sigma_d(a)$ as an average over the Hill sphere, i.e. $\Sigma_d(a) = M_H/(\pi R_H^2)$ where M_H is the dust mass in the Hill sphere, the pebble accretion rate can be expressed as

$$\dot{M}_{\text{peb}} = \frac{2}{\pi} \Omega_K(a) \text{St}^{2/3} \sum_{d \leq R_H} \Sigma_{d,ij} A_{ij}, \quad (2.37)$$

where d is the distance from the planet and A_{ij} is the area of the cell with indices i, j . It is important to account for the fact that the Hill radius will enclose fractions of cells, as can be seen in figure 2.4, in order to provide a smooth change in accretion rate as the Hill sphere increases in radius and thus encloses more cells. To estimate the fraction of a cell that lies partially within the Hill radius, the grid is locally approximated as Cartesian so that each cell can be treated as a rectangle. As shown in figure 2.4, at sufficiently high spatial resolution this is an excellent approximation. The two intercepts that the Hill sphere makes with the boundaries of a given cell are calculated and used to construct a polygon, whose remaining vertices are the cell corners within the Hill radius. The area of this polygon is ultimately evaluated as the sum of areas of a triangle and at most two rectangles, depending on where the Hill sphere crosses the cell. This simple calculation serves as a good approximation for the area of the cell contained within the Hill sphere, yielding a smooth change in accretion rate as the planet’s mass increases and new cells enter its Hill sphere. Figure 2.4 shows an example of the polygon constructed using this method for a cell that lies partially within the planet’s Hill sphere. The value of St used in equation (2.37) is calculated as an azimuthal average at the semimajor axis of the planet, as equation (1.64) implicitly assumes an axisymmetric disc, unperturbed by the presence of the planet.

The planet’s mass is updated according to $\Delta M_P = \dot{M}_{\text{peb}} \Delta t$. The amount of mass ΔM_P is removed from the planet’s Hill sphere, with the mass in each cell reduced in proportion to the mass there at that instant. An additional timestep constraint from the pebble accretion rate is included such that no more than 10% of the mass in any one cell is removed in a single timestep. Note that a planetary embryo for which this is relevant will not be massive enough to accrete a substantial amount of gas, thus the increase in mass is assumed to be due to pebble accretion only.

As a means of verifying that the pebble accretion rate can actually be resolved in a hydrodynamics simulation, I performed a test wherein the pebble accretion rate onto a planet was measured while the planet was held at fixed mass as dust was removed from its Hill sphere, and the pebble accretion rate as expressed in equation (1.64) was varied by factors of 0.5 and 2. The planet was held at a fixed mass because any initial differences in the pebble accretion rate would lead to differences in planet mass and therefore the subsequent accretion rate, making it difficult to distinguish between effects from modifying the accretion rate and effects from differences in planet mass.

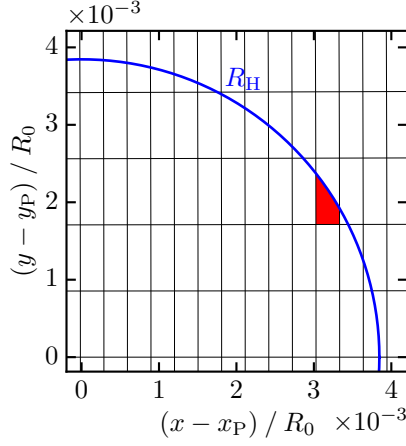


Figure 2.4: A zoom-in on one quadrant of the planet’s Hill sphere, with the coordinate system centred on the location of the planet (x_P, y_P) . The black lines mark cell boundaries and the blue line marks the edge of the planet’s Hill sphere at a given mass. The fraction of a cell only partially contained within the Hill sphere is approximated by constructing a polygon within the cell, an example of which is shown in red.

I performed this test for planets of mass $0.1M_\oplus$, $1.0M_\oplus$ and $10M_\oplus$, the results of which are shown in figure 2.5. If the rate of removal of dust from the Hill sphere was controlling the accretion rate, one would expect these modifications to linearly change the accretion rates. However, if the simulations correctly resolve the physics of pebble accretion, these modifications should instead have no impact on the actual accretion rate as described above. What these tests show is that the effect of the order-unity modification is simply an initial time offset in the accretion rate, due to the dust in the Hill sphere being unperturbed by the embryo at the start of the simulation. Once the surface density in the vicinity of the planet has reached a pseudo steady-state between supply and removal from the Hill sphere, which takes ≈ 4 orbits, the accretion rate is insensitive to the exact choice of the mass removal scheme. These tests demonstrate that the accretion rate naturally adjusts itself to the appropriate value as the planet regulates the mass in its Hill sphere, and hence pebble accretion is resolved as desired.

Given that material is being removed from the disc, it is important to ensure that the surface density cannot reach zero. This is especially important given that the pebble accretion algorithm removes dust from the vicinity of the planet, and planets are known to carve a gap in the disc, thus surface densities can become vanishingly small. To enforce this, I implemented a dust floor such that if the surface density in any cell falls below a specified value, it is reset to that value. If this occurs, the dust velocity components are set to their respective gas velocity components. Figure 2.6 shows the order of each of the main steps in **FARGO3D**, including the modifications to include dust evolution and pebble accretion. Note that the dust floor is applied at the end of each component’s transport step. As the dust velocity is used to determine the upwind direction, during momentum transport the velocity is calculated as the momentum divided by density. Thus, if the mass of dust in any cell falls below zero, an incorrect upwind direction would be obtained for that cell, leading to an incorrect (and unphysical) advection calculation. But by ensuring the density is a positive (albeit small) number, and the dust velocity is equal to that of the gas, no dust drift can occur and the advection is guaranteed to take place in the correct direction.

A planet undergoing rapid pebble accretion will necessarily have a significant accretion luminosity, which may heat the disc locally. However, in using a locally isothermal equation of state, the internal energy of the gas is not evolved. Thus, in order to include the luminosity of an accreting body, the thermal feedback on the surrounding gas is accounted for by re-calculating the sound speed during the pressure calculation step.

The results of the hydrodynamics simulations presented in the following chapters were performed us-

2.1. Numerical Hydrodynamics

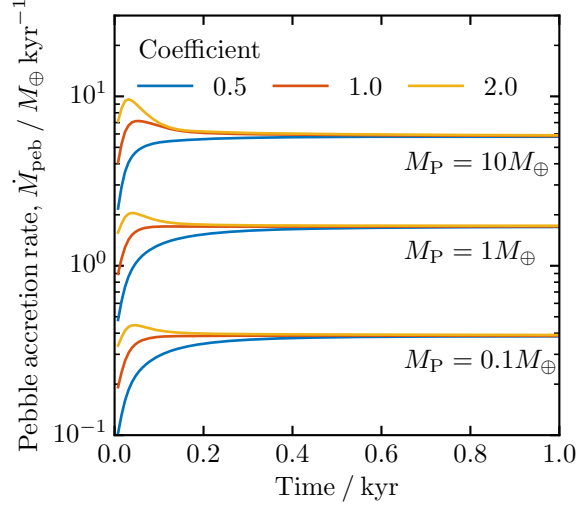


Figure 2.5: A comparison of the pebble accretion rates for the same initial conditions, but with the accretion rate modified by a constant coefficient. The planet mass is held fixed while dust is removed from its Hill sphere according to the pebble accretion rate, and the test is performed for planets of mass $0.1M_{\oplus}$, $1.0M_{\oplus}$ and $10M_{\oplus}$. Despite the modifying factor, the accretion rates quickly converge to a single value for a given planet mass, as the planet naturally sets the accretion rate by regulating the mass in its Hill sphere.

ing **FARGO3D** version 1.3, modified to include the dust dynamics and pebble accretion processes described above, using MPI for parallelisation.

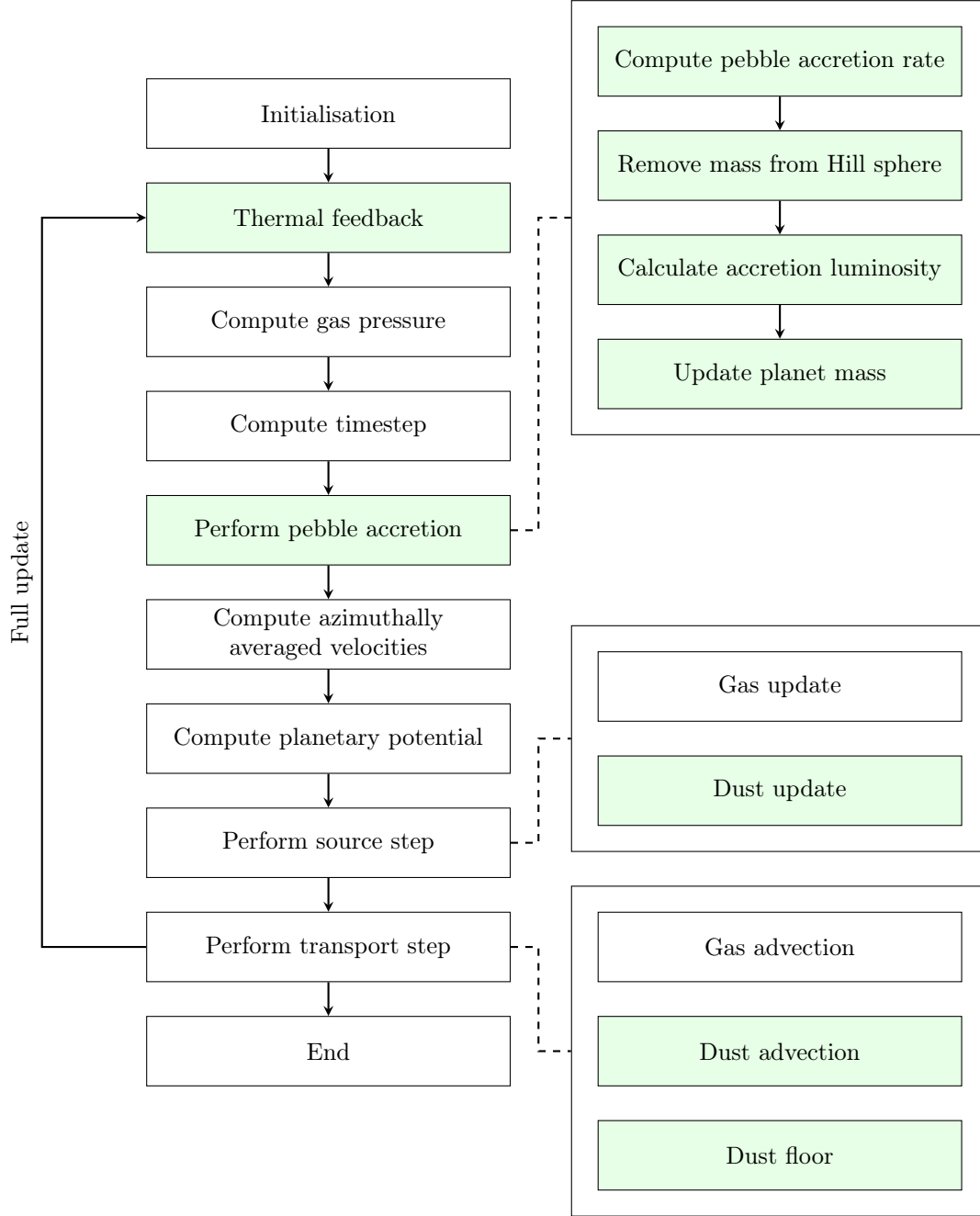


Figure 2.6: Flowchart for the modified **FARGO3D** code, adapted from Benítez-Llambay & Masset (2016). Modifications to the original code are shown in green. At the start of each timestep, the planet’s accretion luminosity is calculated and the gas temperature is updated accordingly. Once the timestep has been determined from the CFL criteria, the amount of dust to remove can be calculated from the pebble accretion rate. The planet mass is then updated before performing the source and transport steps.

2.2 Numerical Radiative Transfer

Hydrodynamics simulations provide a means of calculating the transport of material through a protoplanetary disc, but only by calculating the transport of radiation through the disc is it possible to convert the results of such simulations into quantities that can be observed. Since the only information received from protoplanetary discs is delivered in the form of electromagnetic waves, we are specifically concerned with how the circumstellar gas and dust absorb, re-emit and scatter radiation from the central star.

The amount of energy, E , of a given frequency, ν , delivered per unit area, A , over a solid angle, Ω , per unit time, t , is the specific intensity:

$$I_\nu = \frac{dE}{dt d\nu dA d\Omega}. \quad (2.38)$$

The radiative transfer equation states how this quantity changes when radiation is transported through a medium when accounting for processes that remove energy from a given ray, i.e. extinction, and processes that add energy to the ray, i.e. emission. Extinction can be quantified through the cross section per unit mass of extinguishing material, or “opacity”, κ_ν , such that the change in specific intensity along a distance s due to extinction is

$$dI_\nu = -\kappa_\nu \rho I_\nu ds, \quad (2.39)$$

where ρ is the density of extinguishing material. The total extinction opacity is generally the sum of two components, the scattering opacity $\kappa_{\nu, \text{sca}}$ and the absorption opacity $\kappa_{\nu, \text{abs}}$, both of which act to remove intensity from a given ray.

The quantity $(\kappa_\nu \rho)^{-1}$ is the mean free path of a photon in the medium, from which it can be shown that the probability of a photon undergoing an interaction is $p(s_0, s_1) = \exp(-\tau_\nu(s_0, s_1))$, where $\tau_\nu(s_0, s_1)$ is the optical depth along the ray between points s_0 and s_1 , defined as

$$\tau_\nu(s_0, s_1) = \int_{s_0}^{s_1} \kappa_\nu \rho ds. \quad (2.40)$$

The solution to equation (2.39) can therefore be expressed as

$$I_\nu(s_1) = I_\nu(s_0) \exp(-\tau_\nu(s_0, s_1)). \quad (2.41)$$

While protoplanetary discs are 99% gas by mass, the dominant source of opacity is the dust component. While molecular species in the gas phase absorb and emit only at specific frequencies, the macroscopic structure of dust grains means that they have continuum opacities. The interaction of light with a dust grain is fundamentally a problem of electromagnetism and therefore requires Maxwell’s equations. Mie (1908) considered this problem for spherical, homogeneous particles, the result of which permits the calculation of the extinction and scattering efficiencies as a function of the complex refractive index, $m(\lambda) = n(\lambda) - ik(\lambda)$, of the dust grain. The absorption opacity is the difference between the extinction and scattering opacities. The calculation involves an infinite sum, thus in practice it is calculated numerically. One publicly available code for computing the scattering and absorption efficiencies is Draine’s implementation¹ of the BHMIE code (Bohren & Huffman, 1998). Refractive index data for dust grains are typically obtained from laboratory measurements of spectra of grains thought to be similar in composition to astrophysical dust.

Figure 2.7 shows the absorption and scattering opacities of 1 μm and 1 mm sized dust grains as a function of wavelength. These dust opacities were calculated using the model presented in Birnstiel et al. (2018) to analyse the DSHARP observations. The dust grains were taken to be a mixture of

¹<https://www.astro.princeton.edu/~draine/scattering.html>

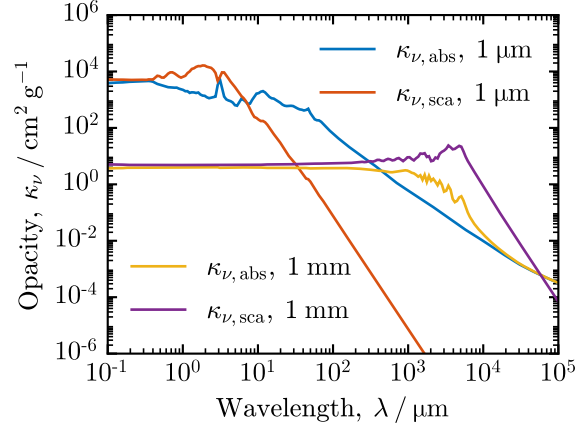


Figure 2.7: Comparison of absorption and scattering opacities for 1 μm and 1 mm dust grains. The scattering falls off as λ^{-4} and the absorption opacity falls off as λ^{-1} once the wavelength exceeds the size of the dust grain. Small dust grains dominate the opacity at short wavelengths, while large grains dominate at long wavelengths. Absorption dominates over scattering at longer wavelengths.

previously established dust species whose optical properties reproduce observations. Specifically, these are “astronomical silicates” (Draine, 2003), troilite and “organics” (Henning & Stognienko, 1996) and water ice (Warren & Brandt, 2008). The water ice mass fraction used was 0.2, and the grains were taken to have zero porosity. The opacity falls off rapidly for wavelengths larger than the grain size. The opacity at short wavelengths is dominated by smaller dust grains. This justifies the heuristic approximations made in §1.2, i.e. that at NIR wavelengths the dominant source of opacity is small dust, while at millimetre wavelengths the large dust is dominant.

Two important quantities that characterise scattering are the dust grain’s albedo, η_ν , defined as

$$\eta_\nu = \kappa_{\nu,\text{sca}} / \kappa_\nu, \quad (2.42)$$

which may be interpreted as the probability of a photon being scattered rather than absorbed, and the scattering angle, θ , which is the angle of deflection of the photon. This is typically cast as $\mu = \cos(\theta)$. Observations suggest that scattering off dust grains is strongly anisotropic (e.g. Witt et al., 1992; Calzetti et al., 1995; Burgh et al., 2002; Gordon, 2004). The probability that the photon will be scattered by an amount μ is given by the scattering phase function, $p(\mu)$. The most widely used (e.g. Steinacker et al., 2013) scattering phase function is the Henyey-Greenstein function (Henyey & Greenstein, 1941), which concisely parametrizes the scattering probability in terms of a single parameter, $g = \langle \mu \rangle$.

As mentioned in §1.2, observations with SPHERE and GPI can exploit the polarisation of dust grains. This can be achieved through a number of techniques, such as polarimetric differential imaging (PDI) or simply measuring the polarised intensity. As light from the central star is expected to be unpolarised, and scattering off dust grains induces polarisation, measuring the polarised flux selects photons that have been scattered off the disc (e.g. Gordon et al., 2001). Thus, given the additional information that the polarisation state of received light provides, it can often be beneficial to calculate this in radiative transfer simulations too. This therefore involves following not just the intensity of the radiation, I , but the full Stokes vector, (I, Q, U, V) . Q and V are linear and circular polarisations respectively, while U is the polarisation at an angle of 45° measured anticlockwise with respect to Q . The total polarised intensity is given by $P = \sqrt{U^2 + V^2 + Q^2}$ and P/I is the polarisation fraction. The inclusion of polarisation requires that the scattering phase function be replaced by a 4×4 scattering matrix. The scattering probability distribution in general then depends not only on the scattering angle, but also on the Stokes vector of the incoming photon. In simplest case of spherical and spherically symmetric dust grains, the scattering

2.2. Numerical Radiative Transfer

matrix takes on a simplified form; this is what will be used throughout this thesis.

While scattering and absorption remove intensity from a ray, contribution to the intensity from the medium via both thermal emission and scattering is captured through the emission coefficient, j_ν ; the total change in intensity along a ray is therefore given by

$$\frac{dI_\nu}{ds} = j_\nu - \kappa_\nu \rho I_\nu. \quad (2.43)$$

This is the radiative transfer equation. As was the case for the fluid equations, the basis for the radiative transfer equation is a statistical one, in that the absorbers are independent and randomly distributed, and that the cross section presented by the absorbers is small compared to the mean inter-particle distance. Given the complex structures of protoplanetary discs, with density, temperature and therefore composition gradients, and the possibility of absorption and scattering processes, the solution of this equation for realistic disc structures can only be obtained through numerical methods. By discretising the density (ρ) over a protoplanetary disc on a grid and defining the disc opacity (κ_ν) as a function of wavelength, the radiative transfer of energy from the central star throughout the disc can be calculated. Note that while the disc temperature varies radially through the disc, the opacity of astrophysical dust is independent of temperature (Bjorkman & Wood, 2001). Temperature differences do, however, mean that dust composition may vary – for example, exterior to ice lines dust grains may acquire icy mantles. It is therefore common to specify a number of dust species which exist separately at different radial locations in the disc, each with their own opacity.

While the density and opacity can be specified, j_ν may not be known *a priori*, however; the injection and removal rates of energy into/out of rays depend on local conditions, which in turn depend on the local radiation field, but this is the goal of the calculation. A single ray therefore cannot be considered in isolation; rays can cross each other, affecting the local conditions and therefore the injection and removal rates of energy into/out of each other. The radiative transfer equation must therefore be solved along all rays simultaneously. One popular and well-established method for achieving this is Monte Carlo radiative transfer, which simulates radiative transfer rather than explicitly solving the radiative transfer equation.

2.2.1 Monte Carlo Radiative Transfer

The method of Monte Carlo radiative transfer employs Monte Carlo techniques, wherein quantities are randomly sampled from probability distribution functions, in order to gradually build up a solution to the radiative transfer equation. More specifically, the energies, propagation distances and directions, and interactions of photons are all drawn from distributions that represent their functional forms. The source luminosity is divided into equal energy, monochromatic “photon packets”, which are stochastically emitted. In the case of a protoplanetary disc, the primary source is the central star, however the disc also emits radiation which must be accounted for.

The photon packets are treated like photons in practice, but in terms of their energy may contain $L_*/(Nh\nu_{\text{IR}}) \sim 10^{38}$ photons (h is Planck’s constant), where typical values for the stellar luminosity L_* and number of photon packets N have been adopted; equal energy packets of different frequencies will therefore represent different numbers of actual photons. Hereafter photon packets will simply be referred to as photons. The photons are emitted in random directions (but isotropically distributed), with random energies sampled from the stellar spectrum and interact – if at all – via scattering or absorption in the protoplanetary disc.

Each photon travels a distance τ without interacting (cf. equation (2.40)), which is randomly sampled from a probability distribution $p(\tau) = \exp(-\tau)$. The photon is thus destined to traverse a linear path (i.e.

a ray) towards the interaction location. Each segment of the ray corresponds to an amount $\Delta\tau = \kappa\rho\Delta s$, thus as the photon crosses a grid cell, the appropriate amount $\Delta\tau$ is subtracted from τ until the interaction location is reached, i.e. when $\Delta\tau > \tau$. Whether scattering or absorption occurs depends on the albedo: a random number is drawn, and if this is greater than the albedo, the photon is absorbed, otherwise it is scattered. The scattering angle depends on scattering phase function $p(\mu)$, so a random angle is drawn from the distribution, whether it is isotropic, Henyey-Greenstein or otherwise. The simulation finishes once all photons have left the computational domain. By considering large numbers (typically $\gtrsim 10^6$) of photon packets, each of the distributions will be sufficiently sampled to build up an accurate solution.

In its simplest form, the star and the disc are treated independently: the stellar photons are first emitted, absorbed by the dust and subsequently re-emitted according only to the absorbed radiation. Then the radiation field of the dust is taken into account; re-emission takes into account the emission from the disc. This is repeated until the re-emission has converged, i.e. the difference between the input and output energies has reduced to within a chosen threshold. This repetition of the dust re-emission is time consuming, thus there have been numerous attempts to devise schemes to avoid this. One could provide an initial guess for the dust temperature, e.g. from hydrostatic equilibrium, but this is not always possible or reasonable. Also, energy is only conserved asymptotically – i.e. once (if) convergence has been achieved.

One notable method to avoid iteration is the “instantaneous re-emission” method of Bjorkman & Wood (2001), which enforces radiative equilibrium and thus prevents the need for iteration. In thermal equilibrium, the laws of thermodynamics require that the radiation from any medium must depend only on its temperature (e.g. Rybicki & Lightman, 1986) (Kirchhoff’s law). The intensity of radiation of a given frequency in such a medium must therefore follow some universal function of temperature, namely the Planck function, B_ν :

$$B_\nu(T) = \frac{2h\nu^3/c^2}{\exp(h\nu/k_B T) - 1}, \quad (2.44)$$

where h is Planck’s constant, c is the speed of light in a vacuum and k_B is Boltzmann’s constant. In local thermodynamic equilibrium, $I_\nu = B_\nu(T)$ and $dI_\nu/ds = 0$. From equation (2.43), this latter condition implies that $j_\nu = \kappa_\nu\rho B_\nu(T)$. In the method of Bjorkman & Wood (2001), N_γ photons are emitted by the star with luminosity L_* over some time interval Δt , representing the duration of the simulation. When absorbed in the disc, a photon has a fraction of its energy E_γ added locally to the absorbing patch of the disc, heating it up. A new photon is instantaneously emitted from the absorbing grid cell with an energy that accounts for the locally modified disc temperature. By keeping track of the number N_i of photons absorbed in the i^{th} cell, the total energy absorbed by the cell can be calculated as

$$E_i^{\text{abs}} = N_i E_\gamma. \quad (2.45)$$

The energy emitted by the cell, whose volume is V_i , is

$$E_i^{\text{em}} = 4\pi\Delta t \int_V dV_i \int_0^\infty j_{\nu,i} d\nu, \quad (2.46)$$

which, in thermal equilibrium where $j_{\nu,i} = \kappa_\nu\rho_i B_\nu(T_i)$, is

$$E_i^{\text{em}} = 4\pi\Delta t m_i \int_0^\infty \kappa_\nu B_\nu(T_i) d\nu, \quad (2.47)$$

where m_i is the total dust mass in the cell. Substituting for the Planck mean opacity,

$$\kappa_P = \frac{\pi \int_0^\infty \kappa_\nu B_\nu(T_i) d\nu}{\sigma T_i^4}, \quad (2.48)$$

2.2. Numerical Radiative Transfer

the energy emitted by the cell can be written in terms of its temperature as

$$E_i^{\text{em}} = 4\Delta t m_i \kappa_{\text{P}}(T_i) \sigma T_i^4. \quad (2.49)$$

Equating equations (2.45) and (2.49), the dust temperature of the i^{th} cell after having absorbed N_i photons is

$$\sigma T_i^4 = \frac{N_i L_*}{4N_{\gamma} m_i \kappa_{\text{P}}(T_i)}. \quad (2.50)$$

Having updated the cell temperature to T_i by an amount ΔT , the absorbed photon must be re-emitted. The cell previously emitted photons at energies according to its emission coefficient $j'_{\nu,i} = \kappa_{\nu} \rho_i B_{\nu}(T_i - \Delta T)$. As such, these were emitted with an incorrect frequency distribution, corresponding to $B_{\nu}(T_i - \Delta T)$. The total energy that this cell *should* emit is given by j_{ν} corresponding to $B_{\nu}(T_i)$, thus the additional energy to radiate away can be determined from

$$\Delta j_{\nu} = j_{\nu} - j'_{\nu} = \kappa_{\nu} \rho [B_{\nu}(T_i) - B_{\nu}(T_i - \Delta T)]. \quad (2.51)$$

For a small temperature change, this can be approximated as

$$\Delta j_{\nu} \approx \kappa_{\nu} \rho \Delta T \frac{dB_{\nu}}{dT}. \quad (2.52)$$

Correcting for the previously emitted spectrum therefore simply entails instantaneously re-emitting the photon at a frequency sampled from Δj_{ν} .

This method is only possible in Monte Carlo radiative transfer, as the paths of the photon packets are followed so their absorption locations are known. It is also only possible for dust whose opacity is independent of temperature. The dust heats up as the stellar photons are absorbed and re-emitted throughout the disc, assuming that each dust grain acquires as much energy as it radiates away. The computed temperature is thus an equilibrium dust temperature. For most cases this should be a very good approximation, as the heating/cooling timescales for dust grains are typically short compared to dynamical timescales in protoplanetary discs (e.g. Draine & Li, 2001). It may not be valid in extremely optically thick regions or near shocks, however.

Using this method, the simulation ends once all photons (the number of which is pre-defined) have left the computational domain rather than once some pre-defined convergence criterion has been met. The only source of error is therefore the sampling error from the number of photon packets used. One issue with this scheme is that in very optically thin regions (where $\tau_{\nu} \ll 1$), the probability of photon absorption is very low, thus it is possible for cells to have $T_i = 0$ if no absorption events take place, which is unphysical. This can be rectified (as suggested in Bjorkman & Wood 2001) by employing the “continuous absorption” method introduced by Lucy (1999). In this method, the photon packet is assumed to deposit a fraction of its energy in all cells it crosses, rather than depositing all of its energy in the cell corresponding to the absorption location. Each cell therefore absorbs an amount

$$E_i^{\text{abs}} = N_i E_{\gamma} \frac{\kappa_{\nu, \text{abs}} \rho_i \Delta s_i}{V_i}, \quad (2.53)$$

where Δs is the length of the path the photon traverses across the cell. Since photon packets represent $\sim 10^{40}$ actual photons, this can be interpreted as that fraction of the photons within the photon packet being absorbed there. In this way, optically thin cells receive small amounts of energy and thus receive temperature increases that they otherwise would not, providing a more accurate and less noisy temperature calculation.

With the temperature of the dust throughout the disc known, it is possible to generate images of

protoplanetary discs, for example to compare models to observations or to generate synthetic observations of the outcomes of hydrodynamics simulations. In principle, this can be achieved by projecting the position of the photons emitted towards a given observation location onto the plane of the sky. However, given that the fraction of photons that will be emitted in this direction, even given some angular tolerance, will likely be very small, a number of techniques have been developed to improve the efficiency of image generation without crudely increasing the number of photons. Since the Monte Carlo approach described above yields the absorption and scattering emission coefficient j_ν , an image can be rendered by “ray-tracing”, i.e. simply integrating the radiative transfer equation along a set of pre-defined rays which end at the observation location.

2.2.2 RADMC-3D

RADMC-3D (Dullemond et al., 2012) is a software package for performing astrophysical radiative transfer calculations via the Monte Carlo method, developed primarily for continuum radiative transfer in dusty media. By providing a tabulated stellar spectrum, dust density distribution and opacity data, **RADMC-3D** can calculate the dust equilibrium temperature as well as generate images and spectra of radiating objects.

Both the instantaneous re-emission method of Bjorkman & Wood (2001) along with the continuous absorption method of Lucy (1999) are used to calculate the equilibrium temperature as described above. In order to generate images (and spectra), **RADMC-3D** performs an all-frequency Monte Carlo radiative transfer simulation given an input stellar spectrum to calculate the dust temperature, assuming that the dust is in radiative equilibrium with its radiation field. With the dust temperature known, ray-tracing can then be performed to generate an image. Accounting for scattering is a more complicated matter, however; while the all-frequency Monte Carlo radiative transfer simulation to determine the dust temperature could have, in principle, included a calculation of the scattering emission coefficient at each location in the grid for all wavelengths and observer position(s), this can require a huge amount of memory and thus is not practical for any realistic disc model. Therefore, to account for scattering, a monochromatic Monte Carlo simulation to calculate the scattering emission coefficient is performed at the desired observation wavelength, which is followed by the ray-tracing procedure towards the observer position.

RADMC-3D uses the pseudo-random number generator of L’Ecuyer (1988), implemented as **ran2** in Press et al. (1992), which generates random numbers uniformly distributed between 0 and 1. Thus, in order to draw random numbers according to the appropriate distribution (e.g. the Planck function for the stellar spectrum, or an exponential decay function for the photon path length etc.), the transformation method (e.g. Press et al., 1992) is used. The code was benchmarked in its original form (**RADMC** Dullemond 2011) by Pascucci et al. (2004).

The results of radiative transfer simulations presented in the following chapters were performed using **RADMC-3D** version 0.41, using OpenMP for parallelisation.

Chapter 3

Vortex Formation from an Accreting Planetary Embryo

3.1 Introduction

One of the most remarkable observational discoveries of the last decade is the wealth of substructure that has been observed in protoplanetary discs. As discussed in §1.4.2, a small fraction of discs show large-scale non-axisymmetric substructure in their continuum emission. However, these same discs do not show corresponding asymmetries in their near-infrared (NIR) emission. It is well understood that (sub-)millimetre and NIR observations probe different dust populations, with a simple rule of thumb being that the observation wavelength probes dust of a similar size; the micron-sized dust grains observed at NIR wavelengths are strongly coupled to the gas, while millimetre grains probed in continuum observations are only intermediately coupled, and rapidly drift up pressure gradients. The leading interpretation is therefore that these asymmetries result from anticyclonic vortices, which contain a pressure maximum at their centre and orbit at roughly the Keplerian speed (Barge & Sommeria, 1995), and can therefore concentrate dust azimuthally.

The most commonly invoked mechanism for vortex formation in protoplanetary discs is the Rossby wave instability (RWI), which is a purely hydrodynamical instability that can be triggered at sufficiently steep density gradients. Hydrodynamics simulations have shown that the density gradients found at the edges of gaps carved by giant planets can become steep enough to trigger the RWI. However, Hammer et al. (2017) pointed out that the short growth timescales of the giant planet prescribed in these simulations result in unphysically large perturbations which are capable of setting up the RWI-unstable density gradient at the gap’s outer edge. When the giant planet is grown over more realistic timescales, vortex production is significantly suppressed because viscosity has time to smooth out any steep gradients, and the planetary torque itself can reshape the gap edge. Hammer et al. (2017) thus found that under the most favourable conditions for vortex formation, vortex lifetimes were found to be limited. One may then argue that if these asymmetries are indeed vortices formed through the RWI, triggered by giant planets, we must be seeing them just after the planet has formed. However, if a RWI-generated vortex typically survives for $\sim 10^3$ orbits and discs are $\sim 10^5$ orbits old, observing the vortex so soon after its formation is unlikely. It is therefore uncertain whether the non-axisymmetric substructures seen in observations can be explained by vortices formed through this mechanism.

Motivated by the recent resolved observations of transition discs, Owen & Kollmeier (2017) proposed an alternative vortex formation process, with vorticity instead sourced by the accretion luminosity of a forming planetary core. Andrews et al. (2011) presented spatially resolved images of twelve transition discs and compared the measured 880 μm fluxes to those measured from previously observed discs in the Taurus and Ophiuchus star-forming regions (Andrews & Williams, 2005, 2007; Andrews et al., 2009, 2010). Intriguingly, they found that $\approx 25\%$ of the upper quartile were transition discs, implying that transition discs are amongst the most massive protoplanetary discs (Owen & Clarke, 2012). The leading interpretation (e.g. Rice et al., 2006; Pinilla et al., 2012b; Dullemond et al., 2018) is that an axisymmetric pressure maximum forms in gas which, as detailed in §1.4.1, can form a dust trap. Drift towards pressure maxima is fastest for dust with $\text{St} \sim 1$, thus dust traps will be able to build up a large supply of these solids; since the inward drift of dust is halted at the outer edge of the cavity, the continued drift of dust

Table 3.1: Transition disc and host properties from Andrews et al. (2011). Column 1 is the name of the host star, column 2 is the host spectral type, column 3 is the stellar mass, column 4 is the inferred dust mass and column 5 provides references from which stellar properties were obtained. Dust masses were obtained by fitting axisymmetric dust surface density profiles to the observed disc structure and integrating radially. Stellar properties have been updated from Andrews et al. (2011) where applicable.

Source	SpT	$M_*/M_\odot$	$M_d/M_\oplus$	References
MWC 758	A7	1.8	27	Acke et al. (2005); Salyk et al. (2013)
SAO 206462	F5	1.6	87	Fairlamb et al. (2015)
LkH α 330	G4	2.4	80	Manara et al. (2014)
SR 21	G4	2.1	20	Manara et al. (2014)
UX Tau A	G8	1.4	23	Espaillet et al. (2010)
SR 24 S	K2	2.0	150	Andrews et al. (2011)
DoAr 44	K2	0.97	23	Manara et al. (2014)
LkCa 15	K2	1.2	183	Manara et al. (2014)
RX J1615-3255	K7	1.2	427	Manara et al. (2014)
GM Aur	K5	1.4	233	Manara et al. (2014)
DM Tau	M2	0.56	133	Manara et al. (2014)
WSB 60	M6	0.24	93	Manara et al. (2014)

from the outer disc into the dust trap will enhance the local dust-to-gas ratio, potentially approaching unity (e.g. Pinilla et al., 2012a). The large dust masses resulting from this process can therefore explain the large millimetre fluxes.

Early high spatial resolution observations at sub-millimetre wavelengths (e.g. Ohashi, 2008; Brown et al., 2009; Mathews et al., 2012) were also beginning to show that the discs which showed large-scale azimuthal asymmetries were also transition discs. Owen & Kollmeier (2017) therefore sought to investigate a possible link between these two observational findings. They proposed that, since transition discs have large cavities surrounded by a dust ring containing $\sim 100M_\oplus$ (see table 3.1), these may be promising sites for planet formation: high dust-to-gas mass ratios are expected to promote planet formation processes, and with the radial drift barrier removed, the dust population can achieve grain growth up to the fragmentation barrier (typically centimetre sizes). It may therefore be possible to build up high enough a dust-to-gas ratio to trigger the streaming instability and thus form planetesimals, and ultimately a planetary embryo.

As a planetary embryo accretes solid material, the potential energy liberated by accretion locally heats the disc. For a sufficiently high accretion rate, the thermal structure of the disc can be modified beyond the embryo’s Hill sphere, making the disc locally baroclinic and therefore unstable to vortex formation. Owen & Kollmeier (2017) performed hydrodynamics simulations of the circumstellar gas, heated in the vicinity of an accreting planetary embryo, showing that a large-scale vortex can form within tens of orbits and survive for hundreds. Furthermore, they proposed a cycle of vortex formation and dissipation in order to explain the fact that vortices that only live for hundreds of orbits can be observed. The vortex will rapidly gather and trap dust from the dust ring, shutting off accretion onto the embryo. However, upon vortex dissipation, this dust will be released, potentially re-forming the dust ring and restarting the pebble accretion process, and thus vortex formation.

However, since these simulations did not explicitly include the dust component of the disc, they required the use of a prescription for the pebble accretion rate and hence the accretion luminosity. The aim of my work is therefore to compute the evolution of the dust along with the gas, enabling a self-consistent calculation of the pebble accretion rate onto the embryo. Specifically, I want to understand if and how the vortex formation process differs when the accretion rate is coupled to the local dust surface density, how long the vortex survives for, and whether it is possible to produce non-axisymmetric dust substructure similar to those present in observations. Not only do I wish to determine whether the

3.2. Physical Mechanism

asymmetry is morphologically similar to those observed, but also whether the asymmetry exists for an observable timescale, thus I also aim to establish whether the cycle of vortex formation and dissipation can occur. One final point of interest is the mass of the planetary core formed through this process, and whether it is possible can make a prediction of what mass of planet one might suspect to be present in discs showing similar asymmetries.

I first provide the theoretical details of this vortex formation process. I then provide a description of the setup for numerical simulations and justify my choice of stellar and disc properties. The results of dust and gas hydrodynamics simulations are then presented, followed by results of Monte Carlo radiative transfer simulations, providing a means of comparison to observations. The outcomes of these simulations are discussed in terms of the disc’s evolution and the planet’s growth. Finally, I address the limitations of my simulations and provide a comparison to Owen & Kollmeier (2017).

3.2 Physical Mechanism

In this section I present the stages of the planet formation scenario to be investigated, beginning with the trapping of dust in an axisymmetric gas pressure maximum and how this promotes the formation of a planetary embryo. I outline how the embryo grows by accreting from the surrounding material in the dust trap, followed by details of how this can lead to the formation of an anticyclonic vortex.

Embryo formation

As detailed in §1.4.1, the thermal pressure gradient provides a force that causes the gas orbital speed to deviate from Keplerian, which can cause the radially-drifting dust to become trapped at a local pressure maximum. With dust trapped at some radius and the continual inward drift of dust from the outer disc, the dust-to-gas mass ratio in the trap is likely to significantly exceed the interstellar medium value of 0.01, potentially reaching values close to unity (e.g. Pinilla et al., 2012a; Yang et al., 2017). It is in such regions of high dust-to-gas ratios that planet formation processes are likely to be enhanced. For example, planetesimal formation can be achieved through the streaming instability (Youdin & Goodman, 2005; Johansen & Youdin, 2007; Bai & Stone, 2010), which has been shown to work in regions of protoplanetary discs where dust-to-gas ratios are expected to be enhanced (e.g. Drażkowska & Dullemond, 2014; Auffinger & Laibe, 2018) including in dust traps and locations with low pressure gradient (e.g. Carrera et al., 2021). Numerical simulations of the streaming instability have demonstrated the formation of planetesimals with masses up to $\sim 10^{-4} M_{\oplus}$ (e.g. Johansen et al., 2015; Simon et al., 2016; Schäfer et al., 2017; Abod et al., 2019), which can subsequently form planetary embryos through a combination of pebble and planetesimal accretion (e.g. Liu et al., 2019a; Voelkel et al., 2021a,b). It is therefore highly likely that a planetary embryo will form in such a dust trap.

It is also conceivable that high dust-to-gas ratios are conducive to the formation of multiple embryos; in such a scenario the system would exhibit oligarchic growth, with the first embryo to form dominating the accretion process (e.g. Kokubo & Ida, 1998), as the accretion rate increases rapidly with embryo mass. Axisymmetric dust traps therefore provide ideal conditions for the formation of a planetary embryo, and I predicate that it is almost inevitable in a sufficiently strong pressure maximum. In the following modelling I therefore simply invoke the formation of an embryo in the peak of the dust trap.

Accretion onto the embryo

An embryo which forms in an axisymmetric dust trap will find itself in a prime position for growth into a planetary core. Since it is dust grains with Stokes numbers of order unity that are most efficiently and rapidly trapped in pressure maxima, the newly formed embryo will be surrounded by a plethora

Chapter 3. Vortex Formation from an Accreting Planetary Embryo

of these pebbles. Furthermore, as discussed in §1.5.1, dust grains with Stokes numbers of order unity also experience enhanced accretion rates by pebble accretion. This therefore naturally gives rise to a rapid and efficient core growth scenario, as the dust which is most efficiently trapped also experiences the highest accretion rate onto the embryo.

Within the typical parameters expected for dust traps, the transition mass from spherical to planar modes of pebble accretion can be obtained from equation (1.65):

$$M_P \gtrsim 2.5 \times 10^{-4} M_\oplus \left(\frac{M_*}{2M_\odot} \right) \left(\frac{H_g/r}{0.05} \right)^3 \left(\frac{\alpha}{10^{-4}} \right)^{3/2}. \quad (3.1)$$

This is more massive than the typical planetesimals formed via the streaming instability (e.g. Johansen et al., 2015; Simon et al., 2016); further growth beyond the streaming instability is therefore required to surpass the transition mass. This could occur, for example, through pebble accretion in the shear regime, or by planetesimal accretion onto the embryo. Further planetesimal formation can occur through, for example, the streaming instability triggered by dust coagulation (Drażkowska & Dullemond, 2014). Alternatively, Windmark et al. (2012a,b) demonstrated how pure coagulation can proceed up to planetesimal sizes, bypassing the fragmentation barrier by “sweep-up”. The success of this mechanism in producing planetesimals in a pressure bump was demonstrated by Drażkowska et al. (2013), although it was questioned by Booth et al. (2018), who found that this can only occur given a mass-ratio-dependent fragmentation velocity.

Estimates of dust masses in ringed discs from their millimetre emission range from a few tens of Earth masses up to a few hundred (e.g. Andrews et al., 2011; Long et al., 2018; Dullemond et al., 2018). Dust surface densities derived from these observations suggest a pebble accretion rate of

$$\dot{M}_{\text{peb}} \sim 1 M_\oplus \text{ kyr}^{-1} \left(\frac{M_P}{1 M_\oplus} \right)^{2/3} \left(\frac{M_*}{2 M_\odot} \right)^{-1/6} \left(\frac{a}{35 \text{ AU}} \right)^{1/2} \left(\frac{\Sigma_d}{5 \text{ g cm}^{-2}} \right). \quad (3.2)$$

Should an embryo form, accretion will therefore be rapid, and the available mass in the trap provides the potential for a massive planetary core to form if accretion is uninhibited.

