Università degli Studi di Milano

PhD course in Physics, Astrophysics and Applied Physics

Substructures in Protoplanetary Discs
Dust and Gas Dynamics

Supervisor:
Prof. Giuseppe Lodato

Ph.D. Candidate:
Giovanni Dipierro
cycle XXIX

Director of the Ph.D. School:
Prof. Francesco Ragusa

Academic Year 2015 - 2016

Contents

Protoplanetary Discs in the ALMA era **v**

Part I Introduction **3**

1 Gas Dynamics in Protoplanetary Discs **3**
1.1 Star and disc formation 3
1.2 Gas dynamics 6
1.3 Angular momentum transport 14
1.4 Thermal structure of the disc 18

2 Dust Dynamics in Protoplanetary Discs **21**
2.1 Dust-gas aerodynamical coupling: Single dust grain is a gas fluid 21
2.2 Application to protoplanetary disc: Dynamics in a dust-gas mixture 24
2.3 Dust dynamics in laminar viscous flows 27
2.4 Dust dynamics in turbulent flows 37
2.5 The radial drift barrier 42

3 Observations of Protoplanetary Discs **47**
3.1 Spectral energy distribution 47
3.2 From near to mid-infrared observations 48
3.3 From far-infrared to radio observations 53

Part II Spirals in Gravitationally Unstable Protostellar Discs **65**

4 Dynamics of Self-Gravitating Accretion Discs **65**
4.1 Self-gravity influence on the disc structure 65
4.2 The onset of gravitational instability 68
4.3 Observational evidences of self-gravitating discs 76

5 How to Detect the Signatures of Self-Gravitating Discs with ALMA **85**
5.1 Disc model 86
5.2 Generation of intensity maps 88
5.3 The ALMA simulator 89
5.4 Results 90
5.5 Conclusion 97

6 Dust Trapping by Spiral Arms in Gravitationally Unstable Protostellar Discs **99**
6.1 Gas and dust disc models 100
6.2 Results 105
6.3 Summary and conclusion 113

Part III Gaps in Protoplanetary Discs **117**

7 Disc-Planet Interaction **117**
7.1 Planetary torque 117
7.2 Planet migration 128
7.3 Gap opening in gas discs 131

8 Two Mechanisms for Dust Gap Opening in Discs: an Opening Criterion **137**
8.1 Methodology 138
8.2 Results 139
8.3 Interpretation of Mechanism I: Dynamics in a viscous disc with a planet 142
8.4 Criterion for gap opening in dusty discs 145
8.5 Testing the criterion: Numerical simulations 151
8.6 Using the criterion 154
8.7 Conclusion 157

9 On Planet Formation in HL Tau **159**
9.1 Methods 160
9.2 Results 163
9.3 Summary 164

Part IV Horseshoes in Transitional Discs **169**

10 Transitional Discs **169**
10.1 Insights from observations 169
10.2 Theoretical models for the gas/dust cavities 172

11 On the Origin of Horseshoes in Transitional Discs **175**
11.1 Methods 175
11.2 Results 177
11.3 Discussion 179
11.4 Conclusions 182

12 Conclusions **185**
12.1 Limitations of the model 187
12.2 Future perspectives 188

Appendices **193**

A The viscous stress tensor **193**
A.1 Stress tensor for a circular shearing flow in cylindrical coordinates 193

B Interferometry and Aperture Synthesis **195**
B.1 Interferometry 196
B.2 Aperture synthesis 200

Bibliography **205**

List of Publications **219**

Protoplanetary Discs in the ALMA era

The question of how our world was created is almost as old as Mankind itself. The mistery about the origins of the Earth and the Solar System has fascinated men since prehistoric times, since they were able to lift their head and watch the skies. For centuries, astronomers were restricted to making visual observations from the Earth's surface, developing the first ideas about our planet and the Solar system architecture.

As long ago as 340 BC, in his book, *On the Heavens*, the Greek philosopher Aristotle presented the first arguments for believing that the Earth was a round sphere rather than a flat plate. Starting from observations of the lunar eclipses, he first realized that these phenomena were caused by the Earth coming between the Sun and the Moon. By studying the Earth's shadow on the Moon during the eclipses, he provided the first evidences of the spherical shape of our planet and the architecture of the Solar System. Aristotle thought the Earth was stationary at the center of the Universe, surrounded by eight spheres corresponding to the orbit of the Moon, the Sun, the stars, and the five planets known at the time, Mercury, Venus, Mars, Jupiter, and Saturn (see the left panel of Fig. 1). This idea was elaborated by Ptolemy in the second century AD into a complete cosmological model that was adopted by the Christian church as the picture of the universe.

After thirteen centuries, the polish astronomers Nicolaus Copernicus introduced a new cosmological model for the geometry of the Solar System. In his book, *De revolutionibus orbium coelestium*, he proposed that the Sun was stationary at the center and that the Earth and the planets moved in circular coplanar orbits around the Sun (see the right panel of Fig. 1). The predictions of this model were confirmed and refined by Tycho Brahe, who first analyzed the motion of Mars, showing that this planet was orbiting around the Sun following an ellipse with the Sun at one focus. After a few years, in 1610, the italian astronomer Galileo Galilei found that the planet Jupiter was surrounded by four small moons that orbited around it. They were the first group of objects found to orbit another planet, implying that everything did not have to orbit directly around the Earth, as Aristotle and Ptolemy had thought.

Once the structure of the Solar System was understood, philosophers and astronomers began wondering about the origin of the Solar System. The architecture and the morphological features of the planet's orbits in our Solar System provided the first constraints on the processes of planet and star formation, leading to the introduction of the first theory on the origin of the Solar System: the *nebula hypoThesis*. The theory was first described in the writings of Emanuel Swedenborg and then refined by Kant (1755) in his book, *Allgemeine Naturgeschichte und Theorie des Himmels*, and later by Laplace (1796). The theory states that the formation of stars are the result of the collapse of a primordial nebula. As a result of this collapse, a gaseous and dusty protoplanetary disc remains around the young star. They also hypothesized that these protoplanetary discs are the birthplaces of planets. The theory was then confirmed by the discovery of planets beyond the Saturn's orbits, Uranus and Neptune. By observing their motion in the sky, astronomers found that these planets have quasi-coplanar orbits, strengthening the belief that disc and planets formation are intimately linked.

The idea of a flattened, rotating structure as the origin of the observed co-planar planetary orbits in the Solar System was physically motivated by Hoyle (1960), who first described the

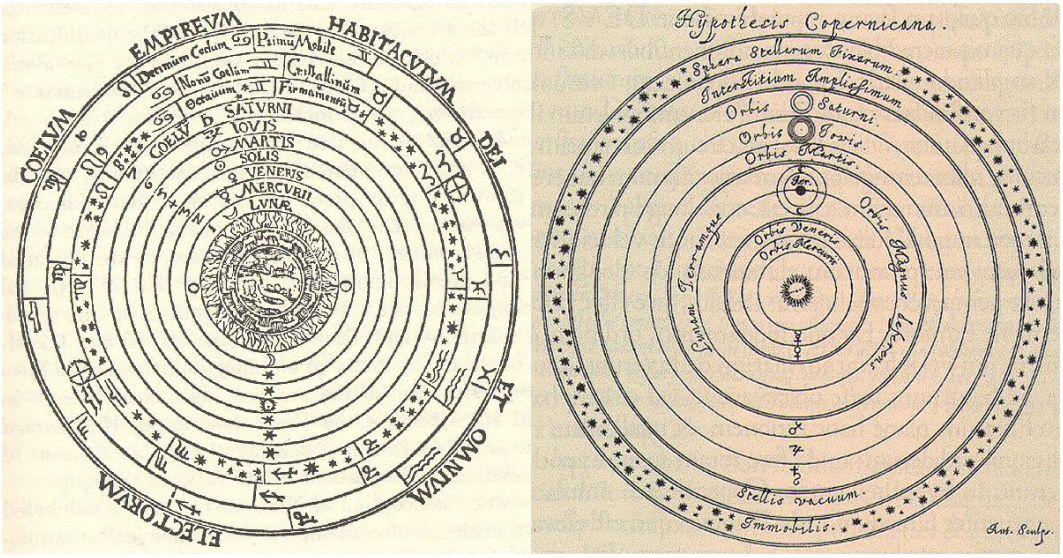

Figure 1: A simple schematic of the two cosmological models: geocentrism (left) and heliocentrism (right).

process of disc formation as the natural consequence of gravitational collapse of a rotating molecular cloud core. The keystone of this theory is that the dynamic evolution of the collapse is deeply affected by phenomena linked to the conservation of angular momentum. This produces, during this initial stage, the formation of an accretion disc whose properties are directly related to the angular momentum content of the molecular cloud core within which the star forms. The accretion disc formation is extremely important for the ongoing process of star formation, since it prevents the radial collapse of the molecular cloud core toward its centre.

Meanwhile, the first indirect observations of a flattened structure of gas and solids around young star were made. The first hints of the existence of circumstellar discs around young stars were obtained by the first surveys of star forming regions (Joy 1945; Herbig 1960, 1962). Importantly, the observation of collimated bipolar outflows were linked to an accretion disc origin (Lada 1985; Bastien & Menard 1988). Moreover, with the launch of the InfraRed Astronomical Satellite (IRAS) in 1983, it becomes evident that most pre-main sequence stars possess a strong far-infrared excess. The optically thin radio flux measurements and the infrared excess in the spectral energy distribution were interpreted as a flattened dusty structure irradiated by the central protostar (Strom et al. 1989; Beckwith et al. 1990; Cabrit et al. 1990). Importantly, the first models on protoplanetary formation were based on the idea that the star forms from the direct collapse of a uniform spherical cloud. The discrepancy between the observed protoplanetary luminosities and the one predicted by model the spherical infall of a typical molecular cloud were explained by the fact that the infalling material mostly lands on a disc at significantly larger radii than the star radius (Hartmann 2009).

The first direct evidences of a circumstellar discs around young protostars were obtained with the launch of the NASA's Hubble Space Telescope (HST) in 1990. The high spatial resolution capabilities of HST allowed to spatially resolve discs around young star. These discs were imaged in "silhouette" in absorption against the bright background of the Orion Nebula (O'dell et al. 1993; McCaughrean & O'dell 1996). Fig. 2 shows a sample of HST observations in the Orion Nebula where can be easily seen a flattened structure around the forming protostar. Importantly, these images revealed that protoplanetary disc are ubiquitous around newly formed star (50% of star were detected with a circumstellar disc), strenghtening the idea that such discs are a common by-product of the star formation process. With the first generation of radio interferometers and the improvements of observational techniques, the following years saw a remarkable increase of disc detection, leading to a substantial maturation of this research area. In particular, the detection of circumstellar discs has been the subject of many (sub-)millimeter wavelengths survey

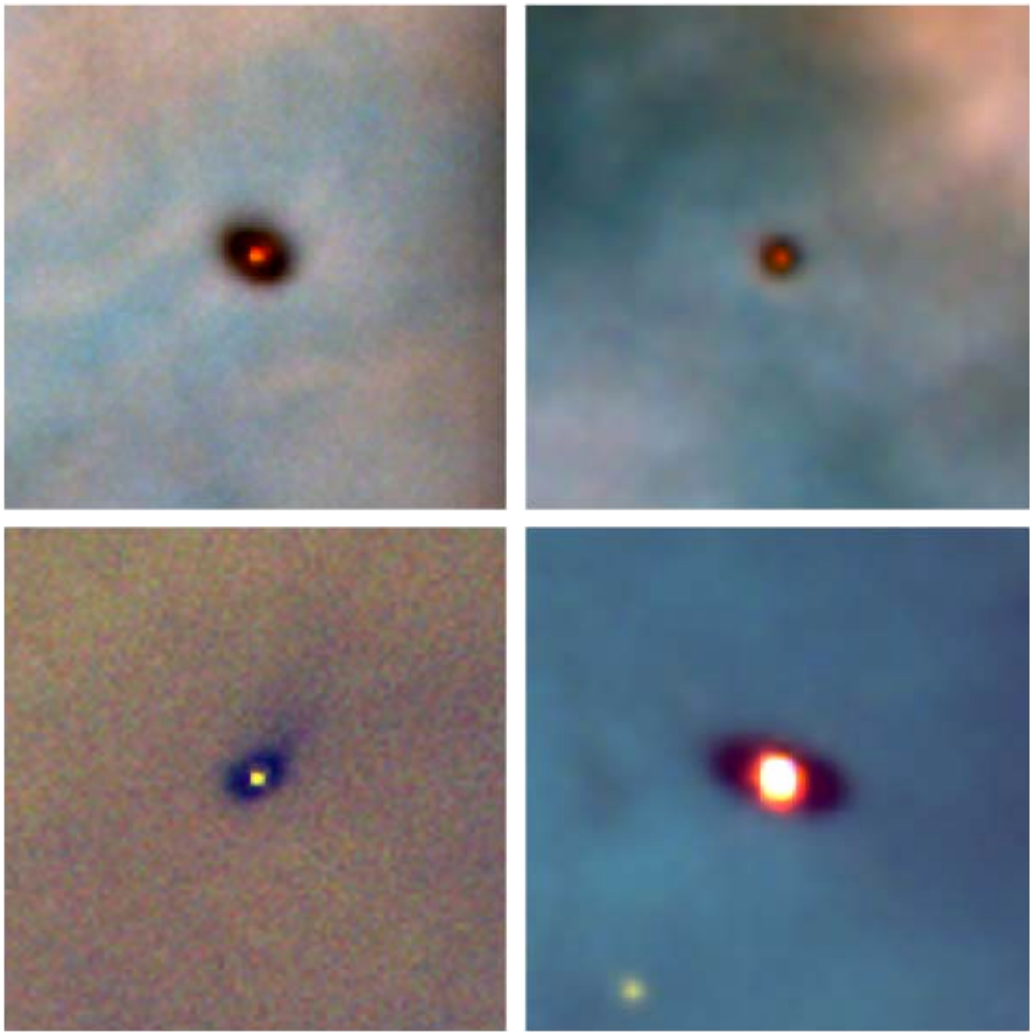

Figure 2: A sample of protoplanetary discs in the Orion nebula, imaged by the Hubble Space Telescope. Image adapted from McCaughrean & O'dell (1996).

(e.g. Andrews & Williams 2007b,a) especially focused on the inference of disc properties in a number of star forming regions.

At the same time, it was soon realized that the dynamics and the chemical structure of protoplanetary discs are very complicated. The physics behind protoplanetary discs involves a large number of research areas: gas and dust dynamics, magnetic fields, radiative transfer and chemistry. They are all closely interlinked, affecting each other in a large number of ways. From an observational point of view, it is not straightforward to interpret observations of circumstellar discs, since the observational measurements is often degenerate, making direct comparison between observations and theory difficult. Most of the lingering uncertainties in protoplanetary discs are linked to our ignorance on the local disc properties and to the limitations in observational techniques.

In the last decades, the investigation on the star formation process has been the subject of several theoretical analysis aimed, on the one hand, at understanding the accretion and angular momentum transport (Lynden-Bell & Pringle 1974; Pringle 1981) and, on the other hand, at exploring the planet formation process (e.g. Lissauer 1993). The process of transport of angular momentum in protoplanetary discs represents one of the unsolved problem in astrophysics. Understanding how the angular momentum is transported (or removed) during the star formation process is the key to figure out the disc evolution and its implications on the planet formation process in the early phases. However, although the last three decades have seen a remarkable progress in the theory of disc evolution, a comprehensive model for protoplanetary

discs formation and evolution is far to be developed. Simplified models of disc evolution have been developed in the past, mostly based on a pseudo-viscous model for the angular momentum transport driven by disc instabilities.

On the other hand, interest in the process of planet formation has remarkably grown in the last 20 years due principally to the discovery of several planetary systems outside the Solar System. In this context, it is worth remarking the significant discoveries carried out by the Kepler spacecraft launched in 2009 (see `http://kepler.nasa.gov/Mission/discoveries/`). The space-based Kepler transit survey (Borucki et al. 2010) and long-term ground-based radial velocity surveys (Cumming et al. 2008; Howard et al. 2010) has been allowing to make statistical studies on the range of outcomes of the planet formation process. These survey are revealing an unexpected diversity in exoplanet systems (Winn & Fabrycky 2015; Hatzes 2016), demonstrating that most of the gross features of the Solar system are only one of the outcomes in a continuum of possibilities. This study is a powerful complement to the work on protoplanetary discs, since it provides insights on the birth environments and to the complex range of outcomes of the planet formation process. Nowadays, excluding the candidate planets, the number of confirmed exoplanet is about 3500 (see `http://exoplanetarchive.ipac.caltech.edu/`), with an intriguing variety of orbital parameters and masses and, remarkably, an expected plethora of circumbinary planets (Doyle et al. 2011; Orosz et al. 2012a,b; Armstrong et al. 2014). The link between the detected exoplanetary system and protoplanetary discs can be the key to figure out the physical mechanisms responsible for the star and disc evolution and the formation of planets in protoplanetary discs.

Fortunately, recent technical advances are offering us an unprecedented opportunity to gain insights on protoplanetary discs and planet formation process. With the completion of ALMA, the Atacama Large Millimeter/Submillimeter Array, and new access to high quality scattered light data, the research in protoplanetary discs will look significantly different in the next future. Current and forthcoming observing facilities are expected to produce a significant step forward in our understanding of the processes regulating the evolution of planetary systems and give the unique possibility to put constraints on the several theoretical models developed in the past. In particular, the capabilities of the ALMA interferometer provide access to the small scales necessary for examining disc substructures with an unprecedent detail.

In this context, a recent active topic of research is the characterization of substructures within protoplanetary discs. Recent high-resolution observations have revealed peculiar disc substructures such as spirals (Muto et al. 2012; Grady et al. 2013; Garufi et al. 2013; Christiaens et al. 2014; Benisty et al. 2015; Wagner et al. 2015; Garufi et al. 2016; Stolker et al. 2016), gaps (ALMA Partnership et al. 2015; Nomura et al. 2016; Andrews et al. 2016; Canovas et al. 2016; van der Plas et al. 2016; de Boer et al. 2016; Ginski et al. 2016; van Boekel et al. 2016) and horseshoes (Brown et al. 2009; Casassus et al. 2012, 2013; Isella et al. 2013; van der Marel et al. 2013; Pérez et al. 2014; Casassus et al. 2015b; van der Marel et al. 2015a) in protoplanetary discs (see the recent review of Casassus 2016). Fig. 3 shows prominent examples of recent high-resolution observations. The link between the morphological features of this substructures and the mechanism behind their formation and evolution may be useful to infer the disc conditions and develop a comprehensive theory of disc and planet formation. These recent breakthroughs in astronomy, made possible by new technologies, suggest answers to some of these longstanding questions. As I will describe in the Thesis, the origin of this structures still remains the subject of intense studies. They could be linked to disc instabilities, or to the presence of a planetary or (sub-)stellar companion orbiting around the central star, or might be induced from variable illumination. Whereas the nature of these structures is still debated, there is hope that they can provide long-awaited questions on the processes responsible for planet formation and disc evolution.

The key to figure out the nature of these substructures is based on the investigation on the spatial distribution of the solid and gas component in the protoplanetary disc. As already mentioned, protoplanetary discs are believed to be formed from the interstellar medium and, as such,

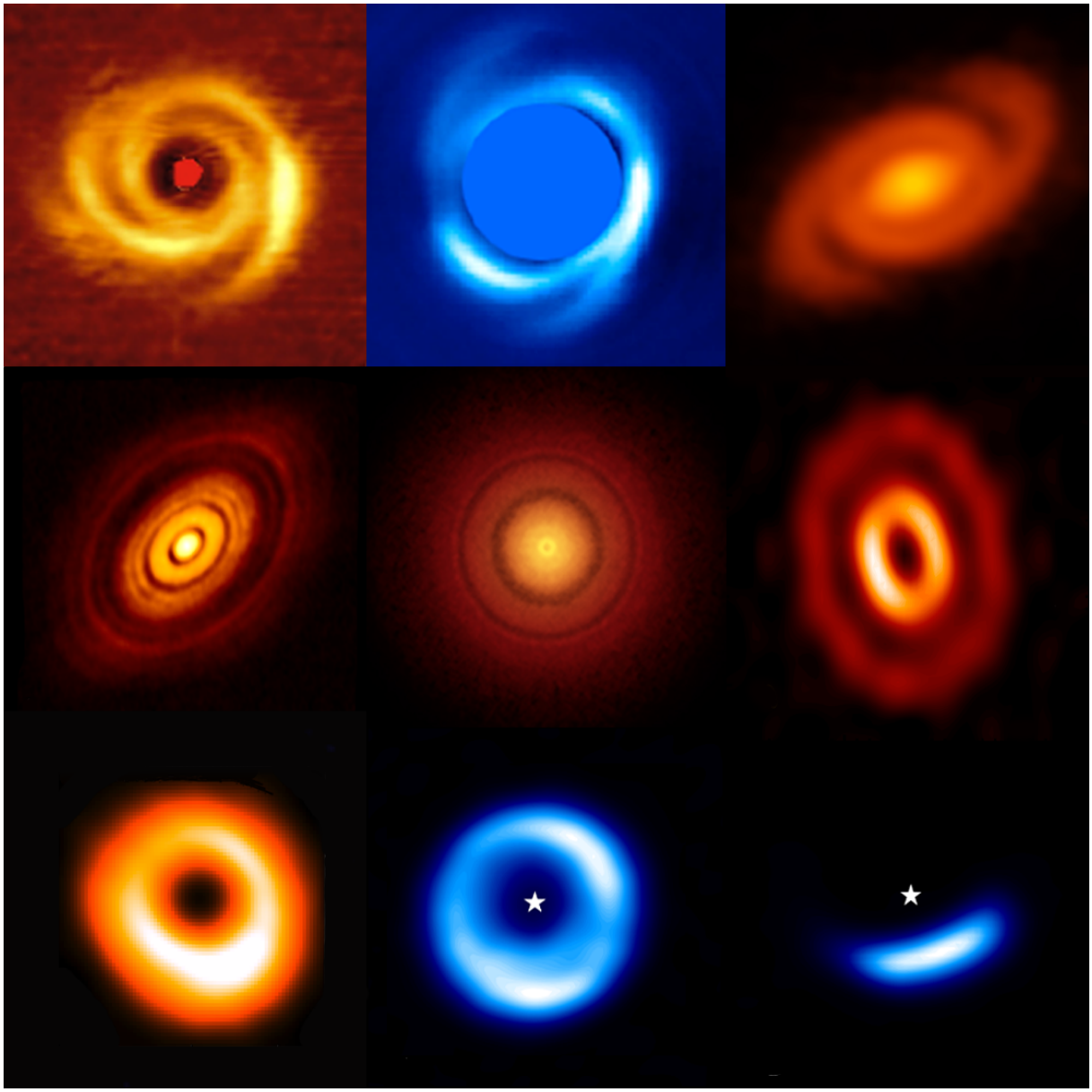

Figure 3: Gallery of disc substructures. Spiral structures are observed in polarized scattered light of (top left) HD 135344B (Garufi et al. 2013), (top center) HD 100453 (Wagner et al. 2015) and (top right) in ALMA band 6 of Elias 2-27 (Pérez et al. 2016). Gap-like features are detected in thermal dust emission at (sub-)millimetre wavelengths in (middle left) HL Tau (ALMA Partnership et al. 2015), (middle center) TW Hydrae (Andrews et al. 2016) and (middle right) HD 97048 (van der Plas et al. 2016). Horseshoe are observed in thermal dust emission at (sub-)millimetre wavelengths in (bottom left) HD 135344B (Pérez et al. 2014), (bottom center) SR 21 (Pinilla et al. 2015) and (bottom right) Oph IRS 48 (van der Marel et al. 2013). All the images are on different scale.

are assumed to be composed of 1% of condensible material (commonly called solids or dust) and 99% of gas *at birth* (Bohlin et al. 1978). Although the mass of protoplanetary discs is dominated by gas, the solid component plays a crucial role for a large range of aspects about their physical and chemical evolution. The most important aspect of the dust in protoplanetary discs is that it dominates the opacity budget. As a result, dust grains play an active role in determining the thermal structure of discs. This aspect is extremely important in interpreting observations, since observations of light scattered on dust grains and their direct thermal emission remain the key diagnostic tools to detect discs and to characterize their structure. Since the emission observed at a given wavelength most efficiently probes particles of a similar size, the key to figure out the nature of these structures is to investigate the spatial distribution of dust grains with different size. It is widely believed that, while the spatial distribution of micron-sized dust grains is similar to the gas distribution, millimeter-sized grains tend to migrate toward the local maximum in the gas pressure distribution and settle toward the disc midplane. The resulting vertical stratification and the size-sorted spatial density distribution produce macroscopic features that can be identified through high angular resolution observations over a wide range of wavelengths. Moreover, an additional complication arises from the fact that dust grains in protoplanetary discs are very different from grains in the diffuse interstellar medium. Observations of dust in discs from (sub-)millimetre to centimetre wavelengths have provided strong evidence for grain growth processes in discs (Testi et al. 2014). Furthermore, it is widely believed that their chemical properties change from object to object and are mostly influenced by the interaction with the gas component. An additional difficulty is that substructures in discs can be caused by a large number of physical mechanism and care should be taken to connect these detections with intrinsic asymmetries in the disc. Any kind of structure might be related to a large range of effects such as optical depth, anisotropic scattering or polarization effects. In the simplest cases, if we assume that the detected substructures are intrinsic asymmetries in the disc, these can be density or temperature fluctuations, but also fluctuations in grain opacity due to variation in grain size or composition, or fluctuations of the local turbulence, which can cause the vertical extent of the disc to vary.

Therefore, the modelling of the dust and gas spatial distribution has to be correctly computed in order to figure out which mechanisms is responsible for such substructures. By identifying these structures and comparing them with theoretical models, it is possible to provide predictive links to the local gas and dust conditions, opening up a novel path for constraining theoretical models developed in the past.

Outline of the Thesis

In light of recent breakthrough observational results, this Thesis work is aimed at bridging the gap between the theoretical and observational approach in investigating some of the complex phenomena occurring in protoplanetary discs by numerically and analytically calculating the dynamical evolution of the dust and gas components and inferring the resulting observational signatures.

This work is focused at reproducing the observational prediction of three classes of discs: gravitationally unstable discs, discs hosting planets and discs around binary system. For each of these models, I investigate the spatial distribution of gas and dust and analyze the observational predictions over a large range of wavelengths in order to verify if recent high resolution observations can be interpreted according to these models. In detail, I simulate the dynamics of gas and dust by performing both grid (Brauer et al. 2008; Birnstiel et al. 2010) and SPH-based (Laibe & Price 2012a,b, 2014; Price & Laibe 2015) simulations adopting the previously mentioned models. Moreover, given the crucial role of the observational signatures of protoplanetary discs, Monte Carlo radiative transfer simulations have been performed to compute the expected emission of disc models over a wide range of wavelengths . Finally, in order to compare the numerical results with observational results, I simulate realistic observations of the disc models

as they may appear with current observing facilities.

This Thesis is organized in 4 parts as follows:

Part I provides an introduction to the research theme of this Thesis, focussing on the theoretical and observational essentials about protoplanetary discs. The aim of this part is to give a background of the theoretical and observational investigations developed in this Thesis. **Chapter 1** and **2** are focused on gas and dust dynamics in protoplanetary discs, while **Chapter 3** offers a introduction about the observational aspects of protoplanetary discs.

Part II explores the scenario of gravitationally unstable discs. **Chapter 4** provides an overview of the theoretical aspects of the development of gravitational instability and offers predictive links to observations, presenting a brief overview of recent observations of spirals in protoplanetary discs. **Chapters 5** and **6** explore the observational predictions of gravitationally unstable discs with different numerical and observative approaches. This work has been published as two articles in *Monthly Notices of the Royal Astronomical Society* (Dipierro et al. 2014, 2015a), on which the **Chapters 5** and **6** are based.

Part III is focused on protoplanetary discs hosting planets. This part explores the scenario in which gaps and ring-like features recently detected in protoplanetary discs are linked to the tidal interaction between the disc and embedded protoplanets. After an introductory part on the topic in **Chapter 7**, **Chapter 8** and **9** focuses on dust dynamics in discs with planets and the observational predictions of this class of discs, giving predictive links to the planet properties responsible for the gaps carved in the HL Tauri protoplanetary disc (ALMA Partnership et al. 2015). This work has been published as two letters in *Monthly Notices of the Royal Astronomical Society Letters* (Dipierro et al. 2015b, 2016), on which the **Chapters 8** and **9** are based. Moreover, the introductory part described in **Chapter 7** is partially based on a third article currently submitted.

Part IV focuses on explaining recent observations of transitional discs, which reveal non-axisymmetric horseshoes in the dust continuum emission. After a brief overview of transitional discs in **Chapter 10**, I will explore in **Chapter 11** the hypoThesis that these asymmetric features are related to the dynamics of gas and dust at the edge of the cavity in circumbinary discs. This works has been published as an article in *Monthly Notices of the Royal Astronomical Society* (Ragusa et al. 2017).

Part I
Introduction

Chapter 1

Gas Dynamics in Protoplanetary Discs

Protostellar discs are the natural outcome of the star formation process. Their formation and evolution are strongly related to the content of angular momentum of their nascent rotating molecular cloud. Importantly, they represent the environment in which planets form and grow and any understanding of the planet formation process is intimately coupled to understanding of the structure and evolution protostellar disc. Since gas and dust coexist in protostellar discs, the evolution of the dust grains strongly depends on the evolving gas and *vice-versa* (see Chapter 2). The study of the dynamics of gas is therefore essential to predict the spatial distribution of dust grains. At the same time, since observations are mostly sensitive to the spatial distribution of dust, understanding the gas dynamics and its interaction with the dust is crucial to interpret observational results in order to infer the background gas structure.

In this Chapter we will review the basic physical processes that determines the formation and evolution of a thin non-magnetic, non-self-gravitating gaseous accretion disc rotating around a single central object. For the time being, we will focus on the gas dynamics, assuming that the dust has a negligible dynamical effect on the gas. In other words, we assume that the mass content of the disc is dominated by the gas phase. This approximation is strictly valid only in the initial part of the star formation process due to the fact that initially the dust coexists with the gas over all the molecular cloud extent and is not segregated in some region (the dust-to-gas ratio in young protostellar discs is typically of the order of 10^{-2}, as found in the interstellar medium, Bohlin et al. 1978; Savage & Mathis 1979). This approximation will be relaxed in the next Chapter where we will describe the dust dynamics and its influence on the gas. In this Chapter we will also summarize the disc structure resulting from hydrostatic equilibrium and viscous accretion with a particular focus on the mechanism of angular momentum transport. The modern theory of thin disc evolution presented in this Chapter was first described in classic papers by Lynden-Bell & Pringle (1974) and Pringle (1981) and recently reviewed by Lodato (2008), to which we refer for more details.

1.1 Star and disc formation

The star formation process has its beginning in the gravitational collapse of molecular clouds cores (Hoyle 1960). In the absence of any rotation of the molecular cloud, the gas in the core could freely collapse towards the centre directly accreting onto the forming protostar (Hartmann 2009). However, several observations of star forming regions have demonstrated that the molecular cloud cores are characterized by rotation (e.g. Goodman et al. 1993). Therefore, angular momentum plays a crucial role in determining the evolution of the system.

A generic cartoon of of the star formation process is shown in the left panels of Fig. 1.1 where it is possible to distinguish the different stages of the process. The collapse of a rotating gaseous cloud leads to the formation of a circumstellar disc with size of the order of hundreds of au in a typical time of $\sim 10^4$ yrs. Its characteristics are essentially dependent on the angular momentum content of the molecular cloud core within which the star forms (Terebey et al. 1984; Shu et al.

1987). Given a generic fluid element of the collapsing cloud, it will accrete onto the protostellar core or settle to the equatorial disc depending on its angular momentum. During this stage the envelope accretes onto the forming disc for approximately $\sim 10^5$ yrs. In addition, due to the angular momentum transport inside the disc, the disc accretes onto a central protostar. This stage is also called *main infall phase* and, using the standard classification of evolutionary stages, this forming protostar/disc system is also referred to as a Class 0/Class I object.

The typical size of molecular cloud cores can be estimated by carrying out a first-order perturbation analysis of the response of the medium to small perturbations (Jeans 1902; Hartmann 2009). In this way, one can estimate the maximum wavelength of small perturbations below which the medium is stable. Assuming that the cloud in which the star forms is characterized by an infinite, static and homogeneous medium with density $\rho_{\rm g}$, the typical length scale $\lambda_{\rm J}$ (typically referred to as *Jeans length*) is given by

$$\lambda_{\rm J} = \frac{c_{\rm s}\sqrt{\pi}}{\sqrt{\mathcal{G}\rho_{\rm g}}}, \tag{1.1}$$

where $c_{\rm s}$ is the sound speed and $\mathcal{G}$ is the gravitational constant. For a typical molecular cloud with numerical density $\sim 10^5$ cm^{-3} and temperature ~ 50 K, the value of $\lambda_{\rm J}$ is of the order of ~ 0.1 pc. As can be seen in the equation above, the density and the temperature play a crucial role in determining the typical mass below which the collapse of the molecular cloud starts. The Jeans mass $\propto \lambda_{\rm J}^3$ decreases with decreasing sound speed (and hence temperature, $T \propto c_{\rm s}^2$) and increasing density.

Before moving to the later evolutionary stages, it is worth remarking that, due to the mass loading of the infalling material from the surrounding envelope, the circumstellar discs (which hereafter can be refereed to as protostellar discs) could be massive enough to be self-gravitating. The development of gravitational instability may provide a source of angular momentum and energy transport which enhances the accretion of the infalling material onto the central protostar (see Chapter 4).

When the infall of material from the molecular cloud has ceased (after approximately 10^5 yrs from the beginning of the star formation), the system is characterized by a protostellar disc (also referred to, in this stage, as protoplanetary disc) that slowly accretes onto the central pre-main-sequece (PMS) star. Systems at this stage of their evolution are also referred to as Classical T-Tauri (CTT) or Class II objects. The final stage of star formation is dominated by the central protostar with a remnant disc consisting almost entirely of solid materials. Systems at this final stage are called Class III object or Weak T-Tauri (WTT) star and the disc component in the system is also referred to as planetary "debris" disc.

The different stages of the star formation process has an observational counterpart that can be obtained through the analysis of Spectral Energy Distribution (SED) of the evolving system (see Sect. 3.1). The SED indicates how the energy (intercepted by stellar radiation and produced by internal stresses) are distributed over frequencies after being processed through the disc material. This SED-based classification is based on the evaluation of the spectral index s of the emitted flux F_λ of the object at wavelength λ, defined as

$$s = \frac{\partial \log(\lambda F_\lambda)}{\partial \log \lambda}, \tag{1.2}$$

and its value is typically measured in the range of wavelengths from $\lambda \sim 2\,\mu$m to $\sim 50 - 100\,\mu$m (André 2002; Armitage 2010). The right panels of Fig. 1.1 show the typical SEDs of the disc+protostar system in the different stages of star formation. Class 0 and younger sources presents a strong emission in the sub-millimetre range of wavelengths which indicates that it is dominated by cool dust in the envelope ($T \lesssim 30 - 50K$). Class I sources are typically visible in the near-infrared with a rising spectrum towards long wavelengths. In addition, they are characterized by a typical "infrared-excess" emission that is due to the reprocessing of star radiation by the circumstellar dust (see Chapter 3, Hartmann 2009). The dusty disc absorbs

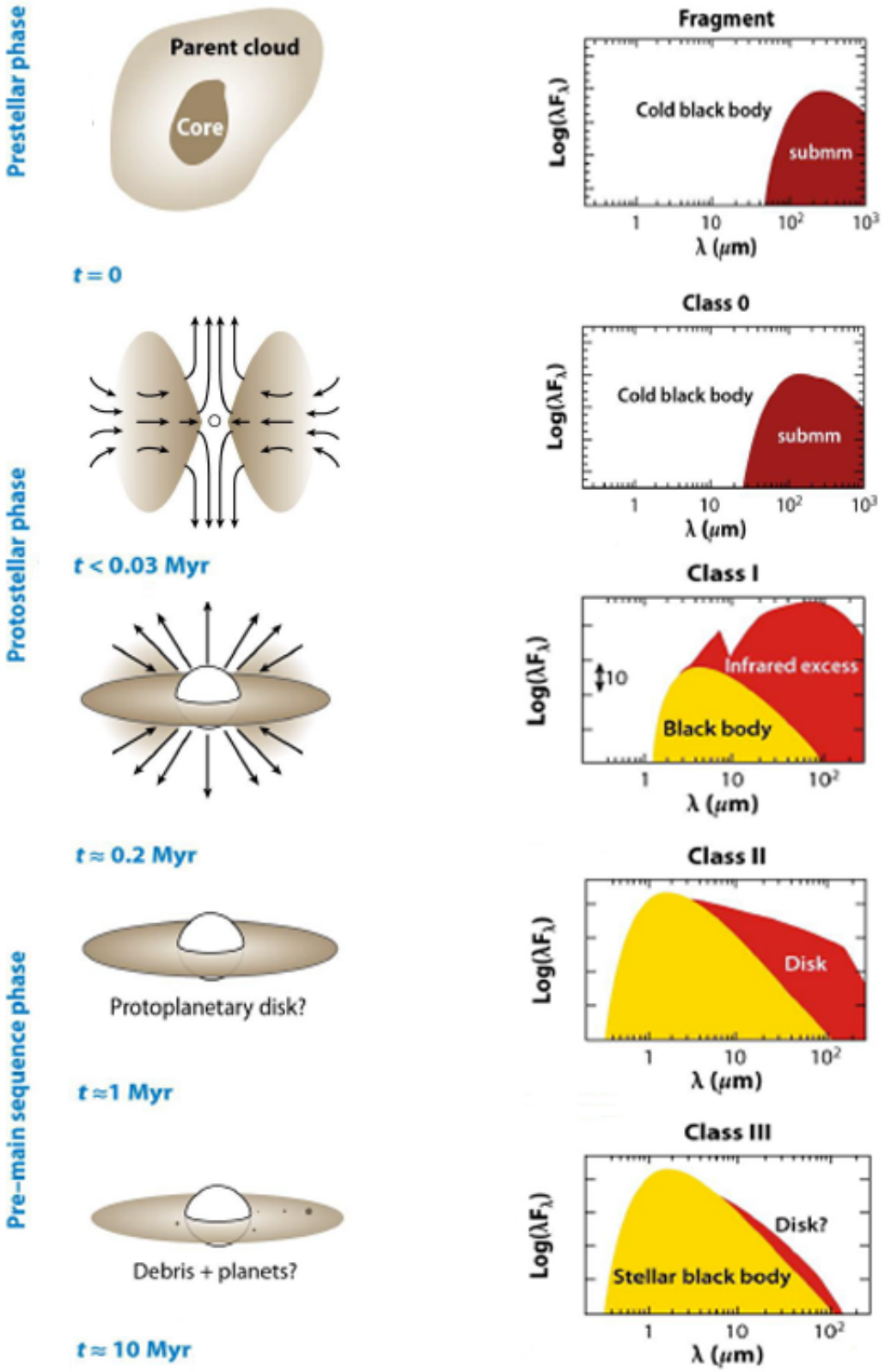

Figure 1.1: Cartoon image illustrating the star and formation process (left) with the expected SEDs (right). A dense molecular cloud collapses under its own gravity and, due to its initial rotation, forms a disc which carries most of the angular momentum of the primordial nebula and slowly accretes onto the central star. Image adapted from André (2002).

energy from the central star at short wavelengths and re-emits it at far-infrared wavelengths where the envelope is sufficiently optically thin for the radiation to escape (see further details in Sect. 3.3). For both the last mentioned objects, the spectral index, since the emission increases with wavelength in the considered waveband, has a positive value, i.e. $s > 0$. For the Class II objects, the SED has a spectral index in the range $-4/3 \lesssim s \lesssim 0$, which is frequently identified as the emission produced from a dusty disc optically thick in the infrared and optically thin in the millimetre range of wavelengths. Finally, the SEDs of Class III sources are dominated by the pre-main sequence star emission without infrared excess. The spectral index is approximately $s \sim -3$.

Coming back to the initial stages of the star formation process, once the disc is formed, the gas needs to redistribute its angular momentum in order to guarantee the further accretion of the central star. The way to accomplish that is provided by the accretion disc. The accretion process has been confirmed by a variety of observations which have shown that the properties of molecular cloud cores are consistent with the formation of accretion discs around young stars. The typical evolution timescale of a protostellar discs can be inferred by analyzing the near-infrared excess of the spectral energy distribution. Haisch et al. (2001) found that the fraction of young stars that show a near-infrared excess as a function of time indicates that at an age of $\lesssim 1$ Myr most young stars harbour a disc, half of the stars within the clusters lose their discs in $\lesssim 3$ Myr, while all the stars older than 10 Myr have completely lost its disc (Strom et al. 1989; Zuckerman et al. 1995; Sicilia-Aguilar et al. 2006; Hernández et al. 2007; Mamajek 2009; Ribas et al. 2014). Moreover, Fedele et al. (2010) investigated the fraction of accreting objects in stellar clusters, showing that no accreting object has been found with age higher than 10 Myr (see also Mendigutía et al. 2012). It is therefore required some mechanism to transport angular momentum within the disc itself.

1.2 Gas dynamics

The basic framework for the construction of accretion disc theory is traditionally based on a few crucial assumptions, which together simplify the theoretical analysis of disc without any loss of generality.

First of all, it is possible to reasonably assume that discs are thin, meaning that the disc radial extent is much greater than the disc thickness $H_{\rm g}$. This assumption is empirically confirmed by observations of discs around active galactic nuclei ($H_{\rm g}/R \approx 0.001 - 0.01$) and protostellar discs ($H_{\rm g}/R \approx 0.1$). The thin-disc approximation allows, therefore, to simplify the calculations using equations integrated in the vertical direction with quantities evaluated per unit surface (such as the surface density) instead of quantities per unit volume (such as the volume density). In addition, it is convenient to adopt cylindrical polar coordinates (R, ϕ, z) with the central object at the origin and with the disc mid-plane lying in the plane $z = 0$. Furthermore, it is useful to assume that the disc is axisymmetric to zeroth order. This assumption allows us to treat all the quantities as independent on the azimuthal coordinate ϕ. Importantly, according to these assumptions, we can safely assume that the azimuthal velocity depends only on the radial coordinate, i.e. $v_{{\rm g},\phi} = v_{{\rm g},\phi}(R)$. It is worth pointing out that the latter assumption will be not valid when we will describe the presence of spiral structures or horseshoes within the disc (see Parts II and IV), which are clearly not axisymmetric. Moreover, since the typical timescale over which the disc evolves is much longer than the dynamical one, it is reasonable to assume that the radial component of velocity is much smaller than the azimuthal component (i.e. $v_{{\rm g},R} \ll v_{{\rm g},\phi}$) as, otherwise, the disc would accrete onto the central object in less than the time required for an orbit. All these approximations allows to simplify the fluid equations.

Additionally, in this Chapter we limit our analysis to non self-gravitating discs. Therefore, we assume that the gravity of the disc does not influence the disc dynamics and structure. In Chapter 4 we will relax this assumption by analyzing the effect of the self-gravity on the disc density and thermal structure with a particular focus on its importance in transporting energy and angular momentum.

1.2.1 Equation of motion for viscous fluids

The disc dynamics can be described through the basic equation of viscous fluid dynamics given by

$$\frac{\partial \rho_{\rm g}}{\partial t} + \nabla \cdot (\rho_{\rm g} \mathbf{v}_{\rm g}) = 0, \tag{1.3}$$

$$\frac{\partial \mathbf{v}_{\rm g}}{\partial t} + (\mathbf{v}_{\rm g} \cdot \nabla)\mathbf{v}_{\rm g} = -\frac{1}{\rho_{\rm g}}(\nabla P - \nabla \cdot \sigma) - \nabla\Phi, \tag{1.4}$$

where the subscript g refer to the gas (to be consistent with the notations used in Chapter 2), P is the pressure, $\rho_{\rm g}$ is the mass density of the gas and σ is the stress tensor. Φ is the gravitational potential that can be calculated by solving the Poisson equation

$$\nabla^2\Phi = 4\pi\mathcal{G}\rho_{\rm g}. \tag{1.5}$$

The gravitational potential may include only the central body alone or, in the case of self-gravitating discs, it is necessary to take into account also the contribution of the disc. Eq. 1.3 is the continuity equation, which express the mass conservation, while Eq. 1.4 is the momentum equation, commonly referred to as Navier-Stokes equation.

As can be noticed in Eq. 1.4, the key role in accretion disc dynamics is covered by the stress tensor. This term describes the effect of viscous forces in transporting angular momentum and dissipating energy across the disc, providing a net conversion of mechanical energy into heat which allows the driving of the accretion flow. The nature of viscosity has historically been a key problem in the accretion disc dynamics. According to the classical view, it can be assumed to be described by the classical shear viscosity. Protostellar discs are characterized by differential rotation, i.e. material at different radii moves with different angular velocity. We can thus consider a flow with parallel azimuthal streamlines in which the velocity gradient is perpendicular to the direction of the streamlines. If there is communication between neighbouring annuli, such as chaotic thermal motion of fluid molecules or turbulent motion of fluid elements, viscous stresses would produce a transfer of angular momentum from fast to slow moving material. In other words, the viscosity is assumed to be ascribed to the shearing flow in the azimuthal direction, thus proportional to the rate of local strain. It is worth remarking that these statements about the transfer of angular momentum are quite general: they are independent on the nature of any particular microscopic model of the angular momentum transport and they can be made with the unique assumption that the stress is related to the local rate of strain between fluid elements. As we will describe below, non-local mechanisms of angular momentum transport might exist in protostellar discs, such as those driven by long range forces. In these cases, the simple linear proportionality between stress and the local gas properties strain breaks down. Importantly, according to these assumptions, we can safely assume that the azimuthal velocity depends only on the radial coordinate, i.e. $v_{{\rm g},\phi} = v_{{\rm g},\phi}(R)$. In the following, we will restrict our analysis to small scale transport processes where the local approximation is valid.

Due to the axisymmetric nature of the system, it is convenient to derive the fluid equations in cylindrical coordinates. In these coordinates, the only non-zero component of the stress tensor in a circular axisymmetric shear flow is $\sigma_{\phi R} = \sigma_{R\phi}$ which indicates the force in the direction ϕ acting on a surface whose normal is in direction R. The non-null viscous stress tensor and its divergence component are given by (see Appendix A)

$$\sigma_{R\phi} = \eta R \frac{\partial \Omega}{dR}, \qquad \nabla \cdot \sigma|_\phi = \left(\frac{\partial}{\partial R} + \frac{2}{R}\right)\sigma_{R\phi}, \tag{1.6}$$

where η is the shear viscosity coefficient. Analyzing separately the vertical, radial and azimuthal components of Eq. 1.4, one can obtain important properties about the disc dynamics and structure.

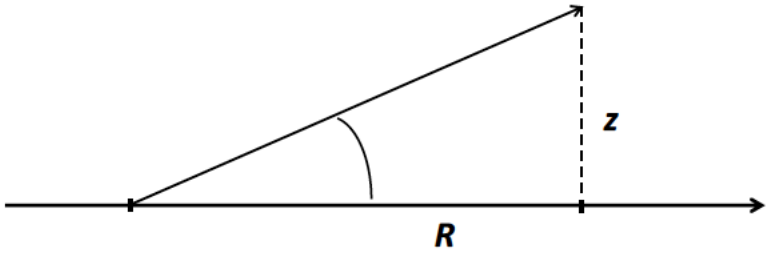

Figure 1.2: Geometrical scheme of the disc section involved in the calculation of the disc vertical structure.

1.2.1.1 Vertical hydrostatic equilibrium

According to the thin-disc approximation, since the disc is assumed to be confined on the midplane, it is expected that the vertical component of the velocity can be neglected. Thus, the left hand side of the vertical component of Eq. 1.4 is null. Since the only non-vanishing component of the stress tensor is in the $R\phi$ direction, the resulting equation reduced to the hydrostatic balance between gravitational and pressure force in the vertical direction:

$$\frac{1}{\rho_{\rm g}}\frac{\partial P}{\partial z} = -\frac{\partial \Phi}{\partial z}. \tag{1.7}$$

Starting from this balance between gravity and pressure, it is possible to infer the vertical structure of the disc. The vertical component of gravity is given by

$$\frac{\partial \Phi}{\partial z} = \frac{\mathcal{G}Mz}{\left(R^2+z^2\right)^{3/2}}, \tag{1.8}$$

where $\mathcal{G}$ is the gravitational constant, M is the mass of the central object and R is the cylindrical radius (see Fig. 1.2). Assuming that the gas is barotropic, i.e. $P = P\left(\rho_{\rm g}\right)$, the first term of Eq. 1.8 can be rewritten, using the definition of sound speed: $c_{\rm s}^2 = dP/d\rho$, as

$$\frac{1}{\rho_{\rm g}}\frac{\partial P}{\partial z} = \frac{c_{\rm s}^2}{\rho_{\rm g}}\frac{\partial \rho_{\rm g}}{\partial z} = -\frac{\mathcal{G}Mz}{\left(R^2+z^2\right)^{3/2}}. \tag{1.9}$$

Assuming that the sound speed does not depend on z, the disc vertical structure, in the limit of thin discs, i.e. $z \ll R$, is thus given by

$$\rho_{\rm g}(R,z) = \rho_{\rm g,0}\exp\left[-\frac{\mathcal{G}M}{c_{\rm s}^2}\left(\frac{1}{R}-\frac{1}{\sqrt{R^2+z^2}}\right)\right] \underset{z\ll R}{\approx} \rho_{\rm g,0}\exp\left[-\frac{z^2}{2H_{\rm g}^2}\right], \tag{1.10}$$

where $\rho_{\rm g,0}$ is the volume density at the mid-plane, i.e. $z = 0$, and where $H_{\rm g}$ is the typical vertical length scale expressed by

$$H_{\rm g} = \frac{c_{\rm s}}{\Omega_{\rm k}}, \tag{1.11}$$

where $\Omega_{\rm k}$ is the Keplerian angular velocity, i.e. $\Omega_{\rm k} \approx \sqrt{\mathcal{G}M/R^3}$. The mid-plane density can be written in terms of the surface density $\Sigma_{\rm g}(R) = \int_{-\infty}^{\infty}\rho_{\rm g}(R,z)\,dz$ as

$$\rho_{\rm g,0} = \frac{1}{\sqrt{2\pi}}\frac{\Sigma_{\rm g}}{H_{\rm g}}. \tag{1.12}$$

Since discs are geometrically thin, deviations from the gaussian profile are negligible. This result shows that the temperature is closely linked to the disc thickness, since the thermal pressure tends to expand the disc in the vertical direction against gravity. In addition, it can be noted that the thin-disc approximation is equivalent to requiring that the flow rotates[1] with a velocity $v_{\rm k} = c_{\rm s}\,R/H_{\rm g} \gg c_{\rm s}$ which indicates that the flow is highly supersonic.

[1] In Sect. 1.2.1.2, we will show that the disc rotation, at a first approximation, is Keplerian, i.e. $v_{{\rm g},\phi} \simeq v_{\rm k}$.

1.2.1.2 Centrifugal Balance

Let us turn to analyze the radial component of the Navier-Stokes equation (Eq. 1.4). It is reasonable to assume that the radial component of velocity is much smaller than the azimuthal component (i.e. $v_{\mathrm{g},R} \ll v_{\mathrm{g},\phi}$) as, otherwise, the disc would accrete onto the central object in less than the time required for an orbit. Therefore, we can neglect all the term in the left-hand side of Eq. 1.4 except for the centrifugal term $v_{\mathrm{g},\phi}^2/R$. On the right-hand side, it can be shown that the viscous term vanishes. Therefore, the centrifugal balance is given by

$$\frac{v_{\mathrm{g},\phi}^2}{R} \simeq \frac{\partial \Phi}{\partial R} + \frac{1}{\rho_{\mathrm{g}}}\frac{\partial P}{\partial R}. \tag{1.13}$$

For the case of a non self-gravitating disc, the central object contribution to the gravitational field, in the limit of thin discs, can be expressed as

$$\frac{\partial \Phi}{\partial R} = \frac{\mathcal{G}MR}{(R^2+z^2)^{3/2}} \underset{z \ll R}{\approx} \frac{\mathcal{G}M}{R^2}. \tag{1.14}$$

Therefore, the disc rotation is characterized by an azimuthal component of velocity given by the Keplerian velocity $v_{\mathrm{k}} = \sqrt{\mathcal{G}M/R}$ with an angular velocity of $\Omega_{\mathrm{k}} = \sqrt{\mathcal{G}M/R^3}$ with a correction term related to the pressure. Since the pressure near the disc mid-plane normally decrease outward, the second term of the right-hand-side of Eq. 1.13 is negative. This implies that the azimuthal velocity of the gas is slightly sub-Keplerian. To quantify this deviation, let us assume power-law profiles for the surface density ($\Sigma_{\mathrm{g}} \propto R^{-p}$) and temperature ($T \propto R^{-q}$). Substituting into Eq. 1.13 yields,

$$v_{\mathrm{g},\phi} = v_{\mathrm{k}} \left[1 + \zeta \left(\frac{H_{\mathrm{g}}}{R}\right)^2\right]^{1/2}, \tag{1.15}$$

where

$$\zeta \equiv \frac{\partial \ln P}{\partial \ln R} = -\left(p + \frac{q}{2} + \frac{3}{2}\right). \tag{1.16}$$

In typical discs, $p = 1$ and $q = 1/2$, $\zeta \simeq -11/4 < 0$. However, the deviation from Keplerian motion is of order $(H_{\mathrm{g}}/R)^2 \ll 1$. The physical origin of this relation is easily understood: the gas does not need to orbit at Keplerian rotation due to the additional support from the thermal pressure forces, whose magnitude is related to the local pressure gradient. In Chapter 2 we will show that, although the deviation from the Keplerian motion appears negligible, the sub-Keplerian azimuthal motion of gas in circumstellar discs plays an important role in the dust dynamics.

1.2.1.3 Angular momentum conservation

Let us finally discuss the azimuthal component of the Navier-Stokes equation, which encapsulate the angular momentum conservation. According to the thin-disc approximation, it is convenient to integrate the fluid equations in the vertical direction. In cylindrical coordinates, since the vertical component of the velocity can be neglected, the continuity equation is given by

$$\frac{\partial \Sigma_{\mathrm{g}}}{\partial t} + \frac{1}{R}\frac{\partial}{\partial R}(R\,\Sigma_{\mathrm{g}}\,v_{\mathrm{g},R}) = 0. \tag{1.17}$$

In the case of accretion disc, the radial velocity should be negative due to the infalling material towards the central object. Integrating the shear viscosity coefficient on the vertical direction, we define the vertical averaged kinematic viscosity ν as

$$\nu\Sigma_{\mathrm{g}} = \int_{-\infty}^{\infty} \eta\, dz, \tag{1.18}$$

which allows to evaluate the vertical integral stress tensor whose only non-vanishing component as

$$T_{R\phi} = \nu \Sigma_{\rm g} R \frac{\partial \Omega}{\partial R}. \tag{1.19}$$

The vertical integral of the Navier-Stokes equation (Eq. 1.4) is given by

$$\Sigma_{\rm g} \left[\frac{\partial \mathbf{v}_{\rm g}}{\partial t} + (\mathbf{v}_{\rm g} \cdot \nabla) \mathbf{v}_{\rm g} \right] = -(\nabla P - \nabla \cdot T) - \Sigma_{\rm g} \nabla \Phi, \tag{1.20}$$

where P indicates here the vertical integral of the pressure.

Along the azimuthal direction, the pressure and gravitational force does not give any contribution due to axisymmetry of the disc. Moreover, the term on the left-hand side $\propto \partial v_{{\rm g},\phi}/\partial \phi = 0$. The azimuthal component of the Navier-Stokes equation integrated in the vertical direction is therefore given by

$$\Sigma_{\rm g} \left[\frac{\partial v_{{\rm g},\phi}}{\partial t} + v_{{\rm g},R} \left(\frac{v_{{\rm g},\phi}}{R} + \frac{\partial v_{{\rm g},\phi}}{\partial R} \right) \right] = \nabla \cdot T|_{\phi}. \tag{1.21}$$

Multiplying each terms by R and using the continuity equation (Eq. 1.17), the equation of the conservation of angular momentum for a viscous fluid disc is

$$\frac{\partial}{\partial t} \left(\Sigma_{\rm g} R v_{{\rm g},\phi} \right) + \frac{1}{R} \frac{\partial}{\partial R} (R v_{{\rm g},R} \Sigma_{\rm g} R v_{{\rm g},\phi}) = \frac{1}{R} \frac{\partial}{\partial R} \left(\nu \Sigma_{\rm g} R^3 \frac{\partial \Omega}{\partial R} \right). \tag{1.22}$$

This equation plays a key role in accretion disc theory and allows us to figure out how viscous forces are crucial for the accretion. The left-hand side term represents the Lagrangian derivative of the angular momentum per unit area $\Sigma_{\rm g} R v_{{\rm g},\phi}$ of an annulus at distance R from the central object. The second term in the left-hand side is the net rate of angular momentum loss from this area due to advection by the radial infalling material. The right-hand side of Eq. 1.22 is the net[2] torque exerted by viscous force or, equivalently, the net rate of angular momentum loss from this unit area. Multiplying the right-hand term by $2\pi R\, dR$, the new term now represents the net torque on an annulus of width dR. Thus, the torque at R or, equivalently, the flux of angular momentum across an annulus at distance R from the central object is given by

$$G(R) = 2\pi R^2\, T_{R\phi} = 2\pi \nu \Sigma_{\rm g} R^3 \frac{\partial \Omega}{\partial R}, \tag{1.23}$$

where the sign convention is such that a positive $G(R)$ implies inward flux. This results is reasonable and can be checked by evaluating the torque at R as the viscous force times R, where the viscous force is the product of the viscous stress and the area of the side of the annulus that rubs against the adjoining annulus (see Clarke & Carswell 2007). As expected, in the case of a solid body rotating disc, Eq. 1.23 states that $G(R)$ would vanish since the shear viscosity does not act in rigidly rotating disc. In the case of Keplerian disc ($\partial \Omega_{\rm k}/\partial R = -3\Omega_{\rm k}/2R$), the angular momentum is transported outward from small to large radii with a net torque at R given by: $G(R) = -3\pi \nu \Sigma_{\rm g} R^2 \Omega_{\rm k}$. Therefore, for a Keplerian disc, the angular momentum is transferred to material at larger distance from the central object and, as a consequence, most of the mass can accrete onto the central object.