Owen & Kollmeier (2017) showed that the growth of a planet in a dust trap can be understood by considering a simple model in which the pebble accretion rate is

$$\frac{dM_P}{dt} = 2 \Omega_K a^2 \left(\frac{M_P}{3 M_*} \right)^{2/3} \Sigma_{\text{peb}}, \quad (3.3)$$

coupled to the rate at which pebbles are depleted from the dust trap,

$$\frac{d\Sigma_{\text{peb}}}{dt} = \left(\frac{H_g}{a} \right)^{-1} \frac{\dot{M}_P}{2\pi a^2}. \quad (3.4)$$

Note that the Stokes number dependence on the accretion rate has been dropped, thus this represents the optimal accretion rate (i.e. $\text{St} = 1$). Another assumption made is that the radial width of dust trap is of order the gas scale height, based on a simple argument that the back-reaction on the gas couples the dust trap width to the gas scale length. Since the pebble accretion rate scales with $R_H^2 \propto M_P^{2/3}$, the accretion timescale scales as $t_{\text{acc}} = M_P / \dot{M}_{\text{peb}} \propto M_P^{1/3}$, thus the planet mass increases as $M_P \propto t^3$. Given expected dust trap properties, then, accretion should proceed rapidly:

$$t_{\text{acc}} \approx 800 \text{ yr} \left(\frac{M_P}{1 M_\oplus} \right)^{1/3} \left(\frac{M_*}{2 M_\odot} \right)^{1/6} \left(\frac{a}{35 \text{ AU}} \right)^{-1/2} \left(\frac{\Sigma_{\text{peb}}}{5 \text{ g cm}^{-2}} \right)^{-1}. \quad (3.5)$$

By numerically solving equations (3.3) and (3.4), the growth of the planet as the dust trap becomes

3.2. Physical Mechanism

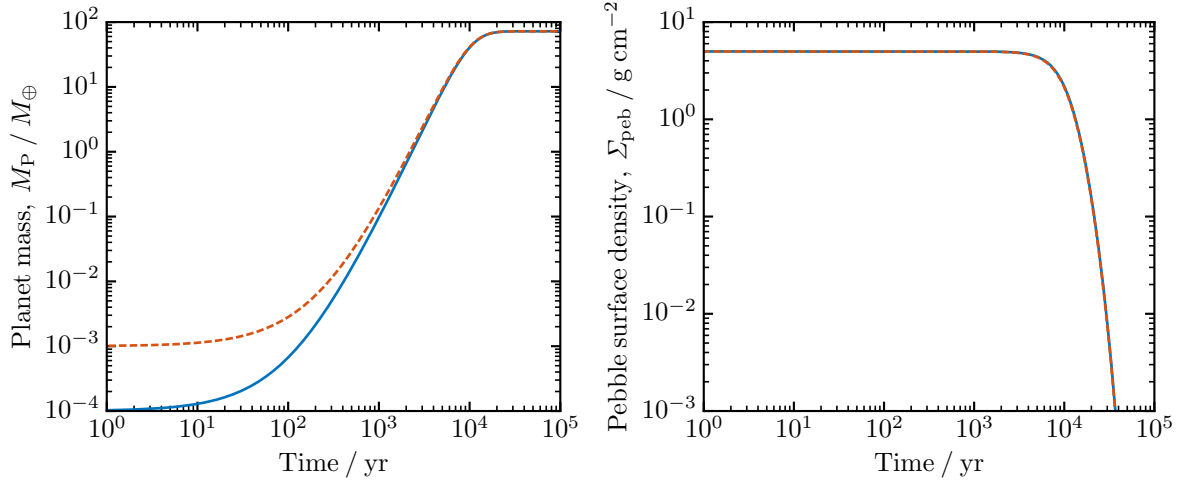


Figure 3.1: Planet growth and pebble depletion in a simple model of pebble accretion in a dust ring, for an initial planetary embryo of mass $10^{-4}M_{\oplus}$ (blue solid line) and $10^{-3}M_{\oplus}$ (orange dashed line). In both cases the planet accretes all of the mass in the ring within ≈ 20 kyr.

depleted can be understood. Figure 3.1 shows the planet mass and pebble surface density over time in this simple scenario for a planet located at 35 AU from its host star. The planet accretes all the mass from the ring ($M_{\text{peb}} \approx 2\pi a \Sigma_{\text{peb}} H_g$), achieving a final mass of $\approx 72M_{\oplus}$. This model demonstrates that the final mass is independent of the initial mass, given the rapid initial accretion rate. This final mass is, however, well above the typical pebble isolation mass. One should expect the planet to carve a gap and prevent further accretion beyond $\approx 20M_{\oplus}$. Thus, while this model is useful for demonstrative purposes, it clearly has its limitations and highlights the need for hydrodynamics simulations.

Vortex formation

The intrinsic axisymmetric nature of smooth Keplerian discs, where gas pressure and density decrease monotonically away from the star, makes them stable to vortex formation. In §1.4.2 I derived an expression for the time evolution of vortensity in an inviscid, compressible fluid – equation (1.56). Specialising to a thin, Keplerian disc, an equivalent derivation can be carried out for the vertically integrated density, pressure and velocities, the result of which is

$$\frac{D}{Dt} \left(\frac{\boldsymbol{\omega}}{\Sigma_g} \right) = \frac{\boldsymbol{\nabla} \Sigma_g \times \boldsymbol{\nabla} \mathcal{P}}{\Sigma_g^3}, \quad (3.6)$$

noting that the term $(\frac{\boldsymbol{\omega}}{\Sigma_g} \cdot \boldsymbol{\nabla})\boldsymbol{v}$ vanishes since, for vertically integrated quantities, $\boldsymbol{\omega}$ and $\boldsymbol{v}$ are perpendicular. In general, unperturbed Keplerian discs are barotropic, with $\mathcal{P}(\Sigma_g)$; their density and pressure gradients are aligned, so there is no forcing term, and their vortensity remains constant. However, the presence of an accreting planetary embryo offers a scenario in which the disc can become locally baroclinic, such that the surface density and pressure gradients are no longer aligned, and therefore provide a source of vortensity (Owen & Kollmeier, 2017).

On the basis that the kinetic energy of the accreted pebbles is converted into the embryo’s thermal energy and ultimately released as radiation, accretion onto the embryo will generate a luminosity of

$$L = \frac{GM_P \dot{M}_{\text{peb}}}{R_P}, \quad (3.7)$$

where R_P is the radius of the planet. Under the assumption that the disc is optically thin to infrared

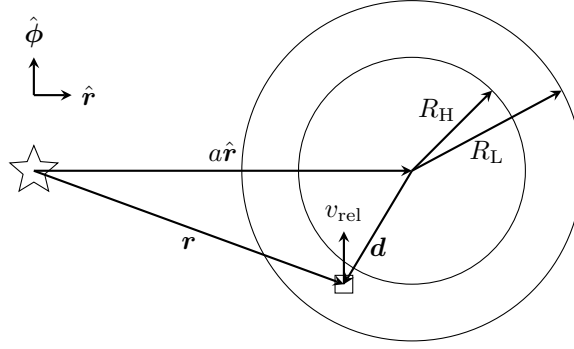


Figure 3.2: Accretion luminosity schematic. The planet is located at position vector $a\hat{r}$, with Hill radius R_H and luminosity radius R_L . The square represents a gas parcel at position vector $\mathbf{r}$, a distance d from the planet, moving with speed v_{rel} relative to the planet. Not to scale.

radiation in the vicinity of the planet (as is expected in the outer regions of the disc), and that the change in luminosity of the planet is instantaneously translated into a change in temperature, the distance at which the blackbody temperature solely due to the planet's radiation would equal the background disc temperature set by the central star, T_b , is

$$R_L = \left(\frac{L}{16\pi\sigma T_b^4} \right)^{1/2}, \quad (3.8)$$

where σ is the Stefan-Boltzmann constant. Thus the accreting embryo will generate a hotspot of radius R_L . The temperature of the disc is therefore approximately given by

$$T = T_b \left[1 + \left(\frac{R_L}{d} \right)^2 \right]^{1/4}, \quad (3.9)$$

where d is the distance from the planet.

Given the accretion timescale calculated above, the accretion luminosity is likely to be significant. However, in order to determine its impact on the disc, it is informative to establish whether the luminosity is sufficient to affect the gas beyond the region in which its gravity dominates, i.e. its Hill sphere. For this to be the case, the luminosity radius must exceed the embryo's Hill radius:

$$\begin{aligned} \frac{R_L}{R_H} &= \left(\frac{GM_P \Omega_K \Sigma_{\text{peb}}}{8\pi R_P \sigma T_b^4} \right)^{1/2} \\ &\approx 1.6 \left(\frac{M_P}{1M_\oplus} \right)^{1/2} \left(\frac{R_P}{1.4R_\oplus} \right)^{-1/2} \left(\frac{\Sigma_{\text{peb}}}{5 \text{ g cm}^{-2}} \right)^{1/2} \left(\frac{T_b}{30 \text{ K}} \right)^{-2} \left(\frac{\Omega_K}{1.4 \times 10^{-9} \text{ s}^{-1}} \right)^{1/2}. \end{aligned} \quad (3.10)$$

Thus, for parameters characteristic of dust traps, this is sufficient to create a hotspot around the embryo that is comparable in size to the embryo's Hill sphere.

For the hotspot to significantly affect the gas in its vicinity, however, the gas must have enough time to be heated by the accreting planet, in spite of the Keplerian shear. The necessary condition for this is that the advective timescale must be longer than the radiative timescale. A schematic of the relevant quantities for the following argument is shown in figure 3.2. The advective timescale can be approximated by considering the relative azimuthal velocity of a gas parcel with respect to the planet:

$$v_{\text{rel}} \approx \frac{3}{2} \Omega_K x_P, \quad (3.11)$$

where $x_P = (a\hat{r} - \mathbf{r}) \cdot \hat{r}$, i.e. the radial component of the distance from the planet (cf. equation 1.59).

3.2. Physical Mechanism

The advective timescale $t_{\text{adv}} = x_P/v_{\text{rel}}$ is therefore

$$t_{\text{adv}} \approx \frac{2}{3\Omega_K}, \quad (3.12)$$

which is independent of the distance from the planet. The radiative timescale can be approximated as

$$t_{\text{rad}} \approx \frac{4\pi d^2 k_B T_b}{\mu L \kappa}, \quad (3.13)$$

which corresponds the time taken to heat up from a temperature of T_b , given an absorption opacity κ at a distance d from the planet. μ is the mean molecular weight of the gas, assumed to be $2.4m_H$. Thus

$$\begin{aligned} \frac{t_{\text{rad}}}{t_{\text{adv}}} &= \frac{3\Omega_K k_B}{8\mu\kappa\sigma T_b^3} \left(\frac{d}{R_L}\right)^2 \\ &\approx 0.12 \left(\frac{d}{R_L}\right)^2 \left(\frac{\Omega_K}{1.4 \times 10^{-9} \text{ s}^{-1}}\right) \left(\frac{T_b}{30 \text{ K}}\right)^{-3} \left(\frac{\kappa}{0.1 \text{ cm}^2 \text{ g}^{-1}}\right)^{-1}. \end{aligned} \quad (3.14)$$

$\kappa \sim 0.1 \text{ cm}^2 \text{ g}^{-1}$ has been adopted for far-infrared wavelengths, based on calculations by D'Alessio et al. (2001). As this ratio is less than unity, gas parcels do have enough time to be heated. The pressure structure associated with this hotspot will therefore cause gas parcels to deviate from the horseshoe streamlines that would otherwise form outside of the Hill sphere.

The behaviour of a gas parcel crossing the hotspot can be analysed by considering the steady-state fluid equations along a streamline with path length $d\ell$:

$$\frac{\partial}{\partial \ell}(\Sigma_g v) = 0, \quad (3.15)$$

$$v \frac{\partial v}{\partial \ell} + \frac{1}{\Sigma_g} \frac{\partial \mathcal{P}}{\partial \ell} = 0. \quad (3.16)$$

Combining these, and adopting an isothermal equation of state $\mathcal{P} = c_s^2 \Sigma_g$, yields an expression for the flow velocity:

$$\left(\frac{v^2}{c_s^2} - 1\right) \frac{\partial \log v}{\partial \ell} = -\frac{\partial \log c_s^2}{\partial \ell}. \quad (3.17)$$

Equation (3.11) shows that, for distances from the planet $d \gtrsim H_g$, the background Keplerian flow becomes supersonic with respect to the planet. The size of the hotspot relative to the gas scale height is

$$\begin{aligned} \frac{R_L}{H_g} &= \left[\frac{GM_P^{5/3} \Omega_K \Sigma_{\text{peb}} a^2}{8\pi (3M_*)^{2/3} R_P \sigma T_b^4 H_g^2} \right]^{1/2} \\ &\approx 0.26 \left(\frac{M_P}{1M_\oplus}\right)^{5/6} \left(\frac{R_P}{1.4R_\oplus}\right)^{-1/2} \left(\frac{M_*}{2M_\odot}\right)^{-1/3} \\ &\quad \times \left(\frac{\Sigma_{\text{peb}}}{5 \text{ g cm}^{-2}}\right)^{1/2} \left(\frac{H_g/a}{0.05}\right)^{-1} \left(\frac{\Omega_K}{1.4 \times 10^{-9} \text{ s}^{-1}}\right)^{1/2} \left(\frac{T_b}{30 \text{ K}}\right)^{-2} \end{aligned} \quad (3.18)$$

The heated gas is therefore expected to be subsonic relative to the planet and the hotspot. This is critical, since the flow time must be shorter than the sound-crossing time in order for the gas to adjust to the new temperature and achieve dynamical equilibrium. Thus, working in the subsonic limit, with the distance from the planet $d < H_g$,

$$\frac{\partial \log v}{\partial \ell} \approx \frac{\partial \log c_s^2}{\partial \ell}. \quad (3.19)$$

Then, by equation (3.15),

$$\frac{\partial \log \Sigma_g}{\partial \ell} \approx -\frac{\partial c_s^2}{\partial \ell}, \quad (3.20)$$

thus the pressure gradient across the hotspot is

$$\frac{\partial \log \mathcal{P}}{\partial \ell} = \frac{\partial \log \Sigma_g}{\partial \ell} + \frac{\partial c_s^2}{\partial \ell} \approx 0. \quad (3.21)$$

This means that the pressure across the hotspot remains approximately constant, while the gas surface density drops, inversely tracking the temperature. Thus $\nabla \Sigma_g \times \nabla \mathcal{P} \neq 0$, which means this process can source vortensity.

Since an anticyclonic vortex contains a pressure maximum at its centre, solids will drift into the vortex and become trapped. This accumulation of dust will occur at a rate dependent on their Stokes number, maximised for $St \sim 1$. As discussed above, the grains that contribute most to the pebble accretion rate are also those with $St \sim 1$. The presence of the vortex therefore has the potential to shut off accretion onto the planet as it traps the pebbles, thereby removing the source of vorticity as well as halting the planet's growth. Similarly, once the vortex dissipates, the dust it accumulated will be released and potentially become available for accretion onto the planet again. This is the basis for the cycle of vortex formation and dissipation proposed by Owen & Kollmeier (2017); explicitly including the dust will allow me to couple the accretion rate of solids onto the planet to the local dust surface density and thus follow the evolution of both the planet and the disc in a self-consistent manner. My simulations should therefore show the accretion cease if and when the vortex has formed and trapped enough dust, as well as any further planet growth and potential vortex formation once the original vortex has dissipated.

3.3 Method

In order to test this vortex formation scenario, I ran a number of hydrodynamics simulations of a protoplanetary disc, initialised with an axisymmetric dust trap containing a planetary embryo. Using FARGO3D with the modifications described in §2.1, the evolution of both the gas and dust components and the pebble accretion rate onto the embryo were calculated. I performed simulations that both include and exclude the thermal feedback on the disc from the accretion onto the planet, in order to isolate the impact of the thermal feedback on the outcome of planet growth and the resulting disc substructure.

Disc setup

In testing this vortex formation mechanism, I used disc parameters similar to those in which the majority of both small- and large-scale, non-axisymmetric substructures have been observed (e.g. IRS 48, HD 143006, HD 163296). I therefore modelled the system on that of a Herbig Ae/Be star of mass $M_* = 2M_\odot$ and a disc that extends out to 175 AU. In order to remain agnostic to the origin of the axisymmetric pressure maximum in the simulation, I set up an initial disc structure that already contained an inner cavity, rather than letting one result from some process (such as planet–disc interactions or photoevaporation) as part of the simulation. I therefore chose a surface density profile that results in a local pressure maximum (at which point η vanishes), but with a gradient shallow enough to prevent RWI onset so as to avoid other vortex formation through this mechanism. The gas surface density profile takes the form

$$\Sigma_g = \frac{\Sigma_0}{2} \left(\frac{r}{R_0} \right)^{-1} \left[1 + \operatorname{erf} \left(\frac{r - R_1}{\sqrt{2}\sigma} \right) + \epsilon \right], \quad (3.22)$$

which mimics an inner cavity and a typical $\Sigma_g \propto r^{-1}$ power-law profile at large radii (cf. figure 2.1). The characteristic radius, R_0 , is set to 35 AU. The corresponding value of gas surface density, Σ_0 , is set by

3.3. Method

the total disc mass, $M_{\text{disc}} = 0.05M_*$, a choice motivated by typical values for discs around similar mass stars when crudely assuming a global dust-to-gas mass ratio of 0.01 (e.g. Andrews et al., 2011; Ansdell et al., 2016; Dullemond et al., 2018). The choice of $R_1 = 1.2R_0$ and $\sigma = 0.663R_0$ results in a surface density maximum at $r \approx 50$ AU. $\epsilon = 10^{-3}$ prevents the surface density from vanishing at small radii in the initial condition. Since I am concerned with regions of the disc $\gtrsim 10$ AU from the central star, the gas density is sufficiently low ($\Sigma_g \sim 5 \text{ g cm}^{-2}$) that the molecular mean free path is larger than the size of dust grains, which justifies my use of the Epstein approximation. For these choices, the minimum value of the Toomre Q parameter is ~ 5 , hence my choice to neglect self-gravity is valid. According to the fragmentation-limited model of Birnstiel et al. (2012), this gas density profile allows for a maximum grain size of approximately 3 cm at the planet location.

The vertical scale height of the disc varies with radial distance from the central star, according to equation (1.16), with $h_0 = 0.05$, which produces a flared disc consistent with a background temperature profile of $T_b \propto r^{-0.5}$. The pressure and density are linked through the locally isothermal equation of state

$$\mathcal{P} = c_s^2 \Sigma_g, \quad (3.23)$$

thus the gas pressure is maximised at $r \approx R_0 \approx 35$ AU.

Perhaps the choice that will make the biggest difference to the outcome of my simulations is the total mass of solids in the dust ring. To stress-test this model for vortex formation, I chose a dust mass comparable to some of the highest dust masses inferred from observations. I used an initial dust mass of $150M_\oplus$, which is comparable to the dust masses in the most massive transition discs (cf. table 3.1; Andrews et al., 2011), noting that the simulations presented here are at the extreme end of the expected outcome of planet formation by pebble accretion in a dust ring. Higher dust masses will naturally produce initially higher accretion rates (equation 1.64) and hence stronger thermal feedback. In the next chapter I will vary the initial dust mass and the disc temperature, mapping out the outcomes of planet formation in such a dust ring.

I used four species of dust in the simulations, each corresponding to a population of fixed size, spherical grains. Each species represents grains of 1 μm , 1 mm, 3.16 mm and 1 cm in radius respectively. All dust species are assumed to have a material density, $\rho_\bullet$, of 3 g cm^{-3} . The evolution of the dust population through fragmentation and coagulation is neglected. My choice of maximum grain size is motivated by the assumption that within a dust trap, dust grows up to the fragmentation barrier ($s_{\text{frag}} \approx 3 \text{ cm}$) and achieves a steady-state distribution of grain sizes given by the MRN distribution (Birnstiel et al., 2011). The number, n , of dust grains within each species is therefore given by $dn/ds \propto s^{-3.5}$ (cf. §1.3.2).

The total mass of dust in the disc is

$$M_d = \int_{s_{\min}}^{s_{\max}} m(s) dn, \quad (3.24)$$

where $m(s)$ is the mass of a dust grain with radius s . Thus for spherical dust grains with an MRN size distribution their mass distribution can be calculated as:

$$M_d \propto \int_{s_{\min}}^{s_{\max}} s^{-1/2} ds = \sqrt{s_{\max}} - \sqrt{s_{\min}}. \quad (3.25)$$

There is therefore a factor $\sqrt{10}$ decrease in the total mass of grains a factor 10 smaller in size. Since the dust used here does not form a continuous size distribution but instead represent four discrete populations (indexed by k in the following) separated by factors of $\sqrt{10}$ in size, their masses are calculated from $M_{k+1}/M_k = \sqrt{s_{k+1}/s_k}$, with M_0 being the mass in the largest dust size (s_0) population, which can be

calculated from

$$M_0 = M_d / \sum_{k=0}^n 10^{-k/4}. \quad (3.26)$$

n is the number of size bins used; note that while four dust species are included in the simulation, n actually represents the total number of bins between the smallest and largest grain sizes in the distribution, i.e.

$$n = \frac{\log_{10}(s_{\max}/s_{\min})}{\log_{10}(s_k/s_{k+1})}. \quad (3.27)$$

The total dust mass in the k th size bin is therefore

$$M_k = M_0 \sqrt{\frac{s_k}{s_0}}. \quad (3.28)$$

The initial dust surface density distribution is determined by numerically solving the steady-state continuity equation for an axisymmetric disc, in an identical manner to that used to test the dust velocity update, described in §2.1.2. I solved equation (2.33) for each dust species, and verified the solutions using FARGO3D by letting the dust evolve from $\Sigma_d(r) = 0.01\Sigma_g(r)$ to a steady state. It is important to note that the dust implementation neglects the feedback on the gas; while this may affect the width of the initial dust ring (e.g. Yang & Zhu, 2020) and the lifetime of any vortices that form, at this stage the aim is to isolate the effects of the thermal feedback. The value of the dust surface density floor used is $10^{-9} \text{ g cm}^{-2}$. The kinematic viscosity of the gas is defined via the alpha viscosity prescription, with $\alpha = 10^{-4}$. The Schmidt number is taken to be 1, such that the diffusion coefficient of the dust is equal to the alpha viscosity parameter.

The simulation domain extends radially from $r_{\text{in}} = 0.54R_0$ to $r_{\text{out}} = 5.0R_0$. I varied the boundary locations to ensure that the area of interest, and thus the results, were not affected by the choice of domain size. I employed an open boundary condition at the inner disc edge, allowing material to flow out of the computational domain towards the host star. At the outer boundary the dust and gas radial velocities are mirrored. Outside of the computational domain the dust densities are set to the floor value, since all the dust is located in the pressure trap in the initial condition, while the gas density is kept at the values defined by equation (3.22). Wave-killing zones were employed near the boundaries as described in §2.1.1. The wave-killing zones extend from the domain boundaries according to equation (2.15), with $a_{\text{damp}} = 1.15$ and the damping is implemented according to equation (2.16), with a damping timescale of 1.5 orbital periods at the respective boundary radii.

In order to resolve the Hill sphere of the planetary embryo at its initial mass, variable radial cell spacing was used such that the resolution was maximised at the location of the planet. Specifically, the radial cell spacing is given by

$$\Delta r = c_1 r \left\{ 1 - c_2 \exp \left[- \left(\frac{r-a}{\sqrt{2}w} \right)^2 \right] \right\}, \quad (3.29)$$

which connects logarithmically-spaced cells ($\Delta r \propto r$) with a Gaussian function of amplitude c_2 and full-width at half-maximum w centered on a , providing a narrow region of increased resolution around the embryo's location. c_2 is set to 0.8 and w is set to $0.1R_0$. c_1 is a constant which is solved for such that

$$\sum_{i=0}^{N_r-1} r_{i+1} - r_i = \frac{r_{\text{out}} - r_{\text{in}}}{N_r}, \quad (3.30)$$

where N_r is the number of cells in the radial direction. In practice this is achieved by solving for c_1 using a simple bisection method. The minimum resolution is set by the requirement of resolving the Hill sphere of the embryo at its initial mass with at least eight cells. While this results in grid cells having a

3.3. Method

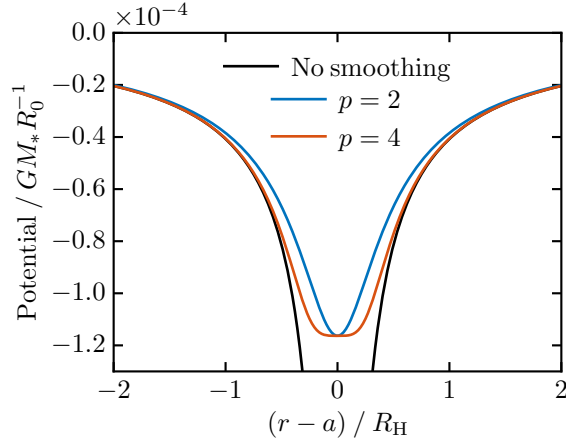


Figure 3.3: The planetary potential is smoothed in order to avoid the singularity at the location of planet. A smoothing index $p = 2$ is typically used, however, $p = 4$ reproduces the potential at the Hill sphere boundary more accurately.

variable aspect ratio, it provides a suitable balance between resolution and numerical expense. In total, 1920 cells were used in the radial direction, 660 of which cover $r \in [0.8, 1.2]a$, and 7680 uniformly-spaced cells in the azimuthal direction; this was found to be a satisfactory compromise between convergence and computational expense (see the resolution study in §3.4.1).

To isolate the physics in this initial work, the planet is kept on a fixed orbit, i.e. torques on the planet from the disc are neglected, preventing the planet from migrating radially through the disc. The potential of the planet is smoothed over a distance $\epsilon = 0.35R_H$ so as to avoid the singularity at the planet location. The smoothed potential, Φ_P , takes the form

$$\Phi_P(d) = -\frac{GM_P}{(d^p + \epsilon^p)^{1/p}}. \quad (3.31)$$

An index of $p = 2$ is typically chosen; here, however, $p = 4$ was used, as this provides a more accurate form of the potential at distances down to the Hill radius, as shown in figure 3.3.

Pebble accretion and thermal feedback

The planetary embryo is initialised with a mass $0.1M_\oplus$, positioned at the centre of the dust trap ($a \approx 35$ AU), since this is the dust surface density peak and thus the most likely site of embryo formation. This choice of initial mass is a compromise between being low enough that it is sufficiently close to the mass threshold for planar pebble accretion calculated in equation (3.1), and high enough that resolving its Hill sphere is feasible. Different initial masses were tested but, as demonstrated in figure 3.1, this choice does not qualitatively change the outcome of the simulations, but merely introduces a time offset.

With an initial mass above the transition mass, the embryo always accretes at a rate roughly given by equation (1.64). The accretion luminosity resulting from pebble accretion is calculated according to equation (3.7), with the planet radius calculated as a function of its mass following the mass-radius relationship of Fortney et al. (2007), assuming an ice mass fraction of 0.5. While this choice of ice mass fraction is motivated purely by the planet’s orbital separation of 35 AU, it allows the simulation to veer on the conservative side of parameter space since a rocky core would yield a higher accretion luminosity.

The accretion luminosity is used to prescribe the radius of the hotspot around the planet. The

temperature within the hotspot is calculated as

$$T = T_b \left(1 + \frac{R_L^2}{d^2 + s^2} \right)^{1/4}, \quad (3.32)$$

where T_b is the background disc temperature due to the central star, set via equation (1.16). s is a smoothing parameter which is set to $0.65R_{H,0}$, where $R_{H,0}$ is the Hill radius of the planet at its initial mass. Since a locally isothermal equation of state is used, the temperature increase due to the hotspot is added by accounting for the resulting increase in the sound speed during the pressure calculation step.

3.4 Results

3.4.1 Hydrodynamics simulations

No planet test

In order to verify that the initial gas density distribution is stable against vortex formation via the RWI or otherwise, a simulation was performed in which the disc is left to evolve without the planetary embryo present. Over 150 kyr (equivalent to ~ 1000 orbits at 35 AU) the gas pressure maximum slowly reduces due to viscosity and the dust trap widens. The dust remains trapped in the axisymmetric pressure maximum, which gradually moves towards the star due to viscous accretion. No substructure forms in the gas or dust, confirming that the density gradient maintaining the dust trap is stable to vortex formation.

No accretion luminosity test

Since I wish to isolate the effect of the accreting embryo's thermal feedback on the disc, I performed a simulation with the accretion luminosity artificially kept at zero while the embryo accretes pebbles. Figure 3.4 shows snapshots of the gas surface density from this simulation. As the planetary embryo accretes the solid material and increases in mass, the strength of its gravitational interaction with the disc increases, launching spiral density waves and carving a gap in the disc. The planet eventually reaches $41M_\oplus$, by which point it has carved a gap deep enough to prevent further drift of dust into its orbit. Note that this is consistent with the standard pebble isolation mass that occurs for planets forming in discs without pressure traps.

After ~ 35 kyr the gap edges become unstable to the RWI, which leads to the formation of vortices at either gap edge, which then dissipate fairly quickly. The result of this can be seen in the final panel in figure 3.4, where the gas surface density immediately outside of the gap is slightly enhanced in the shape of an extended crescent. Dust that was present in the co-orbital region but was not accreted onto the embryo accumulates close to the classical L_5 Lagrange point, though this does not occur at the opposite Lagrange point, L_4 . The location of the L_5 trap varies between dust species, a physical process discussed in §3.5.

Impact of accretion luminosity

With the behaviour of the system due purely to the gravity of the planet established, the results of the simulation where I account for the accretion luminosity of the embryo are now presented. Figure 3.5 shows the evolution of the gas surface density in the disc at a series of snapshots over the ≈ 230 kyr (~ 1600 orbits) simulation, as well as the mass that the embryo has grown to at each snapshot. As the embryo accretes and the energy released heats the surrounding gas, waves are launched from the hotspot

3.4. Results

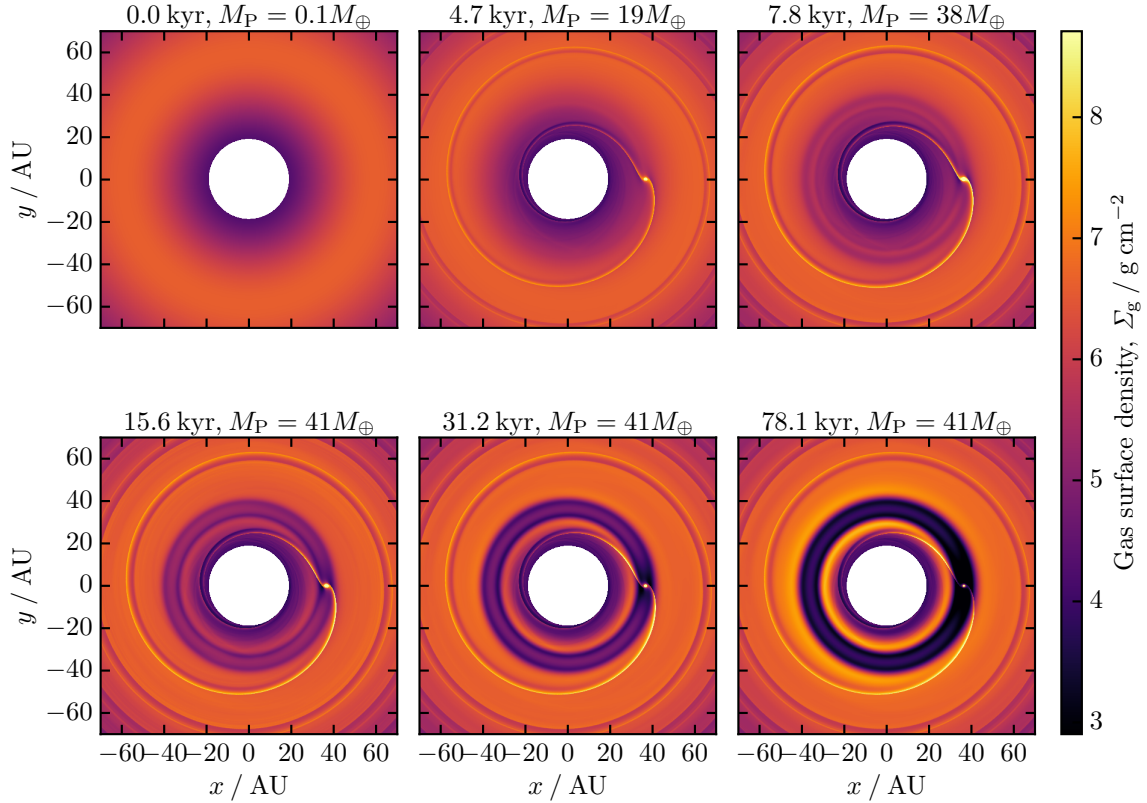


Figure 3.4: Gas surface density snapshots for the simulation in which the planetary embryo undergoes pebble accretion without heating the disc. The disc is viewed in a frame rotating with the planet, which therefore remains fixed at $(35, 0)$. Above each panel is the time of the snapshot since the beginning of the simulation and the mass that the planet has grown to. As the planet grows it creates a shallow gap in the disc which inhibits further inward drift of dust, preventing further accretion. Dust within the co-orbital region is confined to horseshoe orbits, which halts the growth of the planet.

around the embryo. By 4 kyr, an enhancement in gas density in the vicinity of the embryo can be seen, indicating the initial stages of vortex formation are underway. This can be confirmed by inspecting the vortensity, which is shown in figure 3.6, where a region of negative vortensity is associated with this enhanced surface density (note the difference in spatial scale between figures 3.5 and 3.6). By 5 kyr an anticyclonic vortex has formed at the location of the planet. The vortex’s interaction with the disc causes a slight radial migration and subsequent separation of the vortex from the planet’s location. Two smaller, weaker vortices can also be seen (e.g. at 9.4 kyr in figure 3.6), which separate from the planet shortly after the first, largest vortex. The planet becomes massive enough that it carves a deep gap and confines the vortices to horseshoe orbits, before they merge into a single vortex. The resulting vortex survives for ~ 75 kyr.

The inner and outer edges of the gap carved by the planet also become unstable to the RWI for a short period of time. At the outer edge, this results in the growth of the $m = 3$ mode and the subsequent formation of three vortices, at around 22 kyr, while at the inner edge a single vortex forms. By ~ 150 kyr these vortices have dissipated, leaving the disc in its final, stable state, shown in the final panel.

Figure 3.7 shows the surface density of the 1 mm dust grains at the same snapshots in time as in figure 3.5. A high surface density region can be associated with the vortex structures observed in the gas, confirming the ability of the anticyclonic vortices to trap dust. As the planet grows rapidly and carves a gap in the pressure maximum, some dust remains trapped either side of its orbit, unable to drift into the vicinity of the planet to be accreted, hence two axisymmetric rings form either side of the planet’s

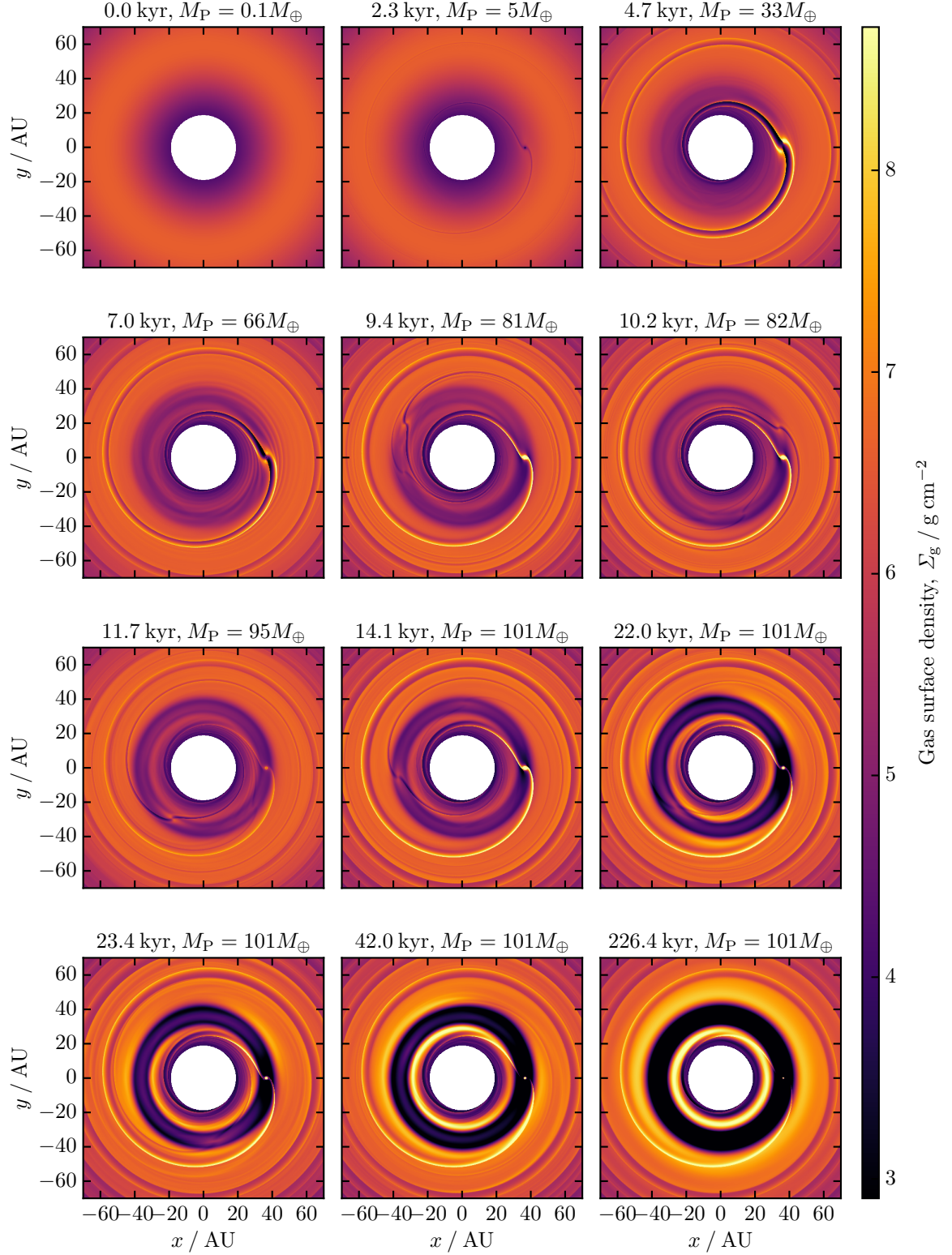


Figure 3.5: Gas surface density snapshots showing the formation of vortices as the planetary embryo accretes pebbles and heats the surrounding gas. Vortex formation can be seen at the planet location via the enhancement in surface density, and vortices can be seen to have separated from the planet after 7 kyr, again identifiable by their surface density enhancements as well as the spiral shocks they generate. The vortices traverse horseshoe orbits due to the planet’s significant mass, and the corotation torque on the vortex from the planet results in material entering the Hill sphere, causing bursts of accretion at ≈ 10.5 kyr and ≈ 12 kyr.

3.4. Results

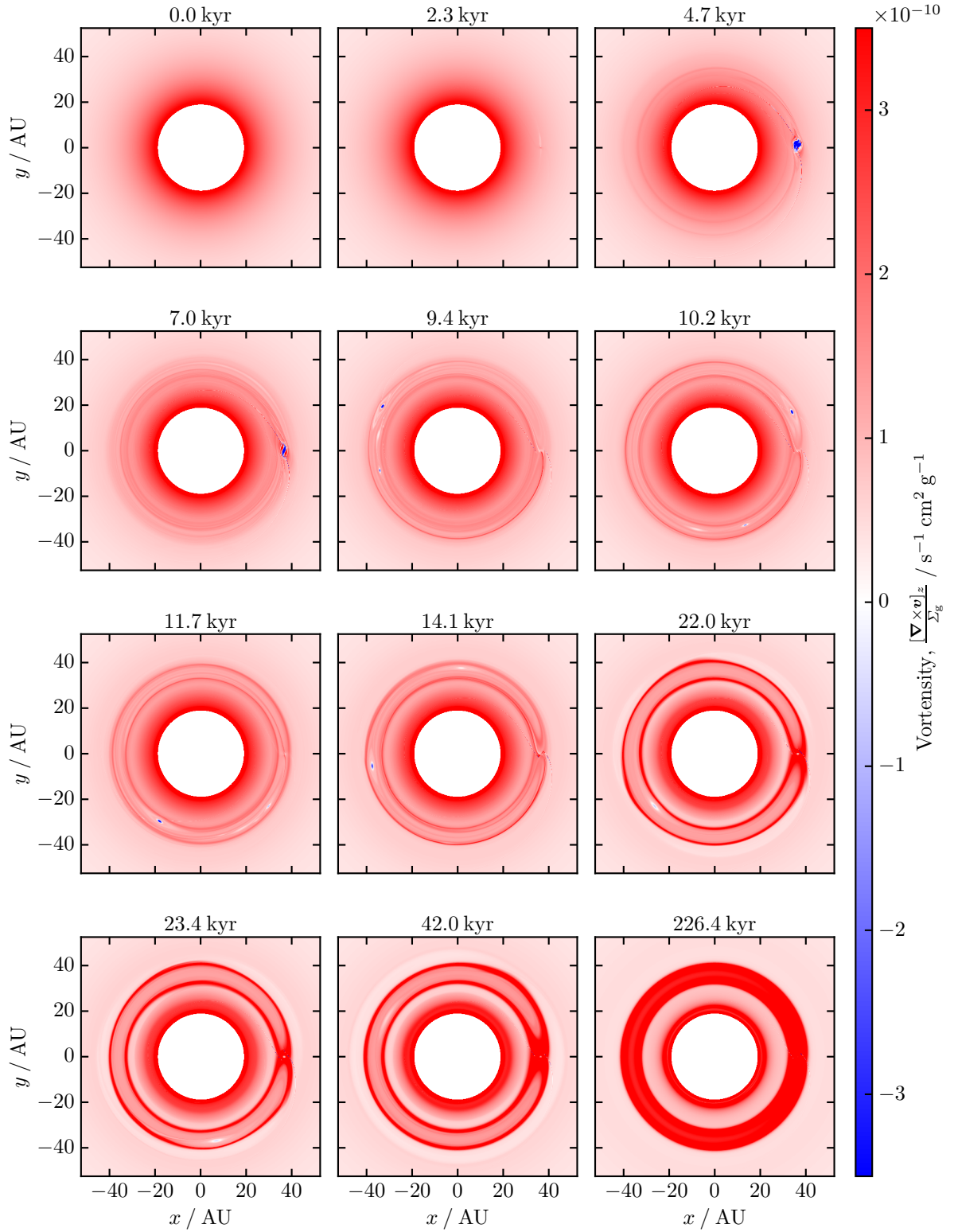


Figure 3.6: Vortensity snapshots for the simulation including the accretion luminosity of the planet. A region of negative vortensity, indicating the anticyclonic nature of the vortex, can be seen to form at the location of the planet before traversing horseshoe orbits. Two smaller, weaker vortices can be identified from the snapshot at 9.4 kyr onward.

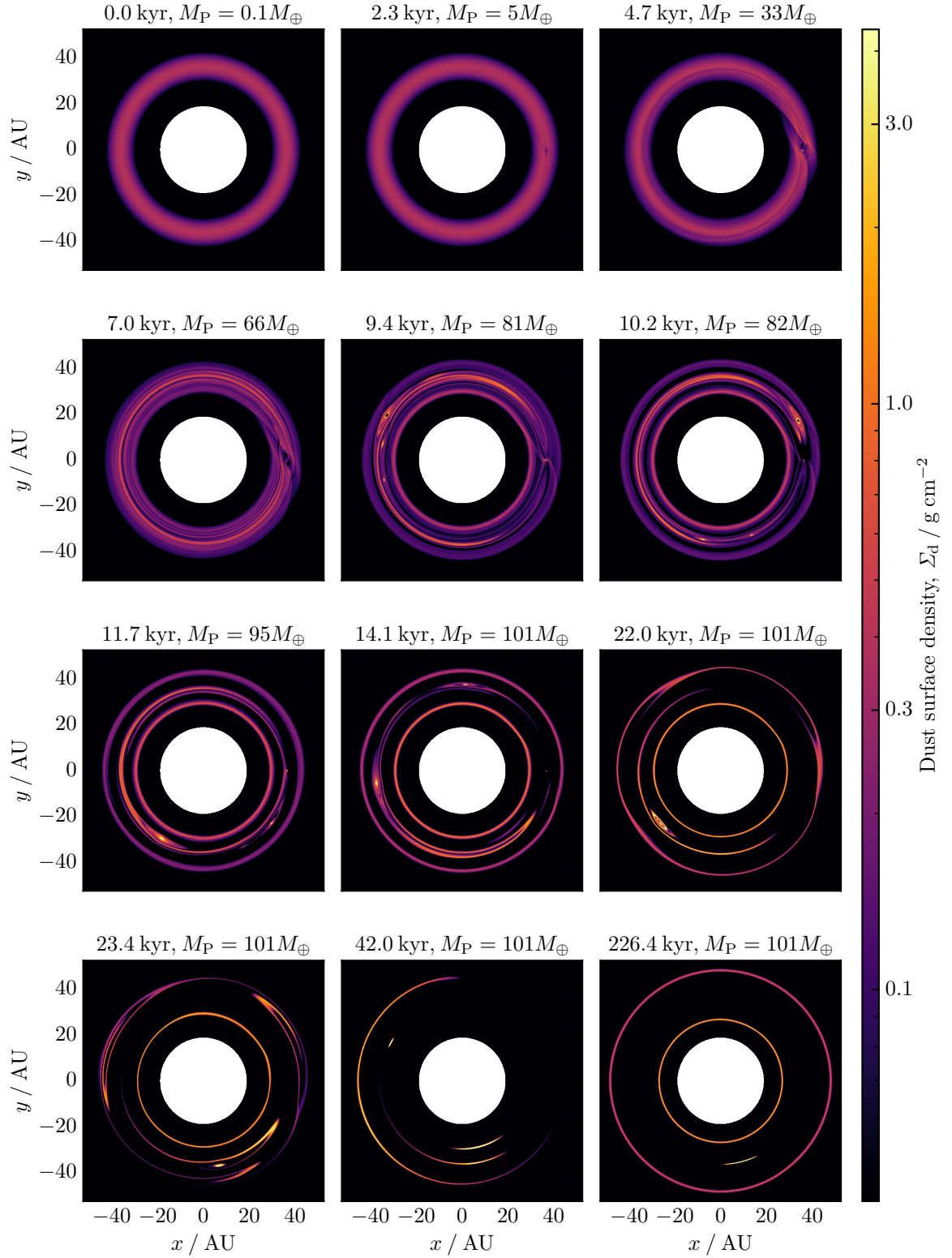


Figure 3.7: Surface density of 1 mm dust for the simulation including the accretion luminosity of the planet. Initially, the shock through the vortex reduces the pressure maximum at its centre, which prevents dust from directly accumulating there. Once the vortensity has reduced, however, dust can drift into the centre. Dust from the initial ring remains trapped in the pressure maxima either side of planet-carved gap, as well as in the vortices. Once the vortices have dissipated, dust resides close to L_5 Lagrange point.

3.4. Results

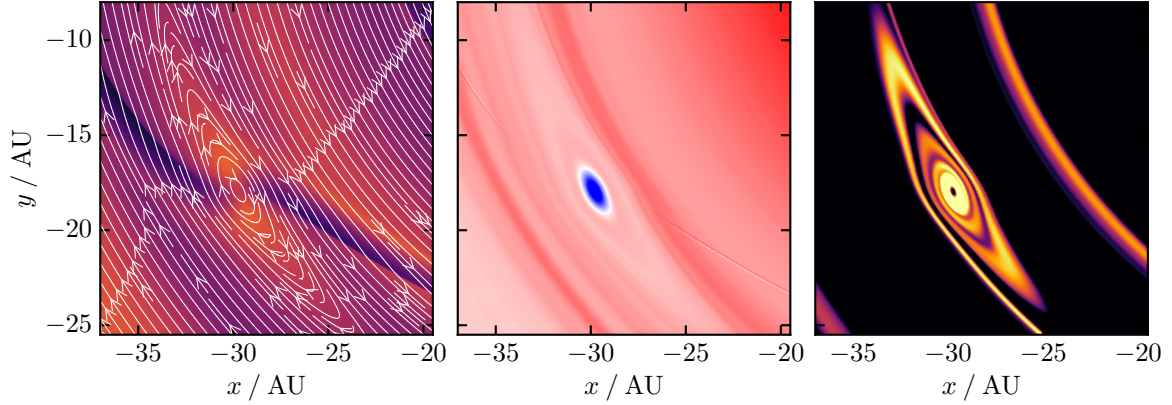


Figure 3.8: Zoom-in on the vortex at ≈ 12 kyr, showing the gas surface density with gas streamlines (left), vortensity (centre) and 3 mm dust surface density (right). The colour scales are the same as in figures 3.5, 3.6 and 3.7 respectively. The radial extent of the vortex is approximately 2.5 AU, corresponding to $4/3H_g$.

orbit. As the vortex begins to dissipate, the dust it releases back into the co-orbital region begins to accumulate around L_5 – the dust is confined to horseshoe orbits due to the planetary potential, and thus cannot enter the Hill sphere to be accreted. The snapshots between 22 and 42 kyr show how the vortex interacts with the dust that settles around L_5 : the vortex attracts this dust to its pressure maximum, but the dust does not have enough time to become trapped in the vortex, and is instead dragged along by it, forming the thin, azimuthally-extended features that can be seen in these later snapshots.

Once the gap edges become RWI-unstable, the axisymmetric dust rings are affected by the vortices that form. The snapshot at 22 kyr shows the dust beginning to concentrate into the $m = 3$ vortex pattern in the outer ring. This results in clear asymmetries in the $1\ \mu\text{m}$ and 1 mm dust, but not in the 3 mm and 1 cm dust, indicating that these vortices are weak (a result confirmed by inspection of the vortensity) and so only the dust most strongly coupled to gas shows asymmetries. Once all the vortices have dissipated, dust settles at L_5 , as was found in the simulation without the thermal feedback. The final state of the disc is therefore two dust rings and dust at L_5 , where the masses remaining in these three distinct substructures are 13.3 , 20.9 and $15.1M_\oplus$ respectively, summed across all four dust species.

Figure 3.8 shows the gas surface density, gas velocity streamlines, vortensity and dust surface density at the location of the vortex ≈ 12 kyr into the simulation. The streamlines confirm the anticyclonic nature of the vortex, with the sense of local rotation in opposition to the global flow around the star, and the gas surface density shows that it is launching spiral density waves. The closed streamlines around the vortex span ≈ 2.5 AU in radial extent, equivalent to $\approx 4/3H_g$. If the radial extent of a vortex exceeds $2/3H_g$ from its centre, the supersonic velocity shear in a Keplerian profile will prevent the vortex from expanding (Barge & Sommeria, 1995), thus this measured width indicates that the vortex has grown to its largest possible radial extent.

A comparison of the planet mass as a function of time for the simulations with and without the thermal feedback is shown in figure 3.9, illustrating the impact of the accretion luminosity on the growth of the planet. Also shown is the ratio of the luminosity radius, R_L , to the Hill radius, R_H . Vortex formation requires the thermal structure of the disc to be disrupted beyond the Hill sphere. The planet growth rates begin to diverge at around 3.5 kyr, corresponding to the time at which the luminosity radius exceeds the Hill radius by a factor ~ 3 . The accretion rate falls off after 7 kyr, once the vortex has separated from the planet location, trapped dust and therefore dramatically reduced the solid material available for accretion onto the planet. There is a second burst in accretion at ≈ 10 kyr, which results from the corotation torque on the vortex from the planet causing a significant amount of material to

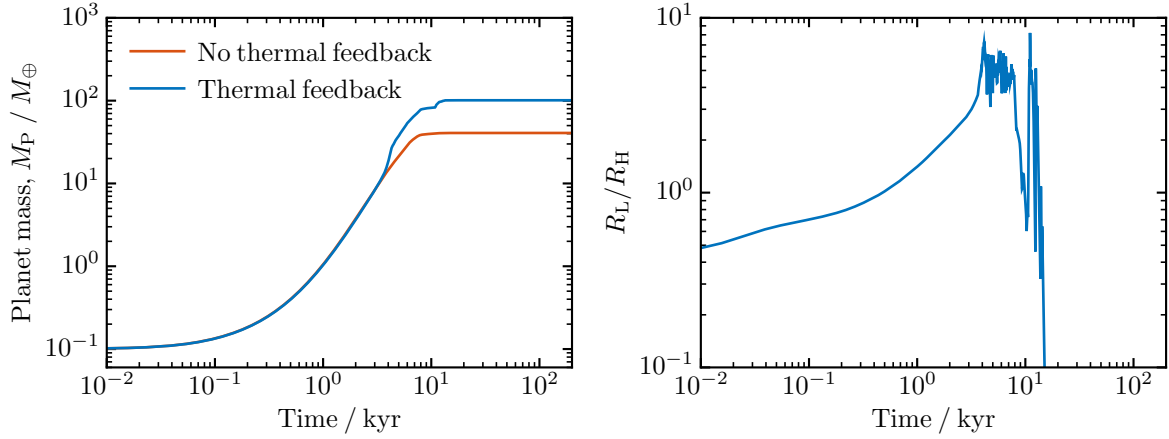


Figure 3.9: The left-hand panel shows the mass of the accreting planet as a function of time for the simulations with and without the thermal feedback. In both cases the embryo initially grows slowly, but the accretion rate increases as its Hill sphere grows. With the inclusion of the thermal feedback, the high accretion rate leads to vortex formation at the planet location at ≈ 3.5 kyr. This initially enhances the accretion rate, as the vortex draws dust toward the planet, but once the vortex breaks away from the planet it quickly traps dust, shutting off accretion onto the planet. The right-hand panel shows the ratio of the luminosity radius to the Hill radius. Once this ratio exceeds ~ 3 , the thermal feedback is sufficient to generate a vortex that can significantly affect the accretion rate. This ratio drops off once accretion ceases, staying below $\sim 10^{-3}$ after 20 kyr.

enter the planet’s Hill sphere. The planet in the simulation with thermal feedback eventually reaches $\approx 100M_\oplus$, more than twice that achieved in the simulation without, which is a naively surprising result, as one might have expected dust trapping by the vortex to suppress the accretion rate. The implication of this result in terms of the physics of pebble accretion and vortex formation is discussed in §3.5.

Figure 3.10 shows the mass of solids in the disc over the duration of the simulation. The largest grain size populations, which have Stokes numbers closest to unity, lose the largest fraction of their mass since they have the highest pebble accretion rate onto the planet. Similarly, since they are the most efficiently trapped in the vortex, a significant amount of dust remains once the dust trapping in the vortex shuts off pebble accretion. This results in an inversion of the grain size distribution between the millimetre to centimetre-sized pebbles, with most mass in the 1 mm population. Clearly visible is the point at which the vortex separates from the planet, marked by the sudden transition to a period of constant mass in the centimetre grain population, until the torque between the planet and vortex during a horseshoe turn leads to a deposition of mass in the planet’s Hill sphere, and thus a short period of further accretion. As expected, there is little change in mass of the micron-sized dust population, as their accretion rate is negligible. The total dust mass remaining in the disc is $\approx 50M_\oplus$, which is trapped in the new pressure maxima that form either side of the planet’s orbit, and in the co-orbital region.

High planet mass test

Given the results obtained in the previous simulations, I performed an additional simulation with the same initial conditions, but with the planet instead initialised with a mass of $100M_\oplus$. This planet is not undergoing pebble accretion, to further discern which behaviour results solely from the accretion luminosity of the planet rather than through gravitational planet–disc interactions. Figure 3.11 shows snapshots of the gas surface density from this simulation. The planet interacts gravitationally with the disc and quickly creates a deep gap, the edges of which become RWI-unstable, leading to the formation of a vortex either side of the gap within 10 kyr. The vortices that form are weak, being narrow in radius and wide in azimuthal extent. As was the case in §3.4.1, the micron and millimetre sized grains also form

3.4. Results

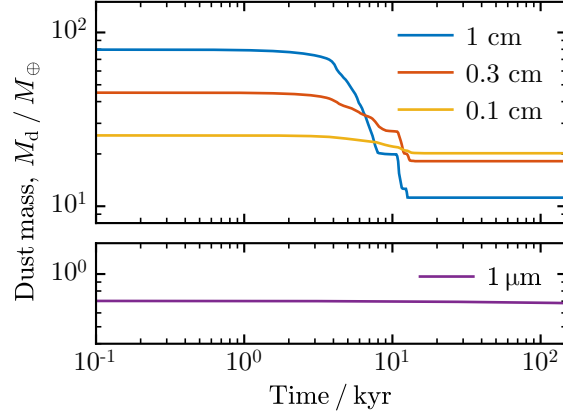


Figure 3.10: Mass remaining in each dust population as a function of time. The 1 cm population loses the largest fraction of its mass since they have the highest pebble accretion rate, and the dust size distribution is inverted once pebble accretion ceases. The micron sized dust has been shown on a separate panel for clarity. The total dust mass remaining is $50M_\oplus$, which is trapped in the dust rings either side of the planet’s orbit, and in the co-orbital region. The steep drops in mass after 10 kyr are due to the corotation torque between the vortex and planet causing dust to be released in to the planet’s Hill sphere.

azimuthally asymmetric substructure due to the vortices, circulating with the gas, but the centimetre grains, being less tightly coupled to the gas, largely maintain a ring structure due to the vortex not being strong enough to trap the dust. Again, dust in the co-orbital region accumulates in the drag-modified L_5 Lagrange point, but not in L_4 .

Resolution tests

I carried out additional simulations at various resolutions to check that the pebble accretion onto the planet had converged, i.e. that my pebble accretion algorithm was not impacted by the grid resolution. Figure 3.12 compares the pebble accretion rate of centimetre-sized dust grains onto the planetary embryo over the first 7 kyr of the simulation – which includes the initial embryo growth and vortex formation – at half and double the original resolution. The accretion onto the embryo from the disc is resolved at half resolution, but once the vortex forms at around 4 kyr, the accumulation of dust in the vortex cannot be sufficiently resolved: concentration of dust in the pressure maximum at the centre of the vortex is impeded only by diffusion, and with a low viscosity of $\alpha = 10^{-4}$, the half resolution case cannot resolve the pressure and density gradients upon which the dust dynamics depend. Since the vortex forms at the planet location and the pebble accretion rate depends on the surface density of solids within its Hill sphere, the unresolved accumulation of dust in the vortex restricts the accretion rate onto the planet.

The pebble accretion rate at double the original resolution is qualitatively similar during vortex formation. Since differences in the accretion rate lead to differences in the planet mass, which in turn determines the subsequent accretion rate, small differences are likely to be amplified. However, given the absolute differences in the accretion rate are small, I considered the simulation at the fiducial resolution to be a reasonable compromise between convergence and computing time.

3.4.2 Radiative transfer simulations

In order to understand how the dust distribution calculated from the hydrodynamics simulation might present in observations, I performed Monte Carlo radiative transfer simulations using RADMC-3D (Dullemond et al., 2012). The dust opacities were calculated using the model presented in Birnstiel et al. (2018): the dust grains are taken to be of a mixture of “astronomical silicates” (Draine, 2003), troilite

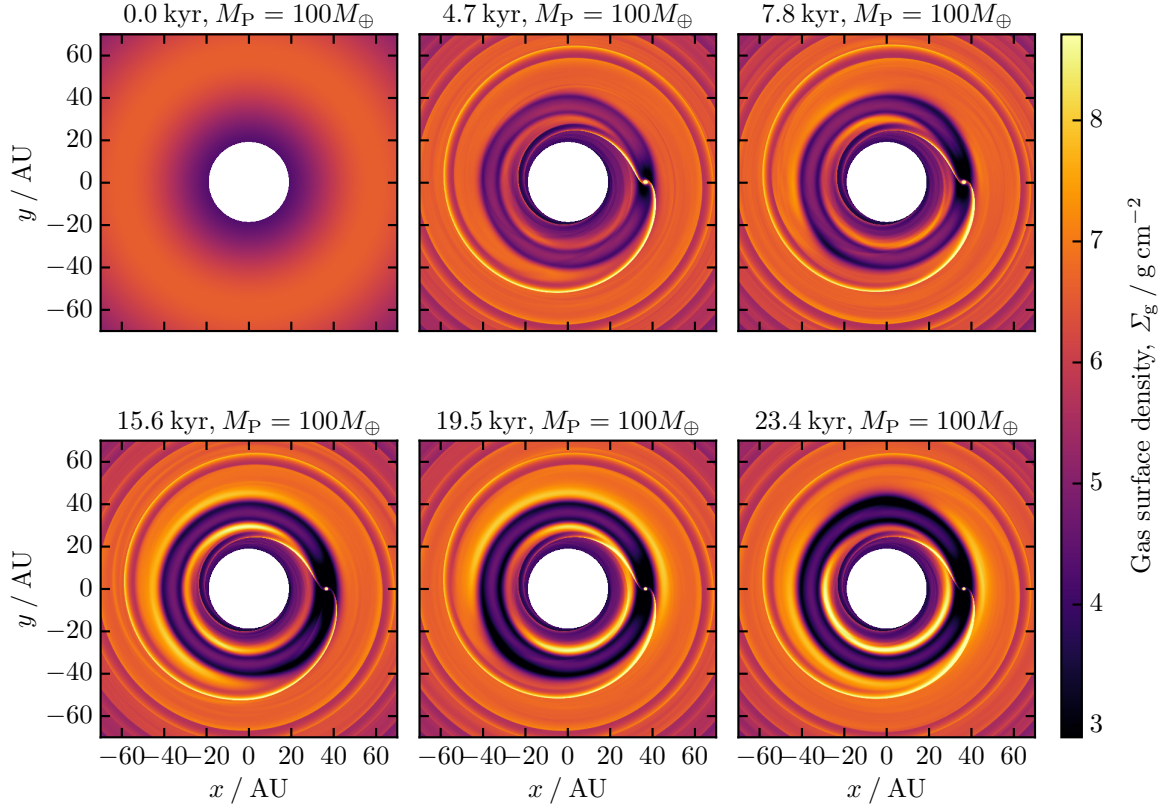


Figure 3.11: Gas surface density snapshots for the simulation in which the planet is initialised at its final mass and no accretion takes place, thus there is no thermal feedback on the disc. The $100M_{\oplus}$ planet carves a deep gap in the disc, the edges of which quickly become RWI-unstable, leading to an azimuthal asymmetry either side of the gap.

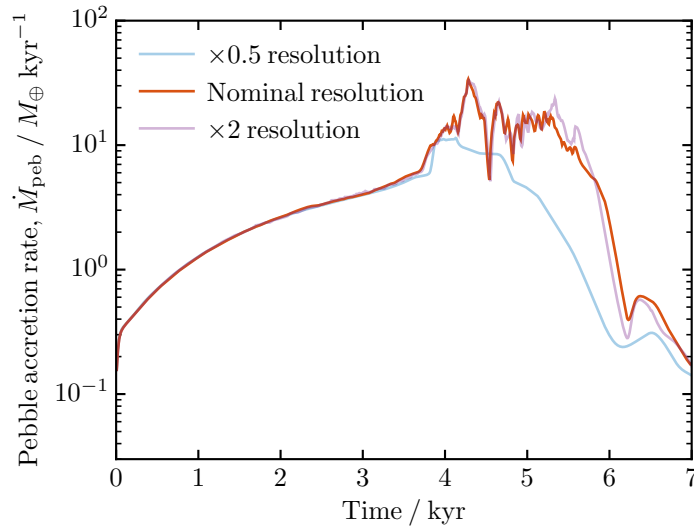


Figure 3.12: Comparison of the pebble accretion rate onto the planetary embryo as a function of time for half and double the original resolution over the first 7 kyr of the simulation. Pebble accretion from the disc is resolved at half the original resolution, but once the vortex forms at around 3.5 kyr, the lowest resolution case cannot accurately resolve the density gradients around the embryo, reducing dust flow into the Hill sphere and therefore impacting the pebble accretion rate.

3.4. Results

and “organics” (Henning & Stognienko, 1996) and water ice (Warren & Brandt, 2008), with a water ice mass fraction of 0.2. The grains are taken to have zero porosity. The optical constants for this mixture were obtained from their constituents’ via the Bruggeman mixing rule (Bruggeman, 1935), which are then used to calculate the absorption and scattering opacities as a function of wavelength via Draine’s implementation of BHMIE. The calculated opacities span the wavelength range $1\ \mu\text{m}$ to $10\ \text{cm}$.

I only considered the region of the disc between 22 and 52 AU, since this contains essentially all of the millimetre to centimetre sized dust. This region is initially covered by 970 cells in the radial direction, but I reduced the resolution over a few narrow annuli where the dust surface density is approximately constant; this reduces the number of radial cells to 435. I reduced the azimuthal resolution by a factor of 6, to 1280 cells. I then performed 2D cubic spline interpolation over the surface density onto this reduced resolution grid. I used 145 cells in the polar direction, with 72 logarithmically spaced between 10^{-3} and $\pi/6$ both above and below the midplane at $\theta = \pi/2$.

Since the gas pressure structure of the disc prevents the radial drift of dust, and the millimetre and centimetre sized grains have short vertical settling timescales, I can assume that each dust species has reached equilibrium between diffusion and settling in the vertical direction. I therefore calculated their three-dimensional density distributions as (e.g. Takeuchi & Lin, 2002)

$$\rho_d = \rho_{d,\text{mid}} \exp \left[-\frac{z^2}{2H_g^2} - \frac{\text{Sc St}_{\text{mid}}}{\alpha} \left(\exp \frac{z^2}{2H_g^2} - 1 \right) \right]. \quad (3.33)$$

The dust density at the disc midplane was calculated via

$$\rho_{d,\text{mid}} = \Sigma_d \left\{ \int_{-\infty}^{\infty} \exp \left[-\frac{z^2}{2H_g^2} - \frac{\text{Sc St}_{\text{mid}}}{\alpha} \left(\exp \frac{z^2}{2H_g^2} - 1 \right) \right] dz \right\}^{-1} \quad (3.34)$$

which I evaluated numerically using quadrature. Again I set $\text{Sc} = 1$, i.e. I am assuming that the vertical and radial diffusion coefficients are equal.

The properties of the central star are chosen to be those of a Herbig Ae/Be star. I used an effective stellar temperature of 9000 K, a mass of $2M_\odot$ and radius of $2.5R_\odot$. I used a template spectrum for an A0V star from Castelli & Kurucz (2003), with a wavelength range spanning $1\ \mu\text{m}$ to $10\ \text{cm}$ and sample this spectrum using 2×10^{10} photon packets. The disc was taken to be at a distance of 140 pc for consistency with sources observed in the Taurus star-forming region, and the resulting image was convolved with a Gaussian beam of FWHM 40 mas ($\approx 5\ \text{AU}$ at 140 pc), such that the resolution is comparable to that achieved in the DSHARP survey (Andrews et al., 2018). I took a representative noise floor from the DSHARP survey of $15\ \mu\text{Jy beam}^{-1}$ and removed any signal below this value from the images. I generated images of the disc face-on, at observation wavelengths of 0.88 mm and 1.25 mm for comparison with ALMA band 7 and 6 observations respectively.

Figures 3.13 and 3.14 show the images generated at observation wavelengths of 0.88 mm and 1.25 mm respectively, from simulation snapshots at 11.7 kyr, 14.1 kyr, 22 kyr and 226 kyr. In the early images, while dust is still present in the co-orbital region, the brightness enhancements resulting from dust trapped in vortices are resolved, as is the lack of emission from the vicinity of the planet. At 22 kyr, there is significant emission from the dust that has been dragged away from L_5 by the vortex, while the compact vortex structure stands out. Also visible is the $m = 3$ vortex pattern in the outer ring, resulting from the RWI.

The general morphology resulting from the band 6 and 7 synthetic observations is similar, though there are flux ratio variations. I calculated the spectral index, $\alpha_\nu = d \log I_\nu / d \log \nu$, over the disc between these two wavelengths. I convolved each of the images with the Gaussian beam before the spectral index calculation. As shown in figure 3.15 the spectral index is lowest in the dust rings and at the vortex

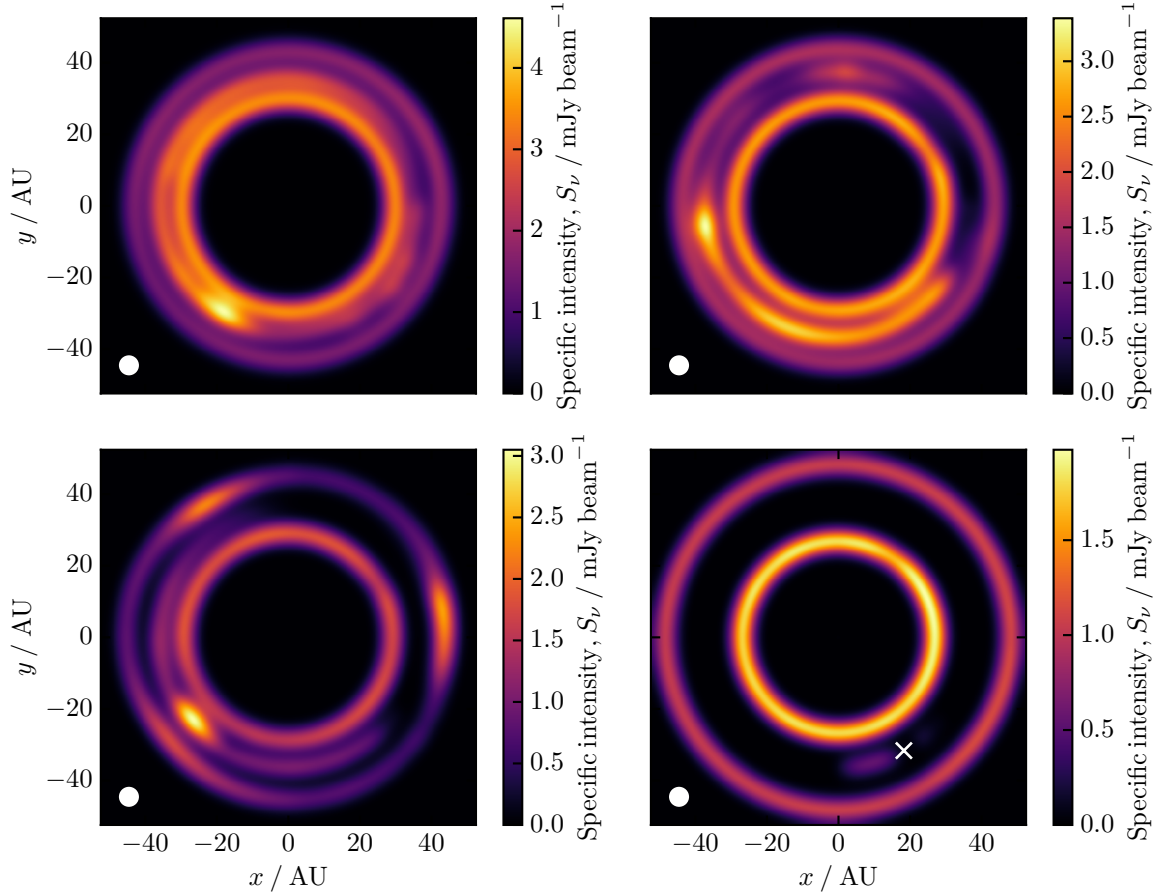


Figure 3.13: Synthetic observations of the disc at 11.7, 14.1, 22 and 226 kyr, using an observation wavelength of 0.88 mm, convolved with a 5 AU FWHM Gaussian beam (shown in the bottom-left of each panel). Flux below my adopted instrumental noise floor of $15 \mu\text{Jy beam}^{-1}$ has been removed. In each image a brightness enhancement can be associated with a dust surface density enhancement due to trapping in vortices. In the final image, the dust residing close to L_5 (marked with a white cross) can be seen, with less flux from the largest dust to the right of L_5 compared to the millimetre sized dust, which has settled to the left of L_5 .

centres, indicating the concentration of the largest dust grains there. Similarly, the spectral index is low around L_5 , where the dust grains are trapped. Due to the fact that different sized dust grains settle at different locations around L_5 , the region of low spectral index has a much wider azimuthal extent than the vortex, which essentially just presents as a single beam-sized region of low spectral index. Note that the separation between the two regions of emission around L_5 in the final image arises due to the use of discrete grain sizes in my hydrodynamics simulations; a continuum of grain sizes would produce a continuous crescent of emission.

3.5 Discussion

Disc morphology

I set out to explore a new link between planet formation and non-axisymmetric substructure in protoplanetary discs. My simulations have shown that a planetary embryo in an axisymmetric dust trap can accrete pebbles at a high enough rate to trigger the formation of vortices and thus create non-axisymmetric substructure in the disc. The vortex that forms spans $\approx 15 - 20$ AU in azimuthal extent at 35 AU from the central star. Since dust drifts towards the pressure maximum located at the centre

3.5. Discussion

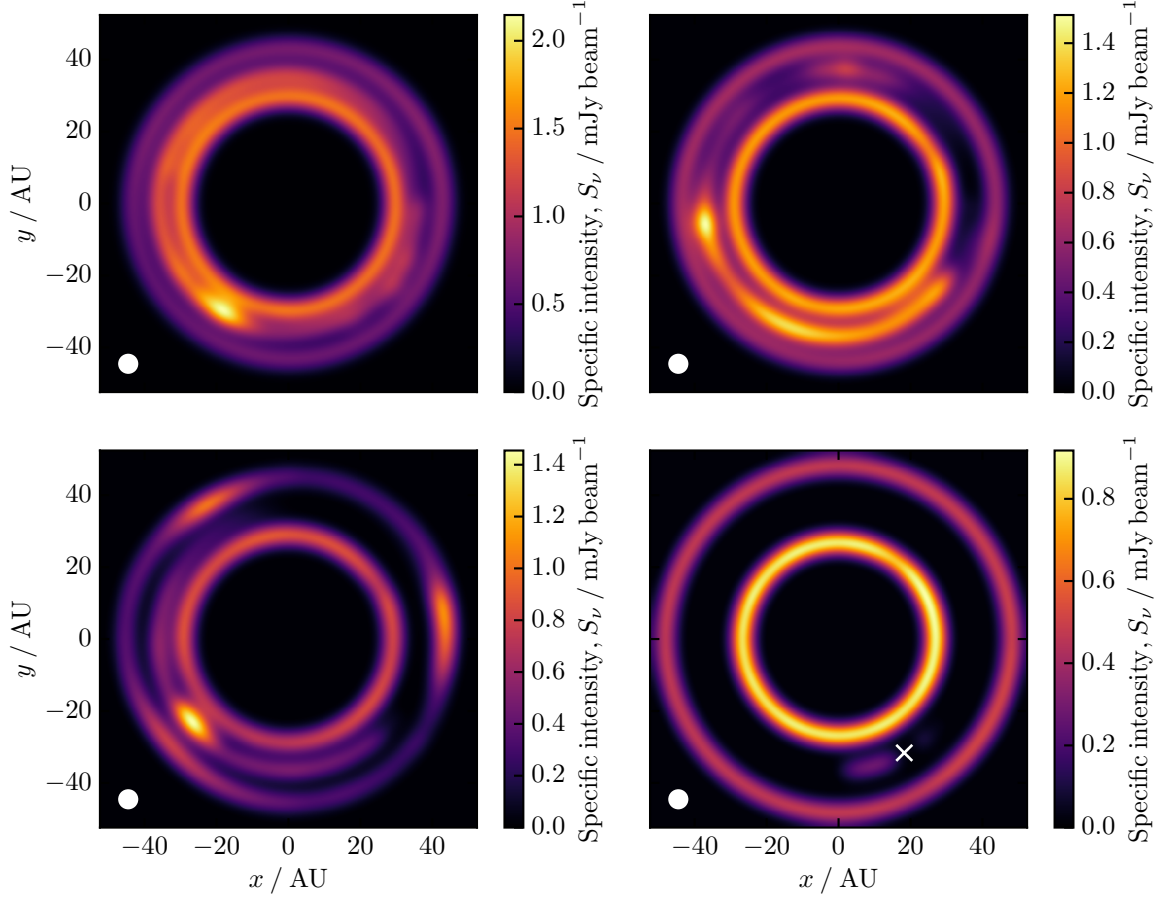


Figure 3.14: Synthetic observations of the disc at 11.7, 14.1, 22 and 226 kyr, using an observation wavelength of 1.25 mm, convolved with a 5 AU FWHM Gaussian beam (shown in the bottom-left of each panel). Flux below my adopted instrumental noise floor of $15 \mu\text{Jy beam}^{-1}$ has been removed. In each image a brightness enhancement can be associated with a dust surface density enhancement due to trapping in vortices. In the final image, the dust residing close to L_5 (marked with a white cross) can be seen, with less flux from the largest dust to the right of L_5 compared to the millimetre sized dust, which has settled to the left of L_5 .

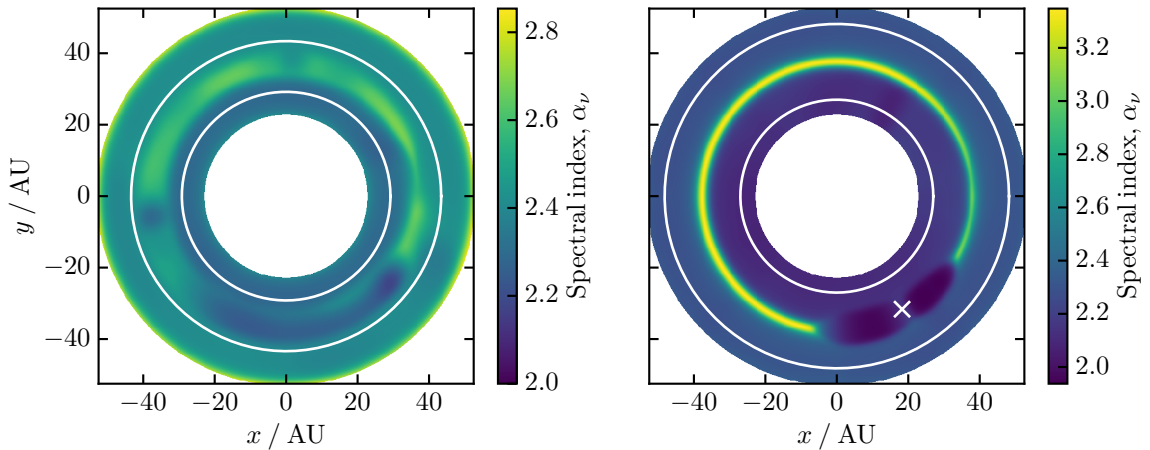


Figure 3.15: Spectral index map of the disc at 14.1 kyr (left) and 226 kyr (right), calculated between 0.88 mm and 1.25 mm. At 14.1 kyr, the low spectral index regions at around (2, 35) and (−35, −5) are due to dust trapping in vortices, while the crescent spanning (−20, −30) to (30, −20) is due to dust beginning to settle around L_5 . At 226 kyr, an extended region of low spectral index can be seen around L_5 , which is marked with a white cross. The white solid lines mark the locations of the pressure maxima.

of the vortex, the size of the non-axisymmetric enhancement in dust will differ over time as well as with grain size. As the strength of the vortex decays it releases dust and its size therefore decreases over time, while dust drift, limited only by diffusion and being dependent on grain size, sets the observable size of the dust asymmetry. My synthetic observations show that the substructure generated in my simulation creates detectable features such as rings and crescents similar to those observed in continuum images (e.g. Andrews et al., 2018), as well as features in spectral index maps (e.g. ALMA Partnership et al., 2015b; Huang et al., 2018a; Carrasco-González et al., 2019; Huang et al., 2020).

Having accreted a significant fraction of the mass from the dust trap, the planet becomes massive enough to carve a gap in the initial pressure maximum, creating new pressure maxima either side of its orbit. Dust becomes trapped in these two pressure maxima, leading to the presence of two dust rings which exist alongside the vortex that forms. The dust remaining in the disc, which primarily resides in the pressure maxima either side of the planet’s orbit, shows an inversion of the pebble size distribution, with most mass in 1 mm grains and the least in the 1 cm grain population. Given time, however, I suspect an MRN distribution could be re-established through collisions via coagulation and fragmentation.

Properties of the planet and a new phase of pebble accretion

In addition to the non-axisymmetric disc substructure produced through this mechanism, I found interesting consequences for the growth of the planet. The simulation including the accretion luminosity produces a $100M_{\oplus}$ core, while in the test case without the accretion luminosity, the planet grows to $\approx 40M_{\oplus}$. This is naively opposite to what one might have imagined, where the vortices trap dust and migrate away from the planet, preventing this dust from accreting onto the planet, resulting in a lower planet mass. However, since the vortex forms around the planet and remains co-located with it for ~ 30 orbits, it acts to significantly increase the area from which the planet can accrete before migrating away and trapping dust.

As the vortex is a pressure maximum, it acts to funnel dust towards the planet from a radial width of $\sim 4/3H_g$, rather than $\sim 2R_H$ in the case of standard pebble accretion. Thus, my simulations have identified a new, short-lived, but extreme form of vortex-mediated pebble accretion. It is the much larger feeding zone that allows the planetary core to reach such a large mass. In this phase the “vortex-assisted” pebble accretion rate can be estimated from the usual capture rate formula:

$$\dot{M} \approx 2v_{\infty}b\Sigma. \quad (3.35)$$

Adopting an impact parameter given by the maximum radial extent of the vortex, i.e. $b = (2/3)H_g$, and the velocity of a pebble approaching the vortex $v_{\infty}(b) \approx (3/2)\Omega_K b = \Omega_K H_g$, yields an accretion rate of

$$\dot{M}_{\text{peb,vort}} \approx \frac{4}{3}\Omega_K H_g^2 \Sigma_{\text{peb}}, \quad (3.36)$$

which is actually independent of the planet mass, and hence time. This explains the sudden increase in, but approximately constant (when averaged over several orbital periods) pebble accretion rate seen in the simulations. Taking the planet mass as $\approx 15M_{\oplus}$ at 4 kyr (figure 3.9), the change in effective accretion area from $2R_H^2$ to $4/3H_g^2$ should result in a jump in accretion rate of ~ 4 . This is illustrated in figure 3.16, where the theoretical classical pebble accretion rate and this theoretical vortex-assisted pebble accretion rate are plotted. The classical pebble accretion rate is obtained by numerically integrating equation (1.64), while the theoretical vortex-assisted accretion rate is obtained from equation (3.36), using the average Σ_d over the period where the vortex is co-located with the planet. This is approximately in agreement with the jump in accretion rate from $\sim 6M_{\oplus} \text{ kyr}^{-1}$ at 4 kyr to a time averaged rate of $\sim 20M_{\oplus} \text{ kyr}^{-1}$ seen during the phase where the vortex is co-located with the planet between 3.5 and

3.5. Discussion

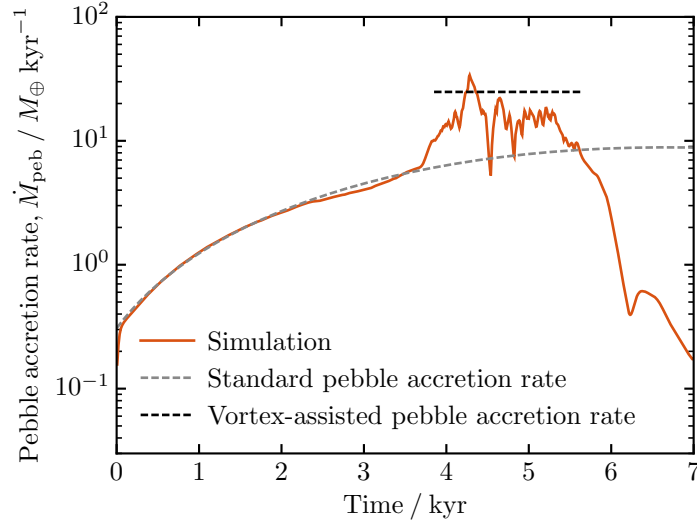


Figure 3.16: Comparison of the pebble accretion rate in the simulation including thermal feedback with the theoretical pebble accretion rate according to the standard pebble accretion scenario, as well as the theoretical accretion rate during the vortex-assisted pebble accretion phase, time-averaged over the period when the vortex is co-located with the planet.

6 kyr. Thus, I suspect this new form of vortex-assisted pebble accretion could play an important role in planet formation inside dust traps, at the extreme end of planet formation.

My simulations, both with and without thermal feedback, result in higher masses than typically found in previous models. This is because the planets form inside a dust trap where there is already a large supply of dust. Standard models of pebble accretion, which typically use a disc with a monotonically decreasing surface density profile, find a maximum core mass that can be formed: the pebble isolation mass. This is the mass at which the core can prevent further drift of pebbles into its orbit; once it is massive enough to carve a gap, the resulting pressure bump exterior to its orbit prevents further inward drift of pebbles (Morbidelli & Nesvorný, 2012; Lambrechts et al., 2014; Bitsch et al., 2018). It is worth noting that the scenario in which this standard pebble isolation mass arises is not directly applicable here, since in a dust trap the planet accretes pebbles within the proximity of its own orbit rather than pebbles continuously drifting in from the outer disc, so halting pebble drift from the outer disc has no impact on the planet’s growth, hence why it can grow so large.

This was also recently demonstrated by Sándor & Regály (2021); they performed 2D hydrodynamics simulations of a core undergoing pebble accretion in a pressure maximum, using the Kley (1999) prescription for gas accretion as their pebble accretion scheme. Neglecting accretion luminosity, they found a rapid pebble accretion rate and core growth up to final masses between $\sim 30 - 50M_{\oplus}$, depending on their prescribed pebble accretion efficiency and the Stokes number used. In their control case without a pressure maximum, cores reached $\approx 20M_{\oplus}$, in agreement with Lambrechts et al. (2014).

While these solid masses seem large, there is strong evidence for large solid masses in the observed exoplanet population. Thermal and structural modelling of transiting giant planets by Thorngren et al. (2016) suggests metal masses $M_Z \gtrsim 100M_{\oplus}$ for $\approx 20\%$ (9/47) of their sample of Jupiter mass planets. Assuming that these planets formed via core accretion, the authors suggest that these may be giant planets with metal-enriched atmospheres, though their modelling is not constraining enough to indicate where the solids are partitioned within the planet. Thus, formation of a massive planetary core in a dust trap, followed by subsequent gas accretion (the core masses in my simulations are well above the critical core mass for runaway accretion, once the pebble accretion has subsided, e.g. Rafikov 2006) and migration is a viable formation scenario of the observed metal enriched giant planets.

Given the high dust surface densities achieved, it is interesting to consider the possibility of direct gravitational collapse into planetesimals. This requires the local dust density to exceed the Roche density, $\rho_R = 3.5M_*/a^3$. Given that the planet and vortex orbit at around 35 AU from the $2M_\odot$ star, this requires the dust surface density to exceed $\rho_R \sim 10^{-10} \text{ g cm}^{-3}$. Calculating the dust density as $\rho_d = \Sigma_d/H_d$, with $\Sigma_d \sim 5 \text{ g cm}^{-2}$ at the vortex core, gives a maximum dust density of $2 \times 10^{-11} \text{ g cm}^{-3}$, which is therefore insufficient to trigger gravitational collapse.

Final trapping of dust at Lagrange points

The simulations presented above, both with and without thermal feedback, show that once the vortex has dissipated, the dust it had accumulated gathers around the classical L_5 Lagrange point. Dust trapping at L_5 has been seen in previous work: for example, Rodenkirch et al. (2021) showed that dust trapped in L_5 may be responsible for non-axisymmetric substructure in observed in HD 163296, which shows a crescent-shaped region of emission in the gap between a dust ring and the inner disc. Their simulations of planets embedded in a disc resulted in a crescent-shaped feature between two rings, similar to the final state of my simulation, thus an asymmetry in the gap between two rings resulting from trapping in L_5 is not unique to my model.

The reason why dust is trapped in L_5 but not in L_4 has been studied analytically within the restricted three-body problem. Murray (1994) showed that the addition of gas drag modifies L_4 , making it an unstable equilibrium point. This is confirmed in my simulations by analysing the forces on dust grains in the vicinity of L_4 . Additionally, Murray (1994) showed that gas drag modifies the location of the stable L_5 point, where the position of L_5 relative to the planet depends on the stopping time of the dust grains, with less strongly coupled dust residing closer to the planet. My simulations generally confirm this trend; however, there are some deviations from the predictions, in that the millimetre-sized dust grains settle behind L_5 , while only the centimetre-sized grains settle ahead of L_5 , closer to the planet, as expected. This discrepancy is attributed to the fact that Murray (1994) did not use an exact gas streamline structure, but instead adopted a simple model.