The steady state radial gas velocity can be obtained by multiplying each term of Eq. 1.21 by R and putting $\partial v_{{\rm g},\phi}/\partial t = 0$, yielding

$$v_{{\rm g},R} = \frac{1}{R \Sigma_{\rm g} \left(R^2 \Omega \right)'} \frac{\partial}{\partial R} \left(\nu \Sigma_{\rm g} R^3 \frac{\partial \Omega}{\partial R} \right), \tag{1.24}$$

where $(R^2\Omega)' = \partial (R^2 \Omega)/\partial R$. For the case of Keplerian disc, the radial gas velocity is given by

$$v_{{\rm g},R} = -\frac{3}{\Sigma_{\rm g} R^{1/2}} \frac{\partial}{\partial R} \left(\nu \Sigma_{\rm g} R^{1/2} \right). \tag{1.25}$$

[2]It is worth pointing out that, assuming rotation velocity $\Omega(R)$ decreasing with radius, the annulus should experience a spin-up torque due to the viscous interaction with faster-moving material at smaller radius and a spin-down torque as a result of the viscous interaction with slower-moving material at larger radius.

Combining the angular momentum equation (Eq. 1.22) with the continuity equation (Eq. 1.17), it is possible to obtain the evolution equation for the surface density,

$$\frac{\partial \Sigma_{\rm g}}{\partial t} = -\frac{1}{R}\frac{\partial}{\partial R}\left[\frac{1}{(R^2\Omega)'}\frac{\partial}{\partial R}\left(\nu\Sigma_{\rm g}R^3\frac{\partial \Omega}{\partial R}\right)\right]. \tag{1.26}$$

This equation describes the temporal evolution of the gas surface density $\Sigma_{\rm g}$, which, as clearly seen, is closely dependent on the disc viscosity expressed by the kinematic viscosity ν. For the case of Keplerian disc, the equation can be rewritten as

$$\frac{\partial \Sigma_{\rm g}}{\partial t} = \frac{3}{R}\frac{\partial}{\partial R}\left[R^{1/2}\frac{\partial}{\partial R}\left(\nu\Sigma_{\rm g}R^{1/2}\right)\right]. \tag{1.27}$$

Generally, assuming that Ω and ν are independent on $\Sigma_{\rm g}$, it is clearly seen from a dimensional analysis of Eq. 1.26 or 1.27 that the disc surface density evolves on a timescale $t_{\rm visc}$ of the order of

$$t_{\rm visc} \equiv \frac{R^2}{\nu}. \tag{1.28}$$

Moreover, it can be shown that, assuming that the viscosity is constant, Eq. 1.27 takes the form a diffusion equation for the surface density evolving under the action of the viscosity with a diffusion timescale expressed in Eq. 1.28. It is instructive to express the latter quantity in units of the dynamical time-scale $t_{\rm dyn} = \Omega^{-1}$. The resulting ratio is given by

$$\frac{t_{\rm visc}}{t_{\rm dyn}} \equiv \frac{R^2\,\Omega}{\nu}. \tag{1.29}$$

This ratio is the Reynolds number of the flow which is essentially a measure of the ratio of inertial forces to viscous forces acting on a fluid. It is mostly used to characterize different flow regimes: at low Reynolds numbers, where viscous forces are dominant, the flow is characterized by a laminar flow while for high Reynolds numbers inertial forces tend to produce turbulence phenomena (i.e. turbolent flow).

1.2.2 Time dependent solutions

We can wonder if there exist some simple solutions for the evolution of the surface density of a Keplerian discs associated to the equations of motion derived above. Typically, by looking at Eq. 1.27, due to the very complex dependencies that the viscosity ν might have with other fluid quantities and time, the general solution should be computed numerically. However, simple solutions can be analytically calculated by assuming that the viscosity ν is a simple function of radius (Pringle 1981; Hartmann 2009).

A first simple diffusive-like solution can be derived by assuming a constant viscosity along radius. Let us adopt an initial condition for the surface density characterised by an infinitesimally thin ring of mass M centered at some radius R_0, given by

$$\Sigma_{\rm g}(R, t=0) = \frac{M}{2\pi R_0}\delta(R - R_0), \tag{1.30}$$

where $\delta(R)$ denotes the Dirac δ-function. The class of solutions of Eq. 1.27 with homogeneous viscosity and this initial condition is known as "spreading ring" solution and has been first derived by (Lynden-Bell & Pringle 1974). The surface density evolution is given by

$$\Sigma_{\rm g}(R,t) = \frac{M}{\pi R_0^2}\frac{x^{-1/4}}{\tau}\exp\left[-\frac{1+x^2}{\tau}\right]I_{1/4}\left(\frac{2x}{\tau}\right), \tag{1.31}$$

where $I_{1/4}$ is a modified Bessel function of the first kind, $x = R/R_0$ and $\tau = 12\nu t/R_0^2$. The time evolution of the surface density described by this solution is plotted in the left panel of Fig. 1.3,

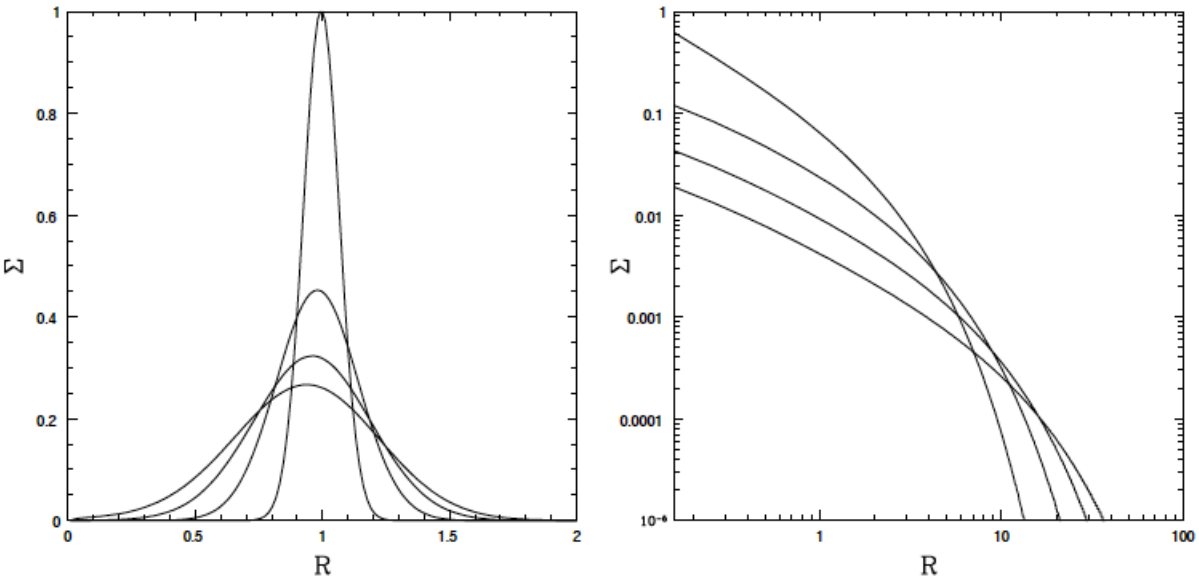

Figure 1.3: Surface density evolution of (left) a Keplerian spreading ring and (right) a radially truncated disc with viscosity $\nu \propto R$ for $\tau = [0.01, 0.05, 0.1, 0.15]$ evaluated in Eqs. 1.31 and 1.33 (with $\gamma = 1$), respectively. Image taken from Lodato (2008).

where can be clearly noticed the diffusive evolution of the surface density evolution. Under the action of the spatially homogeneous viscous forces, the ring spreads both inwards and outwards due to the outward angular momentum transport, as expected by requiring the conservation of total angular momentum. The transition between inward and outward motion occurs at a radius of $\sim R_0\, t/t_{\rm visc}$ with $t_{\rm visc}$ expressed in Eq. 1.28. The location where the mass flux changes sign increases with time, ensuring accretion onto the central star on a timescale $\sim t_{\rm visc}$, as expected on dimensional grounds (see Sect. 1.2.1.3). Therefore, for $t \to \infty$, all the mass is accreted and a negligible amount of mass transports angular momentum to infinitely large radii.

Another class of solutions can be derived in cases where the viscosity has a simple power law profile with radius, i.e. $\nu \propto R^\gamma$. Let us assume an initial density profile given by

$$\Sigma_{\rm g}(R, t=0) = \frac{C}{3\pi\nu(R)} \exp\left[-\left(\frac{R}{R_1}\right)^{2-\gamma}\right], \tag{1.32}$$

where C and γ are constants and R_1 is an arbitrary truncation radius of the disc. In this case, it is possible to derive a self-similarity solution, i.e. with the same functional form at all times, given by (Lynden-Bell & Pringle 1974; Hartmann et al. 1998)

$$\Sigma_{\rm g}(R, t) = \frac{C}{3\pi\nu(R)}\, T^{-(5/2-\gamma)/(2-\gamma)} \exp\left[-\left(\frac{R}{R_1}\right)^{2-\gamma}\right], \tag{1.33}$$

where the non-dimensional time T is defined as

$$T \equiv 1 + \frac{t}{t_\nu} \quad \text{with} \quad t_\nu \equiv \frac{1}{3\,(2-\gamma)^2} \frac{R_1^2}{\nu(R_1)}. \tag{1.34}$$

As for the spreading ring, this solution (see right panel of Fig. 1.3) shows that the disc spreads both inwards and outwards, producing a (time decreasing) accretion onto the central protostar (Hartmann 2009). The transition between inward and outward moving portion of the disc occurs at a radius $\sim R_1\,(t/t_\nu)^{1/(2-\gamma)}$ that moves outwards with time. Note that the typical timescale over which the disc evolves t_ν has the same scaling with viscosity and radius of that expressed in Eq. 1.28. This class of solutions is particularly useful to derive the level of disc viscosity starting by measuring the radial extent of the gaseous disc by means of spatially resolved observations

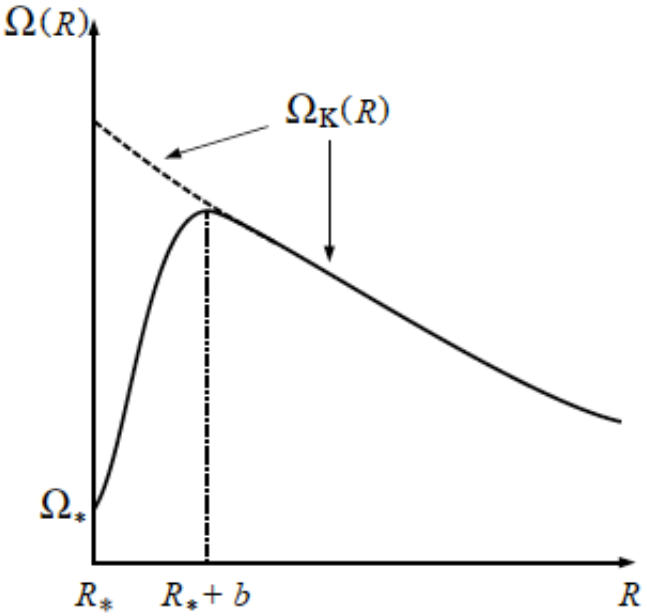

Figure 1.4: Illustration of the angular velocity profile across the boundary layer, where the angular velocity of the gas decreases with decreasing radius to match the speed of the slowly rotating central star. Image taken from Frank et al. (1985).

(such as observations of near-infrared scattering emission or spectral line emission of gas species, see Chapter 3) and an estimate of the viscous timescale. Moreover, this class of solutions has been extensively used to derive the disc properties by comparing high-resolution observations with the predictions of models of flared discs (see further details in Sect. 3.3).

1.2.3 Steady state solutions

Protostellar discs are generally not in steady state. They can be subject to a range of instabilities that can strongly affect the disc structure and evolution. However, it is instructive to derive the steady state solutions in order to get more insight on the angular momentum flux and the mass accretion rate. Multiplying all the terms by $2\pi R$ to refer to annular quantities and setting the derivative to zero, the integration of Eq. 1.17 yields

$$2\pi\Sigma_{\rm g} R v_{{\rm g},R} = -\dot{M}, \tag{1.35}$$

where $\dot{M}$ is the constant (with R) rate of mass flux across an annulus of the disc. The minus sign indicates that $\dot{M} > 0$ when the fluid is moving inwards, i.e. $v_{{\rm g},R} < 0$. Assuming steady state for the angular momentum transport, Eq. 1.22 gives the value of the constant net angular momentum flux across an annulus of the disc,

$$\dot{J} = \dot{M} R^2 \Omega + 2\pi\nu\Sigma_{\rm g} R^3 \frac{\partial\Omega}{\partial R}, \tag{1.36}$$

where the first term indicates the angular momentum flux advected with the inward flow and the second term represents the viscous torque. This term produces, for discs with a decreasing angular velocity with radius such as Keplerian discs ($\partial\Omega_{\rm k}/\partial R = -3\Omega_{\rm k}/2R$), an outward transport of the angular momentum.

To get more insight on the value of the angular momentum flux in steady state, we assume the presence of a boundary layer defined as a region where the disc connects to the central object (see Fig. 1.4). In this layer, the gas angular velocity decreases with decreasing radius to match the speed of the slowly rotating central object. Therefore, at the outer edge of this region, the angular velocity profile flattens, i.e. at radius $R_{\rm in} = R_\star + b \approx R_\star$, where b is the boundary layer

extension, $\Omega'(R_{\rm in}) = 0$. This assumption is mostly known as "no torque" condition due to the fact that viscous flux of angular momentum vanishes at $R_{\rm in}$ (Frank et al. 1985). As a result, $\dot{J} = \dot{M} R_{\rm in}^2 \Omega(R_{\rm in})$ at the inner boundary of the disc. Inserting this in Eq. 1.36, we obtain

$$2 \left| \frac{\partial \ln \Omega}{\partial \ln R} \right| \pi \nu \Sigma_{\rm g} = \dot{M} \left(1 - \sqrt{\frac{R_{\rm in}}{R}} \right). \tag{1.37}$$

At large radii, $R \gg R_{\rm in}$ the accretion rate of a Keplerian disc in steady state is given by

$$\dot{M} = 3\pi \nu \Sigma_{\rm g}. \tag{1.38}$$

Note that the surface density in steady state disc is $\propto \nu^{-1}$. This relation has been extensively used in observational investigations of protostellar discs to derive estimates of the disc viscosity starting from the measured value of the mass accretion rate (Andrews & Williams 2007a,b; Andrews et al. 2009; Pinte et al. 2016). The radial gas velocity in a disc in steady state is therefore given by

$$v_{{\rm g},R} = -\frac{3\nu}{2R} \left(1 - \sqrt{\frac{R_{\rm in}}{R}} \right)^{-1} \approx -\frac{3\nu}{2R}, \tag{1.39}$$

which is of the order of ν/R, as expected from the more detailed analysis shown in Sect. 1.2.1.3.

1.2.4 Energetics of accretion discs

The effect of viscosity on the energy balance in accreting disc is twofold: viscous torques convect energy between neighbouring annuli across the disc and dissipate mechanical energy into heat. As accretion takes place, a significant amount of gravitational potential energy has to be dissipated by viscous forces. The energy dissipation rate per unit surface by both side of the disc at R is therefore given by (Frank et al. 1985)

$$\mathcal{D}(R) = \frac{-G\Omega'}{2\pi R} = \nu \Sigma_{\rm g} \left(R\Omega' \right)^2 = \frac{9}{4} \nu \Sigma_{\rm g} \Omega^2, \tag{1.40}$$

where the last equality holds for Keplerian discs. Assuming that the disc in steady state, we can evaluate the total power emitted by a Keplerian disc with an infinite extension,

$$L_{\rm disc,visc} = \int_{R_{\rm in}}^{\infty} \mathcal{D}(R) 2\pi R dR = \int_{R_{\rm in}}^{\infty} \frac{3}{2} \frac{\mathcal{G}M}{R^2} \dot{M} \left(1 - \sqrt{\frac{R_{\rm in}}{R}} \right) dR = \frac{1}{2} \frac{\mathcal{G}M\dot{M}}{R_{\rm in}}, \tag{1.41}$$

where we have used Eq. 1.37. Importantly, the energy dissipated for accreting mass is equal to just one half of the total potential energy. The remaining energy is used to keep the disc in Keplerian rotation and a fraction of this energy is eventually dissipated in the boundary layer, where the gas loses the kinetic energy and accretes the protostar (Frank et al. 1985).

1.3 Angular momentum transport

The nature of the angular momentum transport covers a key role on the accreting mechanism. Looking back at the derivation of the angular momentum transport and surface evolution, the timescale on which evolution occurs depends directly on the viscosity. Therefore, it is necessary to discuss the origin of the viscous torques to gain insights on the angular momentum transport and the accreting properties.

The first candidates as a source of angular momentum transport is the standard kinetic viscosity due to collisions between gas molecules in the shear flow. It can be shown (e.g. Lodato 2007, 2008) that, for a typical protostellar disc with an outer radius of ≈ 50 au, a mass of $\approx 0.005\, M_\odot$ and an aspect ratio of $H_{\rm g}/R \approx 0.1$, the typical viscous timescale (Eq. 1.28) is of the order of 10^{13} years. According to this evaluation, the accretion of the previously mentioned disc would occur in more that the Hubble time. Therefore, this type of viscosity does not provide accretion at a reasonably rate (Fedele et al. 2010), and thus it is necessary to introduce other sources of transport in accretion discs.

1.3.1 Turbulent transport

Previous analysis has also shown that the ratio of the viscous to dynamical timescale, which is equal to the Reynolds number of the flow, is substantially large (see Eq. 1.29). Therefore, it is reasonable to expect that the flow in accretion disc may be characterized by turbulence. It is well known that the transport mechanism in turbulent flows is more effective with respect to laminar flows due to the different way to exchange angular momentum between fluid elements. While in non turbulent flows the angular momentum is exchanged through collisions of individual gas molecules, the turbulence phenomena produce a mixing of fluid elements which enhances the transport of angular momentum (Turner et al. 2014).

Therefore, assuming that there exist some mechanism that extract energy from the mean flow and inject the latter in the fluctuations, we need to compute directly the torque exerted by these fluctuations in order to measure the flux of angular momentum. Assuming that the disc is inviscid, we can rewrite the Navier-Stokes equation by separating the fluid velocity in two contributions, a secularly varying average component, $v_{\rm g}$, and a rapidly varying fluctuating component, $u_{\rm g}$. It can be shown (e.g. Balbus & Hawley 1998) that in cylindrical coordinates, the vertically integrated Navier-Stokes equation takes the following form

$$\frac{\partial}{\partial t}\left(\Sigma_{\rm g} R v_{{\rm g},\phi}\right) + \frac{1}{R}\frac{\partial}{\partial R}(R v_{{\rm g},R}\Sigma_{\rm g} R v_{{\rm g},\phi}) = -\frac{1}{R}\frac{\partial}{\partial R}\left[\Sigma_{\rm g} R^2 \langle u_{{\rm g},R} u_{{\rm g},\phi}\rangle\right], \tag{1.42}$$

where the angle brackets indicate a vertical and azimuthal average. We can see that, if a positive correlation between u_ϕ and u_R is present, velocity fluctuations provide a stress (the so-called Reynolds stress) given by

$$T^{\rm Reyn}_{R\phi} = -\Sigma_{\rm g}\langle u_{{\rm g},R} u_{{\rm g},\phi}\rangle. \tag{1.43}$$

Turbulent fluctuations directly provide a stress responsible for the transport angular momentum in a purely hydrodynamic flow. For example, shear-driven turbulence is characterized by a high degree of correlation between the radial and azimuthal velocity fluctuations. However, it can be shown that a Keplerian discs are apparently stable in linear analysis to the development of turbulence (Pringle & King 2007). This does not rule out the possibility that the Keplerian shear flow might be non-linearly unstable (although there is no mathematical proof). Importantly, Nath & Mukhopadhyay (2016) found recently that the presence of stochastic noise leads to pure hydrodynamic linear instability in accretion flows.

Additionally, the source of transport of angular momentum can be also caused by an external driver, which might produce disturbances of the velocity field with non-zero correlation between the radial and azimuthal components. In particular, if the disc is magnetized, the perturbed magnetic field $\mathbf{B}$ provides another source of transport, leading to the so-called "Maxwell" stress,

$$T^{\rm M}_{R\phi} = -\Sigma_{\rm g}\langle u_{A,R} u_{A,\phi}\rangle, \tag{1.44}$$

where $\mathbf{u_A} = \mathbf{B}/\sqrt{4\pi\rho_{\rm g}}$ is the Alfvén velocity. If the disc is self-gravitating, the perturbed gravitational field $\mathbf{g}$ provide an additional stress term in the form

$$T^{\rm grav}_{R\phi} = -\Sigma_{\rm g}\langle u^g_R u^g_\phi\rangle, \tag{1.45}$$

where $\mathbf{u^g} = \mathbf{g}/\sqrt{4\pi\mathcal{G}\rho_{\rm g}}$.

In steady state (see Sect. 1.2.3), the angular momentum flux in a magnetized and self-gravitating inviscid disc is therefore given by

$$\dot{J} = \dot{M} R^2\Omega - 2\pi\Sigma_{\rm g} R^2\left(\langle u_{{\rm g},R} u_{{\rm g},\phi}\rangle + \langle u_{A,R} u_{A,\phi}\rangle + \langle u^g_R u^g_\phi\rangle\right). \tag{1.46}$$

This expression accounts for the angular momentum flux advected with the inward flow and generated from internal stresses, including Reynolds stress, Maxwell stress, and stress arising from perturbations in the gravitational fields.

Therefore, since the equation of angular momentum conservation for an inviscid turbulent flow (Eq. 1.42) is similar to Eq. 1.22, the stress provided by turbulent fluctuations plays the same role as a viscous stress from Navier-Stokes equations. However, this does not imply that the angular momentum transport induced by turbulence can be interpreted as a viscous process.

1.3.2 The issue of the "anomalous" viscosity

Even if the equation of angular momentum conservation for an inviscid turbulent flow can be put in a form analogous to Eq. 1.22 as required in a viscous framework, we need to verify if the stress provided by velocity fluctuations is a viscous stress. To answer this question we need to take into account two aspects: First, turbulence associated stresses need to dissipate energy as a viscous stress would do (i.e. according to Eq. 1.40). Second, one needs to understand whether fluctuations are readily dissipated locally or instead whether they are able to travel significantly before being dissipated. Although this issue has not yet been settled completely, Balbus & Papaloizou (1999) have shown that the energy dissipation associated to the Reynolds and Maxwell stress is the same as would be dissipated in a viscous process. However, due to the fact that gravity is a global phenomena, the gravitational stresses might introduce anomalous energy transport (see Sect. 4.2.4 for further details). Numerical investigations performed to date found that the level of non-local transport induced by gravitational instabilities is negligible for low mass discs, thus suggesting that "gravoturbulent" transport does behave to first approximation as an "anomalous" viscosity (Lodato & Rice 2004, 2005; Forgan et al. 2011).

However, the magnitude of turbulent transport in protostellar discs is still largely unknown. In this context, a simple construction of this "anomalous" viscosity can be simply carried out starting from some characteristic features of turbulence. Noting that the maximum size of a turbulent cell is given by the disc thickness $H_{\rm g}$ (in the case of isotropic turbulence) and assuming that the turbulence is subsonic (otherwise, supersonic turbulence would easily dissipate through shocks), Shakura & Sunyaev (1973) introduced a prescription for the "turbulent" kinematical viscosity in the form

$$\nu = \alpha c_{\rm s} H_{\rm g}, \tag{1.47}$$

where α is an unknown dimensionless parameter that embodies all the uncertainties about the nature of the viscous mechanism (see Lodato 2007, 2008 for further details). The form of "turbulent" kinematical viscosity expressed above follows from the assumptions that the magnitude of viscosity driven by local turbulence can be roughly given by $\nu \sim \hat{v} l$, where l is the typical size of the largest eddies ($l \lesssim H_{\rm g}$) and $\hat{v}$ is the typical turbulent velocity which is expected, as mentioned above, to be lower than the local sound speed. Thus, it can be clearly seen that $\alpha < 1$.

In addition, the α-prescription can be expressed in an alternative form based on a dimensional analysis of the stress tensor $T_{R\phi}$. Since the latter has the physical dimension of a pressure, which in turn has the physical dimension of a volume density times the square of a velocity, it is possible to express the vertical integral of the stress tensor $T_{R\phi}$ using the parameter α as

$$T_{R\phi} = \frac{\partial \log \Omega}{\partial \log R} \alpha \Sigma_{\rm g} c_{\rm s}^2, \tag{1.48}$$

where the term $\partial \log \Omega / \partial \ln R$ is related to the proportionality of the viscous stress to the rate of strain. However, this "anomalous" viscosity induced by turbulence does not generally depend on the local rate of strain. In light of this, in order to include all the forms of kinematical viscosities introduced here, the rate of strain can be considered as a multiplicative factor that can be embodied in α. This parameterization of "anomalous" viscosity is called the *α-prescription* and it has been extensively used in a number of studies about accretion discs. In principle, any source of local transport can be expressed in terms of "anomalous" viscosity by using the α-prescription.

Using the prescription for the viscosity, we can estimate how large α should be to reproduce the observed timescale for disc evolution (Fedele et al. 2010). Starting from Eq. 1.28, the viscous

timescale can be expressed as

$$t_{\rm visc} = \frac{R^2}{\nu} = \left(\frac{H_{\rm g}}{R}\right)^{-2} \frac{1}{\alpha\Omega}. \tag{1.49}$$

Thus, in order to reproduce the expected disc lifetime, for typical disc parameters α should be in the range $\left[10^{-4}, 10^{-2}\right]$. From an observational point of view, the measurement of α can be derived by analysing accretion indicators such as ultraviolet excess and/or Hα emission over the stellar photosphere (see Eq. 1.38, Hartmann et al. 1998) and/or by spatially resolved (sub-)millimetre observations (Andrews & Williams 2007a,b; Andrews et al. 2009). Typical values are of the order of $\left[10^{-3}, 10^{-2}\right]$, though there remain uncertainties in its exact value due to observational limitations.

Interestingly, if the disc is in thermal equilibrium, it is possible to derive an useful relation between the viscosity coefficient and the cooling time-scale. Assuming that there is no other source of heating for the disc, the heating rate provided by viscosity (see Eq. 1.40) has to balance the cooling rate. The cooling rate per unit surface is given by

$$C = \frac{\Sigma_{\rm g} c_{\rm s}^2}{\gamma\left(\gamma-1\right)\tau_{\rm cool}}, \tag{1.50}$$

where γ is the ratio of the specific heats and $\tau_{\rm cool}$ is the cooling timescale. Equating the cooling rate with the viscous dissipation term expressed in Eq. 1.40 and using the α-prescription (Eq. 1.48), we obtain

$$\alpha = \left|\frac{\partial \ln \Omega}{\partial \ln R}\right|^{-2} \frac{1}{\gamma\left(\gamma-1\right)\Omega\tau_{\rm cool}}. \tag{1.51}$$

This simple expression has been extensively used in numerical investigations to test the locality of the angular momentum transport induced by gravitational instabilities in cooling massive discs (adopting a simple cooling prescription, see Sect. 4.2.5), once the disc has reached thermal equilibrium (e.g. Lodato & Rice 2004).

However, according to the turbulence theory, it is necessary that a mechanism exists to drive the turbulence through the injection of energy at the largest scales. In this context, the role of disc instabilities may play a crucial role in sustaining the turbulence and in providing a source of transport of angular momentum and energy in accretion discs.

There are essentially two types of fluid instabilities which are more likely to affect the transport mechanism: the magneto-rotational (MRI) and the gravitational instabilities. The former instabilities, firstly identified by Chandrasekhar (1961) and applied to accretion flows by Balbus & Hawley (1991), is thought to be the most likely sources of angular momentum and energy transport in accretion disc. It can be shown that the condition for MRI instability in the regime of ideal magnetohydrodynamics (MHD) is given by

$$\frac{\partial \Omega^2}{\partial R} < 0. \tag{1.52}$$

This condition is generally satisfied in accreting disc (except in the disc boundary layer, see Fig. 1.4), assuming a perfect coupling between the gas and the magnetic field (i.e. it is needed a strong ionization fraction $\gtrsim 10^{-13}$, corresponding to a disc temperature $\gtrsim 1000\,K$). This kind of instability was historically considered the principal source of turbulence in protoplanetary discs, producing a transport of angular momentum and energy at a reasonable rate (Balbus & Hawley 1991). However, in the case of weakly ionized discs, the MRI is likely to be ineffective especially in the outer disc regions close to the midplane (Bai 2011; Bai & Stone 2011), leading to the formation of large "dead zones" in which MRI is not effective (Gammie 1996; Martin & Lubow 2011, 2014). These dead-zones could produce axisymmetric (mostly known as zonal flows) and non-axisymmetric density structures which might be identified with ALMA (Flock et al. 2015; Ruge et al. 2016).

On the other hand, if circumstances allow, it is possible that the accretion discs become unstable to the effect of their self-gravity giving rise to gravitational instabilities which may provide a source of transport and drive accretion. We will focus on this topic in Chapter 4.

1.4 Thermal structure of the disc

Simple energy arguments based on the observed bolometric luminosities indicate that discs are primarily irradiated by the central star, especially in the outer regions (e.g. Adams & Shu 1986). Typically, dusty protostellar envelopes are extremely opaque at the ultraviolet wavelengths at which most of the energy of protostars is emitted. The energy emitted by the central star is therefore absorbed in the optically thick disc. Typically, the dust has lower temperatures than that characteristic of the central star, and thus it reradiates the absorbed energy at longer wavelengths where the dust opacity is lower (see further details in Chapter 3). However, while models of irradiated discs can explain the infrared emission of many protostellar discs (Kenyon & Hartmann 1987; Chiang & Goldreich 1997), some discs are much more luminous at shorter wavelengths with respect to the amount of irradiation radiated from the central star. They exhibit strong excess optical and ultraviolet continuum emission (Basri & Bertout 1989) that can be explained by the accretion of mass from a circumstellar disc (see Sect. 1.2.4). The infall of material onto the central star produces high flux of optical and ultraviolet continuum intensities and strong optical emission lines, which can be used to derive the star accretion rate from the disc in the inner regions. Therefore, we can identify two main source of heating in a protostellar discs:

- radiation intercepted from the star,
- dissipation of gravitational potential energy of accreting mass by the viscous forces.

Both sources contributes significantly to the thermal balance of the disc. Discs for which the thermal structure is dominated by the intercepted stellar radiation are described as passive discs. Otherwise, if the luminosity is due to the internal heating related to the accretion process, discs are described as active discs. Since the accretion is strictly dependent on the mass accretion rate, it is expected that at early epochs discs are dominated by internal heating, while at later times the protostar mostly irradiates the disc. The general picture of a protostellar discs is that the two sources of heating are both active in the discs and their efficiency depends on the disc structure and stellar emission properties (D'Alessio et al. 1998).

1.4.1 Viscous heating

We can get some insights on the internal heating produced by the dissipation of the gravitational potential energy by using the result that the viscosity generates dissipation of energy in the disc at a rate $\mathcal{D}(R)$ per unit time and area (see Eq. 1.40). If we simply assume that all the disc radiates all this power from its two-sided surface like a blackbody, i.e. $\mathcal{D}(R) = 2\sigma_{\rm SB}T^4_{\rm disc}$, the surface temperature of the disc is given by

$$T^4_{\rm disc,visc} = \frac{3\mathcal{G}M\dot{M}}{8\pi\sigma_{\rm SB}R^3}\left(1-\sqrt{\frac{R_{\rm in}}{R}}\right), \tag{1.53}$$

where $\sigma_{\rm SB}$ is the Stefan-Boltzmann constant. The disc total luminosity, assuming that the disc is in thermal equilibrium and optically thick, is expressed in Eq. 1.41. In the outer disc regions, i.e. $R >> R_{\rm in}$, the temperature profile of an actively accreting disc is $T_{\rm disc,visc} \propto R^{-3/4}$. Importantly, the temperature profile does not depend directly on the viscosity. Since the mass accretion rate can be inferred through observations of the ultraviolet excess and/or the Hα emission, it is possible to estimate the typical temperature of an active disc without knowing the origin of the disc angular momentum transport. However, since viscosity regulates the mass accretion rate $\dot{M}$ (see Eq. 1.38), the amount of the angular momentum transport obviously implicitly enters in the determination of the temperature profile.

1.4.2 Stellar irradiation

Starlight absorbed at the disc surface layers heats dust grains, which then re-emit some of that energy deeper into the disc interior that warms the midplane. The exact calculation of the temperature profile requires therefore a detailed knowledge of the shape of the disc surface (see the appendix of Kenyon & Hartmann 1987). Assuming thermodynamic equilibrium, i.e. absorbed flux is equal to the flux radiated by the disc, for a flat thin discs (i.e. $H(R) = \text{const}$), it can be shown that the temperature profile is given by

$$\left(\frac{T_{\rm disc,irr}}{T_\star}\right)^4 = \frac{1}{\pi}\left[\sin^{-1}\left(\frac{R_\star}{R}\right) - \left(\frac{R_\star}{R}\right)\sqrt{1-\left(\frac{R_\star}{R}\right)^2}\right], \tag{1.54}$$

where $T_\star$ is the effective temperature of the central star. For $R >> R_\star$, the temperature profile has the same scaling with R as an active disc, i.e. $T_{\rm disc,irr} \propto R^{-3/4}$ (see Eq. 1.53). The total disc luminosity is given by

$$L_{\rm disc,irr} = 2\int_{R_\star}^{\infty} 2\pi R\,\sigma_{\rm SB} T^4_{\rm disc,irr}(R)dR = \frac{1}{4}L_\star. \tag{1.55}$$

For flared discs the outer regions intercept a larger fraction of stellar emission, flattening the temperature profile and increasing the total disc luminosity. Using simple energetic arguments, it can be shown that the temperature profile of a vertically isothermal flared disc with a shape characterized by grazing angle $\alpha_{\rm irr}$[3] is (Kenyon & Hartmann 1987)

$$\frac{T_{\rm disc,irr}}{T_\star} = \left(\frac{R_\star}{R}\right)^{1/2}\alpha_{\rm irr}^{1/4}. \tag{1.57}$$

Assuming that $T_{\rm disc,irr}$ is a power-law function of R and the disc is vertically isothermal, the grazing angle is proportional to $H_{\rm g}/R = c_{\rm s}/v_{\rm k} \propto T^{1/2}_{\rm disc,irr}R^{1/2}$. Thus, the radial profile of the temperature of an irradiated disc can be well approximated as

$$\frac{T_{\rm disc,irr}}{T_\star} = R_\star^{1/2}R^{-3/7}. \tag{1.58}$$

As expected, this profile is substantially flatter than the one expected for a flared disc ($T_{\rm disc,irr} \propto R^{-3/4}$). By comparing the expected temperature profile of an active discs with the one expected from a flared passive disc, the relative importance scale as $\sim R^{-3/7+3/4} = R^{9/28}$. We then expect viscous heating to be most important in the inner disc, whereas irradiation is the main heating source at sufficiently large radii, as long as the disc is sufficiently opaque to the incoming radiation from the central region (the opacity budget is dominated by small particles which need to be suspended to significant vertical extent in order to absorb radiation). Generally, for low mass accretion rates, accretion heating is often neglected beyond a few au. For example, D'Alessio et al. (1998) found that neglecting accretion heating is good approximation outside 2 au for mass accretion rates less than $\sim 10^{-8}\,{\rm M_\odot\,yr^{-1}}$. It is worth recalling that typical stellar accretion rates range from 10^{-7} to $10^{-10}\,{\rm M_\odot\,yr^{-1}}$ (Gullbring et al. 1998; Muzerolle et al. 2000), further strengthening the idea that both sources of heating are efficient in typical discs.

However, a more realistic model for the disc temperature needs to take into account the opacity provided by the dust and its variation with the wavelengths of the intercepted radiation. Moreover, the disc scale height at hydrostatic equilibrium is linked to the thermal pressure. The

[3]The angle between radiation and the tangent to the disc surface (for not too flared disc, i.e. $H_{\rm g} \lesssim R$) can be well approximated for large radii as

$$\sin\alpha_{\rm irr} \sim \alpha_{\rm irr} \sim \frac{dH_{\rm p}}{dR} - \frac{H_{\rm p}}{R}, \tag{1.56}$$

where $H_{\rm p}$ is the height of the disc's photosphere that can be approximated by $H_{\rm g}$ (Kenyon & Hartmann 1987).

temperature in turn is highly sensitive to the angle of incidence of the starlight. A flared disc intercepts more starlight in its outer regions than a flat one, which will make the disc expand to the hydrostatic equilibrium corresponding to the rising temperature and increase the flaring angle. As we will describe in Chapter 3, small dust grains, which are expected to be well mixed with the gas (see Chapter 2), carry most of the opacity at optical and infrared wavelengths, and therefore control the starlight absorption and the disc's temperature. We therefore need to solve self consistently the shape of the disc by satisfying simultaneously the hydrostatic and thermal equilibrium at every radius taking into account the different conditions of surface layers and disc interiors. Although several analytical models have been introduced in the past (Chiang & Goldreich 1997; Chiang et al. 2001; Garaud & Lin 2007), the full problem requires numerical treatment in order to account for all these effects. A complete solution requires the knowledge of the three-dimensional spatial distribution of dust grains, absorption and scattering opacity of the dust and the source function (ratio of emissivity to opacity of dust), which in turn depends on the intensity coming from any other point in the disc. This complicated treatment is mostly handled numerically, often with Monte Carlo simulations (e.g. Pinte et al. 2006).

Chapter 2

Dust Dynamics in Protoplanetary Discs

Partially based on the paper by ***Giovanni Dipierro*** *and Guillaume Laibe: "An opening criterion for dust gaps in protoplanetary discs" submitted to Monthly Notices of the Royal Astronomical Society.*

Solids in protoplanetary disc are thought to be the primary building blocks for planet formation. In particular, the most widely accepted model for the formation of planets, the *core accretion model*, states that planets form sequentially out of the solid component present in protoplanetary discs. Moreover, as we will describe in Chapter 3, the spatial distribution and optical properties of solids determine the temperature structure and the appearance of the disc over a wide range of wavelengths. Continuum thermal and scattering emission at different wavelengths allow to investigate the three dimensional spatial distribution of solids, providing insights to the dynamics of dust which in turn can be used to infer the spatial distribution of the gas. Dust can potentially be a key observational probe of the gas dynamics since it reacts sensitively to changes in the gas disc. Additionally, the surface of dust grains is the place where surface chemical reactions (such as the formation of complex organics) occur. Importantly, dust grains influence the ionization structure of the disc by sweeping up free electrons and acting as a catalyst for recombination reactions. This influence is crucial to have a comprehensive understanding on how gas couples to magnetic fields in different disc regions and infer where magnetorotational instabilities are likely to occur.

In this Chapter we will review the dynamics of non-growing dust grains and their influence on the gas dynamics. We will therefore relax the approximation of low dust-to-gas ratio assumed in Chapter 1, in order to figure out the dynamical interaction between the two phases. For a detailed description of the dust motion we refer to the classical papers of Whipple (1972); Weidenschilling (1977a); Nakagawa et al. (1986); Dubrulle et al. (1995); Takeuchi & Lin (2002) and Laibe et al. (2012).

2.1 Dust-gas aerodynamical coupling: Single dust grain is a gas fluid

Although protoplanetary discs can be considered a mixture of gas and solids, it is instructive to start the description of dust-gas aerodynamical coupling by considering a system made up of a single particle embedded in a gas fluid. We therefore assume, for the time being, that the dust does not affect significantly the mean gas motion.

Let us consider a compact homogeneous uncharged sphere of fixed radius a, material density $\rho_{\rm p}$ and mass $m_{\rm p} = 4\pi\rho_{\rm p}a^3/3$ moving at a velocity $\mathbf{v}_{\rm rel} = \mathbf{v}_{\rm d} - \mathbf{v}_{\rm g}$ relative to the local velocity

of the gas, where $\mathbf{v}_\mathrm{d}$ and $\mathbf{v}_\mathrm{g}$ are the dust and gas velocities, respectively. Whenever there is a difference in velocity between a solid and the surrounding fluid, the drag force tends to damp this difference on a timescale that depends on the properties of the solid and the surrounding gas. The dust-gas aerodynamical coupling closely depends on the size of the particle relative to the mean free path of the gas molecules, defined as

$$\lambda_\mathrm{mfp} \equiv \frac{\mu m_\mathrm{H}}{\rho_\mathrm{g}\sigma_\mathrm{coll}} \approx 1.15\,\mathrm{cm}\left(\frac{10^{-9}\mathrm{g\,cm^{-3}}}{\rho_\mathrm{g}}\right), \tag{2.1}$$

where $m_\mathrm{H} \approx 10^{-24}$ g is the proton mass, μ is the mean molecular weight (typically taken ~ 2.3 due to the abundances of molecular hydrogen and helium in protoplanetary disc), ρ_g is the gas volume density and $\sigma_\mathrm{coll} \approx 2\times10^{-15}\,\mathrm{cm}^2$ is the molecular collisional cross section (Chapman & Cowling 1970). Another useful quantity to characterize the drag force is the mean thermal velocity, which is the expected mean gas velocity at equilibrium, given by

$$v_\mathrm{th} \equiv \sqrt{\frac{8k_BT}{\pi\mu m_\mathrm{H}}} = \sqrt{\frac{8}{\pi\gamma}}c_\mathrm{s}, \tag{2.2}$$

where T is the gas temperature, k_B is the Boltzmann's constant, γ is the adiabatic index and c_s is the sound speed. The thermal velocity represent the typical impact velocity of molecules onto dust grains. When the grains size is smaller than the mean free path, i.e. $a < 9\lambda_\mathrm{mfp}/4$, the drag force can be computed by considering the transfer of angular momentum between collisions of individual gas molecules at the grains surface (see the top drawing of Fig. 2.1). In this case, grains are small enough not to disturb the Maxwellian distribution of the gas velocity. This regime is usually referred to as Epstein regime (Epstein 1924). In this regime, the gas is considered as a collection of particles colliding with the dust grain. The drag force on a single grain size is given by (Weidenschilling 1977a)

$$\mathbf{F}_\mathrm{D} = -\frac{4}{3}f\pi a^2\rho_\mathrm{g}v_\mathrm{th}\mathbf{v}_\mathrm{rel}. \tag{2.3}$$

This expression can be physically justified since the drag force is given by the volume swept per unit time ($4\pi a^2\mathbf{v}_\mathrm{rel}$) times the momentum per unit volume ($\rho_\mathrm{g}v_\mathrm{th}$) and a geometric factor that depends on the details of the collision. The term f is a correction factor to take into account the case where the relative motion between the gas and dust in protoplanetary discs is supersonic. This term is given by (Kwok 1975, see also Sect. 2.2.1 of Laibe & Price 2012b)

$$f \equiv \sqrt{1 + \frac{9\pi}{128}\left(\frac{|\mathbf{v}_\mathrm{rel}|}{c_\mathrm{s}}\right)^2}. \tag{2.4}$$

Since the relative motion between the gas and dust in protoplanetary discs is typically subsonic, this term is neglected in protoplanetary discs ($f \sim 1$). If the particle's size is larger than the mean free path[1], i.e. $a > 9\lambda_\mathrm{mfp}/4$, we need to consider the flow of gas around the particles, which is clearly dependent on the laminar or turbulent regime of the flow (see the bottom drawing of Fig. 2.1). Therefore, the drag force depends on the Reynolds number for the particle in the fluid, which is a measure of the laminar or turbulent nature of the surrounding flow. The Reynolds number for an object in a fluid is defined as

$$\mathrm{Re}_\mathrm{p} \equiv \frac{vL}{\nu}, \tag{2.5}$$

where v is the characteristic relative velocity between the object and the fluid, L is a characteristic linear dimension and ν is the kinematical viscosity. For a sphere in a fluid, the characteristic

[1]It is worth remarking that no current analytic theory describes exactly the drag force in the intermediate regime, i.e. $a \sim \lambda_\mathrm{mfp}$. Generally, an asymptotic continuous interpolation between the two limiting Epstein and the Stokes regimes is used. The transition $a \sim 9\lambda_\mathrm{mfp}/4$ has been proposed by Stepinski & Valageas (1996) in order to obtain a smooth transition between the two regimes.

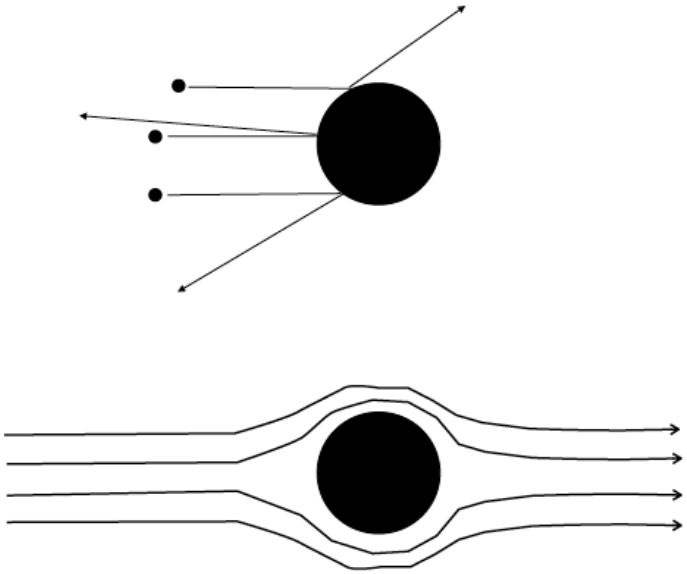

Figure 2.1: Gas-dust aerodynamical interaction in the Epstein regime (top) and in the Stokes regime (bottom).

length-scale is the diameter of the sphere and the characteristic velocity is the relative velocity $|\mathbf{v}_{\rm rel}|$. It is therefore given by

$$\mathrm{Re_p} = \frac{2a\,|\mathbf{v}_{\rm rel}|}{\nu}. \tag{2.6}$$

In this regime (the so-called Stokes regime), the drag force can be written in the form

$$\mathbf{F}_{\rm D} = -\frac{\mathrm{C_D}}{2}\pi a^2 \rho_{\rm g}\,|\mathbf{v}_{\rm rel}|\,\mathbf{v}_{\rm rel}, \tag{2.7}$$

where $\mathrm{C_D}$ is the drag coefficient. The latter can assume three different values, corresponding to the three different Stokes regime, depending on the value of the Reynolds number. While at small Reynolds numbers analytic expressions for the drag force can be derived, at higher Reynolds number, the drag law is mostly inferred from experiments. The coefficient is given by

$$\mathrm{C_D} = \begin{cases} 24\,\mathrm{Re_p}^{-1}, & \text{for } \mathrm{Re_p} < 1, & \text{Stokes regime I,} \\ 24\,\mathrm{Re_p}^{-0.6}, & \text{for } 1 < \mathrm{Re_p} < 800, & \text{Stokes regime II,} \\ 0.44, & \text{for } \mathrm{Re_p} > 800, & \text{Stokes regime III.} \end{cases} \tag{2.8}$$

The drag force is usually expressed as

$$\mathbf{F}_{\rm D} = -K\mathbf{v}_{\rm rel}. \tag{2.9}$$

Importantly, while for the Epstein and Stokes regime I, the drag force remains linear in $\mathbf{v}_{\rm rel}$, i.e. K does not depend on $\mathbf{v}_{\rm rel}$, in other regimes the drag force is non-linear in $\mathbf{v}_{\rm rel}$. Using typical physical quantities to estimate the mean free path in protoplanetary discs, it can be shown that drag forces for particles smaller than $\sim$10 m are well described according to the Epstein regime (Garaud et al. 2004; Laibe et al. 2012). Thus, the relevant drag law from the gas to the solids is the Epstein drag law, though there might be regions close to the star where the gas is sufficiently dense to interact with the dust according to the Stokes regime I drag law. In these cases, the drag force remains linear in $|\mathbf{v}_{\rm rel}|$ (see Eqs. 2.7 and 2.8).

2.1.1 The stopping time for a single grain size

An important concept for the dynamics of dust grains is the stopping time, defined as the typical timescale in which drag modifies the relative velocity between a single grain and the gas significantly. A particle, due to its inertia, does not instantaneously follow the gas motion but requires a certain time. The typical time is given by

$$t_{\rm s} = \frac{m_{\rm p}\,|\mathbf{v}_{\rm rel}|}{|\mathbf{F}_{\rm D}|} = \frac{m_{\rm p}}{K}. \tag{2.10}$$

This represents the time to remove the initial relative momentum of the particle assuming that the force remains constant. Actually, in the linear drag regime (Epstein and Stokes regime I) where $\mathbf{F}_{\rm D} \propto \mathbf{v}_{\rm rel}$, the force decreases during its action. Thus the stopping time has to be considered as the time interval in which the exponentially decreasing relative velocities decreases by a factor of e. In all the regimes previously mentioned the stopping time is given by (Laibe et al. 2012)

$$t_{\rm s} = \begin{cases} \frac{\rho_{\rm p} a}{\rho_{\rm g} c_{\rm s}} \sqrt{\frac{\pi\gamma}{8}}, & \text{Epstein regime,} \\ \frac{2\rho_{\rm p} a^2}{9\nu\rho_{\rm g}}, & \text{Stokes regime I,} \\ \frac{2^{0.6}\rho_{\rm p} a^{1.6}}{9\nu^{0.6}\rho_{\rm g}^{1.4}|\mathbf{v}_{\rm rel}|^{0.4}}, & \text{Stokes regime II,} \\ \frac{6\rho_{\rm p} a}{\rho_{\rm g}|\mathbf{v}_{\rm rel}|}, & \text{Stokes regime III.} \end{cases} \tag{2.11}$$

It can be noticed that in all cases the stopping time increases with the particle size, $t_{\rm s} \propto a^r$ with $r \geq 1$ with a maximum value of $r = 2$ in the Stokes regime I. Thus, larger grains becomes uncoupled with the gas and tend to move with their own velocity. This dependence highlights the complex interplay between the grain growth processing and the dust-gas aerodynamical coupling (Laibe 2014).

2.2 Application to protoplanetary disc: Dynamics in a dust-gas mixture

The motion of dust particles in protoplanetary discs is affected by the aerodynamical interaction with the gas and *vice-versa*. Due to the high value of the gas-to-dust ratio at the initial stage of the disc formation process, the dust-gas aerodynamical coupling can be explained to a first approximation using the framework introduced in Sect. 2.1. However, the previous description does not take into account how the drag influences the gas motion. In detail, the relative velocity between the two phases is damped under the action of drag torque and in turn is spread over the two phases proportionally to the respective density of each phase. Moreover, gas in astrophysical environments is often in a turbulent state of motion, constantly affected by temporally and spatially varying accelerations from eddies characterised by a large variety of scales. We therefore expect that the turbulent gas motion affects the dust dynamics by producing a "drag-induced dust diffusivity" (see Sect. 2.4).

In a protoplanetary disc there are many processes able to create velocity differences between the gas and solids. The different intrinsic nature of the two phases (see next Section) induces remarkable effects on relative velocity field between the two components, such as the different orbital motion around the central star.

2.2.1 Dust as a pressureless and viscousless fluid

The description of the dynamics of solid particles embedded in a gas fluid can be approached in two ways. The most direct way is to assume the dusty disc as an N-body system whose dynamics is governed by the gravity from the star and the aerodynamical interaction with the gas. This method is reliable and precise, since it allows to consider the particle-based nature of the dust component. However, it can only be applied to a finite number of particles, implying that macroscopic quantities reconstructed from the N-body simulations are sometimes inaccurate

due to a low sampling of the phase space. Moreover, the physical and computational complexity to describe the N-body evolution of solid particles in a dust-gas mixture would be prohibitive. The second method, which is the most widely used approach to investigate the dynamics of solids, is to consider the dust as a continuum. This approach allows to treat the dust as an infinite ensemble of particles and compute their macroscopic properties (density, velocity field, etc.) by averaging physical quantities over a control volume large enough to be statistically meaningful but sufficiently small to ignore variations of the fluid quantities across it. These two conditions (see Clarke & Carswell 2007) allow to define the concept of "fluid element" in the context of a continuum description. This method heavily simplifies the analysis, making the "two-fluid" approach (gas and dust) the preferred method of treating the dynamics and growth of solid particles in protoplanetary discs, especially in the star and planet formation community.

A second assumption that is constantly made is to consider the dust as a pressureless fluid. This approximation is justified by the statement that the frequency of collisions between dust particles is very low compared to collisions between particles and gas in protoplanetary discs. Alternatively, this assumption can be justified by comparing the collision timescale and the stopping time[2]. The collision timescale between two equally-sized dust grains is given by

$$t_{\rm coll} = \frac{2\rho_{\rm p} a}{3\rho_{\rm d} v_{\rm coll}}, \tag{2.12}$$

where $v_{\rm coll}$ is the relative velocity between two colliding particles and $\rho_{\rm d}$ is the dust volume density. Comparing the collisional timescale with the stopping time in the Epstein regime we obtain

$$\frac{t_{\rm s}}{t_{\rm coll}} = \frac{3}{2}\frac{\rho_{\rm d} v_{\rm coll}}{\rho_{\rm g} c_{\rm s}}\sqrt{\frac{\pi\gamma}{8}} \simeq \frac{\rho_{\rm d}}{\rho_{\rm g}}\frac{v_{\rm coll}}{c_{\rm s}}. \tag{2.13}$$

This ratio depends on the dust-to-gas ratio, which in typical disc is $\lesssim 1$, and the relative velocities between colliding dust grains. There are a large variety of sources of relative velocities between dust particles such as the turbulent and systematic motion induced by drag (see the recent review of Birnstiel et al. 2016, Sect. 3.1). In typical protoplanetary disc, the relative velocities between equal sized particles induced by turbulence can be approximated by (Cuzzi & Hogan 2003; Ormel & Cuzzi 2007; Birnstiel et al. 2016)

$$v_{\rm coll} \approx \sqrt{\frac{3\alpha}{{\rm t_s}\Omega + ({\rm t_s}\Omega)^{-1}}}\, c_{\rm s}, \tag{2.14}$$

where Ω is the gas angular velocity, α is the Shakura & Sunyaev (1973) viscosity parameter defined in Eq. 1.47. It can clearly be noticed that, for each value of the stopping time, the relative velocities are always subsonic. The relative velocity reaches a maximum for ${\rm t_s}\Omega = 1$, given by $c_{\rm s}\sqrt{3\alpha/2} \ll c_{\rm s}$. We can therefore neglect inter particle collision between dust grains in the Epstein regime. For larger particles in the Stokes regime, collisions become relatively importance due to the higher stopping time. However, for low dust-to-gas ratio, collisions are a small correction to the dust motion induced by drag (see also Sect. 6.1 of Youdin & Lithwick 2007).

In collisionless system, pressure forces can also be created dynamically when external forces, such as gravity, excite velocity fluctuations[3]. However, the dynamical pressure induced by velocity perturbations can be neglected for smaller particles, i.e. ${\rm St} \lesssim 1$, in gas-dominated

[2]It is worth remarking that the stopping time in a dust-gas mixture is slightly different from the stopping time for a single grain size. However, as described in Sect. 2.3.1, for low dust-to-gas ratio we can simply assume that the stopping time is given by Eq. 2.11.

[3]The so-called dynamical pressure can be evaluated starting from the collisionless Boltzmann equation properly modified to include the drag. This equation describes the evolution of the distribution function f in phase space and is given by (Carballido et al. 2006; Youdin & Lithwick 2007; Hersant 2009)

$$\frac{\partial f}{\partial t} + \mathbf{v}\cdot\nabla_x f + \nabla_v\left(\frac{\mathbf{F}}{m} f\right) = 0, \tag{2.15}$$

astrophysical discs (Hersant 2009). This is due to the fact that, since smaller particles are strongly coupled with the gas, any purely dust velocity fluctuations are quickly damped by the drag on a timescale shorter than dynamical time. Therefore, the pressure driven by inter-particle collisions or dynamically-induced can be safely neglected in protoplanetary disc for smaller grain sizes. By contrast, in purely dusty disc where particles have a vey high inertia ($\mathrm{St} \gg 1$) such as grains in Saturn rings, the pressure produced by collisions and velocity dispersion induced by external drivers need to be considered to properly describe the motion of dust grains (Cuzzi et al. 1981).