This dust-size-dependent position of the drag-modified L_5 location means that dust trapped in L_5 and a vortex have a potential observational difference. In the case of the vortex, the peak dust density is always achieved in the same place – the vortex core – irrespective of particle size, resulting in a strongly peaked dust density and therefore a single location where the spectral index drops. In the case of L_5 , however, the particle distribution is spread over a much wider range in azimuth, resulting in a broad region where the spectral index drops, with a gradient such that the spectral index decreases towards the planet (see figure 3.15). This is a potentially important avenue in observationally distinguishing the nature of the observed non-axisymmetric structure.

Impact of simplifications

Given that I aimed to isolate the effects of thermal feedback, there are some physical processes that I have neglected from my investigation, the possible implications of which I now discuss. The first is that I have ignored the feedback on the gas from the drag force on the dust. The peak dust-to-gas ratio $\Sigma_d/\Sigma_g \approx 0.8$ in the initial condition, thus the back-reaction on the gas will be non-negligible here. The effect of this would be to widen the gas pressure maximum and thus weaken the dust trap, which would reduce the dust surface density and hence the initial pebble accretion rate onto the planetary embryo. The feedback on the gas will also become important in the vortices where large amounts of dust become trapped. Fu et al. (2014) found that the back-reaction can reduce the lifetime of the vortex, and a similar result was found by Raettig et al. (2015). However, follow-up work by Lyra et al. (2018) showed that in 3D simulations the vortex lifetime is unaffected by their accumulation of dust. Regardless of whether

3.5. Discussion

the feedback on the gas affects vortex lifetime, it is likely to impact the embryo's growth as described above, and this *does* impact vortex formation, thus this is something that should be addressed in future work. As the simulations I have presented here are 2D, it will be necessary to perform 3D simulations to verify these results regardless, once the back-reaction is taken into account.

Since the dust surface density becomes so large in the vortices, it may also be necessary to include self-gravity and see whether it is possible to achieve gravitational collapse of the solids inside a vortex. For example, Lyra et al. (2009) presented surface densities capable of achieving gravitational collapse inside vortices. Additionally, Lin & Pierens (2018) showed that self-gravity can modify vorticity. While the possibility of gravitational collapse within the vortex is of secondary interest to my investigation, the effect of self-gravity on the lifetime of vortices generated in this way is important to understand.

It is also important to consider the manner in which I remove dust from the embryo's Hill sphere. Due to the treatment of the dust as a fluid, I cannot remove individual dust grains, but instead remove an amount of dust from each cell within the Hill sphere in proportion to the mass currently in that cell. While I have established that my pebble accretion algorithm provides the correct accretion rate, the dynamics of the dust fluid parcels inside the Hill sphere are not explicitly taken into account. I have demonstrated that this is not a problem for standard pebble accretion, but this simplification might have consequences in a situation where the velocity structure differs from that of a standard star-planet system, such as in the presence of a vortex. For example, if a vortex approaches or passes the planet and any dust trapped within the closed streamlines of the vortex enters the planet's Hill sphere, the planet will accrete some fraction of that dust, which may not necessarily represent a physically realistic scenario given the dust dynamics in the vortex. Essentially, my algorithm cannot capture the competition between pebble accretion onto the embryo and drift towards the centre of the vortex. However, a vortex passing through the embryo's Hill sphere is a rare occurrence since they are typically seen to traverse horseshoe orbits in the simulation presented here, so this simplification should have little overall impact. Indeed, the torque between the planet and vortex causes dust to enter the planet's Hill sphere, but the fate of this dust is determined by the equation of motion – I do not force it all to be accreted by virtue of it entering the planet's Hill sphere. Furthermore, I can be confident that this simplification does not invalidate my results during the initial vortex formation phase, where the vortex forms at the planet location, since figure 2.5 shows that the pebble accretion rate during this phase is set by how the embryo regulates the mass in its Hill sphere.

An additional simplification I have made in my simulation is the way in which I set the gas temperature in the vicinity of the accreting planet. The temperature profile instantaneously adjusts to the accretion luminosity, which in turn is instantaneously set by the accretion rate. While the outer regions of the Hill sphere and the disc in the simulation are optically thin, the region where the thermal energy resulting from accretion is deposited, deep in the protoplanet's envelope, is likely to be optically thick, resulting in some time-lag for changes in the accretion rates to be translated into temperature changes. For example, in my simulation, once accretion shuts off, the vicinity of the disc around the planet instantaneously cools and the gas surrounding it contracts. While this effect is physical, it is unlikely to be so extreme in a real disc due to the finite time taken for the envelope to cool. In future, simulations should include an improved calculation of the thermal structure around the planet.

Another potential issue that should be investigated is the thermal ablation of pebbles during their accretion onto the embryo. At the highest luminosities the temperature structure in the envelope, near the core, may exceed the sublimation temperature of the pebbles. If this is the case, (i) the pebbles may not reach the core and (ii) may release their thermal energy at a higher location in the planet's potential; both effects could reduce the accretion luminosity as well as change the envelope's composition, increasing its metallicity (e.g. Hori & Ikoma, 2011), which in turn would reduce the local thermal feedback on the disc. However, this effect would not prevent the pebbles from actually accreting onto the planet in the

first place.

Another effect that I have neglected from my simulations is the migration of the planet, caused by the exchange of its angular momentum with the disc (e.g. Goldreich & Tremaine, 1979). The type I migration rate of planets with masses less than a few tens of an Earth mass scales linearly with planet mass (e.g. Ward, 1997). However, these migration timescales are still longer than the planet growth timescales I find in my simulations.

Guilera & Sándor (2017) performed 1D simulations of a disc containing axisymmetric pressure maxima and a planetary embryo of $0.01M_{\oplus}$, free to migrate as it grew via pebble accretion. The difference between their calculated pressure maxima and zero-torque locations were minimal; as such they found that embryos which formed beyond 2 AU migrated towards the pressure maximum, at which point pebble accretion ensued; at even further separations the planet’s migration was limited ($\Delta a \lesssim 1$ AU) and pebble accretion was still rapid without the planet reaching the pressure maximum. In each case their core mass exceeded $50M_{\oplus}$ before the simulations were terminated.

In addition, Morbidelli (2020) investigated the growth of a planet by pebble accretion in a dust ring with properties motivated by the DSHARP observations. Following the growth of an initially $0.1M_{\oplus}$ planetary embryo over 3 Myr, the torques on the planet were calculated using the Paardekooper et al. (2011) formalism, which resulted in minimal migration of the planet, such that the planet remained close to the pressure maximum. As such its growth was approximately the same as the case without migration, with no difference in the final planet mass.

Recent investigations have revealed further possibilities for planet migration. For example, Benítez-Llambay & Pessah (2018) showed that asymmetries in the dust distribution close to the planet may exert a net torque on the planet, the strength of which increases with the dust mass, and pebble accretion itself may compound this (Regály, 2020). Heating from the planet can give rise to an additional torque contribution, known as the “thermal torque” (e.g. Masset, 2017; Fromenteau & Masset, 2019), which has been shown to modify the migration, including generating outward migration and potentially changing the outcome of pebble accretion (e.g. Baumann & Bitsch, 2020; Guilera et al., 2021). Thus, planetary migration should be explored in future work to see how it might impact both vortex formation and planetary growth.

Considering these simplifications, then, it seems likely that those with the largest impact are those related to the treatment of the embryo’s accretion luminosity. Since the optimal accretion luminosity is assumed, and the gas in the vicinity of the planet is assumed to instantaneously adjust to the hotspot, the optimal vortensity generation rate ensues.

Comparison to previous work

As this work built on that of Owen & Kollmeier (2017), it is pertinent to provide a comparison to their results and thus assess the importance of self-consistently calculating the pebble accretion rate, in particular with respect to the impact on the planet mass achieved and on vortensity generation. In their simulations, Owen & Kollmeier (2017) held the planet mass at a constant $M_P = 1.5 \times 10^{-5}M_*$, a factor of 100 larger than my initial embryo mass. This difference in initial embryo mass is inconsequential since the initial growth of the embryo in my simulation is rapid, reaching $1.5 \times 10^{-5}M_*$ after 20 orbits, which is before vortex formation begins to affect the pebble accretion rate. What does matter is the subsequent growth of the planet, since it reaches $\sim 80M_{\oplus}$ before the vortex migrates from the planet location, which is large enough to carve a gap in the disc. This is in part driven by this new, extreme phase of vortex-assisted pebble accretion my simulations have identified. Gap formation does not occur in Owen & Kollmeier (2017) since their fixed-mass planet is simply not massive enough to strongly perturb the disc, and the extreme vortex-assisted pebble accretion rates were not anticipated. Based on

3.6. Conclusions

their simple model for pebble accretion in a dust trap (equations (3.3) and (3.4); figure 3.1), with values of $\Sigma_{\text{peb}} \sim 1 \text{ g cm}^{-2}$ and $a \sim 20 \text{ AU}$, Owen & Kollmeier (2017) assumed that the planet mass should remain low ($M_{\text{P}} \lesssim 10M_{\oplus}$), and therefore not interact strongly with the disc.

The use of a constant planet mass meant that their accretion luminosity was not coupled to the accretion rate, but was instead increased linearly over time until 50 planetary orbits, after which it was kept constant. This provided a sustained source of vortensity which allowed multiple vortices to form and ultimately merge into a single, large-scale vortex. In contrast, since my accretion rate is coupled to the local dust surface density, the accretion luminosity quickly drops off once the vortex carries dust away from the planet, which prevents further vortex formation.

In Owen & Kollmeier (2017) it was envisioned that once the vortex dissipates and releases the dust it had accumulated, the ring would re-form, restarting accretion onto the planet and potentially the vortex formation process. However, I found that the accretion rate required to initially form a vortex yields a planet massive enough to carve a gap in the disc, which prevents dust from entering its Hill sphere if the ring were to re-form. Moreover, my simulations have shown that dust released from the vortex settles in the (drag-modified) L_5 Lagrange point, and is only removed through diffusion, which is a slow process given expected viscosity values. Dust released from the dissipating vortex would therefore not immediately form a ring, and any further accretion of this material onto the planet would be at a rate insufficient to trigger vortex formation. The high planet mass that results from my simulations is due to the vortex-assisted pebble accretion phase, which was not anticipated by Owen & Kollmeier (2017), who therefore expected a planet below the mass required for gap-opening, thus expecting dust to re-form a ring.

Since my initial simulations have been performed at the extreme end of the parameter space it is still unclear whether the cycle anticipated by Owen & Kollmeier (2017) can occur or not. For example, if at lower initial dust masses a vortex can form but with lower planet mass, dust will be less strongly trapped at L_5 , and may also be able to accrete onto the planet once the vortex has dissipated because of the planet’s reduced gap-carving potential. Alternatively, if the interaction between the embryo and the vortex results in the planet migrating out of the ring, the vortex would no longer be within the planet’s co-orbital region, and would therefore re-form a ring. If the remaining mass in this ring is sufficiently high, it could give rise to a new embryo and begin the process anew.

3.6 Conclusions

I have investigated the growth of a pebble-accreting planetary embryo in an initially axisymmetric dust trap within a protoplanetary disc. I performed global hydrodynamics simulations of the dust and gas at a resolution high enough to resolve the pebble accretion process, allowing me to couple the accretion rate onto the embryo to the local dust surface density. By considering the accretion luminosity of the planetary embryo, I have shown that if the embryo can heat the disc beyond its Hill sphere, the gas becomes locally unstable to the formation of anticyclonic vortices. I found that the formation of a vortex at the location of the planet has significant consequences for the planet’s growth: since an anticyclonic vortex is a local pressure maximum, dust drifts towards the vortex and therefore towards the planet. The vortex thus provides a larger area for dust supply into the embryo’s Hill sphere than permitted by the planetary potential alone. This “vortex-assisted” pebble accretion ultimately leads to a higher final planet mass than can be achieved without considering the embryo’s accretion luminosity.

The planetary embryo accretes $\sim 100M_{\oplus}$ of the $150M_{\oplus}$ in the dust trap, resulting in large amount of solids in the planet, perhaps providing a formation mechanism for giant planets that are inferred to be extremely metal rich (e.g. Thorngren et al., 2016). The primary reason for the growth of such a large

Chapter 3. Vortex Formation from an Accreting Planetary Embryo

core is, as discussed above, the formation of the vortex at the planet location. When not accounting for the embryo’s accretion luminosity however, the embryo grows to $\sim 40M_{\oplus}$, which is still relatively large for a planetary core. The reason for this large core growth is due to the high surface density of pebbles in the dust trap, which is a result of the high disc dust mass and the viscosity assumed. While my disc parameters are observationally motivated, as non-axisymmetric discs are typically around Herbig Ae/Be stars which tend to have more massive discs, it is important to understand the range of parameter space in which this vortex formation process is able to occur. Thus, this provides motivation for a parameter study – this is the subject of the next chapter.

My simulations have also demonstrated a new link between planet formation and non-axisymmetric substructure in protoplanetary discs. While this mechanism does not produce vortices on the scale required to explain the most extreme asymmetries such as those observed in IRS 48 and HD 142527 – at least for the disc parameters I have used here – it may have some promise in explaining the non-axisymmetric features observed in HD 163296 and HD 143006, as well as those observed in recent DM Tau observations. The vortex survives for ~ 75 kyr (~ 500 orbits), though the dust asymmetry lasts for ~ 50 kyr – just a fraction of the typical $\sim$ Myr lifetime of protoplanetary discs. How the size and lifetime of the vortex and observable dust asymmetry varies with initial dust mass and physical disc properties such as temperature and viscosity will be the focus of the next chapter, in order to better understand the potential of this mechanism.

Chapter 4

The Impact of Accretion Luminosity on Planet Growth by Pebble Accretion

4.1 Introduction

In the previous chapter, I demonstrated that pebble accretion taking place in a dust trap with a high surface density of pebbles ($\sim 5 \text{ g cm}^{-2}$) generates substantial accretion luminosity from the planetary embryo. If the accretion luminosity can heat the disc beyond the embryo's Hill sphere above its background temperature, this process leads to the formation of an anticyclonic vortex in the vicinity of the planet. The presence of a vortex at the location of the planet can enhance the pebble accretion rate: being a pressure maximum, the vortex can funnel dust towards the planet if its size exceeds the planet's Hill sphere. However, the vortex can also act to inhibit the planet's growth when it migrates away from the planet, trapping dust and thus reducing the material available for accretion onto the planet.

Owen & Kollmeier (2017) proposed that the formation of a vortex through this mechanism should result in a lower planet mass, hypothesising that the vortex would trap the dust and thus prevent it from being accreted onto the embryo, inhibiting the planet's growth. Indeed, given a lower rate of vortensity generation, the vortex may separate from the planet earlier, and therefore trap dust before the planet becomes too massive. Therefore, an important factor in the outcome of the planet's growth is how long the vortex remains co-located with the planet, which is set by the dynamical interaction between the vortex and both the planet and the disc.

As a vortex is an overdensity in material, it exerts gravitational torques on, and thus exchanges angular momentum with the disc, allowing it to migrate in a similar manner to how a planet does (Paardekooper et al., 2010). Once the planet is large enough to open a gap in the disc, the (approximate) symmetry of the torques is disrupted and the vortex may migrate away from the planet. The interaction between the planet and vortex can be highly non-linear and dynamic, however; as discussed in §1.4.2, and corroborated in §3.4.1, the size of the vortex can only approach $4/3H_g$ in the radial direction before it begins to launch waves due to the supersonic velocity perturbations. These waves disturb the disc and may also contribute to the vortex's migration. However, since the presence of a vortex at the planet location increases the pebble accretion rate, which is the source of vortensity generation, this may result in an extended period over which the vortex and planet are co-located: the instability arising from vortex generation, which enhances the pebble accretion rate and thus further strengthens the vortex, has the potential to result in vigorous disc dynamics, where the vortex launches waves while having its energy replenished by the planet's accretion luminosity.

It is therefore of interest to investigate the impact of the planet's accretion luminosity across a range of parameter space, not just at the extreme end considered previously. In chapter 3, planet formation proceeded within a narrow ($\sim 10 \text{ AU}$) ring containing $150M_\oplus$ of dust and a low background disc temperature (35 K), providing a disc aspect ratio of 0.05 at the planet location; each of these is conducive to strong thermal feedback from the embryo's accretion luminosity, resulting in vortex formation. However, in order to establish how common this process is and how important it might be for planet formation models, it is necessary to understand its potential under more typical protoplanetary disc conditions.

In order to establish what constitutes "typical" protoplanetary disc conditions for this problem, it

Table 4.1: Disc ring properties from Dullemond et al. (2018). Column 1 is the name of the source, column 2 is the number of the ring, column 3 is the distance of the centre of the ring from the central star, column 4 is the width of the ring, column 5 is the estimated midplane disc temperature, column 6 is the aspect ratio of the disc at the ring location, column 7 is the mass of dust in the ring assuming the emission is optically thin and column 8 is the mass of the dust in the ring correcting for the optical depth.

Source	Ring	a_0 / AU	w / AU	T / K	h	$M_d^{\text{thin}}/M_{\oplus}$	$M_d^{\text{true}}/M_{\oplus}$
AS 209	1	74	3.38	15.8	0.08	27.0	31.5
AS 209	2	120	4.11	12.4	0.09	58.7	69.8
Elias 24	1	77	4.57	22.3	0.1	35.4	40.8
HD 163296	1	68	6.84	30.8	0.06	48.3	56.0
HD 163296	2	100	4.67	25.3	0.07	39.0	43.6
GW Lup	1	86	4.80	10.2	0.09	33.2	37.0
HD 143006	1	41	3.90	27.2	0.05	9.2	9.9
HD 143006	2	65	7.31	21.6	0.06	24.0	25.6

is instructive to examine the most comprehensive study of the distribution of dust in protoplanetary discs in this regard to date: the DSHARP survey (Andrews et al., 2018). While this survey specifically targeted bright and therefore massive discs, and intentionally avoided “transition” discs, it seems evident that it is these most massive discs that contain pronounced rings (e.g. Long et al., 2018; Andrews et al., 2018; Cieza et al., 2019, 2021; Kurtovic et al., 2021), and it is planet formation within such rings that I am concerned with. Note, however, that there is a possible observational bias given the tendency for surveys to focus on bright discs, and brighter discs tend to be larger, making substructure easier to resolve.

Dullemond et al. (2018) studied the properties of eight rings across five discs from the DSHARP survey; these specific discs were focussed on for the particularly strong brightness contrasts between their rings, allowing their widths and amplitudes to be clearly defined. Table 4.1 lists the properties of these rings as determined by Dullemond et al. (2018). Other notable examples of discs with well-defined, narrow rings include HD 169142 (Fedele et al., 2017; Pérez et al., 2019), HD 97048 (Walsh et al., 2016; van der Plas et al., 2017), GM Aur (Sierra et al., 2021), MWC 480 (Long et al., 2018; Sierra et al., 2021), ISO-Oph 17 and ISO-Oph 54 (Cieza et al., 2021) but these have not been analysed in as much detail on a per-ring basis as those in Dullemond et al. (2018). In the ALMA survey presented in Long et al. (2018), they estimated ring widths of ~ 10 to 30 AU, though at a lower spatial resolution of $\gtrsim 0.1''$ (cf. $\sim 0.035''$ in the DSHARP survey), so it is not clear whether the rings were adequately resolved.

The most significant difference between these discs and the system considered in chapter 3 is that these discs contain multiple concentric rings, with widths of ~ 5 AU and dust masses of $\sim 50M_{\oplus}$, rather than the “transition” discs with a single ring (i.e. the outer disc) spanning $\gtrsim 10$ AU and containing $\gtrsim 100M_{\oplus}$ in solids (it is also worth noting that the wide rings observed in some transition discs have been resolved into multiple rings when observed at higher spatial resolution, e.g. GM Aur: Macías et al. 2018; Sierra et al. 2021; LkCa 15: Facchini et al. 2020). Thus the rings listed in table 4.1 are more narrow and less massive than those used in the simulations presented in chapter 3, however they are further out in the disc, which bears consequences for the strength of the thermal feedback on the disc. The aspect ratio of the disc can be related to its temperature by

$$h(r) = \sqrt{\frac{k_B T_b r}{\mu m_H G M_*}}. \quad (4.1)$$

Assuming a background disc temperature profile of $T_b \propto r^{-1/2}$, the aspect ratio follows $h \propto r^{1/4}$. Thus, while many of the rings listed above are in cold regions of the disc, the disc flaring gives them

4.1. Introduction

large gas pressure scale heights. In order for the accretion luminosity to significantly impact the gas dynamics, the luminosity radius must exceed the accreting embryo's Hill radius; the background disc temperature is therefore critical when considering the impact of the accretion luminosity. This can be seen in equation (3.10), where the ratio of the luminosity radius to the Hill radius depends on the inverse square of the background disc temperature. Expressing this instead in terms of the aspect ratio,

$$\frac{R_L}{R_H} \propto \frac{M_P^{1/2} \Sigma_{\text{peb}}^{1/2} a^{1/4}}{R_P^{1/2} M_*^2 h^4}, \quad (4.2)$$

shows that the scaling with semimajor axis is weak, but strong with the inverse of the aspect ratio. The accretion luminosity of a forming planet is therefore much more significant in a thinner disc, making it easier to form a vortex.

The rings listed in table 4.1 have temperatures ~ 30 K, but aspect ratios between 0.05 and 0.1. Dullemond et al. (2018) determined the midplane disc temperature via

$$T_b(r) = \left(\frac{\phi L_*}{8\pi r^2 \sigma} \right)^{1/4}, \quad (4.3)$$

(e.g. Chiang & Goldreich, 1997), where $\phi = 0.02$ is the flaring angle of the disc (i.e. $h = a \tan \phi$) and σ is the Stefan-Boltzmann constant, allowing an estimate of the disc aspect ratio via equation (4.1). By fitting a Gaussian to each of the peaks in the azimuthally averaged intensity profile for each disc, and deconvolving the Gaussian beam, they obtained absolute widths of the rings. They estimated the mass of dust in each ring using the optically thin approximation (§1.4.1), but also calculated optical-depth-corrected dust masses.

The dust mass in a ring at a semimajor axis a , described by a Gaussian of width (standard deviation) w and amplitude A is given by

$$M_d = \int_{r_{\text{in}}}^{r_{\text{out}}} 2\pi r \Sigma_d dr \approx \frac{(2\pi)^{3/2} A a w}{\kappa_{\nu, \text{abs}} B_\nu(T_b)}, \quad (4.4)$$

where r_{in} and r_{out} enclose the ring radially. The dominant source of uncertainty in this dust mass estimate is the opacity, $\kappa_{\nu, \text{abs}}$, which depends on the maximum grain size assumed, the grain size distribution, as well as their composition and shape, and could be uncertain by factor of 10 (Birnstiel et al., 2018). Dullemond et al. (2018) adopted a grain size of 1 mm based on the width of the rings and an estimate of the dust's Stokes number, and thus used an opacity of $2 \text{ cm}^2 \text{ g}^{-1}$, given by the opacity model presented in Birnstiel et al. (2018). However, if we instead assume that the maximum grain size is given by the fragmentation limit – since a dust population trapped in a pressure maximum should reach fragmentation-coagulation equilibrium – the dust masses will be modified by a factor $\kappa_{\nu, \text{abs}}(s_{\text{max}} = 1 \text{ mm}) / \kappa_{\nu, \text{abs}}(s_{\text{max}} = s_{\text{frag}})$.

Determining the fragmentation-limited grain size requires an estimate of the gas density, which is difficult to do observationally (as discussed in §1.3.2). Dullemond et al. (2018) placed upper and lower boundaries on the gas surface density based on gravitational stability and dust-to-gas ratio arguments respectively, resulting in typical values of Σ_g between ~ 0.1 and 10 g cm^{-2} . Similarly, an estimate of the strength of turbulence is required; again, while this is difficult to constrain, a typically adopted value is

$\alpha \sim 10^{-4}$. Using equation (1.29), the maximum grain size can be expressed as

$$s_{\text{frag}} \approx \frac{2\Sigma_g}{\pi\alpha\rho_\bullet} \frac{v_{\text{frag}}^2}{GM_*} \frac{r}{h^2} \approx 1 \text{ cm} \left(\frac{\Sigma_g}{1 \text{ g cm}^{-2}} \right) \left(\frac{\alpha}{10^{-4}} \right)^{-1} \left(\frac{M_*}{2M_\odot} \right)^{-1} \left(\frac{r}{80 \text{ AU}} \right) \left(\frac{h}{0.1} \right)^{-2} \left(\frac{v_{\text{frag}}}{10 \text{ m s}^{-1}} \right)^2. \quad (4.5)$$

Since I am considering regions of the disc beyond the water iceline, grains are assumed to be icy and thus have fragmentation velocities of $v_{\text{frag}} \sim 10 \text{ m s}^{-1}$, which permits maximum grain sizes $s_{\text{frag}} \sim 1 \text{ cm}$. Using the Birnstiel et al. (2018) opacity model, $\kappa_{\nu, \text{abs}}(s_{\text{max}} = 1 \text{ cm}) \sim 0.5 \text{ cm}^2 \text{ g}^{-1}$, suggesting that the dust mass in these rings may well be up to a factor ~ 4 higher than those reported when assuming a maximum grain size of 1 mm. Note that this is comparable to the upper bound of the uncertainty ranges calculated by Dullemond et al. (2018). While there is uncertainty in the exact masses in these rings, the values observed from observations can be used as a guide and supplemented by theoretical considerations, thus it is these supplemented values that I consider in this work.

One important question to answer is how much of the mass in the dust trap the planet will accrete. Many of the observed rings contain vast amounts of large grains, and rings seem to be ubiquitous in discs observed in high spatial resolution observations. In chapter 3 I showed that, both with and without the thermal feedback from its accretion luminosity, a planet can grow beyond the classical pebble isolation mass in a dust trap. Thus, if planet formation is almost inevitable in these rings, as argued in chapter 3, one should expect to see many planets with large metal masses, i.e. planets with solid masses well above the typical mass required for runaway gas accretion, if they can accrete a large fraction of the dust within the ring. It is therefore of interest to determine how this varies with both the aspect ratio, upon which the classical isolation mass depends, as well as the accretion rate, set by the dust mass. The latter may be important if one reason that the isolation mass can be exceeded is simply that to the accretion time-scale is shorter than the gap opening time-scale.

Following Lin & Papaloizou (1986b), the gap-opening time-scale at a distance H_g from the planet can be estimated from $\Sigma_g/\dot{\Sigma}_g$, which gives

$$t_{\text{gap}} \approx \left(\frac{M_P}{M_*} \right)^{-2} \left(\frac{H_g}{a} \right)^5 \Omega_K^{-1} \sim 8 \text{ kyr} \left(\frac{M_P}{20M_\oplus} \right)^{-2} \left(\frac{a}{35 \text{ AU}} \right)^{3/2} \left(\frac{h}{0.05} \right)^5 \left(\frac{M_*}{2M_\odot} \right)^{3/2} \quad (4.6)$$

for a typical pebble isolation mass of $20M_\oplus$. This is longer than the growth time-scale of the planet seen in figure 3.4 where, without considering the accretion luminosity (since the waves launched by both the temperature disturbance and the vortex would also contribute to gap formation), the planet continued to accrete from $\sim 20M_\oplus$ at 5 kyr to $\sim 40M_\oplus$ at 10 kyr, at which point pebble accretion ceased. The time-scale for halting pebble accretion in the classical scenario will be shorter, however, as a pressure bump will form exterior to the planet's orbit, halting pebble drift, before the co-orbital region is significantly depleted of material, thus the above cannot be used to accurately determine the pebble isolation mass. This was also argued by Bitsch et al. (2018) in their investigation into the pebble isolation mass, and they showed that pebble drift is halted when gap depths are only around 10–20%. This can be seen in figure 3.4, where pebble accretion is halted despite the planet only having created a shallow gap.

Another issue to address is whether this mechanism can really produce large-scale vortices, i.e. comparable in size to those inferred from the observations of IRS 48 and HD 142527, as found by Owen & Kollmeier (2017). It is worth noting that in their simulations they used a disc aspect ratio of 0.1 at the planet location, and since discs with a larger scale height are able to support vortices with a larger radial

4.2. Method

extent, their simulations produced an asymmetry similar in size to those produced by other mechanisms, and similar to the scales inferred from observations. However, their vortensity generation was not calculated self-consistently, and the planet was kept at a low mass, which allowed multiple small-scale vortices to form and eventually merge to form one large vortex. It is not *a priori* obvious how this will vary with different dust masses, and thus different accretion rates, however. The strength of the generated vortex can be approximated by $\frac{D\omega}{Dt}\Delta t$, thus while a high accretion rate may generate vorticity quickly, it will produce a high mass planet which will quickly create a gap, preventing further accretion and therefore vorticity generation. Similarly, the highly dynamic interaction between the vortex and the planet may lead to the vortex separating from the planet quickly, preventing any further vorticity generation. However, a lower initial pebble accretion rate would allow vorticity to be generated more slowly, over a longer period of time, and may result in a stronger, larger vortex. Alternatively, a reduced vortex generation rate may permit the formation of multiple small-scale vortices, each of which separate from the planet in turn and eventually merge into one large-scale vortex.

In addition to the effect of accretion luminosity enhancing the pebble accretion rate, there were some additional outcomes of the simulations in chapter 3 that can be explored further by performing simulations across a wide range of parameter space. For example, I showed that since the dust populations most efficiently accreted onto the embryo were also those most rapidly trapped in the vortex, an inversion of the size distribution of pebbles was produced. I therefore wish to see whether this is a general outcome of this process, or whether it only occurs in the most extreme of cases. This could potentially be an observable indicator that this mechanism has taken place: if vortices are short-lived and we *are* seeing them shortly after formation, then the dust population within the residual ring(s) may not have had time to re-equilibrate into an MRN distribution.

Another outcome I wish to further investigate is how the vortex-assisted pebble accretion rate varies with the scale height of the disc. As discussed in the previous chapter, I showed that the vortex-assisted pebble accretion rate can be interpreted as the capture rate of pebbles by the vortex, which therefore depends on the vortex’s radial extent. The accretion rate is enhanced by a factor $2H_g^2/3R_H^2$ and thus depends on the aspect ratio of the disc, as the maximum radial extent of a vortex allowed by the Keplerian shear is $4/3H_g$; the explicit dependence on the aspect ratio will allow me to see whether this relation holds across parameter space.

Even if it doesn’t result in vortex formation, the accretion luminosity may affect the accretion rate simply through altering the velocity streamlines around the planet. There may therefore be a narrow region of parameter space where thermal feedback is insufficient to produce vortex-assisted pebble accretion, but can still affect the accretion rate, ultimately resulting in a difference in planet mass. I therefore want to determine how the accretion luminosity of a pebble-accreting planetary embryo affects its growth in a general sense, rather than just its ability to source vorticity and generate a vortex.

I first provide details of the dust masses and disc aspect ratios to be used in this parameter study, and recap the details of the hydrodynamics simulations performed. I then present results of simulations, focussing on the final planet mass reached as well as the consequences for the evolution of the disc for the various disc parameters used. I then discuss what these results imply about planet formation in dust rings, both in terms of the planet’s growth and observable disc substructure.

4.2 Method

In order to investigate the impact of an accreting embryo’s luminosity on its growth across a range of disc properties, I repeated the hydrodynamics simulations presented in chapter 3, with and without the thermal feedback from pebble accretion, but varied the initial mass of dust trapped in the axisymmetric

pressure maximum as well as the background disc temperature, motivated by more typical properties of ringed discs. I varied the initial mass in the dust trap from $50M_{\oplus}$ to $150M_{\oplus}$ in steps of $25M_{\oplus}$. For each dust mass, I also varied the disc temperature by using an aspect ratio $h(a)$ of 0.035, 0.05 and 0.1, with the gas pressure scale height given by equation (1.16). While an aspect ratio of $h(a) = 0.1$ requires a relatively high temperature at the semimajor axis of the planet, it is the scale height that determines the impact of the thermal feedback.

Apart from the dust mass and background temperature, the disc setup is otherwise the same as that in chapter 3, which I summarise here for completeness. I used a central star with mass $M_* = 2M_{\odot}$, hosting a disc of mass $M_{\text{disc}} = 0.05M_*$, 175 AU in radial extent and with a constant alpha viscosity of $\alpha = 10^{-4}$. The gas surface density is initialised according to

$$\Sigma_g = \frac{\Sigma_0}{2} \left(\frac{r}{R_0} \right)^{-1} \left[1 + \text{erf} \left(\frac{r - R_1}{\sqrt{2}\sigma} \right) + \epsilon \right], \quad (4.7)$$

with $R_1 = 1.2R_0$, $\sigma = 0.663R_0$ and $\epsilon = 10^{-3}$. I chose $R_0 = 35$ AU, and $\Sigma_g \approx 14.5 \text{ g cm}^{-2}$ to set the disc mass. Note that while the dust mass is varied between simulations, the initial gas mass is the same, thus the results from tests presented in chapter 3 to show that the initial gas distribution is stable both against self-gravity and the Rossby wave instability remain valid. I used a locally isothermal equation of state, which maximises the vertically integrated gas pressure at $r \approx R_0$ with the chosen gas surface density and pressure scale height profiles.

I used four dust species, with sizes 1 μm , 1 mm, 3.16 mm and 1 cm in an MRN size distribution and densities of 3 g cm^{-3} , with the mass in each species scaled in accordance with the various disc masses for each simulation. As the gas surface density is the same for each simulation, the dynamical properties of the dust grains, i.e. their Stokes numbers, are preserved. The initial dust density distribution for each dust species was calculated by solving the steady state advection-diffusion equation (cf. §2.1.2).

In each simulation I used an initial embryo mass of $M_P = 0.1M_{\oplus}$, positioned at the centre of the dust trap. The embryo is subject only to the potential from the central star, and thus cannot migrate due to the potential from the disc. The embryo's own potential is smoothed according to equation (3.31).

The accretion luminosity was calculated using equation (3.7), with the planet radius calculated using Fortney et al. (2007) with an ice mass fraction of 0.5. The hotspot temperature was calculated using equation (3.32), with a temperature smoothing length of $0.65R_{\text{H},0}$. I kept the grid resolution the same, i.e. 1920 cells in radius and 7680 in azimuth, spanning $r \in [0.54, 5.0]R_0$ and $\phi \in [0, 2\pi]$. The resolution is maximised at the planet location by using variable cell sizes in radius, defined by equation (3.29). I used the same boundary conditions, and the same damping zone sizes and damping timescales.

4.3 Results

I present the results of simulations in which the initial mass within the dust ring was varied between $50M_{\oplus}$ and $150M_{\oplus}$, for aspect ratios of $h(a) = 0.035, 0.05$ and 0.1 , first showing the impact on the final planet masses achieved, before focussing on the evolution of the discs and their substructure.

4.3.1 Planet growth

Final planet mass

Figure 4.1 shows the mass that the embryo reached via pebble accretion after 60 kyr (~ 400 orbits) within each dust ring, as a function of the mass within the ring and the disc aspect ratio at the planet location, with and without including the thermal feedback on the disc. In each simulation the planet

4.3. Results

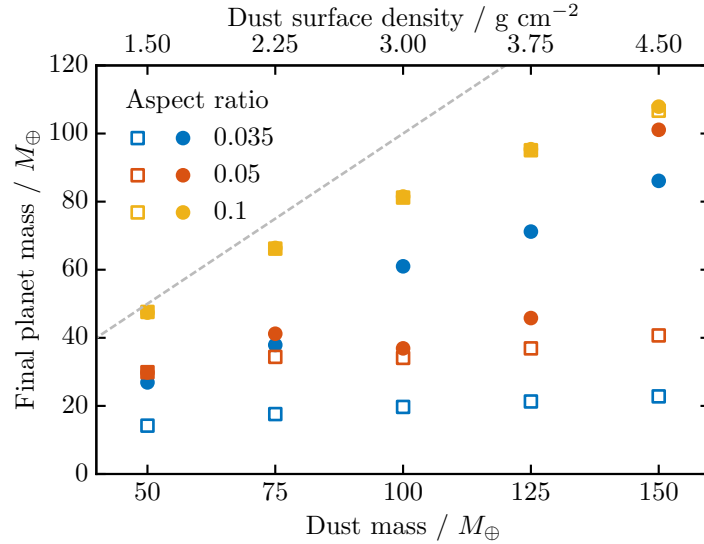


Figure 4.1: Final planet mass for each ring mass and aspect ratio. Filled circles are for simulations with thermal feedback from pebble accretion, open squares are for simulations without. The grey dashed line shows the mass the planet would reach if it accreted all dust within the ring. Also shown is the peak dust surface density at the location of the planet, which scales linearly with the initial dust mass.

reaches its final mass within 60 kyr.

The final planet mass generally increases with both the scale height of the disc as well as the initial mass in the ring. The thermal feedback results in a planet mass that is always greater than or equal to that reached when neglecting this effect. This is most apparent in the coldest discs, where there is a factor of ≈ 2 increase in the least massive ring, up to a factor ≈ 4 increase in the most massive ring. The increase in final planet mass with scale height holds for growth with and without thermal feedback on the disc.

In the $h(a) = 0.05$ discs, there is a non-monotonic increase in planet mass with increasing ring mass: the $100M_{\oplus}$ ring produces a lower mass planet than the $75M_{\oplus}$ ring, otherwise the planet mass increases with ring mass. Similarly, the increase in final planet mass due to thermal feedback increases non-monotonically with increasing ring mass, with no difference in the least massive ring, up to a factor of 2.5 in the most massive ring. In the discs with the highest aspect ratio, thermal feedback only affects the planet’s growth in the most massive ring, with an increase of $1.2M_{\oplus}$ ($\approx 1\%$) over that reached when neglecting accretion luminosity, otherwise there is no discernible difference when including the thermal feedback. Despite the embryo’s accretion luminosity having little-to-no effect on its growth, its final mass increases approximately linearly with the ring mass – a stronger scaling than one might expect given the results at lower aspect ratios.

Accretion rates

How the planet’s final mass varies with the initial dust mass and aspect ratio can be better understood by considering the pebble accretion rates over time. Figure 4.2 shows the growth of the planet over 60 kyr for each ring mass and aspect ratio. For the lowest aspect ratio, and thus the lowest background disc temperature, the impact of the thermal feedback is clearly apparent for all initial ring masses: there is a distinct jump in accretion rate seen after ~ 1 kyr, at progressively later times for lower dust masses (~ 3 kyr for the lowest dust mass). In each case the strong thermal feedback results in the formation of a vortex which enhances the pebble accretion rate, producing a significantly more massive planet than when neglecting the thermal feedback.

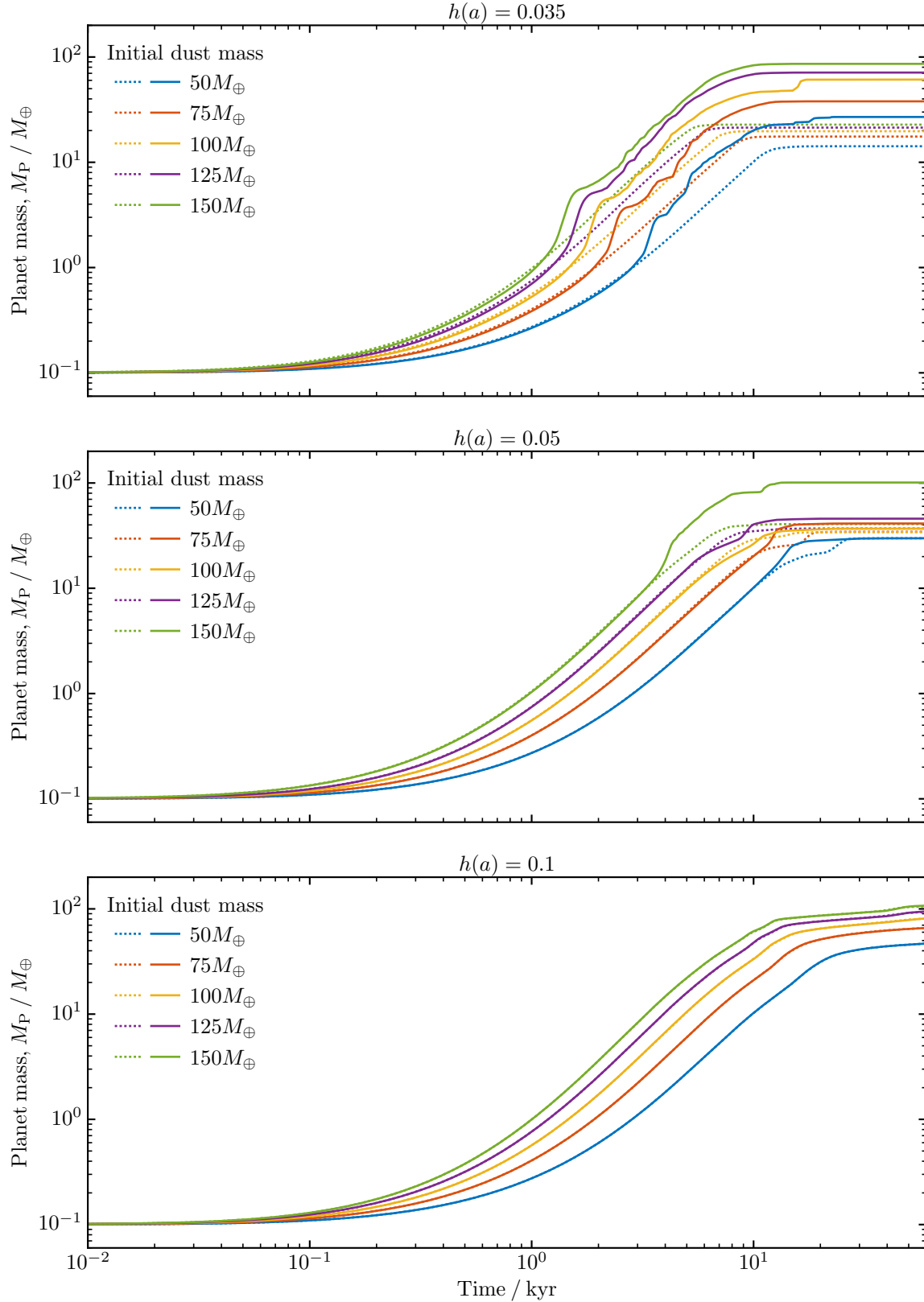


Figure 4.2: Planet mass as a function of time for a range of initial dust masses and disc temperatures. Each panel shows a comparison of the planet’s growth for different dust masses in a dust trap with the same scale height. The dashed lines show the planet’s growth without accounting for the thermal feedback on the disc from its accretion luminosity, while the solid lines show the growth when accounting for this effect.

4.3. Results

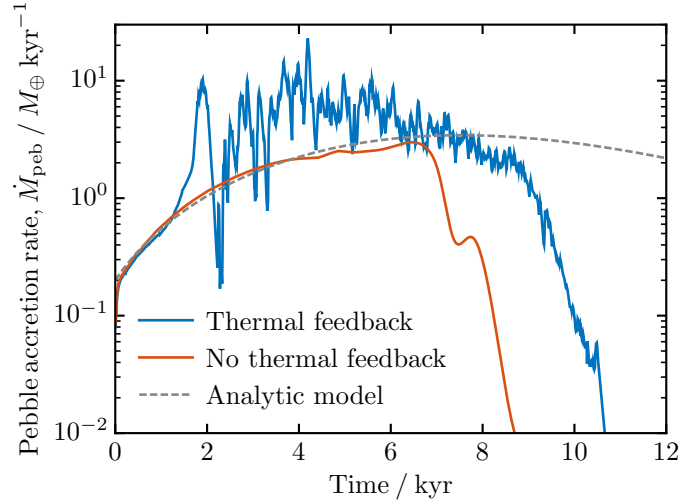


Figure 4.3: Pebble accretion rates onto the embryo, with and without thermal feedback, in the $100M_{\oplus}$, $h(a) = 0.035$ ring. The blue line shows the pebble accretion rate in the simulation with thermal feedback. The red line shows the pebble accretion rate in the simulation without thermal feedback. The grey dashed line shows the pebble accretion rate from a simple analytic calculation, which does not account for gap-opening or the embryo’s accretion luminosity.

The pebble accretion rates show similar behaviour for each dust mass in the $h = 0.035$ rings, closely following the accretion rate when neglecting thermal feedback until the vortex forms, at which point it rapidly increases. While the vortex is co-located with the planet, the pebble accretion rate, although showing significant fluctuations, has a reasonably steady time-averaged value, as was seen in chapter 3. The pebble accretion rate then begins to decay after ~ 5 to 9 kyr, depending on (but not necessarily scaling with) the ring mass, before rapidly dropping off once the vortex separates from the planet. This can be seen in figure 4.3, which shows the pebble accretion rate over the first 12 kyr onto the planet accreting in the $100M_{\oplus}$ ring. This is largely representative of the pebble accretion rates for each ring mass in the $h(a) = 0.035$ discs, however, the increase in accretion rate due to the formation of the vortex begins earlier, and the pebble accretion rate is higher, with increasing ring mass. Furthermore, only in the $50M_{\oplus}$ case does the accretion rate sharply drop off *before* it does in the simulation without thermal feedback – this is likely to be due to the fact that the drop-off in accretion rate without thermal feedback is shallower for lower dust masses, as the planet mass and therefore pebble accretion rate is lower.

The simulations for the intermediate scale height with $h(a) = 0.05$ appear to show a range of behaviours, suggesting that they may cross the region of parameter space over which the thermal feedback becomes important for the planet’s growth. As discussed in the previous chapter, the $150M_{\oplus}$ dust trap provides an accretion rate sufficient to result in vortex formation and therefore a phase of vortex-assisted pebble accretion. For the $125M_{\oplus}$ ring however, an initial enhancement in the accretion rate with respect to the case without thermal feedback is seen, followed by a reduction in accretion between 6 and 9 kyr, before the accretion rate increases again. This does not follow the same course of events as in the $150M_{\oplus}$ ring, purely due to the fact that the rate of vortensity generation is reduced and the vortex that forms is not large enough to enhance the accretion rate. Similar behaviour is seen at $100M_{\oplus}$, although without the initial increase in accretion rate, due to even weaker vortex formation. Moving to even lower dust masses, the behaviour of the accretion rate and thus the planet’s growth over time changes further still. At $75M_{\oplus}$ there is an enhancement in accretion rate after 10 kyr and the planet ends up more massive than in the $100M_{\oplus}$ case – this is the only instance of planet growth curves crossing; incidentally, the final planet mass in the simulation without thermal feedback is also higher in the $75M_{\oplus}$ ring than the $100M_{\oplus}$ ring, suggesting that there must be another effect at play here. This could perhaps relate to

the accretion timescale and gap-opening timescales being comparable, allowing the embryo to continue accreting while it opens a gap. For the lowest dust mass of $50M_{\oplus}$, similar behaviour is observed, with an increase in accretion rate with respect to the control simulation, but the planet’s final mass is the same as without the accretion luminosity.

For the largest scale height, the thermal feedback from the embryo’s accretion luminosity is negligible and as such there are no significant differences between the planet’s growth with and without including its accretion luminosity. This can also be seen in figure 4.4, which shows the pebble accretion rates for this aspect ratio. The accretion rates with and without thermal feedback are indistinguishable for each ring mass until around 10 kyr, where slight departures can be seen in the most massive rings. Given that the pebble accretion rate scales with the square of the planet’s Hill radius, it is no surprise that slight differences in accretion rate become amplified at later times for the most massive planets. However, as shown in figure 4.2, this has little effect on the planet’s final mass.

An interesting feature is the bump in pebble accretion rate onto the planets in the $100M_{\oplus}$, $125M_{\oplus}$ and $150M_{\oplus}$ rings with $h = 0.1$ between 30–55 kyr. Only in these rings, at this scale height, does the planet become massive enough to create a pressure bump outside of its orbit – albeit a weak one – which concentrates pebbles into a ring. As the planet keeps on growing during this gap-opening phase, however, its Hill sphere extends into to this ring and thus accretes from it. This process is discussed further in §4.4.2.

Weak vortices do in fact form in the most massive rings with $h = 0.1$, but quickly separate from the planet and thus have a negligible impact on the evolution of the disc or the planet’s growth; the final planet mass is set purely by planet–disc interactions. A more detailed look at the vortex formation process and the conditions under which the vortex affects the accretion rate are given in §4.4.1. The fluctuations in pebble accretion rate onto the most massive planets seen beyond ~ 20 kyr in figure 4.4 are due to the accumulation of dust around L_5 ; the planet does not reach a high enough mass to carve a gap in the disc, thus as the smaller (millimetre-sized) dust grains are less strongly concentrated around L_5 they are able to enter the planet’s Hill sphere as a vortex attracts them on passing. Given that this happens in the simulations both with and without thermal feedback, this vortex cannot have formed via the embryo’s accretion luminosity, and thus likely forms due to the planet’s spiral wake: the shock generated by the planet is a source of vorticity (e.g. Dong et al., 2011; de Val-Borro et al., 2007; Lin & Papaloizou, 2010), which is stronger in the leading shock, thus once the planet opens a gap, conservation of vortensity means that this small vortex structure survives until dissipation due to viscosity.

Quantifying the thermal feedback from accretion luminosity

The planet’s accretion luminosity can only have a significant effect on the disc if its luminosity radius exceeds its Hill radius. The impact of the accretion luminosity on the planet’s growth can therefore be understood by considering how the ratio of these two quantities varies between simulations. Figure 4.5 shows the time-evolution of this ratio for each initial dust mass and gas pressure scale height. This ratio is proportional to the square root of the pebble accretion rate and thus increases with the mass in the ring. Similarly, it is inversely proportional to the square of the background disc temperature, and thus decreases with increasing aspect ratio.

The luminosity radius, being the distance from the planet at which the temperature due to its accretion luminosity equals that due to the stellar luminosity alone, is therefore smaller for hotter discs (cf. equation (3.10)). For the coldest disc, the embryo’s luminosity radius exceeds its Hill radius from the beginning of the simulation for all dust masses, resulting in strong thermal feedback and vortex formation, until the vortex separates from the planet, at which point the ratio falls below unity. For the disc with $h(a) = 0.05$, the luminosity radius is initially smaller than the Hill radius, but as the embryo

4.3. Results

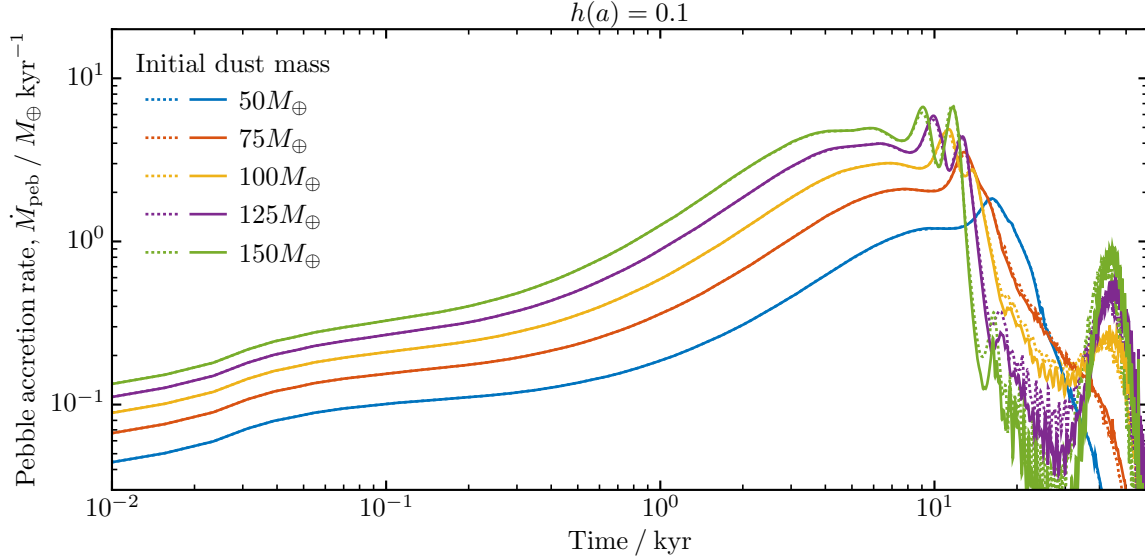


Figure 4.4: Pebble accretion rate as a function of time for a range of initial ring masses for the disc with the largest scale height. The dashed lines show the accretion rate without accounting for the thermal feedback on the disc from its accretion luminosity, while the solid lines show the accretion rate when accounting for this effect.

grows its Hill sphere expands, enabling it to accrete from a greater distance. This increases its accretion rate and this ratio eventually exceeds unity for all ring masses, at times ranging from ~ 500 yr in the most massive ring to ~ 4 kyr in the least massive. In the $150M_{\oplus}$ ring, this results in vortex formation and thus an enhancement in accretion rate at around 3.5 kyr, once this ratio exceeds ≈ 3.5 ; this ratio also reaches ≈ 3 in the $125M_{\oplus}$ ring, and a vortex does form, however it quickly separates from the planet, preventing a phase of vortex-assisted pebble accretion from occurring. Vortex formation does not occur in the less massive rings at this aspect ratio. In the disc with the largest aspect ratio, the luminosity radius never exceeds the Hill radius. The maximum value their ratio reaches decreases with ring mass, from ≈ 0.25 in the most massive ring, to ≈ 0.1 in the least massive.

4.3.2 Disc evolution

In addition to its effect on the planet's growth, the thermal feedback on the disc has consequences for the evolution and structure of the disc, with the formation of vortices producing non-axisymmetric features as well as the axisymmetric substructure formed as the planet carves a gap in the initial ring. The formation of substructure also affects the size distribution of dust grains, as the efficiency of trapping in pressure maxima is grain-size-dependent.

Dust distribution

In the disc with the largest scale height, the only dust remaining once pebble accretion has ceased is that in the co-orbital region. Only in the $150M_{\oplus}$ ring does the planet open a partial gap and create a ring outside of its orbit but, as described above, the planet is able to accrete the dust from this ring, making it a short-lived feature. The vortices that form in the most massive rings are weak and therefore don't trap significant amounts of dust. The remaining dust therefore gradually settles towards L_5 . For the $h(a) = 0.05$ discs, regardless of ring mass, the remaining dust resides in the new axisymmetric pressure maxima that form either side of the planet's orbit. However, the largest dust grains are not present in the inner ring – they are only found in the outer ring as well as in L_5 . The millimetre-sized grains

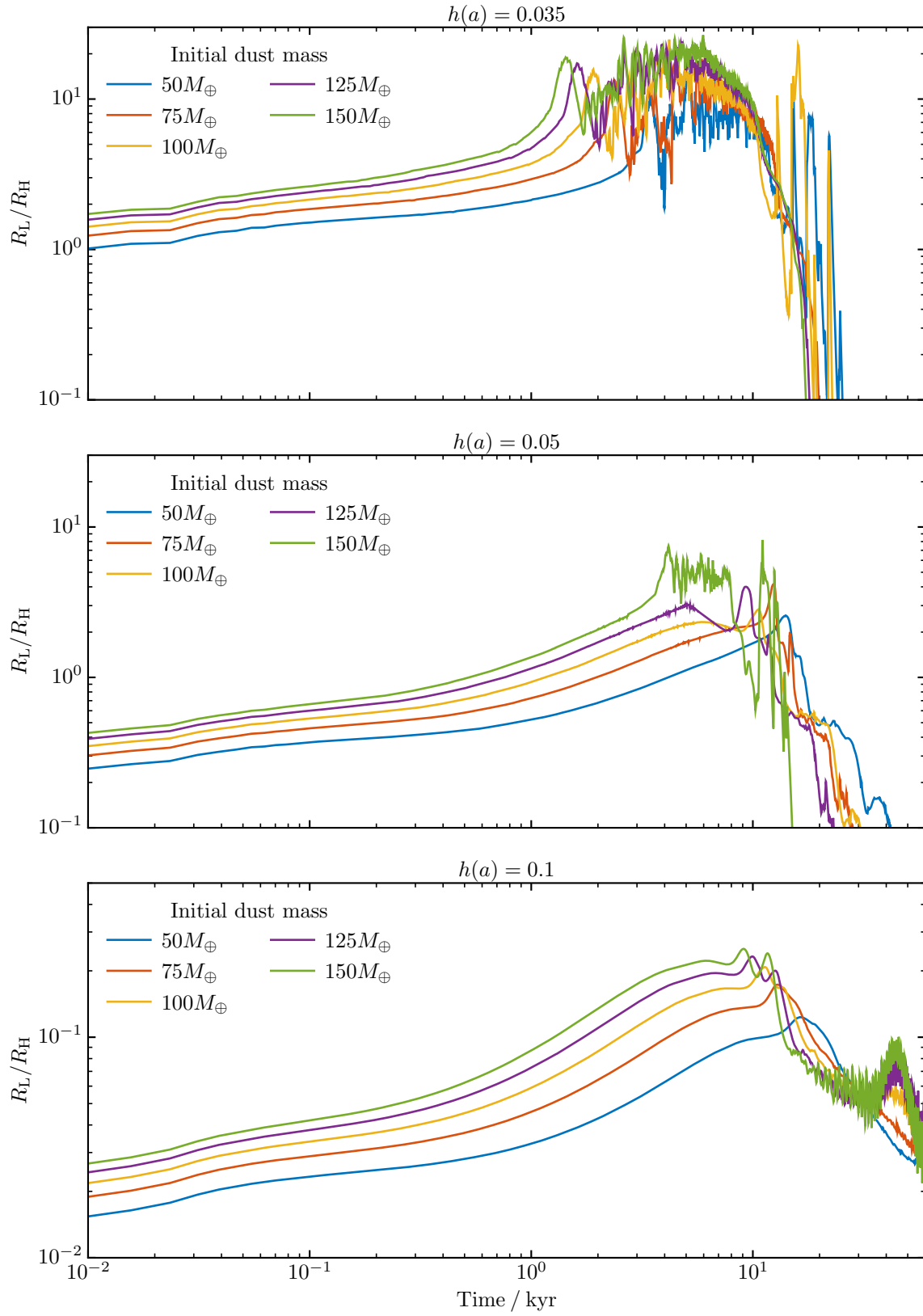


Figure 4.5: Ratio of the accreting embryo's luminosity radius to its Hill radius as a function of time, for a range of initial dust masses. Each panel shows a comparison of this ratio for different dust masses in a dust trap with the same aspect ratio, which is indicated above each panel.

4.3. Results

are found in both rings and around L_5 . In the coldest rings all remaining dust is trapped in both rings enclosing the planet's orbit as well as L_5 .

The dust remaining in the co-orbital region therefore accumulates around the L_5 Lagrange point for all aspect ratios, even in the least massive rings. This is of particular importance in relation to the hypothesis put forward by Owen & Kollmeier (2017), that dust released by the vortex upon its eventual dissipation would reform a ring, restarting pebble accretion. In chapter 3 I speculated that perhaps, if the final planet mass is not too high (i.e. below the gap-opening mass), then the remaining dust would not accumulate in L_5 but reform a ring. However, the results of these simulations suggest that even in the least massive rings, at this viscosity, the planets produced are still massive enough ($\sim 10\text{--}20M_\oplus$), to break the azimuthal symmetry of the ring, leaving dust to settle at L_5 .

In all simulations the larger dust grain populations lose the largest fractions of their mass. In the $150M_\oplus$ ring with an aspect ratio of 0.05 this results in an inversion of the size distribution between the millimetre and centimetre grains, as described in the previous chapter, however this is the only instance in which this occurs. Because these largest grains have Stokes numbers closest to unity they are preferentially accreted onto the embryo, and the thermal feedback on the disc can act to emphasize this effect: as higher pebble accretion rates generally result in stronger thermal feedback, then if a vortex forms this increases the pebble accretion rate and reduces the mass in the largest dust species as they drift towards the vortex the fastest. An example of this is shown in figure 4.6, which shows the mass in each dust species in the $125M_\oplus$, $h = 0.035$ ring, with and without thermal feedback. This illustrates how the centimetre-sized grains lose the largest fraction of their mass, without the size distribution being inverted.

In the $h = 0.1$ rings, no inversion of the size distribution results due to the initial pebble accretion phase. However, as the planet forms a weak pressure maximum outside of its orbit, only the highest Stokes number grains are trapped there while the millimetre grains remain in the co-orbital region, destined for L_5 . Thus, once the planet accretes from the resulting ring (as described above), the reserve of centimetre grains is depleted, and the millimetre grains dominate the final disc mass. This suggests that the formation of pressure maxima of any kind by the planet helps preserve the largest dust grains. While these are accreted onto the planet most efficiently, then unless the planet accretes all of the dust before opening a gap, the rest end up in the vortex or the resulting rings. If trapped in a vortex, they end up in L_5 once it dissipates; if trapped in rings, they stay there. If no pressure maxima form, or only weak maxima form, such as those formed by the planet in the $h = 0.1$ rings, the population of centimetre grains are almost entirely depleted.

Vortex lifetimes

The lifetimes of the vortices formed from the thermal feedback on the disc range from ~ 250 to 450 orbits, with a tendency towards longer lifetimes in lower mass rings. I measured the vortex lifetime as the time between formation and the point at which the vortensity reduces to the background level, determined by eye. As vortices only form consistently via this process in the $h = 0.035$ rings, a comparison as a function of ring mass can only be made for this aspect ratio.

In the most massive ring the vortex survives for 250 orbits. The vortex lifetime increases to around 300 and 450 orbits in the $125M_\oplus$ and $100M_\oplus$ rings respectively. In the $75M_\oplus$ ring there are two significant vortices that form: the initial vortex that forms around the planet separates from the planet after ≈ 40 orbits, but the embryo continues to accrete, sourcing vortensity and eventually a second vortex forms, which separates after 110 orbits. Both vortices dissipate after a further ~ 200 orbits. A similar outcome is seen in the least massive ring, although these survive for ~ 400 orbits.

The three main effects that can weaken the vortices are viscosity, shocks and the increased shear

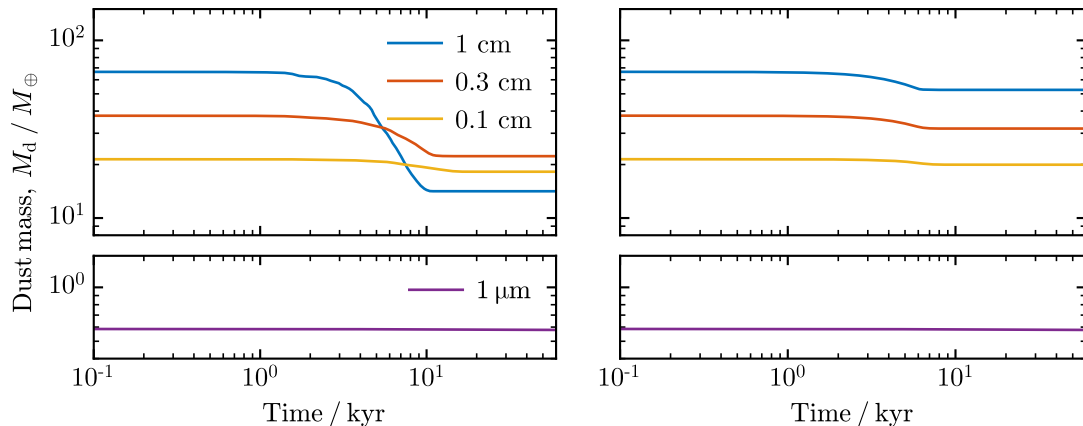


Figure 4.6: Mass in each species of dust over the duration of the simulation, for the $125M_{\oplus}$, $h = 0.035$ ring. The left and right-hand panels are for the simulation with and without thermal feedback. When vortex formation increases the pebble accretion rate onto the embryo, the centimetre-sized grains, are both preferentially attracted towards the vortex and accreted onto the embryo.

that the vortices experience during horseshoe turns. Since the viscosity is relatively low, shocks and the shear during horseshoe turns are the dominant effects. This can explain why the vortex lifetime in the $75M_{\oplus}$ ring is an outlier from the trend: there are two vortices on horseshoe orbits, each generating spiral shocks, which are therefore constantly disrupting each other. This is also the case in the $50M_{\oplus}$ ring, however, as the planet is less massive than in the $75M_{\oplus}$ ring, the horseshoe region is narrower, thus there is less shear, allowing the vortices to survive longer. In the $h = 0.05$, $150M_{\oplus}$ ring, the vortex survives for ~ 500 orbits, approximately twice as long as in the corresponding $h = 0.035$ ring. While the higher aspect ratio ring produces a more massive planet, it permits a more massive and therefore a stronger vortex, thus increasing its lifetime. However, since there are many effects at work here, from those determining the strength of the vortex when it separates from the planet to those which act to dissipate the vortex, the impacts of which are difficult to disentangle, this remains hypothetical and thus requires further investigation.

Vortices also form via the Rossby wave instability at the edges of the gap created by the planet in the lower aspect ratio rings. This does not occur in the $h = 0.1$ rings, or in the lowest mass $h = 0.05$ ring, as the planet is unable to carve a deep enough gap to trigger this instability. These vortices are generally weak and only the smaller dust grains, being strongly coupled to the gas, tend to form asymmetries; the larger dust grains, having longer stopping times, typically maintain their axisymmetric ring structures. Gap formation is even easier in the coldest rings, thus even the planet in the $50M_{\oplus}$ ring is able to carve a gap whose edges are unstable to the RWI. However, only in the more massive rings are the vortices that form at the gap edges able to trap dust significantly. These vortices eventually dissipate and the dust rings reform within the axisymmetric pressure maxima, and thus have no impact on the planet or final disc structure.

4.4 Discussion

4.4.1 Vortex-assisted pebble accretion

In chapter 3 I hypothesised that the jump in pebble accretion rate seen when the vortex initially forms around the planet can be interpreted as arising due to the vortex providing a larger accretion cross-section than the embryo's Hill sphere, with its maximum semiminor axis being $\approx 2H_g/3$. The resulting

4.4. Discussion

expression for the accretion rate, equation (3.36), shows the dependence on both the aspect ratio and dust surface density, which means that this hypothesis can be directly tested with the simulations I have performed here. As the simulations with the same aspect ratio vary only by dust mass, and therefore dust surface density in the ring, the accretion rates should vary linearly with surface density between simulations. While this is true for both the vortex-assisted pebble accretion rate and the standard pebble accretion rates, in the case of vortex-assisted accretion the accretion rate is independent of planet mass, enabling a potential means of discriminating between the two accretion scenarios. Similarly, simulations performed with the initial same dust mass vary only by the disc aspect ratio, though only the $150M_{\oplus}$ rings provided an enhancement in pebble accretion rate due to the presence of the vortex, and only for aspect ratios $h(a) = 0.035$ and 0.05 , thus this scaling is more difficult to reliably test with this set of simulations.

Figure 4.7 shows the total pebble accretion rates in the $50M_{\oplus}$, $100M_{\oplus}$, and $150M_{\oplus}$ rings with $h = 0.035$ during the first 12 kyr of the simulations. These should scale linearly with the local dust surface density, and thus increase with the initial mass of dust mass in the ring. Furthermore, according to equation (3.36), they should be constant with time during the vortex-assisted phase. The initial peak observed in each of the accretion rates due to vortex formation around the planet agrees well with this expectation, scaling accordingly with initial dust mass. In each case the accretion rate drops significantly after this peak, before rising up again on a time-scale of ~ 0.5 kyr and reaching a roughly constant value. After a period of time which is dependent on the initial dust mass, the pebble accretion rate begins to decay, before sharply falling once the vortex separates from the planet. The steady decay in pebble accretion rate, particularly apparent in the $100M_{\oplus}$ and $150M_{\oplus}$ rings (starting around 6 kyr) will be discussed in the following sections. The time-averages of the pebble accretion rate during the vortex-assisted phase, however, do not scale linearly with the initial dust mass. These are $2.6M_{\oplus} \text{ kyr}^{-1}$, $5.1M_{\oplus} \text{ kyr}^{-1}$ and $10.8M_{\oplus} \text{ kyr}^{-1}$ for the $50M_{\oplus}$, $100M_{\oplus}$ and $150M_{\oplus}$ rings respectively, thus the pebble accretion rate in the most massive ring is higher than expected. Note, however, that the initial dust mass serves only as an estimate of the surface density of dust available for pebble accretion: higher pebble accretion rates will lead to stronger vortex formation, which in turn leads to higher surface density in the vicinity of the planet due to dust trapping by the vortex.

Figure 4.8 shows the pebble accretion rates in the $150M_{\oplus}$ rings over the first 12 kyr of the simulations. The accretion rates onto the embryo are very similar for the first 1000 years, as expected, before a vortex forms in coldest disc and the accretion rate increases by a factor ~ 10 as a result. After 3.5 kyr a vortex forms in the $h(a) = 0.05$ disc, and the accretion rate increases by a factor ~ 5 . Given the difference in scale height, the accretion rates during the period while the vortex is co-located with the planet should vary by a factor ≈ 2 , which is not observed – the accretion rates are approximately equal. A direct comparison of these accretion rates is not straight-forward, however. Vortex-assisted pebble accretion requires $2H_g^2/3 > R_H^2$, which sets a limit on the planet mass for which this can occur, at a given aspect ratio:

$$M_P \lesssim 3M_* \left(\frac{2h^2}{3} \right)^{3/2}. \quad (4.8)$$

Beyond this mass, which is comparable to the thermal mass, standard pebble accretion in the Hill regime takes over, as the planet’s Hill radius exceeds the maximum semiminor axis of the vortex. For an aspect ratio of 0.035 , this threshold is reached once the embryo grows to $\approx 47M_{\oplus}$, which occurs shortly after 5 kyr in the $150M_{\oplus}$ ring, while the vortex-assisted pebble accretion phase in the $h = 0.05$ ring lasts from 3.5 to 8 kyr, after which the vortex separates from the planet. The time-averaged pebble accretion rate during the vortex-assisted phase for the $h = 0.035$ ring, approximately $20M_{\oplus} \text{ kyr}^{-1}$, is a factor of 1.4 lower than that during the equivalent phase in the $h = 0.05$ ring ($\approx 28M_{\oplus} \text{ kyr}^{-1}$), rather than the predicted factor of 2. This can be seen in the right-hand panel of figure 4.8, which shows the accretion rate

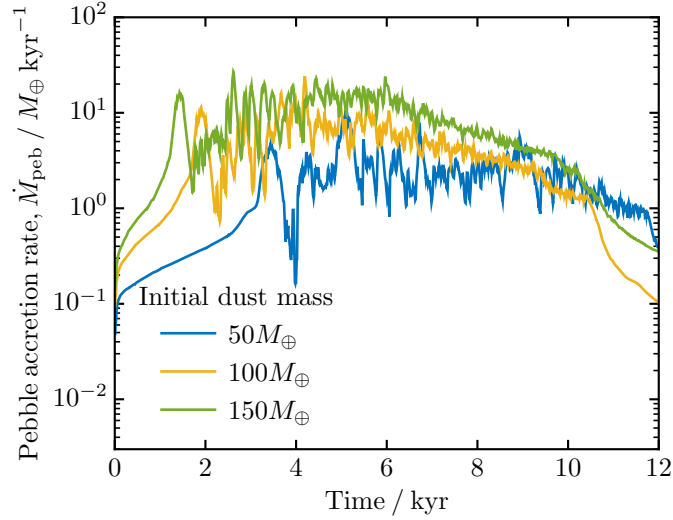


Figure 4.7: Pebble accretion rate onto the embryo growing in the $50M_{\oplus}$, $100M_{\oplus}$ and $150M_{\oplus}$, $h(a) = 0.035$ rings over the first 12 kyr. Analytically, the pebble accretion rate is expected to scale linearly with the dust surface density, and therefore the initial dust mass. The initial peaks in accretion rate due to vortex formation reach 5, 11 and $16M_{\oplus} \text{ kyr}^{-1}$ respectively, which do scale approximately linearly with the initial dust mass, while the time-averaged accretion rates (2.6 , 5.1 and $10.8M_{\oplus} \text{ kyr}^{-1}$ respectively) during the vortex-assisted pebble accretion phase scale logarithmically with the initial dust mass.

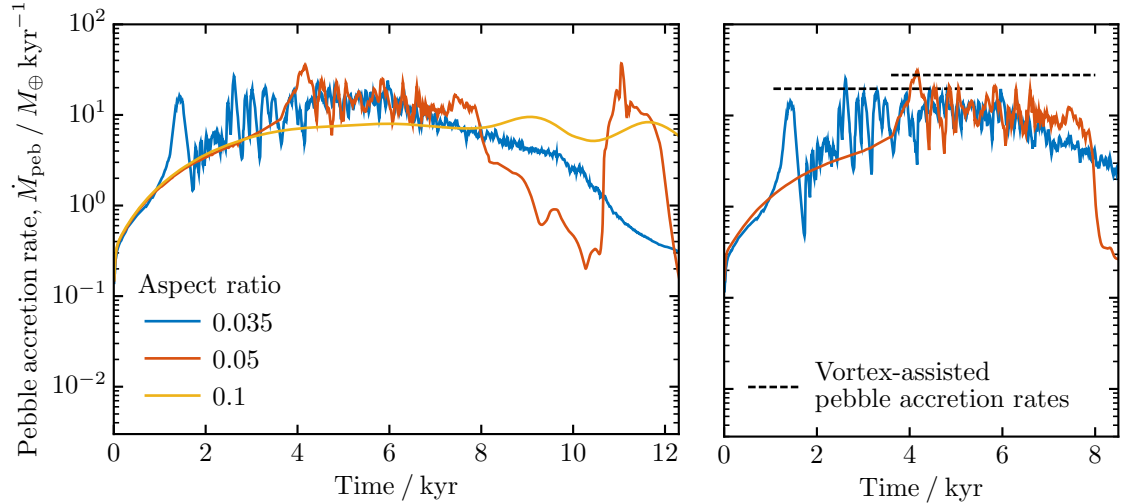


Figure 4.8: Pebble accretion rates in the $150M_{\oplus}$ rings with aspect ratios of 0.035, 0.05 and 0.1 over the first 12 kyr of the simulations. Both the $h = 0.035$ and $h = 0.05$ rings lead to vortex formation, enhancing the accretion rate, while the thermal feedback is insufficient in the highest temperature ring. The right-hand panel shows the accretion rates of the centimetre-sized pebbles only, for the $h = 0.035$ and $h = 0.05$ rings over first 8 kyr, with the analytic time-averaged vortex-assisted pebble accretion rates.

4.4. Discussion

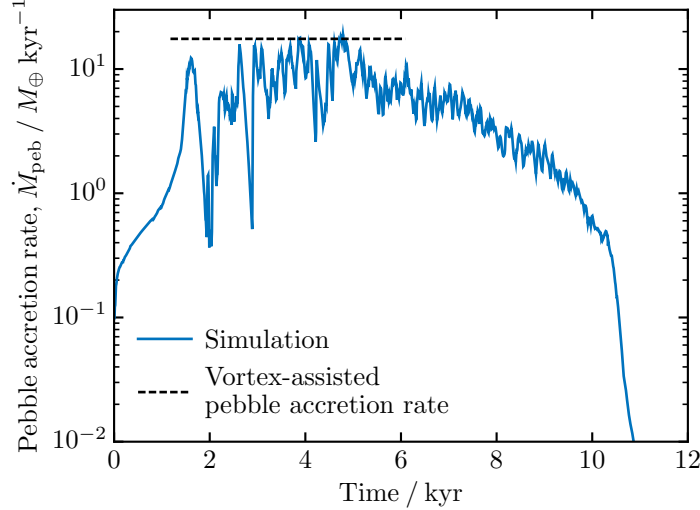


Figure 4.9: Accretion rate of the centimetre-sized pebbles onto the embryo in the $125M_{\oplus}$, $h(a) = 0.035$ during the vortex-forming phase. The blue line shows the pebble accretion rate from the simulation with thermal feedback. The black dashed line shows the analytic time-averaged vortex assisted pebble accretion rate, between the period when the vortex size exceeds the embryo’s Hill radius.

of the centimetre-sized pebbles onto the embryo, as well as the time-averaged analytic accretion rates. Only the accretion rates for centimetre-sized dust grains are shown as the Stokes number dependence is not accounted for in the analytic expression of the vortex-assisted pebble accretion rate, and since these largest grains have Stokes numbers closest to unity, this minimises the expected impact of this additional factor. However, while a discrepancy of order unity between the simulated and analytic accretion rates is not surprising, neglecting the Stokes number dependence should not impact the ratio between the accretion rates in different aspect ratio rings.

The fact that the embryo’s growth in the more massive $h = 0.035$ rings transitions out of the regime where the vortex size exceeds the embryo’s Hill sphere motivates taking a closer look at the pebble accretion rates during this period. Figure 4.9 shows the accretion rate of the centimetre-sized pebbles onto the embryo growing in the $125M_{\oplus}$ $h = 0.035$ ring, for the first 12 kyr of the simulation, as well as the time-averaged pebble accretion rate given by equation (3.36) between the time that the vortex forms and the time at which the vortex no longer provides a greater accretion cross-section than the embryo’s Hill sphere. Between ~ 2 and 5 kyr, the accretion rate shows large fluctuations, but about a value that is roughly constant in time. After around 5 kyr the accretion rate begins to decline; coincidentally this is around the time that the embryo’s Hill radius exceeds the vortex size, but this is not a consistent outcome for all dust masses – this decline begins earlier with increasing ring mass, except for in the most massive ring (see figures 4.7 and 4.8). This can be explained by the fact that more massive rings generate a more massive planet, but the flat (averaged) accretion phase delays the point at which the decline in surface density offsets the increase in the planet’s Hill radius.

Vortex dynamics

Given the large fluctuations observed in the pebble accretion rate while the vortex is co-located with the planetary embryo, it is worth considering how to interpret this “vortex-assisted” accretion rate. These fluctuations arise from motion of the vortex relative to the embryo: their mutual interaction results in slight ($a\Delta\phi \lesssim 3$ AU) azimuthal oscillations in the vortex’s position relative to the embryo. This moves dust in and out of the embryo’s Hill sphere, and thus causes fluctuations in the pebble accretion rate.

An additional component may be due to the growth and destruction of the vortex: as the vortex

grows it exerts stronger torques on the disc, exchanging angular momentum with it, which leads to radial migration. However, once it reaches its maximum size, the supersonic flow creates a shock, and the energy lost in this process weakens the vortex, which in turn slows down its migration. If the vortex remains in the vicinity of embryo, the embryo continues to accrete, sourcing vortensity and thus allowing the vortex to grow again. If its migration doesn't slow down sufficiently, the vortex separates from the planet.

In this case the vortex-assisted pebble accretion rate should estimate the accretion rate the vortex can provide when at its limiting size, and the fluctuations should average to this value. There will be a short duration over which the accretion rate reaches this value while the vortex is growing from the planet's Hill radius up to $2H_g/3$, but the accretion rate will likely exceed this rate as the vortex briefly exceeds this size, after which it loses energy and thus decreases in size, causing the pebble accretion rate to fall below this value. This process repeats until the vortex finally separates from the planet. This seems consistent with the accretion rates shown in, for example, figures 4.8 and 4.9, when accounting for the missing Stokes number dependence in the analytic expression, which likely introduces a factor of ~ 0.5 – 0.7 , given the Stokes numbers of the largest grains are ~ 0.5 .

This then raises the issue of what drives the vortex to ultimately separate from the planet. One might expect the vortex to separate once the planet reaches its gap-opening mass, due to the sudden change in the torques experienced by the vortex, causing it to migrate. In this case, the vortex should separate earlier for higher ring masses, as the gap-opening mass should be reached faster. However, there does not appear to be a clear trend with ring mass, as I now describe, considering the $h = 0.035$ rings.

In the $50M_\oplus$ ring, the first vortex separates from the planet after 76 orbits, while the planet's mass is approximately $22M_\oplus$, and the second shortly after (82 orbits). This planet mass is comparable to the maximum mass reached in the simulations without thermal feedback (i.e. related to the isolation mass); the vortices still enter and cross the planet's Hill sphere until it has reached $\approx 26M_\oplus$, after which they become confined to horseshoe orbits, indicating that gap formation is underway. If we consider the thermal mass to be the mass at which the embryo's Bondi radius, $R_B = GM_P/c_s^2$, equals the gas scale height, we find a thermal (and therefore gap-opening) mass of $\approx 28M_\oplus$ – this is close to the mass at which the vortices are confined to horseshoe orbits.

In the $75M_\oplus$ ring, a small vortex separates at 40 orbits, when the planet has reached $21M_\oplus$, and the main vortex separates at 110 orbits, by which point the planet has reached $38M_\oplus$. The first vortex is already confined to horseshoe orbits before the second vortex separates, indicating that the vortex remains co-located with the planet even after the planet has reached the gap-opening mass. Note also that this is approximately the mass at which the embryo's Hill radius exceeds the vortex size, thus the second vortex is contained within the embryo's Hill sphere and is therefore gravitationally bound to it. During this phase the strength of the vortex decays until it ultimately separates from the planet.

In the $100M_\oplus$ ring, a vortex separates from the embryo at 29 orbits ($M_P \approx 20M_\oplus$), while pebble accretion and thus vortex formation at the planet location persist. This vortex re-enters the embryo's Hill sphere upon their next encounter at 58 orbits and merges with the newly-formed vortex. The combined vortex separates from the planet at 69 orbits, once the planet has reached $47M_\oplus$. Note that this is approximately equal to the mass at which Hill pebble accretion takes over from vortex-assisted, according to equation (4.8). In the $125M_\oplus$ and $150M_\oplus$ rings, the vortex separates after 101 orbits ($M_P = 73M_\oplus$) and 96 orbits ($M_P = 87M_\oplus$) respectively. In both of these cases the planet's Hill radius exceeds the vortex size, while the vortex is still co-located with the planet, thus the vortex is contained within the planet's Hill sphere. The vortex therefore remains bound to the planet beyond the planet's gap-opening mass, however, the pebble accretion rates may be sufficiently high that the Hill sphere encloses vortex faster than the vortex can separate. Clearly, gap-opening by the planet plays a role in the vortex's separation, however, performing simulations with finer spacing in dust mass may elucidate

4.4. Discussion

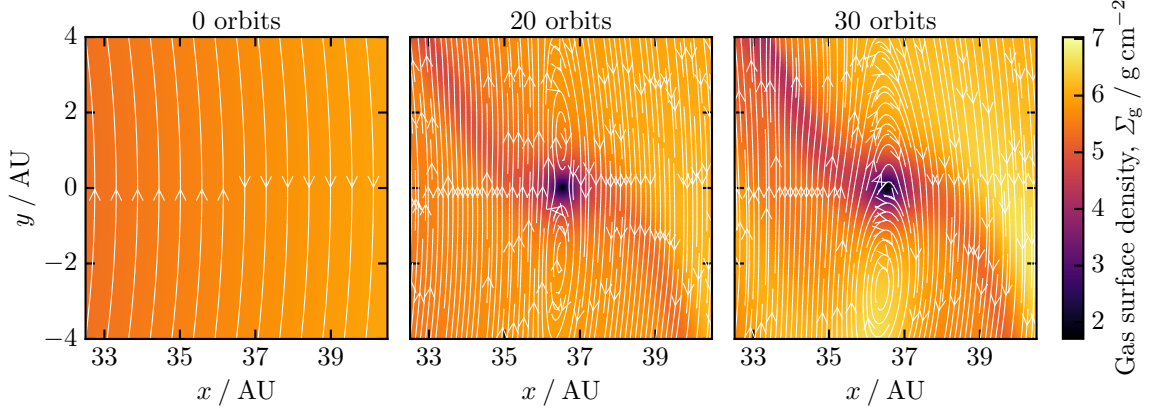


Figure 4.10: Gas surface density and velocity streamlines showing the evolution of the vortex structure around the planet, which is located at (36.5, 0). The snapshots are viewed in the frame rotating with the planet. Two vortices initially form, one ahead of and one behind the planet, which grow and merge to form one large vortex surrounding the planet.

this.

An interesting feature of the pebble accretion rates is that, in the rings with properties most conducive to strong thermal feedback, before the vortex-assisted pebble accretion phase, they are slightly reduced with respect to the corresponding simulations ignoring the thermal feedback. Without any accretion luminosity, the embryo’s location is a local density maximum (see e.g. figure 3.4). However, as its accretion luminosity forms a hotspot, in order for the gas in its vicinity to maintain hydrostatic equilibrium, the increase in temperature due to the hotspot is accompanied by a reduction in gas density. As vortensity generation depends on the gas surface density and pressure gradients, this local pressure structure is critical – the result is that there are two vortices that form, one ahead of and one behind planet. Being pressure maxima, these therefore slightly reduce the dust surface density in the embryo’s Hill sphere until they ultimately merge to form one large vortex surrounding the planet.

To demonstrate this process I performed a simulation in which the planet’s mass was set to and fixed at zero, and artificially introduced a hotspot around the planet, so that any feedback on the disc was purely due to the thermal structure and not the planet’s gravity. Figure 4.10 shows three snapshots of the gas surface density from this simulation, overlaid with the gas velocity streamlines. This shows the process of vortex formation, with vortices initially forming ahead of and behind the planet, which ultimately merge to form one large vortex. This explains the edge-case simulations (e.g. $h(a) = 0.05$, $100M_{\oplus}$ and $125M_{\oplus}$), in which vortices were formed but unable to produce the distinctive enhancement in pebble accretion rate, as the vortices were not large enough to merge and form a single vortex surrounding the planet. The embryo’s luminosity radius exceeded its Hill radius in each of the $h = 0.05$ discs, but only in the most massive ring did the vortices merge. While I have argued that $R_L > R_H$ for vortex formation, figure 4.5 suggests that $R_L \gtrsim 3R_H$ to trigger vortex-assisted pebble accretion. Closer inspection of the vortensity structure around the embryo as the luminosity radius grows suggests that the luminosity radius must extend to the centres of each of the vortices in order for them to merge.

The generation of vortensity in these two regions may also contribute to the oscillations of the vortex about the planet location: since vortensity is generated at either end of the single, merged vortex, as the smaller vortex structures become incorporated into the larger vortex, its morphology will likely vary. Clearly, the small-scale dynamics in the vicinity of the accreting embryo are complex, so this requires further investigation.

4.4.2 Pebble isolation masses in axisymmetric dust traps

In order to understand what sets the final masses for planetary embryos forming in pressure-generated dust rings, it is useful to first consider the final masses reached in the simulations ignoring the embryo's accretion luminosity, in order to isolate the effects of the thermal feedback on the disc. Based on the concept of the pebble isolation mass, one should expect the pebble accretion rate (without thermal feedback) to drop off once the planet reaches a critical mass, perturbing the disc and thus preventing further accretion. Formulations of the pebble isolation mass show a scaling with the cube of the aspect ratio, and the results obtained here do indeed show that larger planets form in higher aspect ratio discs, for all initial dust masses. However, there is also a dependence on the dust mass in the ring.

Bitsch et al. (2018) determined empirical scaling laws for the pebble isolation mass as functions of viscosity and the pressure gradient, finding

$$M_{\text{iso}} = 50M_{\oplus} \left(\frac{h}{0.05} \right)^3 \left[0.34 \left(\frac{-3}{\log_{10} \alpha} \right)^4 + 0.66 \right] \left[1 - \frac{1}{6} \left(\frac{\partial \log P}{\partial \log r} + 2.5 \right) \right]. \quad (4.9)$$

Note that I have scaled the pre-factor with the stellar mass used here (cf. equation 1.67), since the pebble isolation mass is expected to scale linearly with stellar mass (e.g. Liu et al., 2019b). For a ring with an aspect ratio of 0.035, and zero pressure gradient¹ (by the nature of it being a dust trap), the pebble isolation mass according to this formulation is $M_{\text{iso}} \approx 7M_{\oplus}$, while the thermal mass is $M_{\text{th}} \approx 85M_{\oplus}$. Neglecting thermal feedback, the embryo reaches just over $14M_{\oplus}$ in the least massive ring with this scale height, and approximately $23M_{\oplus}$ in the most massive ring.