Additionally, since the interaction between dust particles are neglected, dust in protoplanetary discs can be considered an inviscid fluid, since there is no way to transfer the momentum associated with velocity differences between different elements.

All these assumptions allows to analytically (and, in most cases, computationally) treat the dust in protoplanetary discs as a free-streaming collisionless fluid made of homogeneous, uncharged and gravitationally non-interacting dust grains whose motion is controlled solely by gravitational forces from the central star and the aerodynamical interaction with the gas. It is also assumed that the volume occupied by the dust grains is negligible. Since in young protoplenatary discs, whose typically extension is $\sim 10^2 - 10^3$ au (Eisner et al. 2008; Andrews et al. 2011), dust consists of grains with micron to kilometre in size, this is an extremely good approximation, but it can be important in non-astrophysical problems (Fan & Zhu 1988; Laibe & Price 2012a).

Finally, in order to simplify the analytical investigation we will focus only on the Epstein and Stokes I regime, which are the two relevant drag regimes in typical protoplanetary discs (see Sect. 2.1). However, in some cases such as gravitationally unstable discs with a large dust grain population, all the cases needs to be taken into account since high relative velocity between the two phases may occur (Rice et al. 2004, 2006b).

2.2.2 Timescale for the dust evolution

In laminar flows, the dynamics of solids depends on both the magnitude of the drag and on its relative contribution with respect to the gravity of the central star. However, protoplanetary discs are expected to be highly turbulent (see Sect. 1.3.1). The stirring of particles by turbulent gas is a fundamental problem in fluid dynamics, since it strongly affects the ability of particles to grow and form planetesimals (Voelk et al. 1980; Ormel & Cuzzi 2007). The dimensionless parameters to describe the coupling of solids with the laminar and the turbulent gas flow are given by

$$\tau_\mathrm{s} \equiv t_\mathrm{s}\Omega_\mathrm{k}, \tag{2.16}$$

$$\tau_\mathrm{e} \equiv t_\mathrm{e}\Omega_\mathrm{k}, \tag{2.17}$$

$$\mathrm{St} \equiv \frac{t_\mathrm{s}}{t_\mathrm{e}} = \frac{\tau_\mathrm{s}}{\tau_\mathrm{e}}. \tag{2.18}$$

The dimensionless quantity τ_s measures the coupling (t_s is the stopping time defined in Eq. 2.10 for a single grain and will be then generalized in Eq. 2.36 for dust in a mixture) of the particle with the unperturbed orbital gas flow. The effect of orbital shear on turbulent eddies is described by the parameter τ_e, where the eddy time, t_e, is the correlation time of turbulent fluctuations. It is worth pointing out that, for Kolmogorov-like turbulence, t_e is the turnover time of the largest scale eddies. When this correlation time is much smaller than the orbital period, i.e. $\tau_\mathrm{e} \ll 1$, and particles are decoupled from the unperturbed gas flow, i.e. $\tau_\mathrm{s} \gg 1$, dust is stocastically forced by small kicks in all directions, leading to random impulsive changes to the velocity of the particle (Youdin & Lithwick 2007). The Stokes number St measures the coupling of the

where x is the space coordinate, v is the velocity coordinate and $\mathbf{F}/m$ is the acceleration due to the gravity and drag. Once the equation is solved for f, it is possible to evaluate macroscopic quantities such as density and velocity field at equilibrium by computing the moments of the distribution function f (Landau & Lifshitz 1980).

particle to turbulent fluctuations. This number is a dimensionless quantity that describes the aerodynamic properties of a dust particle, i.e. particles with different sizes, densities, or shapes but the same Stokes number are aerodynamically identical.

Importantly, a variety of numerical investigations found that the correlation time of turbulent fluctuations driven by fluid instabilities can be approximated by the dynamical timescale $\Omega_{\rm k}^{-1}$ (e.g. Johansen & Klahr 2005; Carballido et al. 2006; Zhu et al. 2015b). Therefore, $\tau_{\rm e}$ is usually assumed equal to unity and the Stokes number can therefore be given by

$$\mathrm{St} = \tau_{\rm s} = t_{\rm s}\Omega_{\rm k}. \tag{2.19}$$

This allows to relate directly the stopping time to the unperturbed orbital timescale. However, as described in Youdin & Lithwick (2007), deviations from the favored case $\tau_{\rm e} = 1$ might be important for the dust motion.

2.3 Dust dynamics in laminar viscous flows

Let us start our analysis on dust dynamics by assuming that the gas flow is laminar. The system consists of a dust-gas mixture in a thin, axisymmetric, non-magnetic, non-self-graviting, viscous and vertically isothermal disc. We treat the dust phase as a pressureless and viscousless fluid of equally-sized solids[4]. According to the previously mentioned assumptions to analytically treat the dust-gas dynamics in protoplanetary disc, the two-fluid approach is implemented by introducing two Euler equations, coupled through the drag force. The equations of motion for the gas and the dust are (e.g. Nakagawa et al. 1986; Laibe & Price 2012a)

$$\frac{\partial \rho_{\rm g}}{\partial t} + \nabla \cdot (\rho_{\rm g}\mathbf{v}_{\rm g}) = 0, \tag{2.20}$$

$$\frac{\partial \rho_{\rm d}}{\partial t} + \nabla \cdot (\rho_{\rm d}\mathbf{v}_{\rm d}) = 0, \tag{2.21}$$

$$\frac{\partial \mathbf{v}_{\rm g}}{\partial t} + (\mathbf{v}_{\rm g}\cdot\nabla)\,\mathbf{v}_{\rm g} = -\frac{1}{\rho_{\rm g}}\mathbf{F}_{\mathbf{D}}^{\mathbf{V}} - \frac{1}{\rho_{\rm g}}(\nabla P - \nabla\cdot\sigma) - \nabla\Phi, \tag{2.22}$$

$$\frac{\partial \mathbf{v}_{\rm d}}{\partial t} + (\mathbf{v}_{\rm d}\cdot\nabla)\,\mathbf{v}_{\rm d} = \frac{1}{\rho_{\rm d}}\mathbf{F}_{\mathbf{D}}^{\mathbf{V}} - \nabla\Phi, \tag{2.23}$$

where the indices g and d refer to the gas and the dust phases, Φ denotes the gravitational potentials of the star and P and σ denote the pressure and the viscous tensor of the gas. Note that Eq. 2.22 is the Navier-Stokes equation already discussed in Sect. 1.2.1 (see Eq. 1.4) with an additional force given by the dust back-reaction on the gas flow. Note that the drag force $\mathbf{F}_{\mathbf{D}}^{\mathbf{V}}$ is a force per unit volume defined as

$$\mathbf{F}_{\mathbf{D}}^{\mathbf{V}} = -K^V(\mathbf{v}_{\rm d} - \mathbf{v}_{\rm g}), \tag{2.24}$$

where the new drag coefficient K^V has dimensions of mass per unit volume per unit time and it is related to the drag coefficient on a single grain (Eq. 2.9) by

$$K^V = \frac{\rho_{\rm d}}{m_{\rm p}}K. \tag{2.25}$$

Importantly, we will show that, in typical protoplanetary discs, there is a marked separation between radial and vertical motion of solids, since the typical radial drift timescale is much longer than the vertical settling timescale (Garaud et al. 2004; Laibe et al. 2014). The analysis on the motion of dust and gas in a mixture will be therefore separated into two parts: the radial motion on the disc plane and the vertical settling.

[4]The general description for multi-species dust particles can be found in Bai & Stone (2010).

2.3.1 Radial drift in laminar gas flow

To reduce Eqs. 2.20 – 2.23 to a system of equations that describes a steady steady solution for the gas and the dust, we follow the approach introduced in Nakagawa et al. (1986) (hereafter NSH86). As already assumed in Chapter 1 for the gas dynamics, we consider classical shear viscosity, i.e. the only non-zero component of σ for the circular sheared flow of the disc is (rewriting Eq. 1.6)

$$\sigma_{R\phi} = \eta R \frac{\partial}{\partial R}\left(\frac{v_{\mathrm{g},\phi}}{R}\right), \tag{2.26}$$

where $\eta = \nu\rho_{\mathrm{g}}$ denotes the shear viscosity coefficient of the gas (see Appendix A) and

$$\frac{v_{\mathrm{g},\phi}^2}{R} = \frac{v_{\mathrm{k}}^2}{R} + \frac{1}{\rho_{\mathrm{g}}}\frac{\partial P}{\partial R} + \mathcal{O}\left(H_{\mathrm{g}}/R\right)^2, \tag{2.27}$$

is the unperturbed gas orbital velocity (see also Eq. 1.15), where H_{g} is the vertical scale height of the gas at hydrostatic equilibrium (see Sect. 1.2.1.1). However, the deviation of $\sigma_{R\phi}$ induced by the sub-Keplerian rotation (the second term on the right hand side of Eq. 2.27) is only of order $(H_{\mathrm{g}}/R)^2 \ll 1$ and can be neglected. We therefore assume that $\sigma_{R\phi}$ can be approximated by putting $v_{\mathrm{g},\phi} = v_{\mathrm{k}}$.

Making the usual assumptions to describe the disc dynamics (see Sect. 1.2), Eqs. 2.20 – 2.23 can be expressed in polar coordinates as

$$\begin{aligned}
\frac{\partial v_{\mathrm{g},R}}{\partial t} - \frac{{v_{\mathrm{g},\phi}}^2}{R} &= \frac{K^V}{\rho_{\mathrm{g}}}(v_{\mathrm{d},R} - v_{\mathrm{g},R}) - \frac{1}{\rho_{\mathrm{g}}}\frac{\partial P}{\partial R} - \Omega_{\mathrm{k}}^2 R, && (2.28)\\
\frac{\partial v_{\mathrm{g},\phi}}{\partial t} + v_{\mathrm{g},R}\left(\frac{v_{\mathrm{g},\phi}}{R} + \frac{\partial v_{\mathrm{g},\phi}}{\partial R}\right) &= \frac{K^V}{\rho_{\mathrm{g}}}(v_{\mathrm{d},\phi} - v_{\mathrm{g},\phi}) + \frac{1}{\rho_g}\nabla\cdot\sigma|_\theta, && (2.29)\\
\frac{\partial v_{\mathrm{d},R}}{\partial t} - \frac{{v_{\mathrm{d},\phi}}^2}{R} &= -\frac{K^V}{\rho_{\mathrm{d}}}(v_{\mathrm{d},R} - v_{\mathrm{g},R}) - \Omega_{\mathrm{k}}^2 R, && (2.30)\\
\frac{\partial v_{\mathrm{d},\phi}}{\partial t} + v_{\mathrm{d},R}\left(\frac{v_{\mathrm{d},\phi}}{R} + \frac{\partial v_{\mathrm{d},\phi}}{\partial R}\right) &= -\frac{K^V}{\rho_{\mathrm{d}}}(v_{\mathrm{d},\phi} - v_{\mathrm{g},\phi}). && (2.31)
\end{aligned}$$

Let us perform a perturbative expansion relative to the Keplerian velocity $\mathbf{v}_{\mathrm{k}} = (0, R\Omega_{\mathrm{k}}, 0)$ in both phases, and look for stationary solutions (i.e. $\partial/\partial t = 0$). Note that from here, the notation v_ϕ will be used to refer to the *perturbed* azimuthal velocities for sake of simplicity. The equations of motion for the perturbed velocities are

$$\begin{aligned}
\frac{\partial v_{\mathrm{g},R}}{\partial t} &= \frac{K^V}{\rho_{\mathrm{g}}}(v_{\mathrm{d},R} - v_{\mathrm{g},R}) - \frac{1}{\rho_{\mathrm{g}}}\frac{\partial P}{\partial r} + 2\Omega_{\mathrm{k}} v_{\mathrm{g},\phi} = 0, && (2.32)\\
\frac{\partial v_{\mathrm{g},\phi}}{\partial t} &= \frac{K^V}{\rho_{\mathrm{g}}}(v_{\mathrm{d},\phi} - v_{\mathrm{g},\phi}) - \frac{\Omega_{\mathrm{k}}}{2} v_{\mathrm{g},R} + \frac{1}{\rho_{\mathrm{g}}}\nabla\cdot\sigma|_\phi = 0, && (2.33)\\
\frac{\partial v_{\mathrm{d},R}}{\partial t} &= -\frac{K^V}{\rho_{\mathrm{d}}}(v_{\mathrm{d},R} - v_{\mathrm{g},R}) + 2\Omega_{\mathrm{k}} v_{\mathrm{d},\phi} = 0, && (2.34)\\
\frac{\partial v_{\mathrm{d},\phi}}{\partial t} &= -\frac{K^V}{\rho_{\mathrm{d}}}(v_{\mathrm{d},\phi} - v_{\mathrm{g},\phi}) - \frac{\Omega_{\mathrm{k}}}{2} v_{\mathrm{d},R} = 0. && (2.35)
\end{aligned}$$

This system can be visualized as a non-homogeneous differential equations of the form $X'(t) = -AX(t) + B$, where A and B are two constant matrices. The stationary solution correspond to $X_{\mathrm{stat}} = A^{-1}B$. The typical timescale it takes to reach the stationary solution is usually referred to as the stopping time in a gas-dust mixture and is given by

$$t_{\mathrm{s,m}} = \frac{\rho_{\mathrm{d}}\rho_{\mathrm{g}}}{K^V(\rho_{\mathrm{d}} + \rho_{\mathrm{g}})} = \frac{\rho_{\mathrm{d}}}{K^V(1+\epsilon)} = \frac{m_{\mathrm{p}}}{K(1+\epsilon)} = \frac{t_{\mathrm{s}}}{1+\epsilon}, \tag{2.36}$$

where we have used Eqs. 2.10 and 2.25 and $\epsilon = \rho_{\mathrm{d}}/\rho_{\mathrm{g}}$ is the dust-to-gas density ratio. This expression for the stopping time extends the previous formulation based on the interaction

between a single grain and the gas flow (Eq. 2.10). The latter can be obtained by assuming that the dust component in the mass balance of the mixture in negligible or, in other words, if the dust does not affect the motion of the gas. In this case, $\epsilon \ll 1$ and the stopping time is $t_{\rm s,m} = t_{\rm s}$. Thus, the stopping time $t_{\rm s}$ is correct in the limit where the density of one phase is negligible compared to the density of the other phase, but becomes erroneous in the case where the two fluids have densities of the same order of magnitude, as expected in the midplane of the inner disc regions (see Sect. 2.5.1).

In order to evaluate how the dust is coupled with the gas flow, it is convenient to use the Stokes number ${\rm St} \equiv \Omega_{\rm k} t_{\rm s,m}$, introduced in Sect. 2.2.2. In the Epstein regime, the Stokes number is given by

$$ {\rm St} = t_{\rm s,m}\Omega_{\rm k} = \frac{\rho_{\rm p} a}{(1+\epsilon)\rho_{\rm g}} \frac{\Omega_{\rm k}}{c_{\rm s}} \sqrt{\frac{\pi\gamma}{8}}. \tag{2.37} $$

The stationary solutions of the the linear system Eqs. 2.32 – 2.35 are then given by

$$ v_{{\rm g},R} = -\frac{1}{1+\epsilon}\left[\frac{\epsilon\,\Delta v}{{\rm St}+{\rm St}^{-1}} - \left(1+\epsilon\frac{{\rm St}^2}{1+{\rm St}^2}\right) v_{\rm visc}\right], \tag{2.38} $$

$$ v_{{\rm d},R} = \frac{1}{1+\epsilon}\left[\frac{\Delta v}{{\rm St}+{\rm St}^{-1}} + \frac{v_{\rm visc}}{1+{\rm St}^2}\right], \tag{2.39} $$

$$ v_{{\rm g},\phi} = \frac{1}{2(1+\epsilon)}\left[\left(1+\epsilon\frac{{\rm St}^2}{1+{\rm St}^2}\right)\Delta v + \frac{\epsilon}{{\rm St}+{\rm St}^{-1}} v_{\rm visc}\right], \tag{2.40} $$

$$ v_{{\rm d},\phi} = \frac{1}{2(1+\epsilon)}\left[\frac{\Delta v}{1+{\rm St}^2} - \frac{v_{\rm visc}}{{\rm St}+{\rm St}^{-1}}\right], \tag{2.41} $$

where

$$ \Delta v \equiv \frac{1}{\rho_{\rm g}\Omega_{\rm k}}\frac{\partial P}{\partial R}, \tag{2.42} $$

is the typical optimal drift velocity, and

$$ v_{\rm visc} \equiv \frac{2}{\Omega_{\rm k}\rho_{\rm g}}\nabla\cdot\sigma|_{\phi} = \frac{1}{R\rho_{\rm g}\frac{\partial}{\partial R}\left(R v_{\rm k}\right)}\frac{\partial}{\partial R}\left(\eta R^3\frac{\partial\Omega_{\rm k}}{\partial R}\right) = -\frac{3}{\rho_{\rm g}R^{1/2}}\frac{\partial}{\partial R}\left(\nu\rho_{\rm g}R^{1/2}\right), \tag{2.43} $$

is the radial gas velocity driven by viscous accretion (see also Eq. 1.25). The steady-state motion of the dust and gas in a two-fluid mixture provides deep insight on how the viscous and orbital motion of the gas influences the dust radial and azimuthal motion and vice-versa.

2.3.1.1 Physical interpretation

As observed in Eqs. 2.38-2.41, the dust motion induced by the drag can be ascribed to the different motions between the two phase both in the radial and azimuthal direction. These different orbital and radial velocities are linked to the different nature of the two fluids in the mixture. The gas feels pressure and viscous forces, leading to a secular viscous radial evolution and, for standard disc geometries ($\partial P/\partial R < 0$, $\Delta v < 0$), a sub-Keplerian orbital motion. By contrast, a freely orbiting dust particle feels only centrifugal forces and gravity, and should therefore be in a Keplerian orbit without any radial motion.

Hence, the dust particle feels a continuous headwind from the gas in the azimuthal direction and a crosswind toward inner radii due to the radial inward motion of the gas (if the gas is moving inward, i.e. $v_{\rm visc} < 0$). The headwind causes the dust particle to lose its angular momentum and spiral towards the star. This motion is mostly referred to as *radial drift*. The crosswind forces the dust particles to move toward inner radii due to the induced direct pushing. This component is mostly referred to as *drag*. The latter component (the second term in the right hand side of Eq. 2.39) shows that drag makes very small grains stick to the gas, so that they are

carried radially by the viscous flow. Hence, dust evolves "viscously" under the indirect effect of the viscous evolution of the gas, with a so-called "drag-induced dust viscosity" given by

$$\nu_{\rm d,eff} \equiv \frac{1}{1+\epsilon}\frac{1}{1+{\rm St}^2}\nu. \tag{2.44}$$

Note that although Eq. 2.44 resembles the same scaling with St as the radial dust diffusivity derived in Youdin & Lithwick (2007), the drag-induced dust viscosity does not describe the motion induced by turbulence (the turbulence-induced dust motion will be described in Sect. 2.4) over dust grains, but models how the viscous evolution of the gas affects the radial dynamics of grains. Moreover, Eq. 2.44 shows, as expected, that for tightly coupled particles and low dust-to-gas ratio, the dust viscously evolves as the gas, since in this case the dust and gas share the same motion. Both the drag and drift components, if the gas is moving inward (i.e. $v_{\rm visc} < 0$), lead the particle to spiral inward at a rate that depends on the Stokes number and dust-to-gas ratio. On the other hand, the gas feel a tailwind and a crosswind toward outer radii, leading to a gain of angular momentum and a drift toward the disc outer region with a timescale that depends on Stokes number and dust-to-gas ratio as well.

Therefore, the two principal source of differential motion are due to the pressure and viscous forces. Comparing the orders of magnitude of the different velocity terms, we obtain

$$|\Delta v| \sim \left(\frac{H_{\rm g}}{R}\right)^2 v_{\rm k}, \tag{2.45}$$

$$|v_{\rm visc}| \sim \alpha\left(\frac{H_{\rm g}}{R}\right)^2 v_{\rm k}, \tag{2.46}$$

where we have used the α-prescription (see Eq. 1.47) for the turbulent viscosity $\nu = \alpha c_{\rm s} H_{\rm g}$ (Shakura & Sunyaev 1973). Thus, the relative contribution of the radial drift to the drag component scales with $|\Delta v| \,/\, |v_{\rm visc}| \sim \alpha^{-1} \gg 1$. Therefore, to a first approximation, since in typical protoplanetary discs $\alpha^{-1} \sim \left[10^2, 10^4\right]$, the radial drift component is expected to dominate the dust motion for a large range of Stokes number.

Before moving on describing Eqs. 2.38 – 2.41 in a general view, it is instructive to evaluate the dust and gas motion in a mixture dominated by the gas component, i.e. $\epsilon \ll 1$. In this limit, the dust and gas velocities are

$$v_{{\rm g},R} = v_{\rm visc}, \tag{2.47}$$

$$v_{{\rm d},R} = \frac{\Delta v}{{\rm St}+{\rm St}^{-1}} + \frac{v_{\rm visc}}{1+{\rm St}^2}, \tag{2.48}$$

$$v_{{\rm g},\phi} = \frac{\Delta v}{2}, \tag{2.49}$$

$$v_{{\rm d},\phi} = \frac{1}{2}\left[\frac{\Delta v}{1+{\rm St}^2} - \frac{v_{\rm visc}}{{\rm St}+{\rm St}^{-1}}\right]. \tag{2.50}$$

Importantly, the gas velocities resembles the same results shown in Sect. 1.2, as expected. While the radial gas velocity is governed by viscous forces, the gas azimuthal velocity is related to the pressure supported orbital motion described in Sect. 1.2.1.2. In detail, the value of $\Delta v/2$ can be derived by expanding Eq. 1.15 (or Eq. 2.27) to the first order in $H_{\rm g}/R$. The dust motion will be described below in the general view.

In Fig. 2.2 are shown the radial and azimuthal dust and gas velocity expressed in Eqs. 2.38 – 2.41 as a function of the Stokes number for different values of the dust-to-gas ratio. It can be clearly noticed that in the range of Stokes number close to unity, the dust and gas velocity are characterised by a large variation. Let us discuss the dust and gas motion in the three main regimes of dust-particle aerodynamical interaction: well coupled grains (${\rm St} \ll 1$), marginally coupled grains (${\rm St} = 1$) and poorly coupled grains (${\rm St} \gg 1$).

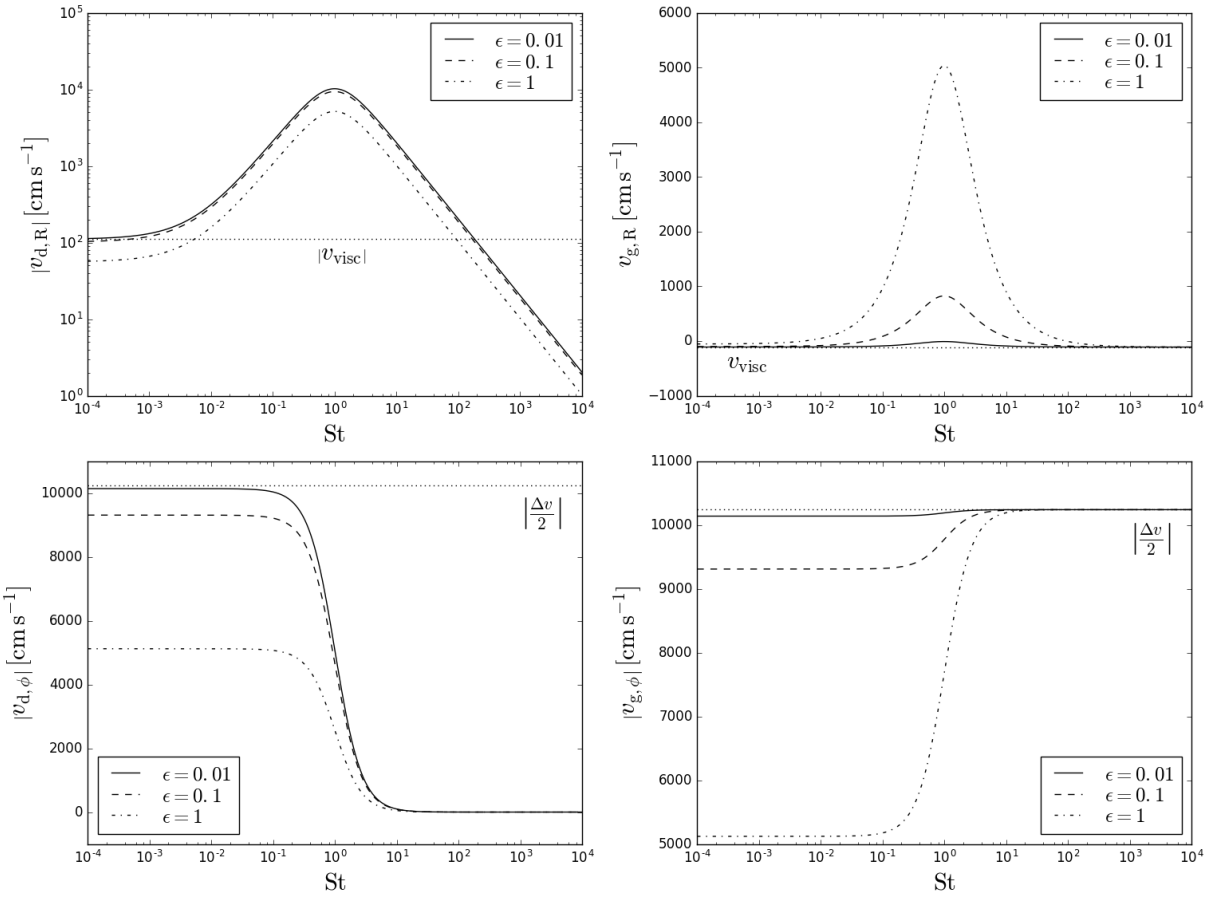

Figure 2.2: The radial (upper panels) and azimuthal (bottom panels) dust and gas velocity as a function of Stokes number St for three values of the dust-to-gas ratio: 0.01 (solid), 0.1 (dashed) and 1 (dotted-dashed) as expressed in Eqs. 2.38 – 2.41. The dotted lines indicates the viscous radial gas velocity $v_{\rm visc}$ (upper panels) and half of the optimal drift velocity Δv (bottom panels). The adopted disc model assumes a disc with mass $0.01 M_\odot$ between 1 and 160 au orbiting around a central star with mass $1\,M_\odot$, power law radial profiles of the surface density and sound speed as $\Sigma_{\rm g} \propto R^{-1}$, $c_{\rm s} \propto R^{-1/4}$, a disc aspect ratio $H_{\rm g}/R = 0.05$ at 1 au and $\alpha = 10^{-2}$. The plots are computed on the disc midplane at $R = 10$ au where $H_{\rm g}/R = 0.09$.

- **Well coupled particles:** $\mathrm{St} \ll 1$

  In this regime, the stopping time is much smaller than the Keplerian timescale. The dust and gas velocities are approximated by

$$v_{{\rm g},R} \sim -\frac{1}{1+\epsilon}\left(\mathrm{St}\,\epsilon\,\Delta v - v_{\rm visc}\right), \tag{2.51}$$

$$v_{{\rm d},R} \sim \frac{1}{1+\epsilon}\left(\mathrm{St}\,\Delta v + v_{\rm visc}\right), \tag{2.52}$$

$$v_{{\rm g},\phi} \sim \frac{1}{2(1+\epsilon)}\left(\Delta v + \mathrm{St}\,\epsilon\, v_{\rm visc}\right), \tag{2.53}$$

$$v_{{\rm d},\phi} \sim \frac{1}{2(1+\epsilon)}\left(\Delta v - \mathrm{St}\, v_{\rm visc}\right). \tag{2.54}$$

  The radial velocity of small grains is mostly governed by the direct radial drag induced by viscous gas motion with a small correction term due to the back reaction $\propto 1/(1+\epsilon)$. For $\mathrm{St} = 0$, both gas and dust moves with $v_{{\rm d},R} = v_{{\rm g},R} = v_{\rm visc}/(1+\epsilon)$ as can be seen in Fig. 2.2 (upper panels). Small deviations $\propto \mathrm{St}\Delta v$ from these motions are linked to the exchange of angular momentum due to the different orbital motion $\propto \Delta v$. The orbital velocity

of small grains is forced by the gas drag to become sub-Keplerian in just a few stopping times. As a result, the centrifugal acceleration is not efficient enough to counterbalance the gravitational attraction of the central star, and the grain feels an inward radial differential acceleration. In turn, the gas drag counterbalances almost instantaneously this radial motion and the grain reaches a local limit velocity in a few stopping times. While for smaller particles with $\mathrm{St} \ll |v_{\rm visc}| \,/\, |\Delta v| \sim \alpha$, the direct drag dominates the radial motion, larger particles, i.e. $\alpha \ll \mathrm{St} \ll 1$, move radially inward faster than the gas under the effect of the headwind from the gas. The radial velocities of these particles increase linearly with St and reach a maximum for $\mathrm{St} = 1$, as described below.

In order to investigate the dust influence on the gas motion, we need to highlight the role played by the dust-to-gas ratio. From Eq. 2.51, if $\epsilon \gtrsim \alpha/\mathrm{St}$, the dust can dynamically affect the gas motion, pushing it back in the regions of lower pressure to conserve the global angular momentum (see the upper right panel of Fig. 2.2). Otherwise, if $\epsilon \lesssim \alpha/\mathrm{St}$ the dust back reaction is negligible and the gas moves with the unperturbed radial velocity $v_{\rm visc}$ decreased by a factor given by $(1+\epsilon)$.

Eqs. 2.53 – 2.54 show that for $\mathrm{St} \ll \alpha^{-1}$ the dust orbital motion is mostly dominated by the unperturbed gas motion, while for $\mathrm{St} \ll (\alpha\epsilon)^{-1}$, the gas does not feel any influence from the dust. For larger grains $\alpha^{-1} \ll \mathrm{St} \ll 1$, the orbital motion of dust (relative to the Keplerian velocity) is affected by the radial viscous flow while, for $(\alpha\epsilon)^{-1} \ll \mathrm{St} \ll 1$, the back reaction from the dust affects the gas orbital motion. However, since in typical protoplanetary discs $\alpha^{-1} \sim \left[10^2, 10^4\right]$ and $\epsilon^{-1} \sim \left[1, 10^2\right]$, we can safely assume that the dust and gas share the same sub-Keplerian orbital motion, $v_{\mathrm{d},\phi} = v_{\mathrm{g},\phi} = \Delta v/2(1+\epsilon)$ (see the lower panels of Fig. 2.2), which is essentially the unperturbed gas orbital motion decelerated by a term $\propto (1+\epsilon)^{-1}$. For $\epsilon \ll 1$, we obtain the typical gas orbital motion $\Delta v/2$. This highlights the role covered by the dust-to-gas ratio in the well-coupled regime.

- **Marginally coupled particles:** $\mathrm{St} = \mathbf{1}$

  This regime is particularly interesting because some terms in Eqs. 2.38-2.41 reach the maximum value and the orbital velocities change very rapidly. The dust and gas velocities are given by

$$v_{\mathrm{g},R} = -\frac{1}{1+\epsilon}\left[\frac{\epsilon\,\Delta v}{2} - \left(1+\frac{\epsilon}{2}\right) v_{\rm visc}\right], \tag{2.55}$$

$$v_{\mathrm{d},R} = \frac{1}{2\,(1+\epsilon)}\left[\Delta v + v_{\rm visc}\right], \tag{2.56}$$

$$v_{\mathrm{g},\phi} = \frac{1}{2(1+\epsilon)}\left[\left(1+\frac{\epsilon}{2}\right)\Delta v + \frac{\epsilon}{2} v_{\rm visc}\right], \tag{2.57}$$

$$v_{\mathrm{d},\phi} = \frac{1}{4(1+\epsilon)}\left[\Delta v - v_{\rm visc}\right]. \tag{2.58}$$

  In this regime, particles most efficiently exchange their angular momentum with the gas and drift at the highest rate. From Eq. 2.56, since $|\Delta v| \gg |v_{\rm visc}|$, the radial motion is given by

$$v_{\mathrm{d},R,\mathrm{peak}} \simeq \frac{\Delta v}{2\,(1+\epsilon)} = \frac{\zeta v_{\rm k}}{2\,(1+\epsilon)}\left(\frac{H_{\rm g}}{R}\right)^2, \tag{2.59}$$

  where ζ is defined in Eq. 1.16. On the other hand, the gas radial velocity reaches the maximum deviation from the radial viscous flow, leading to a fast drift toward outer regions. This effect becomes significant for $\epsilon \gg 2\alpha/(1-\alpha) \sim 2\alpha$ and cannot clearly be neglected to investigate the gas radial evolution (see the dashed and dotted-dashed line in the upper right panel of Fig. 2.2). As expected, the orbital motion of the dust increases from a value around half of the unperturbed gas orbital motion to the Keplerian value $v_{\mathrm{d},\phi} = 0$, while the gas, gaining angular momentum from the inward dust flow, accelerates

its orbital motion approaching the unperturbed value of the orbital velocity $\Delta v/2$ (see the lower panels of Fig. 2.2).

- **Poorly coupled particles:** $\mathrm{St} \gg 1$

  In this regime, the dust inertia dominates and does not dynamically interact with the gas and vice-versa. The dust and gas velocities are given by

$$v_{\mathrm{g},R} \sim -\frac{1}{1+\epsilon}\left[\frac{\epsilon\,\Delta v}{\mathrm{St}} - (1+\epsilon)\,v_{\mathrm{visc}}\right] \approx v_{\mathrm{visc}}, \tag{2.60}$$

$$v_{\mathrm{d},R} \sim \frac{1}{1+\epsilon}\left[\frac{\Delta v}{\mathrm{St}}\right] \approx 0, \tag{2.61}$$

$$v_{\mathrm{g},\phi} \sim \frac{1}{2(1+\epsilon)}\left[(1+\epsilon)\,\Delta v + \frac{\epsilon}{\mathrm{St}} v_{\mathrm{visc}}\right] \approx \frac{\Delta v}{2}, \tag{2.62}$$

$$v_{\mathrm{d},\phi} \sim \frac{1}{2(1+\epsilon)}\left[-\frac{v_{\mathrm{visc}}}{\mathrm{St}}\right] \approx 0. \tag{2.63}$$

  In this mode of migration, the stopping time is much larger than the Keplerian timescale. The gas and dust motions are therefore aerodynamically decoupled and their velocities no longer depend on the dust-to-gas ratio. Importantly, unlike the case $\mathrm{St} \ll 1$ where the pressure and viscous contribution need to be taken into account properly, in this regime $\mathrm{St} \gg \epsilon\alpha/(1+\epsilon)$ for typical values of α and ϵ. Therefore, the influence of the pressure (viscous) term in the radial (orbital) gas motion is negligible. The orbital velocity of a grain around the central star is almost the Keplerian velocity ($v_{\mathrm{d},\phi} \approx 0$ where we recall that $v_{\mathrm{d},\phi}$ is the relative Keplerian velocity) and the dust does not migrate since the centrifugal force equates the gravitational force ($v_{\mathrm{d},R} \approx 0$, see the upper panels of Fig. 2.2). As expected, the gas radial motion is governed by viscous force and its orbital motion is sub-Keplerian due to the pressure support ($v_{\mathrm{g},R} \approx 0$ and $v_{\mathrm{g},\phi} = \Delta v/2$, see the upper panels of Fig. 2.2).

The main results of this analysis is that the radial drift produces a fast migration of large grains, i.e. $\mathrm{St} \gg \alpha$, toward the inner regions at a rate that, in units of orbital timescale, given by

$$t_{\mathrm{drift}}\Omega_{\mathrm{k}} = \frac{R\Omega_{\mathrm{k}}}{|v_{\mathrm{d},R}|} = \frac{2\,(1+\epsilon)\,(\mathrm{St}+\mathrm{St}^{-1})}{|\zeta|}\left(\frac{H_{\mathrm{g}}}{R}\right)^{-2} \approx (1+\epsilon)\,\frac{1+\mathrm{St}^2}{\mathrm{St}}\left(\frac{H_{\mathrm{g}}}{R}\right)^{-2}. \tag{2.64}$$

Typical values of the minimum radial drift timescale occurring for particles with $\mathrm{St} = 1$ are of the order of $\sim 10^2\Omega_{\mathrm{k}}$, depending on the local value of the disc aspect ratio. Therefore, the radial velocity induced by gas drag can be high enough to produce a rapid depletion of marginally coupled particles. We will focus more on the radial drift motion of dust in Sect. 2.5.

2.3.2 Dust vertical settling in laminar gas flow

The vertical motion of dust grains is particularly important because, as we will discuss in Chapter 3, the vertical distribution of grains with different sizes heavily affects the appearance of discs and the slope of the SEDs.

Here we derive approximate solutions for the vertical motion of dust grains under the action of the gravitational force and the drag force. Due to the intrinsic dependencies of the grains aerodynamical properties and gravitational force to the radial distance, the dust vertical settling is expected to produce a size-sorted vertical distribution of dust grains at different radii, affecting the radial dust motion.

Writing Eq. 2.23 in cylindrical coordinates and assuming $\epsilon \ll 1$ (i.e. no back-reaction and $t_{\mathrm{s,m}} = t_{\mathrm{s}}$), the vertical motion of dust grains is given by the following Langevin equation (e.g.

Youdin & Lithwick 2007; Zhu et al. 2015b)

$$\frac{\partial v_{\mathrm{d},z}}{\partial t} = -\Omega_{\mathrm{k}}^2 z + \frac{v_{\mathrm{g},z} - v_{\mathrm{d},z}}{t_{\mathrm{s}}}, \tag{2.65}$$

where we have assumed that the disc is geometrically thin. This describes a damped oscillator stochastically forced by gas turbulence (embodied in $v_{\mathrm{g},z}/t_{\mathrm{s}}$, where $v_{\mathrm{g},z}$ is the vertical component of the gas velocity which is fluctuating stochastically due to turbulence) on the vertical direction.

Assuming that the the gas flow is stationary (i.e. $v_{\mathrm{g},z} = 0$), the equation of motion of dust grains is equivalent to the damped harmonic oscillator equation

$$\frac{\partial^2 z}{\partial t^2} + \Gamma \frac{\partial z}{\partial t} + \Omega_{\mathrm{k}}^2 z = 0, \tag{2.66}$$

where $v_{\mathrm{d},z} = \partial z/\partial t$ and $\Gamma \equiv 1/t_{\mathrm{s}} = \Omega_{\mathrm{k}}/\mathrm{St}$. We limit our analysis very close to the midplane, where we can assume that ρ_{g} (and thus St, see Eq. 2.37 for $\epsilon \ll 1$) is constant along the vertical direction. The equation for the vertical motion of dust grains is therefore in the form of a homogeneous second order differential equation with constant coefficients, a simple damped harmonic oscillator equation. It can be solved by looking for trial solutions of the form $z(t) = \exp(pt)$. Plugging this in Eq. 2.66 we obtain the characteristic equation with roots given by

$$p^{\pm} = \frac{1}{2}\left(-\Gamma \pm \sqrt{\Gamma^2 - 4\Omega_{\mathrm{k}}^2}\right). \tag{2.67}$$

There are three cases depending on the sign of the expression under the square root, which can be restated in terms of the Stokes number.

- **Underdamped oscillator:** $\mathrm{St} > \mathbf{1/2}$

  This motion characterises the settling of large dust grains, which are loosely coupled with the gas and are therefore expected to be able to oscillate around the midplane with a decaying amplitude. In this case, the term under the square root is negative and the characteristic roots are not real,

  $$p^{\pm} = -\frac{\Gamma}{2} \pm i\lambda \quad \text{with} \quad \lambda = \Omega_{\mathrm{k}}\sqrt{1 - \left(\frac{1}{2\mathrm{St}}\right)^2}. \tag{2.68}$$

  The general real solution is give by

  $$z(t) = \exp\left(-\frac{\Gamma}{2}t\right)\left[A\cos(\lambda t) + B\sin(\lambda t)\right], \tag{2.69}$$

  where A and B can be expressed in terms of the initial condition $z(0)$ and $\dot{z}(0)$ as

  $$A = z(0), \tag{2.70}$$

  $$B = \frac{1}{\lambda}\left(\frac{\Gamma}{2}z(0) + \dot{z}(0)\right). \tag{2.71}$$

  This solution represents an oscillation about the midplane with frequency

  $$\lambda = \Omega_{\mathrm{k}}\sqrt{1 - \left(\frac{1}{2\mathrm{St}}\right)^2} \approx \Omega_{\mathrm{k}}, \tag{2.72}$$

  and with a decaying amplitude $\propto \exp(-\Omega_{\mathrm{k}} t/2\mathrm{St})$ due to the decelerating effect given by the drag. The vertical oscillations of loosely coupled particles damp in a time, identified as *settling time*, which can be expressed in units of orbital timescale as

  $$\Omega_{\mathrm{k}} t_{\mathrm{settle}} \simeq \mathrm{St}. \tag{2.73}$$

  Loosely coupled particle at a height z above the disc mid-plane would therefore orbit on an inclined Keplerian orbit, thus oscillating around the midplane.

- **Critical damping:** $\mathrm{St} = \mathbf{1/2}$

  In this case, the term under the square root is null and the characteristic polynomial has repeated roots, $p = -\Gamma/2 = -\Omega_\mathrm{k}$. Since the roots are real, the solution no longer has an oscillatory part. The general solution is therefore given by

$$z(t) = (A + Bt)\exp\left(-\Omega_\mathrm{k} t\right), \tag{2.74}$$

  where A and B can be expressed in terms of the initial condition $z(0)$ and $\dot{z}(0)$ as

$$A = z(0), \tag{2.75}$$
$$B = \dot{z}(0) + \Omega_\mathrm{k} z(0). \tag{2.76}$$

  In this regime, the dust settling occurs on the fastest timescale Ω_k^{-1}. The motion is not oscillatory and goes quickly to the steady state on a orbital timescale.

- **Overdamped oscillator:** $\mathrm{St} < \mathbf{1/2}$

  In this case, the term under the square root is positive and the characteristic roots are real and distinct. The solutions are both real and can be linearly combined to obtain the general solution

$$z(t) = A\exp\left(p^+ t\right) + B\exp\left(p^- t\right), \tag{2.77}$$

  where

$$p^\pm = -\frac{\Gamma}{2} \pm \frac{\Gamma}{2}\sqrt{1 - (2\mathrm{St})^2}. \tag{2.78}$$

  The constants A and B are given by

$$A = z(0) + \frac{p^+ z(0) - \dot{z}(0)}{p^- - p^+}, \tag{2.79}$$
$$B = -\frac{p^+ z(0) - \dot{z}(0)}{p^- - p^+}. \tag{2.80}$$

  In order to have simple illustration of the motion of smaller grains, let us consider typical initial conditions. Assuming $z(0) = H_\mathrm{d}$ and $\dot{z}(0) = 0$, the characteristic roots and the constants can be expanded to the first order in St, yielding

$$p^+ = -\frac{\Gamma}{2}\left\{1 - \left[1 - (2\mathrm{St})^2\right]^{1/2}\right\} \approx -\mathrm{St}\Omega_\mathrm{k}, \tag{2.81}$$
$$p^- = -\frac{\Gamma}{2}\left\{1 + \left[1 - (2\mathrm{St})^2\right]^{1/2}\right\} \approx -\Gamma\left(1 - \mathrm{St}^2\right) \approx -\frac{\Omega_\mathrm{k}}{\mathrm{St}}, \tag{2.82}$$
$$A = H_\mathrm{d}\left(1 - \mathrm{St}^2\right) \approx H_\mathrm{d}, \tag{2.83}$$
$$B = -\mathrm{St}^2 H_\mathrm{d}, \tag{2.84}$$

  where it can be clearly noticed that $|p^+| << |p^-|$ and $|B| \ll |A|$. Thus, the first term of Eq. 2.77 reaches the null value at longer timescale compared to the second term and, additionally, it always dominates the motion in this regime, assuming that the oscillator is far from the critical regime where the two terms both dominates. The second term of Eq. 2.77 is therefore a fast transient that rapidly decays to zero. The solution can be therefore well approximated by

$$z(t) \approx H_\mathrm{d}\exp\left(-\mathrm{St}\Omega_\mathrm{k} t\right). \tag{2.85}$$

  The settling timescale is therefore given by

$$\Omega_\mathrm{k} t_\mathrm{settle} \simeq \frac{1}{\mathrm{St}}, \tag{2.86}$$

  much larger than the orbital period. Instead of oscillating about the midplane, these grains slowly settle toward the midplane on a timescale of a few hundred thousand years, in typical disc conditions.

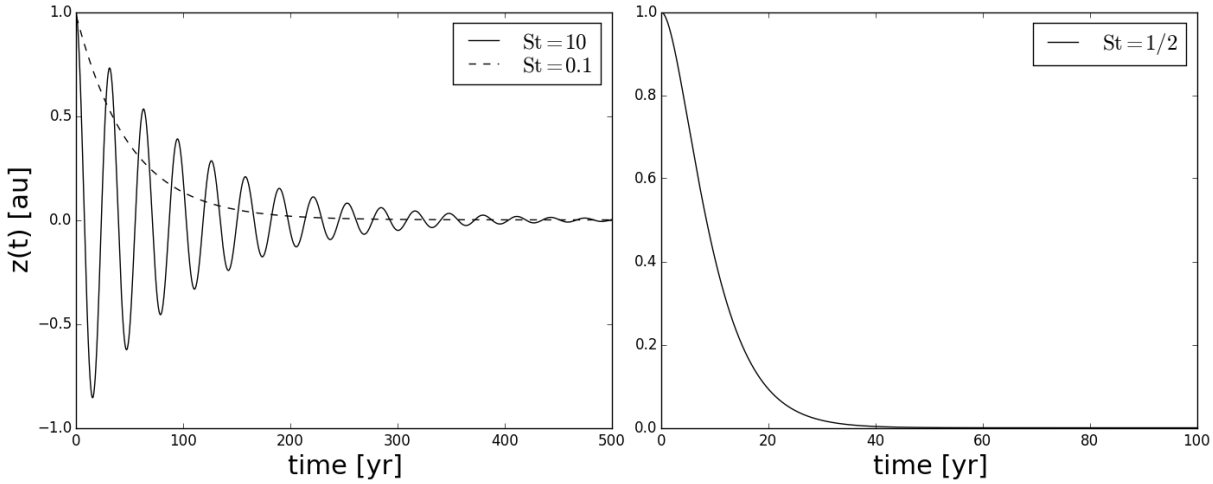

Figure 2.3: Vertical motion of a dust particle starting at $z(0) = 1$ au with $\mathrm{St} = 10$ (left panel - solid line), $\mathrm{St} = 0.1$ (left panel - dashed line) and $\mathrm{St} = 1/2$ (right panel) expressed in Eqs. 2.69, 2.74 and 2.77 at a radial distance of 10 au from a central star with mass $1\,M_\odot$.

Therefore, dust grains settle towards the midplane with a timescale that depends on their size. For small (large) grains, the dimensionless settling time is $1/\mathrm{St}$ (St), while the fastest settling time ($\sim \Omega_\mathrm{k}^{-1}$) occurs for grains with $\mathrm{St} = 0.5$. Fig. 2.3 shows the vertical motion of a dust particle starting at $z(0) = 1$ au with $\mathrm{St} = 10$ (left panel - solid line), $\mathrm{St} = 0.1$ (left panel - dashed line) and $\mathrm{St} = 1/2$ (right panel) in a typical disc at a radial distance of 10 au. Encompassing all the three cases, the settling timescale can be well approximated by (e.g. Youdin & Lithwick 2007; Laibe et al. 2014)

$$\Omega_\mathrm{k} t_\mathrm{settle} \simeq \mathrm{St} + \frac{1}{\mathrm{St}}. \tag{2.87}$$

As mentioned above, the vertical motion of dust grains can be analyzed regardless the radial drift motion. This is justified by the fact that the settling timescale is much shorter compared to the drift timescale. The ratio between the two timescales are given by

$$\frac{t_\mathrm{settle}}{t_\mathrm{drift}} \approx \left(\frac{H_\mathrm{g}}{R}\right)^2 \ll 1, \tag{2.88}$$

where we have used Eqs. 2.64 and 2.87. Thus, the vertical settling is almost a hundred times faster than radial migration and both the vertical and the radial motion can be considered independent.

Importantly due to the low density of the gas at the surface layers, even micron sized grains would have a large Stokes number. The stopping time is also a function of z due to its dependence on the gas volume density (see Eq. 2.37). Assuming that the disc is in vertical hydrostatic equilibrium, the vertical profile of the Stokes number can be expressed by

$$\mathrm{St} = \frac{\rho_\mathrm{p} a}{\Sigma_\mathrm{g}} \frac{\pi}{2} \sqrt{\gamma} \exp\left[\frac{z^2}{2H_\mathrm{g}^2}\right] \equiv \mathrm{St}_0 \exp\left[\frac{z^2}{2H_\mathrm{g}^2}\right], \tag{2.89}$$

where we have used Eqs. 1.10, 1.11, 1.12 and 2.37 and St_0 indicates the value of the Stokes number on the midplane that, for typical disc parameters, has the value of

$$\mathrm{St}_0 = \sqrt{\gamma} \left(\frac{\Sigma_\mathrm{g}}{0.5\,\mathrm{g/cm^2}}\right)^{-1} \left(\frac{\rho_\mathrm{p}}{3\,\mathrm{g/cm^3}}\right) \left(\frac{a}{1\,\mathrm{mm}}\right). \tag{2.90}$$

Eq. 2.89 shows that particles that are well coupled with the gas in the midplane can be poorly coupled in the disc upper layers. Fig. 2.4 shows the Stokes number of millimetre grains at

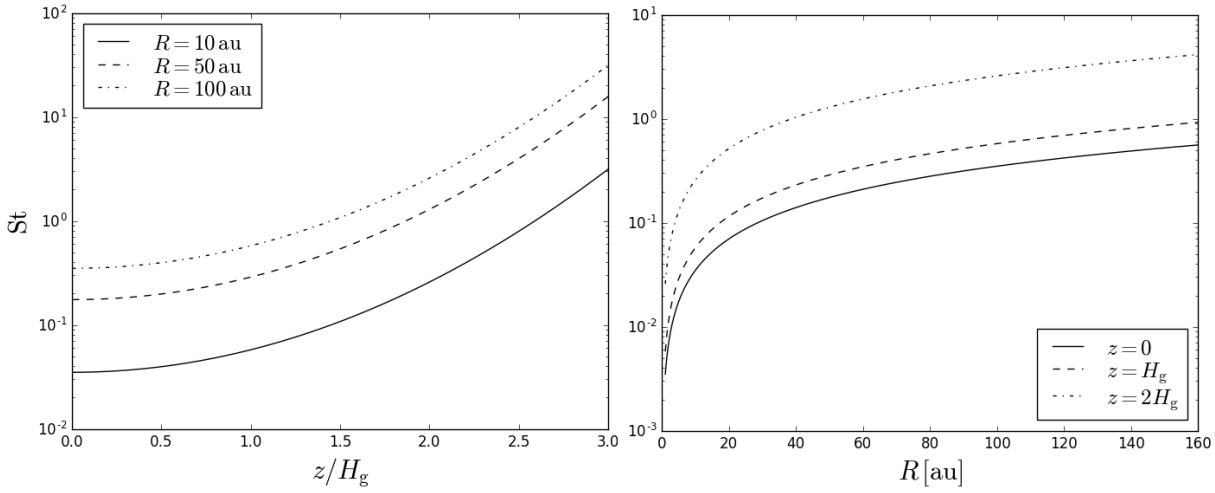

Figure 2.4: Left (Right) panel: Stokes number of millimetre sized grains as a function of z (resp., R) at different distances (altitudes) from the star (midplane) as expressed in Eq. 2.64. The different lines in the plot on the left (right) refer to different distances (altitudes). In the left panel the quantity on the x axis denotes the distance from the midplane z divided by the disc scale height evaluated at each radial distance shown in the plot. The disc model is described in the caption of Fig. 2.2, assuming a dust-to-gas ratio $\epsilon = 0.01$.

different altitudes above the midplane and at different distances from the star adopting typical disc parameters (see the Figure caption).

While dust settles to the midplane, its Stokes number will decrease due to the increased gas densities (see the left panel of Fig. 2.4). Therefore, non-growing grains will decelerate their motion. Otherwise, if they can grow while settling down, their increase in size can partly counteract the increase in gas density. If the settling and growth timescale are locally comparable, the Stokes number of the falling grains can stay large enough, such that they continue to settle towards the mid-plane, reaching macroscopic sizes and forming planetesimals (Dullemond & Dominik 2004, 2005; Laibe et al. 2014).

We can get more insights on the typical settling by assuming a steady state for the vertical motion. This assumption is a good approximation for smaller particles motion due to the high coupling with the gas, i.e. small dust particles reach very quickly the terminal velocity. In steady state, the terminal velocity can be estimated by setting the stellar gravity acceleration equal to the drag (or using Eq. 2.85). A dust grain located at height z settles towards the midplane at the velocity

$$\Omega_\mathrm{k} z = \frac{|v_\mathrm{settle}|}{\mathrm{St}} \quad \rightarrow \quad |v_\mathrm{settle}| = \mathrm{St}\Omega_\mathrm{k} z. \tag{2.91}$$

In the Epstein regime, the time for a grain to settle to a certain height z is therefore given by

$$t_\mathrm{settle} = \frac{z}{|v_\mathrm{settle}|} = \frac{1}{\Omega_\mathrm{k}\mathrm{St}} = \sqrt{\frac{8}{\pi}}\frac{c_\mathrm{s}\rho_g}{\Omega_\mathrm{k}^2 a \rho_\mathrm{p}} = \frac{2\Sigma_g}{\pi\Omega_\mathrm{k} a\rho_\mathrm{p}} \exp\left[-\frac{z^2}{2H_\mathrm{g}^2}\right], \tag{2.92}$$

where we have used Eq. 1.10. Therefore, dust settling depletes the upper layers of a disc surface much faster compared to the lower layers, especially in the outer disc region where the gas density is lower and St is higher (see the right panel of Fig. 2.4). Unless a mixing process counteracts the settling (see Sect. 2.4.2), the upper disc layer should become depleted of dust quickly, heavily affecting the radiative processes at the disc surface and, thus, the disc appearance (Dullemond & Dominik 2004).

2.4 Dust dynamics in turbulent flows

Gas in protoplanetary discs is highly turbulent (see Sect. 1.3.1) and, therefore, it is expected that the dust will be strongly affected by the turbulent gas motion. Previous analysis (see

Sect. 2.3) has focused on the radial and vertical motion of solids induced by drag assuming that the gas flow is laminar. It is therefore instructive to include radial and vertical diffusion in the dust motion analysis due to stochastic forcing by turbulent gas. The dust turbulent stirring induced by the drag is clearly dependent on the velocity fluctuations of the gas flow and their effect on the dust motion clearly depends on how strong the dust is coupled with the turbulent flow (quantified by the Stokes number, see Eq. 2.18). The gas and drag-induced dust turbulent diffusivity have been derived and numerically measured in a number of works (Voelk et al. 1980; Dubrulle et al. 1995; Schräpler & Henning 2004; Johansen & Klahr 2005; Carballido et al. 2006; Youdin & Lithwick 2007; Zhu et al. 2015b).

2.4.0.1 Drag-induced dust diffusivity

The aerodynamical coupling between gas and dust produce a drag-induced turbulent motion that can be modeled as an additional stochastic forcing on the equation of motion of dust particles (as an example, the term $v_{\mathrm{g},z}/t_{\mathrm{s}}$ in the dust vertical motion, Eq. 2.65). It is known that while the dust turbulent velocity on the vertical direction is due to stochastic vertical oscillations around the disc midplane, the turbulent dust velocities in the radial and azimuthal direction originates from stochastic epicyclic oscillations on the disc midplane (Youdin & Lithwick 2007). This suggests that the turbulent stirring induced by the gas is clearly anisotropic, and thus requires a proper treatment.

First of all, we need to consider the diffusivities of the gas in vertical and radial directions $D_{\mathrm{g},R}$ and $D_{\mathrm{g},z}$. Numerical simulations of turbulence flow driven by ideal magnetorotational instabilities (MRI) shows that both diffusivities can be well approximated close to the disc midplane by the kinematical viscosity, $D_{\mathrm{g},R} \sim D_{\mathrm{g},z} \sim \nu$ (Johansen & Klahr 2005; Carballido et al. 2006; Zhu et al. 2015b). Assuming the the dust back-reaction in negligible, Youdin & Lithwick (2007) found that the drag-induced dust diffusivity induced by drag on the vertical and radial direction in a unstratified gaseous disc are given by

$$D_{\mathrm{d},z} \simeq \nu \frac{1+\mathrm{St}}{1+2\mathrm{St}}, \qquad D_{\mathrm{d},R} \simeq \nu \frac{1+4\mathrm{St}^2}{\left(1+\mathrm{St}^2\right)^2} \simeq \frac{\nu}{1+\mathrm{St}^2}. \tag{2.93}$$

For tightly coupled particles, i.e. $\mathrm{St} \ll 1$, the dust and gas diffusivities are approximately equal while, for loosely coupled particles, i.e. $\mathrm{St} \gg 1$, $D_{\mathrm{d},z} = \nu/2$ and $D_{\mathrm{d},R} \ll \nu$. This result improves the previous findings of Dubrulle et al. (1995) which were valid for tightly coupled particles.