The pebble isolation masses for aspect ratios of 0.05 and 0.1 are $22M_{\oplus}$ and $179M_{\oplus}$ respectively. In the $h(a) = 0.05$ rings, the final planet masses all exceed this isolation mass, while in the $h(a) = 0.1$ rings the planet cannot reach this mass as it exceeds the total mass of dust in the simulation. While figure 4.1 shows that more massive rings enable more massive planets to form, the pebble accretion rates (e.g. figure 4.2) show that accretion ceases earlier in more massive rings, as planets reach their isolation mass earlier, and the subsequent increase in planet mass is due to the finite time it takes to open the gap.

The situation is further complicated by the fact that even though the planet doesn't reach the pebble isolation mass predicted by equation (4.9) in the $h = 0.1$ rings, it does cease accretion. While there is a shallow dependence on ring mass for the smallest aspect ratios, it is almost linear for $h(a) = 0.1$. The fraction of dust mass accreted from the ring decreases with increasing dust mass, from 95% in the least massive ring to 71% in the most massive ring. This suggests that the accumulated torque on the disc from the planet throughout its growth is important, rather than there being a well-defined critical mass below which the planet can accrete, and beyond which no accretion can take place.

In general, larger scale heights result in a higher final planet mass due to the fact that the mass required for gap-opening increases with the disc aspect ratio. In the simple picture of pebble accretion without any thermal feedback on the disc, the planet will simply accrete until it begins to open a gap and pebbles can no longer drift into its orbit, becoming trapped at the pressure maximum at the outer gap edge. While this expectation is verified in my simulations, there is an additional consideration in that the planet is accreting in a dust trap, the width of which is set by drift-diffusion equilibrium. This width is independent of the disc aspect ratio. However, the width of the gap opened by the planet depends strongly on the aspect ratio, as the planetary wakes form shocks which deposit angular momentum further from the planet with increasing aspect ratio (Goodman & Rafikov, 2001). Each dust species becomes more strongly trapped as their Stokes numbers increase towards unity. With the gas pressure profile used in these simulations, the millimetre grains are trapped in a ring of full-width half-maximum

¹Although the simulations performed by Bitsch et al. (2018) were 3D, they considered a gas surface density profile $\Sigma_g \propto r^{1/2}$, which provides a local maximum in vertically-integrated pressure if $h \propto r^{1/4}$. Note also that they found differences in pebble isolation mass of up to a factor of 1.5 lower in 2D than in 3D simulations.

4.4. Discussion

Table 4.2: Gap full-width half-maxima after 60 kyr. All widths are in units of AU. Values in parentheses are for simulations without thermal feedback. Gap widths for the $h(a) = 0.1$ rings were not possible to measure due to their negligible or shallow depths.

		Ring mass / $M_{\oplus}$				
		50	75	100	125	150
$h(a)$	0.035	7.8 (6.8)	8.5 (7.2)	9.4 (7.4)	9.9 (7.5)	10.5 (7.6)
	0.05	9.4 (9.4)	10.0 (9.6)	9.7 (9.7)	10.3 (9.9)	12.2 (10.1)
	0.1	– (–)	– (–)	– (–)	– (–)	~ 15 (~ 15)

(FWHM) ≈ 7 AU, the 3 mm grains within 4 AU and the centimetre grains within 3 AU. Furthermore, the fact that the dust grains have an MRN size distribution means that most of the mass is in large grains, which are not only the most strongly trapped but also have the highest pebble accretion rate.

Table 4.2 lists the widths of the gaps formed by the planet in each simulation after 60 kyr, i.e. after the planets have all reached their final mass. The gap widths are reported as full-width at half maxima, where in each case the maximum is the difference between the top and bottom of the gap, and the full-width measured at the midpoint of these values. The “top” refers to the average of the local maxima either side of the gap, and the “bottom” refers to the average of the local minima immediately either side of the planet. The planets in the rings with an aspect ratio of 0.1 do not form gaps within the simulated time-scale. The planet in the most massive ring may eventually form a deep gap, but since it has ceased accreting by this point gap formation would have no impact on the results presented here. In the rings with aspect ratios of 0.035 and 0.05 each planet forms a gap: the least massive planet, i.e. that which formed in the $50M_{\oplus}$ ring, forms a gap that is comparable in width to the extent of the millimetre grains, but in all other rings the planet forms a gap that is wider than the dust ring. The planets can therefore grow beyond the classical pebble isolation mass as the mass within the dust trap already exceeds this, and the planet can accrete a significant fraction of this mass before it begins to repel material onto horseshoe orbits and dust can no longer enter its Hill sphere to be accreted.

As each of the planets in the $h = 0.035$ rings generated strong thermal feedback on the disc, these provide a means of understanding what role this plays in setting the final mass. As discussed above, the planet continues to accrete past the mass attained without thermal feedback, due to the presence of the vortex around the planet. The masses at which the accretion rate rapidly drops off (e.g. around 9 kyr in figure 4.3) also vary with the initial dust mass, as these correspond to the times at which the vortex separates from the planet. For the most massive ring, this occurs when the planet has reached $83M_{\oplus}$, and for the least massive ring this occurs when the planet has reached $22M_{\oplus}$. These significantly exceed the mass reached without thermal feedback, where the accretion rate begins to fall off once planet reaches the thermal mass, repelling material onto horseshoe orbits.

The gap widths are increased when including thermal feedback, though only because the planet is more massive. Torques from the planet dominate gap-opening, rather than torques from vortices. As evidenced by the $50M_{\oplus}$ and $75M_{\oplus}$ simulations, reaching the thermal mass appears to be important in the vortex separating from the planet. However, if the dust mass in the ring is high enough that the embryo can continue to accrete and source vortensity, and the aspect ratio is low enough that the embryo’s Hill radius can exceed the vortex size, the vortex remains bound to planet until the gap is opened and the vortex ultimately separates. The growth time-scale and gap-opening time-scale are important, and the growth history of the planet also plays a role: the $150M_{\oplus}$ ring produces a more massive planet than the $125M_{\oplus}$ ring, and the vortex separates from the planet at a higher mass.

Thermal feedback also plays a significant role in the planet’s growth in the $h = 0.05$, $150M_{\oplus}$ ring. The higher aspect ratio ring generates a more massive planet than in the $h = 0.035$ ring of the same mass, corroborating that larger scale heights result in higher masses, even when vortex assistance is

involved, as the vortex is larger and can thus accumulate dust from a larger distance. Furthermore, at larger scale heights, gap-opening requires a higher planet mass, the gap-opening time-scale is longer and the gap formed is wider.

4.5 Conclusions

In this chapter I have investigated the impact of a planetary embryo’s accretion luminosity on its growth via pebble accretion in an axisymmetric dust trap, for a range of dust masses and disc temperatures. I have shown that the process of vortex formation via the thermal feedback on the disc from the embryo’s accretion luminosity is robust, and can have a significant effect on the embryo’s growth. Under the most favourable conditions for strong thermal feedback, the planet’s mass can be increased by a factor of four with respect to the same initial conditions but neglecting the thermal feedback. This effect is most pronounced in dust traps located in cold regions of discs, such that their aspect ratios are $\lesssim 0.05$, otherwise large amounts of dust – in excess of $100M_{\oplus}$ – are required in the trap to have an impact.

The rings considered here as the formation sites of these planets were motivated by those observed in the DSHARP survey, with aspect ratios between 0.035 and 0.1, and dust masses between $50M_{\oplus}$ and $150M_{\oplus}$. In the discs with the highest aspect ratio, the background temperature is relatively high and the thermal feedback was therefore not strong enough to source vortensity sufficient for vortex formation, and thus had no impact on the embryo’s growth. In the 0.05 aspect ratio rings, only those with the highest mass provided an accretion rate sufficient to produce strong thermal feedback. Of the rings studied by Dullemond et al. (2018), only one had an estimated scale height as low as 0.05, which was also the least massive, thus this process is not likely to take place within those rings. Given the short time-scales on which the pebble accretion and vortex formation processes take place, one could plausibly speculate that this process has already taken place in some of the observed ringed discs, and that the resulting planet is as yet undetected, residing between the rings formed as a consequence. Of course, an unambiguous observational signature would be required to make this claim. The final state of the disc in my simulations is typically two rings and dust trapped in L_5 , which is due only to the process of planet formation, rather than vortex formation during the pebble accretion phase. The vortex formation phase does impact the size distribution of dust, however, with a significant depletion of centimetre-sized ($St \sim 0.5$) dust grains. Multi-wavelength observations of discs can provide constraints on the sizes of dust grains present, although these are still subject to the uncertainties in the physical properties of the dust grains. Similarly, both the maximum possible grain size (e.g. the fragmentation size) and their Stokes numbers are difficult to determine without knowledge of gas densities and turbulence levels.

In the discs with the lowest scale height considered, even the least massive ring provided a pebble accretion rate which led to vortex formation. This motivates further investigation into the ring mass at which the pebble accretion rate is no longer sufficient to lead to vortex formation.

I have also shown that planet formation in a pressure-enabled dust trap can lead to planets with masses in excess of the classical pebble isolation mass if the dust is confined to a ring more narrow than the gap opened by the planet. This is especially so for high pebble accretion rates, such that the growth time-scale is shorter than the gap-opening time-scale. If the radial extent of the dust within the ring is less than the gas pressure scale height, the entire dust distribution will be contained within the gap opened by the planet, since the spiral wake it drives shocks at a distance $\approx H_g$, creating a gap of order a few times the scale height. A long-lived pressure bump cannot be narrower in radial extent than $\sim H_g$, otherwise it cannot establish vertical hydrostatic equilibrium, and can also become Rossby-unstable (Ono et al., 2016; Dullemond et al., 2018). However, the radial extent of pebbles within a pressure bump is likely to be significantly narrower than the gas given their expected Stokes numbers. The rings analysed

4.5. Conclusions

in the DSHARP survey were estimated to have dust widths $w_d \sim 0.5H_g$ to $2H_g$, thus this effect could be important for consideration in planet formation models which may otherwise assume that pebble accretion stops at the classical isolation mass.

Regarding the prospect of the ring reforming once the vortex dissipates, my results suggest that even in the least massive rings, the planets produced trap dust released by the vortex in L_5 . Higher viscosities may mitigate this, as a more massive planet would be required. However, higher viscosities may also prevent vortex formation in the first place, as they would widen the dust trap, reducing the dust surface density and therefore the pebble accretion rate. Similarly, a higher viscosity has consequences for the initiation of the streaming instability to form the embryo: estimating the pebble scale height as $H_d/H_g \sim \sqrt{\alpha/St}$, the vertical extent of pebbles is greater for more turbulent discs, which may prevent the dust-to-gas ratio reaching the required value of unity. This therefore reinforces the need to include planetary migration in future simulations, as the interaction between the embryo and vortex could cause the planet to migrate away from the vortex, therefore allowing the ring to reform once the vortex dissipates.

By examining the gas velocities in the vicinity of the planet during the vortex formation phase, I have shown that the pressure structure generates a baroclinic term that results in the formation of two vortices – one ahead of and one behind the planet. Only if the accretion luminosity of the planet is sufficient for these two vortices to merge will the vortex-assisted pebble accretion phase proceed. I have hypothesised that the luminosity radius must be large enough for the vortices to merge, specifically that the luminosity radius must extend to the centres of these two vortices, rather than simply being larger than the embryo’s Hill radius. Further analysis is required to qualify this. Furthermore, I have attempted to determine the factors contributing to the dynamics of the vortex during its formation and while it is co-located with the planet, as well as what causes it to ultimately separate from planet. The embryo reaching the thermal mass can cause vortex migration, however, it is also around this mass that the embryo’s Hill radius exceeds the vortex’s radial extent, causing the vortex to become bound to the planet for an extended period. What ultimately drives the vortex to separate after this phase remains unclear, but may be possible to determine by performing simulations at a fixed aspect ratio with more closely-spaced ring masses.

Chapter 5

Shadowing in Warped Transition Discs

5.1 Introduction

I have thus far focussed on non-axisymmetric substructure observed in the thermal emission from large dust grains, and will now turn to the fact that the majority of discs with large-scale crescent-shaped features in their (sub-)millimetre emission do not show corresponding asymmetries in near-infrared (NIR) observations. Curiously, many of these discs actually show additional substructure at NIR wavelengths, such as spiral arms and brightness variations thought to result from self-shadowing due to misalignments – or warps – within the disc, which are not seen in continuum emission. Perhaps the most well-known example of this is the disc around HD 142527, a Herbig Ae/Be (F6IIIe, Waelkens et al., 1996) star, whose SED shows a large infrared excess in the far IR, but is suppressed at near-to-mid IR wavelengths, suggesting that this is a transition disc (Malfait et al., 1999).

Near-infrared imaging of HD 142527

HD 142527 was first imaged in near-IR (*H*-band, 1.6 μm and *K*-band, 2.2 μm) observations with the Coronagraphic Imager with Adaptive Optics (CIAO) on the Subaru Telescope by Fukagawa et al. (2006), who resolved a large cavity separating an inner and outer disc, confirming the transitional nature of this object. In fact, this is an extreme example of a transition disc, with a huge cavity spanning ~ 100 AU. They also noted that the centre of the cavity is not aligned with the position of the star, and that the outer disc is not continuous, but resembles two elliptical arcs facing each other, with a spiral arm extending to the north from the western arc. They suggested that the spiral arm hints at the presence of an unseen eccentric companion, or perhaps a recent stellar encounter. By studying the brightness of the light scattered from different regions of the disc, they proposed that the system is inclined by $\sim 30^\circ$, with the southwestern side being closest to us.

Mid-IR (18.8 μm and 24.5 μm) observations with Cooled Mid-Infrared Camera and Spectrometer (COMICS) on the Subaru Telescope were reported by Fujiwara et al. (2006). They detected a central compact component at both wavelengths, which they determined to be thermal radiation from hot dust close to the star; they inferred an inner disc extent of 80 AU. They did not detect the spiral arm reported by Fukagawa et al. (2006). They further explained the inclination and orientation of the system based on the brightness distribution in the outer disc, with back-scattering off the inner rim of the far (eastern) side, and forward scattering off the surface of the near (western) side, but provided no further constraint on the inclination angle.

Verhoeff et al. (2011) observed HD 142527 at multiple wavelengths using SEST (Swedish/ESO Submillimetre Telescope; 1.2 mm), ATCA (Australia Telescope Compact Array; 3.5 mm), VISIR (VLT Imager and Spectrometer for Mid-IR; *N*- and *Q*-bands). They determined the dust mass using the 1.2 mm flux to be $10^{-3}M_\odot$, suggesting a high total disc mass of $0.1M_\odot$ when assuming a global dust-to-gas mass ratio of 10^{-2} . VISIR imaging constrained the angular extent of the central component to be 50 AU in *N*-band and 30 AU in *Q*-band. Using their measured fluxes, which they supplemented with data from previous observations, they developed a model of the disc and, through radiative transfer simulations, tuned parameters such as the scale height, grain size and composition and cavity size to fit the measured SED. The result of their modelling suggested an inclination of 20° , an inner disc size of

5.1. Introduction

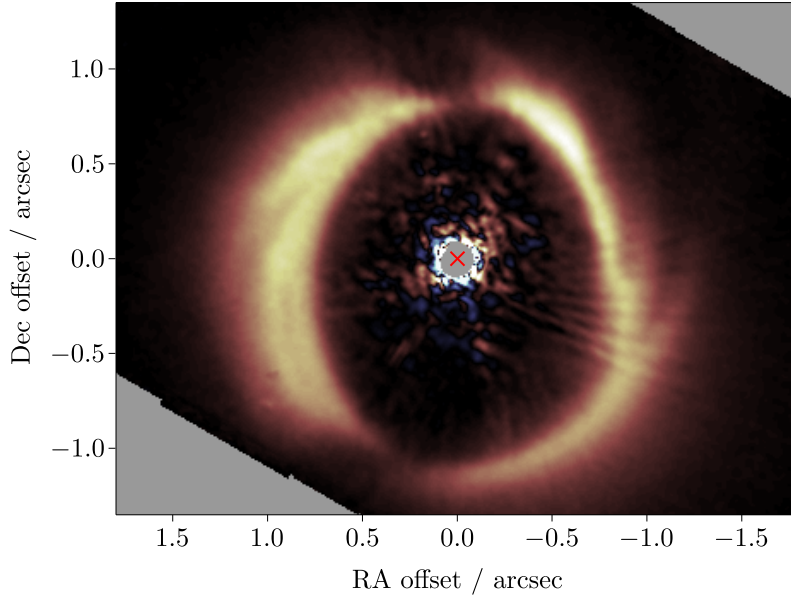


Figure 5.1: NaCo PDI H -band $P_{\perp}$ observation of HD 142527. North is up, east is left. The red cross marks the position of the central star. The colour scales linearly with brightness (blue -ve). Areas where no data is available due to saturation effects or the polarimetric mask are marked in grey. The disc is inclined with respect to the plane of the sky such that the eastern side of the disc is the far side. Drops in brightness are clearly visible around 0° and 160° east of north, which separate the outer disc into the two opposing arcs previously reported. A number of spiral arms on each side of the outer disc can also be seen. Adapted from Avenhaus et al. (2014) preprint.

30 AU and a cavity extending out to ~ 130 AU. They also suggested a very thin inner disc but a large scale height for the outer disc to account for the extremely high far-IR flux, since a high vertical wall would be able to reprocess a large fraction of the flux from the central star.

Subsequent IR adaptive optics observations with the NIR Coronagraphic Imager (NICI) on the Gemini telescope by Casassus et al. (2012) revealed multiple spiral arms on the western side of the disc, one of which coincided with that observed by Fukagawa et al. (2006). Their observations showed that the outer disc appears in fact to be composed of multiple overlapping spiral arms. They also considered the origin of the vast cavity; based on the significant accretion rate ($7 \times 10^{-8} M_{\odot} \text{ yr}^{-1}$, Garcia Lopez et al. 2006), they argued against photoevaporation of the disc being the cause, and thus proposed the existence of a giant planet within the cavity. They performed a 2D, gas-only hydrodynamics simulation of a disc modelled on HD 142527 containing a $10 M_{\text{Jup}}$ planet on a fixed, circular orbit at a semimajor axis of 90 AU, which they showed can carve a large cavity. However, the inner disc extended out to ~ 60 AU, which is inconsistent with observations. Furthermore, Rameau et al. (2012) used the NaCo instrument on the Very Large Telescope – a high contrast adaptive optics system and imager – to observe HD 142527 in angular differential imaging (ADI) mode in L' -band ($3.8 \mu\text{m}$). Their detection limits rule out the existence of a $10 M_{\text{Jup}}$ planet at around 90 AU.

Avenhaus et al. (2014) also observed HD 142527 with VLT/NaCo, using polarimetric differential imaging (PDI) at a spatial resolution of $0.1''$ (~ 15 AU, assuming a distance of 145 pc), in both the H -band (image shown in figure 5.1) and the K -band. They further resolved the two arcs of the outer disc into three spiral arms each, and confirmed the intensity nulls around 0° and 160° east of north. No emission was observed from within the inner $0.1''$, implying that the inner disc cannot extend beyond ~ 15 AU. Their modelling suggests an inclination of 23° , with a position angle of -2° .

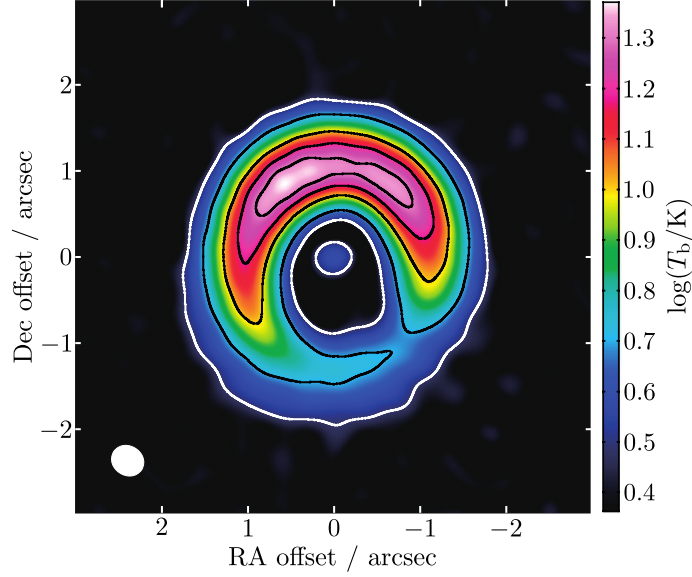


Figure 5.2: Submillimetre observation of HD 142527 (ALMA 0.89 mm). The colour scale traces the brightness temperature. The emission presents as a horseshoe shape, with emission peaking on the northern side of disc, and a lack of emission from the south. An unresolved inner disc can be seen, separated from the outer disc by a large cavity. There are no features corresponding to the nulls seen in near-IR observations. Adapted from Fukagawa et al. (2013) preprint.

Sub-millimetre imaging of HD 142527

Sub-millimetre (880 μm) observations with the Sub-Millimeter Array (SMA) by Ohashi (2008) revealed a large-scale crescent feature on the northern side of the disc, with a lack of emission from the southern side. They also showed that the peaks in the NIR emission were shifted southwards relative to the sub-millimetre emission peak; this is indicative that the asymmetry is due to dust trapping in a vortex, with the larger grains being more strongly concentrated. No counterpart to the spiral arm reported by Fukagawa et al. (2006) was observed, nor was any emission from close to the star. Their measured CO emission confirmed the orientation of the outer disc as proposed by Fukagawa et al. (2006).

HD 142527 was one of the first targets to be observed with ALMA. Observations by Casassus et al. (2013) and Fukagawa et al. (2013) in band 7 (890 μm), with spatial resolutions of ~ 70 AU and ~ 50 AU respectively, confirmed the highly non-axisymmetric continuum emission initially reported by Ohashi (2008), with emission from the south depleted by a factor ≈ 30 with respect to the north, as shown in figure 5.2. This is particularly intriguing as the northern intensity null seen in near-IR images coincides with the peak of this horseshoe structure, thus it cannot be due simply to a lack of material. They noted that the continuum emission is better described by a Gaussian than a power law profile, at all azimuthal angles. By fitting a Gaussian to the radial profile of the disc, they found an anti-correlation between the peak intensity and the deprojected distance from the star, thus ruling out the observed crescent resulting from eccentric orbits of the dust grains – if on eccentric orbits, they would concentrate around apastron. If this feature is instead due to dust trapping of grains in a vortex, one would hope to see a signature of this in emission from the gas.

Casassus et al. (2013) reported a reduction in dust surface density by factor of at least 300 between the peak in the outer disc and the cavity. They detected emission in the CO $J = 3 - 2$ (where J is the angular momentum quantum number; such transitions are hereafter referred to using the notation CO(3–2), for example) line within the cavity, and no such azimuthal asymmetry. Likely being optically thick, however, this traces the gas temperature rather than its density. They also presented an emission map of the $\text{HCO}^+(4-3)$ line, which shows an outer ring and a central component separated by a cavity,

5.1. Introduction

with filamentary structures connecting the two from opposing sides of the disc. These “streamers” are accompanied by non-Keplerian flows in the velocity channel maps (see figure 5.3 in the upcoming section), and were also marginally detected in the continuum emission. They interpreted these features as planet-induced, gap-crossing accretion flows, i.e. bridges between the inner and outer discs. The peak in continuum may be optically thick as it coincides with a decrease in HCO^+ emission.

Fukagawa et al. (2013) detected unresolved emission from the inner disc, suggesting the presence of large grains in the inner disc, but did not detect the filamentary structures reported by Casassus et al. (2013); it is worth noting that Fukagawa et al. (2013) had higher sensitivity and spatial resolution, and the filaments were aligned with the synthesized beam and so may have been artefacts. Fukagawa et al. (2013) also presented $^{13}\text{CO}(3-2)$ and $\text{C}^{18}\text{O}(3-2)$ line emission, neither of which showed any significant azimuthal asymmetry, rather they each presented as a ring surrounding a cavity, with intensity fluctuations less than a factor of 2 around the ring. Emission from both isotopologues were inferred to be optically thick in the outer disc. A tentative spiral arm was detected in the north-west, resembling that observed in near-IR observations, but farther out by $\approx 0.7''$. This could be an artefact of the peak of the dust horseshoe being optically thick and absorbing CO emission. No spiral arms were observed in continuum emission. The system inclination was determined to be $27 \pm 2^\circ$ based on the ^{13}CO channel maps. They estimated the disc mass to be $0.09M_\odot$, consistent with the Verhoeff et al. (2011) estimate of $0.1M_\odot$.

Further support for a dust trapping scenario came from Casassus et al. (2015b), who performed observations at 8.8 mm with ATCA. The crescent feature was shown to be more compact as a function of observation wavelength, suggesting that larger grains, responsible for longer wavelength emission, are concentrated more strongly. Similarly, Perez et al. (2015a) reported ALMA band 6 (1.2 mm) observations, and the continuum emission showed a more compact crescent than at band 7, but less concentrated than the 8.8 mm emission. The $\text{C}^{18}\text{O}(2-1)$ emission resembles the near-IR emission, but also features a minimum in the outer disc coinciding with that of the continuum emission, i.e. between the ends of the horseshoe. However, the emission doesn’t peak at the same location as the continuum, as might be expected for a dust-trapping vortex.

Detection of a companion: HD 142527B

Biller et al. (2012) reported the detection of a low-mass stellar companion ($0.1 - 0.4M_\odot$) at a separation of approximately 13 AU from the central star using high contrast imaging within $0.1''$ with VLT/NaCo. This claim was strengthened by the detection of $\text{H}\alpha$ emission – a signature of accretion, as hydrogen is shock-heated – at the expected location of HD 142527B by Close et al. (2014). Rodigas et al. (2014) observed this system with the Gemini Planet Imager (GPI) and detected a point source in polarised light close to the reported position of HD 142527B. Lacour et al. (2016) performed follow-up observations with both VLT/NaCo and GPI and, using the previously reported positions of HD 142527B (Biller et al., 2012; Close et al., 2014; Rodigas et al., 2014), attempted to determine the companion’s orbital parameters. The dataset covered two years and thus only a small arc of the companion’s orbit, but their fitting suggested one of two eccentric orbits, both of which lay in a plane inclined with respect to the outer disc. Balmer et al. (2022) provided the latest study of the companion’s orbit, with measurements up to 2018, finding that the companion is on an orbit significantly misaligned with respect to the inclination of the inner and outer discs.

Evidence for a warp

The first suggestion for a change in inclination throughout the disc, i.e. a warp, was by Casassus et al. (2013), who proposed that a misaligned inner disc might explain the high velocity radial flows observed in the HCO^+ kinematics. The twisted channel map, which is shown in figure 5.3, implies that the line-

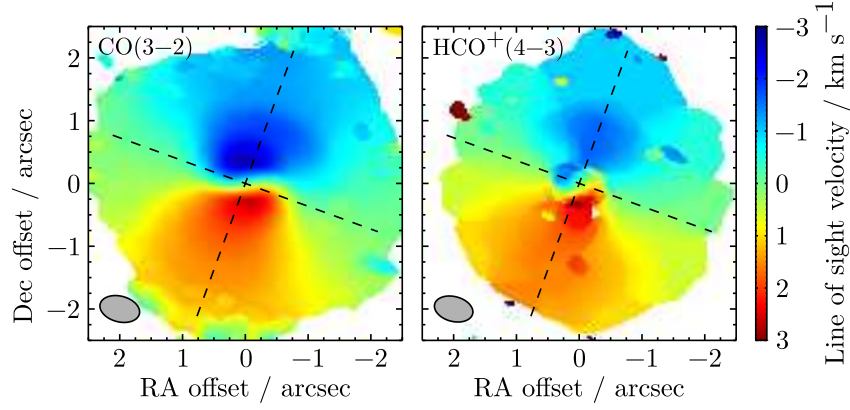


Figure 5.3: Intensity-weighted velocity maps of the HD 142527 $^{12}\text{CO}(3-2)$ (left) and $\text{HCO}^+(4-3)$ (right) line emission. The velocities are presented as relative to the systemic velocity of the system. The black dashed lines show the major and minor axes of the outer disc; the velocities appear twisted from the inner to outer disc. The grey ellipse in lower left corner is the beam. The $\text{CO}(3-2)$ has higher signal to noise. Adapted from Rosenfeld et al. (2014) preprint.

of-sight velocity deviates significantly from a single Keplerian velocity field. Similarly, Casassus et al. (2015a) performed ALMA band 9 (0.4 mm) observations and presented $\text{CO}(6-5)$ kinematics, which reveal a change in the disc position angle from the outer disc at -20° to 60° at $r \sim 0.25''$, and further to 110° inside $0.1''$. Rosenfeld et al. (2014), developed a framework for studying such fast radial inflows inferred from channel maps, and applied this to HD 142527. They modelled the CO and HCO^+ kinematics from the Casassus et al. (2013) observations with a warped disc, finding that this has the same observational signature as a radial inflow, since the line of sight velocity changes with radius.

Perhaps the strongest evidence for a warped disc came from Marino et al. (2015), who showed that the intensity nulls seen in the outer disc in scattered light (see figure 5.1) can be explained as the result of shadowing from an inclined inner disc. They modelled the dust density distribution in the disc, matching the outer disc inclination to observations, trialling inner disc inclinations that might reproduce shadows in the locations of the observed nulls. By performing radiative transfer simulations with the generated disc structure, they generated scattered light synthetic observations for various inner disc inclinations and were able to reproduce the general morphology and positions of the intensity nulls. They concluded that there is a mutual inclination of 70° between the inner and outer discs. Their best fit (by eye) to the Avenhaus et al. (2014) observation (figure 5.1) was with an inner disc position angle of -8° and an outer disc position angle of -20° , with the outer disc at an inclination of 24° and the inner disc at an inclination of 46° . The inclination of the inner disc was reduced to that of the outer disc over 10 AU to 15 AU from the central star, but no twist was considered.

Bohn et al. (2022) determined the inclinations and position angles of a sample of transition discs which had been previously suggested to contain misalignments. They used the GRAVITY instrument on the VLTI to measure the emission from the inner disc and modelled the K -band complex visibilities and closure phases to determine the inner disc orientations. For HD 142527, they found an inner disc inclination of 24° at a position angle of 15° and, refitting ALMA CO line emission, an outer disc inclination of 38° at a position angle of 163° , thus confirming the misalignment. These outer disc position angles and inclinations ($\text{PA} = 163^\circ$, $i = 38^\circ$) are somewhat consistent with those determined by Marino et al. (2015) ($\text{PA} = -20^\circ$, $i = 24^\circ$): the two transformations $i \rightarrow -i$ and $\text{PA} \rightarrow \text{PA} + 180^\circ$ together leave the orientation unchanged (but does reverse the direction of orbital motion), thus this appears to simply be a difference in convention. While the position angles are consistent ($163^\circ \approx -20^\circ + 180^\circ$), there appears to be contention between the inclination angles of 24° and 38° . This discrepancy is likely

5.1. Introduction

due to the difference in method used to obtain the inclination: there is a degeneracy between the disc scale height and inclination determined from near-IR images, since these measure light scattered from the surface of the disc, which varies with both scale height and inclination. Estimating the inclination from the gas kinematics requires constraining the emitting layers of relevant gas species, and thus intimate knowledge of the disc’s physical and chemical conditions. Furthermore, the poorly constrained stellar mass introduces uncertainty in the radial velocity measurements of the molecular lines.

While the exact inclination of the outer disc is of secondary interest, the inclination of the *inner* disc and the relative misalignment are critical in understanding the extent of the warp. Bohn et al. (2022) found a less inclined inner disc (24° vs. 46°) and a significant difference in its position angle (15° vs. -8°). The overall misalignment, i.e. the angle between the vectors normal to the inner and outer discs, is 59° , while the geometry proposed by Marino et al. (2015) gives a misalignment angle of 70° . The evidence for a warp is therefore strong, however, the exact details remain uncertain and dependent on the analysis method.

Such a large misalignment is surprising, and has largely been put down to the presence of the companion (e.g. Marino et al., 2015; Casassus et al., 2015a; Lacour et al., 2016; Owen & Lai, 2017). Regardless of its origin, the relative misalignment of the inner disc is an elegant explanation of how the scattered light and thermal emission observations can appear so different, with a null in scattered light coinciding with the peak in thermal emission. There, however, are more features observed in the near-IR images that do not appear in thermal emission that cannot be explained by this simple picture, namely the spiral arms. Particularly compelling is the fact that the spiral arms appear to originate from the intensity nulls. This is also the case for the disc around HD 100453 (Wagner et al., 2015).

Only two discs have had spiral arms observed in both NIR and sub-millimetre observations (e.g. Wagner et al. 2015; Rosotti et al. 2020a: HD 100453; Brown-Sevilla et al. 2021; Huang et al. 2018b: WaOph 6). In addition to showing spiral arms, HD 100453 also shows intensity nulls on opposite sides of the disc, which have been interpreted as shadows cast by an inclined inner disc (Benisty et al., 2017). Similar to HD 142527, the spiral arms start from the nulls. Can it then be pure coincidence that, in the two cases in which such features have been observed, the inner disc is misaligned such that it casts shadows on the points from which spiral arms originate? Perhaps the shadowing itself causes spiral arms to be launched, e.g. due to the temperature difference between shadowed and unshadowed regions (Montesinos et al., 2016). The issue with this hypothesis, however, is that a temperature drop is not observed in the optically CO emission, which traces the gas temperature. This has only been explored through hydrodynamics simulations, thus detailed modelling of the CO emission is required to further test this.

ALMA observations of HD 142527 by Christiaens et al. (2014) show tentative spiral arms in CO(2–1) and CO(3–2), albeit at low signal-to-noise, which are not aligned with the spiral arms seen in NIR observations. Furthermore, their fits to analytic models of spiral arms launched by planets were poor, and none can be linked to the sub-stellar companion detected by Biller et al. (2012). Estimates of the Toomre stability parameter find $Q \sim 2$, suggesting that the disc is marginally stable against gravitational collapse (Fukagawa et al., 2013; Christiaens et al., 2014). They therefore proposed that a recent close stellar encounter might be responsible for the spiral arms.

Marino et al. (2015) were able to reproduce the general morphology of the disc and the positions of the intensity nulls, but not the smaller scale details such as the brightness variations and spiral arms. Reproducing the brightness distribution is challenging as it depends on the dust grain size, composition and other physical properties which are difficult to constrain. However, following the hypothesis that the spiral arms are related to the intensity nulls, in that they result purely from shadowing due to changes in inclination throughout the disc, it should be possible to generate spiral arms by adding additional structure to the disc.

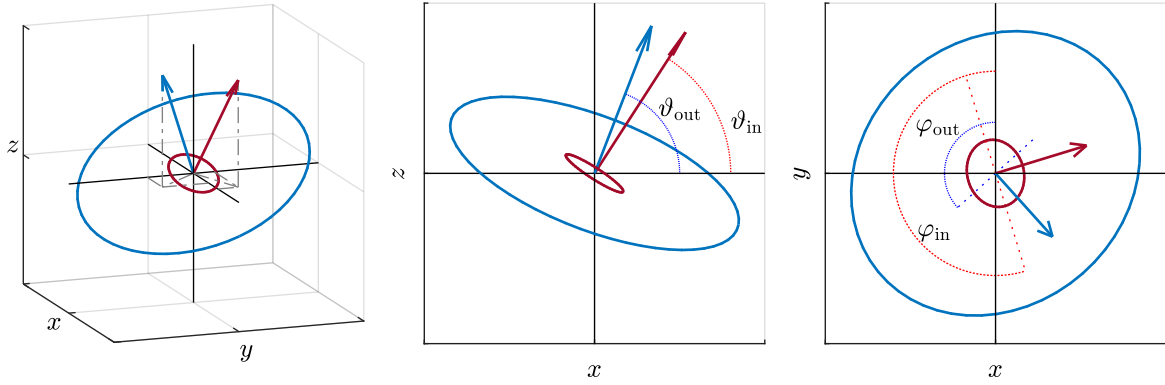


Figure 5.4: Schematic of a misaligned transition disc to demonstrate the inner and outer disc geometries. The left-hand panel shows the 3D orientations of the inner (red) and outer (blue) discs, with their respective normal vectors, as well as their x , y and z components and projection onto the x - y plane shown in dot-dashed grey. The middle panel shows the system's projection onto the x - z plane to illustrate their inclination angles, ϑ , marked with dotted arcs. The right-hand panel shows the projection onto x - y plane to illustrate their position angles: the blue and red dotted lines are parallel to each disc's major axis, and the position angles, φ , are marked with dotted arcs of the corresponding colours. Keplerian rotation is assumed to be in the anti-clockwise direction.

In this chapter I explore this in a more general context. Building on the model presented in Marino et al. (2015), of a mutual inclination between the inner and outer discs, I impose a warp and twist on the inner and outer discs and generate synthetic observations to understand whether the more complex morphology of the inner disc can produce additional substructure in the outer disc. I first present details of a warp and twist in a protoplanetary disc. I then outline the method used to set up static models of warped and twisted discs, including the parametrization of disc inclination as a function of radius. I then present results of radiative transfer simulations, with a comparison to observations.

5.2 Method

5.2.1 Warp and Twist

A warped disc has an inclination, and therefore an angular momentum vector, that varies as a function of its radial distance from the star. A twisted disc has the angular momentum vector also vary azimuthally. Figure 5.4 shows the inner and outermost annuli of a model warped and twisted disc to demonstrate the system geometry, and to illustrate the inclination and position angle definitions. A warp is simply a change in inclination angle, ϑ , while a twist is a change in position angle, φ , with cylindrical radius. A face-on disc has an inclination of zero. I adopt the convention that position angles are measured east of north towards the blueshifted part of the major axis of the disc. Following the convention that north points up the page and east points to the left, the position angle represents the angle measured anticlockwise from north to the axis along which the disc is inclined, projected onto the plane of the sky.

While 3D hydrodynamics simulations can be used to generate the density structure of a warped and twisted disc, these are extremely expensive to perform computationally, thus it is not feasible to do so to investigate a range of disc structures. I therefore use static models, imposing a warp and twist into the disc's dust density, and perform radiative transfer simulations to generate synthetic scattered light observations.

5.2. Method

5.2.2 Disc setup

The aim of this study is to investigate whether a warped and twisted disc can generate substructure similar to that seen in HD 142527 and HD 100453, namely the intensity nulls and spiral arms, through shadowing alone. Given the success of the Marino et al. (2015) model of HD 142527 in explaining the intensity nulls around 0° and 160° , I focus on this system and attempt to reproduce the additional substructure, specifically the spiral arms in the outer disc, as well as the intensity nulls. I use the geometry suggested by Marino et al. (2015) as a starting point, however, with so many properties of the disc unconstrained, as well as a desire to maintain a level of generality without fine-tuning the disc and dust properties, I use a fairly generic disc model and aim to reproduce the specific features rather than the exact observations themselves.

I only model the dust component of the disc, considering a single population of micron-sized dust grains – the dominant source of opacity at NIR wavelengths – to compare to scattered light observations. In order to keep the disc structure generic, I set up axisymmetric inner and outer discs separated by a cavity. I use a typical power-law background surface density profile,

$$\Sigma_b = \Sigma_{d,0} \left(\frac{r}{R_0} \right)^{-1}, \quad (5.1)$$

extending from 0.5 AU to 150 AU, and then impose a cavity to create a gap between the inner and outer discs via

$$\Sigma_d = \frac{\Sigma_b}{2} \left[1 + \operatorname{erf} \left(\frac{r - r_{\sigma_1}}{\sqrt{2}\sigma_1} \right) + \operatorname{erf} \left(\frac{r_{\sigma_2} - r}{\sqrt{2}\sigma_2} \right) + \operatorname{erf} \left(\frac{r - r_{\sigma_3}}{\sqrt{2}\sigma_3} \right) + \epsilon \right]. \quad (5.2)$$

The error functions provide smooth transitions between the inner/outer discs and the cavity, which must each have a finite (and sufficiently shallow) density gradient in order to be hydrodynamically stable. This is also for the benefit of the fidelity of the simulation, as sharp transitions between optically thin and thick regions can result in erroneous results if they are not sufficiently resolved. I set the widths of these transitions via $\sigma_1 = 0.07$ AU, $\sigma_2 = 0.7$ AU and $\sigma_3 = 2.8$ AU, and the transition radii are $r_{\sigma_1} = 0.8$ AU, $r_{\sigma_2} = 12$ AU and $r_{\sigma_3} = 110$ AU. The constant $\epsilon = 10^{-10}$ provides a small amount of dust in the cavity as well as at the very inner disc edge. While this provides a deeper cavity than likely exists in HD 142527, it ensures that any substructure generated due to shadowing arises only from the inner and outer discs. I adopt a total disc mass of $0.1M_\odot$ (Verhoeff et al., 2011) and therefore a total dust mass of $10^{-3}M_\odot$. I assume an MRN distribution of dust grain sizes and five logarithmically spaced size bins between $1 \mu\text{m}$ and 1cm , which yields a mass of approximately $2.5 \times 10^{-5}M_\odot$ in $1 \mu\text{m}$ grains. I therefore set $\Sigma_{d,0} = 7 \times 10^{-2} \text{ g cm}^{-2}$ and $R_0 = 10$ AU. The resulting surface density profile is depicted in figure 5.5.

Since I am primarily concerned with the distribution of micron-sized dust grains, which will be strongly coupled to the gas, I assume that the dust is vertically distributed following a typical flared disc profile set by the vertical hydrostatic equilibrium of the gas, with a scale height corresponding to $T \propto r^{-1/2}$. The scale height of the dust is therefore given by

$$H_d = h_0 R_0 \left(\frac{r}{R_0} \right)^{5/4}, \quad (5.3)$$

where I choose $h_0 = 0.05$. The dust density is thus calculated as

$$\rho_d = \frac{\Sigma_d}{\sqrt{2\pi} H_d}. \quad (5.4)$$

In order to impose a warp in the disc, I parametrize the inclination of each annulus comprising the

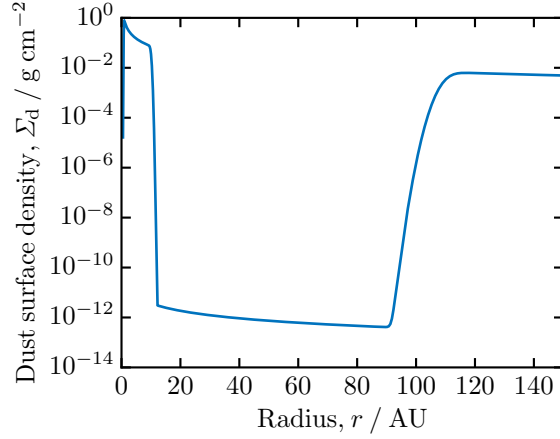


Figure 5.5: Surface density profile of the micron-sized dust grains in the model transition disc, with an inner disc extending to 12 AU, a cavity spanning ~ 100 AU and an outer disc extending out to 150 AU.

disc as a function of cylindrical radius from the star. The warp in the inner disc is parametrized as

$$\vartheta_{\text{in}}(r) = \begin{cases} \vartheta_{\text{in},0} & r_{\text{in},0} \leq r < r_{\vartheta_0} \\ \frac{\vartheta_{\text{in},1} - \vartheta_{\text{in},0}}{(r_{\text{in},1} - r_{\vartheta_0})^2} (r - r_{\vartheta_0})^2 + \vartheta_{\text{in},0} & r_{\vartheta_0} \leq r < r_{\text{in},1}, \end{cases} \quad (5.5)$$

such that the disc is at inclination $\vartheta_{\text{in},0}$ from $r_{\text{in},0}$ up to r_{ϑ_0} , after which the inclination varies quadratically in r from $\vartheta_{\text{in},0}$ to $\vartheta_{\text{in},1}$ between r_{ϑ_0} and the outer edge of the inner disc $r_{\text{in},1}$. Similarly, the outer disc is warped according to

$$\vartheta_{\text{out}}(r) = \begin{cases} \frac{\vartheta_{\text{out},0} - \vartheta_{\text{out},1}}{(r_{\vartheta_1} - r_{\text{out},0})^2} (r - r_{\vartheta_1})^2 + \vartheta_{\text{out},1} & r_{\text{out},0} < r \leq r_{\vartheta_1} \\ \vartheta_{\text{out},0} & r_{\vartheta_1} < r \leq r_{\text{out},1}, \end{cases} \quad (5.6)$$

such that its inclination varies quadratically from $\vartheta_{\text{out},0}$ to $\vartheta_{\text{out},1}$ from $r_{\text{out},0}$ to r_{ϑ_1} , after which it is fixed at $\vartheta_{\text{out},1}$. The relative inclination of 70° derived from observations thus corresponds to $\vartheta_{\text{in},0} - \vartheta_{\text{out},1}$. This is shown schematically in figure 5.6.

For the inner disc to create additional apparent substructure in the outer disc due to its shadow, it must cross the plane of the outer disc at multiple position angles, not just at the position angles responsible for the intensity nulls. Given the proposed extreme misalignment between the inner and outer discs of $\sim 70^\circ$, this requires the relative inclination to vary from $\sim 70^\circ$ to at least zero while the position angle also varies with radius, otherwise the top surface of the eastern side of the outer disc would be completely in shadow, rather than having distinct features at specific position angles. This change in position angle is achieved by twisting the disc. The twist is parametrized in an identical manner:

$$\varphi_{\text{in}}(r) = \begin{cases} \varphi_{\text{in},0} & r_{\text{in},0} \leq r < r_{\vartheta_0} \\ \frac{\varphi_{\text{in},1} - \varphi_{\text{in},0}}{(r_{\text{in},1} - r_{\vartheta_0})^2} (r - r_{\vartheta_0})^2 + \varphi_{\text{in},0} & r_{\vartheta_0} \leq r < r_{\text{in},1}, \end{cases} \quad (5.7)$$

$$\varphi_{\text{out}}(r) = \begin{cases} \frac{\varphi_{\text{out},0} - \varphi_{\text{out},1}}{(r_{\vartheta_1} - r_{\text{out},0})^2} (r - r_{\vartheta_1})^2 + \varphi_{\text{out},1} & r_{\text{out},0} < r \leq r_{\vartheta_1} \\ \varphi_{\text{out},0} & r_{\vartheta_1} < r \leq r_{\text{out},1}. \end{cases} \quad (5.8)$$

This provides a convenient means of varying the warp and twist throughout the disc, setting the regions over which the disc is warped and twisted, and by how much.

In order to investigate how the warped and twisted inner regions of the disc alter the shadows cast

5.2. Method

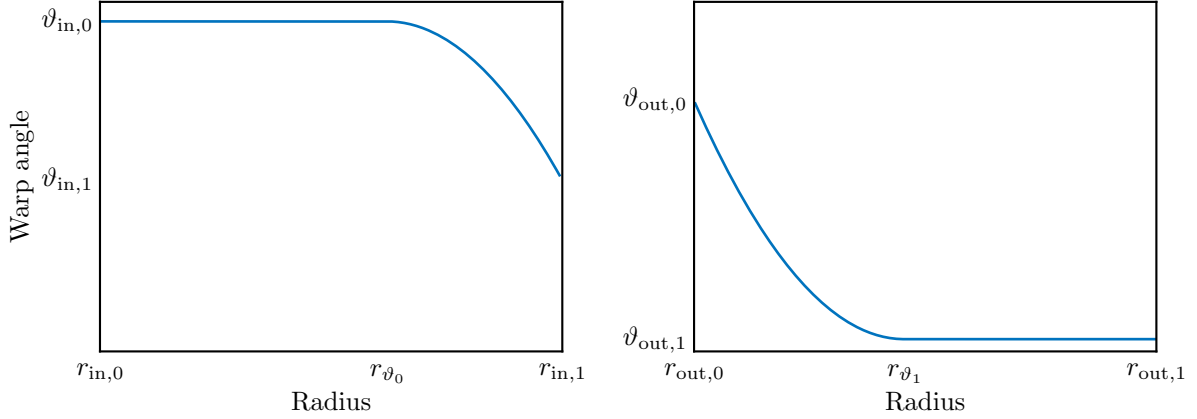


Figure 5.6: Parametrization of the inner and outer discs’ warp angle as a function of orbital distance from the star. The inner (outer) disc is unwarped interior (exterior) to r_{ϑ_0} (r_{ϑ_1}), after (before) which its inclination angle varies quadratically. The twist angle is parametrized using the same functional form.

on the outer disc, I ran a series of simulations in which I varied the warp and twist of the inner and outer discs, separately. My initial configuration is motivated by the disc geometry proposed by Marino et al. (2015), though having used a more general setup than theirs, and given the discrepant geometry reported by Bohn et al. (2022), I varied the inner and outer disc inclinations and position angles to find something that, by eye, matched the large scale morphology of the disc.

Starting with a purely misaligned inner disc, I introduced a warp to the outer edge of the inner disc according to equation (5.5), and the twist according to equation (5.7), in order to understand how altering the inner disc structure affects the shadows cast on the outer disc, while attempting to preserve the intensity nulls caused by the misalignment. I also investigated the possibility of a warped and twisted outer disc, with a misaligned inner disc, thus I introduced a warp to the inner edge of the outer disc according to equation (5.6), and a twist according to equation (5.8), in order to understand how altering the outer disc structure affects the observed substructure. In each case the warp and twist angles throughout the cavity are set to $\vartheta_{in,1}$ and $\varphi_{in,1}$ respectively, however the surface density is so low that the results are independent of the value used.

For each static disc structure generated, I performed a Monte Carlo radiative transfer simulation with RADMC-3D in order to generate a synthetic observation of the disc. In each case I used a spherical polar grid, with 576, 384 and 256 cells covering 0.5 AU to 150 AU in radius, 0 to 2π in azimuth and 0 to π in colatitude respectively. I used uniform, linear spacing in azimuth and colatitude. I used linear cell spacing in radius though with different resolutions in each section: 256 cells spanning the inner disc, 128 across the cavity and 192 for the outer disc. The central star was modelled on HD 142527, with a mass of $2M_{\odot}$ (Fukagawa et al., 2006; Verhoeff et al., 2011; Bohn et al., 2022), a radius of $3.8R_{\odot}$ (Verhoeff et al., 2011) and an effective temperature of 6300 K (Waelkens et al., 1996). The stellar spectrum is assumed to be that of a black body, and extends from 0.1 μm to 1 cm. Dust opacities and the full scattering matrix were calculated with BHMIE, using optical constants of the pyroxene $\text{Mg}_{0.8}\text{Fe}_{0.2}\text{SiO}_3$ provided in Jaeger et al. (1994); Dorschner et al. (1995). A dust material density of 3 g cm^{-3} was adopted. The optical constants span the wavelength range 0.2 μm to 300 μm and therefore don’t cover the full stellar spectrum; the opacities were therefore extrapolated to cover the range of the stellar spectrum. In order to generate a scattered light image, I set the dust temperature to zero everywhere in the disc so as to not produce any thermal emission. With the full scattering matrix available, I used anisotropic scattering, including the change in polarization of light from scattering events. I used 4×10^7 photon packets for the scattering Monte Carlo simulations and generated synthetic *H*-band images.

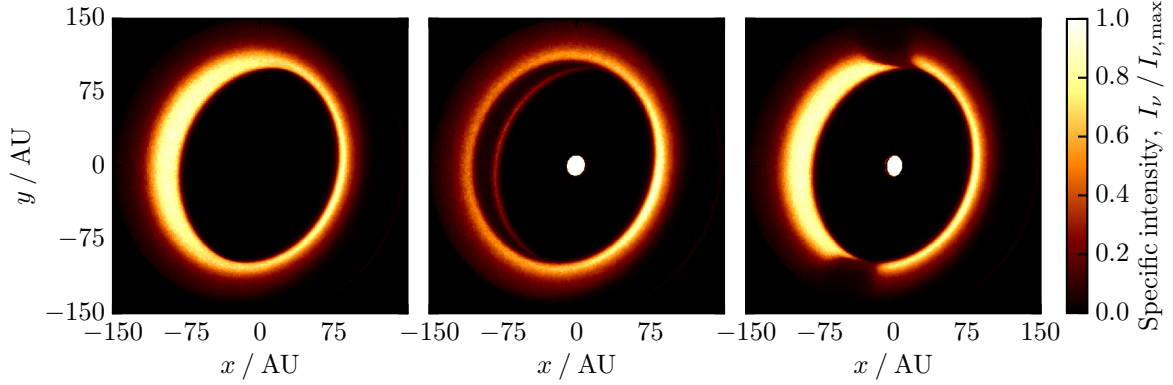


Figure 5.7: H -band synthetic observations of a model transition disc for three test cases. The left-hand panel shows the light scattered off the outer disc without the presence of the inner disc. The middle panel shows the scattered light when the inner and outer discs are aligned. The right-hand panel shows the scattered light when the inner disc is inclined by 43° at a position angle of -5° . In each case the outer disc is inclined by -27° at a position angle -17° . The colour scale in each image is normalised to the maximum in each image. In the middle and right-hand panels the inner disc is saturated to the maximum brightness of the outer disc.

5.3 Results

As a reference point, it is useful to see the effect of a purely misaligned inner disc with no continuous warp or twist, similar to that considered by Marino et al. (2015). Figure 5.7 shows the H -band synthetic observation for three test cases, demonstrating the effect of the inner disc's inclination on the light scattered off the outer disc. The system is inclined by -27° with respect to the plane of the sky at a position angle of -17° . If the inner disc is aligned with the outer disc, the wall of the outer disc on the eastern side is shadowed, and the western side of the disc appears brightest due to the strong forward-scattering properties of dust grains. Part of the back side of the outer disc remains visible in the east as the disc flares, causing the scale height of the outer disc to be larger than the shadow subtended by the inner disc. If the inner disc is misaligned with respect to the outer disc, however, it casts shadows on the outer disc at two distinct locations on opposite sides of the outer disc, with the exact locations dependent on the specific geometry. The right-hand panel shows the result of a misalignment of 70° as proposed by Marino et al. (2015). The inner disc is at an inclination of 43° at a position angle of -5° , which reproduces the two nulls at around 0° and 160° . As expected, the spiral arms in the outer disc are not reproduced.

Warp and twist in the inner disc

The effect of introducing a warp and twist to the outer edge of the inner disc can be seen in figure 5.8, which shows the result of warping and twisting the inner disc by increasing amounts, for increasing portions of the inner disc, according to equations (5.5) and (5.7), while keeping the outer disc at a fixed inclination and position angle. I varied the radial extent of the warp and twist by decreasing the distance at which they begin (r_{ϑ_0}) linearly from the inner disc's outer edge ($r_{\text{in},1}$) to its inner edge ($r_{\text{in},0}$), i.e. from 12 AU to 0.5 AU. For each value of r_{ϑ_0} , I varied $\vartheta_{\text{in},1}$ logarithmically from 43° to -27° , according to $\vartheta_{\text{in},1} = \vartheta_{\text{out},1} + \frac{70^\circ}{2^n}$ for integers $1 \leq n \leq 4$, as well as the limiting case of $n \rightarrow \infty$, where the warp meets the inclination of the outer disc ($n = 0$ corresponds to the case of a purely misaligned disc). I simultaneously varied $\varphi_{\text{in},1}$ logarithmically from -5° to -17° according to $\varphi_{\text{in},1} = \varphi_{\text{out},1} + \frac{12^\circ}{2^n}$ for integers $1 \leq n \leq 4$, and the limiting case of $n \rightarrow \infty$, where the twist meets the position angle of the

5.3. Results

outer disc (again, $n = 0$ corresponds to the case of a purely misaligned disc, with a discrete change in position angle). Each column shows the results of varying r_{ϑ_0} , and therefore the distance over which the inner disc is warped, while maintaining the warp and twist amplitudes ($\vartheta_{\text{in},1} - \vartheta_{\text{in},0}$ and $\varphi_{\text{in},1} - \varphi_{\text{in},0}$), and each row thus shows the results of changing the warp and twist amplitudes for a fixed r_{ϑ_0} .

The first column shows the scattered light images with $n = 1$, thus the inner disc inclination varies from $\vartheta_{\text{in},0} = 43^\circ$ to $\vartheta_{\text{in},1} = 8^\circ$, and the position angle varies from $\varphi_{\text{in},0} = -5^\circ$ to $\vartheta_{\text{in},1} = -11^\circ$. Moving down the column shows that, despite a larger portion of the inner disc becoming warped and twisted, the fact that the change in inclination is small (35°) compared to the initial misalignment (70°) means that only a small part of the outer region of the inner disc crosses the line of sight from the star to the outer disc. As such, there is very little change in the resulting image – only the widths of the outer regions of the nulls increase.

The first row shows the images with $r_{\vartheta_0} = r_{\text{in},1} = 12$ AU, where a change in inclination and position angle has no effect. However, as r_{ϑ_0} is moved inwards, closer towards the inner edge ($r_{\text{in},0}$) (corresponding to lower rows in figure 5.8), a larger fraction of the inner disc becomes warped and twisted, and the morphology of the outer disc begins to change, particularly on the eastern side. The general result is that a larger portion of the outer disc becomes shadowed: the front of the disc on the eastern side, and the back of the disc on the western side, are in shadow due to the large change in inclination of the inner disc. This is the true for increasing both the radial extent as well as the amplitude of the warp and twist. For the maximum warp and twist, where the inclination changes from 43° to -27° at a position angle of -5° to -17° across the whole inner disc, the eastern side of the disc appears as a crescent – this is simply the wall of the outer disc. Comparison with the middle panel of figure 5.7 shows that this is the crescent that is in shadow in that image, caused by the alignment of the inner disc.

Warp and twist in the outer disc

The effect of introducing a warp and twist to the inner edge of the outer disc can be seen in figure 5.9, which shows the result of warping and twisting the outer disc by increasing amounts, for increasing portions of the outer disc, according to equations (5.6) and (5.8). While keeping the inner disc unwarped, but inclined at 43° at a position angle of -5° , I varied the radial extent of the warp and twist by increasing the radius at which they end (r_{ϑ_1}) linearly from the outer disc inner edge ($r_{\text{out},0}$) to its outer edge ($r_{\text{out},1}$), i.e. from 105 AU to 150 AU. For each value of r_{ϑ_1} , I varied $\vartheta_{\text{out},0}$ logarithmically from -27° to 43° , according to $\vartheta_{\text{out},0} = \vartheta_{\text{in},0} - \frac{70^\circ}{2^n}$ for integers $1 \leq n \leq 4$, as well as the limiting case of $n \rightarrow \infty$, where the warp meets the inclination of the inner disc ($n = 0$ corresponds to the case of a purely misaligned disc). I simultaneously varied $\varphi_{\text{out},0}$ logarithmically from -17° to -5° according to $\varphi_{\text{out},0} = \varphi_{\text{in},0} - \frac{12^\circ}{2^n}$ for integers $1 \leq n \leq 4$, and the limiting case of $n \rightarrow \infty$, where the twist meets the position angle of the inner disc (again, $n = 0$ corresponds to the case of a purely misaligned disc, with a discrete change in position angle). Each column shows the results of varying r_{ϑ_1} , and therefore the distance over which the outer disc is warped, while maintaining the warp and twist amplitudes ($\vartheta_{\text{out},0} - \vartheta_{\text{out},1}$ and $\varphi_{\text{out},0} - \varphi_{\text{out},1}$), and each row thus shows the results of changing the warp and twist amplitudes for a fixed r_{ϑ_1} .

The first column shows the scattered light images with $n = 1$, thus the outer disc inclination varies from $\vartheta_{\text{out},0} = 8^\circ$ at its inner edge to $\vartheta_{\text{out},1} = -27^\circ$ at its outer edge, while the position angle varies from $\varphi_{\text{out},0} = -11^\circ$ to $\vartheta_{\text{out},1} = -17^\circ$. Moving down the column shows that as the warp becomes more extended over the outer disc, bringing it closer to being face-on, the inner wall begins to shadow the outermost regions on the eastern side, while more of the western side comes into view.

The first row shows the images with $r_{\vartheta_1} = r_{\text{out},0} = 105$ AU, thus the outer disc is all at the same inclination. As r_{ϑ_1} is moved further out towards 150 AU (corresponding to lower rows in figure 5.9), a greater proportion of the outer disc becomes warped and twisted, resulting in a significant change to

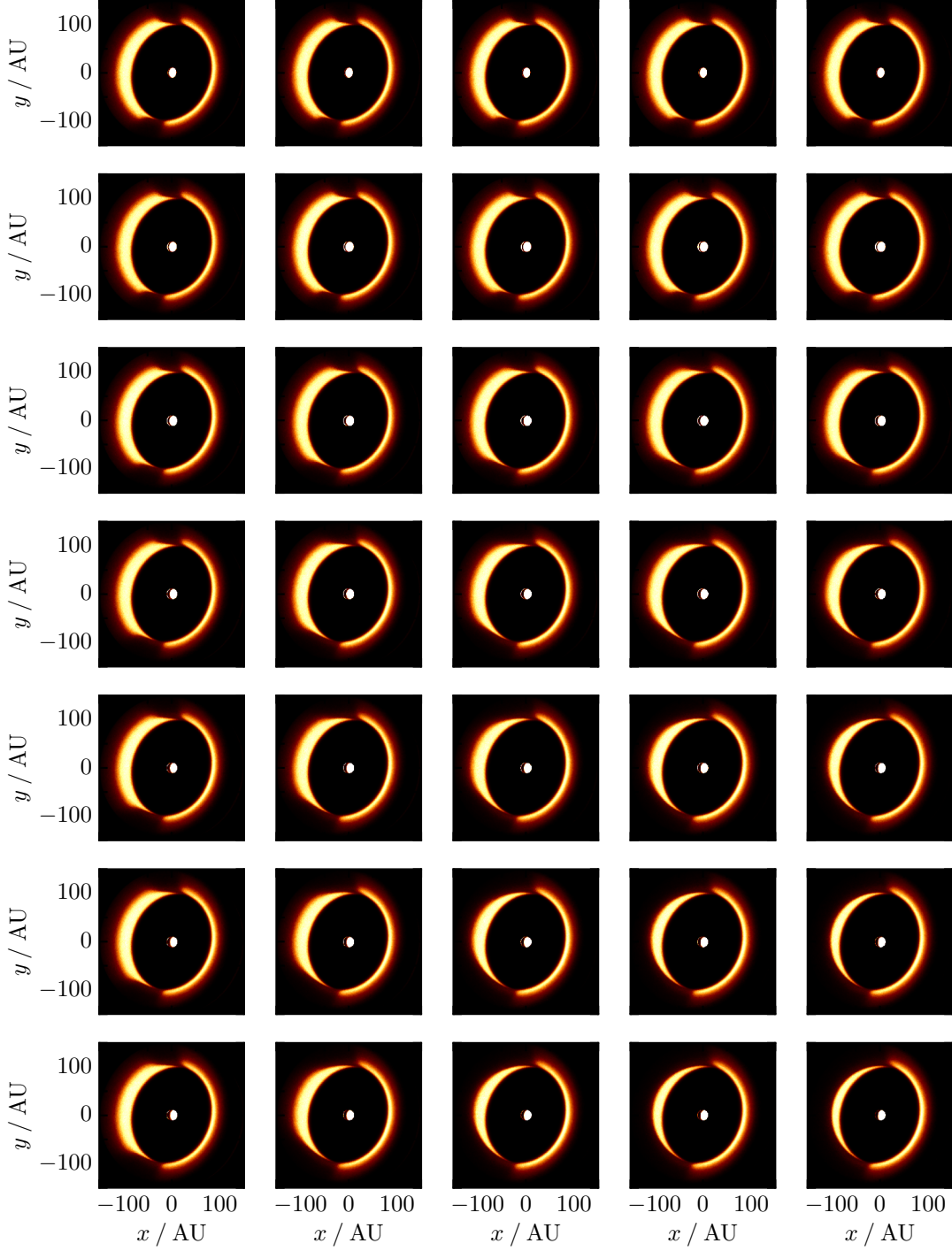


Figure 5.8: Synthetic observations of a transition disc with a warped and twisted inner disc. From top to bottom increases the amount of the inner disc that is warped and twisted (moving r_{ϑ_0} from $r_{\text{in},1}$ to $r_{\text{in},0}$). From left to right increases the amplitude of the warp and twist ($\vartheta_{\text{in},1}$ decreases from $\vartheta_{\text{out},1}$ to $\vartheta_{\text{in},0}$ and $\varphi_{\text{in},1}$ from $\varphi_{\text{out},1}$ to $\varphi_{\text{in},0}$) both by factors of 1/2, 3/4, 7/8, 15/16 and 1 times the maximum misalignment inferred from observations. The colour scales with specific intensity at $1.6 \mu\text{m}$, normalised to the maximum in the outer disc in each panel. The inner disc is saturated at that maximum value.

5.4. Discussion

the appearance of both the eastern and western sides of the outer disc. On the eastern side, the outer edge of the disc becomes increasingly less visible due to the fact that the inner edge is being raised in inclination and therefore shadowing the outer regions. The same effect occurs as the amplitude of the warp increases. The appearance of the western side also changes, as the inner wall on that side is being lowered, leaving more of the outer disc exposed to light from the central star.

In the middle rows, where the outer disc is warped and twisted between 105 AU and ~ 115 –135 AU, particularly for the largest changes in inclination and position angle, a radially narrow, dark band appears immediately behind the inner wall on the eastern side. This arises from the fact that the raised inner wall is shadowing the region immediately behind it, but the flared surface of the disc farther out is still visible. This feature disappears both as more of the outer disc becomes warped and as the amplitude of the warp increases, once all of the eastern side is shadowed by the wall (the outer edge of the disc remains fixed at an inclination of -27° , position angle -17°).

The nulls around 0° and 160° also change in appearance, losing their wedge-like shape and eventually just appear as part of the outer disc's inner wall, albeit with reduced brightness due to the shadow cast by the inner disc. The reduced misalignment between the inner disc and the outer edge of the inner disc causes the width of these shadows to increase.

An alternative warp profile

Neither a warp in the inner nor outer disc, using the parametrization described above, are capable of creating the appearance of spiral arms in the outer disc due to shadowing. Thus, with the improved understanding of how shadowing on the outer disc varies as the morphology of the inner disc changes, I also considered an alternative geometry for the inner disc, such that the warp and twist lie further into inner disc, i.e. closer to the central star, rather than at its outer edge. I used a profile similar to that described above, but instead with the change in inclination and position angle taking place over the central region of the inner disc, while the inner and outer regions are at constant inclination and position angle as a function of radius, resulting in a sigmoid shape. The warp profile for the inner disc and the resulting image are shown in figure 5.10. As before, the twist profile follows the same parametrization as that of the warp.

This produces two sharp features on each side of the outer disc. Since the innermost region of the inner disc is unchanged from the previous models, the nulls are reproduced at position angles 0° and 160° , albeit less prominently, and two further notches arise at around 50° and 220° . These latter features arise from the outer portion of the inner disc being at constant inclination and position angle. However, being at a lower inclination than the innermost region, it produces a shallower feature in the outer disc, i.e. more extended in azimuth, than the nulls at 0° and 160° , and the twist causes it to appear at a separate position angle. In the north-east, this gives the region between these two features (i.e. from 0° to 50°) an appearance somewhat resembling a spiral arm. However, there is a reduction in brightness between the two features, caused by shadowing from the warped region. From 50° to 160° , the central star has an unrestricted view of the outer disc, thus the wall is bright in scattered light. The effect on the south-western side is less pronounced, as we are viewing the back of that portion of the disc, thus the feature manifests simply as a reduction in brightness between 160° and 220° .

5.4 Discussion

The structure of the inner disc in transition discs can drastically affect the appearance of the outer disc when observed at near-infrared wavelengths, which trace scattered (and/or polarised) light. While previous studies have shown that a misalignment between the inner and outer discs can create radially

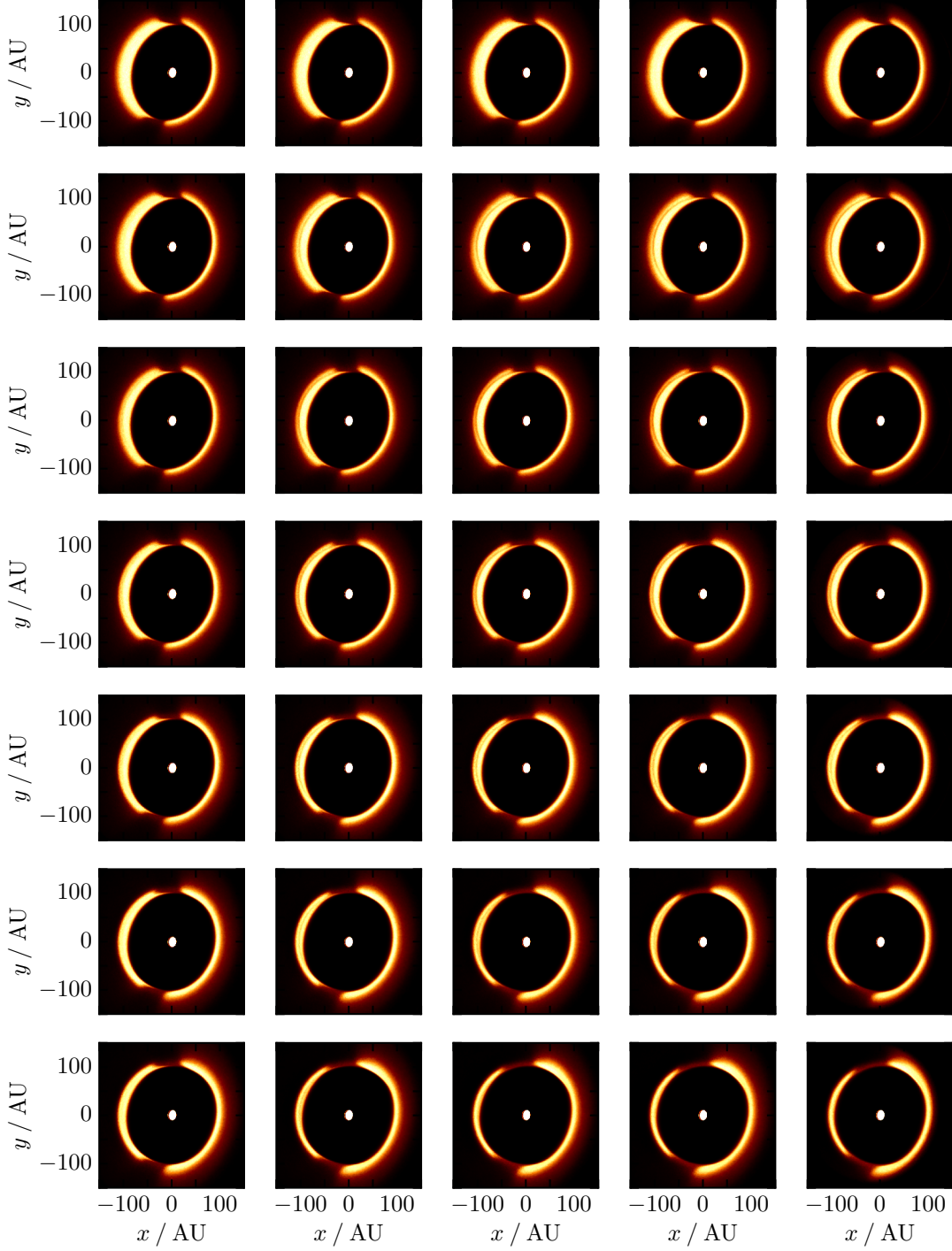


Figure 5.9: Synthetic observations of a transition disc with a warped and twisted outer disc. From top to bottom increases the amount of the outer disc that is warped and twisted (moving r_{ϑ_1} from $r_{\text{out},0}$ to $r_{\text{out},1}$). From left to right increases the amplitude of the warp and twist ($\vartheta_{\text{out},1}$ from $\vartheta_{\text{out},0}$ to $\vartheta_{\text{in},0}$ and $\varphi_{\text{out},0}$ from $\varphi_{\text{out},1}$ to $\varphi_{\text{in},0}$) both by factors of $1/2$, $3/4$, $7/8$, $15/16$ and 1 times the maximum misalignment inferred from observations. The colour scales with specific intensity at $1.6 \mu\text{m}$, normalised to the maximum in outer disc in each panel. The inner disc is saturated at that maximum value.

5.4. Discussion

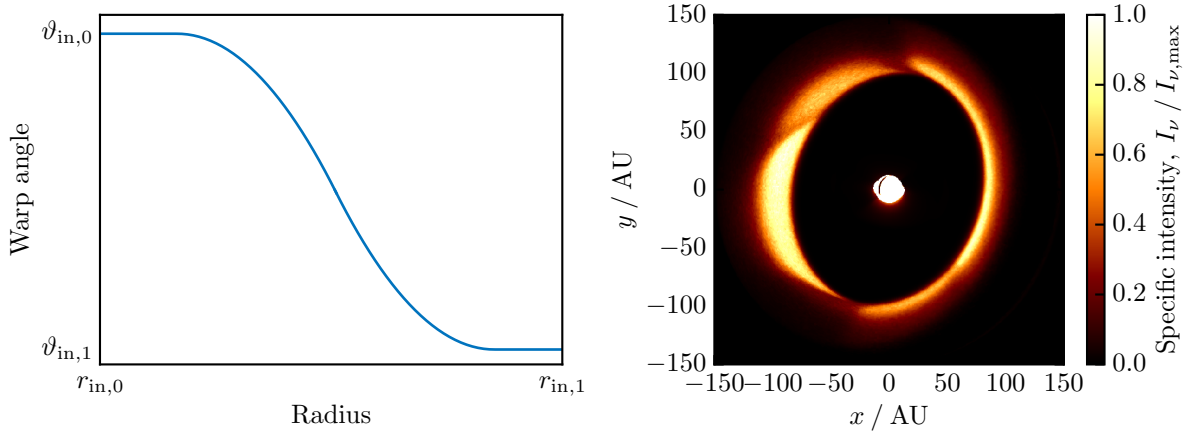


Figure 5.10: Synthetic observation (right) of a transition disc with a warped and twisted inner disc, parametrized by a sigmoid profile (left). The two regions of fixed inclination produce narrow shadows in the outer disc, separated by a region over which the brightness is reduced, caused by the warped and twisted section of the inner disc.

extended but azimuthally narrow dark features, where the inner disc intercepts rays from the central star that would otherwise reach these regions, these have typically considered the geometry of the inner disc as having a single inclination, where the inner and outer discs occupy distinct and separate planes. Having considered the possibility of warped and twisted discs, i.e. with continuous, radially varying inclinations, I have shown how the substructure and overall appearance of the outer disc varies as the inclination of the inner disc departs from a single value.

Assuming that the inner and outer discs, despite having different inclinations and position angles at their extrema, represent one continuous system, they must be connected by a warped and twisted region. If this region is contained entirely within the inner disc, the inclination will be matched at either edge of the cavity. Depending on the system geometry from a given observational location, this can result in the outer wall of the cavity being exposed and therefore being bright in scattered light, while the flared surface of the disc on that side is in shadow due to the large change in inclination of the inner disc. This also affects the width of the nulls, as the surface of the disc immediately either side of the null also begins to be shadowed.

If instead the warp and twist take place over the outer disc, the inner wall will become raised on one side of the disc and lowered on the other, causing the outermost regions to become either shadowed or further exposed depending on the direction of the warp. This can give rise to an interesting and potentially observable feature, where the wall shadows the region immediately exterior to it, but the flared disc profile enables the surface even farther out to be visible. For a large enough twist, this itself may give the appearance of spiral arms, as the shadow extends in radius as well as azimuth. Furthermore, this is likely to be emphasized in real discs, where the unshielded wall of outer disc can become “puffed-up” due to the increased flux, raising its temperature. This effect was not taken into account in my simple model for the density and scale height of the disc. Although originally considered at the inner rim of the inner disc (Natta et al., 2001; Isella & Natta, 2005), this has been shown to be relevant for discs with inner holes, i.e. lacking an inner disc (Dullemond et al., 2001), and will therefore be important in systems with a misaligned inner disc, where most of the outer disc is exposed. Dong (2015) showed that this effect fails to produce observable features in scattered light images, but with the additional contribution from a warp and twist, the features may become more pronounced.

Application to HD 142527

Whether the effects described above manifest as observable substructure is largely dependent on the viewing geometry, as well as the overall misalignment between the outer discs. It is therefore useful to apply this to the specific case of HD 142527, where these are well constrained. The large overall misalignment of 70° between the inner and outer discs, as well as the ~ 100 AU cavity, gives little information on how the structure of the disc varies between these two regions, however. While observations show that the cavity is largely devoid of both small and large dust grains, optically thick gas is present in the cavity, so the gas structure between the inner and outer discs remains uncertain.

A polynomial parametrization of the warp across the inner disc causes the surface of the disc on the eastern side to be completely shadowed, leaving just the inner edge of the outer disc visible in scattered light. Similarly, the nulls are best reproduced when the change in inclination over the inner disc is small. Together, these suggest that the inner disc is unlikely to contain a significant fraction of the warp. Given the wealth of infrared observations of HD 142527, the large-scale geometry of the outer disc is reasonably well understood, albeit subject to degeneracies in inclination and scale height.

An equivalent warp and twist in the outer disc changes the large-scale appearance of the outer disc due to the inner wall shadowing or revealing the outermost regions. This brings the outer disc closer to being face-on, thus making the cavity more circular, which is not in agreement with observations. Note however that I have explicitly assumed a circular disc (and therefore cavity) in my simple model for the disc, but it is possible that the cavity itself is eccentric given the mass of the resident companion. Indeed, the inner wall of the outer disc on the western certainly shows conspicuous changes in distance to the star. A warp in the outer disc also changes the appearance of the nulls, creating further discrepancy with the observations.

The overall effect of the twist is minimal as the difference in position angles between the inner and outer discs is small, and there is no reason to propose that the twist (and warp, for that matter) is non-monotonic. Furthermore, the twist can't be too large since, for such a large difference in inclination, a large twist amplitude would result in the disc wrapping around on itself, causing most of the outer disc to be shadowed.

This parametrization, for the range of parameters that I have considered, therefore fails to explain the spiral arms in the outer disc as an artefact of shadowing. An alternative profile, such as a sigmoid, offers some promise as it permits multiple sharp features to appear at different position angles, thus mimicking the appearance of spiral arms. However, this is only successful on the side of the disc exposed to the observer. In the case of HD 142527, where we see the back side of the disc in the west, the spiral arms on the inner wall are not visible. This is inconsistent with the observation of multiple spiral arms on the western side. It is possible that a warp in the outer disc could expose those features to the observer, however, proposing a significant warp in both the inner and outer discs is more difficult to justify physically.

The sigmoid warp and twist profile also introduces differences in brightness around the outer disc. Despite there being challenges in explaining brightnesses in scattered or polarised light images due to uncertainties in dust grain size, shape and composition, this is difficult to reconcile with observations, given that the disc appears brightest in the north, while this model predicts the opposite. A less spatially extended warp of the same amplitude would produce sharper features and a less-shadowed region in between, but again this is harder to justify.

If we assume that the inner and outer discs of the HD 142527 system are connected, as suggested by the detection of gas in the cavity, and that the warp and twist are continuous, we can therefore place a constraint on the extent of the warp. Firstly, the outer regions of the inner disc must be more-or-less at constant inclination, otherwise we would only see the inner wall of the outer disc, and not its surface,

5.4. Discussion

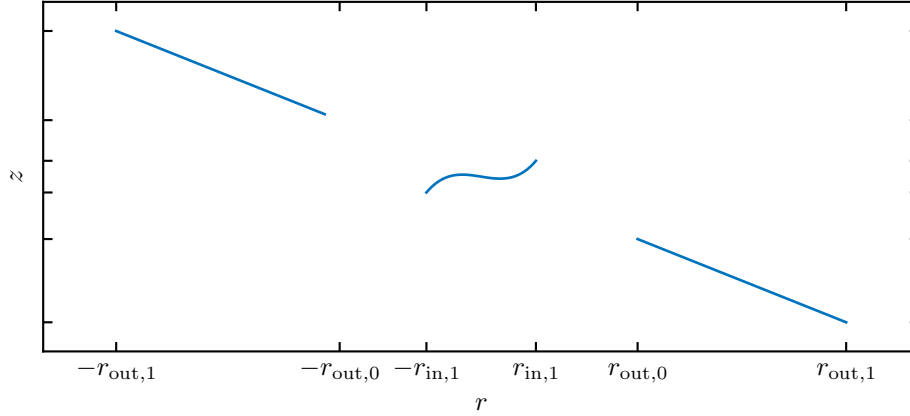


Figure 5.11: Warped disc cartoon showing the planes of the inner and outer discs as if viewed from edge-on. The innermost region of inner disc is aligned with outer disc, and stellar rotation axis. The inner disc is warped due to companion whose orbit is misaligned with respect to the outer disc, which changes the inclination of the outer regions of the inner disc. Not to scale.

on the eastern side. This then implies that the warp must extend across the cavity or be contained within the outer disc. Secondly, the fact that a warp in the outer disc shadows/exposes the outer regions uniformly, rather than creating multiple distinct shadows, as well as the effect this has on the shape of the nulls, appears to rule out this possibility. This therefore leaves the cavity as the most likely place for the change in inclination and position angle to occur. Furthermore, if this is the case, the spiral arms cannot be due to shadowing, as the dust density is likely too low to cast shadows on the outer disc.

The geometries considered above all result in shadows being cast on the inner wall of the outer disc, which is not visible from our point of view. Given that the western side of the image shows the back of the disc, i.e. we are looking towards the outer wall of the outer disc, it is difficult to see how a continuous, monotonic warp from the inner to outer disc can create observable shadows on that region. If substructure created by shadowing is to be seen on the western side, it must originate from part of inner disc that can obscure the front side. This must therefore originate from close to star, where a small difference in vertical structure can create potentially large features in the outer disc due to the angle subtended. Indeed, the observations do show features in the north-west, where the top surface of the disc is more exposed to the observer. We may therefore speculate that, to cast a shadow on both sides, the innermost region of inner disc must be warped and twisted away from the region responsible for the nulls.

A simple sketch of such a geometry is shown in figure 5.11. The physical reasoning for this setup is as follows: the outer disc angular momentum vector is aligned with that of the central star, and therefore the innermost region of the inner disc also shares this alignment due to torques from the magnetic field of the star (e.g. Armitage & Clarke, 1996). If the companion is misaligned with respect to this plane, it torques and therefore warps the outer edge of the inner disc. The misaligned outer regions of the inner disc therefore produce the nulls at 0° and 160° , while the innermost regions cast shadows which give the appearance of spiral arms. As shown by Balmer et al. (2022), the companion’s orbit is likely misaligned with respect to both the inner and outer discs. The outer disc may therefore also be warped in this case, however, given the results obtained above, it is unlikely that the amplitude of this warp is large.

This is potentially consistent with the mechanism proposed by Owen & Lai (2017), whereby the inclination of the inner and outer discs and the companion gives rise to a secular precession resonance. The slight misalignment in the system causes the inner disc to become strongly inclined, but as the innermost region of the inner disc is coupled to the star through its magnetic field, only the outer region

of the inner disc is warped; further investigation is required to confirm this scenario, however.

There is clearly a need to determine the geometry of transition discs driven by companions using 3D hydrodynamics simulations. Aside from the computational expense in performing such simulations, this is typically performed with smoothed-particle hydrodynamics codes due to the lack of a requirement for a fixed grid, which causes problems due to the high viscosities required when using this method. As discussed in §1.4.2, viscosity determines whether the warp propagates in the diffusive ($\alpha > H_g/r$) or wave-like ($\alpha < H_g/r$) regime. The necessarily high viscosities required in SPH simulations drive warp propagation in the diffusive regime, which may not be appropriate for protoplanetary discs. One interesting outcome of such simulations, however, is that the inner disc can break into separate annuli, each occupying a different orbital plane (e.g. Facchini et al., 2013, 2018; Zhu, 2019). If the inner disc of HD 142527 is broken into multiple planes, each with a different inclination and position angle, this could produce multiple narrow features in the outer disc, giving the appearance of spiral arms.

5.5 Conclusions

In this chapter I have investigated the substructure observed in scattered light images of warped and twisted transition discs. Motivated by the evidence for misalignments between the inner and outer regions of transition discs, which causes the inner disc to cast shadows on the outer, I have attempted to determine whether a warped and twisted inner disc can create the appearance of spiral arms in the outer disc. I have reproduced the result of Marino et al. (2015), showing that the proposed geometry of the HD 142527 disc can explain the nulls seen at near-infrared wavelengths. I have shown that a more complex disc structure, wherein the inner and outer discs are not just misaligned but have radially varying inclinations can further alter the appearance of the outer disc, but producing additional substructure by smoothly connecting the inclinations of the inner and outer discs is not possible.

Such modelling can, however, be used to constrain properties of the warp in discs such as HD 142527 and HD 100453. Application of my results to the HD 142527 disc suggest that the inner and outer discs are at a fixed inclination, and that if they are connected then the warp and twist must take place across the cavity. The origin of the spiral arms thus remains unclear. Given the considerable degree of freedom in choosing the extent of the warp, constrained only by the overall misalignment between the inner and outer discs, it is likely that one may be able to construct a warp profile that *can* produce the nulls and give the appearance of spiral arms, thus there is scope for exploring alternative warp profiles. However, these need to be physically-motivated and informed by reliable 3D simulations.

Such a warp profile must also be able to reproduce the observations at longer wavelengths, tracing the larger dust grains, and also the line emission from molecular gas species. Simulations informed by an improved understanding of the host, companion and disc parameters, rather than just modelling, are required to definitively rule the possibility of a warped disc. More extreme scenarios such as disc tearing may be able to produce apparent spiral arms through shadowing, but confirmation of this requires 3D hydrodynamics simulations.

Chapter 6

Summary

In this thesis I have investigated the connection between non-axisymmetric substructure and massive bodies within protoplanetary discs. Specifically, I have considered the role of accretion luminosity onto a planetary embryo in generating large-scale vortices, producing crescent features, as well as the warped and twisted structure of a transition disc, likely driven by a sub-stellar companion within the cavity, creating the appearance of spiral arms in the outer disc.

In chapter 3 I considered that the rings observed in protoplanetary discs, inferred to contain masses of millimetre-to-centimetre-sized dust, are prime locations for the formation of a planetary embryo. The properties of these pebbles that allow them to be efficiently trapped in rings also allows them to efficiently accrete onto a planetary embryo. I therefore argued that the pebble accretion process that should necessarily proceed will do so at such a rate that the accretion luminosity of the embryo should be considerable, and that the thermal feedback on the disc can make it unstable to the formation of vortices. I confirmed this using dust and gas hydrodynamics simulations, showing that a vortex forms around the planet, trapping dust and thereby producing an observable azimuthal asymmetry in the disc. Being limited by the scale height of the disc, however, the size of the asymmetry is less than those observed in transition discs, but is more comparable to the crescent features observed in HD 143006 and HD 163296 at high spatial resolution.

I also demonstrated that the vortex acts to increase the pebble accretion rate onto the embryo over their brief period co-location, by effectively increasing the accretion cross section. Thus, in aiming to explore a link between planet formation and non-axisymmetric substructure, I identified a new phase of pebble accretion, possible only if the accretion luminosity is sufficiently high to trigger vortex formation. Aware that this might therefore only occur in the most massive rings, such as those observed in transition discs, in chapter 4 I explored whether this process might occur in discs such as those observed in the DSHARP survey. These discs contain multiple narrow rings, each containing significant quantities of pebbles but typically less than in transition discs, thus providing comparable surface densities. I showed that the aspect ratio of the disc is a critical factor, with lower aspect ratios favouring vortex formation and that unless the available mass of pebbles is sufficiently high, the thermal feedback on the disc from the embryo's accretion luminosity is insufficient to trigger vortex formation. Nevertheless, when vortex formation does occur, it always acts to increase the planet's mass. I also substantiated how pebble accretion in dust rings can proceed past the classical pebble isolation mass if the width of the ring is less than the width of the gap opened by the planet.