2.4.1 Radial drift in turbulent flows

The principal effect of diffusivity in a two-fluid mixture is to equalize the concentration of the dust relative to the dominant gas component. If gradients in concentration are present in the mixture, the diffusion will produce a mixing of both fluids in order to recover the homogenous concentration, smoothing out the dust-to-gas ratio. Let us limit our analysis on two-fluid mixtures with a dominant component in mass given by the gas (i.e. $\epsilon \ll 1$) and treat solids as passive trace species, i.e. the particles have no influence on the gas motion and spatially coexist with the gas (Morfill 1985; Clarke & Pringle 1988). The dust surface density radial evolution can be described in terms of an advective part, due to the systematic dust motions induced by the viscous and orbital motion of the gas (see Eq. 2.48), and a diffusive part proportional to the radial gradients of the dust-to-gas ratio. It is given by (e.g. Morfill & Voelk 1984; Clarke & Pringle 1988; Alexander & Armitage 2007; Brauer et al. 2008; Birnstiel et al. 2010)

$$\frac{\partial \Sigma_{\mathrm{d}}}{\partial t} + \frac{1}{R}\frac{\partial}{\partial R}\left\{R\left[\Sigma_{\mathrm{d}}\, v_{\mathrm{d},R} - D_{\mathrm{d},R}\,\Sigma_{\mathrm{g}}\,\frac{\partial}{\partial R}\left(\frac{\Sigma_{\mathrm{d}}}{\Sigma_{\mathrm{g}}}\right)\right]\right\} = 0. \tag{2.94}$$

It is worth remarking that the diffusive term in Eq. 2.94 was derived by Clarke & Pringle (1988) for the case where $\epsilon \ll 1$. The terms $v_{\mathrm{d},R}$ and $D_{\mathrm{d},R}$ in Eq. 2.94 are expressed in Eqs. 2.48 and 2.93,

respectively. We have also assumed that there are no sources or sinks of dust. Generally, this analysis can be extended by adding a source term on the right hand side of Eq. 2.94, which can be used to model the infall, re-condensation, dust evaporation or outflows. The equation for the dust density evolution can be solved analytically by making a number of assumptions related to the background gas density radial evolution and the dust diffusivity. In detail, it is worth recalling that $v_{\mathrm{d},R}$ and $D_{\mathrm{d},R}$ depend on the Stokes number (see Eqs. 2.48 and 2.93) and, therefore, the advecting and diffusive terms depend on the local value of the gas surface density. Generally, the solution of this equation is computed numerically using different techniques such as a standard first-order explicit scheme (e.g. Pringle et al. 1986; Alexander & Armitage 2007) or a flux-conserving donor-cell scheme (e.g. Birnstiel et al. 2010; Dipierro et al. 2015a). Moreover, the solution for the dust density evolution needs to be simultaneously computed with the evolution of the gas surface density (Eq. 1.26, Alexander & Armitage 2007; Birnstiel et al. 2010). However, the assumption of a constant gas surface density might be well justified by computing the dust surface density evolution on timescale smaller than the typical timescale of gas density evolution (Dipierro et al. 2015a), but might lead to mistakes in interpreting results computed on longer timescale (Pinilla et al. 2012b, 2015).

2.4.2 Dust settling in turbulent flows

In Sect. 2.3.2 we have shown that dust settling produces a fast depletion of the gas upper layers, producing a very steep vertical gradient of the dust density, especially for grain with $\mathrm{St} \sim 1$. In that analysis, we assumed that the gas vertical velocity is small compared to the particle vertical velocity, which is only valid in absence of sustained turbulent motions. However, the gas disc is expected to be highly turbulent, especially in the gas upper layers. The turbulent motion of the gas is expected to stir dust particles up to high altitude and prevent dust sedimentation.

As previously remarked, dust diffusions tends to counteract concentration gradients between the dust and gas phase. An equilibrium distribution of particles in the vertical direction can be achieved by imposing the balance between the sedimentation and the diffusion. The dust density evolution due to the dust settling and diffusion can be computed for low dust-to-gas ratio $\epsilon \ll 1$, in terms of an advection-diffusion equation, as already done for the radial motion,

$$\frac{\partial \rho_{\mathrm{d}}}{\partial t} - \frac{\partial}{\partial z}\left(z\Omega_{\mathrm{k}}\mathrm{St}\rho_{\mathrm{d}}\right) - \frac{\partial}{\partial z}\left[D_{\mathrm{d},z}\rho_{\mathrm{g}}\frac{\partial}{\partial z}\left(\frac{\rho_{\mathrm{d}}}{\rho_{\mathrm{g}}}\right)\right] = 0. \tag{2.95}$$

The advective term is related to the vertical sedimentation velocity that can be well approximated by Eq. 2.91, especially for small grains. Upon integrating once and rearranging terms, we obtain

$$\frac{\partial}{\partial z}\left(\ln\frac{\rho_{\mathrm{d}}}{\rho_{\mathrm{g}}}\right) = -\frac{\Omega_{\mathrm{k}}\mathrm{St}_0}{D_{\mathrm{d},z}} z \exp\left(\frac{z^2}{2H_{\mathrm{g}}^2}\right), \tag{2.96}$$

where we have used the Stokes number vertical profile expressed in Eq. 2.89, The integration of this equation requires the knowledge of the diffusion coefficient $D_{\mathrm{d},z}$ as a function of z, which is largely unknown. The dust diffusivity on the vertical direction can be well approximated by the gas diffusivity, that in turn can be approximated on the disc midplane by the kinematical viscosity, $D_{\mathrm{d},z} \sim D_{\mathrm{g},z} \sim \nu$, regardless of the value of the Stokes number (see Eq. 2.93). Although the gas diffusivity is expected to be a strong function of z, we can assume in our analysis that it is constant along the small vertical extension of the dust layer. Eq. 2.96 can be simply integrated to give (Carballido et al. 2006; Fromang & Nelson 2009)

$$\rho_{\mathrm{d}} = \rho_{\mathrm{d},0} \exp\left\{-\frac{z^2}{2H_{\mathrm{g}}^2} - \frac{\mathrm{St}_0}{\alpha}\left[\exp\left(\frac{z^2}{2H_{\mathrm{g}}^2}\right) - 1\right]\right\}, \tag{2.97}$$

where we have used the α-prescription: $\nu = \alpha H_{\mathrm{g}}^2 \Omega_{\mathrm{k}}$ assuming discs at vertical hydrostatic equilibrium (see Eq. 1.47 and 1.11, Shakura & Sunyaev 1973). Fig. 2.5 shows, in the left panel, the dust-to-gas ratio of millimeter particles (corresponding to Stokes numbers shown in the left

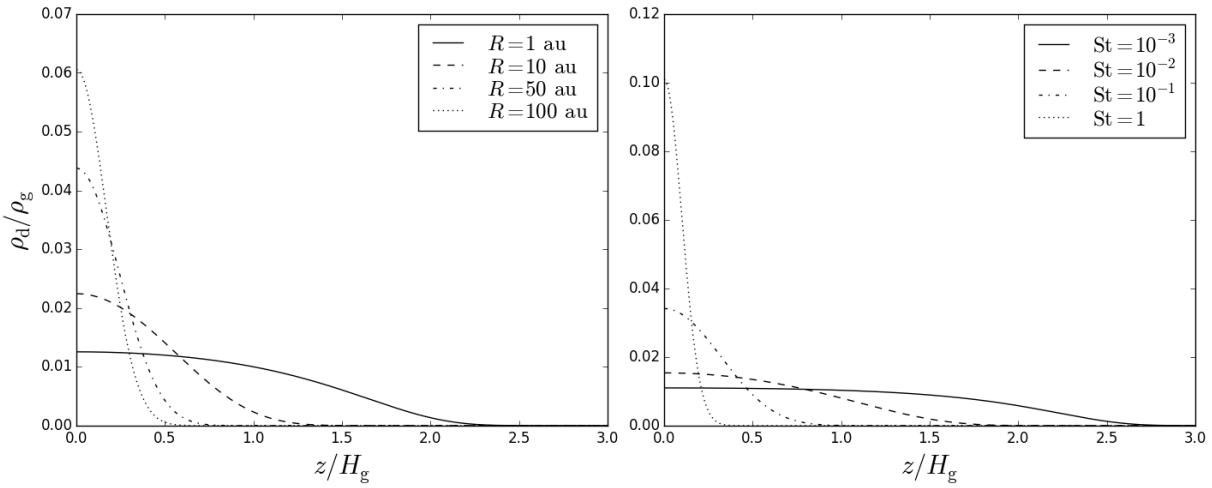

Figure 2.5: Left panel: Vertical profile of the dust-to-gas ratio of millimetre sized grains (their Stokes number is plotted in the left panel of Fig. 2.4) at different distances from the star. Right Panel: Vertical profile of the dust-to-gas ratio at a distance of 10 au for different Stokes numbers. The different lines in the plot on the left (right) refer to different distances (Stokes numbers). In the left panel the distance from the midplane z is normalized by the disc scale height at each radial distance shown in the plot. The vertical integral of the dust-to-gas ratio is assumed to be 0.01. The disc model is described in the caption of Fig. 2.2.

panel of Fig. 2.4) as a function of the altitude above the midplane at different distances from the central star. The right panel of Fig. 2.5 shows the dust-to-gas ratio vertical profile for different Stokes number. Decoupled particles, which are most likely to populate the outer parts of the disc surface layers (see Fig. 2.4) are more concentrated around the midplane. Importantly, close to the mid-plane ($z \ll H_g$ and $\mathrm{St}_0 \sim \mathrm{St}$), this profile approaches a Gaussian profile with a scale height given by

$$H_d = H_g \sqrt{\frac{\alpha}{\alpha + \mathrm{St}}}. \tag{2.98}$$

This equation clearly shows that dust settling towards the midplane is balanced by turbulent diffusion, which determines the thickness of the dust disc (Carballido et al. 2006; Fromang & Nelson 2009). It is worth emphasizing that estimates of the dust disc scale height obtained using a Gaussian fit for the dust density vertical distribution may lead to incorrect conclusions (see Fig. 2.5 for particles with low Stokes number or close to the central star), since The gaussian profile is more consistent with the vertical distribution of larger particles (Fromang & Nelson 2009), since they are expected to be strongly settled on the midplane (see the right panel of Fig. 2.5).

For $\mathrm{St}_0 \gg \alpha$, we obtain the well known relation for the dusty disc scale height (Dubrulle et al. 1995), which can be simply evaluated solving Eq. 2.95 assuming unstratified flows ($\rho_g(z) = \mathrm{const}$ and, thus, $\mathrm{St}(z) = \mathrm{const}$),

$$H_d = H_g \sqrt{\frac{\alpha}{\mathrm{St}}}. \tag{2.99}$$

Therefore, solid particles become strongly concentrated if their Stokes number exceeds α. Additionally, these results shows that the sedimentation of particles is more effective in the outer part of the disc, where the Stokes number is higher (see the right panel of Fig. 2.4), so the gas drag is weaker and turbulence cannot stir the particles up to high altitudes (Takeuchi & Lin 2002).

From an observational point of view, the measure of the degree of the dust settling though high-resolution (sub-)millimetre observations of non-face-on discs provides an estimate of the value of α, assuming that the emitting dust has a fixed value of Stokes number and assuming a value for the disc scale height of the gas (e.g. Pinte et al. 2016). The left panel of Fig. 2.5

shows that the dust close to the central star is spread out to large scale heights. The differential settling at different distances from the central star would produce a disc photosphere with a partly self-shadowed shape in the outer disc. Therefore the outer regions of the disc would be not anymore directly exposed to the stellar radiation, but only indirectly heated by infrared emission from other parts of the disc. Therefore, as we will describe in Sect. 3.3, the dust settling is expected to strongly affect the observational appearance of the disc, producing a much weaker mid to far-infrared excess of the SED and reducing the scattered light emission (Dullemond & Dominik 2004, 2005).

2.4.2.1 Implications on planet formation

In principle, dust settling in a low turbulent disc could produce a massive thin dust layer which can be gravitationally unstable, giving rise to gravitationally bound clumps and eventually leading to planetesimal formation (Safronov 1969; Goldreich & Ward 1973). This mechanism for the production of bound planetesimals, usually known as *Goldreich-Ward mechanism*, requires a low velocity dispersion of the particles[5] and an enhanced planetesimal surface density (around a factor $\sim 10 - 100$). These conditions may be sufficient to make the planetesimal sub-disc gravitationally unstable (Youdin & Shu 2002). However, as described above, the gravitational collapse of smaller solids into planetesimals is opposed by turbulence (Weidenschilling 1980; Youdin & Shu 2002). Even low levels of turbulence will inhibit sedimentation of solids to a sufficiently dense midplane layer. Moreover, the coupling between the gas and a the subdisc of solids induces turbulent mixing that might produce high dispersion of the dust velocities (especially for tightly coupled particles, i.e. $\mathrm{St} \ll 1$), preventing dust grains from continuously settling well before gravitational instability is able to operate and, thus, stabilizing the dust subdisc (Weidenschilling 1980; Supulver & Lin 2000; Youdin & Shu 2002). Additionally, also loosely coupled grains, due to their rapid infall, could generate sufficient velocity dispersions in such a population of bodies and, therefore, inhibit the onset of any instability.

Furthermore, the key requirement for the Goldreich-Ward mechanism is that the local particle density has to exceed the gas density. In this case, as described in Sect. 2.3.1, the dust feedback produces a significant influence on the gas on the dust particle layer. The dust-feedback on the gas motion would accelerate the sub-Keplerian velocities of the gas up to the Keplerian values within the dust layer, while dust in the upper gas layer would not affect the sub-Keplerian gas orbital motion (see Sect. 2.3.1, Weidenschilling 1977a; Nakagawa et al. 1986). The resulting vertical stratification of the gas velocity inside and above the dust disc would lead (if the vertical shear is large enough) to the development of Kelvin-Helmotz instabilities, which may induce a turbulent stirring of the gas up to velocities of several metres per second. In turn, this self-excited turbulence in the gas flow might produce additional velocity dispersions in the dust layer due to drag. These arguments suggest that a self-induced limit to the scale height of the dust layer exists (Cuzzi et al. 1993), further challenging the Goldreich & Ward's mechanism to create bound objects. Alternative mechanism, such as particle trapping in overpressure regions induced by long-lived vortices (e.g. Barge & Sommeria 1995; Birnstiel et al. 2013; Lyra & Lin 2013) or gravitationally-induced spiral modes (see Chapter 6, Rice et al. 2004, 2006b), might trap particles and produce high dust densities, accelerating the collisional growth and triggering a fast gravitational collapse (see the next Section, Johansen et al. 2006, 2007).

[5]It can be derived a condition for the linear stability of a massive thin dust layer against the effect of the disc self-gravity. This is given by the parameter (Toomre 1964; Bertin 2014),

$$Q \equiv \frac{\sigma \Omega}{\pi \mathcal{G} \Sigma_{\mathrm{d}}} < 1, \tag{2.100}$$

where σ is the particle velocity dispersion, Ω is the angular velocity of the dust, $\mathcal{G}$ is the gravitational constant and Σ_{d} is the local dust density at equilibrium. This condition will be derived and discussed in Sect. 4.2.1 for the gas disc.

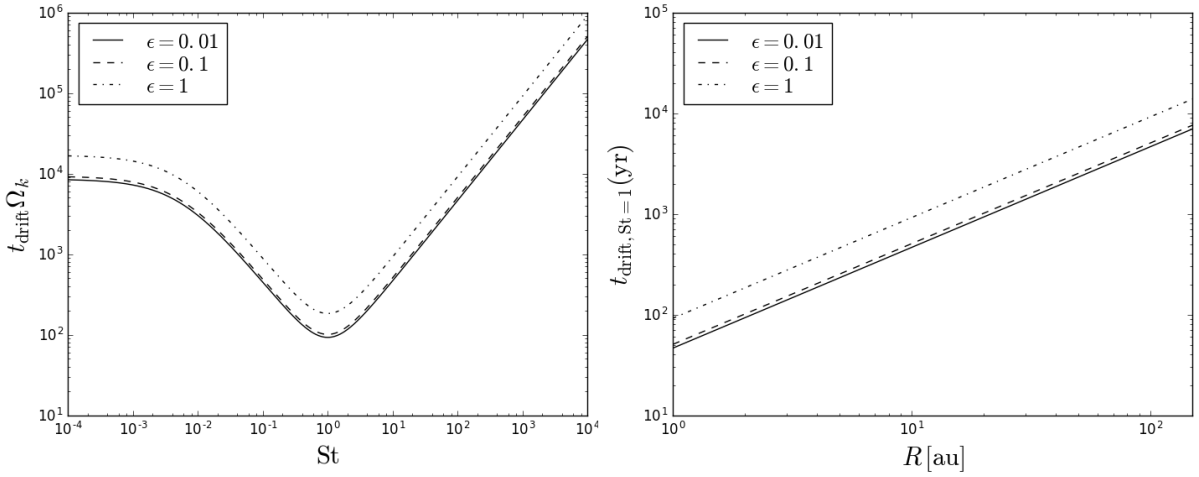

Figure 2.6: Left panel: Radial drift timescale in units of orbital time as a function of Stokes number St as expressed in Eq. 2.64. Right panel: The minimum timescale for the radial drift of solid particles as a function of radius. The plots show the timescale for three values of the dust-to-gas ratio: 0.01 (solid), 0.1 (dashed) and 1 (dotted-dashed). The disc model is described in the caption of Fig. 2.2. The plot on the left shows the drift timescale computed at a radial distance of 10 au.

2.5 The radial drift barrier

The main results of our analysis (see Sect. 2.3.1) is that the radial drift produces a fast migration of large grains (typically millimetre-centimetre grains in the outer disc regions, see the right panel of Fig. 2.4) toward the central star at a rate that, in units of orbital timescale, is expressed in Eq. 2.64. The left panel of Fig. 2.6 shows that the drift timescale can be extremely short, of the order of 10^2 orbital timescale. Particles with St = 1 experience the highest radial drift and migrate towards the star at the minimum timescale. The right panel of Fig. 2.6 shows that particles with St = 1 at 1 au drift toward the star on timescale as short as 100 years, for reasonable disc parameters. For typical discs, particles with sizes of the order $\sim$ mm are expected to be marginally coupled with the gas in the outer disc regions and drift quickly toward the star. Moreover, since discs are expected to be flared at large radii, the minimum radial drift timescale increases with radius as $\propto (H_{\rm g}/R)^{-2} \sim R^{q+1/2}$ (see Eq. 2.64), assuming a disc model with $c_{\rm s} \propto R^{-q/2}$. Since the temperature in the inner disc regions is sufficiently high ($\gtrsim 1000K$) to evaporate solid material, the radial drift of dust particles leads to a continuous loss of solids, preventing the growth of planetesimals required for the formation of the planetary cores. This is what is known as *radial drift barrier*, and it still represents one of the main unsolved problem of the early phases of planet formation (Weidenschilling 1977a; Nakagawa et al. 1986; Brauer et al. 2008).

However, during the radial drift, since the gas density increases with decreasing radius, particles tend to couple with the gas and decrease their motion. Moreover, the dust concentration in the inner regions produces an increase of the dust-to-gas ratio (see Sect. 2.5.1). Therefore, (non-)growing dust grains could decelerate and, in some cases, stop their migration (Laibe et al. 2012; Laibe 2014).

One of the possible ways to prevent loss of solids is to assume that the grain growth processing has to be efficient enough to decouple the grains from the gas and stop their migration. However, the growth of solid particles has to circumvent another hurdle closely linked to the sticking efficiency of colliding dust grains. It is expected that, while low-velocity collisions result in grain growth, high-velocity impacts between particles are destructive (Poppe et al. 1999; Blum 2006; Blum & Wurm 2008). From laboratory and numerical studies it has been shown that, as particles become larger, they tend to collide at higher impact velocities (Güttler et al. 2010; Zsom et al. 2010). Therefore, it is expected that the velocities of millimetre-sized objects in the outer disc are larger than the critical threshold for sticking, so that collisions lead to fragmentation and thus

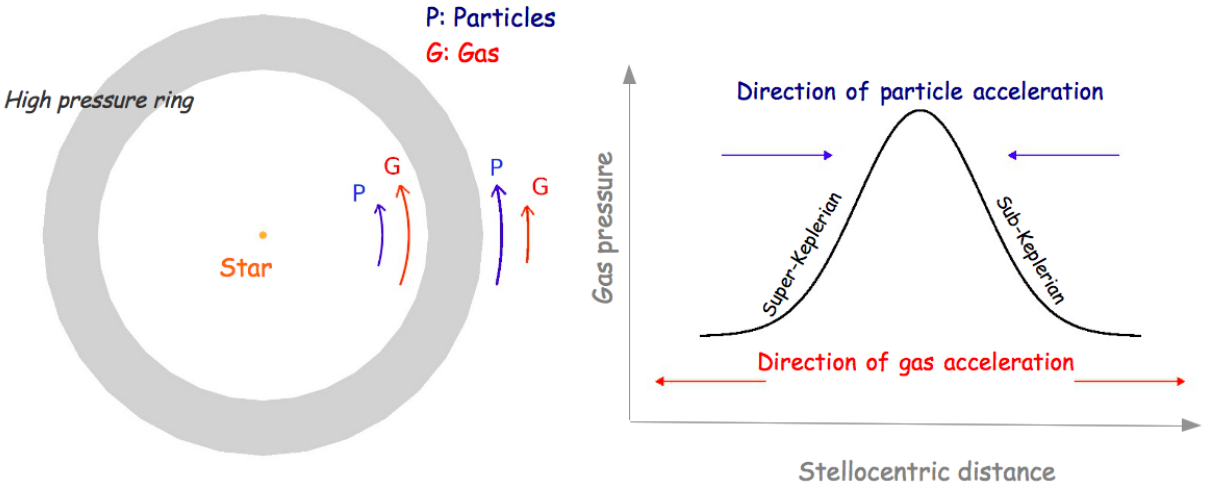

Figure 2.7: Particle trapping into a over pressure regions due to the different drag force between the rising and falling trend of the radial pressure profile. Image adapted from Whipple (1972).

prevent dust particles from forming larger bodies (Brauer et al. 2008). As a result, the growth of dust particles toward meter sizes needs to overcome two obstacles: the rapid depletion of material due to the radial drift and the particle fragmentation produced by destructive collisions.

From an observational point of view, the evidence that grains in discs are significantly larger than those in the diffuse interstellar medium has been obtained from a variety of observational techniques (see Chapter 3 and the review of Testi et al. 2014). However, multi-wavelengths (sub-)millimetre observations of protostellar discs have shown that grain growth is efficient enough to quickly produce millimetre-sized particles which are retained in the outer disc for a relatively long time (for more details see Sects. 3.3.1 and 3.3.2, Ricci et al. 2010; Pérez et al. 2012). Therefore, how to theoretically explain the retention of millimetre-sized particles in the outer regions of discs remains an open question.

Several ideas of plausible mechanisms aimed at slowing down the rapid inward drift have been proposed, such as vortices, planet-disc interactions, snowlines, zonal flows or dead zones (e.g. Johansen et al. 2009a, 2011; Pinilla et al. 2012b; Lyra & Lin 2013; Ruge et al. 2016). The common feature of these mechanisms is to create long-lived radial and/or azimuthal inhomogeneites in the gas density structure in order to "trap" particles into the overpressure regions (Haghighipour & Boss 2003b,a). In detail, since the gas pressure gradient changes from positive to negative on one side of the inhomogeneity to negative on the other, the gas orbital velocity changes from super- to sub-Keplerian causing dust grains to drift toward the density/pressure maxima (see Fig. 2.7). On the other hand, the dust back-reaction, which is most effective for particles with $\mathrm{St} = 1$ and $\epsilon = 1$, tends to move the gas out of the overpressure region towards pressure minima (see Sect. 2.3.1), deforming the pressure bump and avoiding further trapping (Taki et al. 2016). The timescale to accumulate dust particles in the over pressure region is faster than the global drift timescale (see Eq. 2.64). In detail, if the pressure inhomogeneity has a scale length ΔR, the local radial dust velocity is $\sim \Delta v \propto \partial P/\partial R \sim P/\Delta R$, exceeding the unperturbed radial dust motion $\propto P/R$. The timescale for solids to pile-up at the pressure maximum is further reduced by factor $\Delta R/R$ because the planetesimals have to move only a distance of order ΔR to reach the density maxima. Typical density/pressure inhomogeneities are characterized by a typical length scale $\sim H_{\mathrm{g}}$. Therefore, the timescale to concentrate solids at the pressure maximum is shorter than the global drift timescale by a factor of $(\Delta R/R)^2 \sim (H_{\mathrm{g}}/R)^2 \sim 0.01$, approaching values of the order of $\sim$ one year for marginally coupled particles in typical discs (see Fig. 2.6).

The dust concentration in pressure bumps is the result of the competition between the radial drift, that forces the particle to migrate towards pressure maxima, and the turbulent radial diffusion that tends to recover the dust density distribution into the initial condition where the dust-to-gas ratio is equal in the overall disc. In order to figure out the relative

importance of advection and diffusion, we can compare the advection timescale, that can be estimated by the minimum time required to concentrate solid material in a pressure maximum, $t_{\rm adv} \sim \Delta R/\left|v_{{\rm d},R}\right|$, and the diffusion timescale $t_{\rm diff} \sim \Delta R^2/D_{\rm d,R}$, where ΔR is the scalelength of the density inhomogeneities and $D_{\rm d,R}$ is the radial dust diffusivity expressed in Eq. 2.93. In the limit $\epsilon \ll 1$, the ratio of these timescale is given by

$$\frac{t_{\rm diff}}{t_{\rm adv}} = \frac{\rm St}{\alpha}\frac{\Delta R}{\rho_{\rm g}}\frac{\partial \rho_{\rm g}}{\partial R} \eqsim \frac{\rm St}{\alpha}\frac{\Delta \rho_{\rm g}}{\rho_{\rm g}}. \tag{2.101}$$

Thus, for particles with ${\rm St} \gtrsim \alpha\,(\Delta\rho_{\rm g}/\rho_{\rm g})^{-1}$, the radial migration towards pressure maximum dominates over turbulent diffusion. For smaller particles, diffusion dominates over advection, implying that variations in the dust-to-gas ratio are efficiently smeared out.

2.5.1 The role of the dust back-reaction

Dust feedback on the gas motion can produce remarkable effects on the gas evolution and might solve the problem of the radial drift barrier. As previously described, the key role in the dust feedback mechanism is covered by the dust-to-gas ratio.

In literature, the dust feedback is usually neglected. This is justified by the fact that the adopted value of the dust-to-gas ratio is 10^{-2}, similarly to what derived in the interstellar medium (Bohlin et al. 1978; Savage & Mathis 1979). However, the dust motion in protoplanetary discs leads to the formation of a more compact concentration of dust grains both radially and vertically. To gain more insights on the importance of the dust feedback on the gas motion, let us derive the typical dust-to-gas ratio in protoplanetary discs due to the two main drivers of dust concentration: dust settling and radial drift.

A simple estimation of the dust-to-gas ratio enhancement induced by radial drift can be obtained by comparing the radial extension of the gas and dust disc inferred by observations. Since most of the mass is concentrated in larger grains $\gtrsim 1$ mm, recent observations of dust continuum emission might be used to infer the dust mass. As an example, de Gregorio-Monsalvo et al. (2013) presented dust continuum observations of the source HD 163296 at 850 μm and CO spectral line emission which revealed a compact dusty disc with an outer radius of ∼240 au and a gaseous disc with an outer radius of ∼580 au. If we assume that the initial (when the gas and dust were well-mixed) dust-to-gas ratio in millimetre grain was $\epsilon_0 = 10^{-2}$, the current dust-to-gas ratio in the compact dusty disc is given by

$$\epsilon_{\rm drift} = \epsilon_0 \left(\frac{R_{\rm g,out}}{R_{\rm d,out}}\right)^2 \approx 6\,\epsilon_0. \tag{2.102}$$

Therefore, the dust radial migration increases the dust-to-gas ratio by concentrating large particles in the inner disc.

Moreover, the settling produce a fast sedimentation of large grains. If we assume a Stokes number of unity and a turbulence parameter $\alpha \sim 10^{-3}$, the dust-to-gas ratio on the midplane of the dusty disc due to the combined effect of drift and settling is given by

$$\epsilon_{\rm settl+drift} = \epsilon_{\rm drift}\frac{H_{\rm g}}{H_{\rm d}} \sim \epsilon_{\rm drift}\sqrt{\frac{\rm St}{\alpha}} \approx 2, \tag{2.103}$$

where we have used Eq. 2.99 as an estimate for the disc scale height.

Therefore, since the radial drift velocity is proportional to $(1+\epsilon)^{-1}$, the increased dust-to-gas ratio produces a decrease of the radial dust velocity by a factor of ~ 3. On the other hand, the gas acquires angular momentum from the dust and its radial velocity increases due to pressure effect by the factor $\epsilon/(1+\epsilon) \sim 10^2$, as expected. It is worth recalling that, as already discussed in Sect. 2.3.1, particles with ${\rm St} \ll 1$ affect the gas motion if $\epsilon \gg \alpha/{\rm St}$, while for marginally coupled particles (${\rm St} = 1$), the dust back reaction cannot be neglected for $\epsilon \gg 2\alpha$.

Additionally, the increase of the dust-to-gas ratio produces an enhanced collision rate. This can be simply illustrated by considering the the mean growth time of monomers with mass m at a fixed distance from the star as given by (Brauer et al. 2008; Okuzumi et al. 2011)

$$\tau_{\rm growth} = \left(\frac{1}{m} \frac{\partial m}{\partial t} \right)^{-1}, \tag{2.104}$$

where $\partial m/\partial t = \rho_{\rm d} \sigma_{\rm coll} v_{\rm coll}$ with $\sigma_{\rm coll} = \pi a^2$ is the collisional cross section and $v_{\rm coll}$ is the collisional velocity. Assuming that the collision velocity is mostly induced by turbulence (see Eq. 2.14), the collisional velocity for two equally sized particles with ${\rm St} \lesssim 1$ is $v_{\rm coll} \sim \sqrt{\alpha {\rm St}} c_{\rm s}$ (Cuzzi & Hogan 2003; Ormel & Cuzzi 2007; Birnstiel et al. 2016). After a little algebra, it can be shown that the growth timescale is proportional to (e.g. Brauer et al. 2008)

$$\tau_{\rm growth} \propto \frac{1}{\epsilon \Omega_{\rm k}}, \tag{2.105}$$

where it can be clearly noticed that an increase of the dust-to-gas ratio leads to an increase of the growth rate.

This analysis highlights the role of the dust-to-gas ratio in the grain growth processing rate and how the radial drift and dust settling have a remarkable effect on the formation of rocky planets. From an observational point of view, recent detections of gas and dust emission are revealing an enhanced dust-to-gas ratio on the midplane of the dusty disc (Muto et al. 2015; Pinte et al. 2016; Ansdell et al. 2016), though uncertainties remain in the exact value of the local gas mass (see Sect. 3.3.3).

Finally, the dust-to-gas ratio enhancement covers a crucial role for the growing rate of perturbations in the mixture. Youdin & Goodman (2005) demonstrated that perturbing the steady state solutions of dust and gas motion (Eqs. 2.38 – 2.41) can eventually lead to instabilities, known as *streaming instabilities* (Youdin & Johansen 2007; Johansen et al. 2007; Jacquet et al. 2011). The energy to sustain the development of these perturbations is tapped from the background differential rotation between the two phases. This instability is particularly relevant for planet formation, since it provides a mechanism to accumulate solid material, enhancing local particle density by a factor of up to 10^3. With the aid of self-gravity, these perturbations could eventually collapse and form planetesimals locally (Johansen et al. 2009b). Importantly, the growth rate of these instabilities increases with ϵ, due to the increased dust feedback on the dust motion (Youdin & Johansen 2007). Moreover, the streaming instabilities can trigger dust grain growth from metre-sized boulders to kilometre-scale, producing a gravitational collapse on timescale much faster than radial drift, offering a possible path to planetesimal formation in accreting circumstellar discs (Johansen et al. 2007, 2009b).

Chapter 3

Observations of Protoplanetary Discs

Disc imaging at different wavelengths provides insights into a large number of fundamental physical properties of the disc population, which allows to put constraints on models of disc evolution and planet formation. The link between observations of protostellar discs and their physical properties is complicated: a variety of optical effects such as anisotropic scattering, extreme optical depths or asymmetries in the structural composition, temperature and spatial density of dust grains need to be properly taken into account in interpreting observations.

Although the disc is dominated by the gas component in mass, the opacity budget is controlled almost everywhere by the dust particles at most wavelengths. The widely used diagnostic tools to detect discs and characterize their structure is essentially based on the analysis of the light scattered by dust grains, spectral dust features and the thermal emission of dust. The key to obtain quantitative insights on the disc properties is to study the way in which dust particles interact with starlight and acquire energy directly from the gas.

In this Chapter we review the basic framework to derive disc properties starting from observations over a wide range of wavelengths with different techniques. For a detailed description on this topic we refer to the recent reviews of Natta et al. (2007); Testi et al. (2014); Andrews (2015) and Woitke et al. (2016).

3.1 Spectral energy distribution

The most extensively used technique to characterise the structure of protostellar discs is the measure of the Spectral Energy Distribution (SED). Each part of the SED cointains informations about the surface density structure, temperature and optical properties of dust, that might be combined to build a disc model and gain insights on the grain growth processing and dust and gas dynamics.

Fig. 3.1 shows an example of an SED fitted by a simple disc and protostar model (Tilling et al. 2012). The SED can generally be divided in 5 parts:

1. The spectral region at ultraviolet and optical wavelengths is dominated by the emission from the protostar. This region of the SED is mostly analyzed to obtain the star properties by fitting the SED with a black body spectrum.

2. The near-infrared region is dominated by the thermal emission from the hottest dust (ranging from the dust sublimation temperature, from ∼1500 K to ∼ 150 K) very close to the star. Since the disc is typically optically thick at these wavelengths, its emission is directly related to the temperature and the area of the emitting region (see Sect. 3.2). Moreover, at these wavelengths, the scattering emission provides another key resolved diagnostic to infer structural properties of discs (see Sect. 3.2.1).

3. The silicate emission region in the mid-infrared part of the SED is extraordinarily informative on optical properties of dust grains, which are in turn related to their physical properties such as composition, size and shape. These features may also be connected to the grain growth processing in protoplanetary discs (see Sect. 3.2.2).

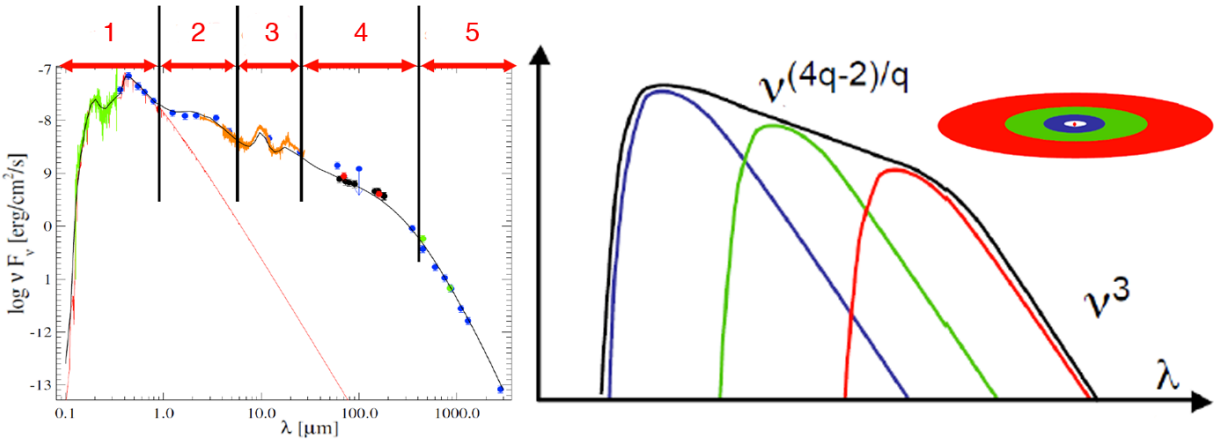

Figure 3.1: (left) The SED of HD163296 measured by different instruments over a large range of wavelengths (Tilling et al. 2012). The SED is fitted by a simple disc model (black line) and a photospheric model (red line). (right) Multicolor black body emission model.

4. The so-called multicolor region at far-infrared wavelenghts probes the thermal emission of dust grains at the disc surface at a few au radial scale. This emission is typically optically thick and can be used to infer the temperature radial profile.

5. Sub-mm part of the SED is typically optically thin and can be directly related to the dust temperature, density and opacity, which in turn is related to the grain properties (see Sect. 3.3).

Importantly, assuming that the disc is in thermodynamical equilibrium (the specific intensity is proportional to the Planck function for the black body emission, Rybicki & Lightman 1979), the thermal emission trace colder (warmer) material at longer (shorter) wavelengths. Therefore, since warm dust is expected to be closer to the star (see Sect. 1.4), SEDs offers a $\lambda \rightarrow R$ mapping, tracing different radial region of the disc at different wavelengths (Dullemond et al. 2007).

In the typical SED modeling approach, the first step is to assume a set of simple geometrically thin and optically thick disc models with an optically thin surface layer (e.g. Chiang & Goldreich 1997). The SEDs computed from these models through radiative transfer models are then compared against the observed SEDs to find the best-fitting parameters. However, disc models based on SED-only fitting are usually affected by large uncertainties because the SED samples the emission from the whole disc. Protostellar discs are described by varying quantities across all the disc extent such as density, temperature and, importantly, the opacity of dust grains. Moreover, non-axisymmetric density features in specific locations in the disc might produce SED features that cannot be interpreted in a proper way. To improve the reliability of this analysis and break some key structural degeneracies, it is necessary to merge the informations derived from the SED analysis with the structural constraints obtained by performing high-resolution observations (see Sect. 3.2.1 and 3.3).

3.2 From near to mid-infrared observations

The spectral range from near to mid-infrared wavelengths is particularly interesting because the disc emission is typically optically thick, i.e. its luminosity reflects the temperature at the disc surface and the area of the emitting region and its slope is related to the radial temperature gradient (Rybicki & Lightman 1979; Kenyon & Hartmann 1987). The disc heating is strongly dominated by the absorption of light coming from the central source (whose spectrum is expected to peak at a wavelengths of $\sim$ μm for typical T-Tauri star) by the small dust grains (from $\sim 0.1\,\mu$m to $\sim 20\,\mu$m, Miyake & Nakagawa 1993). The irradiated energy coming from the central protostar is partly scattered outward and inward (see Sect. 3.2.1), and partly absorbed

in lower disc layers which reradiate energy at longer wavelengths where the disc is transparent (see Sect. 3.3).

Assuming that the material is in local thermodynamic equilibrium and optically thick, it is possible to simply evaluate the spectral energy distribution. Each annulus in the discs radiates as a blackbody at the local temperature of the dust grains $T(R)$. The total flux at frequency ν for an observer at a distance D whose line of sight makes an angle i to the normal of the disc plane (where $i = 0^\circ$ indicates the case of a face-on disc) is expressed by

$$F_\nu = \frac{2\pi \cos i}{D^2} \int_{R_{\rm in}}^{R_{\rm out}} B_\nu(T(R)) R \, dR, \tag{3.1}$$

where $R_{\rm in}$ and $R_{\rm out}$ are the inner and outer radius of the disc and B_ν is the Planck function for the black body emission,

$$B_\nu(T) = \frac{2h\nu^3}{c^2} \frac{1}{e^{h\nu/k_B T} - 1}. \tag{3.2}$$

where h is the Planck's constant, c is the speed of light and k_B is the Boltzmann's constant. This spectrum is usually referred to as mutlicolor black body spectrum (see the right panel of Fig. 3.1). For low frequencies the Planck function takes the Rayleigh-Jeans form,

$$\nu F_\nu \propto \nu^3, \tag{3.3}$$

while, for high frequencies, it recovers the Wien form with an exponential cut-off related to the hottest annulus of the disc (at the disc inner radius $R_{\rm in}$),

$$\nu F_\nu \propto \nu^4 \exp\left[-\frac{h\nu}{K_{\rm B} T(R_{\rm in})}\right]. \tag{3.4}$$

For intermediate wavelengths, assuming that the temperature profile is given by a power law $T(R) \propto R^{-q}$, it can be shown that (e.g. Hartmann 2009)

$$\nu F_\nu \propto \nu^{4-2/q}. \tag{3.5}$$

These equations provides a realistic first picture of the disc SEDs at short wavelengths.

3.2.1 Scattered light

Due to the steep wavelength dependence of the scattering opacity of small dust grains ($\kappa_{\rm scat}(\lambda) \propto \lambda^{-4}$), the scattering emission from small dust grains at the disc surface layer dominates at short wavelengths, roughly from optical to near-infrared (NIR) part of the spectrum. Small dust grains carry most of the opacity at optical and near-infrared wavelengths, and therefore control the stellar absorption and the disc's temperature (Hartmann 2009). As described in Chapter 2, small dust grains are expected to be well coupled with the gas in the inner and intermediate disc regions and populate the full vertical extent of the gas disc. Therefore, such observations can be considered a fundamental probe of the disc vertical structure (see the review of Watson et al. 2007) and can be used to infer the disc inclination, the shape of the surface and detect features at the surface layer (e.g. induced by gaps carved by planets, Jang-Condell & Turner 2012). Importantly, even small features on the disc surface layers can cast long shadows, heavily affecting the resulting scattered light emission. As an example, if a planet embedded in the disc produces a gap in the gas structure (and, consequently, in the micron-sized dust grains distribution) around its orbit, the gap's outer wall receives extra starlight heating, leading to a puff up of the gas surface layer and throwing a shadow across the disc beyond. In this context, Isella & Turner (2016) have shown that this shadow can appear in scattered light as an additional dark ring, which could be mistaken for a gap opened by another more distant planet. Care should therefore be taken in interpreting scattered light images of protoplanetary discs.

Direct imaging of edge-on discs in scattered light can be used to derive the disc scale height (Perrin et al. 2006; Glauser et al. 2008). A resolved map of scattering emission typically appears as a bipolar pair of conical nebulae which represent the layers on each side of the disc. These observations may reveal a strong variable inclination, usually known as warps (Golimowski et al. 2006; Boccaletti et al. 2015; Kasper et al. 2015), whose dynamical evolution is closely dependent on the gas viscosity and disc aspect ratio (Lodato & Pringle 2007; Lodato & Price 2010; Facchini et al. 2013, 2014). For intermediate inclinations (most likely to occur in disc observations), the inference of the disc scale height suffers some degeneracies from the viewing geometry (Grady et al. 1999; Pinte et al. 2008; Takami et al. 2014) and the directional optical properties of the scatterers (Min et al. 2016). For face-on discs, the scattering intensity profile can be used to infer the disc scale height and the shape of the surface layer by comparing the scattering emission with that expected from theoretical investigations (Jang-Condell & Sasselov 2003; Inoue et al. 2008; Akiyama et al. 2015). Importantly, warped inner discs can produce azimuthal surface brightness reductions in the scattering emission of nearly-face-on discs (Marino et al. 2015, see also Sect. 4.3.2.4).

Interestingly, since the scattering radiation at a certain wavelength λ probe particles with size comparable with λ, multi-wavelength imaging in scattered light images is a powerful tool to detect dust settling, especially for edge-on discs (Dullemond & Dominik 2004; Pinte et al. 2007, 2008; Duchêne et al. 2010). As described in Section 2.4.2, dust grains with different sizes settle at different heights above the midplane, i.e. larger grains are expected settle much faster and deeper than the smaller ones (Laibe et al. 2014). Importantly, multi-wavelength images in scattered light confirm the presence of stratification, with smaller grains found higher up in the disc atmosphere while larger grains have settled about the disc midplane (Pinte et al. 2008; Duchêne et al. 2010).

One of the most intriguing characteristic of disc scattering emission is its anisotropy (Min et al. 2016). The scattering phase function is strongly dependent on the grain size: While smaller grains scatter radiation uniformly in all directions, the scattering phase function of large grains (grains beyond the Rayleigh limit: $2\pi a > \lambda$ where a is the grain size and λ is the observing wavelength) is dominated by forward scattering. Therefore, since large grains scatter stellar photons mostly into the disc, the light scattered by large grains disappears into the disc, making the tracing of larger grains in non-edge-on discs challenging (Mulders et al. 2013). Additionally, scattering images of inclined disc might reveal an asymmetry in brightness between the part of the disc tilted toward the observer and the part tilted away (Mishchenko et al. 2000; Kudo et al. 2008). This is due to the forward scattering of starlight by large dust particles in an inclined disc. Thus, a brighter front side is often interpreted as a sign of grain growth (Quanz et al. 2011; Thalmann et al. 2014). Moreover, an additional indicator of the particle size in surface layers of protostellar discs is obtained by analyzing the degree of polarization of the scattering emission. The full and polarized intensity changes due to interaction with dust grains are angular-dependent functions of wavelength, particle size, shape, and material. It can be shown that while particles much smaller than the wavelength of radiation reach a 100% degree of polarization at 90° scattering angle, larger grains ($2\pi a > \lambda$) do not polarize the incoming radiation from the star (Min et al. 2016).

From a technical point of view, NIR imaging of protoplanetary discs can achieve high angular resolution due to the small observing wavelength[1]. This allows to detect potential disc substructure at the surface layer on sub-au scales in nearby protoplanetary discs. Recently, a large number of sources has been observed with the high-resolution near-IR NAOS/CONICA (NACO, Lenzen et al. 2003; Rousset et al. 2003) instrument installed on the Very Large Telescope (VLT) and by the High-Contrast Coronographic Imager for Adaptive Optics (HiCIAO, Tamura 2009) installed on the Subaru telescope (Quanz et al. 2011; Muto et al. 2012; Garufi

[1]The highest angular resolution achievable at wavelength λ for a single-dish telescope of diameter D is $\sim 1.22\lambda/D$. NIR observation can achieve high angular resolution $\lesssim 0.1''$ which corresponds to a scale of 13 au for a source at the distance of Taurus star forming region ($\sim$130 pc).

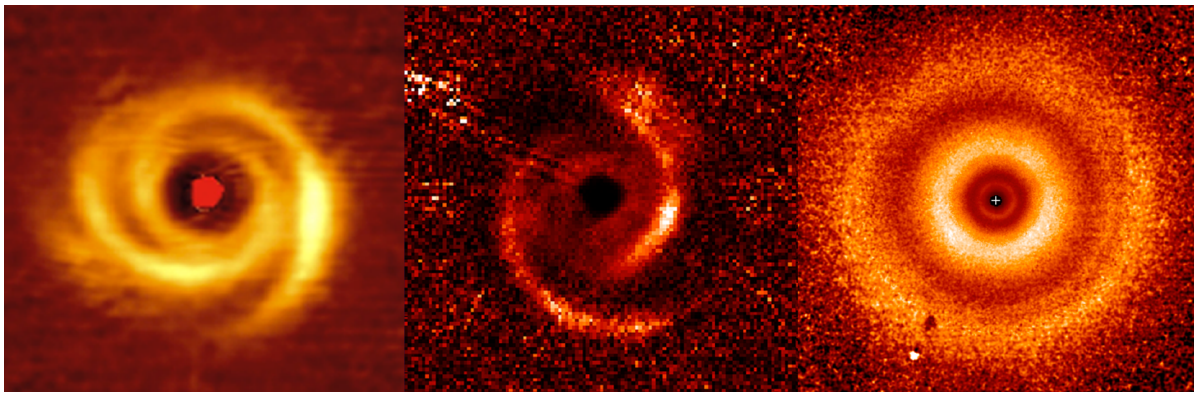

Figure 3.2: Gallery of spiral and gap-like structures detected in NIR high-resolution observations. (top left) polarized scattered light emission of HD 135344B (Garufi et al. 2013), (center) MWC 758 (Benisty et al. 2015) and (right) TW Hydrae (van Boekel et al. 2016). All the images are on different scale.

et al. 2013; Follette et al. 2013; Avenhaus et al. 2014; Akiyama et al. 2015). These observations took advantage of the Polarimetric Differential Imaging technique (PDI, e.g. Perrin et al. 2004) to effectively remove the unpolarized light. Assuming that the signal from the central star is unpolarized, the scattering images retain the polarized component in the scattered light from the dust grains in discs, achieving the high contrast ratios required to detect potential forming protoplanets and probe the circumstellar environment very close to the star (as close as 0.1-0.2" to the central star). With the advent of the Spectro-Polarimetric High-contrast Exoplanet REsearch (SPHERE, Beuzit 2013) installed on VLT and Gemini Planet Imager (GPI, Macintosh et al. 2008; Perrin et al. 2014) aboard the Gemini South Telescope in Chile, we expect to obtain images with a better resolution and a sensitivity never before obtained. New observing facilities will also offer a better direct detection and spectral characterization of giant extra-solar planets due to the improvements in adaptive optics and coronagraphic technique.

Importantly, new observations are revealing smaller scale structures at the disc surface such as spiral arms and gap-like structures that might be induced by embedded protoplanets or by the development of instabilities (Debes et al. 2013; Garufi et al. 2013; Akiyama et al. 2015; Benisty et al. 2015; Ginski et al. 2016; van Boekel et al. 2016). Some examples of disc substructures in scattering light emission are shown in Fig. 3.2. Since the intensity of these images is approximately linked to density and/or scale height variations of well coupled dust grains, it is possible to gain a wide range of insights on the gas density distribution and figure out the physical mechanism behind these substructures. We will focus our analysis on specific cases of disc substructures detected in NIR observations of protoplanetary discs in the next Chapters.

3.2.2 Dust features at mid-infrared wavelengths

Infrared spectroscopy roughly probes dust particles with sizes of up to a few microns in the disc surface layers. This part of the spectra, roughly between 3 and 100 μm, contains several spectral resonances whose features can be related to the shape and composition of dust grains. These resonances, known as solid state features, are locations in the spectrum where the refractive index locally increases. The first direct evidences of the solid state features in mid-infrared spectroscopy of discs were used as a probe of the existence of a warm, optically thin layer of dust over a cold thick midplane. Since discs are usually optically thick at these wavelengths, the presence of *emission* features, i.e. where the opacity locally increases, would suggest that the emission feature comes from an optically thin layer at the disc surface. This has led to the introduction of the so-called two-layer models for the thermal and physical structure of discs (Chiang & Goldreich 1997; Chiang et al. 2001; Dullemond et al. 2001).

Among the several features that can be observed in the disc spectra, the most common one is the 10 μm amorphous silicate features. The morphology of this feature is strongly dependent on

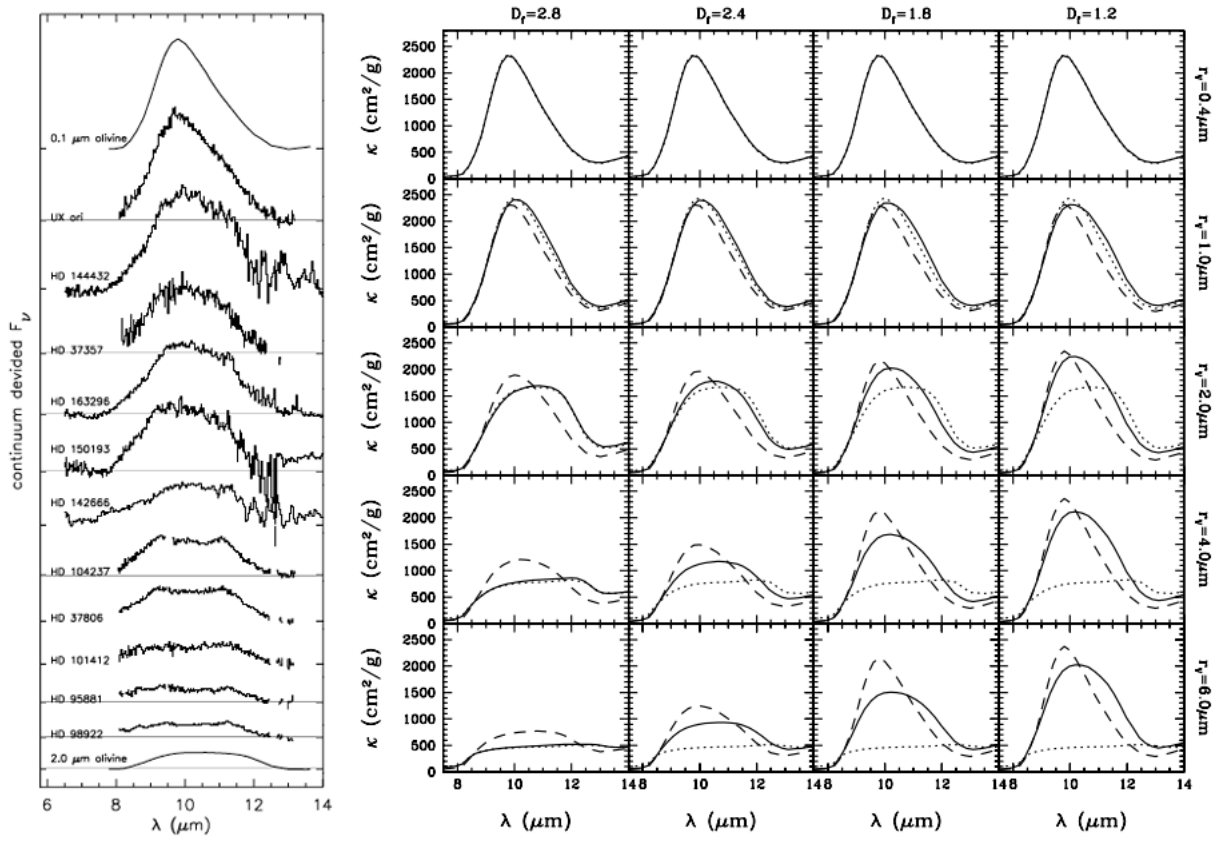

Figure 3.3: (left) Mid-IR spectra of a sample of Herbig Ae/Be stars. The two solid curves on the top and on the bottom panel show the theoretically expected 10 μm feature for olivine grains with size 0.1 μm and 2 μm, respectively. Image taken from van Boekel et al. (2003). (right) 10 μm solid feature of (solid) aggregate particles, (dotted) homogeneous sphere and (dashed) porous sphere for different volume equivalent radiii r_v and fractal dimension D_f. Image taken from Min et al. (2006). These spectra show that the grain size, structural composition and the fluffiness of dust grains strongly affect the shape of the 10 μm solid state feature.

the size and shape of dust grains and has been extensively used as a diagnostic to identify grain growth (Bouwman et al. 2001; Kessler-Silacci et al. 2006; Sicilia-Aguilar et al. 2007). Bouwman et al. (2001) and van Boekel et al. (2003) noted that the shape and strength of the silicate band are correlated: weaker features are also flatter (see the left panel of Fig. 3.3). This result is consistent with the growth of silicate grains from sub-micron to micron sizes, since dust particle growth to radii of several μm leads to flatter silicate emission features (Min et al. 2016). However, this effect might be induced by structural or compositional features of dust grains. Moreover, this region of the SED is crucial to gain insight on the composition and the fluffiness of dust grains (see Min et al. 2006 for details). The right panel of Fig. 3.3 shows that for the same grain size (that for non-spherical particles is usually referred to as volume equivalent radius r_v), the solid feature is stronger for fluffy aggregates (lower fractal dimension D_f). However, the lack of the 10 μm feature can be also ascribed to the absence of micrometer warm dust grains in the disc surface layers from 3-5 to a few of au (the exact value of the distance range of course depends on the luminosity of the central star). Moreover, dust settling might be identified indirectly by analyzing the strength of the silicate features because settling removes the large grains from the disc surface, leading to an enhanced silicate emission features (Kessler-Silacci et al. 2006).

Observations of 10 μm solid feature have provided evidence for silicate processing during the disc phase. This spectral feature has been detected in a number of discs around Herbig Ae-Be stars and T-Tauri stars with the ESA's Infrared Space Observatory (ISO), allowing grain growth to be studied for the first time (Waelkens et al. 1996). The sample of observed discs was extended by including discs around low-mass and Sun-like stars by the Infrared Spectrograph

(IRS) aboard the Spitzer Space Telescope (Hillenbrand et al. 2008; Evans et al. 2009).

3.3 From far-infrared to radio observations

Observations at longer wavelengths are dominated by the thermal emission from large dust grains. The emission is optically thin at almost all disc radii (especially intermediate and outer regions, i.e. $\gtrsim 10$ au) and essentially comes from the dense regions in the midplane. The study of the disc midplane covers a key role in our understanding on the planet formation process, since it is the place where most of the mass of the disc is confined and where planets are expected to form. For this reason, high angular resolution observations at these wavelengths can provide information about the spatial density distribution of dust in discs, which is essential to have a complete understanding of the disc evolution models.

Assuming that the material is optically thin and in local thermodynamical equilibrium, the intensity $I_\nu(R,\theta)$ at frequency ν in cylindrical coordinates can be related to physical quantity of the disc (Rybicki & Lightman 1979)

$$I_\nu(R,\theta) \propto B_\nu\left(1-\exp\left(-\kappa_\nu \Sigma_\mathrm{d}\right)\right) \approx B_\nu\left(T(R,\theta)\right)\kappa_\nu(R,\theta)\,\Sigma_\mathrm{d}(R,\theta)\,, \tag{3.6}$$

where κ_ν is the dust opacity per gram of dust and Σ_d is the dust surface density. Although the size of dust grains in protoplanetary disc is expected to cover a large range of values, it can be shown that observation at wavelength λ essentially trace particles of a similar size $\sim\lambda$ with a maximum size of $\sim 3\lambda$ (Draine 2006).

Generally, due to the intrinsic degeneracy of the optically thin dust thermal emission to the density, temperature and opacity of dust grains, the standard approach to derive disc properties is to compare observations with the predictions of models of flared discs. The most common choice for the disc physical and thermal structure is the two-layer disc model: the disc is essentially modeled as a surface layer (defined as the locus of points in the disc where the optical depth is less than unity) directly heated by the radiation of the central star and a midplane heated by the radiation reprocessed by the surface layer (Chiang & Goldreich 1997; Dullemond et al. 2001). In order to reduce the parameters involved in the fitting procedure, it is assumed that the dust opacity is spatially homogeneous by defining a homogenous chemical and structural composition (e.g., shape, porosity) and a grain size distribution (typically a truncated power-law distribution). The dust surface density profile is typically taken as a power law with an exponential taper at large radii,

$$\Sigma_\mathrm{d}(R) = \Sigma_{\mathrm{d},0}\left(\frac{R}{R_c}\right)^{-\gamma}\exp\left[-\left(\frac{R}{R_c}\right)^{2-\gamma}\right], \tag{3.7}$$

where $\Sigma_{\mathrm{d},0}$ is the normalization parameter, R_c is a characteristic scaling radius and γ is the gradient parameter. This choice for the dust surface density is motivated by assuming that the distribution of large grains is described by a class of solution of the viscous diffusion equation for the gas surface density (see Eq. 1.27) assuming a time-independent viscosity that follows a simple power-law dependence with radius, $\nu \propto R^\gamma$ (see Sect. 1.2.2, Lynden-Bell & Pringle 1974 and Pringle 1981 for details). Although this assumed profile for the dust density structure is physically inaccurate due to the strongly different dynamics between large dust grains and gas, in discs with a smooth radial profile of fluxes (i.e. without gap, ring-like features and any deviation from the axisymmetry) this choice for the dust surface density profile allows to gain a large number of insights on the disc structure. The temperature structure can be self-consistently calculated by adopting the two-layer approximation (Chiang & Goldreich 1997) or through radiative transfer models by assuming that the disc is irradiated by a central star of fixed temperature and luminosity. Then, starting from the temperature structure, a set of millimeter continuum measurements is computed and are simultaneously compared with the observations for each observing wavelengths. The process is iterated until the fit converges on a minimum

value of the χ^2 between the simulated continuum observations and the real observations using a Bayesian approach and adopting Markov chain Monte Carlo methods to explore parameter space in an efficient way. This approach has been extensively used in recent years (e.g. Andrews & Williams 2007a,b; Andrews et al. 2010; Ricci et al. 2010; Pérez et al. 2012). The intrinsic degeneracies introduced in disc quantities evaluated with this approach can be mitigated by including external constraints on the thermal structure and geometry inferred by observation at different wavelengths (e.g., scattering observation at NIR wavelengths). Thanks to advances in numerical modelling, this procedure is being constantly refined by including new features in the fitting procedures in order to break the degeneracy and thus converging to a self-consistent physical description of the disc structure. In this context, Tazzari et al. (2015) derived a new multi-wavelength analysis to constrain self-consistently the disc structure and the radial variation of the dust properties through a simultaneous fit of interferometric (sub-)millimetre observations at several wavelengths (Guidi et al. 2016; Testi et al. 2016).

The shape of the SEDs at larger wavelengths might be used to gain some insights on geometrical properties of the disc. The disc aspect ratio profile heavily affects the area directly irradiated by the central protostar, leading to an enhanced or reduced heating in the outer discs. As expected, flared discs intercept more starlight than their flat counterparts and therefore the dust in the outer disc reaches warmer temperature, leading to a reddening of the SED, i.e. an increase of the dust emission at longer wavelengths (e.g. Kenyon & Hartmann 1987; Williams & Cieza 2011). On the other hand, most T-Tauri stars exhibit less mid-infrared emission than expected for a disc in hydrostatic equilibrium. This can be understood in terms of dust settling which reduces the scale height and flaring angle of the disc (Dullemond & Dominik 2004). Actually, it is very challenging to distinguish between settling and an intrinsically low gas pressure scale height. However, the shape of the dust and gas line emission at far-infrared wavelengths (see Sect. 5.2.3 of Woitke et al. 2016) and multi-wavelengths scattering images (Pinte et al. 2008) can be useful to discern between different mechanisms. Moreover, if small dust grains (which are expected to be well coupled with the gas) at a certain location in the disc are distributed to sufficiently scale heights at the inner edge, they can attenuate the radiation field at larger radii (see Sect. 2.4.2 and Fig. 2.5). This leads to a reduction of the effective surface area directly irradiated by the star in the outer disc and therefore make the disc cooler, lowering the infrared emission and thus affecting the shape of the SED (Kenyon & Hartmann 1987; Dullemond & Dominik 2004; Espaillat et al. 2012). This shadowing effect is expected to be related to the heating of the inner disc edge due to the direct exposure to the stellar radiation or due to the viscous dissipation in cases of highly accreting discs (D'Alessio et al. 1998; Natta et al. 2001; Calvet et al. 2005).

3.3.1 Observation of grain growth

One of the greatest uncertainties in studies of circumstellar discs is related to the estimate of dust opacity. The standard procedure to interpret observations is to assume a specific dust composition, structure and grain size distribution across all the disc. However, it is reasonably expected that the grains responsible for such emission may have different physical and chemical properties at different locations in the disc (Pérez et al. 2015). In the high-density environment of a circumstellar disc, μm dust grains are expected to grow via collisional agglomeration and gravitationally settle toward the midplane of the disc. The evidence that grains in discs are significantly larger than grains in the diffuse interstellar medium (ISM) has been obtained from a variety of observational techniques (Natta et al. 2007; Testi et al. 2014). In this scenario, the dust opacity is obviously influenced by grain processing and therefore its value depends on grain sizes, chemical compositions and shapes which can not be easily probed by observations.

In reference to the grain size, following the analytical approach of Draine (2006), it is convenient to consider a power-law size distribution given by (Mathis et al. 1977)

$$n(a) \propto a^{-q} \qquad \text{for} \quad a_{\min} < a < a_{\max}, \tag{3.8}$$

where $a_{\rm min}$ and $a_{\rm max}$ represent the minimum and maximum size of the grains. By analyzing extinction and scattering of star light from the interstellar dust, Draine & Lee (1984) found that $a_{\rm min}$ is a few tens of Å, $a_{\rm max} \sim 0.1 - 0.2\,\mu m$ and $q = 3.5$ (see also Mathis et al. 1977). In a protoplanetary disc the grain processing generally modifies the typical ISM values (Testi et al. 2001; Natta & Testi 2004; Testi et al. 2014). Assuming that the dust grains are sphere-shaped particles in the Rayleigh limit $a \ll \lambda$ where λ is the radiation wavelength, it is possible to analytically infer the dust opacity law in terms of a power-law function of the wavelength given by

$$\kappa_\lambda = \kappa_0 \left(\frac{\lambda}{\lambda_0}\right)^{-\beta}, \tag{3.9}$$

where κ_0 is a fiducial dust opacity at wavelength λ_0. The parameter β is a very good indicator of the level of grain growth. While the typical value for ISM grains is $\beta_{\rm ism} = 1.7$, several observations (Testi et al. 2001, 2003; Natta & Testi 2004; Ricci et al. 2010) have found that $\beta < \beta_{\rm ism}$, which is naturally interpreted in terms of grain growth (see e.g. Draine 2006). Since $F_\nu \propto \kappa_\nu \nu^2$, using multi-frequency measurements of F_ν with the same observational properties (same beam, pixel size and centered on the peak of the emission) in a range where the disc is optically thin, it is possible to easily deduce the spectral index and, consequently, the β value as

$$\beta_{\nu_1-\nu_2} = \left[\log\left(\frac{\nu_1}{\nu_2}\right)\right]^{-1} \left[\log\left(\frac{F_{\nu_1}}{F_{\nu_2}}\right) - \log\left(\frac{B_{\nu_1}}{B_{\nu_2}}\right)\right], \tag{3.10}$$

where, under the Rayleigh-Jeans approximation, the second term in brackets is equal to 2 (see Eq. 3.3). Therefore, for an optically thin disc in the Rayleigh-Jeans limit, the submillimeter spectral energy distribution (SED) has a power-law behavior with frequency expressed by: $F_\nu \propto \kappa_\nu\, \nu^2 \propto \nu^\alpha$ where α is the spectral index. Thus, from eq. 3.9, the β value is given by: $\beta = \alpha - 2$. Increased grain growth is therefore expected to lead to grayer opacities and more efficient submillimeter emission (i.e., a shallower SED slope).

However, taking into account the possibility that the emission is not completely thin and not in the Rayleigh-Jeans limit, the simple fit of the disc SEDs by a power-law in frequency does not provide a unique interpretation about the level of grain growth (Beckwith et al. 1990). For example, low values of β might be interpreted with the presence of regions of high optical depth, which will drive the dust emission spectrum to be close to $F_\nu \propto \nu^2$ (see Eq. 3.3), leading to an incorrect evaluation of the grain growth processing. However, Ricci et al. (2012) found that this effect would be important only for the brightest and most massive discs.

Fig. 3.4 (right panel) shows the distribution of the power law index β obtained by Ricci et al. (2010) for a large sample circumstellar discs around T-Tauri star in the Taurus-Auriga star formation region. They found a median for β nearly less than 1 that is consistent with model predictions for the collisional growth of solids to cm/mm-size scales in the outer disc. Once the estimate of β has been obtained, one can infer the level of grain growth, i.e. the value of $a_{\rm max}$, simply comparing β with the one computed with appropriate codes according to the adopted dust model. For instance Ricci et al. (2010), by computing the dust opacity values (shown in the left panel of Fig. 3.4 in terms of κ_{1mm} and β_{1-3mm} using Eq. 3.9) as functions of $a_{\rm max}$ for different values of q adopting the dust model proposed by Pollack et al. (1994), found that dust grains with sizes larger than 1 mm are likely to be present in the disc.

However, the milimetre opacity slope might be linked to a different composition rather than an increasing size of dust grains. Since the most important absorber in protoplanetary discs is represented by carbonaceous material, it can be shown that an increased amount of the amorphous carbon in size-fixed dust grains leads to a shallower profile (Min et al. 2016). This can have very important consequences in connecting the spectral slope to the level of grain growth, since, as previously described, a shallow slope is usually interpreted as a sign of grain growth. However, for a sensible set of dust properties the most significant influence on the opacity spectral index arises from an increase in grain size (Draine 2006; Testi et al. 2014).

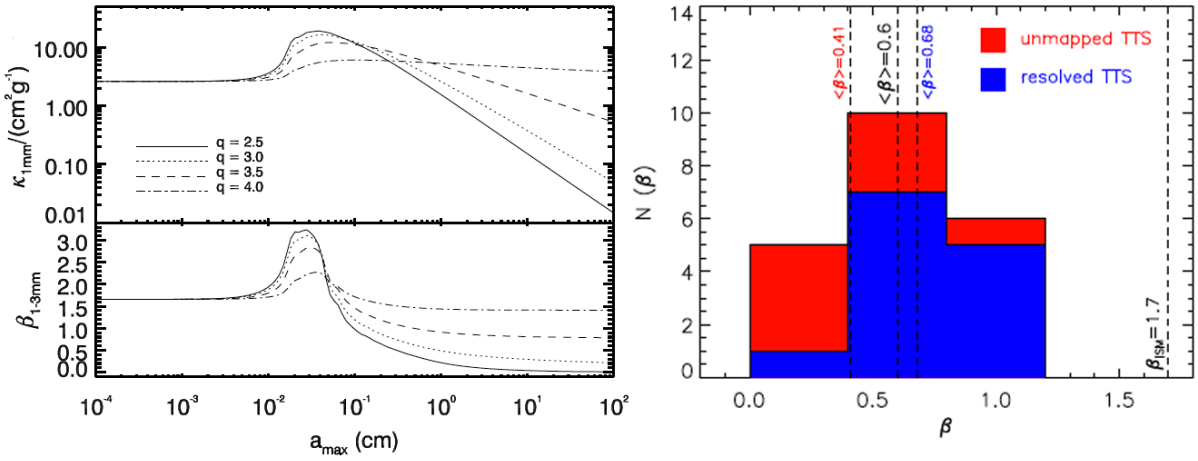

Figure 3.4: (left) Millimeter dust opacity for the adopted dust model. Top panel: dust opacity at 1 mm as function of $a_{\max}$ for a grain size distribution $n(a) \propto a^{-q}$ between $a_{\min} = 0.1\,\mu$m and $a_{\max}$. Different curves are for different values of q, as labelled. The dust grains adopted here are porous composite spheres made of astronomical silicates (10% in volume), carbonaceous materials (20%) and water ice (30%). Bottom panel: β between 1 and 3 mm as a function of $a_{\max}$ for the same grain distributions; (right) Distribution of the dust opacity spectral index inferred by multi wavelengths observations. In blue the values for the spatially resolved discs are indicated, whereas the red is for the unmapped ones. The average value for all the sources in the sample (black), for the unmapped sources only (red), for the resolved sources only (blue). The value of $\beta = 1.7$ for the ISM dust is indicated as dashed vertical lines. Images taken from Ricci et al. 2010.

Moreover, (sub-)millimetre observation in polarized light might also be useful to infer the grain growth level in protoplanetary disc. It is widely believed that the polarization in the dust thermal emission is due to the grain alignment with magnetic field (Girart et al. 2009; Hull et al. 2014). Recently, Kataoka et al. (2015b) proposed that, if a strong asymmetry in the dust continuum emission occurs in a protoplanetary disc, the dust thermal emission might be polarized by the scattering of its own emission on large grains. This mechanism occurs if grains are sufficently large to have a high albedo but sufficiently small to be in the Rayleigh scattering regime ($a \lesssim \lambda/2\pi$) at (sub-)millimetre wavelengths. In this regime, there is a significant amount of scattered light component in dust continuum emission that might be polarized. This method can be used to constrain the maximum grain size in protoplanetary discs by performing observations in polarization at millimeter wavelengths, due to the strong dependence of the maximum grain size on polarization (Kataoka et al. 2015a; Pohl et al. 2016).

3.3.2 Summary of recent observations

Far-infrared and (sub-)millimetre observations has been performed in a large sample of protoplanetary discs using different observing facilities, such as the Submillimeter Array (SMA, e.g. Brown et al. 2009; Andrews & Williams 2007a; Andrews et al. 2011), the Combined Array for Research in Millimeter-wave Astronomy (CARMA, e.g. Isella et al. 2009, 2010a), the Plateau de Bure Interferometer (PBI, e.g. Piétu et al. 2006; Ricci et al. 2010) and ALMA (e.g. Casassus et al. 2013; van der Marel et al. 2013; Pérez et al. 2014; ALMA Partnership et al. 2015; Muto et al. 2015; Andrews et al. 2016; Ansdell et al. 2016). Importantly, recent studies have successfully extended the wavelength range to cm wavelengths using the Very Large Array (VLA, Guidi et al. 2016; Carrasco-Gonzalez et al. 2016). Long-wavelengths observations are performed by means of an observational technique called *interferometry*, that allow to produce, through a complex technique called *aperture synthesis*, high-resolution and high-sensitive observations (see Appendix B for further details).

Some of the recent observations have shown that large dust grains appear to be more concentrated close to the star with respect to the gas phase (e.g. de Gregorio-Monsalvo et al. 2013;

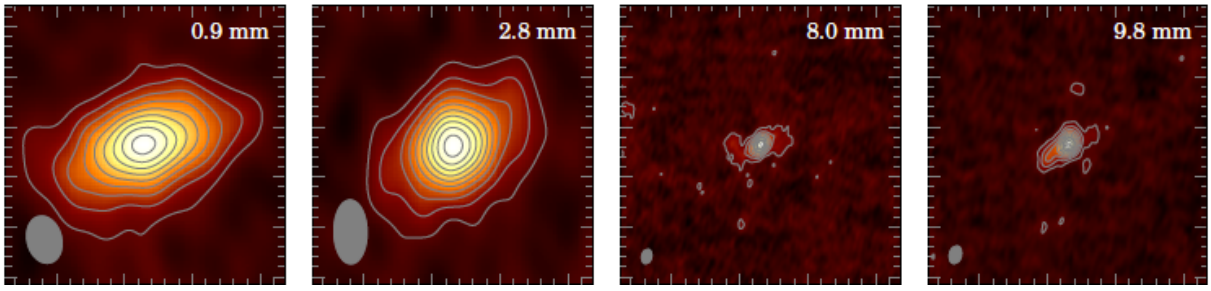

Figure 3.5: Multi-wavelength observations of the continuum emission from DoAr 25, observed at wavelengths (from left to right) of 0.9, 2.8, 8.0, and 9.8 mm. The contours refer to different multiplies of the noise. The filled ellipse in the lower-left corner indicates the size of the half-power contour of the synthesized beam. Image taken from Pérez et al. (2015).

Walsh et al. 2014; Canovas et al. 2016). Resolved multi-wavelengths observations have revealed an anti-correlation between the wavelength and the radial extent of the continuum emission: the disc emission observed at short wavelengths is more extended than the disc emission observed at long wavelengths (see Fig. 3.5, Banzatti et al. 2011; Pérez et al. 2012, 2015; Guidi et al. 2016). While submillimetre dust grains are detected over all the disc extent, larger grains are distributed closer to the central star. These observations show an increasing $\beta(R)$ with increasing radii (see Fig. 3.6). Since larger particles have a lower opacity spectral slope (i.e. lower β, see Sect. 3.3.1), these results can be simply interpreted as a sign of the radial drift of large particles. As shown in Section 2.3.1, since gas and dust particles orbit at different velocities around the central protostar, for standard disc geometries (i.e. negative radial gradient profile of the pressure), dust particles are expected to move towards the inner radii with a timescale that depends on their sizes. The radial drift of dust particles produce a strong concentration of larger grains towards the gas pressure maxima close to the central star, leading to an enhanced grain growth processing (see Sect. 2.5, Birnstiel et al. 2010). These results could be used to infer the level of aerodynamical coupling between gas and dust with different sizes. At the low gas densities in the very outer gas disc, even micron sized particles can be marginally coupled from the gas and start spiralling inwards, until they become strongly coupled at smaller radii and, thus, well mixed from the gas. The drift of small particles can produce a steep radial decrease in the dust-to-gas ratio (see Sect. 2.5.1, Birnstiel & Andrews 2014), as recently observed (e.g. de Gregorio-Monsalvo et al. 2013; Walsh et al. 2014; Canovas et al. 2016). Since the gas density increases toward the star, the size associated to the marginal coupling decreases in the direction toward the star. Additionally, a particle that drifts inwards would have an increasing Stokes number and progressively reduces its radial drift. Importantly, since particles grow while drifting, it is highly expected that larger grains populate the inner disc due to the combined effect of the radial gradient of the radial drift and the grain growth efficiency (which typically increases with decreasing R, i.e. $\tau_{\rm grow} \propto \Omega_{\rm k}^{-1}$, see Sect. 2.5.1, Brauer et al. 2008). Moreover, there are indications of a strong variation in the dust-to-gas ratio and in the efficiency of grain growth processes across the disc extent (see Sect. 2.5.1), which in turn leads to a variation of the optical properties of dust grains emitting at these wavelengths. Importantly, the concentration of large dust grains in the inner disc regions might explain the abundances of rocky exoplanets close to the central star (Chatterjee & Tan 2014).

Additionally, the unprecedented angular resolution and sensitivity of ALMA observations has led to the discovery of several disc substructures in nearby dusty protoplanetary discs in the form of horseshoe, gaps, cavities and spirals (ALMA Partnership et al. 2015; Andrews et al. 2016; Canovas et al. 2016; Pérez et al. 2016, see the recent review of Casassus 2016). The modelling of this substructures by simple disc models described above is simplistic and can lead to serious errors in interpreting observation. Spatial dependent variations of density of grains with different sizes, different level of grain growth and chemical composition in disc substructures need to be taken into account properly in order to constraint disc properties. Fig. 3.7

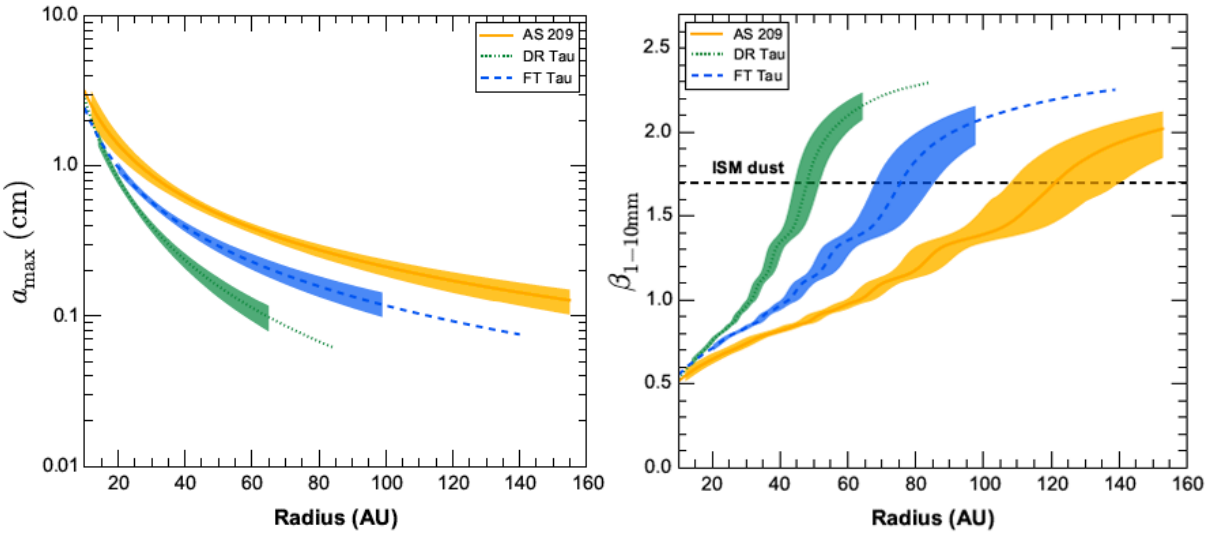

Figure 3.6: (left) radial profile of the maximum dust grain size a_{max} constrained from the multi-wavelength observations of three protoplanetary discs. (right) radial profile of the dust opacity spectral slope $\beta(R)$ between 1 mm and 10 mm. The value of $\beta = 1.7$ for the ISM dust are indicated as dashed vertical lines. The tick lines represent the best-fit models wile the shaded area indicate the 1σ confidence intervals. Image taken from Tazzari et al. (2015).

shows an example of recent detection of disc substructures in protoplanetary discs. A variety of scenarios has been proposed to explain the mechanism that induce ring-like structures in protoplanetary disc, such as interaction with planets (see Part III, Dong et al. 2015c; Picogna & Kley 2015; Rosotti et al. 2016), zonal flows (Ruge et al. 2013; Flock et al. 2015; Béthune et al. 2016; Ruge et al. 2016), rapid pebble growth around condensation fronts (Zhang et al. 2015), self-induced dust pile-ups (Gonzalez et al. 2015), aggregate sintering (Okuzumi et al. 2015), baroclinic instability arising due to dust settling (Lorén-Aguilar & Bate 2016) or secular gravitational instabilties (Takahashi & Inutsuka 2016). Although several gap-carving mechanisms can produce structures potentially detectable over a wide range of wavelengths, the observational features of gaps detected in the most investigated source so far, the HL Tau disc (ALMA Partnership et al. 2015), suggest however that the pattern of gaps and rings might be produced by embedded protoplanets (see Chapter 9, Tamayo et al. 2015; Dong et al. 2015c; Picogna & Kley 2015; Jin et al. 2016). For what concerns the spiral structures recently detected in NIR (see Fig. 3.2) and (sub-)millimetre observations, the most widely accepted explanations for this kind of substructures is the presence of a planet orbiting around the central star (Muto et al. 2012; Dong et al. 2015b) or the development of gravitational instabilities (see Part II, Cossins et al. 2010; Dong et al. 2015a). However, while the planetary origin might explain a limited sample of this substructures and a large range of assumptions needs to be made to interpret these observations, gravitational instabilities might be unlikely due to the relatively old age of the system where the spiral structure has been detected. Specific models to reproduce spiral- and gap-like features will be discussed in the Part II and III, respectively.

3.3.3 Disc mass estimate

The most fundamental quantity to understand the origin and diversity of exoplanetary systems, and thereby the formation of the Solar System, is the disc mass. The bulk constituent of the gas, H_2, is very difficult to observe due to its weak emission. Therefore, in order to measure the mass of a protoplanetary disc, it is necessary to find another constituent of the disc that best traces the gas in the region where is more concentrated, i.e. the midplane. As previously mentioned, the material that makes up the protoplanetary disc is sourced from the interstellar medium. It is therefore assumed to be composed of 1% dust and 99% of gas at birth (Bohlin

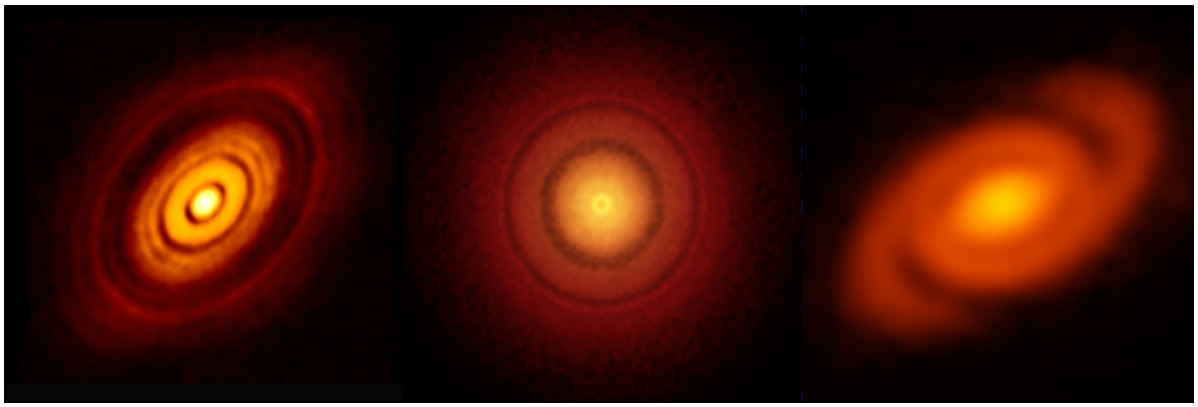

Figure 3.7: Gallery of disc substructures detected at (sub-)millimetre wavelengths. Gap-like features are detected in (left) HL Tau (ALMA Partnership et al. 2015) and (center) TW Hydrae (Andrews et al. 2016). The first spiral structure at (sub-)millimetre wavelengths is observed in Elias 2-27 (Pérez et al. 2016). All the images are on different scale.

et al. 1978). Due to the different dynamical and chemical process of the two phases, the fraction of dust in the mixture is expected to evolve strongly in space and time during disc lifetime. Moreover, as already mentioned dust and gas do not spatially coexist in protoplanetary discs (e.g. de Gregorio-Monsalvo et al. 2013).

Due to the dominant opacity of the dust component and the optically thin emission at (sub-)millimetre wavelengths, the most extensively used method to infer the gas mass is to measure the mass of the dust component. Alternatively, it is possible to estimate the gas mass by detecting molecular line observations of less abundant chemical species, such as HD or carbon monoxide, CO, and its isotopologue. In both cases, the dust-to-gas ratio and the chemical abundances with respect to H_2 need to be well constrained in order to best estimate the gas mass. In the following we briefly summarize the two important methods to infer the gas disc mass from disc observations.

3.3.3.1 From optically thin emission at (sub-)millimetre wavelengths

The most used method to obtain the mass of a circumstellar disc is by the measure of the dust component of the disc through observations in the (sub-)millimetre range of the disc SEDs. Since in the spectral region the emission typically becomes optically thin, the flux can be used as a tracer for the mass of the dust in the disc, which can in turn be used to obtain the total disc mass. For a vertically isothermal disc in thermal equilibrium located at a distance D and inclined with an angle i with respect to the line of sight of the observer, the total flux at frequency ν is expressed by

$$F_\nu = \frac{2\pi \cos i}{D^2} \int_{R_{\rm in}}^{R_{\rm out}} B_\nu(T(R)) \left[1 - \exp\left(-\frac{\tau_\nu}{\cos i}\right)\right] R\, dR, \tag{3.11}$$

where τ_ν is the optical depth at frequency ν evaluated along a given line of sight perpendicular to the plane of the disc,

$$\tau_\nu = \int_{-\infty}^{\infty} \kappa_\nu(z)\rho(z)\, dz, \tag{3.12}$$

where κ_ν is the opacity at frequency ν and dz is a differential element of length along the line of sight through the disc. Given that the optical depth is dominated by the optically thin dust fraction and assuming that the dust opacity κ_ν is invariant with z, the optical depth τ_ν is related to the dust surface density $\Sigma_{\rm d}$ via $\tau_\nu = \Sigma_{\rm d}\, \kappa_\nu$. The gas surface density $\Sigma_{\rm g}$ can be obtained by $\Sigma_{\rm d}$ adopting a value for the dust-to-gas mass ratio ϵ (~ 0.01 for the interstellar medium, Bohlin et al. 1978) as $\Sigma_{\rm g} = \Sigma_{\rm d}/\epsilon$. Since the dust is optically thin in this waveband, it is possible to use

the following approximation,

$$1 - \exp\left(-\frac{\tau_\nu}{\cos i}\right) \approx \frac{\kappa_\nu \epsilon \Sigma_{\rm g}}{\cos i}. \tag{3.13}$$

Since we are interested in the disc emission at long wavelengths, we can use the Rayleigh-Jeans approximation for the Planck's law: $B_\nu = (2\nu^2 k_B T)/c^2$. Thus, Eq. 3.11 can be rewritten as

$$F_\nu = \frac{2\pi}{D^2} \int_{R_{\rm in}}^{R_{\rm out}} \frac{2\nu^2 k_B T(R)}{c^2} \kappa_\nu \epsilon \Sigma_{\rm g}(R) R\, dR. \tag{3.14}$$

Furthermore, let us assume the following relations for the temperature and surface density profiles,

$$T(R) = T_0 \left(\frac{R}{R_0}\right)^{-q}, \qquad \Sigma_{\rm g}(R) = \Sigma_{\rm g,0} \left(\frac{R}{R_0}\right)^{-p}, \tag{3.15}$$

where T_0 and $\Sigma_{\rm g,0}$ are the temperature and the surface density at some given radius R_0. These assumptions have been extensively used in a number of disc models in order to evaluate the disc properties by fitting the expected emission of disc models with observed data (e.g. Andrews & Williams 2007a). Given the power-law of the gas surface density $\Sigma_{\rm g}(R)$, let us write the disc mass as

$$M_{\rm disc} = 2\pi \int_{R_{\rm in}}^{R_{\rm out}} \Sigma_{\rm g} R\, dR = 2\pi \Sigma_{\rm g,0} R_0^2 \int_{x_{\rm in}}^{x_{\rm out}} x^{1-p}\, dx, \tag{3.16}$$

where $x \equiv R/R_0$. Assuming that $R_{\rm in} \ll R_{\rm out}$, the last equation can be rewritten as

$$M_{\rm disc} \approx \frac{2\pi \Sigma_{\rm g,0} R_0^2}{2-p} \left(\frac{R_{\rm out}}{R_0}\right)^{2-p}. \tag{3.17}$$

Integrating Eq. 3.14 using previous equations, it is possible to obtain the flux emitted by the disc as

$$F_\nu = \frac{2\, k_B\, \nu^2\, \kappa_\nu\, \epsilon}{c^2\, D^2}\, M_{\rm disc}\, T(R_{\rm out})\, \frac{2-p}{2-p-q}. \tag{3.18}$$

Thus, using the observed flux in the (sub-)millimetre waveband, it is possible to obtain the mass of the disc.

The value of disc mass inferred by this procedure suffers from large uncertainties. First of all, as previously mentioned, the uncertainty in the value of the dust opacity represent one of the primary sources of error in the disc mass estimate. This occurs because the mass of the dust is directly related to the dust opacity, that in turn is dependent on the particular model considered. Secondly, Eq. 3.18 has been obtained using the approximation of optically thin emission in the Rayleigh-Jeans limit. Generally, the optically thick emission cannot be completely neglected and the Rayleigh-Jeans approximation is not very accurate even when the disc temperature is $T \lesssim 15K$. Therefore, it is convenient to include these effects on the estimate of the disc mass (Beckwith et al. 1990). Importantly, the emission coming from the inner regions, where most of the mass resides, is expected to be optically thick and applying this procedure might underestimate the gas mass. Thirdly, since millimeter observations are insensitive to the emission of bodies much larger than $\sim$ cm due to their low opacities at these wavelengths, an additional major uncertainty in the mass estimate is the hidden mass in large grains. Centimetre and even larger objects can hold considerable mass in a small solid angle with negligible effect on the SED. The dust mass is therefore intended to be the mass in "(sub-)millimetre grains" only, implying that it is a lower limit to the total mass in dust. Finally, in order to convert the dust mass into the total gas mass we need to assume a value for the gas-to-dust ratio. It is worth noting that with gas-to-dust ratio we here indicate the mass ratio between the gas component and the dust responsible of the detected emission at (sub-)millimetre wavelengths (generally millimetre particles, Draine 2006). The standard approach is to assume a value equal to what measured in the ISM ($\epsilon \sim 0.01$, Bohlin et al. 1978) across the whole disc. However, as shown in Section 2.3.1

and observationally confirmed in Sect. 3.3.2, the dust and gas components are characterized by a different dynamics and evolutionary timescales. This ratio is therefore a function of time and space, and using the standard ISM value would be inappropriate to estimate the gas disc mass for evolved protoplanetary disc.

3.3.3.2 From CO isotopologue line observations

The closest chemical counterpart of H_2, HD, was observed by the Hershel Space Observatory in a limited sample of protoplanetary discs: TW Hydra (Bergin et al. 2013), DM Tau and GM Aur (McClure et al. 2016). These results, although rather statistically insignificant, actually suggests that the discs might be more massive than previously thought ($> 0.01\,M_\odot$). However, since Herschel is no longer operating, measuring the disc masses using this technique is no longer possible for other objects.

An alternative method to infer the gas mass independently of the dust content is by means of molecular lines observations. After H_2, CO is the most abundant molecule and, in contrast with H_2 and HD, rotational lines emission at millimeter (mm-)wavelengths are readily detectable with a high signal-to-noise ratio (e.g. Dent et al. 2005). However, due to the optically thick emission of CO lines, it is not possible to measure the gas mass. Fortunately, less abundant CO isotopologues such as ^{12}CO, ^{13}CO and C^{18}O are optically thin at the wavelengths of their line emission and trace the gas down to the midplane of the disc (e.g. Williams & Best 2014).

The main uncertainties on this method concern the H_2-CO abundance and the isotopologue ratios. There are two processes that decrease the isotopologue abundance, ultraviolet (UV) photo-dissociation and freeze-out. Photodissociation is the main process that regulates the abundance of gas phase CO in the emitting layer of discs. One of the key features of this process is that the UV absorption lines can become optically thick and CO can shield itself from photodissociating radiation. The location where CO and its isotopologue become optically thick is different for different isotopologues and, as shown in Miotello et al. (2014b), there could be regions where ^{12}CO is self-shielded and able to survive while less abundant isotopologues are still photodissociated (Visser et al. 2009). This makes photodissociation an isotopologue-selective process (van Dishoeck & Black 1988) and this fact needs to be taken into account to infer the gas mass. In this context, Miotello et al. (2014b) showed that the disc mass may be underestimated by up to two orders of magnitude if only a single CO isotopologue line is observed and isotope selective effects are not properly taken into account. The second effect that depletes the CO and its isotopologues is the condensation from the gas phase and the consequent freezing onto dust grains. This effect is most likely to occur in the cold ($\lesssim 20$ K) outer parts of the disc at the midplane, although the exact value of the freeze-out temperature may vary from disc to disc (e.g. Qi et al. 2015). Due to photo-dissociation and freeze-out processes, these ratios are expected to be a function of space and time, making the CO isotopologues emission line a weak tracer of the gas phase. Therefore, the conversion from CO isotopologue molecular emission to H_2 mass hampered by a large range of uncertainties in the exact values of local density and temperature in gas and dust phase, which strongly affect the chemical reactions (Aikawa et al. 2002; Qi et al. 2015).

The first survey aimed at measuring the gas disc mass through CO isotopologues was recently carried out in the Lupus star forming region by Ansdell et al. (2016). They found that most discs has a gas mass lower than a Jupiter mass, well below the typical gas masses and the disc mass inferred from HD detection of other sources (Bergin et al. 2013; McClure et al. 2016). Importantly, they found, as expected, a gas-to-dust ratios lower than the ISM. The low values of gas masses may indicate that giant planet formation is largely complete by a few Myr. However, these masses might be underestimated due to their assumption of an ISM-like H_2-CO abundances and, generally, due to the ignorance on the photodissociation and freeze-out processes. Alternatively, the different masses inferred by means of HD and CO detections might suggest a gas phase carbon depletion of up to two orders of magnitude.

Importantly, Manara et al. (2016) combines the accretion rates onto the central star deter-

mined from UV excess (using the X-Shooter spectrograph installed on VLT) and disc masses from both sub-mm continuum and CO line emission measured by ALMA (Ansdell et al. 2016) in order to test the predictions of the viscous evolution theory (Pringle 1981; Jones et al. 2012). They found a significant correlation between the mass accretion rates and the disc mass evaluated from the dust mass estimate inferred by (sub-)millimetre emission. This result suggests that the millimetre-sized dust mass better traces the total disc mass and that masses derived from CO might be underestimated.

Part II

Spirals in Gravitationally Unstable Protostellar Discs

Chapter 4

Dynamics of Self-Gravitating Accretion Discs

The accretion disc theory described so far has been focused on non-self gravitating discs. However, massive discs where self-gravity is non negligible are found in a number of different context in astrophysics, from the relatively small scale of Saturn's rings to the pc scale of discs around active galactic nuclei up to kpc scales of the disc of spiral galaxies. One of the most important area where the development of gravitational instabilities affects the evolution of a disc-like geometry system is represented by the process of star and planet formation.

In the early stages of star formation, the disc could be massive enough to have a non negligible dynamical effect on the evolution of the overall system. In this context, the disc self-gravity may affect the disc dynamics through the propagation of density waves which lead to the formation of a prominent structure in the form of a spiral. These density waves are thought to provide a non negligible contribution to the angular momentum transport and may have a crucial role in the genesis of gas giant planets through the direct fragmentation of spiral overdensities into bound objects. Moreover, if circumstances allow, a gravitationally unstable disc might fragment and produce bound clumps, offering a new scenario to form massive gaseous planets.

In this Chapter we will review the basic physical processes that determine the formation and evolution of a self-gravitating gaseous accretion disc. We will also summarize the disc structure as it result from the influence of the disc self-gravity and present a brief overview of recent detections of disc substructure which might be linked to the these instabilities. The modern theory of gravitationally unstable discs was first described by the seminal papers of Paczynski (1978) and Lin & Pringle (1987) and recently reviewed by Lodato (2007) and Kratter & Lodato (2016), to which we refer for more details.

4.1 Self-gravity influence on the disc structure

As in the case of disc formation as a result of the gravitational collapse of a nebular cloud, the primary result of the disc self-gravity is the formation of overdense clumps in protoplanetary discs. The (long-range) interactions between gas fluids in different parts of the disc make the disc susceptible to the development of non-axisymmetric perturbations, leading to the production of torques that redistribute angular momentum across the disc. Before moving onto details about the angular momentum transport induced by these perturbations, let us focus on the general effect induced by self-gravity: collapse. A simple analysis on the collapse of a self-gravitating disc-like structure provides useful insights on the conditions under which self-gravity wins against stabilizing factors. The process of collapse is counteracted by the shear motion on large scales and by local pressure effects on small scales. Let us consider a forming collapsing clump of scale ΔR. A simple analysis on the outcome of this process is based on timescale arguments: if the free-fall time of the self-gravitating clump,

$$t_{\rm ff} \sim \sqrt{\frac{\Delta R}{\pi \mathcal{G} \Sigma_{\rm g}}}, \tag{4.1}$$

is short compared to the timescale on which sound waves crosses the clump,

$$t_{\rm p} = \frac{\Delta R}{c_{\rm s}}, \tag{4.2}$$

and to the timescale on which radial shear can destroy it (by separating neighboring fluid elements),

$$t_{\rm sh} = \frac{1}{R}\left(\frac{\partial \Omega}{\partial R}\right)^{-1} \sim \Omega^{-1}, \tag{4.3}$$

the clump collapses under its own self-gravity. Setting the three time-scale equal, it is possible to infer the condition for the disc to be marginally unstable to clump formation. This is given by

$$\mathcal{Q} \equiv \frac{c_{\rm s}\Omega}{\pi \mathcal{G}\Sigma_{\rm g}} \approx 1. \tag{4.4}$$

The parameter $\mathcal{Q}$ (Toomre 1964), controls the local axisymmetric stability of a gaseous infinitesimally thin disc with respect to gravitationally instabilities (see Sect. 4.2.1 for the analytical derivation). If $\mathcal{Q} \lesssim 1$, the stabilizing terms are not sufficient enough to prevent collapse and the disc will be gravitationally unstable. A first look at this condition suggests that a massive and/or a cold disc will be unstable. This condition highlights the crucial roles played by the gas surface density, the temperature and the local orbital timescale.

It is convenient to rearrange terms in order to get more insight on the regions of the disc mostly prone to be unstable. Let us consider a Keplerian ($\Omega = \Omega_{\rm k}$) vertically isothermal disc at hydrostatic vertical equilibrium, i.e. $H_{\rm g} = c_{\rm s}/\Omega_{\rm k}$. Assuming a power-law profiles for the mass surface density and the sound speed, $\Sigma_{\rm g} \propto R^{-p}$ and $c_{\rm s} \propto R^{-q/2}$. The parameter $\mathcal{Q}$ scales with radius R as

$$\mathcal{Q} \propto R^{p-(q+3)/2}. \tag{4.5}$$

If $p < (q+3)/2$, the right hand side of Eq. 4.5 decreases with radius. Therefore, the condition $\mathcal{Q} \lesssim 1$ is more likely to be satisfied in the outer regions of a disc. Otherwise, if $p > (q+3)/2$, the disc will be more likely gravitationally unstable in the inner disc regions. Assuming a typical power-law exponent for the temperature, $q = 1/2$, the condition is likely to be satisfied in the outer (inner) disc regions for $p < 7/4$ ($p > 7/4$). Since the typical power-law exponents inferred from disc observations are in the range $[0, 2]$ (e.g. Andrews & Williams 2007a), the condition $\mathcal{Q} \lesssim 1$ is more likely to be satisfied in the outer regions of a disc.

4.1.1 Vertical hydrostatic equilibrium in self-gravitating discs

Let us consider the role of self-gravity for the vertical hydrostatic equilibrium. To simplify the calculation, let us assume a radially homogenous self-gravitating disc (i.e., $\rho = \rho(z)$), and infinite in extent and the central point mass is negligible. In this case the hydrostatic equilibrium condition for a vertically isothermal disc is given by (see Eq. 1.7)

$$\frac{c_{\rm s}^2}{\rho_{\rm g}}\frac{\partial \rho_{\rm g}}{\partial z} = -2\pi \mathcal{G}\Sigma_{\rm g}(z), \tag{4.6}$$

where $\Sigma_{\rm g}(z) = \int_{-z}^{z} \rho_{\rm g}(z')\, dz'$. In the case of vertically isothermal disc, it is possible to analytically derive the vertical density profile, which is given by (Spitzer 1942)

$$\rho_{\rm g}(z) = \rho_{\rm g,0}\, \frac{1}{\cosh^2(z/H_{\rm sg})}, \tag{4.7}$$

where $H_{\rm sg}$ is the vertical scale height in the fully self-gravitating case

$$H_{\rm sg} = \frac{c_{\rm s}^2}{\pi \mathcal{G}\Sigma_{\rm g}}. \tag{4.8}$$

In addition, it is also possible to solve the vertical hydrostatic balance equation accounting for both the central object and the disc contributions to the gravitational field (Bertin & Lodato 1999; Lodato 2007).

The analysis presented above has introduced a new different scale height of circumstellar disc related to the case of fully self-gravitating disc. It is instructive to compare this scale height to the one related to the case when gravity is dominated by the central object, i.e. $H_{\rm nsg} = c_{\rm s}/\Omega_{\rm k}$ (see Eq. 1.11), and figure out how massive has the disc to be in order to not neglect the effect of the disc self-gravity on the vertical disc structure. To this end, let us first analyze the two contributions to the vertical component of the gravitational field. For the self-gravitating disc, the term $2\pi\mathcal{G}\Sigma_{\rm g}$ is of the order of $\sim \mathcal{G}M_{\rm disc}/R^2$ where $M_{\rm disc}$ is the disc mass contained with the radius R. For the non self-gravitating case, the vertical component of the star's gravitational field is given by $\sim \mathcal{G}M_\star/R^2 \cdot H_{\rm nsg}/R$. Therefore, comparing the last expressions, the transition from the non self-gravitating to self-gravitating case occurs when the disc mass is of the order of

$$\frac{M_{\rm disc}}{M_\star} \approx \frac{H_{\rm nsg}}{R} \ll 1. \tag{4.9}$$

It should be noticed that this condition is equivalent to the one expressed in Eq. 4.4, by approximating the surface mass density as $\Sigma_{\rm g} \approx M_{\rm disc}/(\pi R^2)$. Thus, according to thin-disc approximation, it is expected that the role of self-gravity in affecting the vertical structure of a circumstellar disc should not be neglected even when the disc mass is much smaller than the central object mass.

In addition, it is possible to infer the condition that marks the transition between the two regimes, by equating the two height scales

$$\frac{H_{\rm sg}}{H_{\rm nsg}} = \frac{c_{\rm s}\Omega_{\rm k}}{\pi\mathcal{G}\Sigma_{\rm g}}. \tag{4.10}$$

It can be noted that, if the right-hand side of the Eq. 4.10 is of order unity, it is instructive to recast the equation in the following form

$$1 \approx \frac{c_{\rm s}\Omega_{\rm k}}{\pi\mathcal{G}\Sigma_{\rm g}} = \frac{c_{\rm s}}{\Omega_{\rm k}R}\,\frac{\mathcal{G}M_\star}{\pi\mathcal{G}\Sigma_{\rm g}R^2} \approx \frac{H}{R}\frac{M_\star}{M_{\rm disc}}, \tag{4.11}$$

where $H \approx H_{\rm nsg} \approx H_{\rm sg}$. Thus, the condition $H_{\rm nsg} \approx H_{\rm sg}$ is equivalent to $M_{\rm disc}/M_\star \approx H_{\rm nsg}/R$ which is the same condition obtained by the previous analysis (see Eq. 4.9) and, in turn, is equivalent to the condition expressed in Eq. 4.4. This implies that whenever self-gravity is important in the vertical direction, the disc is close to the onset of gravitational instability.

4.1.2 Centrifugal balance in self-gravitating discs

To calculate the gas orbital velocity in a self-gravitating disc we need to take into account the gravitational field of the disc. In a self-gravitating disc, the gravitational potential associated to the disc self-gravity depends on the matter distribution through Poisson's equation

$$\nabla^2\Phi_{\rm sg} = 4\pi\mathcal{G}\Sigma_{\rm g}\,\delta(z), \tag{4.12}$$

where $\delta(z)$ is the Dirac δ-function which is introduced as a consequence of the thin-disc condition to indicate that the disc is confined to the mid-plane. The relation between the local surface density and the disc gravitational field is rather complicated (Bertin & Lodato 1999; Lodato 2007). The relation is strongly simplified if the surface density is in the form $\Sigma_{\rm g} = \Sigma_{\rm g,0}R_0/R$. In these discs, known as Mestel discs (Mestel 1963), the gravitational field is given by

$$\frac{\partial\Phi_{\rm sg}}{\partial R} = 2\pi\mathcal{G}\Sigma_{\rm g}. \tag{4.13}$$

Therefore, in a fully self-gravitating Mestel discs, the rotation curve is constant,

$$v^2_{{\rm g},\phi} = R\frac{\partial\Phi_{\rm sg}}{\partial R} = 2\pi\mathcal{G}\Sigma_{\rm g,0}R_0. \tag{4.14}$$

Further details about the derivation of the rotation curve in the case of dominant disc self-gravity can be found in Bertin & Lodato (1999) and Lodato (2007). In addition, for a complete description of gravitational potentials for a range of disc configurations see Binney & Tremaine (1987).

Generally, it is expected that the rotation curve may be affected by self-gravity when the disc mass becomes comparable to the central object mass $M_\star$. Thus, the Keplerian approximation is generally appropriate for protostellar disc because the disc mass is generally much lower than that of the central object, even thought the disc might be vertically self-gravitating and subject to gravitational instability.

4.2 The onset of gravitational instability

The study of fluid instabilities due to gravitational forces dates back to the seminal work of Jeans (1902), who established that for an infinite, static and homogeneous medium at a fixed temperature, there is a critical mass above which gravity overcomes the thermal pressure support, and the cloud becomes unstable to small perturbations. In the case of a rotating thin disc, the stability analysis to small perturbations gets more complicated, due to the different geometry and the additional stabilizing effect: the shear. While in the Jeans stability criteria, the stabilizing effect was related to pressure forces, mostly effective at small scales, here the disc rotation provides another stabilizing effect, which is more effective at large scales.

In this section we analyze the response of an infinitesimally thin disc under the action of small perturbations. To this end, it is possible to carry out standard linear perturbation local analysis of the relevant dynamical equations previously described (see Sect. 1.2) with the equation of state (which is obtained, in the case of barotropic gas, by the definition of sound speed: $c_\mathrm{s}^2 = \partial P/\partial \rho_\mathrm{g}$) and the Poisson equation (Eq. 4.12). Obviously, we also assume that the density perturbations imposed on the unperturbed disc are small, so we can analyze the dynamics using linear perturbation theory.

The propagating density waves are characterised by a radial wavenumber k and the azimuthal one m. In addition, it is possible to define the opening angle i (or pitch angle) of the density waves by the following expression

$$\tan i = \frac{m}{|kR|}. \tag{4.15}$$

This expression allows us to define the idea of *tightly wound perturbations*, which have the property that the pitch angle: $m/|kR| \ll 1$. In light of the local analysis adopted here, we can infer the features of propagating radial and azimuthal waves by adopting the tight-winding, or Wentzel - Kramers - Brillouin (WKB), approximation in order to simplify the analysis and obtain the dispersion relation.

4.2.1 Dispersion relation

To simplify our analysis, we consider the case of axisymmetric disturbance ($m = 0$) and we assume an underlying, unperturbed solution to the Euler and continuity equations. It should be noted that axisymmetric waves ($m = 0$) have zero pitch angle and hence they always satisfy the WKB approximation. However, the analysis shown below is valid for axisymmetric waves only if they also have short radial wavelengths, i.e. $|kR| \gg 1$.

To further simplify our analysis, we perform the linear perturbation analysis in a small box around a point (R_0, ϕ_0) with angular velocity $\mathbf{\Omega}|_{R_0} = \Omega \mathbf{e_z}$ where $\mathbf{e_z}$ is the unit vector identifying the vertical direction. A set of local cartesian coordinates is attached to the frame, consisting of two axis x and y directed along the local radial and the azimuthal directions, respectively. We denote the velocity in the frame of the box $\mathbf{v} = v_x \mathbf{e_x} + v_y \mathbf{e_y}$. The equations for the mass and

the momentum conservation in this frame are given by

$$\frac{\partial \Sigma_{\rm g}}{\partial t} + \nabla \cdot (\Sigma_{\rm g} \mathbf{v}) = 0 \tag{4.16}$$

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{v} = -\mathbf{x}\left[R\frac{\partial \Omega^2}{\partial R}\right]_{R_0} - \frac{\nabla P}{\Sigma_{\rm g}} - \nabla\Phi - 2\mathbf{\Omega} \times \mathbf{v}. \tag{4.17}$$

The first term of the right hand side of Eq. 4.17 is the centrifugal force, the second term is the contribution from the pressure forces, the third is the gravity from the central star and the third term is the Coriolis force. These equations are called the shearing box equations of a rotating thin disc. To obtain the stationary solutions, we set $\partial/\partial t = 0$ and neglect the pressure forces, which produce a small perturbation ($\sim (H/R)^2$, see Eq. 1.15) to the gas velocity in steady state. The velocities are given by

$$v_{x,\rm eq} = 0, \tag{4.18}$$

$$v_{y,\rm eq} = \frac{1}{2}\left[\frac{R}{\Omega}\frac{\partial \Omega^2}{\partial R}\right]_{R_0} x. \tag{4.19}$$

We need also to evaluate a density distribution for the initial hydrostatic configuration that satisfy Eqs. 4.12, 4.16 and 4.17. As observed in the classical problem of Jean instabilies (Jeans 1902; Binney & Tremaine 1987; Clarke & Carswell 2007), all the equation cannot be satisfied simultaneously. We therefore adopt the usual approach by assuming the existence of a uniform surface density distribution at equilibrium $\Sigma_{\rm g,eq}$. Having now established a framework for our analysis, we can now slightly perturb the equilibrium, i.e.

$$\Sigma_{\rm g} = \Sigma_{\rm g,eq} + \Sigma'_{\rm g}, \tag{4.20}$$

$$v_x = 0 + v'_x, \tag{4.21}$$

$$v_y = v_{y,\rm eq} + v'_y, \tag{4.22}$$

$$P = P_{\rm eq} + P', \tag{4.23}$$

$$\Phi = \Phi_{\rm eq} + \Phi'. \tag{4.24}$$

Substituting in the fluid equations (Eqs. 4.12, 4.16 and 4.17), and retaining only first order terms in the perturbed quantities, we obtain, assuming a barotropic gas, $P' = c_{\rm s}^2 \Sigma'_{\rm g}$,

$$\frac{\partial \Sigma'}{\partial t} + \Sigma_{\rm g,0}\frac{\partial v'_x}{\partial x} = 0, \tag{4.25}$$

$$\frac{\partial v'_x}{\partial t} = 2\Omega v'_y - \frac{c_{\rm s}^2}{\Sigma_{\rm g,eq}}\frac{\partial \Sigma'}{\partial x} - \frac{\partial \Phi'}{\partial x}, \tag{4.26}$$

$$\frac{\partial v'_y}{\partial t} + \frac{1}{2}\left[\frac{R}{\Omega}\frac{\partial \Omega^2}{\partial R}\right]_{R_0} v'_x = -2\Omega v'_x, \tag{4.27}$$

$$\nabla^2 \Phi' = 4\pi \mathcal{G} \Sigma'_{\rm g} \delta(z). \tag{4.28}$$

Let us assume monochromatic perturbations for all the fluid quantities of the form

$$\Sigma'_{\rm g} = \Sigma'_{\rm g,0} \exp\left[i\left(kx - \omega t\right)\right], \tag{4.29}$$

where ω is the angular frequency and k is the wavenumber of the perturbation. For simplicity and with no loss of generality, we have chosen the x axis to be parallel to k, so $\mathbf{k} = k\mathbf{e_x}$. This assumptions for the solutions for the perturbed quantities implies that

$$\frac{\partial}{\partial t} \to -i\omega, \qquad \frac{\partial}{\partial x} \to ik. \tag{4.30}$$

Let us now focus on the Poisson's equation. It can be shown that the only solution to Eq. 4.28 is given by

$$\Phi' = \Phi'_0 \exp\left[i\left(kx - \omega t\right) - |kz|\right], \tag{4.31}$$

that, as required, approaches 0 at large z. In order to relate Φ' with $\Sigma'_{\rm g}$ we compute the integral of Eq. 4.28 around the plane $z = 0$ from ϵ to $-\epsilon$ where ϵ is a positive constant, and evaluate its limit for $\epsilon \to 0$. Since the components of the Laplacian on the plane, i.e. $\partial^2\Phi'/\partial x^2$ and $\partial^2\Phi'/\partial y^2$, are continuous and even function around $z = 0$, we obtain

$$\lim_{\epsilon\to 0}\int_{-\epsilon}^{\epsilon}\nabla^2\Phi' dz' = \lim_{\epsilon\to 0}\left.\frac{\partial\Phi'}{\partial z'}\right|_{-\epsilon}^{\epsilon} = -2\,|k|\,\Phi'_0\exp\left[i\,(kx-\omega t)\right]. \tag{4.32}$$

The right hand side of Eq. 4.28 is given by

$$4\pi\mathcal{G}\lim_{\epsilon\to 0}\int_{-\epsilon}^{\epsilon}\Sigma'_{\rm g}\delta(z')dz' = 4\pi\mathcal{G}\Sigma'_{\rm g,0}\exp\left[i\,(kx-\omega t)\right]. \tag{4.33}$$

Thus, the relation between Φ'_0 and $\Sigma'_{\rm g,0}$ is

$$\Phi'_0 = -\frac{2\pi\mathcal{G}\Sigma'_{\rm g,0}}{|k|}. \tag{4.34}$$

Assuming solutions with the form expressed in Eqs. 4.29 and putting these into Eqs. 4.25-4.28 and using 4.30 and 4.34, we obtain the following linear system of three equations in $\Sigma'_{\rm g,0}$, $v'_{x,0}$ and $v'_{y,0}$

$$\begin{pmatrix} -i\omega & -ik\Sigma_{\rm g,eq} & 0 \\ \frac{ikc_{\rm s}^2}{\Sigma_{\rm g,eq}} - \frac{i2\pi\mathcal{G}k}{|k|} & -i\omega & -2\Omega \\ 0 & 2\Omega + \frac{1}{2}\left[\frac{R}{\Omega}\frac{\partial\Omega^2}{\partial R}\right]_{R_0} & -i\omega \end{pmatrix}\begin{pmatrix}\Sigma'_{\rm g,0}\\ v'_{x,0}\\ v'_{y,0}\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix}. \tag{4.35}$$

This linear system admits a non zero solution if and only if the determinant of the matrix on the left hand side of Eq. 4.35 is null. This condition provides

$$\omega^2 = c_{\rm s}^2\,k^2 - 2\pi\mathcal{G}\Sigma_{\rm g}|k| + \kappa^2, \tag{4.36}$$

where κ is the epicyclic frequency defined as

$$\kappa^2 = \frac{2\Omega}{R}\frac{\partial\left(\Omega R^2\right)}{\partial R}, \tag{4.37}$$

which represents the oscillation frequency of a particle in a pressureless shearing non self-gravitating flow subject to a small radial displacement from equilibrium. For non axisymmetric disturbances ($m \neq 0$), it can be shown that the general form of the dispersion relation for tightly wound axisymmetric disturbances, is given by (Lin & Shu 1964; Binney & Tremaine 1987; Bertin 2014):

$$(\omega - m\Omega)^2 = c_{\rm s}^2\,k^2 - 2\pi\mathcal{G}\Sigma_{\rm g}|k| + \kappa^2. \tag{4.38}$$

This relation is very similar to the dispersion relation obtained for the non-rotating, uniform, infinite and infinitely thin sheet of matter (e.g. Binney & Tremaine 1987). The only difference is in the term involving the epicyclic frequency which is obviously related to rotation.