In chapter 5 I investigated the possibility that transition discs inferred to contain large misalignments between their inner and outer discs are warped and twisted. By generating static models of transition discs containing a warp and twist, varying in both amplitude and radial extent, I performed synthetic scattered light observations of such systems, in order to determine the degree to which shadowing from the warped and twisted regions create the appearance of non-axisymmetric substructure. Applying this to the HD 142527 system, I showed that a simple, monotonic warp profile across the inner disc, reducing its inclination to that of the outer disc, results in the visible surface of the outer disc being shadowed, leaving only the cavity wall visible. Considering instead that the outer disc is warped and twisted, I showed that the resulting change in inclination changes the shape of the cavity as well as the widths of the nulls created by the inner disc. I explored alternative profiles for the warp and twist, demonstrating that this can indeed create the appearance of spiral arms in the outer disc, but whether such warp profiles

are physical requires verification. It is therefore compelling to consider future routes of investigation for this work.

6.1 Future Directions

Pebble accretion in dust rings

The number of possibilities for further exploration is a strong function of our ignorance of the physical processes at play in protoplanetary discs and inability to break degeneracies when interpreting observations. In exploring the outcome of pebble accretion in dust traps, in particular the consequence of the embryo’s accretion luminosity, I considered possible variations in dust mass and disc temperature, but viscosity is likely to play an equally important role in this. Firstly, the viscosity determines the width of the dust trap, and therefore the amount of pebbles available for accretion onto the embryo. Although a higher viscosity and therefore wider dust ring would reduce the pebble accretion rate and therefore the accretion luminosity, the resulting reduction in dust at the peak of the dust trap will increase the drift of pebbles from the edges of the ring towards the centre, providing a more sustained pebble accretion rate, perhaps sourcing vorticity over a longer time. Secondly, viscosity determines the mass at which gap-opening occurs, and as shown in chapter 4, gap-opening plays a role in the vortex’s separation from the planet. If a higher viscosity delays the vortex’s separation, while also reducing but prolonging the pebble accretion rate, this could ultimately produce massive planetary cores, provided the initial surface density is sufficient to trigger vortex formation.

It is also worth considering the width of the dust ring independently of viscosity. If we can determine the mechanism that creates the dust rings in the first place, the pressure structure and thus the ring structure can be constrained. In my simulations I adopted a gas surface density profile similar to that expected of a transition disc, with an inner cavity but power-law profile for outer disc. Although the width of the dust trap was calculated from the drift–diffusion equilibrium, this depends on the pressure structure. Rather than choosing an initial condition containing a pressure maximum, if we can instead self-consistently evolve the disc into a state containing dust rings, we can further understand the potential of this mechanism, and thus its importance in planet formation models.

One additional process to consider is that of planetary migration. Although previous works have shown that migration appears to be suppressed in pressure maxima (e.g. Morbidelli, 2020; Guilera et al., 2020, see also Johansen et al. 2019), these works did not consider the possibility of a planet–vortex interaction. In this case, the planet may migrate out of the pressure trap, in which case the vortex would no longer be within the planet’s coorbital region, which has two interesting consequences. The first is that once the vortex dissipates, the dust released will no longer accumulate in L_5 , but may form a ring. This will depend on the structure of the gap created by the planet, as it may now be a pressure minimum and thus an unstable equilibrium. If this is the case, then the dust will likely drift towards the newly created pressure maxima either side of the gap. If not, the re-formed ring may give rise to a new planetary embryo, potentially starting the process again. The second consequence is that this may impact the survival of the vortex. The vortex is continuously weakened by the planet on horseshoe turns, thus with the planet at a different orbital separation, this process will not occur. However, the vortex will now be subject to the stationary spiral shock pattern created by the planet, which will disrupt the vortex.

At this stage the process of vortex-assisted pebble accretion is not fully understood, both from a theoretical perspective and as seen in hydrodynamics simulations. It will be important to theoretically establish a Stokes number dependence of the pebble accretion rate in order to compare the rates observed in simulations with theoretical expectations, but also to determine the fate of pebbles entering the vortex,

6.2. Outlook

i.e. how the velocity structure in the vortex affects their accretion onto the embryo. This could be achieved by performing simulations in which the dust is treated as test particles, directly calculating their trajectory and therefore following them to the planetary core. Similarly, an improved pebble accretion algorithm that accounts for the velocity of dust fluid parcels may aid in resolving the attraction of pebbles towards either the planet or vortex core, rather than simply having pebbles accrete because they are in Hill sphere. Note that the pebble algorithm employed was conceived from the need to couple the pebble accretion rate to the local dust surface density and thus self-consistently calculate the accretion luminosity. Based on simulations performed by Owen & Kollmeier (2017), which I reproduced, I did not predict the extreme vortex–planet interactions that occur, thus I did not anticipate needing to resolve between dust trapping in vortex or ring vs accretion onto embryo.

Warped and twisted transition discs

As discussed in chapter 5, there is considerable freedom in choosing a parametrization for how the warp and twist vary across the disc, simply due to the fact that they cannot as yet be obtained from observations. It is therefore necessary to perform 3D hydrodynamics simulations to determine whether a given warp profile is physical. Although this is a computationally expensive and numerically challenging problem, this may be the only route to determine the extent to which a sub-stellar companion warps and twists the inner and outer discs between which it orbits.

A deeper understanding of the process(es) responsible for the misalignments these discs would also aid in interpreting these observations. For example, estimating the precession rate of the inner disc would allow us to confirm one mechanism over another, thus providing constraints on the warp profile.

It would be interesting to attempt to reproduce observations of other discs inferred to contain large misalignments, such as HD 100453, as well as discs which appear to show wide shadowed regions in their scattered light images, such as TW Hya (Roberge et al., 2005; Debes et al., 2017; Poteet et al., 2018; Rosenfeld et al., 2012), which would aid in understanding the two extremes of small versus large misalignments, and the processes that drive them.

6.2 Outlook

Just as *Spitzer* drastically improved our understanding of young stellar objects, as the SMA improved our understanding of their protoplanetary discs, and as ALMA continues to improve our understanding of the substructure within them, the inevitable improvements in observational facilities and techniques will improve our understanding of the physical processes at play in these planet-forming environments, helping to confirm or rule out many of the competing ideas that continue to be debated.

In the mean time, the high spatial resolution afforded by ALMA can help with some of the issues addressed through this thesis. For example, the transition discs IRS 48 and HD 152527, with their large-scale azimuthal asymmetries, are yet to be observed with the longest baselines, providing the highest spatial resolution. As such, the scale of the crescent-shaped features is smeared out by the large beam (IRS 48: $\sim 0.3''$ in band 9; HD 142527: $\sim 0.4''$ in band 7). As anticyclonic vortices should have a maximum size set by the gas scale height, spatially resolving these features would allow us to determine whether these are indeed vortices and thus further understand these systems. At higher spatial resolution, we only see smaller scale asymmetries such as HD 143006 and HD 163296, although few transition discs have been observed at such high resolution.

Similarly, deeper ALMA observations of discs inferred to contain misalignments would provide information about their dynamics; higher signal-to-noise data of optically thin gas species would help constrain their kinematics and potentially spatially resolve the warp and twist. As these systems are observed in

the future, we may be able to detect changes in the positions of shadows, potentially enabling us to determine precession rates for comparison to theory.

In closing, the work I have presented here is thus a contribution to the great and ever-changing endeavour to understand the formation of planets and planetary systems. The complex nature of protoplanetary discs and the enormous range of spatial scales and timescales over which physical processes occur makes this especially challenging, therefore enormous efforts in the theoretical, numerical and observational domains are required. But this is an effort to understand whether our Solar System, and therefore the planet on which we live, is unique, which I would argue is amongst the most worthwhile of human endeavours.

Bibliography

- ALMA Partnership et al., 2015a, ApJ, 808, L1
- ALMA Partnership et al., 2015b, ApJ, 808, L3
- Abod C. P., Simon J. B., Li R., Armitage P. J., Youdin A. N., Kretke K. A., 2019, ApJ, 883, 192
- Acke B., van den Ancker M. E., Dullemond C. P., 2005, A&A, 436, 209
- Adachi I., Hayashi C., Nakazawa K., 1976, Progress of Theoretical Physics, 56, 1756
- Adams F. C., Shu F. H., 1985, ApJ, 296, 655
- Adams F. C., Shu F. H., 1986, ApJ, 308, 836
- Adams F. C., Lada C. J., Shu F. H., 1987, ApJ, 312, 788
- Akiyama E., et al., 2016, AJ, 152, 222
- Alexander R. D., Armitage P. J., 2007, MNRAS, 375, 500
- Alexander R. D., Clarke C. J., Pringle J. E., 2006, MNRAS, 369, 229
- Andersson B. G., Lazarian A., Vaillancourt J. E., 2015, ARA&A, 53, 501
- Andrews S. M., Williams J. P., 2005, ApJ, 631, 1134
- Andrews S. M., Williams J. P., 2007, ApJ, 671, 1800
- Andrews S. M., Wilner D. J., Hughes A. M., Qi C., Dullemond C. P., 2009, ApJ, 700, 1502
- Andrews S. M., Wilner D. J., Hughes A. M., Qi C., Dullemond C. P., 2010, ApJ, 723, 1241
- Andrews S. M., Wilner D. J., Espaillat C., Hughes A. M., Dullemond C. P., McClure M. K., Qi C., Brown J. M., 2011, ApJ, 732, 42
- Andrews S. M., et al., 2016, ApJ, 820, L40
- Andrews S. M., et al., 2018, ApJ, 869, L41
- Ansdell M., et al., 2016, ApJ, 828, 46
- Armitage P. J., Clarke C. J., 1996, MNRAS, 280, 458
- Ataiee S., Pinilla P., Zsom A., Dullemond C. P., Dominik C., Ghanbari J., 2013, A&A, 553, L3
- Ataiee S., Baruteau C., Alibert Y., Benz W., 2018, A&A, 615, A110
- Auffinger J., Laibe G., 2018, MNRAS, 473, 796
- Avenhaus H., Quanz S. P., Schmid H. M., Meyer M. R., Garufi A., Wolf S., Dominik C., 2014, ApJ, 781, 87
- Avenhaus H., et al., 2018, ApJ, 863, 44
- Bai X.-N., 2016, ApJ, 821, 80
- Bai X.-N., Stone J. M., 2010, ApJ, 722, 1437

- Bai X.-N., Stone J. M., 2014, *ApJ*, 796, 31
- Balbus S. A., Hawley J. F., 1991, *ApJ*, 376, 214
- Balbus S. A., Hawley J. F., 1998, *Reviews of Modern Physics*, 70, 1
- Balmer W. O., et al., 2022, *AJ*, 164, 29
- Bardeen J. M., Petterson J. A., 1975, *ApJ*, 195, L65
- Barge P., Sommeria J., 1995, *A&A*, 295, L1
- Barranco J. A., 2009, *ApJ*, 691, 907
- Baruteau C., Zhu Z., 2016, *MNRAS*, 458, 3927
- Bate M. R., Bonnell I. A., Bromm V., 2003, *MNRAS*, 339, 577
- Batygin K., Adams F. C., 2013, *ApJ*, 778, 169
- Baumann T., Bitsch B., 2020, *A&A*, 637, A11
- Beckwith S., Zuckerman B., Skrutskie M. F., Dyck H. M., 1984, *ApJ*, 287, 793
- Beckwith S. V. W., Sargent A. I., Chini R. S., Guesten R., 1990, *AJ*, 99, 924
- Benisty M., et al., 2017, *A&A*, 597, A42
- Benítez-Llambay P., Masset F. S., 2016, *ApJS*, 223, 11
- Benítez-Llambay P., Pessah M. E., 2018, *ApJ*, 855, L28
- Benítez-Llambay P., Ramos X. S., Beaugé C., Masset F. S., 2016, *ApJ*, 826, 13
- Beuzit J.-L., et al., 2008, in McLean I. S., Casali M. M., eds, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 7014, Ground-based and Airborne Instrumentation for Astronomy II*. p. 701418, doi:10.1117/12.790120
- Biller B., et al., 2012, *ApJ*, 753, L38
- Birnstiel T., Ormel C. W., Dullemond C. P., 2011, *A&A*, 525, A11
- Birnstiel T., Klahr H., Ercolano B., 2012, *A&A*, 539, A148
- Birnstiel T., et al., 2018, *ApJ*, 869, L45
- Bitsch B., Morbidelli A., Johansen A., Lega E., Lambrechts M., Crida A., 2018, *A&A*, 612, A30
- Bjorkman J. E., Wood K., 2001, *ApJ*, 554, 615
- Blum J., 2018, *Space Sci. Rev.*, 214, 52
- Blum J., Wurm G., 2008, *ARA&A*, 46, 21
- Bohlin R. C., Savage B. D., Drake J. F., 1978, *ApJ*, 224, 132
- Bohn A. J., et al., 2022, *A&A*, 658, A183
- Bohren C. F., Huffman D. R., 1998, *Absorption and Scattering of Light by Small Particles*
- Booth R. A., Meru F., Lee M. H., Clarke C. J., 2018, *MNRAS*, 475, 167

BIBLIOGRAPHY

- Boss A. P., 1997, *Science*, 276, 1836
- Briceño C., Luhman K. L., Hartmann L., Stauffer J. R., Kirkpatrick J. D., 2002, *ApJ*, 580, 317
- Brown-Sevilla S. B., et al., 2021, *A&A*, 654, A35
- Brown J. M., Blake G. A., Qi C., Dullemond C. P., Wilner D. J., Williams J. P., 2009, *ApJ*, 704, 496
- Bruggeman D. A. G., 1935, *Annalen der Physik*, 416, 636
- Bryden G., Chen X., Lin D. N. C., Nelson R. P., Papaloizou J. C. B., 1999, *ApJ*, 514, 344
- Burgh E. B., McCandliss S. R., Feldman P. D., 2002, *ApJ*, 575, 240
- Calvet N., et al., 2005, *ApJ*, 630, L185
- Calzetti D., Bohlin R. C., Gordon K. D., Witt A. N., Bianchi L., 1995, *ApJ*, 446, L97
- Carrasco-González C., et al., 2019, *ApJ*, 883, 71
- Carrera D., Simon J. B., Li R., Kretke K. A., Klahr H., 2021, *AJ*, 161, 96
- Casassus S., Perez M. S., Jordán A., Ménard F., Cuadra J., Schreiber M. R., Hales A. S., Ercolano B., 2012, *ApJ*, 754, L31
- Casassus S., et al., 2013, *Nature*, 493, 191
- Casassus S., et al., 2015a, *ApJ*, 811, 92
- Casassus S., et al., 2015b, *ApJ*, 812, 126
- Cash J. R., Karp A. H., 1990, *ACM Trans. Math. Softw.*, 16, 201–222
- Castelli F., Kurucz R. L., 2003, in Piskunov N., Weiss W. W., Gray D. F., eds, *IAU Symposium Vol. 210, Modelling of Stellar Atmospheres*. p. A20 ([arXiv:astro-ph/0405087](#))
- Chiang E. I., Goldreich P., 1997, *ApJ*, 490, 368
- Christiaens V., Casassus S., Perez S., van der Plas G., Ménard F., 2014, *ApJ*, 785, L12
- Cieza L. A., et al., 2010, *ApJ*, 712, 925
- Cieza L. A., et al., 2019, *MNRAS*, 482, 698
- Cieza L. A., et al., 2021, *MNRAS*, 501, 2934
- Clarke C. J., Pringle J. E., 1988, *MNRAS*, 235, 365
- Clarke C. J., Gendrin A., Sotomayor M., 2001, *MNRAS*, 328, 485
- Close L. M., et al., 2014, *ApJ*, 781, L30
- Colella P., Woodward P. R., 1984, *Journal of Computational Physics*, 54, 174
- Courant R., Friedrichs K., Lewy H., 1928, *Mathematische Annalen*, 100, 32
- Courant R., Isaacson E., Rees M., 1952, *Communications on Pure and Applied Mathematics*, 5, 243
- Crida A., Morbidelli A., Masset F., 2006, *Icarus*, 181, 587
- Cuzzi J. N., Zahnle K. J., 2004, *ApJ*, 614, 490

- Cuzzi J. N., Dobrovolskis A. R., Champney J. M., 1993, *Icarus*, 106, 102
- D'Alessio P., Calvet N., Hartmann L., 2001, *ApJ*, 553, 321
- Debes J. H., et al., 2017, *ApJ*, 835, 205
- Dohnanyi J. S., 1969, *J. Geophys. Res.*, 74, 2531
- Dominik C., Tielens A. G. G. M., 1997, *ApJ*, 480, 647
- Dong R., 2015, *ApJ*, 810, 6
- Dong R., Rafikov R. R., Stone J. M., 2011, *ApJ*, 741, 57
- Dong R., et al., 2018a, *ApJ*, 860, 124
- Dong R., Najita J. R., Brittain S., 2018b, *ApJ*, 862, 103
- Dorschner J., Begemann B., Henning T., Jaeger C., Mutschke H., 1995, *A&A*, 300, 503
- Draine B. T., 2003, *ARA&A*, 41, 241
- Draine B. T., Li A., 2001, *ApJ*, 551, 807
- Drażkowska J., Dullemond C. P., 2014, *A&A*, 572, A78
- Drażkowska J., Windmark F., Dullemond C. P., 2013, *A&A*, 556, A37
- Dubrulle B., Morfill G., Sterzik M., 1995, *Icarus*, 114, 237
- Duffell P. C., 2015, *ApJ*, 807, L11
- Duffell P. C., MacFadyen A. I., 2012, *ApJ*, 755, 7
- Duffell P. C., MacFadyen A. I., 2013, *ApJ*, 769, 41
- Dullemond C. P., 2011, RADMC: A 2-D Continuum Radiative Transfer Tool, Astrophysics Source Code Library, record ascl:1108.016 (ascl:1108.016)
- Dullemond C. P., Penzlin A. B. T., 2018, *A&A*, 609, A50
- Dullemond C. P., Dominik C., Natta A., 2001, *ApJ*, 560, 957
- Dullemond C. P., Juhasz A., Pohl A., Sereshti F., Shetty R., Peters T., Commercon B., Flock M., 2012, RADMC-3D: A multi-purpose radiative transfer tool (ascl:1202.015)
- Dullemond C. P., et al., 2018, *ApJ*, 869, L46
- Epstein P. S., 1924, *Physical Review*, 23, 710
- Espaillat C., Calvet N., D'Alessio P., Hernández J., Qi C., Hartmann L., Furlan E., Watson D. M., 2007, *ApJ*, 670, L135
- Espaillat C., et al., 2010, *ApJ*, 717, 441
- Facchini S., Lodato G., Price D. J., 2013, *MNRAS*, 433, 2142
- Facchini S., Juhász A., Lodato G., 2018, *MNRAS*, 473, 4459
- Facchini S., et al., 2020, *A&A*, 639, A121

BIBLIOGRAPHY

- Fairlamb J. R., Oudmaijer R. D., Mendigutía I., Ilee J. D., van den Ancker M. E., 2015, *MNRAS*, 453, 976
- Fedele D., et al., 2017, *A&A*, 600, A72
- Flaherty K. M., Hughes A. M., Teague R., Simon J. B., Andrews S. M., Wilner D. J., 2018, *ApJ*, 856, 117
- Follette K. B., et al., 2013, *ApJ*, 767, 10
- Fortney J. J., Marley M. S., Barnes J. W., 2007, *ApJ*, 659, 1661
- Frank J., King A., Raine D. J., 2002, *Accretion Power in Astrophysics: Third Edition*
- Fromenteau S., Masset F. S., 2019, *MNRAS*, 485, 5035
- Fu W., Li H., Lubow S., Li S., Liang E., 2014, *ApJ*, 795, L39
- Fujiwara H., et al., 2006, *ApJ*, 644, L133
- Fukagawa M., Tamura M., Itoh Y., Kudo T., Imaeda Y., Oasa Y., Hayashi S. S., Hayashi M., 2006, *ApJ*, 636, L153
- Fukagawa M., et al., 2013, *PASJ*, 65, L14
- Fung J., Shi J.-M., Chiang E., 2014, *ApJ*, 782, 88
- Gammie C. F., 1996, *ApJ*, 457, 355
- Garcia Lopez R., Natta A., Testi L., Habart E., 2006, *A&A*, 459, 837
- Gingold R. A., Monaghan J. J., 1977, *MNRAS*, 181, 375
- Godon P., Livio M., 1999, *ApJ*, 523, 350
- Godon P., Livio M., 2000, *ApJ*, 537, 396
- Goldreich P., Lynden-Bell D., 1965, *MNRAS*, 130, 125
- Goldreich P., Tremaine S., 1979, *ApJ*, 233, 857
- Goldreich P., Ward W. R., 1973, *ApJ*, 183, 1051
- Goodman J., Rafikov R. R., 2001, *ApJ*, 552, 793
- Gordon K. D., 2004, in Witt A. N., Clayton G. C., Draine B. T., eds, *Astronomical Society of the Pacific Conference Series Vol. 309, Astrophysics of Dust*. p. 77 ([arXiv:astro-ph/0309709](https://arxiv.org/abs/astro-ph/0309709))
- Gordon K. D., Misselt K. A., Witt A. N., Clayton G. C., 2001, *ApJ*, 551, 269
- Grady C. A., et al., 2013, *ApJ*, 762, 48
- Grasdalen G. L., Strom S. E., Strom K. M., Capps R. W., Thompson D., Castelaz M., 1984, *ApJ*, 283, L57
- Guilera O. M., Sándor Z., 2017, *A&A*, 604, A10
- Guilera O. M., Sándor Z., Ronco M. P., Venturini J., Miller Bertolami M. M., 2020, *A&A*, 642, A140

- Guilera O. M., Miller Bertolami M. M., Masset F., Cuadra J., Venturini J., Ronco M. P., 2021, *MNRAS*, 507, 3638
- Guillot T., Burrows A., Hubbard W. B., Lunine J. I., Saumon D., 1996, *ApJ*, 459, L35
- Guillot T., Ida S., Ormel C. W., 2014, *A&A*, 572, A72
- Gundlach B., Blum J., 2015, *ApJ*, 798, 34
- Haisch Karl E. J., Lada E. A., Lada C. J., 2001, *ApJ*, 553, L153
- Hall C., et al., 2020, *ApJ*, 904, 148
- Hammer M., Kratter K. M., Lin M.-K., 2017, *MNRAS*, 466, 3533
- Hartmann L., Calvet N., Gullbring E., D'Alessio P., 1998, *ApJ*, 495, 385
- Hashimoto J., Muto T., Dong R., Liu H. B., van der Marel N., Francis L., Hasegawa Y., Tsukagoshi T., 2021, *ApJ*, 911, 5
- Hawley J. F., Smarr L. L., Wilson J. R., 1984, *ApJS*, 55, 211
- Hayashi C., 1981, *Progress of Theoretical Physics Supplement*, 70, 35
- Henning T., Stognienko R., 1996, *A&A*, 311, 291
- Henry L. G., Greenstein J. L., 1941, *ApJ*, 93, 70
- Hildebrand R. H., 1983, *QJRAS*, 24, 267
- Hori Y., Ikoma M., 2011, *MNRAS*, 416, 1419
- Huang J., et al., 2018a, *ApJ*, 852, 122
- Huang J., et al., 2018b, *ApJ*, 869, L43
- Huang J., et al., 2020, *ApJ*, 891, 48
- Hughes A. M., et al., 2009, *ApJ*, 698, 131
- Ida S., Guillot T., Morbidelli A., 2016, *A&A*, 591, A72
- Isella A., Natta A., 2005, *A&A*, 438, 899
- Isella A., Natta A., Wilner D., Carpenter J. M., Testi L., 2010, *ApJ*, 725, 1735
- Jaeger C., Mutschke H., Begemann B., Dorschner J., Henning T., 1994, *A&A*, 292, 641
- Johansen A., Klahr H., 2005, *ApJ*, 634, 1353
- Johansen A., Lacerda P., 2010, *MNRAS*, 404, 475
- Johansen A., Youdin A., 2007, *ApJ*, 662, 627
- Johansen A., Klahr H., Henning T., 2006, *ApJ*, 636, 1121
- Johansen A., Youdin A., Klahr H., 2009, *ApJ*, 697, 1269
- Johansen A., Kato M., Sano T., 2011, in Bonanno A., de Gouveia Dal Pino E., Kosovichev A. G., eds, Vol. 274, *Advances in Plasma Astrophysics*. pp 50–55, doi:10.1017/S1743921311006569

BIBLIOGRAPHY

- Johansen A., Mac Low M.-M., Lacerda P., Bizzarro M., 2015, *Science Advances*, 1, 1500109
- Johansen A., Ida S., Brasser R., 2019, *A&A*, 622, A202
- Joy A. H., 1945, *ApJ*, 102, 168
- Juhász A., Benisty M., Pohl A., Dullemond C. P., Dominik C., Paardekooper S. J., 2015, *MNRAS*, 451, 1147
- Kanagawa K. D., Muto T., Tanaka H., Tanigawa T., Takeuchi T., Tsukagoshi T., Momose M., 2015, *ApJ*, 806, L15
- Kant I., 1755, *Allgemeine Naturgeschichte und Theorie des Himmels*
- Kataoka A., et al., 2015, *ApJ*, 809, 78
- Kataoka A., Tsukagoshi T., Pohl A., Muto T., Nagai H., Stephens I. W., Tomisaka K., Momose M., 2017, *ApJ*, 844, L5
- Kenyon S. J., Hartmann L., 1987, *ApJ*, 323, 714
- Kerswell R. R., 2002, *Annual Review of Fluid Mechanics*, 34, 83
- Klahr H. H., Bodenheimer P., 2003, *ApJ*, 582, 869
- Klahr H. H., Henning T., 1997, *Icarus*, 128, 213
- Kley W., 1999, *MNRAS*, 303, 696
- Kokubo E., Ida S., 1998, *Icarus*, 131, 171
- Kretke K. A., Lin D. N. C., 2007, *ApJ*, 664, L55
- Kurtovic N. T., et al., 2021, *A&A*, 645, A139
- Kwon W., Looney L. W., Mundy L. G., 2011, *ApJ*, 741, 3
- L’Ecuyer P., 1988, *Commun. ACM*, 31, 742–751
- Lacour S., et al., 2016, *A&A*, 590, A90
- Lada C. J., Wilking B. A., 1984, *ApJ*, 287, 610
- Lambrechts M., Johansen A., 2012, *A&A*, 544, A32
- Lambrechts M., Johansen A., Morbidelli A., 2014, *A&A*, 572, A35
- Law C. J., et al., 2021a, *ApJS*, 257, 3
- Law C. J., et al., 2021b, *ApJS*, 257, 4
- Lazarian A., 2007, *J. Quant. Spec. Radiat. Transf.*, 106, 225
- Lesur G., Papaloizou J. C. B., 2009, *A&A*, 498, 1
- Li H., Finn J. M., Lovelace R. V. E., Colgate S. A., 2000, *ApJ*, 533, 1023
- Li H., Colgate S. A., Wendroff B., Liska R., 2001, *ApJ*, 551, 874
- Lin M.-K., 2012, *MNRAS*, 426, 3211

- Lin D. N. C., Papaloizou J., 1979, MNRAS, 186, 799
- Lin D. N. C., Papaloizou J., 1986a, ApJ, 307, 395
- Lin D. N. C., Papaloizou J., 1986b, ApJ, 309, 846
- Lin M.-K., Papaloizou J. C. B., 2010, MNRAS, 405, 1473
- Lin M.-K., Papaloizou J. C. B., 2011, MNRAS, 415, 1445
- Lin M.-K., Pierens A., 2018, MNRAS, 478, 575
- Lin C. C., Shu F. H., 1964, ApJ, 140, 646
- Lissauer J. J., Stevenson D. J., 2007, in Reipurth B., Jewitt D., Keil K., eds, Protostars and Planets V. p. 591
- Liu B., Ormel C. W., Johansen A., 2019a, A&A, 624, A114
- Liu B., Lambrechts M., Johansen A., Liu F., 2019b, A&A, 632, A7
- Long F., et al., 2018, ApJ, 869, 17
- Lorek S., Johansen A., 2022, arXiv e-prints, p. arXiv:2208.01902
- Lovelace R. V. E., Li H., Colgate S. A., Nelson A. F., 1999, ApJ, 513, 805
- Lucy L. B., 1977, AJ, 82, 1013
- Lucy L. B., 1999, A&A, 344, 282
- Lynden-Bell D., Pringle J. E., 1974, MNRAS, 168, 603
- Lyra W., Klahr H., 2011, A&A, 527, A138
- Lyra W., Lin M.-K., 2013, ApJ, 775, 17
- Lyra W., Mac Low M.-M., 2012, ApJ, 756, 62
- Lyra W., Johansen A., Klahr H., Piskunov N., 2008, A&A, 491, L41
- Lyra W., Johansen A., Zsom A., Klahr H., Piskunov N., 2009, A&A, 497, 869
- Lyra W., Raettig N., Klahr H., 2018, Research Notes of the American Astronomical Society, 2, 195
- Macías E., et al., 2018, ApJ, 865, 37
- Macintosh B., et al., 2006, in Ellerbroek B. L., Bonaccini Calia D., eds, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 6272, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series. p. 62720L, doi:10.1117/12.672430
- Malfait K., Waelkens C., Bouwman J., de Koter A., Waters L. B. F. M., 1999, A&A, 345, 181
- Mamajek E. E., 2009, in Usuda T., Tamura M., Ishii M., eds, American Institute of Physics Conference Series Vol. 1158, Exoplanets and Disks: Their Formation and Diversity. pp 3–10 (arXiv:0906.5011), doi:10.1063/1.3215910
- Manara C. F., Testi L., Natta A., Rosotti G., Benisty M., Ercolano B., Ricci L., 2014, A&A, 568, A18
- Marino S., Perez S., Casassus S., 2015, ApJ, 798, L44

BIBLIOGRAPHY

- Marquis de Laplace P.-S., 1798, Exposition du systeme du monde
- Masset F., 2000, A&AS, 141, 165
- Masset F. S., 2017, MNRAS, 472, 4204
- Masset F. S., Morbidelli A., Crida A., Ferreira J., 2006, ApJ, 642, 478
- Mathews G. S., Williams J. P., Ménard F., 2012, ApJ, 753, 59
- Mathis J. S., Rumpl W., Nordsieck K. H., 1977, ApJ, 217, 425
- Mayor M., Queloz D., 1995, Nature, 378, 355
- Meakin P., Donn B., 1988, ApJ, 329, L39
- Meheut H., Casse F., Varniere P., Tagger M., 2010, A&A, 516, A31
- Meheut H., Keppens R., Casse F., Benz W., 2012, A&A, 542, A9
- Mendoza V. E. E., 1968, ApJ, 151, 977
- Mendoza Villareal E. E., 1966, ApJ, 143, 1010
- Mie G., 1908, Annalen der Physik, 330, 377
- Mittal T., Chiang E., 2015, ApJ, 798, L25
- Miyake K., Nakagawa Y., 1993, Icarus, 106, 20
- Mizuno H., 1980, Progress of Theoretical Physics, 64, 544
- Montesinos M., Perez S., Casassus S., Marino S., Cuadra J., Christiaens V., 2016, ApJ, 823, L8
- Morbidelli A., 2020, A&A, 638, A1
- Morbidelli A., Nesvorny D., 2012, A&A, 546, A18
- Morbidelli A., Lambrechts M., Bitsch B., Jacobson S., 2015, in IAU General Assembly. p. 2248393
- Mori T., Kataoka A., 2021, ApJ, 908, 153
- Muro-Arena G. A., et al., 2020, A&A, 636, L4
- Murray C. D., 1994, Icarus, 112, 465
- Muto T., et al., 2012, ApJ, 748, L22
- Najita J. R., Strom S. E., Muzerolle J., 2007, MNRAS, 378, 369
- Nakagawa Y., Sekiya M., Hayashi C., 1986, Icarus, 67, 375
- Natta A., Prusti T., Neri R., Wooden D., Grinin V. P., Mannings V., 2001, A&A, 371, 186
- Natta A., Testi L., Calvet N., Henning T., Waters R., Wilner D., 2007, in Reipurth B., Jewitt D., Keil K., eds, Protostars and Planets V. p. 767 ([arXiv:astro-ph/0602041](https://arxiv.org/abs/astro-ph/0602041))
- Nelson R. P., Gressel O., Umurhan O. M., 2013, MNRAS, 435, 2610
- Neugebauer G., et al., 1984, ApJ, 278, L1

- Norman M. L., Winkler K. H. A., 1986, in Winkler K.-H. A., Norman M. L., eds, NATO Advanced Study Institute (ASI) Series C Vol. 188, NATO Advanced Study Institute (ASI) Series C. p. 187
- Öberg K. I., et al., 2021, *ApJS*, 257, 1
- Ogilvie G. I., Lubow S. H., 2002, *MNRAS*, 330, 950
- Ohashi N., 2008, *Ap&SS*, 313, 101
- Ohashi S., Kataoka A., 2019, *ApJ*, 886, 103
- Ohashi S., et al., 2020, *ApJ*, 900, 81
- Ohta Y., et al., 2016, *PASJ*, 68, 53
- Okuzumi S., Momose M., Sirono S.-i., Kobayashi H., Tanaka H., 2016, *ApJ*, 821, 82
- Ono T., Muto T., Takeuchi T., Nomura H., 2016, *ApJ*, 823, 84
- Ormel C. W., 2017, in Pessah M., Gressel O., eds, *Astrophysics and Space Science Library* Vol. 445, Formation, Evolution, and Dynamics of Young Solar Systems. p. 197, doi:10.1007/978-3-319-60609-5_7
- Ormel C. W., Klahr H. H., 2010, *A&A*, 520, A43
- Owen J. E., 2020, *MNRAS*, 495, 3160
- Owen J. E., Clarke C. J., 2012, *MNRAS*, 426, L96
- Owen J. E., Kollmeier J. A., 2017, *MNRAS*, 467, 3379
- Owen J. E., Kollmeier J. A., 2019, *MNRAS*, 487, 3702
- Owen J. E., Lai D., 2017, *MNRAS*, 469, 2834
- Paardekooper S.-J., Lesur G., Papaloizou J. C. B., 2010, *ApJ*, 725, 146
- Paardekooper S. J., Baruteau C., Kley W., 2011, *MNRAS*, 410, 293
- Paneque-Carreño T., et al., 2021, *ApJ*, 914, 88
- Papaloizou J., Lin D. N. C., 1984, *ApJ*, 285, 818
- Papaloizou J. C. B., Pringle J. E., 1983, *MNRAS*, 202, 1181
- Pascucci I., Wolf S., Steinacker J., Dullemond C. P., Henning T., Niccolini G., Woitke P., Lopez B., 2004, *A&A*, 417, 793
- Pérez L. M., et al., 2012, *ApJ*, 760, L17
- Pérez L. M., Isella A., Carpenter J. M., Chandler C. J., 2014, *ApJ*, 783, L13
- Perez S., et al., 2015a, *ApJ*, 798, 85
- Pérez L. M., et al., 2015b, *ApJ*, 813, 41
- Pérez S., Casassus S., Baruteau C., Dong R., Hales A., Cieza L., 2019, *AJ*, 158, 15
- Pinilla P., Birnstiel T., Ricci L., Dullemond C. P., Uribe A. L., Testi L., Natta A., 2012a, *A&A*, 538, A114

BIBLIOGRAPHY

- Pinilla P., Benisty M., Birnstiel T., 2012b, *A&A*, 545, A81
- Pinte C., Dent W. R. F., Ménard F., Hales A., Hill T., Cortes P., de Gregorio-Monsalvo I., 2016, *ApJ*, 816, 25
- Poblete P. P., Cuello N., Cuadra J., 2019, *MNRAS*, 489, 2204
- Pollack J. B., Hubickyj O., Bodenheimer P., Lissauer J. J., Podolak M., Greenzweig Y., 1996, *Icarus*, 124, 62
- Poteet C. A., et al., 2018, *ApJ*, 860, 115
- Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P., 1992, *Numerical recipes in FORTRAN. The art of scientific computing*
- Pringle J. E., 1981, *ARA&A*, 19, 137
- Raettig N., Klahr H., Lyra W., 2015, *ApJ*, 804, 35
- Rafikov R. R., 2002, *ApJ*, 572, 566
- Rafikov R. R., 2006, *ApJ*, 648, 666
- Ragusa E., Dipierro G., Lodato G., Laibe G., Price D. J., 2017, *MNRAS*, 464, 1449
- Ragusa E., Alexander R., Calcino J., Hirsh K., Price D. J., 2020, *MNRAS*, 499, 3362
- Rameau J., Chauvin G., Lagrange A. M., Thébault P., Milli J., Girard J. H., Bonnefoy M., 2012, *A&A*, 546, A24
- Rebull L. M., Wolff S. C., Strom S. E., 2004, *AJ*, 127, 1029
- Regály Z., 2020, *MNRAS*, 497, 5540
- Rice W. K. M., Armitage P. J., Wood K., Lodato G., 2006, *MNRAS*, 373, 1619
- Roberge A., Weinberger A. J., Malumuth E. M., 2005, *ApJ*, 622, 1171
- Rodenkirch P. J., Rometsch T., Dullemond C. P., Weber P., Kley W., 2021, *A&A*, 647, A174
- Rodigas T. J., Follette K. B., Weinberger A., Close L., Hines D. C., 2014, *ApJ*, 791, L37
- Rosenfeld K. A., et al., 2012, *ApJ*, 757, 129
- Rosenfeld K. A., Chiang E., Andrews S. M., 2014, *ApJ*, 782, 62
- Rosotti G. P., Juhasz A., Booth R. A., Clarke C. J., 2016, *MNRAS*, 459, 2790
- Rosotti G. P., et al., 2020a, *MNRAS*, 491, 1335
- Rosotti G. P., Teague R., Dullemond C., Booth R. A., Clarke C. J., 2020b, *MNRAS*, 495, 173
- Rybicki G. B., Lightman A. P., 1986, *Radiative Processes in Astrophysics*
- Safronov V. S., 1972, *Evolution of the protoplanetary cloud and formation of the earth and planets.*
- Salyk C., Herczeg G. J., Brown J. M., Blake G. A., Pontoppidan K. M., van Dishoeck E. F., 2013, *ApJ*, 769, 21
- Sándor Z., Regály Z., 2021, *MNRAS*, 503, L67

- Sargent A. I., Beckwith S., 1987, *ApJ*, 323, 294
- Schäfer U., Yang C.-C., Johansen A., 2017, *A&A*, 597, A69
- Shakura N. I., Sunyaev R. A., 1973, *A&A*, 24, 337
- Sierra A., et al., 2021, *ApJS*, 257, 14
- Simon J. B., Armitage P. J., Li R., Youdin A. N., 2016, *ApJ*, 822, 55
- Skrutskie M. F., Dutkevitch D., Strom S. E., Edwards S., Strom K. M., Shure M. A., 1990, *AJ*, 99, 1187
- Sod G. A., 1978, *Journal of Computational Physics*, 27, 1
- Speedie J., Booth R. A., Dong R., 2022, *ApJ*, 930, 40
- Steinacker J., Baes M., Gordon K. D., 2013, *ARA&A*, 51, 63
- Stevenson D. J., 1982, *Planet. Space Sci.*, 30, 755
- Stoll M. H. R., Kley W., 2014, *A&A*, 572, A77
- Stone J. M., Norman M. L., 1992, *ApJS*, 80, 753
- Strom K. M., Strom S. E., Edwards S., Cabrit S., Skrutskie M. F., 1989, *AJ*, 97, 1451
- Swedenborg E., 1734, *Opera Philosophica et Mineralia*
- Takeuchi T., Lin D. N. C., 2002, *ApJ*, 581, 1344
- Tang Y. W., Guilloteau S., Piétu V., Dutrey A., Ohashi N., Ho P. T. P., 2012, *A&A*, 547, A84
- Tanga P., Babiano A., Dubrulle B., Provenzale A., 1996, *Icarus*, 121, 158
- Tazzari M., et al., 2016, *A&A*, 588, A53
- Tazzari M., Clarke C. J., Testi L., Williams J. P., Facchini S., Manara C. F., Natta A., Rosotti G., 2021a, *MNRAS*, 506, 2804
- Tazzari M., et al., 2021b, *MNRAS*, 506, 5117
- Teague R., et al., 2016, *A&A*, 592, A49
- Teague R., Bae J., Bergin E. A., 2019, *Nature*, 574, 378
- Terry J. P., Hall C., Longarini C., Lodato G., Toci C., Veronesi B., Paneque-Carreño T., Pinte C., 2022, *MNRAS*, 510, 1671
- Thorngren D. P., Fortney J. J., Murray-Clay R. A., Lopez E. D., 2016, *ApJ*, 831, 64
- Toomre A., 1964, *ApJ*, 139, 1217
- Uribe A. L., Klahr H., Flock M., Henning T., 2011, *ApJ*, 736, 85
- Uyama T., et al., 2018, *AJ*, 156, 63
- Uyama T., et al., 2020, *AJ*, 159, 118
- Varnière P., Tagger M., 2006, *A&A*, 446, L13
- Verhoeff A. P., et al., 2011, *A&A*, 528, A91

BIBLIOGRAPHY

- Villenave M., et al., 2020, *A&A*, 642, A164
- Villenave M., et al., 2022, *ApJ*, 930, 11
- Voelkel O., Deienno R., Kretke K., Klahr H., 2021a, *A&A*, 645, A131
- Voelkel O., Deienno R., Kretke K., Klahr H., 2021b, *A&A*, 645, A132
- Von Neumann J., Richtmyer R. D., 1950, *Journal of Applied Physics*, 21, 232
- Waelkens C., et al., 1996, *A&A*, 315, L245
- Wagner K., Apai D., Kasper M., Robberto M., 2015, *ApJ*, 813, L2
- Walsh C., Juhász A., Meeus G., Dent W. R. F., Maud L. T., Aikawa Y., Millar T. J., Nomura H., 2016, *ApJ*, 831, 200
- Ward W. R., 1997, *Icarus*, 126, 261
- Warren S. G., Brandt R. E., 2008, *Journal of Geophysical Research (Atmospheres)*, 113, D14220
- Weidenschilling S. J., 1977a, *Ap&SS*, 51, 153
- Weidenschilling S. J., 1977b, *MNRAS*, 180, 57
- Weidenschilling S. J., 1980, *Icarus*, 44, 172
- Weidenschilling S. J., 1995, *Icarus*, 116, 433
- Weidenschilling S. J., Cuzzi J. N., 1993, in Levy E. H., Lunine J. I., eds, *Protostars and Planets III*. p. 1031
- Whipple F. L., 1972, in Elvius A., ed., *From Plasma to Planet*. p. 211
- Windmark F., Birnstiel T., Güttler C., Blum J., Dullemond C. P., Henning T., 2012a, *A&A*, 540, A73
- Windmark F., Birnstiel T., Ormel C. W., Dullemond C. P., 2012b, *A&A*, 544, L16
- Winn J. N., Fabrycky D. C., 2015, *ARA&A*, 53, 409
- Witt A. N., Petersohn J. K., Bohlin R. C., O’Connell R. W., Roberts M. S., Smith A. M., Stecher T. P., 1992, *ApJ*, 395, L5
- Wolszczan A., Frail D. A., 1992, *Nature*, 355, 145
- Yang C.-C., Zhu Z., 2020, *MNRAS*, 491, 4702
- Yang C.-C., Johansen A., Carrera D., 2017, *A&A*, 606, A80
- Youdin A. N., Goodman J., 2005, *ApJ*, 620, 459
- Youdin A. N., Lithwick Y., 2007, *Icarus*, 192, 588
- Zhang S., et al., 2018, *ApJ*, 869, L47
- Zhu Z., 2019, *MNRAS*, 483, 4221
- Zhu Z., Baruteau C., 2016, *MNRAS*, 458, 3918
- Zhu Z., Nelson R. P., Dong R., Espaillat C., Hartmann L., 2012, *ApJ*, 755, 6

- Zhu Z., Dong R., Stone J. M., Rafikov R. R., 2015, *ApJ*, 813, 88
- de Val-Borro M., et al., 2006, *MNRAS*, 370, 529
- de Val-Borro M., Artymowicz P., D’Angelo G., Peplinski A., 2007, *A&A*, 471, 1043
- van Leer B., 1977, *Journal of Computational Physics*, 23, 276
- van Leeuwen F., 2007, *A&A*, 474, 653
- van der Marel N., et al., 2013, *Science*, 340, 1199
- van der Marel N., Pinilla P., Tobin J., van Kempen T., Andrews S., Ricci L., Birnstiel T., 2015, *ApJ*, 810, L7
- van der Plas G., et al., 2017, *A&A*, 597, A32