4.2.2 Stability condition

Let us now obtain the stability condition by analyzing the dispersion relation. It is well known that, whenever ω^2 is positive, a given perturbation simply propagates in a wave-like way. Thus, as can be easily seen in the right-hand side of Eq. 4.38, the first term, corresponding to thermal pressure force, and the third, related to rotation, are stabilizing terms. The pressure term $\omega^2 = c_{\rm s}^2k^2$ is the standard dispersion relation for acoustic waves in a non-shearing non-self-gravitating fluid. The pressure waves tends to stabilize the disc and restore the pressure balance on a timescale given by $c_{\rm s}^{-1}l$, where l is the length scale of the perturbation ($\sim k^{-1}$). Pressure is therefore more efficient to restore equilibrium over short lengths (i.e. large k). The

differential rotation term $\omega^2 = \kappa^2$ has a stabilising effect for small perturbations if $\kappa^2 > 0$, i.e. $\partial\left(\Omega R^2\right)/\partial R > 0$ in non-self-gravitating pressureless flows. In this case, the particles oscillate around equilibrium position (this motion is mostly known as *epicyclic oscillation*, see Sect. 7.1.2) and restore their unperturbed state on a timescale of the order of κ^{-1}, regardless the size of the perturbations (ω does not depend on k). Thus, rotation is less efficient than pressure to stabilise perturbations over short scales but more efficient to stabilise large scale perturbations. There exist a scale $l_{\rm crit}$ over which the shear restore the equilibrium on lower timescale compared to pressure forces, i.e. $l_{\rm crit} = c_{\rm s}\kappa^{-1}$. As an example, in Keplerian discs, i.e. $\kappa = \Omega_{\rm k}$, the critical lenght scale is $c_{\rm s}\Omega_{\rm k}^{-1} = H_{\rm nsg}$, i.e. the disc's scale height (see Sect. 1.2.1.1).

Importantly, the instability arises when ω^2 is negative. The negative value of ω^2 implies that ω is imaginary. Thus, in this case, there are perturbations characterized by an exponentially decaying and an exponentially increasing amplitude. The latter solution corresponds to the development of gravitational instability in the form of a spiral structure. As expected, the destabilizing term in the dispersion relation is that related to the self-gravity: $-2\pi\mathcal{G}\Sigma_{\rm g}|k|$.

To simplify the evaluation of the stability condition, let us consider only axisymmetric disturbance ($m = 0$). In this case, since the resulting dispersion relation is a simple quadratic expression for the wave number k (Eq. 4.36), it can be easily seen that ω^2 is positive if

$$\mathcal{Q} = \frac{c_{\rm s}\kappa}{\pi\mathcal{G}\Sigma_{\rm g}} > 1. \tag{4.39}$$

The parameter $\mathcal{Q}$, firstly introduced in the simple timescale analysis (see Eq. 4.4), is therefore essential in determining the local axisymmetric susceptibility of an infinitesimally thin disc to gravitational instabilities. As can be clearly seen in Eq. 4.39, this parameter encapsulates the balance between the stabilising effects of rotation and of the local thermal pressure, characterized by $\kappa(R)$ and $c_{\rm s}(R)$ respectively, and the destabilising effect of the disc self-gravity expressed via the surface density $\Sigma_{\rm g}(R)$.

For the case of Keplerian disc ($M_{\rm disc} \ll M$), the epicyclic frequency is approximately equal to the Keplerian angular velocity $\Omega_{\rm k}$. As expressed in Eq. 4.11, the condition $\mathcal{Q} \sim 1$ for Keplerian discs is roughly equivalent to $M_{\rm disc}/M_\star \approx H_{\rm nsg}/R$ which indicates, as previously mentioned, that even a relatively light thin disc can be gravitationally unstable.

For gravitationally unstable discs, it is possible to evaluate the most unstable radial wavelength λ_0 for axisymmetric perturbations. It can be obtained by calculating the wave number k_0 for which the right-hand side of Eq. 4.38 attains a minimum. It is given by

$$\lambda_0 = \frac{2\pi}{k_0} = \frac{2\pi c_{\rm s}^2}{\pi\mathcal{G}\Sigma_{\rm g}} = 2\pi H_{\rm sg}, \tag{4.40}$$

which is of the order of the disc thickness scale $H_{\rm sg}$.

Furthermore, numerical simulations of self-gravitating discs have shown that the stability criterion remains fairly valid also for non-axisymmetric disturbance (e.g. Lodato & Rice 2004, 2005). However, even discs that are locally stable according to the stability criterion (Eq. 4.39), might generate large scale perturbations. This case is not included in this treatment due to the fact that such global modes are not captured by the WKB tightly wound approximation. However, it is possible to generalize the analysis by including the description of large scale perturbations in order to infer the global dispersion relation (Lau & Bertin 1978; Kratter & Lodato 2016).

4.2.3 Self-regulation

Let us turn to analyze in more detail the stability parameter $\mathcal{Q}$. As can be seen by Eq. 4.39, since the parameter is proportional to the sound speed, it is naturally expected that colder discs are more gravitationally unstable. Therefore, a complete description of the evolution of gravitational instabilities has to take into account disc properties involving the temperature and, especially, the cooling rate.

Let us consider a disc initially hot enough to ensure that $\mathcal{Q} > 1$ at all radii. Assuming that there are no transport and heating mechanisms, it is expected that the disc cools down due to the radiative cooling. This produces a decrease in the value of $\mathcal{Q}$ until it becomes of the order of unity. At this point, the disc is cold enough to develop gravitational instabilities which leads to the propagation of spiral density waves within the disc providing angular momentum and energy transport. In addition, the propagation of density waves produces heating mechanisms such as compressions and shocks which reheat the disc. As a consequence, the value of $\mathcal{Q}$ increases. Therefore, it is expected that the disc reaches the so-called *marginal stability*, where $\mathcal{Q} \sim \bar{\mathcal{Q}} \approx 1$. The disc is maintained close to the marginal stability as a consequence of the heating (induced by gravitational instabilities) that powers up if the disc is colder than a given temperature. Therefore, the stability condition works as a kind of "thermostat" which forces the instability to self-regulate and the disc to reach a thermal equilibrium state where cooling is balanced by heating generated through gravitational instabilities. This peculiar behaviour of gravitational instability has been extensively probed in a number of numerical simulations (Lodato & Rice 2004; Vorobyov & Basu 2005, 2006; Cossins et al. 2009). Numerical simulations of cooling self-gravitating discs (one of these will be analyzed in Chapter 5) have shown that, after an initial transient, the disc settles down in a quasi-steady state characterized by the presence of a spiral structure.

4.2.4 Transport induced by gravitational instabilities

In recent years, one of the well-studied aspect of the development of gravitational instabilities in discs has been the locality of the transport or, in other words, whereas the gravitational instability can be described as a pseudo viscous process (e.g. Lodato & Rice 2004, 2005; Clarke 2009; Cossins et al. 2009; Forgan et al. 2011).

The introduction of gravitational instability as a source of angular momentum and energy transport in circumstellar discs was usually linked to the investigation of the so-called "anomalous" viscosity whose magnitude has to be sufficient to ensure a reasonable viscous timescale for the circumstellar disc evolution (see Sect. 1.3.1). The development of non-axisymmetric modes induced by self-gravity produce torques that redistribute angular momentum within the disc. Analogously to the viscous torque, the work done by the stress provided by these "gravitoturbulent" fluctuations is dissipated in the disc, providing a net conversion of mechanical energy into heat which permits the driving of an accretion flow. In light of this effect of the gravitational instability onto the angular momentum and energetic contents of the disc, the action of self-gravity is analogous to that of viscosity.

However, it is expected that a gravitationally unstable disc cannot be treated as a viscous disc because gravity can produce transport over long distances. The main problem is to verify if these waves are able to propagate over significant radial distances before being dissipated. This issue is closely linked to the use of the α prescription which requires that the stress tensor $T_{R\phi}$ is dependent on the local value of pressure (see Eq. 1.48). In other words, if we want to place the gravitationally unstable disc in the context of accreting viscous disc (or α-discs), it is necessary to guarantee the rate of work done on the flow locally is entirely specified by the local torque. Balbus & Papaloizou (1999) found that the global transport of perturbations can be measured by the quantity $|\Omega - \Omega_{\rm p}|/\Omega$, i.e. the fractional deviation between the local flow speed and the mode's pattern speed $\Omega_{\rm p}$. By measuring this quantity in numerical simulations of marginally stable self-gravitating discs, Cossins et al. (2009) found that the flow speed of the gas into spiral arms is marginally sonic,

$$|\Omega - \Omega_{\rm p}| \approx \frac{c_{\rm s}}{R}, \tag{4.41}$$

which indicates that waves are not allowed to propagate far from where they are produced, since they almost instantaneously dissipates their energy through shocks. The fraction of non-local

energy density transport is thus given by (Cossins et al. 2009)

$$\frac{|\Omega - \Omega_{\rm p}|}{\Omega} \approx \frac{H}{R} \sim \frac{M_{\rm disc}}{M_\star} \lesssim 1, \tag{4.42}$$

where we have used Eq. 4.11. Thus, a gravitationally unstable disc in marginal state ($\mathcal{Q} \sim 1$) is characterized by a self-adjustment of the pattern speed of perturbations in order to ensure that the normal flow speed into the arms is sonic and minimizing the fraction of energy non-locally dissipated.

Alternatively, the major part of studies about this issue is based on the purpose to compare the dissipation power provided by gravitational instabilities with the expectations based on a viscous theory of disc by running numerical simulation of cooling discs using a simple cooling prescription (see next Section,). A number of numerical simulations have been carried out to compute the total torques (see Sect. 1.3.1) induced by gravitational instability for self-gravitating disc once the thermal equilibrium is reached ($\mathcal{Q} \sim 1$),

$$T_{R\phi}^{\rm grav+Reyn} = -\Sigma_{\rm g} \left(\langle u_{{\rm g},R} u_{{\rm g},\phi} \rangle + \langle u_R^g u_\phi^g \rangle \right), \tag{4.43}$$

where $u_{{\rm g},R}$, $u_{{\rm g},\phi}$, u_R^g and u_ϕ^g are the varying fluctuating components of the velocity field due to purely hydrodynamic fluctuations and disc self-gravity (see Eqs. 1.43 and 1.45). Assuming that the torque can be expressed using an α-like approach (as a viscous process would be), the "gravitorbulent" α parameter can be expressed by (see Eq. 1.48)

$$\alpha_{\rm grav} = \frac{1}{\Sigma_{\rm g} c_{\rm s}^2} \left| \frac{\partial \ln \Omega}{\partial \ln R} \right|^{-1} \left| T_{R\phi}^{\rm grav+Reyn} \right|, \tag{4.44}$$

Comparing $\alpha_{\rm grav}$ with the α value provided by viscous processes in thermal equilibrium (Eq. 1.51, see Lodato & Rice 2004; Forgan et al. 2011 and Kratter & Lodato 2016 for further details), the general picture that emerges from these studies is that the gravitational stresses can be described as a viscous stresses for low mass gravitationally unstable disc at thermal equilibrium. In this context, it is worth remarking that Lodato & Rice (2004) have found that for disc-to-star mass ratio $M_{\rm disc}/M_\star \lesssim 0.25$, the gravitational instability is well described using a local analysis. Otherwise, in massive disc where $M_{\rm disc}/M_\star \gtrsim 0.25$, global effects start to become dynamically important (Lodato & Rice 2005). In this case, the disc does not reach the quasi-steady state but its dynamics is dominated by strong transient low-m spiral modes which induce a strong angular momentum transport. Therefore, the properties of the transport mechanism induced by the development of gravitational instability depends generally on the parameter $M_{\rm disc}/M_\star$. It indicates how well the transport induced by gravitational instability can be described as a viscous (and hence local) process.

Since the angular momentum transport induced by gravitational instabilities can be included in the framework of viscous accreting discs, we can infer the magnitude of the gravitational pseudo-viscosity (which can be measured by α, Shakura & Sunyaev 1973) when the disc is in a state of marginal gravitational stability (e.g. Rafikov 2005, 2007, 2009; Clarke 2009; Rafikov 2015). Recently Rafikov (2015) developed a self-consistent analytical formalism for the evolution of a gravitoturbulent accretion disc externally irradiated by a central source. He found that a marginally gravitationally unstable discs ($\mathcal{Q} = \bar{\mathcal{Q}}$) at thermal equilibrium can be described as a viscous disc with

$$\alpha_{\rm grav} = \frac{8\sigma}{9} \frac{\left(\pi \mathcal{G} \bar{\mathcal{Q}}\right)^6}{f(\tau)} \left(\frac{\mu m_{\rm H}}{k_B} \right)^4 \frac{\Sigma_{\rm g}^5}{\Omega^7} \left(1 - \frac{T_{\rm irr}^4}{T_{\mathcal{Q}}^4} \right), \tag{4.45}$$

where $m_{\rm H} \approx 10^{-24}$ g is the proton mass, μ is the mean molecular weight, σ is the Stephan-Boltzmann constant, k_B is the Boltzmann constant, $T_{\mathcal{Q}}$ is the temperature of equilibrium derived by the marginal condition of the instability ($\mathcal{Q} = \bar{\mathcal{Q}}$) and $T_{\rm irr}$ is the temperature that characterise

the black body emission of the central source. The function $f(\tau)$ is a function of the optical depth $\tau \approx \kappa\Sigma_g/2$ where κ is the Rosseland mean opacity, given by

$$f(\tau) = \tau + \frac{1}{\tau}, \tag{4.46}$$

that captures how the accretion energy diffuses from the midplane in the optically thick and thin regimes. Eq. 4.45 shows that $\alpha_{\rm grav}$ is is simply a constant times a function of Σ_g and Ω. Typically, the value of the gravitoturbulent $\alpha_{\rm grav}$ increases steeply with radius, showing that the angular momentum transport induced by gravitational instabilities is highly effective in the outer disc regions.

4.2.5 The role of the cooling rate

For a given cooling rate, the disc provides, through the development of gravitational instabilities, a balancing heating term in order to reach the previously described state of thermal equilibrium. In other words, the amplitude of the spiral structure induced by gravitational instability has to provide a stress large enough to balance the effective cooling rate. Thus, it is expected that the strength of the modes are proportional to the cooling rate.

A numerical investigation about the relation between the strength of the modes induced by gravitational instability and the cooling rate has been carried out by Cossins et al. (2009). They performed high-resolution three-dimensional SPH simulations of self-gravitating discs evolving under an analytically prescribed cooling function,

$$\frac{{\rm d}u}{{\rm d}t} = -\frac{u}{\tau_{\rm cool}}, \tag{4.47}$$

where u is the specific internal energy of the fluid and $\tau_{\rm cool}$ is the cooling timescale. The latter is typically fixed using the simple parametrization $\Omega\tau_{\rm cool} = \beta$ where β is held constant. It is worth remembering that this cooling law does not represent the realistic thermodynamical description of the disc but it is only a convenient way to simply relate the development of gravitational instabilities with the cooling rate $\tau_{\rm cool}$. However, more recent three-dimensional hydrodynamical simulations have improved the thermodynamical description by the use of radiative transfer model based on the opacity properties of the gas appropriately computed (see Forgan et al. 2011 and reference therein).

In this way, once the disc has reached the self-regulated state, Cossins et al. (2009) computed the radial and azimuthal average of the amplitude of the density perturbations with respect to the imposing cooling (parameter β), obtaining the following relationship

$$\left\langle \frac{\delta\Sigma_g}{\Sigma_g} \right\rangle \approx \frac{1}{\sqrt{\beta}}. \tag{4.48}$$

This equation shows that the strength of perturbations increases with decreasing cooling time for disc at thermal equilibrium. This relation can be theoretically justified by remarking that the energy density of sound waves is proportional to $(\delta\Sigma_g/\Sigma_g)^2$. In order to mantain the dynamic thermal equilibrium in marginally stable discs, the rate at which energy is lost due to cooling $\propto \beta^{-1}$ must be balanced by the energy dissipated by the waves, that, as described in Sect. 4.2.4, for low massive gravitationally unstable discs dissipate their total energy through shocks at the locations where they are generated. Therefore, Eq. 4.48 is strictly related to the condition of marginal stability. Moreover, this result can be also justified by considering the properties of weak adiabatic shocks. By calculating the entropy change ΔS across a weak adiabatic shock and recalling that the heat added to the disc due to wave dissipation is proportional to ΔS, one can show that, equating the heating and cooling rate, the strength of perturbations is proportional to $\beta^{-1/2}$.

Therefore, it can be clearly inferred that the disc develops high amplitude perturbations in density which may lead to the formation of gravitationally unstable overdensities. At this

stage, these overdensities may become Jeans unstable and collapse under their own gravity. This can be easily understood by recalling that the Jeans mass decreases, for a given density, with decreasing temperature (see Sect. 1.1, Jeans 1902). Therefore, if the gravitational heating timescale is longer than the cooling timescale, gravity has not enough time to prevent the formation of bound objects.

Due to the relative importance of the disc thermal process in the formation of bound objects, it is reasonably expected that external irradiation may wipe out the spiral structure, since the gravitational heating induced by perturbations is not the unique responsible for setting the equilibrium temperature. In this case, the disc does not self-regulate as described in Sect. 4.2.3 (Kratter & Murray-Clay 2011; Rice et al. 2011; Hall et al. 2016). Essentially, the increased temperature induced by the external heating provides extra pressure support against gravitational collapse. Therefore, irradiation from the star, which is expected to highly affect the flared outer disc regions, weakens the strength of disc self-gravity. However, the effect of irradiation on a gravitationally unstable disc is still debated. While the irradiation from the star is expected to stabilize the marginally unstable discs weakening the spiral density perturbations, a very strongly irradiated disc will behave like an isothermal disc, due to the fact that irradiated discs have cooling times lower than the one required to ensure thermal equilibrium in a viscously-heated disc (e.g. Rafikov 2005; Kratter & Murray-Clay 2011; Rafikov 2015). This is because strongly irradiated disc are heated to higher temperatures by irradiation than by dissipation of accretion energy induced by gravitational stresses. In this regime, the over-densities induced by gravitational perturbation can never become pressure supported and so collapses in free-fall, since any extra heat generated by compression of the collapsing over-density is quickly radiated away. In this case, irradiation has the effect of destabilizing the disc against fragmentation, because smaller amplitude perturbations could lead to collapse (Rice et al. 2011).

Importantly, Kratter & Lodato (2016) have shown that gravitationally unstable discs might be encapsulated in three cases, depending on the relative efficience of the heating and cooling sources. Very compact discs, such as Class 0 objects, can be gravitationally unstable and regulated via thermal saturation. In other cases, the heating is mostly controlled by stellar irradiation, resulting in a self-regulated or fragmenting disc depending of the infall rate (see further details in Kratter & Lodato 2016).

4.2.6 Implications on planet formation

In recent years, several studies (e.g. Rice et al. 2005; Meru & Bate 2011a,b, 2012; Young & Clarke 2015) have been carried out in order to establish at which cooling time the disc, rather than achieving the previously mentioned self-regulated state, undergoes fragmentation in (non-)irradiated discs. The minimum cooling time below which fragmentation occurs is equivalent to a maximum α that a stable self-gravitating disc at thermal equilibrium can sustain ($\beta \propto \alpha^{-1}$, see Eq. 1.51). Alternatively, since the mass accretion rate is related to the value of α for steady state disc (see Eq. 1.38 and Eq. 1.47), it therefore exists an absolute maximum accretion rate that a marginal stable disc can sustain before fragmenting (e.g. Hall et al. 2016). Rice et al. (2005) found a value of α above which an isolated, non-irradiated gravitationally unstable disc fragment of the order ~ 0.06, though there remain uncertainties on its exact value due to limitations in numerical computations (Lodato & Clarke 2011; Meru & Bate 2012). This process of fragmentation of a self-gravitating gaseous disc is thought to be one of the possible mechanisms which lead to the formation of gaseous giant planet (Boss 2007; Clarke 2009; Boss 2011).

Since the disc is able to sustain density perturbations characterised by a maximum α, it is possible to define a critical radius $R_{\rm crit}$ outside which the disc is expected to fragment using the "gravitoturbulent" α expressed in Eq. 4.45. In order to have an estimate of typical values of $\alpha_{\rm grav}(R)$ and infer $R_{\rm crit}$, we can simplify our analysis by assuming a non-irradiated ($T_{\rm irr} = 0$) optically thick ($f(\tau) = \tau$) disc with opacities dominated by ice grains, i.e. $\kappa(T_{\mathcal{Q}}) = \kappa_0 T_{\mathcal{Q}}^2$ with $\kappa_0 = 2 \times 10^{-4}\,{\rm cm^2 g^{-1} K^{-2}}$ (Bell & Lin 1994). Assuming $\bar{\mathcal{Q}} = 1.5$, $\alpha_{\rm grav}$ is given by (e.g. Clarke

2009; Armitage 2015)

$$\alpha_{\rm grav} \sim 0.3 \left(\frac{R}{50\,{\rm au}} \right)^{9/2} . \tag{4.49}$$

Assuming that the maximum α that the disc can sustain before fragmentation is of the order ~ 0.1, typical values of $R_{\rm crit}$ are expected to be of the order of 40 - 100 au (Rafikov 2009; Clarke 2009).

Therefore, in the context of planet formation models, while the core accretion scenario is thought to occur out to $\sim$ 10 au due to high densities (and, thus, high stiking efficiencies, see Eq. 2.105), gravitational instability is historically thought to operate in the outer regions. The typical masses of these clumps can be estimated by evaluating the amount of mass involved in perturbations of the most unstable wavelengths expressed in Eq. 4.40,

$$M_{\rm fragm} \approx \Sigma_{\rm g} \lambda_0^2 = \frac{4\pi}{\bar{\mathcal{Q}}} \left(\frac{H}{R} \right)^3 M_\star, \tag{4.50}$$

where we have used Eqs. 4.4 and 4.40. Assuming $\bar{\mathcal{Q}} \approx 1$, for typical values of the disc aspect ratio, the mass of fragments are of the order of $\sim 10\,M_{\rm J}$, where $M_{\rm J}$ is the mass of Jupiter.

Importantly, the development of gravitational instabilities is also expected to play a key role in the growth of planet-building rocky cores. Most of the uncertainties in the core accretion model are related to our ignorance about how micrometer sized dust grains grow to form kilometer sized planetesimals. As described in Sect. 2.5, coagulation of millimetre grains is hampered by the fast radial drift and fragmentation of colliding grains. One possible mechanism to stop their migration and increase the likelihood of constructive collisions is that grains become concentrated in these gravitationally-induced spiral density structures (Rice et al. 2004, 2006b; Gibbons et al. 2012, 2014, 2015; Booth & Clarke 2016). The dynamical effect of spiral density perturbation on non-growing dust grains will be investigated in Chapter 6.

4.3 Observational evidences of self-gravitating discs

The development of gravitationally instabilities in young disc produces a variety of effects which might be observationally identified with current observing facilities over a wide range of wavelengths. Since the mass of the disc is expected to be higher in the initial stage of the disc formation process, young objects are the best candidate to be gravitationally unstable. The observational consequences of the instability can be splitted in two groups: direct and indirect evidences.

4.3.0.1 Direct evidences

- The macroscopic feature introduced by self-gravity is the creation of large-scale density fluctuations in the form of a spiral pattern. These features can be observationally and unambiguously identified in the near-infrared scattering emission (Dong et al. 2015a; Dipierro et al. 2015a; Pohl et al. 2015; Dong et al. 2016a) and through (sub-)millimetre observations (Cossins et al. 2010; Vorobyov et al. 2013; Dipierro et al. 2014, 2015a) where the disc is mostly optically thin. See also Chapter 5 and 6.

- The development of gravitational instability may lead to the formation of bound objects in the outer disc regions (Boss 2007, 2011), which might be observationally detected (Kraus & Ireland 2012; Vorobyov et al. 2013; Testi et al. 2015). It is worth remarking that bound objects wich might be formed via fragmentation induced by gravitational instabilities has been recently detected in the system L1448 IRS3B (Tobin et al. 2016).

The detection of the spiral density waves in self-gravitating circumstellar discs can provide morphological informations in terms of the pitch angle, size and number of the spiral arms. In addition, informations about the intensity of the spiral arms with respect to the inter-arm regions

can be obtained by high-resolution observations. In this context, it is worth remarking that Cossins et al. (2009) have shown that the morphology and the intensity of the spiral structure can be related to the disc-to-star mass ratio and to the cooling time. Using a 3D global SPH simulation with a simple cooling model (expressed by Eq. 4.47), they have investigated the radial and azimuthal structure of a self-gravitating disc in the marginal state. They found that, for a given star mass, the spiral structure is characterized by a decreasing number and an increasing size and pitch angle of the spiral arms as the disc mass increases. Moreover, the sharpness of the spiral structure can related to the disc cooling time, as described in Sect. 4.2.5. In addition, the spiral shocks occurring in arm regions of gravitationally unstable discs, produce clear chemical signatures of the disc dynamics which may be detected using ALMA observations with the aim to infer the physical structure of self-gravitating discs (Ilee et al. 2011; Douglas et al. 2013).

4.3.0.2 Indirect evidences

- Some clues about the importance of the disc self-gravity can be obtained by the measure of the mass and radial extent of protostellar discs. While the estimates of the radial disc extents appear robust, there are still large uncertainties in the value of the disc mass inferred from observations (see Sect. 3.3.3).

- The disc self-gravity affects the rotation of the infalling material leading to a significant deviation from Keplerian rotation (Rosenfeld et al. 2012, 2014). In this context, it is worth remarking that Lodato & Bertin (2003) have proposed a method aimed at evaluating significant deviation from Keplerian rotation starting from the SEDs of FU Orionis objects.

- Gravitational instabilities can trigger large, short-lived accretion events, producing outbursts which can be detectable as changes in the accretion luminosity (Armitage et al. 2001; Vorobyov & Basu 2005, 2006, 2015).

In the following we present a brief summary of the observed mass and radial extent of very young disc, which might be an indirect evidence of gravitationally unstable discs.

4.3.1 Observed properties for Class 0/Class I objects

In order to put realistic constraints on self-gravitating disc properties, it is necessary to consider observation of Class0/Class I objects which more likely go through a self-gravitating phase. However, they are very difficult to identify because they are still heavily embedded in their nascent molecular cloud cores and are surrounded by power energetic outflows. As a consequence, the emission from the surrounding envelope is entangled with that of the disc. In addition, since these are short-lived ($\sim 10^5$ yr) and thus relatively rare objects, solid observational evidence is difficult to find. Sensitive interferometers are needed to achieve high angular resolution and spatially/spectrally disentangle the various disc, envelope and outflow components. A recent survey of Class I sources in Taurus with CARMA has shown that the masses of these young objects are typically higher than in the Class II phase by a factor of two and, importantly, $\sim$ 10% of the sample have disc masses higher than $0.01 M_\odot$ (Eisner 2012; Sheehan & Eisner 2014).

Moreover some individual sources have been measured to have relatively massive and compact discs. One of the most intriguing case is represented by the disc around the Class I object WL12 (Miotello et al. 2014b). Recent ALMA observations suggest that WL12 is a very small disc with a 30-50 au radius with a disc mass of $\gtrsim 0.3\, M_\odot$ orbiting around a star with mass $\sim 0.6\, M_\odot$. The high disc-to-star mass ratio ensures that WL12 is an excellent candidate to be a gravitationally unstable disc. Moreover, Rodríguez et al. (2005) presented mm observations of IRAS 16293-2422B, a Class 0 object located in the Ophiuchus molecular complex at a distance of 120 pc, revealing a characteristic circumstellar disc radius of < 30 au. Recently, high resolution (~ 0.2 arcsec) sub-millimeter continuum and line observations with ALMA carried out by Zapata et al. (2013) of the same disc confirmed this result and, moreover, they estimated a lower limit for

the mass of the disc of about $0.03\,M_{\odot}$. However, they did not reveal any inner structure inside of its diameter maybe because the dust emission is optically thick at the selected wavelengths. Moreover, Ward-Thompson et al. (2011), using 7 mm and 4.4 cm observations of the Class 0 protostar VLA 1623 taken with the Multiple Element Radio Linked Interferometer Network (MERLIN), estimated an upper limit for the disc mass of $< 0.13 M_{\odot}$ and for its radius of ~ 50 au. Moreover, recent observations (Tobin et al. 2012) of the Class 0 protostar L1527 IRS in the Taurus cloud revealed an edge-on disc with an outer radius of at least 100 au about a protostellar mass of $\sim 0.2\,M_{\odot}$. This last result indicates that these discs can be also reasonably extended and their size can be comparable to or smaller than the size of older discs.

However, most of the mass estimates in young discs are expected to be highly biased by the typical high optical depth of self-gravitating discs. Importantly, Forgan et al. (2016) constructed one dimensional self-gravitating disc models for five disc candidates around massive stars: G11.92 - 0.61 MM1 (Ilee et al. 2016), NGC 6334 I(N) SMA1b (Hunter et al. 2014), AFGL 4176 mm1 (Johnston et al. 2015), IRAS 16547-4247 (Zapata et al. 2015) and IRAS 20126+4104 (Chen et al. 2016). They found a disc-to-star mass ratios for most of the sources of the order of ~ 0.2, implying that these discs are likely to be gravitationally unstable.

4.3.2 Recent detections of spiral structure

Spiral structures has recently been observed in a quite large number of discs, typically in the late stage of their evolution. The structure has been identified mostly by imaging the distribution of scattered light at near-infrared wavelengths (Muto et al. 2012; Grady et al. 2013; Garufi et al. 2013; Wagner et al. 2015) and only one case has been detected by high-resolution ALMA observations of molecular CO line emission (Christiaens et al. 2014). However, a new spiral structure in dust continuum submillimeter emission in the disc around the star Elias 2-27 has recently been detected by ALMA (Pérez et al. 2016).

Two spiral arms seem to be most common feature in all the disc where a spiral structure is detected, though the number of spiral arms might be an artefact resulting from the observational limitations (see Chapter 5). These arms are rather open, with pitch angles of the order of 10° to 20°. Interestingly, some objects reveal a different shape of disc substructures at different wavelengths, further challenging the modelling of these discs and the comparison with theoretical models. Since spiral-like features has been mostly observed in transitional discs (see Chapter 10) around Herbig stars, the self-gravitating origin of this structure might be unlikely due to the relatively old age and the low disc-to-star mass ratio. Several scenarios have been proposed for the occurrence of spirals in transitional discs (Boccaletti et al. 2013). The most widely accepted interpretation is the presence of a planet or sub-stellar companion inside or outside the cavity of the transitional discs (see Sect. 10.2.0.2). However, this hypothesis might explain a limited sample of this substructures and a large range of assumptions needs to be made to compare the observations with theoretical models. Importantly, initial attempts to fit the shape of the observed spirals using linear density waves theory (see Sect. 7.1.2) were consistent with a disc model characterised by a relatively high temperature at the location of the spirals (see Sects. 4.3.2.1 and 4.3.2.2, Muto et al. 2012; Benisty et al. 2015), i.e. several hundred Kelvin, at least ten times greater than the temperature calculated based on the stellar flux received (Rafikov 2002; Kley & Nelson 2012). However, massive companions produce a different morphology of the tidally-induced spiral structure (Zhu et al. 2015a), which might be well fitted with the observed spirals (Dong et al. 2015b). Importantly, Isella & Turner (2016) have shown that a massive planet able to carve a gap in the gas density structure (see Chapter 7) generates spiral density waves which produce weak signatures at near-infrared wavelengths. This is because the outer edge of the gap cleared by the planet casts a shadow on the outer disc, strongly reducing the visibility of any spiral feature located outside the gap (see also Juhász et al. 2015). Importantly, numerical simulations have shown that while the spiral structure induced by planet-disc interaction are unlikely to trap dust particles, a gravitationally induced spiral structure will result in a high concentration of large grains (Rice et al. 2004, 2006b; Gibbons

et al. 2012, 2014, 2015; Dipierro et al. 2015a). This produces a remarkable difference in the disc appearance in scattering emission (which trace small grains at the disc surface, see Sect. 3.2.1) and (sub-)millimetre emission (which trace large grains on the midplane, see Sect. 3.3) that can be a valuable tool to distinguish between the origin of the spiral structure.

The importance of gravitational instabilities in protoplanetary discs cannot be ruled out by analyzing the limited sample of spiral structure observed to date. As already remarked, young objects, which are the best candidate to be gravitationally unstable, are very difficult to observe since they are still heavily embedded in their nascent molecular cloud cores. With the advance of observational techniques, the sample of discs with spiral structures can be broadened by including Class 0/I sources. In the following we briefly summarize the main features of the spiral structures detected to date in protoplanetary discs.

4.3.2.1 HD 135344B

HD 135344B (also known as SAO 206462) is an extensively studied F4Ve star (Dunkin et al. 1997) with estimated mass and age of about 1.7 $M_{\odot}$ and 8 Myr, respectively. The system hosts a transitional disc and is located in the Upper Centaurus Lupus star forming region at a distance of 140 pc (van Boekel et al. 2005). One of the most intriguing characteristics of this disc is the different disc substructures revealed in near-infrared and (sub-)millimetre observations. While scattered light observations in the near-infrared have revealed spiral structure in the outer disc (see the upper left panel of Fig. 4.1, Muto et al. 2012; Garufi et al. 2013; Stolker et al. 2016), (sub-)millimetre ALMA observations have revealed large-scale horseshoe in the dust continuum (see the lower left panel of Fig. 3, Pérez et al. 2014; Pinilla et al. 2015; van der Marel et al. 2016a). This object shows a large region of highly depleted dust close to the star surrounded by bright rim located at $\sim$ 45 au (Brown et al. 2009) and extends out to 220 au and has a mass of 0.026 $M_{\odot}$ (Lyo et al. 2011; Andrews et al. 2011). Adopting these parameters, Garufi et al. (2013) found a value for the Toomre $\mathcal{Q} \gtrsim 25$ across all the disc, meaning that disc would have to be more than twenty times more massive in order to be gravitationally unstable and develop spiral density perturbations. Moreover, Hall et al. (2016) found that, assuming that the accretion rate of the disc is triggered by gravitational stresses, a disc-to-star mass ratio of ~ 0.1 is needed in order for the disc to be in a quasi-steady, self gravitating state.

By fitting the shape of the observed spirals at near-infrared wavelengths using linear density waves theory (see Sect. 7.1.2), Muto et al. (2012) found that the shape is consistent with a disc model with an aspect ratio of 0.2 at a distance of 100 au, which is too large for any realistic disc structure. Interestingly, the substructure detected in scattered light observation might be explained as perturbations of the disc's pressure scale height (Juhász et al. 2015). The horseshoes observed at (sub-)millimetre wavelengths are often assumed to be due to a vortex induced by a low-mass companion in the cavity. Models of dust trapping by a vortex predict that larger grains would be more azimuthally concentrated in the centre of the vortex, producing a different azimuthal extent of the horseshoe detected at different wavelengths. However, in this disc smaller grains appear to be more trapped than larger grains, challenging the vortex model (Pinilla et al. 2015). An alternative explanation for horseshoes in transitional discs is based on the eccentricity development at the edge of the cavity surrounding a binary system characterised by a high mass ratio between the companion and the star $\gtrsim 0.05$ (see Chapter 11 for details). The relation between the spiral arms observed in near-infrared scattered light and the structures seen in (sub-)millimetre emission was recently investigated by van der Marel et al. (2016a). They proposed that this structure can be explained by the presence of a massive planet at 30 au that produces a spiral density waves outwards. In this picture, a second spiral arm is triggered by the dust concentration on the pressure bump at the outer edge of the gap carved by the planet.

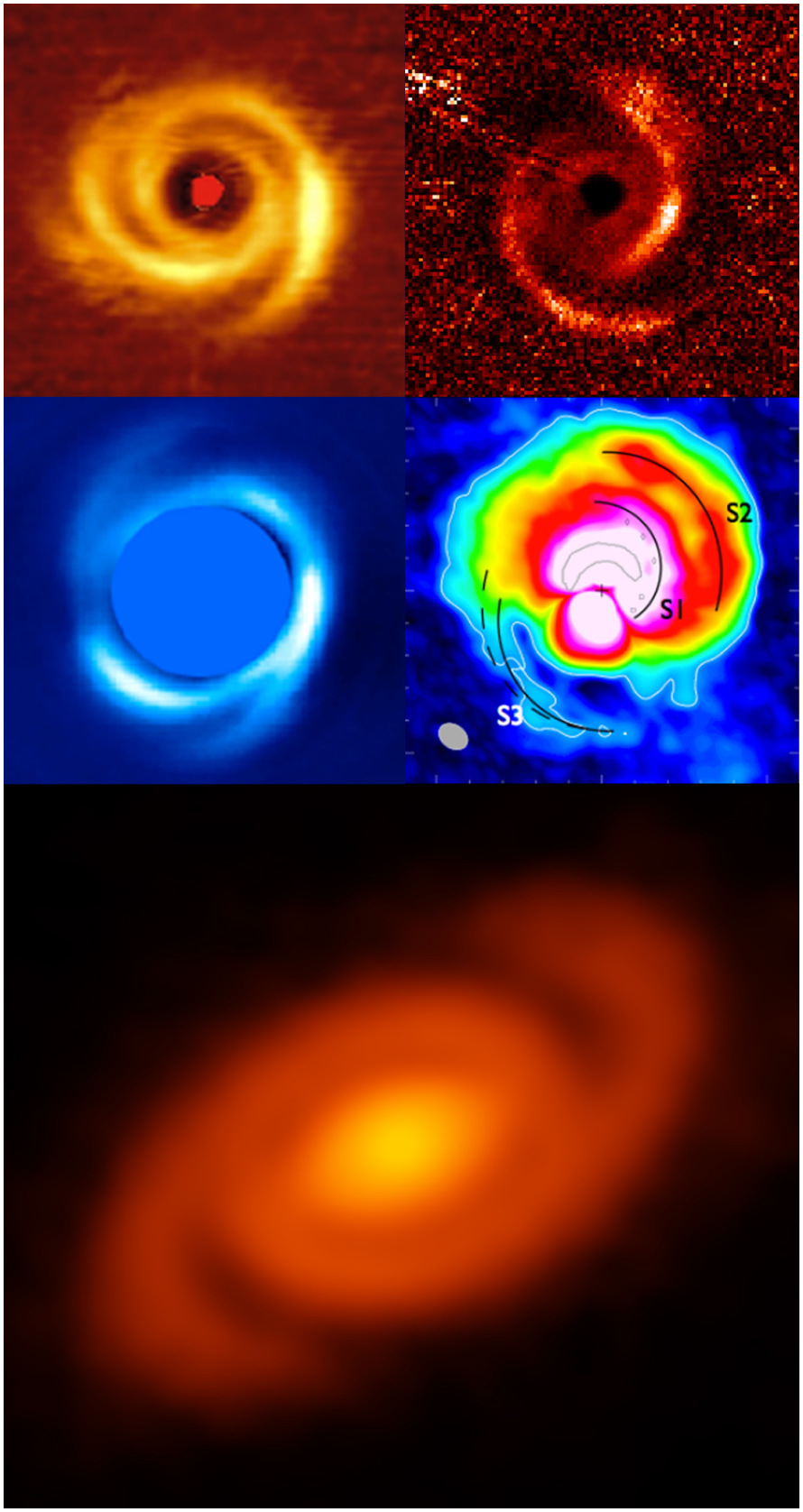

Figure 4.1: Gallery of spiral structure observed to date in protoplanetary discs. (top left) polarized scattered light emission of HD 135344B (Garufi et al. 2013), (top right) MWC 758 (Benisty et al. 2015), (middle left) HD 100453 (Wagner et al. 2015), (middle right) ^{12}CO J=2-1 line emission in HD 142527 (Christiaens et al. 2014) and (bottom) dust continuum emission at 1.3 mm of Elias 2-27 (Pérez et al. 2016). All the images are on different scale.

4.3.2.2 MWC 758

The spiral structure in the transitional disc around the Herbig A5 star MWC 758 was first imaged in near-infrared scattered light image by Grady et al. (2013). The central star has a mass of $\sim 1.8\,M_{\odot}$ and an estimated age of 3.5 Myr (Meeus et al. 2012). The disc has also been identified as hosting a partially cleared cavity detected in the (sub-)millimetre with a radial extent between 70 and 100 au (Isella et al. 2008). The mass of the disc is $\sim 0.01 M_{\odot}$, measured from sub-millimetre observations (Andrews et al. 2011). The relative old age of the source ($\sim 3.5 \pm 2$ Myr, Meeus et al. 2012) and the high mass of the central star ($\sim 2 M_{\odot}$, Isella et al. 2010b) suggest that the spiral structure is probably triggered by a mechanism different from the gravitational instabilities. If we assume that the disc mass measurement are inaccurate, the spiral structure can be explained by gravitational instabilities if the mass of the disc surrounding MWC 758 has been underestimated by a factor of 50, since a two-armed spiral structure typically require a disc-to-star mass ratio of $\gtrsim 0.5$ (Lodato & Rice 2005; Cossins et al. 2009; Dong et al. 2015a).

The spiral structure was then further imaged in scattered light by Benisty et al. (2015), using SPHERE/VLT (see the upper right panel of Fig. 4.1). The image revealed a grand design $m = 2$ spiral structure that asymmetrically shadow the outer disc with a pitch angle of $\sim$10° – 15° with an azimuthal extent of $\sim$180° – 270°, and a emission ratio between arm and interarm region of $\sim$1.5. A likely explanation for the spiral structure in this object is that it is triggered by a massive planet. Although the planetary hypothesis has been usually invoked to explained one armed spiral structure with a low pitch angle and contrast between the arm and the background disc, for massive planet the morphology of the perturbations changes and cannot be explained using the linear theory. Therefore, fitting the shape of the observed spirals using linear density waves (see Sect. 7.1.2) might lead to serious errors in inferring the local disc properties (Benisty et al. 2015). As recently investigated by Zhu et al. (2015a), a secondary spiral arm, especially for the inner arms, is excited by a massive planet. The whole spiral has an increasing pitch angle and it is characterised by an increased separation in the azimuthal direction between the primary and secondary arms with increasing planetary mass. We will focus on the morphology of the planet-induced spiral in Chapter 7. Dong et al. (2015b) have suggested that the observed spiral waves might be excited by a massive planet ($\sim 6\,M_{\rm J}$) orbiting outside the cavities at 160 au. Although they well reproduced the main morphological features of the detected spiral, they needed to assume that the planet is in the initial phases of opening the gap. In particular, they stopped their simulation after 20 orbits at the orbital radius of 160 au, a time lower than the time required to open a gap at the planet location. This is due to the fact that once the gap is fully opened, the part of the spiral density wave with the largest pitch angle and contrast with respect to the surrounding fluid will disappear, since it will be located within the gas/dust depleted gap (Isella & Turner 2016). To deepen understand the origin of these spiral features, high resolution images at (sub-)millimetre wavelengths with ALMA would be needed. In this context, Juhász et al. (2015) demonstrated that spirals generated by planet-disc interaction more likely results from changes in the vertical structure rather than the density structure, which are hard to detect in millimeter emission. By contrast, spiral induced by gravitational instabilities might trap large dust particles, strongly affecting the resulting (sub-)millimetre disc emission (Rice et al. 2004, 2006b; Gibbons et al. 2012, 2014, 2015; Dipierro et al. 2015a).

4.3.2.3 HD 100453

HD 100453 is an Herbig Ae star with a mass of $1.7\,M_{\odot}$ and an estimated age of 10 Myr (Collins et al. 2009) located in the Lower Centaurus-Crux association at a distance of $\sim$ 112 pc (Dominik et al. 2003). Recent SPHERE/VLT near-infrared images revealed a ring of dust at 21-29 au projected separation from the star, with two highly symmetric spiral arms in the outer disc extending out to $\sim$ 45 au (see the middle left panel of Fig. 4.1, Wagner et al. 2015). Moreover, NACO/VLT observations identified a companion candidate at a projected separation of 120 au

from the primary (Chen et al. 2006). The companion, HD 100453 B, was found to be a co-moving M dwarf with an estimated age and mass of 10 Myr and 0.3 $M_\odot$ (Chen et al. 2006; Collins et al. 2009). The disc around HD 100453 is rather compact and has a mass of a few tenths of Jupiter mass of H_2 gas at 150 K, and less than a few Earth masses of optically thin H_2 gas at 300 K and higher temperatures (Carmona et al. 2008). The low disc mass and the old age of this system suggests that gravitational instabilities cannot explain the observed spiral structure. The widely accepted explanation of the the disc truncation and the two-armed spiral structure is the tidal interaction with the companion. Simulations in the context of protoplanetary discs have shown that two-armed spirals could be induced by stellar encounters (Quillen et al. 2005). Importantly, Dong et al. (2016b) demonstrates that the disc truncation at ∼45 au and the double spiral arm patterns in near-infrared scattered light images can be produced by the companion.

4.3.2.4 HD 142527

HD 142527 is a transitional disc around a relatively old (∼ 5 Myr, Fukagawa et al. 2006) Herbig Fe star with a mass of ∼ 2.2 $M_\odot$ (Verhoeff et al. 2011) at a distance of 140 pc. Due to its unique morphology, this object is one of the best studied transitional discs in literature. The disc boast the largest inner cavity known, with a radial extent of $\gtrsim$ 100 au, and presents a sharp asymmetry in the (sub-)millimetre emission (see Sect. 11.3.1.5, Casassus et al. 2012; Fukagawa et al. 2013; Casassus et al. 2015b,a; Muto et al. 2015). The near-infrared scattered light image of the disc shows extended emission up to 300 au with two bright arcs facing one another along the east-west direction and one spiral arm extending to the north from the western arc (Fukagawa et al. 2006; Canovas et al. 2013; Rodigas et al. 2014; Avenhaus et al. 2014). Two more spiral arms extending over 600 au are observed in ^{12}CO J=2-1 line emission by ALMA (see middle right panel of Fig. 4.1, Christiaens et al. 2014). Recently, Marino et al. (2015) proposed that the two intensity peaks seen along the outer disc of the HD 142527 system in near-infrared observations might be shadows cast by a warped inner disc produced by the tidal interaction with a sub-stellar companion orbiting on a plane inclined by ∼ 70° with respect to the outer disc (Lacour et al. 2016). By contrast, ALMA observations of dust continuum emission show a significant azimuthal asymmetry with a bright horseshoe emission in the northern part of the disc with a contrast of ∼ 30 at 336 GHz between the northern peak and the south western minimum (Casassus et al. 2013; Perez et al. 2015) which may be caused by a massive close companion (Lacour et al. 2016, see also Chapter 11). Focussing our attention on the extended spiral structures in the outer disc, Fukagawa et al. (2013) and Christiaens et al. (2014) found that the Toomre parameter is $\mathcal{Q} \sim 2$ in the outer disc regions, which might indicate that the self-gravitating hypothesis would be valid to explain the spiral structure. According to the model presented in Hall et al. (2016), in order to match the observed accretion rate of HD 142527, such a disc would require an incredibly high disc-to-star mass ratio of ∼ 0.75. Such a disc would probably produce strong transient low-m spiral modes never reaching the quasi-steady state. This make the gravitational instability an unlikely explanation for the observed spiral structure.

4.3.2.5 Elias 2-27

A particular focus needs to be given to the case of the disc around the youngest source in our sample: Elias 2-27. The source Elias 2-27 is classified as a class II young stellar object and is located in the $\rho-$Ophuicus star forming complex at a distance of ∼ 140 pc. From (sub-)millimetre low-resolution observation performed in the past, it is thought that the disc around Elias 2-27 has a mass in the range $[0.04, 0.14]\,M_\odot$ and orbits around a star with a mass of ∼ 0.5 $M_\odot$ (Mamajek 2008; Andrews et al. 2009). The new ALMA observation at 1.3 mm reveals a grand-design two-armed spiral structure in the radial range between 100 and 300 au (see bottom panel of Fig. 4.1, Pérez et al. 2016). The relative young age and the high dust-to-gas ratio of this system suggests that the disc can be massive enough to be gravitationally unstable. Using the disc properties inferred by recent (sub-)millimetre observations (Andrews

et al. 2009), the disc is expected to be marginally unstable, due to the low accretion rates and the absence of fragments in the outer disc regions (Hall et al. 2016). This result represents the first case of spirals detected at (sub-)millimetre wavelengths, providing unique insights on the disc midplane where most of the mass is concentrated and where planet formation takes place. While in previous cases the spiral structure is detected at short wavelengths, at which the disc is mostly optically thick, the emission detected in Elias 2-27 at radius $\gtrsim 10$ au is optically thin. Therefore the spiral structure has to be linked with true density, temperature and/or grains' optical properties variations (see Chapter 3). New high-resolution multi-wavelengths ALMA observations might potentially reveal a differential grain growth processing in the arm and inter arm regions by measuring the spectral index variations, further strengthening the idea that the detected spirals are density enhancements induced by gravitational instability. Moreover, comparing the size of spiral arms at different wavelengths it is possible to see if large grains are trapped inside the arms (see Chapter 6), offering a robust path to figure out which is the mechanism behind the formation of this structure.

Chapter 5

How to Detect the Signatures of Self-Gravitating Discs with ALMA

Based on the paper by ***Giovanni Dipierro****, Giuseppe Lodato, Leonardo Testi and Itziar de Gregorio Monsalvo: "How to detect the signatures of self-gravitating circumstellar discs with the Atacama Large Millimeter/sub-millimeter Array" published in Monthly Notices of the Royal Astronomical Society,* ***444****, 1919-1929 (2014).*

As already remarked in the previous Chapter, it is believed that in the early stage of star formation (Class 0/Class I objects), the disc could be massive enough to have a non-negligible dynamical effect on the evolution of the overall system. The disc self-gravity may affect the disc dynamics through the propagation of density waves which lead to the formation of a prominent spiral structure and provide a non-negligible contribution to the angular momentum transport (Lodato & Rice 2004, 2005). The macroscopic feature introduced by self-gravity is the creation of large-scale density fluctuations (in the form of a spiral pattern). To detect these features observationally and unambiguously identify them as density enhancement, it is necessary to observe at (sub-)mm wavelengths where the disc is mostly optically thin in the vertical direction and the emission from the mid-plane can be traced directly (Dullemond et al. 2007). It is thus very timely to investigate how ALMA observations could be optimized to study the expected effects of self-gravity in discs.

A first attempt at assessing, based on the expected properties of the gravitational instability, the detectability of a spiral structure in the disc has been made by Cossins et al. (2010) who have found that spiral structures in compact self-gravitating discs are visible with ALMA over a wide range of wavelengths. In this context, high-resolution, direct imaging observations of circumstellar/protoplanetary discs have already revealed non-axisymmetric structures which are believed to be due to embedded planets or external tidal interactions (see Sect. 4.3.2).

The theoretical characterization of gravitational instability conducted by Cossins et al. (2009) has shown a number of interesting features which could be empirically confirmed by high-resolution observations. For example, features such as the opening angle and the size of spiral arms are related to the disc-to-star mass ratio. Thus, the detection of spiral arms may be useful to infer some physical characteristics about the star-disc system (see Sect. 4.3). In addition, it is expected that the density inhomogeneities induced by the development of gravitational instabilities affects both the radial migration and the growth of the dust grains (Rice et al. 2004, 2006b, see also Chapter 6). In this Chapter we demonstrate whether the spiral structure excited by the gravitational instability can be detected using ALMA. This work extends the previous

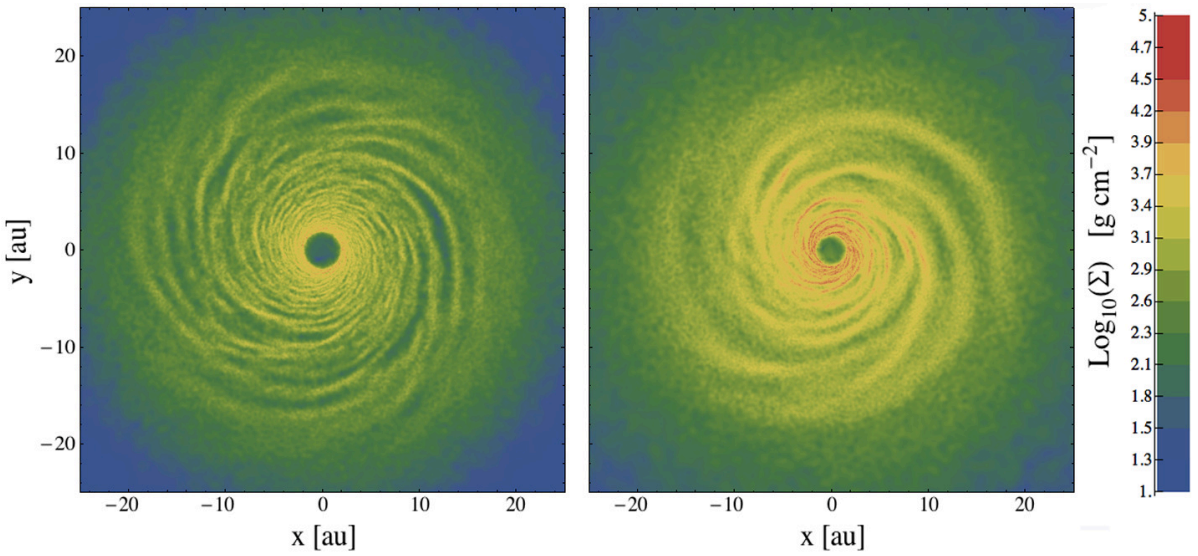

Figure 5.1: Surface density structure at the end of the simulations with radius 25 au and $M_\star = 1\,\mathrm{M}_\odot$ and two different disc-to-star mass ratio: (left) $q = 0.1$; (right) $q = 0.25$.

analysis of Cossins et al. (2010) through the evaluation of ALMA simulated observations of a wide set of both face-on and non face-on discs with different properties in size and mass and a more accurate computation of the dust opacity.

5.1 Disc model

The simulations used to generate the expected disc emission have been taken from Lodato & Rice (2004). They performed simulations using a 3D smoothed particle hydrodynamics (SPH) code, a Lagrangian hydrodynamics code (see reviews of Monaghan 2005 and Price 2012). The system comprises a central star modelled as a point mass $M_\star$ surrounded by a gaseous disc of mass $M_{\rm disc}$ which extends from $R_{\rm in} = 0.25$ to $R_{\rm out} = 25$ in code units. The simulations were performed using two values for the disc-to-central object mass ratio $q = M_{\rm disc}/M_\star = 0.1$ and $q = 0.25$. Regarding the thermal aspect of disc evolution, the heating is governed by both PdV work and viscous dissipation while the cooling is implemented using a simple cooling law given by

$$\frac{\mathrm{d}u_{\rm i}}{\mathrm{d}t} = -\frac{u_{\rm i}}{t_{\rm cool,i}}, \tag{5.1}$$

where $u_{\rm i}$ is the specific internal energy and $t_{\rm cool,i}$ is the cooling time associated with the i^{th} particle. The latter is determined using the simple parametrization $\Omega_{\rm i}\, t_{\rm cool,i} = \beta_{\rm cool}$, where $\beta_{\rm cool}$ is held constant and equal to 7.5. This form of cooling has been used extensively in simulations of discs in various context, as described in Sect. 4.2.5. The choice of $\beta_{\rm cool}$ value is critical because the cooling modifies the dynamical evolution of the disc leading to the formation of bound objects in sufficiently high-resolution simulations (Meru & Bate 2011b; Lodato & Clarke 2011, see Sect. 4.2.5). However, our analysis is aimed at investigating the observability of the "grand design" spiral structure of a self-gravitating disc and we do not consider phenomena occurring at small scales. We have thus used a value of $\beta_{\rm cool}$ that, at the numerical resolution used in Cossins et al. (2009), does not lead to fragmentation.

In order to consider discs with realistic different properties in mass and radial extension, we modify the length and the mass scale of the data output of the SPH simulations. In this way, we can produce discs with properties consistent with the expected typical values obtained in observational campaigns (see the review of Williams & Cieza 2011). We therefore consider two different values of the code unit length such that the radial extensions of the disc are 25 au or 100 au. Regarding the mass of the disc-star system, we use three different value of the mass

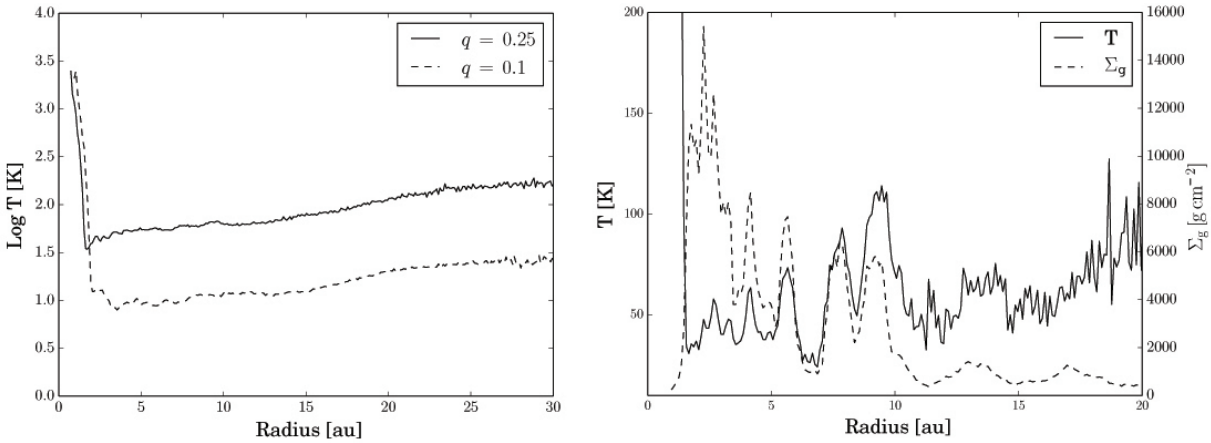

Figure 5.2: (left) Azimuthally averaged temperature profiles of the two discs with radius 25 au, $M_\star = 1\,\mathrm{M}_\odot$ and $q = 0.25$ (solid line) and $q = 0.1$ (dashed line) once they are settled into the dynamic thermal equilibrium state. (right) Temperature (solid line) and density (dashed line) profiles for the disc with radius 25 au, $M_\star = 1\,\mathrm{M}_\odot$ and $q = 0.25$. The x-axis indicates the distance from the star along a line across the disc.

scale such that the central mass star is equal to $0.3\,\mathrm{M}_\odot$, $1\,\mathrm{M}_\odot$ and $3\,\mathrm{M}_\odot$ and with disc mass covering a range from $0.075\,\mathrm{M}_\odot$ to $0.75\,\mathrm{M}_\odot$.

5.1.1 Disc evolution

Once the thermal equilibrium state is reached, the disc is characterized by a spiral structure induced by self-gravity. The spiral features are clearly seen in Fig. 5.1 where the (logarithmic) surface density is shown for the two simulations with radius 25 au and $M_\star = 1\,\mathrm{M}_\odot$ and $q = 0.1$ (left) and $q = 0.25$ (right). It can be noted that, as the disc-to-star mass ratio q increases, the pattern of the instability becomes more dominated by low number of arms and, in addition, the spiral structures tends to become more open. Also, as the disc mass increases, a marginally stable disc becomes thicker (since $H/R \sim q$) and the typical size of the spiral arms, which is of the order of H, thus increases. A detailed Fourier analysis of the expected morphology of the spiral can be found in Cossins et al. (2009).

In the left panel of Fig. 5.2 we show the azimuthally averaged temperature profiles associated with the two discs under consideration. The radial trend and the difference between profiles are linked to the thermal equilibrium state reached in the discs in which the cooling is balanced by heating generated through gravitational instabilities.

To highlight the effect of the spiral on the gas and temperature structure within the disc, we show in the right panel of Fig. 5.2 the temperature and density profiles related to the disc with radius 25 au, $M_\star = 1\,\mathrm{M}_\odot$ and $q = 0.25$ along a horizontal slice across the disc. As could have been expected, the peaks in temperature occur at the same positions of the peaks in density showing that the overdensity regions, which correspond to the arm regions, are characterized by relatively higher temperatures. This is due to gas compression in the spiral arms. From an observational point of view, the density and temperature bumps occurring in spiral arms allow the peculiar spiral structure of self-gravitating discs to be more easily detectable.

Note that the self-regulation of the gravitational instability, which is established by the thermal saturation of the instability, results in a slightly increasing azimuthally averaged temperature profile (see Fig. 5.2), which is not expected for a typical passive irradiated disc around pre-main-sequence stars. In fact, the disc model adopted here does not take into account the irradiation from the central star. Assessing the influence of this additional heating in the evolution of a self-gravitating disc is beyond the scope of this thesis. In general, we expect that such extra heating source would weaken the amplitude of the spiral perturbations somewhat (see Cossins et al. 2009 and Rice et al. 2011). Thus, from an observational point of view, the

detection of the spiral structure would probably be more challenging due to the decrease of the arm-interarm density contrast.

5.2 Generation of intensity maps

Starting from the output of the SPH simulations described above, it is then necessary to use the disc parameters to create intensity maps of the objects previously simulated. Following the approach of Cossins et al. (2010), we assume that the disc is in thermal equilibrium at temperature T and is vertically isothermal. Therefore the source function at frequency ν is given by Eq. 3.2. The temperature T is obtained by vertically averaging the temperatures of all the particles inside a cylindrical radius equal to the smoothing length h associated with the particle under consideration.

At sub-mm wavelengths, dust emission is generally optically thin, meaning that it is necessary to estimate the optical depth in order to create intensity maps. The optical depth τ_ν at frequency ν is defined by:

$$\tau_\nu = \int_{-\infty}^{\infty} \kappa_\nu(z)\rho(z)\, dz, \tag{5.2}$$

where $\kappa_\nu(z)$ is the opacity of the disc at frequency ν. Assuming that z is the vertical component of the disc coordinate system and that the opacity is invariant with z, we can express the optical depth, for a face-on disc, as: $\tau_\nu = \kappa_\nu \Sigma$, where Σ is the total column density estimated by projecting the particles on the mid-plane and using a standard SPH formalism in 2D. Finally, we calculate intensity maps of discs with different properties in terms of mass, size and inclination using the following expression:

$$I_\nu = B_\nu\,(1 - \exp(-\tau_\nu)), \tag{5.3}$$

where the optical depth is appropriately computed taking into account also the case of non face-on discs.

5.2.1 Dust opacity

In the high-density environment of a circumstellar disc, microns dust grains are expected to grow via collisional agglomeration and gravitationally settle towards the mid-plane of the disc (Natta et al. 2007; Testi et al. 2014). The evidence that grains in discs are significantly larger than grains in the diffuse interstellar medium (ISM) has been obtained from a variety of observational techniques (see Sect. 3.3.1 and the reviews of Testi et al. 2014 and Andrews 2015). In this scenario, the dust opacity is obviously influenced by grain processing and therefore its value depends on the grain sizes, chemical compositions and shapes that cannot be easily probed by observations.

Often, the dust opacity is expressed as a simple power-law function of wavelength appropriate for dust grains with a size distribution of the form expressed in Eq. 3.8, (Draine 2006). Assuming that the dust grains are sphere-shaped particles in the limit $a \ll \lambda$ where λ is the radiation wavelength, the dust opacity law can be expressed by Eq. 3.9 where the parameter β represents, as described in Sect. 3.3.1, a very good indicator of the level of grain growth.

Generally, through sub-millimetre observations and making some assumptions on the chemical composition, shape and size distribution of the dust grains, it is possible to put constraints on the level of grain growth. For an optically thin disc in the Rayleigh-Jeans limit, the sub-millimetre spectral energy distribution (SED) has a power-law behaviour with frequency expressed by $F_\nu \propto \kappa_\nu\, \nu^2 \propto \nu^\alpha$ where α is the spectral index. Thus, from Eq. 3.9, the indices β and α are related by $\beta = \alpha - 2$.

In light of the previous discussion, we compute the dust opacity using a particular dust model (Ricci et al. 2010) with an appropriate choice of model parameters. In Fig. 5.3, the dust opacity values in units of $\mathrm{cm^2 g^{-1}}$ of gas are shown for the range of wavelengths under

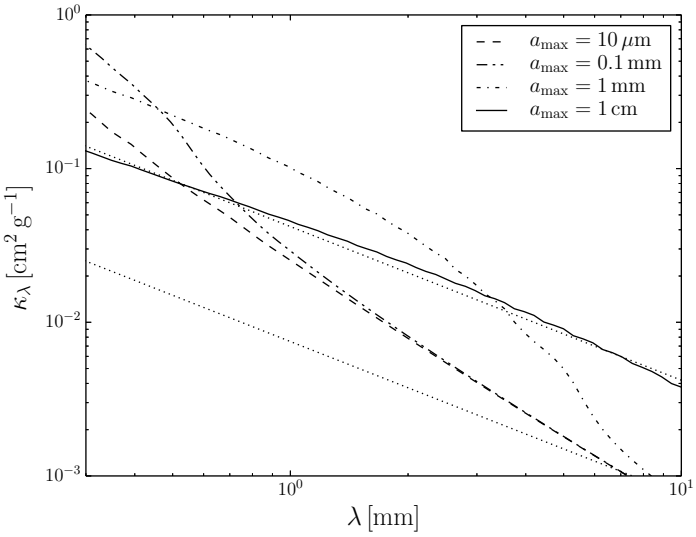

Figure 5.3: Dust opacity for $a_{\max}$ between $10\,\mu\mathrm{m} - 1\,\mathrm{cm}$. Note that the power-law assumption, $k_\lambda \propto \lambda^{-\beta}$ breaks down for $a_{\max} \gtrsim \lambda$. The upper dotted line plots a power-law fit to the opacity data adopted here (solid line) and the lower dotted line represents the opacity values used by Cossins et al. (2010).

consideration computed for a grain size distribution $n(a) \propto a^{-3.5}$ between $a_{\min} = 0.1\,\mu\mathrm{m}$ and different values of $a_{\max}$. We use opacity values computed with $a_{\max} = 1\,\mathrm{cm}$ (solid line) for all the disc in order to take into account the grain growth processing. It can be noted that the dust opacity values adopted here approximately follow a power law (upper dotted line in Fig. 5.3). In this respect, it is instructive to calculate the implied values of $\kappa_{1\mathrm{mm}}$ and $\beta_{1-3\mathrm{mm}}$ in this case: $\kappa_{1\mathrm{mm}} \approx 0.04\ \mathrm{cm^2 g^{-1}}$ and $\beta_{1-3\mathrm{mm}} \approx 1$. The value of β indicates that the dust opacities we use here take into account the grain growth. Note that Cossins et al. (2010) assume $\kappa_\lambda \approx 0.0075\,(\lambda/1\mathrm{mm})^{-1}\,\mathrm{cm^2 g^{-1}}$ (lower dotted line in Fig. 5.3), a factor ~ 5 smaller than ours.

5.3 The ALMA simulator

We use the ALMA simulator in the Common Astronomy Software Application (CASA) package (version 4.1). We used the version 4.1 of CASA, which includes simulation capabilities developed to simulate realistic ALMA observations over a wide range of wavelengths that would be measured by one of the full set of antenna configurations of ALMA. The simulation tool comes with 28 ALMA array configuration files containing the position of every antenna for each configuration. In addition, it is also possible to add the expected receiver noise and the contribution due to the atmosphere that affect the visibility measurements. Using the Atmospheric Transmission at Microwaves (ATM) code (Pardo et al. 2002), the task `simobserve` calculates signal atmospheric corruption effects into the visibilities measurements by evaluating the atmospheric profile starting from average site conditions and input values for the ground temperature and the Precipitable Water Vapour (PWV) provided by the user. Here, we assume a perfect calibration of the visibility measurements and thus only include the effects of thermal noise from receivers and the atmosphere. This implies that for the highest angular resolution and highest frequency maps presented here, the simulations may be optimistic. Nevertheless, at this stage, the experience with real ALMA observations on long baselines and high frequencies is extremely limited and it is difficult to assess how to properly account for these in the simulations.

Furthermore, having obtained the visibility measurements, the task `clean` operates a variety of algorithms in order to compute the deconvolved image from the visibility measurements. At the end of the process, it generates a cleaned image which represents the ALMA simulated observation (see Sect. B.2.1 for further details on the cleaning procedure).

Frequency (GHz)	PWV (mm)
45	2.748
100	2.748
220	2.748
345	1.262
680	0.472
870	0.472

Table 5.1: PWV values for each observing frequency.

5.3.1 Image analysis

The ALMA simulated observation can be analysed in terms of the ability to spatially resolve the spiral features of the circumstellar discs with an acceptable signal-to-noise ratio.

The noise of the cleaned image is obtained as the root mean square (rms) of the emission in regions where no signal is present, where we assume that the signal fluctuations are due to noise. Typically, it is used as a threshold for the signal detection and to calculate the signal-to-noise ratio. In this work, we use the typical minimum acceptable signal-to-noise ratio used for signal detection, i.e. equal to 3. It is worth remarking that each simulated observation has been performed with an appropriate choice of the antenna configuration and observation duration. The aim is to ensure enough spatial resolution and sensitivity to resolve and detect spiral arms in the intermediate and external region of the discs with an acceptable signal-to-noise ratio. In particular, having evaluated the noise level in the resulting simulated observation images, we generate contour plots at different multiples of the evaluated rms noise which allow us to demonstrate that the spiral structure are detectable with an acceptable signal-to-noise ratio.

5.4 Results

In this section, we present a sub-set of simulated ALMA observations of the discs previously described as they would be observed if they were located in four of the best-known and surveyed star-forming regions. In order to consider a wide range of distances and different sky positions, we locate the discs in the TW Hydrae ($\sim 50\,\mathrm{pc}$), Taurus - Auriga, Ophiucus ($\sim 140\,\mathrm{pc}$) and Orion ($\sim 400\,\mathrm{pc}$) star-forming regions. In this work, we use six observing frequencies chosen at the centre of some of the ALMA observation bands: 45 GHz in band 1, 100 GHz in band 3, 220 GHz in band 6, 345 GHz in band 7, 680 GHz in band 9 and 870 GHz in band 10. It is worth noting that, although the ALMA bands 1 and 10 are not yet available, we include these in our simulations as they are currently being deployed on the array (band 10) or in an advanced status of prototype development in view of full production (band 1). The array configuration used for each simulated observation varies for each frequency with the goal to use the configuration that offers the best compromise between resolution and sensitivity. For the atmospheric conditions of the observations, as discussed previously, we neglect the effect of signal decorrelation, assuming excellent calibration, but we include the effect of the atmospheric optical depth and emission parametrized through the amount of PWV. The values of the PWV that we use (see Table 5.1) are typical of the maximum matching constraints used at the observatory for the frequencies of our choice. The choice of the PWV values does not play a crucial role at low frequencies because the atmospheric transmission is always very high at the centre of the band within the selected range of PWV. However, at higher frequencies, the PWV has to be carefully chosen due to the strong variation of the atmospheric noise with the PWV value. Thus, in real observation at higher frequencies (> 200 GHz), excellent observing conditions may be required for significant detection. We note that the values of the PWV assumed at 680 and 870 GHz are excellent, but better conditions are expected at Chajnantor for 15% of the time.

Disc model parameters				ALMA configurations						Duration	Total fluxes (mJy)					
q	$M_\star(\mathrm{M}_\odot)$	$R_{\rm out}$(au)	Incl.(°)	45	100	220	345	680	870	(min)	45	100	220	345	680	870
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)
			0								6	37	158	325	752	717
0.25	0.3	25	45	28	28	25	22	20	18	30	3	20	88	172	420	403
			65								1	7	32	59	162	151
			0								26	139	647	1466	4595	5837
0.25	1	25	45	28	28	25	22	20	18	10	13	74	341	788	2531	3357
			65								4	27	125	274	932	1244
			0								85	454	2200	4903	16186	21284
0.25	3	25	45	28	28	25	22	20	18	10	44	239	1092	2558	8890	13167
			65								16	85	398	902	3276	4998
0.1	3	25	0	28	28	27	24	22	20	120	14	71	364	749	1232	2529
			45								7	37	183	346	704	1490
			0								50	306	1324	3173	5745	7317
0.25	1	100	45	27	22	18	15	11	9	30	28	169	712	1705	2993	4703
			65								12	68	282	645	1229	2045
			0								228	1352	5553	14257	24417	30786
0.25	3	100	45	27	22	18	15	11	9	10	138	655	2733	7417	13751	17692
			65								62	277	1151	2810	6913	9907

Table 5.2: Disc model, observation parameters and total resulting fluxes for the simulated observations of discs located in TW Hydrae star-forming region. Column (1): disc-to-star mass ratio. Column (2): stellar mass. Column (3): outer radius. Column (4): disc inclination. Columns (5)-(10): ALMA configuration (see `http://casa.nrao.edu/` for details) for the six observing frequencies (expressed in GHz). Column (11): transit duration. Columns (12)-(17): total resulting flux at the various frequencies (expressed in GHz).

5.4.1 Simulated ALMA images

In this subsection we present different simulations as a function of the frequency, disc-to-star mass ratio, stellar mass, inclination and coordinates of the discs. We analyse how the different values of the parameters mentioned above affects the detection of the spiral structures using ALMA. As discussed above, there are several physical and observational parameters that determine the appearance of the discs with ALMA. We have selected some specific configurations and parameter choices that best show the possible outcome. The various parameters adopted in the images presented here are summarized in Tables 5.2, 5.3 and 5.4, referred to discs located in TW Hydrae, Taurus-Auriga and Ophiucus star-forming region, respectively. In particular, the first four columns specify the physical parameters of the disc (disc-to-star mass ratio, stellar mass, outer radius and inclination), columns (5)-(10) specify the ALMA configuration for the six bands that we have considered (where the number indicates the CASA identifier for the configuration of the various antennas, as described in `http://casa.nrao.edu/`), column (11) indicates the transit duration for the simulated image, while columns (12)-(17) show the total resulting flux at the various frequencies (expressed in GHz).

First, we show a representative set of ALMA images for all the observing frequencies of a disc with a given set of model parameters. In Fig. 5.4 we show the result of our simulations for non face-on discs (inclination angle of 45°) of radius 100 au located in TW Hydrae for a stellar mass of $1\,\mathrm{M}_\odot$ and disc-to-star mass ratio $q = 0.25$ for the six selected observing frequencies (the original density structure of the disc is shown in Fig. 5.1, right panel). In order to reach an

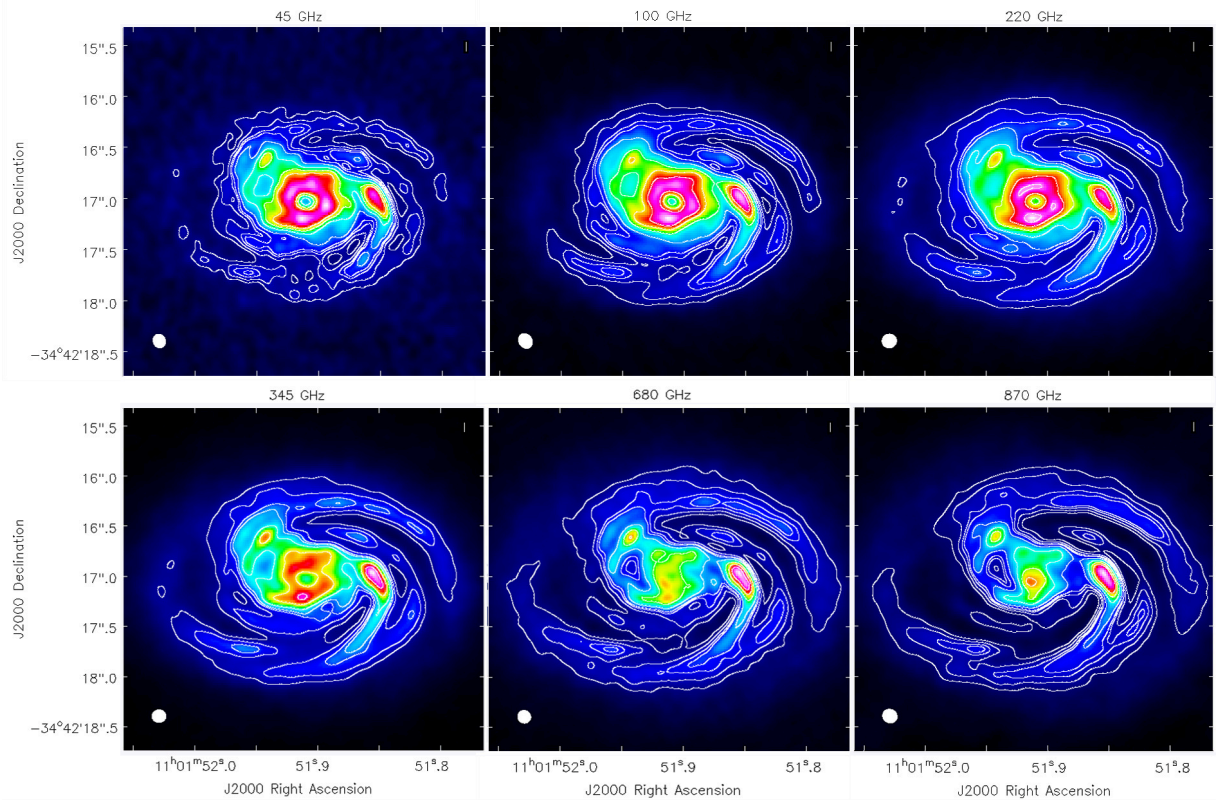

Figure 5.4: ALMA images for the disc of radius 100 au with mass equal to $0.25\,M_\odot$ and a disc-to-star mass ratio $q = 0.25$ located in TW Hydrae. Contours are 7, 10, 13, 15, 20, 30, 40, 50 $\times$ the rms noise at 45 GHz, 100 and 220 GHz (7, 18, 27 μJy/beam); 4, 7, 10, 13, 15, 20, 30, 40 $\times$ the rms noise at 345, 680 and 870 GHz (42, 260, 400 μJy/beam). The white colour in the filled ellipse in the lower left corner indicates the size of the half-power contour of the synthesized beam: $6.6'' \times 6.1''$ (45 GHz), $7.2'' \times 6''$ (100 GHz), $7.3'' \times 6.7''$ (220 GHz), $7'' \times 6.2''$ (345 GHz), $6.5'' \times 6.2''$ (680 GHz) and $7.3'' \times 6.7''$ (870 GHz).

optimal signal-to-noise ratio, the ALMA observations are simulated with a transit duration of 30 min. The resolution and sensitivity of ALMA are amply sufficient to spatially resolve and detect the substructures of the simulated discs over a wide range of wavelengths (confirming and extending the results of Cossins et al. 2010).

By analysing the ALMA images at different frequencies, it can be noted that the relative intensity between arm and interarm regions varies with frequency. In this respect, there are essentially two important effects to consider in order to specify the trend of the contrast between arm and interarm regions with frequency: the optical depth and the disc temperature. To illustrate the effects of varying the frequency, in Fig. 5.5 is shown the optical depth at frequencies of 100 and 345 GHz for the disc under consideration. It can be noted that, while the dense regions corresponding to spiral arms are generally optically thick, the interarm regions transition from being optically thin at frequencies below ~ 345 GHz, to optically thick above 345 GHz. Thus, for $\nu \lesssim 345$ GHz, the contrast between arm and interarm regions tends to decrease with increasing frequency. On the contrary, at higher frequencies ($\gtrsim 345$ GHz), the contrast increases (see Fig. 5.4) because of a temperature effect. In the interarm regions the temperature reaches values low enough to ensure that the Planck law is in the Wien limit for frequencies greater than 345 GHz. For this reason, the contrast between arm and interarm regions increases with frequency in the range 345 - 870 GHz which makes it easier to detect the spiral features. This characteristic trend of contrast between arm and interarm regions with frequency can be observed also for compact discs (outer radius of 25 au) with a stellar mass of $0.3\,M_\odot$ and $q = 0.25$. Other discs in our sample are characterized by a temperature high enough to ensure that the Planck law is in the Rayleigh-Jeans limit. In those cases, the variation in intensity with increasing frequency is only due to the different optical regimes in different parts of the disc.

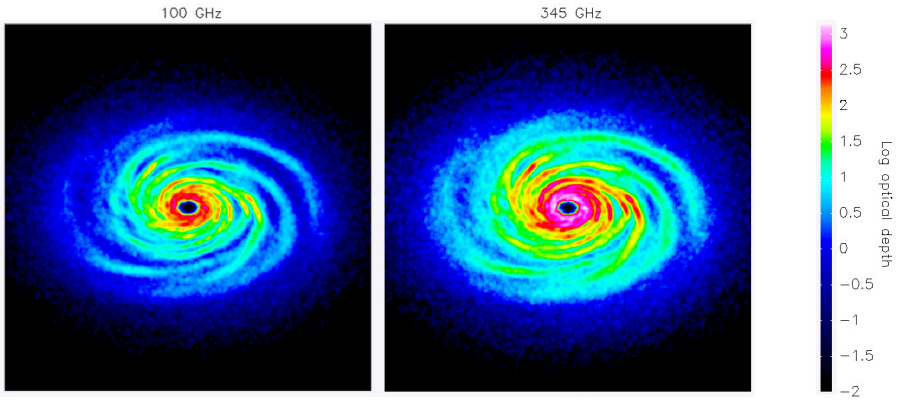

Figure 5.5: Comparison of the optical depth across the disc face at frequencies of 100 GHz (right) and 345 GHz (right) for the disc of radius 100 au with mass equal to $0.25\,M_\odot$ and a disc-to-star mass ratio $q = 0.25$.

There are two interesting features that can be seen from the morphology apparent in the ALMA images. First, note that, while the original density structure of the disc is characterized by a relatively large number of arms, only two spiral arms are visible in the images, because the smaller-sized arms have been washed out by the limited resolution of the observations. Care should thus be taken when interpreting observed morphologies: for example, a clean two-armed spiral is often interpreted as the natural effect of a tidal interaction with a companion star or massive planets, while we show here that it can be easily reproduced in the context of gravitational instability, even though the underlying disc structure is more complex than a two-armed spiral.

Additionally, one can also note a clean clump on the right of the star, along the spiral arm. Now, when comparing the location of the clump with the original density structure, one can notice that it actually corresponds to a mild overdensity at the intersection of two arms. Note that the original simulation did not have any clump, and in fact it is a non fragmenting simulation. The detection of an image such as that in Fig. 5.4 might be easily mistaken as evidence of fragmentation, while in fact the clump is an artefact resulting from the observational limitations, that in this case overemphasize a mild overdensity in the disc.

5.4.1.1 Changing the disc-to-star mass ratio

As previously mentioned (see Section 5.1.1), discs with different disc-to-star mass ratio are characterized by a different surface density structure. The number of arms increases and the spiral structure tends to become more close with decreasing disc-to-star mass ratio making it therefore more difficult to detect the spiral features (see Sect. 4.3). Thus, in order to test the ALMA capabilities in detecting the spiral structure with different intensity and morphology, we show in Fig. 5.6 the ALMA images at 220 GHz of two non face-on (inclination angle of 45°) compact discs (outer radius of 25 au) located in TW Hydrae around a star of $3\,M_\odot$ with two different disc-to-star mass ratio: $q = 0.25$ (left) and $q = 0.1$ (right). As previously mentioned, the temperature profile is closely linked to the mass of the disc (see the left panel of Fig. 5.2). Therefore, for a cooler disc the intensity is lower and we thus need a long observation time to detect spiral features with an appropriate signal-to-noise ratio. Thus, we set the observation time equal to 10 min for the discs with $q = 0.25$ and 2 h for the other one (see Table 5.2). It can be noted that, using an appropriate choice of input parameters, ALMA provides a spatial resolution sufficient to detect the spiral structure with different arm sizes with an acceptable signal-to-noise ratio, even for a more tightly wound spiral.

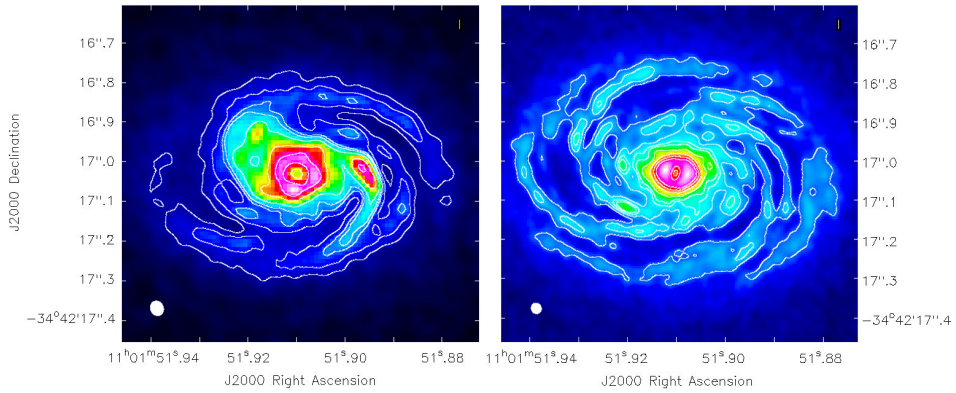

Figure 5.6: ALMA images at 220 GHz of two non face-on (inclination angle of 45°) compact discs (outer radius of 25 au) located in TW Hydrae around a star of $3\,M_\odot$ with two different disc-to-star mass ratio: $q = 0.25$ (left) and $q = 0.1$ (right). Contours are at different multiples (> 4) of the evaluated rms noise: $38\,\mu$Jy/beam (left) and $10\,\mu$Jy/beam (right). The white colour in the filled ellipse in the lower left corner indicates the size of the half-power contour of the synthesized beam: $0.033'' \times 0.033''$ (left) and $0.027'' \times 0.025''$ (right).

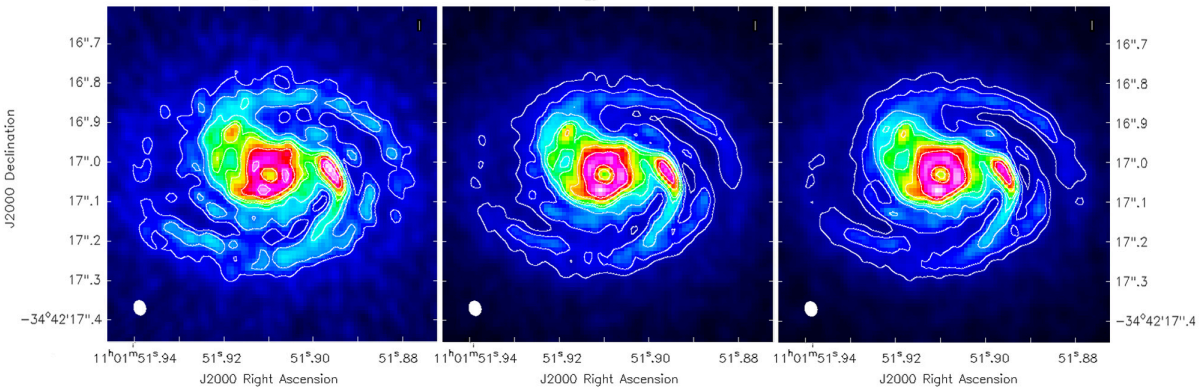

Figure 5.7: ALMA images at 220 GHz for compact discs (outer radius of 25 au) with different star and disc mass located in TW Hydrae. The three images are related to discs with $q = 0.25$ with stellar mass of $0.3\,M_\odot$ (left), $1\,M_\odot$ (centre) and $3\,M_\odot$ (right). Contours are at different multiples (> 4) of the evaluated rms noise: $21\,\mu$Jy/beam (left), $40\,\mu$Jy/beam (centre) and $38\,\mu$Jy/beam (right). The white colour in the filled ellipse in the lower left corner indicates the size of the half-power contour of the synthesized beam: $0.033'' \times 0.033''$.

5.4.1.2 Changing the stellar mass

Here, we show ALMA images for different values of stellar mass for a given value of the disc-to-star mass ratio. In Fig. 5.7 we show the predicted observability of all the compact discs (outer radius of 25 au) in our sample with $q = 0.25$ at the frequency of 220 GHz and located in TW Hydrae star-forming region. The three panels are related to non face-on discs (inclination of 45°) with stellar mass of $0.3\,M_\odot$ (left), $1\,M_\odot$ (centre) and $3\,M_\odot$ (right). Taking into account the required sensitivity for detection, the ALMA observations are performed with the same choice of antenna configuration and a transit duration of 30 min for the disc with stellar mass $0.3\,M_\odot$, while for more massive star, due to the greater intensity, the observation time is fixed to 10 min. Moreover, comparing the total resulting flux of discs with 1 and 3 $M_\odot$ (see Table 5.2), note that, as expected, the flux increases with the disc (stellar) mass. In addition, it can be noted that, due to the lower emission of the less massive disc (left-hand panel in Fig. 5.7), the signal-to-noise ratio on the spiral arms is lower with respect to other cases. However, all the simulations demonstrate that resolution and sensitivity of ALMA are amply sufficient to spatially resolve and detect the substructures of the simulated discs with an appropriate choice of observation parameters.

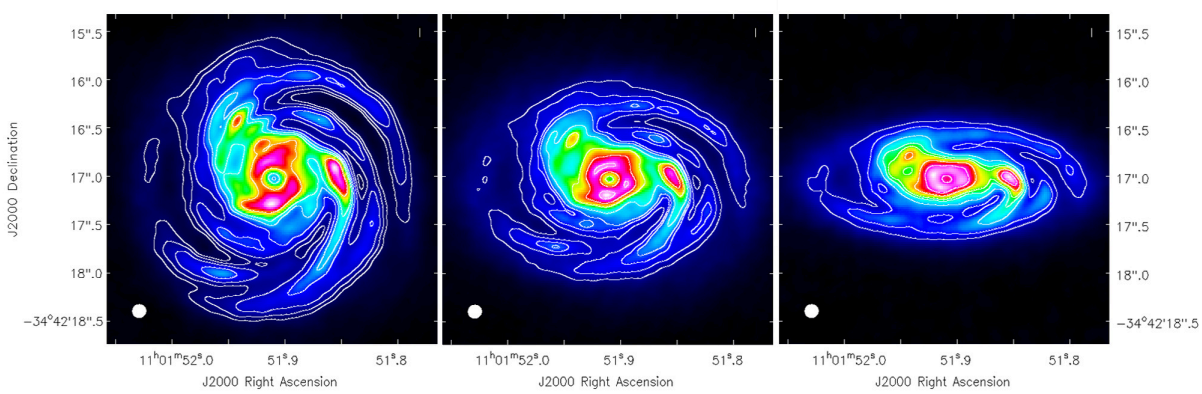

Figure 5.8: ALMA images at 220 GHz for larger discs (outer radius of 100 au) and central mass star of $1\,M_\odot$ with $q = 0.25$ and with three values of inclination angle: 0°(left), 45°(centre) and 65°(right). Contours are at different multiples (> 4) of the evaluated rms noise: $\sim 27\,\mu$Jy/beam. The white colour in the filled ellipse in the lower left corner indicates the size of the half-power contour of the synthesized beam: $0.147'' \times 0.135''$.

Disc model parameters				ALMA configurations						Duration	Total fluxes (mJy)					
q	$M_\star(M_\odot)$	R_{out}(au)	Incl.(°)	45	100	220	345	680	870	(min)	45	100	220	345	680	870
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)
0.25	1	25	0	28	28	28	26	22	21	30	3	18	85	203	708	1087
			45								2	10	46	112	373	571
0.25	1	100	0	28	26	21	20	15	13	60	6	46	209	394	787	1225
			45								4	25	120	212	398	645

Table 5.3: Disc model, observation parameters and total resulting fluxes for the simulated observations of discs located in Taurus - Auriga star-forming region. See Table 5.2 for details.

5.4.1.3 Changing the disc inclination

In addition, for what concerns the inclination angle of circumstellar discs, it is expected that the spiral features of the disc becomes less prominent with increasing angle of inclination. This effect makes it more difficult to observe the spiral structures of the accretion disc. We show in Fig. 5.8 ALMA images at 220 GHz of the larger disc (outer radius of 100 au) in our sample with a central mass star of $1\,M_\odot$, $q = 0.25$ and with three values of inclination angle: 0° (left), 45° (centre) and 65° (right). It can be noted that, using the same antenna configuration, the number of the detected spiral arms decreases with inclination angle. As expected, the contrast between arm and interarm regions decreases with increasing angle. In addition, we have also tested the limit of ALMA capability to spatially resolve the spiral pattern. We find that for discs with inclination angle >75°, the detected separation between arm and interarm regions vanishes. Assuming that protostellar discs are characterized by a uniform distribution in inclination angle, we conclude that the detection of spiral structure is acceptable for ~ 80 per cent of discs.

5.4.1.4 Changing the disc location

Finally, in order to create a full 'atlas' of simulated ALMA images of self-gravitating accretion discs, it is necessary to evaluate the detection ability of ALMA for discs located in different star-forming regions. To this aim, we simulate ALMA observations of a sub-set of discs in our sample located in TW Hydrae, Taurus - Auriga, Ophiucus and Orion star-forming regions. In Fig. 5.9 are shown the predicted images at the frequency of 220 GHz of all the larger discs (outer radius of 100 au) in our sample with a stellar mass of $1M_\odot$, $q = 0.25$ with an inclination angle

Disc model parameters				ALMA configurations						Duration	Total fluxes (mJy)					
q	$M_\star(M_\odot)$	R_{out}(au)	Incl.(°)	45	100	220	345	680	870	(min)	45	100	220	345	680	870
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)
0.25	1	25	0	28	28	28	26	22	21	30	4	21	112	232	810	1121
			45								2	12	55	141	417	594
0.25	1	100	0	28	26	21	20	15	13	60	7	55	223	452	812	1266
			45								4	28	127	257	435	673

Table 5.4: Disc model, observation parameters and total resulting fluxes for the simulated observations of discs located in Ophiucus star-forming region. See Table 5.2 for details.

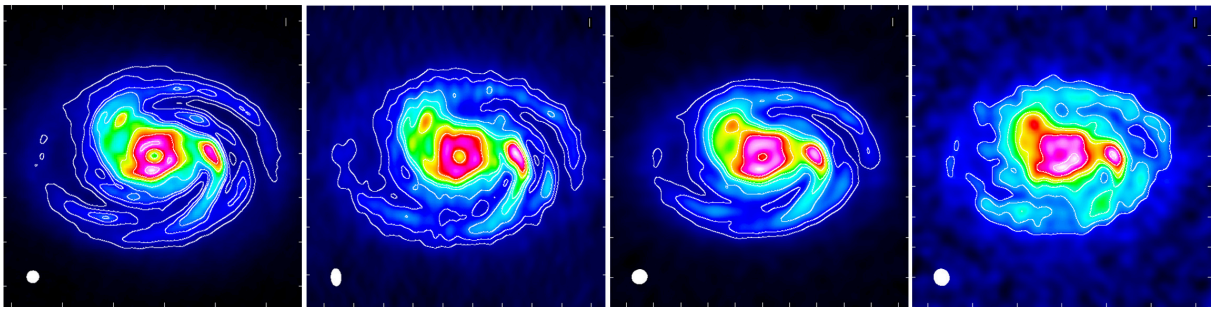

Figure 5.9: ALMA images at 220 GHz for larger discs (outer radius of 100 au) and central mass star of $1M_\odot$ with $q = 0.25$. From left to right we show observations for discs located in TW Hydrae, Taurus - Auriga, Ophiucus and Orion star-forming region. Contours are at different multiples (> 4) of the evaluated rms noise: 27, 20, 15 and 10 μJy/beam. The white colour in the filled ellipse in the lower left corner indicates the size of the half-power contour of the synthesized beam: $0.147'' \times 0.135''$, $0.070'' \times 0.044''$, $0.077'' \times 0.070''$ and $0.025'' \times 0.010''$.

of 45° located in the four previously mentioned star-forming regions.

The first panel in Fig. 5.9 refers to the case of the TW Hydrae star-forming region. Due to its closeness to the Solar system (~ 50 pc), the resolution capabilities of ALMA are largely sufficient to spatially resolve the substructures for all the discs in the sample over the whole range of observing wavelengths.

The second and the third panel are related to discs in Taurus - Auriga and Ophiucus (~ 140 pc) star-forming regions. In these cases, the ALMA observations are performed with a transit duration of one hour. Since these regions are located at nearly the same distance, it is instructive to compare the image reconstruction capabilities of ALMA for disc located in different sky regions with the same choice of observation parameters (see Tables 5.3 and 5.4). For what concerns the case of discs located in Taurus - Auriga, it can be noted that the sky position of this source (declination: $\delta = +30°$) during all the observation duration, produces two important effects involving atmospheric noise and spatial resolution. First, the observation of discs located in this star-forming region is more corrupted by atmospheric noise than other cases. This occurs because of the greater atmospheric emission at lower elevations. As a consequence, the evaluated rms noise of the ALMA images is greater than what obtained when the source is at lower declination observed with the same input parameter (see the third panel in Fig. 5.9). In addition, the atmospheric noise is enhanced especially for higher frequency where the atmospheric transmission plays a crucial role for the quality of observations. As can be seen by comparing the total resulting fluxes for the same discs located in the two considered star-forming regions (see Tables 5.3 and 5.4), the effect of the atmospheric noise is essentially a decrease of the detected signal. The second consequence of the particular source position is related to the spatial resolution provided by an antenna ALMA configurations. The angular resolution of an interferometer is determined by the length of the baselines projected on to

a plane normal to the direction of the source. Given that the source elevation is low for all the transit duration, the spatial frequency plane is characterized by an irregular coverage. For this reason, the synthesized beam is characterized by an half-power contour with two very different values of minor and major axes. As a consequence, the spatial resolution sufficient to resolve spiral features can be reached only in one direction. Comparing the second and the third panels of Fig. 5.9, it is clearly visible that ALMA provides a better reconstruction of image in terms of spatial resolution and signal-to-noise ratio for the disc located in Ophiucus (declination: $\delta = -24°$). It is worth remarking that the ALMA images for the discs in the two different regions have been carried out using the same antenna configurations and the same total observation durations. Furthermore, comparing the rms noise at each frequency for the same discs located in the two previously mentioned star-forming regions, we find, as expected, a higher rms for the lower elevation sources, especially for intermediate and higher frequencies.

Finally, the fourth panel of Fig. 5.9 refers to discs in the Orion star-forming region ($\sim 400\,\mathrm{pc}$). We find that the spiral structure for larger discs is detectable at frequencies greater than 100 GHz with an acceptable signal-to-noise ratio. In addition, ALMA observation of the disc of radius 25 au (not presented here) show that the spiral structure is not well resolved at low and intermediate frequencies (45 - 345 GHz), whereas at higher frequencies, as noted above, the simulation are probably over-optimistic. Furthermore, the antenna configuration used to simulate ALMA observation for frequencies lower than 680 GHz are characterized by the longest baseline available. As a consequence, calibration uncertainties and residual phase noise play a crucial role in a realistic image reconstruction. Therefore, the spiral structure on compact discs in the sample is not conclusively detectable using ALMA.

5.5 Conclusion

In this Chapter we have shown simulated ALMA observations of a variety of non face-on self-gravitating circumstellar discs models with different properties in mass and radial extension with the aim to demonstrate that the peculiar spiral structure should be detectable using ALMA.

We conclude that, using a careful choice of observation parameters such as antenna configuration and observation duration, for discs at distances comparable to TW Hydrae ($\sim 50\,\mathrm{pc}$), Taurus - Auriga and Ophiucus ($\sim 140\,\mathrm{pc}$) star-forming regions, the spiral structure is readily detectable by ALMA over a wide range of wavelengths. However, for discs located in Orion complex ($\sim 400\,\mathrm{pc}$), only the largest discs in the parameter ranges we have explored could be spatially resolved while the smaller ones are characterized by a spiral structure that should not be detectable with ALMA. Our simulations also confirm the well-known fact that great care should be taken when interpreting morphological features, such as clumps, in interferometric observations of complex sources. Indeed, in our simulated observations, an apparent, well-detectable clump is only an artefact due to finite resolution effects, where the underlying disc structure only has a mild and transient density enhancement due to the intersection of two spiral arms.

Spiral and non-axisymmetric structures in protostellar discs have started to be observed relatively frequently in the past few years (see Sect. 4.3.2). In particular, several Class 0 sources, that are more likely to be subject to gravitational instabilities (Rodríguez et al. 2005; Eisner 2012), have been detected to show non-axisymmetric features. For example, Rodríguez et al. (2005) have presented mm observations of IRAS 16293-2422B, a Class 0 object located in the Ophiuchus molecular complex, revealing that the gravitational instability might be the origin of the non axisymmetric structures in the outer disc region. These papers, however, did not fully separate the disc and the envelope contribution to the emission. Recently, Miotello et al. (2014b) have reported mm observations of Class 0/I sources where the disc and envelope contribution could be identified. In some cases, the resulting disc appeared to be massive enough to be gravitationally unstable. We should note, however, that the envelope contribution has not been included in our modelling.

Additionally, as described in Sect. 4.3.2, spiral structures has also been observed recently in transitional discs by imaging the distribution of scattered light at near-infrared wavelengths

(Muto et al. 2012; Grady et al. 2013; Garufi et al. 2013) and by high-resolution ALMA observations of molecular line emission (Christiaens et al. 2014). For such evolved discs, one might not expect the disc to be massive enough to be self-gravitating, and thus the more common explanation for the origin of the spiral is the dynamical interaction with a third body, like a companion star or an embedded planet. However, note that Fukagawa et al. (2013) and Christiaens et al. (2014) have shown that the outer regions of the transitional disc of the Herbig Fe star HD 142527 is very close to the instability regime, which might indicate that the self-gravitating phase might in some cases last well into the Class II phase.

Moreover, the simplicistic cooling prescription used in the SPH simulations does not take into account the irradiation emitted by the central star. The effect of including irradiation is essentially a decrease in the amplitude of the spiral structure (and thus of the arm-interarm density contrast), which means that our simulations can be somewhat optimistic for strongly irradiated discs.

Furthermore, one of the assumptions of the emission model states that the gas and dust component of the disc are perfectly mixed within the disc. In addition, the dust opacities have been computed using the same level of the grain growth in the whole disc. In this respect, it is thought that the spiral structure induced by the development of gravitational instabilities affects both the radial migration and the grain growth of the dust grains. Since the gas density presents local maxima corresponding to the spiral arms, the radial motion of the dust particles is affected (Rice et al. 2004, 2006b, see Chapter 6). In addition, the enhanced dust density favors and speeds up collisional growth leading the formation of kilometre-sized planetesimal that can easily collapse and form gravitationally bound larger objects (Rice et al. 2006b). We will explore the dynamics of non-growing dust grains in gravitationally unstable discs in the next Chapter.

Chapter 6

Dust Trapping by Spiral Arms in Gravitationally Unstable Protostellar Discs

Based on the paper by ***Giovanni Dipierro****, Paola Pinilla, Giuseppe Lodato and Leonardo Testi: "Dust trapping by spiral arms in gravitationally unstable protostellar discs" published in Monthly Notices of the Royal Astronomical Society,* ***451****, 974-986 (2015).*

According to the most widely accepted scenario of planet formation, the "core accretion" model, in the high-density environment of a circumstellar disc, micron sized dust grains are expected to grow via collisional agglomeration to kilometre-sized planetesimals (Testi et al. 2014). In recent years, the investigation on the dust evolution in protostellar discs has been the subject of several theoretical (e.g. Brauer et al. 2008; Birnstiel et al. 2010) and laboratory studies (e.g. Blum 2006; Blum & Wurm 2008) aimed at understanding dust dynamics and its growth. The general picture that emerges from these investigations is that the growth from submicron sized particles to larger objects is a complex process that contains many physical challenges (see Sect. 2.5, Zsom et al. 2010; Okuzumi et al. 2011; Birnstiel et al. 2010).

In this Chapter we discuss the influence of the spiral density waves induced by gravitational instabilities on the dust dynamics. As already remarked, create long-lived radial and/or azimuthal inhomogeneites in the gas density structure in order to "trap" particles into the overpressure regions, solving the problem of the radial drift and fragmentation barrier (Haghighipour & Boss 2003b,a; Pinilla et al. 2012b). Recent work has already suggested that grain growth may be accelerated by concentrating the grains in the peak of the self-gravitating structures, possibly overcoming the issue of the rapid grain migration (Rice et al. 2004, 2006b; Gibbons et al. 2012, 2014). Moreover, the dust retention in spiral arms implies low relative velocities between particles, which suggests that collisions are more likely to be constructive (Gibbons et al. 2012, 2014).

In addition, the dust migration towards spiral arms can produce macroscopic features that can be identified through high angular resolution observations (see Sect. 4.3). In the last Chapter (see also Cossins et al. 2010) we have found that, assuming that the dust is perfectly mixed with the gas, spiral structures of a variety of non-face-on self-gravitating circumstellar discs models with different properties in mass and radial extension are readily detectable by ALMA over a wide range of wavelengths. One of the aims of the present work is to improve the observational predictions of self-gravitating circumstellar discs presented in Cossins et al. (2010) and Dipierro et al. (2014) (see Chapter 5) by including the treatment of dust dynamics (in particular, by relaxing the hypothesis that gas and dust are perfectly mixed) and a more detailed procedure for the generation of emission maps.

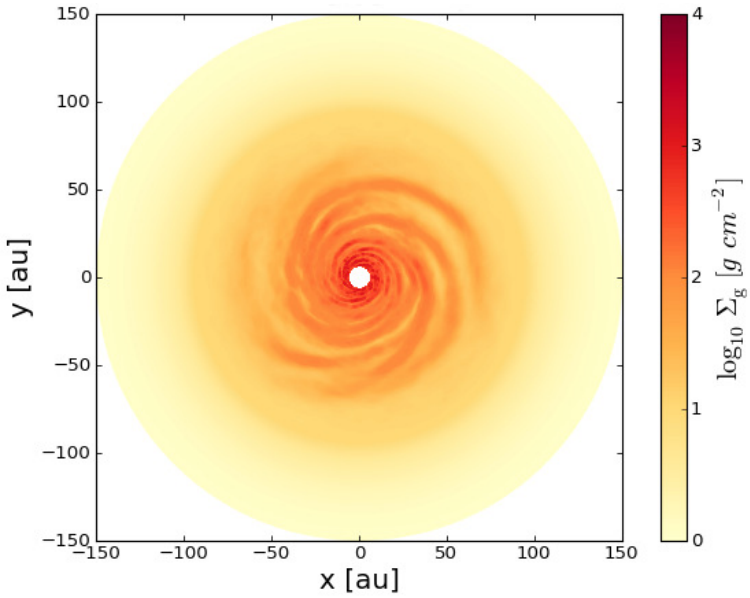

Figure 6.1: Gas surface density structure of a $0.25 M_\odot$ disc around a $1 M_\odot$ star that has reached a long-lived, self-gravitating state.

In this Chapter, by combining the results of hydrodynamic dust and gas simulations with 3D Monte Carlo radiative transfer calculations (using RADMC-3D), we study if the occurrence of local pressure maxima induced by gravitational instability can trap dust particles and investigate the detectability of these inhomogeneites at near-infrared and (sub-)millimetre wavelengths by the High-Contrast Coronographic Imager for Adaptive Optics (HiCIAO) and ALMA, respectively.

6.1 Gas and dust disc models

The simulations presented here calculate the evolution of the dust surface density in a quasi-static gaseous self-gravitating disc. The system comprises two components discs: a gas disc, characterized by a spiral surface density structure induced by gravitational instability, and a dusty disc that evolves under the action of drag forces and turbulent mixing induced by the gas-dust aerodynamical coupling. The self-gravitating gas disc model has been taken from a snapshot of a three-dimensional Smoothed Particle Hydrodynamics (SPH) simulation performed by Lodato & Rice (2004). This model can be considered as representative of a self-gravitating gas disc in a quasi-steady state characterized by a spiral surface density structure produced by a combined effect of self-gravity and cooling. We compute the dust surface density evolution in such fixed gaseous background in the radial direction along different azimuthal cuts across the disc. Afterwards, starting from the resulting surface density structure we create 3D models using a representative estimator for the tickness of the dusty disc. We carry out 3D radiative transfer simulations in order to create the expected emission maps at (sub-)millimetre and near-infrared wavelengths of our models which are used to simulate realistic observations.

6.1.1 Gas disc model

The gas disc model used as an input for the dust simulations has been taken from Lodato & Rice (2004) and has already been described in Section 5.1. In order to consider a gas disc model with realistic properties in mass and radial extension, we modify the length and the mass scale of the data output of the SPH simulations such that the radial extent of the disc is 150 au with a disc mass equal to $0.25\,M_\odot$. Once the thermal equilibrium state is reached, the gas disc is characterized by a spiral structure (shown in Fig. 6.1) where the heating generated through gravitational instability balances the external cooling. This structure remains quasi-steady,

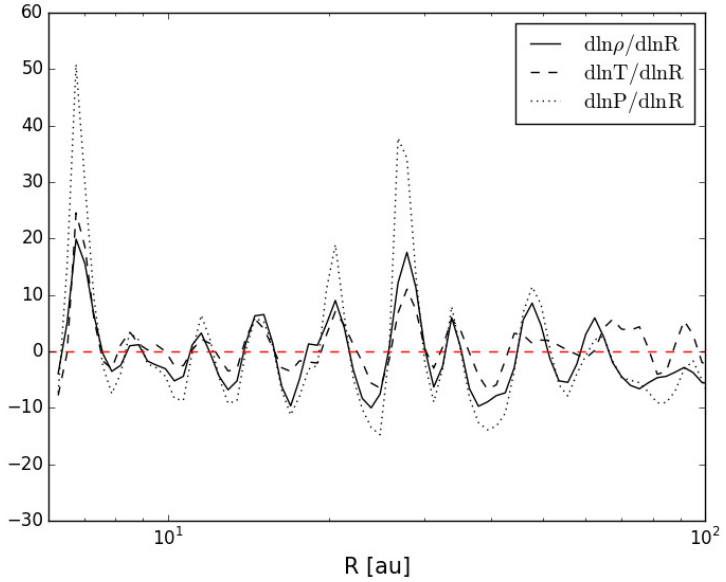

Figure 6.2: Gas volume density, temperature and pressure gradient along a slice across the disc.

evolving secularly on the viscous timescale (10^5-10^6 yrs in this case). Individual spiral arms are recurrent structures evolving on the dynamical timescale ($\sim 10^3$ yrs in the outer disc).

Fig. 6.2 shows the volume density, temperature and pressure gradient along a slice across the disc. It can be noticed that all the gradients are characterized by a similar trend and are null at the same radial positions. Therefore, the overdensity regions, which correspond to the arm regions, are characterized by relatively higher temperatures while in the interarm region the temperature is lower (see also the right panel of Fig. 5.2). This is due to the increase in the pressure caused by a shock driving a compression of gas and, as a consequence, a growth of the temperature. In the interarm regions, the increase in the pressure provides an expansion of gas parallel to the shock and, as a consequence, a decrease of the temperature (Cossins et al. 2009). Moreover, since the pressure depends on the volume density and the sound speed, the peaks of the pressure gradient are higher than those related to the temperature and volume density profile.

6.1.2 Dust evolution model

To describe the evolution of the dust due to drag forces and turbulent diffusion it is convenient to define how well coupled the particles are to the gas. The gas-dust aerodynamical coupling is expressed by the Stokes number, usually denoted by St, defined as the ratio of the stopping time of a particle and the local dynamical timescale (Ω^{-1}). Generally, as described in Chapter 2, there are four different regimes for the dust-to-gas aerodynamical coupling depending on the value of the Reynolds number and the ratio between mean free path $\lambda_{\rm mfp}$ of the gas molecules and the dust particle size, a. In the Epstein regime, for $\lambda_{\rm mfp}/a \geq 4/9$, the Stokes number for particles close to the disc mid-plane, is given by (see Eq. 2.89)

$$\mathrm{St} = \frac{a\rho_{\rm p}}{\Sigma_{\rm g}}\frac{\pi}{2}, \tag{6.1}$$

where $\Sigma_{\rm g}$ is the gas surface density and $\rho_{\rm p}$ is the internal density of dust grains. The radial velocity of dust particles can be determined by self-consistently solving the radial and azimuthal components of the momentum equation for low dust to gas ratio (Eq. 2.39, Weidenschilling 1977a)

$$v_{{\rm d},R} = v_{\rm drag} + v_{\rm drift} = \frac{v_{{\rm g},R}}{1+\mathrm{St}^2} + \frac{1}{\mathrm{St}+\mathrm{St}^{-1}}\frac{1}{\rho_{\rm g}\Omega}\frac{\partial P}{\partial R}, \tag{6.2}$$

where Ω is the Keplerian angular velocity, P is the pressure and ρ_g is the gas volume density. The first term is the gas drag term which is produced by the coupling of the dust particle to gas moving with a radial velocity of $v_{g,R}$ given by viscous evolution (see Eq. 1.25). The second term is the radial drift caused by the angular momentum exchange between dust particles and the non-Keplerian orbital motion of the gas. If the Stokes number is much smaller than unity, the dust particle is strongly coupled to the gas, which implies that the gas and dust share the same motion. On the other hand, for St >> 1, the particles are unaffected by the gas and consequently move on Keplerian orbits to a first approximation. Particles in the intermediate size range, i.e. St $\sim$ 1, have large drift velocities and are expected to experience the highest concentration in pressure maxima.

Since the gas is expected to be turbulent, the dust is mixed by the gas due to aerodynamical coupling (see Sect. 2.4). The turbulent radial transport depends on the gas radial diffusion, which is usually described by the canonical turbulence prescription (Shakura & Sunyaev 1973). The dust diffusivity induced by the gas-dust aerodynamical coupling is defined as (Eq. 2.93, Youdin & Lithwick 2007)

$$D_{d,R} = \frac{D_{g,R}}{1+\mathrm{St}^2} = \frac{\alpha \langle c_s \rangle H_g}{1+\mathrm{St}^2}, \tag{6.3}$$

where H_g is the gas scale height, $\langle c_s \rangle$ is the azimuthal average of the sound speed and α is an unknown dimensionless parameter which embodies all the uncertainties about the nature of the diffusion mechanism (see Sect. 1.3.2). Since the diffusion time scale is longer than the lifetime of individual spiral features in self-gravitating discs ($\sim \Omega^{-1}$), the dust diffusivity is computed using the azimuthal average of the sound speed.

Generally, it is expected that a gravitationally unstable disc cannot be treated as a viscous disc because gravity can affect the transport mechanism over long distances. In this context, a number of numerical simulations have been recently carried out with the aim to test the locality of transport provided by gravitational instability. Lodato & Rice (2004), Cossins et al. (2009) and Forgan et al. (2011), using an α-like approach, have compared the dissipation power provided by gravitational instability with the expectations based on a viscous theory of disc. They have found that for disc to star mass ratio $M_{disc}/M_\star \lesssim 0.25$, the gravitational instability is well described using a viscous approach.

Encompassing all the systematic and random motions described above, the time evolution for the dust surface density Σ_d^i of the i^{th} dust species with size a^i can be described by an advection-diffusion equation (see Eq. 2.94, Brauer et al. 2008; Birnstiel et al. 2010)

$$\frac{\partial \Sigma_d^i}{\partial t} + \frac{1}{R}\frac{\partial}{\partial R}\left\{ R\left[\Sigma_d^i\, v_{d,R}^i - D_{d,R}^i\, \Sigma_g \frac{\partial}{\partial R}\left(\frac{\Sigma_d^i}{\Sigma_g}\right)\right]\right\} = 0, \tag{6.4}$$

where, as previously described, the diffusion coefficient (see Eq. 6.3) and the radial velocity (see Eq. 6.2) depends on the Stokes number and therefore on the particle size.

In this work, the dust evolution for each grain size in our sample is computed using the numerical method adopted in Birnstiel et al. (2010). All the simulations performed here are evolved for one outer orbital period: 10^3 yrs. Such timescale is long enough to ensure that the drag forces and turbulent diffusion affect the dust density structure but short enough that we do not expect a significant evolution of the spiral structure. We cut the disc into 720 radial directions. For each cut, we compute the gas pressure needed in Eq. 6.2. We then solve, for each cut, Eq. 6.4 using 100 radial cells, so as to obtain the 2D surface density of the dust.

6.1.3 Initial conditions

The evolution of the dust density and the resulting observational signatures depend on the initial conditions of the simulations. Generally, the distribution of grain size in a protostellar disc is affected by the grain growth and fragmentation processes which in turn are closely linked to the relative velocities between colliding particles. In this work, we assume that the dusty disc

has settled into a coagulation-fragmentation steady state. Since the self-gravitating phase of protostellar disc is more likely to occur in the early stage of star formation, such initial condition for the dust density ditribution might appear simplistic. However, recent observational results have shown that large dust aggregates can form during the disc formation stage in the infalling envelope (Miotello et al. 2014b). The distribution of grain size is assumed to be a standard power-law size distribution, introduced in Sect. 3.3.1 and repeated here for clarity,

$$n(a) \propto a^{-q} \qquad \text{for} \quad a_{\min} < a < a_{\max}, \tag{6.5}$$

where $a_{\min}$ and $a_{\max}$ represent the minimum and maximum size of the grains. The dust size distribution is modelled using 100 sizes and it is normalized such that the total initial dust density is one per cent of the gas density in each cell of the grid. The minimum size of the grains is assumed to be equal to $0.1\,\mu$m.

Regarding the value of $a_{\max}$, Birnstiel et al. (2010) have found that an acceptable estimator for the maximum grain size of the dust distributions can be computed by assuming the random turbulence motion as the only source of relative motions (for a detailed discussion of the turbulent relative velocities see Ormel & Cuzzi 2007). In this context, Rice et al. (2005) have found that the turbulent diffusion induced by gravitational instability for the gas disc model used here, is characterized by a "gravoturbolent" parameter $\alpha = 0.05$. Since the turbulence parameter related to gravitational instability is reasonably larger than the typical turbulence parameters, the previously mentioned assumption for the maximum grain size of the dust size distribution can be considered a valid approximation. The value of the maximum grain size is therefore computed by equating the relative velocities between colliding particles induced by "gravoturbolent" diffusion and the threshold velocity for fragmentation (Birnstiel et al. 2010). In the Epstein regime, this is given by

$$a_{\max} = \frac{4\langle\Sigma_{\rm g}\rangle}{3\pi\alpha\rho_{\rm p}} \frac{v_{\rm frag}^2}{\langle c_{\rm s}\rangle^2}, \tag{6.6}$$

where $v_{\rm frag}$ is the fragmentation threshold velocity and $\langle\Sigma_{\rm g}\rangle$ is the azimuthal average of the gas surface density. Similarly to the computation of the dust diffusivity (Eq. 6.3), since the typical grain growth timescales is longer than the lifetime of individual spiral features in self-gravitating discs (Brauer et al. 2008; Okuzumi et al. 2011), $a_{\max}$ is computed using the azimuthal average of the gas surface density and sound speed. Regarding the value of the fragmentation velocity, it has been shown through theoretical and experimental studies that it depends on dust grains material properties. Typically, the fragmentation velocity for silicates dust aggregates is of the order of a few m s^{-1} (Blum & Wurm 2008) while icy aggregates are able to grow at collisions with velocities up to 50 m s^{-1} (Wada et al. 2009; Gundlach & Blum 2015).

Moreover, the dust size distribution is characterized by the value of the power-law exponent q that depends on the dust dynamics and the processes of grain coagulation and fragmentation that occur in the disc. While the typical value in the ISM is $q_{\rm ism} = 3.5$ (Mathis et al. 1977), in a protoplanetary disc if coagulation (fragmentation) processes dominate over fragmentation (coagulation), a lower (higher) value for q than the ISM one is expected.

In order to cover a wide range for the initial dust size distributions, we perform simulations using different values for the fragmentation threshold velocity and the power-law exponent. We consider two values for the fragmentation threshold velocity: $v_{\rm frag} = [10, 30]$ m s^{-1} and three values for the power-law exponent: $q = [3.0, 3.5, 4.0]$.

6.1.4 Monte Carlo radiative transfer simulation

We calculate the expected emission maps of the self-gravitating disc model adopted here using the 3D radiative transfer code RADMC-3D[1]. We focus on images at H-band (1.65 μm) and ALMA band 6 (220 GHz), band 7 (345 GHz), band 8 (460 GHz) and band 9 (680 GHz). Our aim is to compute HiCIAO scattered light images in H-band polarized intensity and ALMA

[1]http://www.ita.uni-heidelberg.de/ dullemond/software/radmc-3d

(sub-)millimetre images in order to study the detectability of the peculiar spiral structure induced by gravitational instability both in scattering and thermal dust emission.

It is known that the scattering intensity is closely linked to the behaviour of the scattering phase function of dust grains (see Sect. 3.2.1). While small grains scatter photons uniformly, the scattering phase function of large grains is dominated by forward scattering. Therefore, since large grains scatter stellar photons mostly into the disc, tracing larger grains by analyzing scattering images of non-edge-on discs results challenging (Mulders et al. 2013). Thus, polarized scattering images at short wavelengths trace small (0.1-10 μm) dust grains at the surface of the disc (where stellar photons get absorbed or scattered). At (sub-)millimetre wavelengths, since the disc is generally optically thin in the vertical direction, the emission probes the mid-plane disc surface density in large grains (0.1-10 mm), due to their high opacity at these wavelengths (see Sect. 3.3, Dullemond et al. 2007; Williams & Cieza 2011).

In order to perform RADMC-3D simulation, we need to create 3D structures starting from the 2D dust density distributions obtained with the procedure described in Sect. 6.1. In a protostellar disc the vertical structure of the dust is determined by the balance between the dust settling toward the disc mid-plane due to the gravity and the vertical turbulent stirring induced by the gas drag (see Sect. 2.4.2). The dust density profile of the i$^{\rm th}$ dust species with size $a^{\rm i}$ is assumed to be Gaussian

$$\rho_{\rm d}^{\rm i}\left(z\right) = \frac{\Sigma_{\rm d}^{\rm i}}{H_{\rm d}^{\rm i}\sqrt{2\pi}}\,\exp\left(-\frac{z^2}{2{H_{\rm d}^{\rm i}}^2}\right), \tag{6.7}$$

where the scale height $H_{\rm d}^{i}$ is computed assuming that the gravitational settling balances the turbulent diffusion (Brauer et al. 2008; Birnstiel et al. 2010)

$$H_{\rm d}^{\rm i}(R) = H_{\rm g}(R)\min\left(1, \sqrt{\frac{\alpha}{\min\left({\rm St}^{\rm i}, 1/2\right)\left(1+{\rm St}^{\rm i\,2}\right)}}\right), \tag{6.8}$$

where, bearing in mind that the disc is in marginal self-gravitating state, the gas scale height is given by (Eq. 4.8. Cossins et al. 2009)

$$H_{\rm g}(R) = \frac{\pi\Sigma_{\rm g}\left(R\right)R^3}{M_\star}. \tag{6.9}$$

Having constructed the 3D distribution for all the dust species in our sample, we compute the temperature structure of the dust via thermal Monte Carlo simulations. The source of radiation is assumed to be the central star, located at the centre of the coordinate system, with $M_\star = M_\odot$, $R_\star = R_\odot$ and $T_{\rm eff} = 6000$ K. The dust in our model consists of spherical composite porous grains with chemical abundances adopted in Ricci et al. (2010). The absorption and scattering cross sections for each dust species in our sample are calculated using the routine developed by Bohren & Huffman (1983). In addition, in order to compute near infrared scattered light images in the H-band, we include the full treatment of polarization by computing the scattering matrix for each dust species.

Starting from the dust temperature structure of the model, the expected emission maps are computed via raytracing assuming that the disc is non-face-on (inclination angle of 30°). The polarized intensity image (PI=$\sqrt{\rm Q^2+U^2}$) is calculated starting from the full Stokes vector (I,Q,U,V) computed by RADMC-3D. The full-resolution (sub-)mm images directly produced by RADMC-3D are used as input sky models to simulate realistic ALMA observations using the Common Astronomy Software Application (CASA) ALMA simulator. The HiCIAO images are computed by convolving the full-resolution images in H-band with a measured HiCIAO point spread function taken from the ACORNS-ADI SEEDS Data Reduction Pipeline software (Brandt et al. 2013).

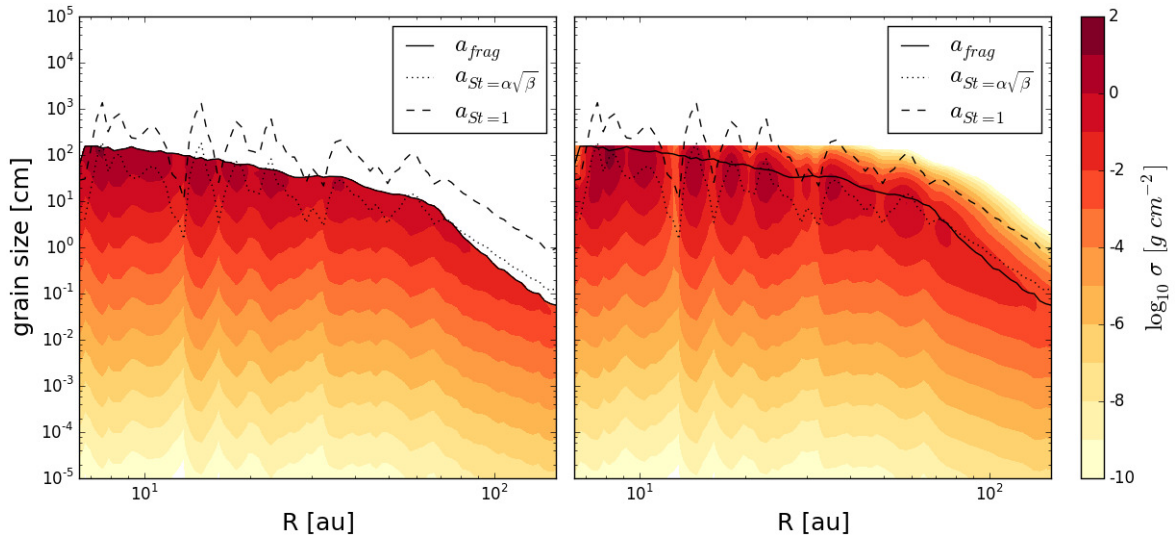

Figure 6.3: Dust density distribution at the beginning (left) and at the end (right) of simulation ($t = 10^3$ yrs) along a slice across the disc. The simulation starts from an initial size distribution with exponent $q = 3.5$ and a fragmentation velocity equal to 30 m s^{-1}. The dashed line represents the particle size corresponding to a Stokes number of unity while the dotted line show the particle size where $St = \alpha\sqrt{\beta_{\rm cool}}$. The solid line shows the fragmentation barrier, i.e. the maximum value of the initial grain size distribution (see Eq. 6.6).

6.2 Results

In this section, we present results of the dust simulations and the observational predictions of the resulting disc models. We will focus on the level of the arm-interarm contrast induced by gas-dust coupling for small and large grains and how the dust evolution and observational signatures change with different initial grain size distributions.

6.2.1 Dust density structure

First, we show a representative set of results with a given set of model parameters. Fig. 6.3 shows the initial (left) and the final (right) dust density distribution along a slice across the disc from a simulation starting from an initial size distribution with exponent $q = 3.5$ and a fragmentation velocity equal to 30 m s^{-1}. The colour contours denote the dust surface density distribution $\sigma(a, R)$ per logarithmic size bin, defined through:

$$\Sigma_{\rm d}(R) = \int_{a_{\rm min}}^{a_{\rm max}(R)} \sigma(a, R)\, {\rm d}\ln a, \tag{6.10}$$

where $\Sigma_{\rm d}$ is the total dust surface density. The solid line shows the fragmentation barrier, i.e. the maximum value of the grain size distribution (see Eq. 6.6). The dashed line represents the particle size corresponding to a Stokes number of unity, which represent the size of particles that experience the fastest radial drift. Since the dust is in the Epstein regime, this line reflects the shape of the surface density, as expressed in Eq. 6.1 so that peaks in this quantity refer to the high density arms, while valleys indicate the interarm regions. The dotted line in Fig. 6.3 shows the particle size corresponding to a Stokes number of $\alpha\sqrt{\beta_{\rm cool}}$, which, as described below, represent the size above which the radial migration towards pressure maximum dominates over turbulent diffusion. It can be noticed that particles with Stokes number St $\sim$ 1 tend to be affected by spiral density waves most effectively and exhibit the largest concentration in wave crests. The dust migration toward pressure maxima is negligible for smaller particles. This is due to the higher coupling between the dust and the slowly evolving gas and also because the diffusion is strong enough such that variations in the dust-to-gas ratio are being smeared out. In addition, since the maximum value of the initial grain size distribution decreases with

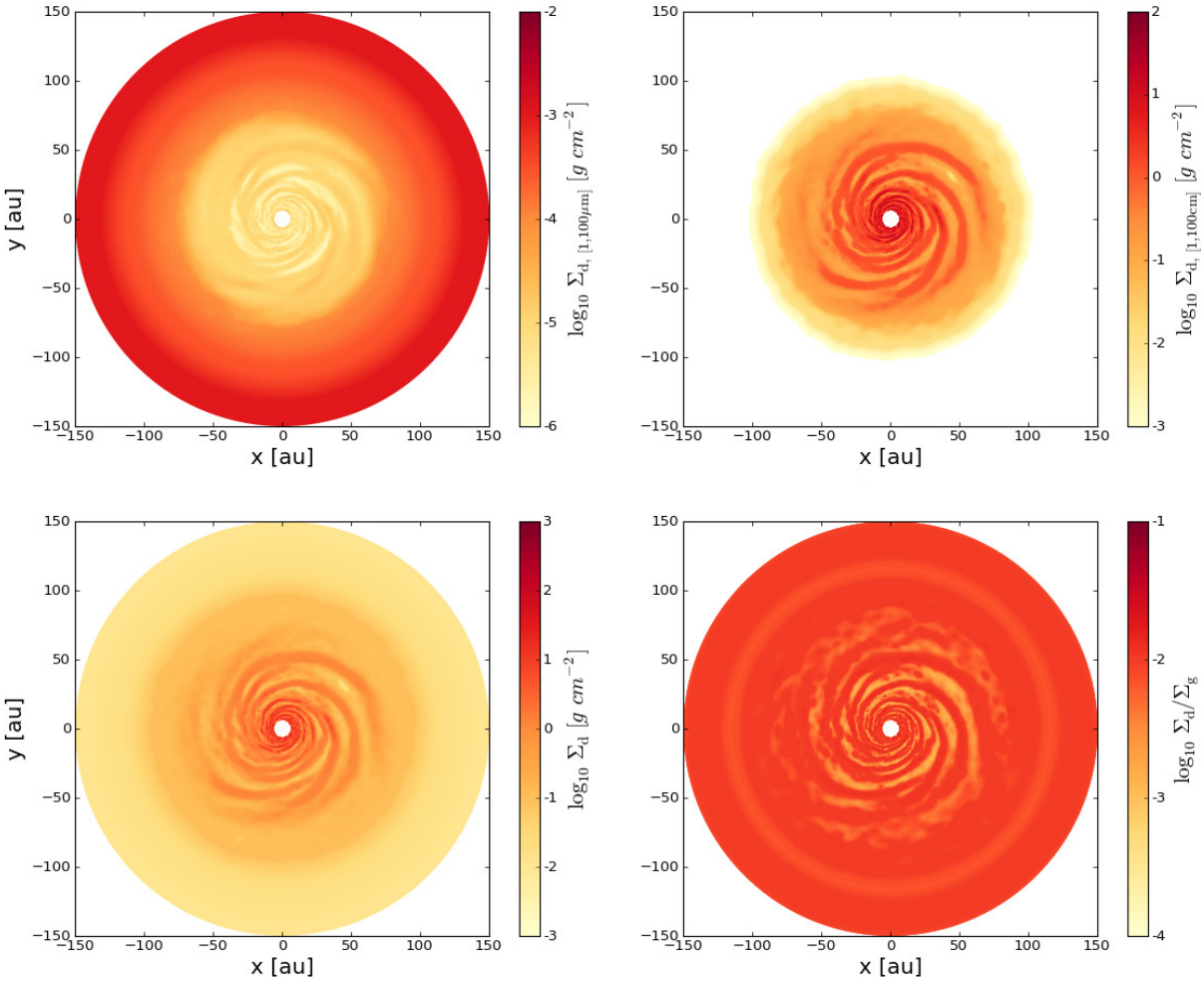

Figure 6.4: The upper panels show the surface density at the end of simulation ($t = 10^3$ yrs) integrated in size in the range $[1\,\mu\mathrm{m}, 100\,\mu\mathrm{m}]$ (left) and $[1\,\mathrm{cm}, 100\,\mathrm{cm}]$ (right). The lower panels show the total dust density distribution (left) and the dust-to-gas ratio (right) at the end of simulation.

radius, there is a strong density gradient of the largest particles between adjacent cells along the radial grid. Therefore, the gravoturbulent diffusion produces a fast migration of larger particles towards outer regions leading to an additional density enhancement in the spiral arm as can be easily noticed in the right panel of Fig. 6.3, where a large population of particles above the fragmentation limits is present. The motion towards outer regions induced by diffusion is more evident when the slope of maximum grain size increases (for $R > 70$ au) because a higher number of particles diffuse towards the outer radius. However, it is expected that largest particles beyond the fragmentation threshold tend to experience disruptive collisions on longer timescales.

It is known that the dust concentration in pressure bumps is the result of the competition between the radial drift, that forces the particle to migrate towards pressure maxima, and the turbulent diffusion that tends to recover the dust density distribution into the initial condition where the dust-to-gas ratio is equal in the overall disc. As already discussed in Section 2.5, in order to figure out the relative importance of advection and diffusion, we compare the advection time scale, i.e. the minimum time required to concentrate solid material in a pressure maximum: $t_{\mathrm{adv}} \sim \lambda/u_{\mathrm{drift}}$, and the diffusion time scale $t_{\mathrm{diff}} \sim \lambda^2/D_{\mathrm{d}}$, where λ is the scale length of the density inhomogeneities which, in a gravitationally unstable disc, is of the order of the disc scale height H_{g}. The ratio of these timescale is given by (see Sect. 2.5)

$$\frac{t_{\mathrm{diff}}}{t_{\mathrm{adv}}} = \frac{\mathrm{St}}{\alpha}\frac{\Delta\Sigma_{\mathrm{g}}}{\Sigma_{\mathrm{g}}} \tag{6.11}$$

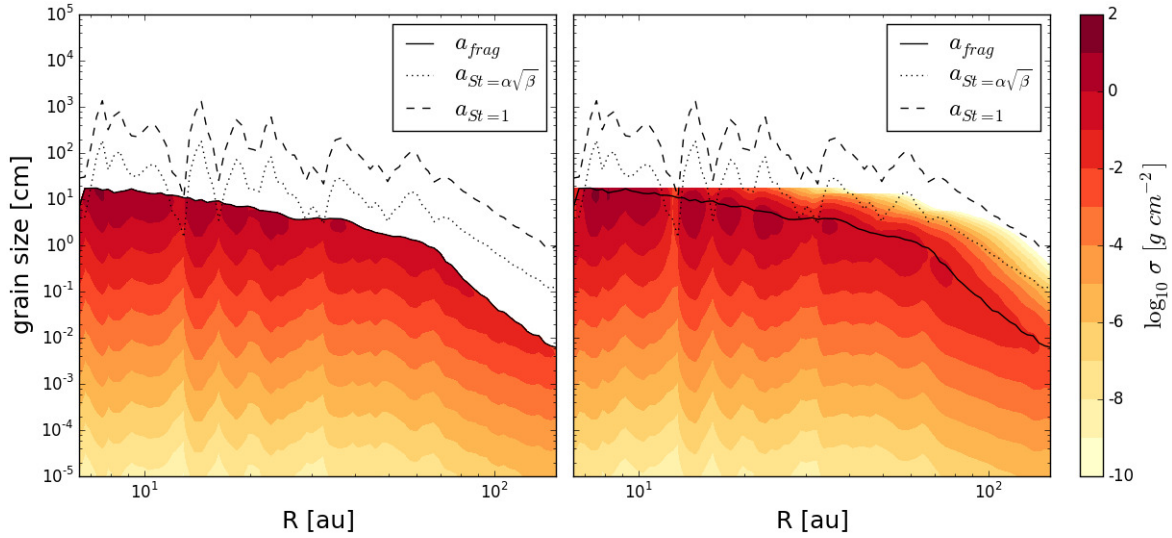

Figure 6.5: Dust density distribution at the beginning (left) and at the end (right) of simulation ($t = 10^3$ yrs) along a slice across the disc. The simulation starts from an initial size distribution with exponent $q = 3.5$ and a fragmentation velocity equal to 10 m s^{-1}. The dashed line represents the particle size corresponding to a Stokes number of unity while the dotted line show the particle size where $St = \alpha\sqrt{\beta_{\rm cool}}$. The solid line shows the fragmentation barrier, i.e. the maximum value of the initial grain size distribution (see Eq. 6.6).

where, since we have a surface density enhancement $\sim \Delta\Sigma_{\rm g}$ over a radial length scale $H_{\rm g}$, we approximate $(1/\rho_{\rm g})(\partial P/\partial R)$ as $\sim (c_{\rm s}^2/H_{\rm g})(\Delta\Sigma_{\rm g}/\Sigma_{\rm g})$. Thus, for particles with $St \gtrsim \alpha\,(\Delta\Sigma_{\rm g}/\Sigma_{\rm g})^{-1}$, the radial migration towards pressure maximum dominates over turbulent diffusion. In a self-gravitating disc in thermal equilibrium, the amplitude of the spiral structure induced by gravitational instability has to provide a stress large enough to balance the effective cooling rate (see Sect. 4.2.5). This balance is expressed in Eq. 4.48. Moreover, the value of the gravoturbulent parameter α depends on the imposed cooling (Eq. 1.51). Thus, the minimum Stokes number above which the advection towards pressure maxima dominates over gravoturbulent diffusion varies with the cooling as $\beta_{\rm cool}^{-1/2}$. In other words, the range extension of particle size that show the concentration into pressure maxima decreases with decreasing cooling time due to the combined effect of radial drift and turbulent mixing.

In Fig. 6.4 we show the final surface density rendered images of the dust for all whole disc. In order to show the differential distribution from different ranges of dust grain size, the upper panels show the surface density integrated in size in the range $[1\,\mu{\rm m}, 100\,\mu{\rm m}]$ (left) and $[1\,{\rm cm}, 100\,{\rm cm}]$ (right). As expected, the dust density distribution for particles with Stokes number close to unity (see the upper right panel of Fig. 6.4) show the highest concentration in spiral arms. In this case, the spiral arms are very thin and display a strong contrast between arm and interarm regions, while the structure shown by smaller particles is broader and essentially follows the gas structure (see Fig. 5.1). Note that beyond ~ 50 au the density of large particles drops because we have crossed the fragmentation threshold. The lower panels of Fig. 6.4 show the total dust density distribution (left) and the dust-to-gas ratio (right). It can be noticed that the dust transport process in a self-gravitating accretion discs quickly modify the radial and azimuthal distribution of the dust-to-gas ratio. Starting from an initial uniform dust-to-gas ratio (10^{-2}), dust particles migrate quickly towards pressure maxima leading to a creation of a spiral structure in the dust-to-gas ratio. The local density of particles in interarm regions reaches levels a factor of $\sim 10^{-2}$ times that of the initial dust density. It is worth noting that the evolution of the dust-to-gas ratio shows a remarkable depletion of dust grains from the interarm regions since larger particles are more subject to radial drift due to the lower gas density in interarm regions that produce a better coupling between dust and gas (see Eq. 6.1).

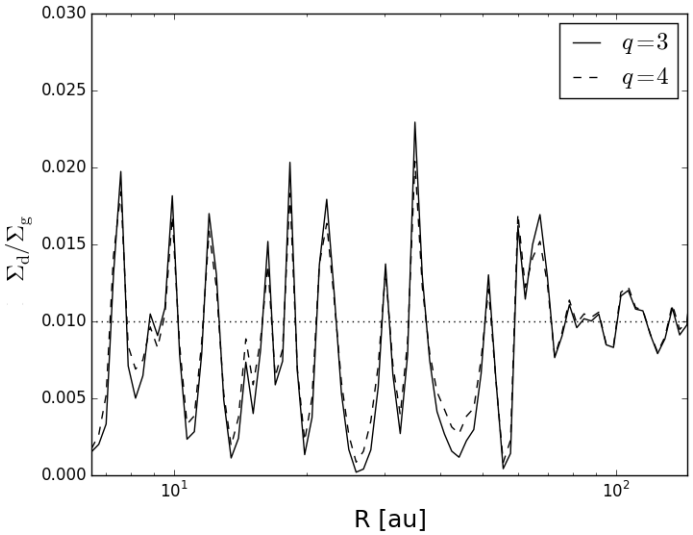

Figure 6.6: Dust-to-gas mass ratio at the end of the simulation ($t = 10^3$ yrs) along a slice across the disc for two values of the power-law exponent of the grain size distribution: $q = 3, 4$. The dotted line shows the initial dust-to-gas ratio.

6.2.1.1 Influence of fragmentation velocity

In order to cover a wide range of initial dust density distribution, we run the same simulations, but we now change the value of the fragmentation threshold velocity. As mentioned above, dust aggregates are able to grow by collisions with velocities up to a value that depends on the material properties (Blum & Wurm 2008; Wada et al. 2009).

In Fig. 6.5 we show the initial (left) and the final (right) dust density distribution along a slice across the disc using a fragmentation velocity equal to 10 m s^{-1} and an initial size distribution with exponent $q = 3.5$. If the velocity at which particles fragment decreases from 30 to 10 m s^{-1}, grains are not expected to grow to large sizes so that the condition St ~ 1 is not expected to be met anywhere in the disc. In this case, the efficiency of the dust trapping is reduced since most of the particles are characterised by a Stokes number less than unity. As a result, variations in the dust-to-gas ratio are being smeared out by the turbulent diffusion and the radial drift has only a minor effect in the dust dynamics.

6.2.1.2 Influence of the initial power-law grain size distribution

As mentioned above, the grain coagulation and fragmentation processes can affect the size distribution quite significantly. If the fragmentation process dominates over coagulation, a constant replenishment of smaller grains occurs in the disc producing a steeper size distribution, approaching $n(a) \propto a^{-4}$. In this case, most of the mass will be concentrated in small grains which are less affected by the radial drift toward pressure maxima. By contrast, if the grain coagulation dominates over fragmentation, the resulting grain size distribution is expected to be flatter (e.g. $n(a) \propto a^{-3}$). In this case, the grain coagulation process would produce large grains which migrate quickly toward pressure maxima. Fig. 6.6 shows the radial dependence of the final dust-to-gas mass ratio along a slice across the disc for two values of the power-law exponent of the grain size distribution and using a value of the fragmentation velocity equal to 30 m s^{-1}. For $q = 3$, the radial distribution of dust is strongly affected by the migration toward pressure maxima induced by gas drag since most of the dust mass is in large grains. As expected, a steeper size distribution reduce the concentration but only marginally, so that the dust-to-gas ratio in the interarm regions is slightly enhanced with respect to the previous case.

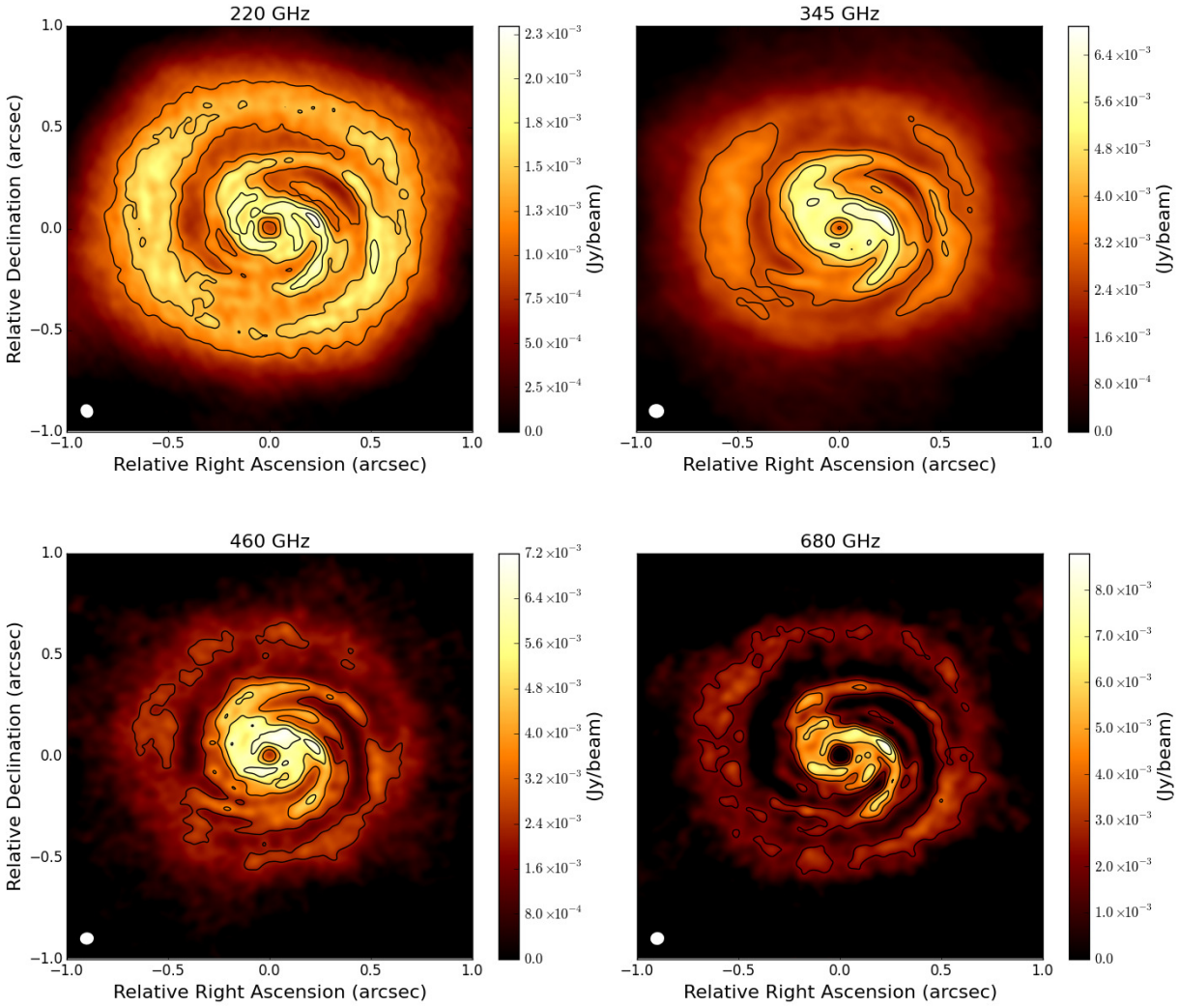

Figure 6.7: ALMA simulated images of the self-gravitating disc model obtained starting from an initial dust grain size distribution with exponent $q = 3.5$ and a fragmentation velocity equal to 30 m s^{-1}. Observing frequencies are reported at the top of each image. Contours are 4, 6, 8, 10 $\times$ the corresponding rms noise. The white colour in the filled ellipse in the lower left corner indicates the size of the half-power contour of the synthesized beam (see Table 6.1 for details).

6.2.2 Observational predictions

In this section, we present results from simulated ALMA and HiCIAO observations of the sky models computed via RADMC-3D simulations starting from the disc models obtained from hydro simulations. We translate hydro simulations into model observations using the method explained in Sect. 6.1.4. Once having computed the expected full-resolution emission maps, we simulate realistic ALMA and HiCIAO images at (sub-)mm and near-infrared wavelengths. The synthetic images are analysed in terms of the ability to spatially resolve the spiral features of our models with the aim to test the detectability of gravitationally-induced spiral density waves.

6.2.2.1 ALMA simulated observations

The resulting emission maps at (sub-)mm wavelength are used as input sky models for the ALMA simulator (version 4.1) in order to produce realistic ALMA observations of self-gravitating disc models. The selection of observing parameters is chosen to ensure enough spatial resolution and sensitivity to resolve and detect spiral arms in the intermediate and external region of the discs. Assuming a perfect calibration of the visibility measurements, the computation of ALMA simulated observations include the effects of thermal noise from the receivers and atmosphere.

Frequency (GHz)	PWV (mm)	beam size (arcsec × arcsec)	Total flux (Jy)	rms (mJy/beam)
220	2.75	0.063 × 0.055	0.71	0.18
345	1.26	0.070 × 0.063	0.93	0.49
460	1.26	0.062 × 0.054	1.12	0.84
680	0.47	0.062 × 0.057	1.31	1.15

Table 6.1: Atmospheric conditions, beam size, total flux and rms for the simulated observations shown in Fig. 6.7.

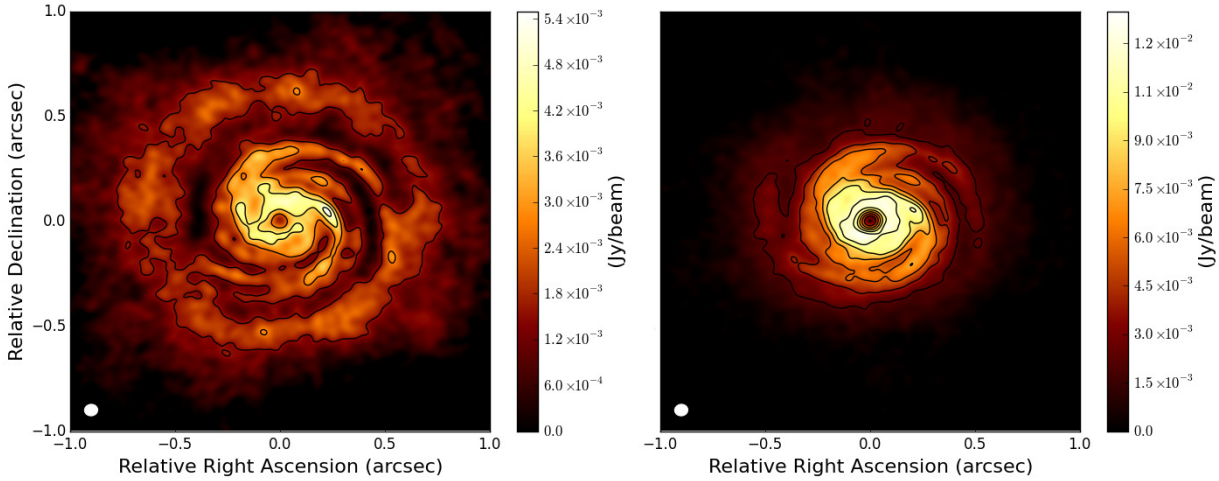

Figure 6.8: Comparison of ALMA simulated images at 460 GHz of disc models obtained from simulations with an initial dust grain size distribution with power-law exponent $q = 3$ (left) and $q = 4$ (right) and a fragmentation velocity equal to 30 m s^{-1}. Contours are 4, 6, 8, 10 × the corresponding rms noise: 0.78 mJy/beam (left) and 0.72 mJy/beam (right). The white colour in the filled ellipse in the lower left corner indicates the size of the half-power contour of the synthesized beam: 0.062 arcsec × 0.054 arcsec.

The signal atmospheric corruption effects into the visibilities measurements is evaluated taking into account the atmospheric transparency computed using the Atmospheric Transmission at Microwaves (ATM) code (Pardo et al. 2002) starting from average site conditions and input values for the ground temperature and the Precipitable Water Vapour (PWV) provided by the user. All the sources are located in the Ophiucus star-forming region ($\sim$ 140 pc) and are observed with a transit duration of one hour.

First, we focus on a representative disc model obtained from a given set of model parameters. In Fig. 6.7 we show ALMA simulated observations at different observing wavelengths of the disc models obtained starting from an initial dust grain size distribution with exponent $q = 3.5$ and a fragmentation velocity equal to 30 m s^{-1}. The observation parameters adopted in the images presented here and the measured total fluxes are summarized in Table 6.1. It can be noticed that the spiral structure is readily detectable by ALMA over a wide range of wavelengths. Note that relative intensity between the spiral arms and the external region decreases with increasing frequency. This variation is linked to the optical regime of the spiral arms at different frequencies. It can be shown that, while the interarm regions are optically thin, the arm regions became optically thick for frequencies $\gtrsim$ 345 GHz. Thus, for $\nu \gtrsim$ 345 GHz, the radiation coming from the central protostar is more likely to be captured by the arm regions, producing a fainter emission from the outer regions.

In addition, as already noted by Dipierro et al. (2014), while the original density structure of the disc is characterized by a relatively large number of arms (see Fig. 5.1), the ALMA image show a spiral pattern with two spiral arms since the smaller-sized arms have been smeared out by the limited resolution of the observations.

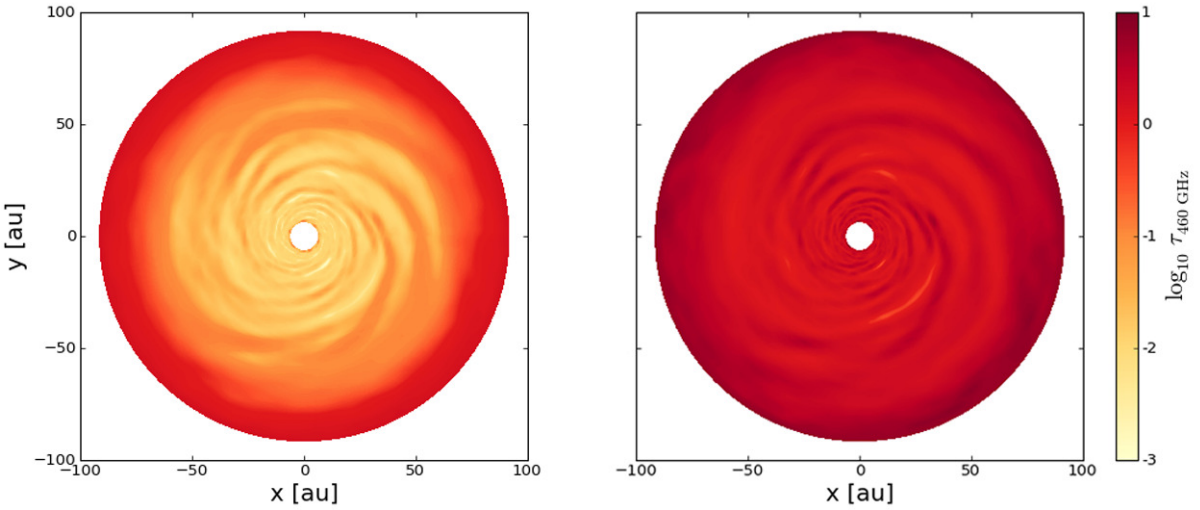

Figure 6.9: Comparison of optical depth maps at 460 GHz of disc models obtained from simulations with an initial dust grain size distribution with power-law exponent $q = 3$ (left) and $q = 4$ (right).

In Fig. 6.8 is shown the comparison of two ALMA simulated observations at 460 GHz of the disc models obtained from simulations with an initial dust grain size distribution with power-law exponent $q = 3$ (left) and $q = 4$ (right). As expected, the slope of the dust grain distribution affects the resulting observation. For $q = 4$, most of the dust mass is concentrated in small grains. This produces an optically thick inner region that capture all the radiation coming from the central protostar. As a result, due to the remarkable decrease of the contrast between arm and interarm region, the spiral pattern is only barely with ALMA. To illustrate the effects of varying the slope of the size grain distribution, in Fig. 6.9 we show the optical depth maps at 460 GHz of the disc models obtained from simulations with an initial dust grain size distribution with power-law exponent $q = 3$ (left) and $q = 4$ (right). The optical depth is evaluated locally using the following expression:

$$\tau_\nu = \int_{a_{\min}}^{a_{\max}} \frac{\partial \Sigma_{\mathrm{d}}}{\partial a} \kappa_\nu(a)\, da, \tag{6.12}$$

where $\kappa_\nu(a)$ represents the absorption opacity at frequency ν for the dust species with size a. It can be noted that, for grain size distribution with $q = 3$, the spiral pattern region is optically thin. On the contrary, for steeper grain size distribution the optical regime is thick for all the disc.

Observation at (sub-)millimetre wavelengths can probe some dust properties in the disc interior. By making some assumptions on the chemical composition, shape and size distribution of the dust grains, it is possible to put constraints on the level of grain growth (see Sect. 3.3.1 for details). In the context of the present work, we calculate the value of the β parameter by computing the ratio of the ALMA simulated images at two different frequencies, with the aim to infer the different level of dust concentration of large grains in arm and interarm regions. The β parameter can be obtained as (Eq. 3.10)

$$\beta = \left[\log\left(\frac{\nu_1}{\nu_2}\right)\right]^{-1} \log\left(\frac{F_{\nu_1}}{F_{\nu_2}}\right), \tag{6.13}$$

where F_{ν_1} and F_{ν_2} are the fluxes at frequencies ν_1 and ν_2, respectively. In Fig. 6.10 we show the map of β computed from the model data (left) and the β parameter calculated from two simulated ALMA observations at 650 μm and 3 mm (right). We consider the disc model obtained starting from an initial dust grain size distribution with exponent $q = 3$ and a fragmentation velocity equal to 30 m s^{-1}. At wavelengths mentioned above we verify that the disc is optically thin and the dust temperature is high enough to ensure that the disc is in the Rayleigh-Jeans

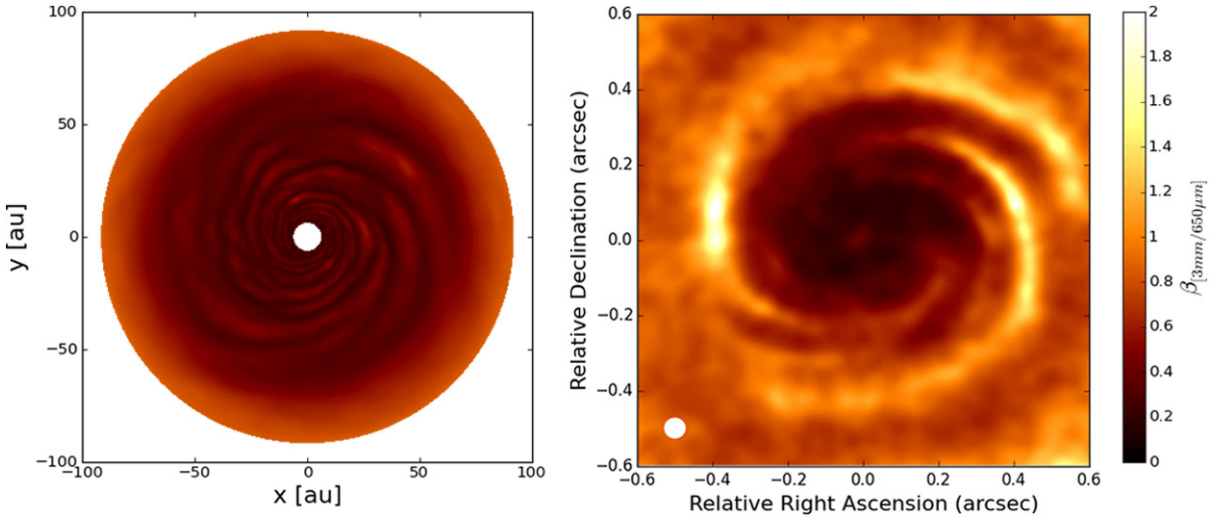

Figure 6.10: The β parameter computed from (left) the model data and from (right) two simulated ALMA observations at 650 μm and 3 mm of the disc model obtained starting from an initial dust grain size distribution with exponent $q = 3$ and a fragmentation velocity equal to 30 m s^{-1}. The white colour in the filled ellipse in the lower left corner of the right panel indicates the size of the half-power contour of the synthesized beam: 0.054 arcsec $\times$ 0.052 arcsec.

limit. Therefore, the variations of β can be interpreted as a sign of dust trapping regions. From Fig. 6.10 it can be noticed that the beta index map has a shape that resemble the spiral structure of the disc model. The spiral arms, where most of the centimetre sized grains are concentrated due to the radial drift, are characterized by a lower value of β. By contrast, the β parameter is relatively higher in the interarm regions, where large grains are depleted due to the dust trapping mechanism. Thus, taking into account all the assumptions about the dust grain properties, the dust trapping induced by gravitational instability produces a migration of large dust grains into spiral arms region that can be potentially investigated through multi-wavelenght observations.

6.2.2.2 Polarized H-band scattered light images

In this subsection we show synthetic HiCIAO H-band polarized intensity image of the disc model obtained starting from an initial dust grain size distribution with exponent $q = 3.5$ and a fragmentation velocity equal to 30 m s^{-1}. As previously mentioned, the polarized intensity images essentially trace the gas disc structure since small dust grains are less subject to the radial drift. The calculated scattered light image convolved with the measured HiCIAO point spread function is shown in Fig. 6.11. The H-band image reveals an asymmetry in brightness between the far (top) and near (low) disc regions (see Sect. 3.2.1). This is due to forward scattering of starlight by large dust particles in a non-face-on disc (inclination of 30°). In addition, the outer ring-like structure is very bright since the density of smaller particles increases with radius due to the lower value of the fragmentation threshold size in outer disc regions (see Fig. 6.3). Moreover, the significative brightness of the outer ring is also caused by the fact that scattered light is mostly determined by the curvature of the disc surface that, due to the flaring of the disc, is very enhanced in the outer regions. However, due to the higher brightness of the inner regions and the instrument resolution, the spiral structure is not readily detected. The spiral arms and cavities clearly detected in (sub-)millimetre ALMA images (see Fig. 6.7) are smeared out and are thus marginally visible in the HiCIAO H-band image. It is worth remarking that the shape of the HiCIAO Point Spread Function (PSF) is characterized by large sidelobes which produce a decrease of the dynamic range in the convolved images and, consequently, a decrease of the detected contrast between arm and interarm regions (see Appendix B for further details on the technical aspect of aperture synthesis technique).

We also consider a new set of simulations with a gas disc model two times bigger. In Fig. 6.12

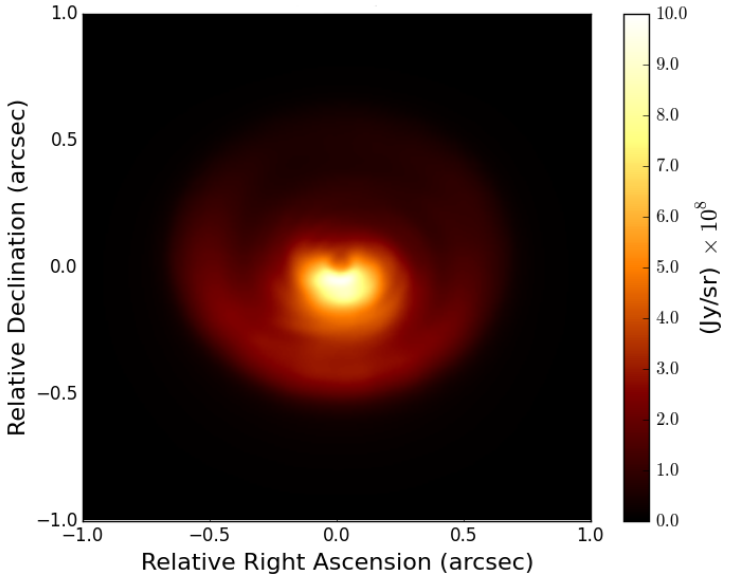

Figure 6.11: Synthetic HiCIAO H-band polarized intensity image of the non-face-on disc model (inclination of 30°) obtained starting from an initial dust grain size distribution with exponent $q = 3.5$ and a fragmentation velocity equal to 30 m s^{-1}.

we present the synthetic HiCIAO H-band polarized intensity image of this specific disc model with two different inclinations: 10° (left) and 30° (right). A remarkable improvement of the detection of spiral arms can be noticed with comparison to the model previously considered (see Fig. 6.11), due to the fact that the HiCIAO PSF is small enough to preserve spiral features in scattering images. As expected, the asymmetry in brightness between the far (top) and near (low) disc regions due to the forward scattering of starlight by large dust particles becomes more prominent with increasing angle of inclination. In addition, since the separation between arm and interarm regions decreases with increasing angle of inclination, the detected spiral structure results sharper for face-on discs. Moreover, the contrast between arm and interarm region appears lower in scattering emission than in thermal emission. This is caused mostly by the smaller contrast in dust density of smaller particles with respect to larger particles between arm and interarm regions. All these effect makes it more challenging to observe the spiral structures in inclined disc through scattering intensity observations.

6.3 Summary and conclusion

In this Chapter, by combining hydrodynamical and dust evolution models, we investigate the influence of gravitational instabilities on the dynamics of dust grains. Using representative grains size distributions, we model the dust grains dynamics under the action of systematic and random motion induced by the gas-dust aerodynamical coupling. Once having computed the resulting 2D dust density distribution, we perform 3D Monte Carlo radiative transfer simulation in order to make predictions of self-gravitating discs for future observations.

The spiral density waves induced by gravitational instability produce a remarkable effect on the dust dynamics leading to the formation of significant dust overdensities in the spiral arms regions. The dynamical effect produced by the occurrence of pressure bumps in discs closely depends on the size of the particles. We found that, assuming that particles are able to grow through collisions with velocities up to 10 and 30 m s^{-1}, the dust density structure of centimetre sized particles is affected by the presence of the spiral overdensities. While smaller particles closely trace the overall gaseous spiral structure, larger particles experience the highest concentration in density maxima. The influence of the dust trapping mechanism has been tested by computing the dust-to-gas ratio. Since most of the mass is concentrated in the largest grains, the dust transport process modifies the radial and azimuthal distribution of the dust-to-gas ratio

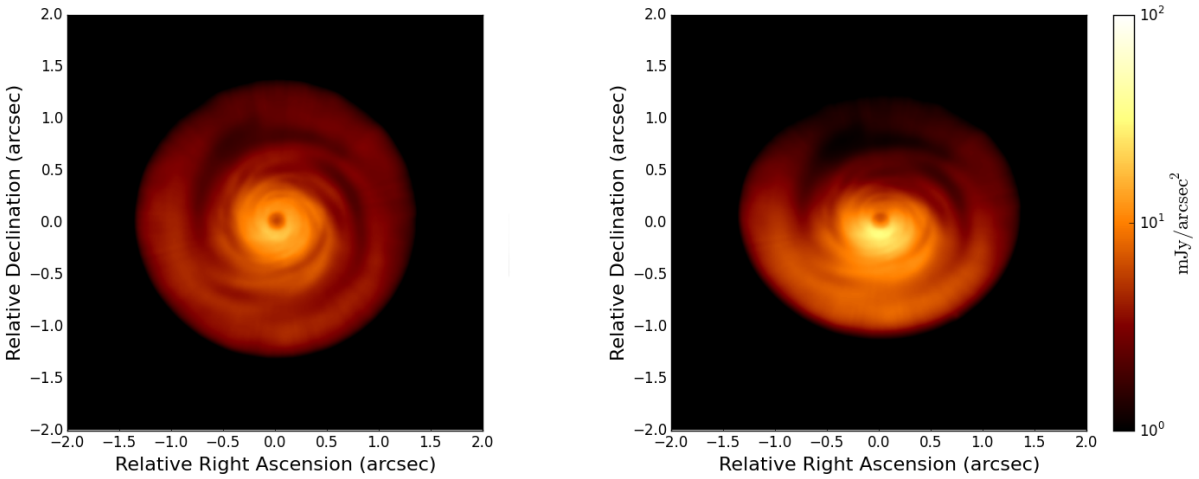

Figure 6.12: Synthetic HiCIAO H-band polarized intensity image of a disc model two times bigger than the model of Fig. 6.11 (i.e. $R_{\rm out} = 300$ au) with two different inclinations: 10° (left) and 30° (right).

leading to a creation of a spiral structure.

The observational predictions of the resulting models show that the resolution capabilities and sensitivity of ALMA are sufficient to spatially resolve the peculiar spiral structure of gravitationally unstable discs with an acceptable signal-to-noise ratio for non-face-on discs located in the Ophiucus ($\sim$ 140 pc) star-forming region. Through multi-wavelenght (sub-)millimetre observations, we have investigated whether the density enhancement of the larger dust grains into spiral arm regions produces signatures in the spectral index map that can be detected using ALMA. We find that the migration of larger grains into spiral arms region produces clear variations in the spectral index that can be interpreted in terms of different level of grain growth between arm and interarm regions.

In addition, our simulations show that the development of gravitational instability can create strong enough surface density perturbation that could be detected in near-infrared scattered light with HiCIAO. Therefore, the spiral arms observed to date in protoplanetary disc in polarized scattered light (see Sect. 4.3.2, Muto et al. 2012; Grady et al. 2013; Garufi et al. 2013) might be the result of spiral density perturbations induced by the development of gravitational instabilities.

The simulations presented here contain a number of simplifications that need to be addressed. First, since the lifetime of the individual spiral features in self-gravitating discs is of the order of Ω^{-1}, the approximation of a stationary gas density structure for as long as one outer orbital period appears to be simplistic (Clarke & Lodato 2009). Moreover, in this work we neglect the grain growth and fragmentation processing that occur in the disc. This can be considered an acceptable approximation since the timescale of our simulations is shorter than the typical grain growth timescale (Brauer et al. 2008; Okuzumi et al. 2011) in the intermediate and outer part of the disc. In addition, it is expected that the occurrence of long-lived non-axisymmetric structures in the disc affects the azimuthal dynamics of dust grains leading to an additional systematic drift motion of the dust particles towards the pressure maximum (Nakagawa et al. 1986; Birnstiel et al. 2013). Assuming azimuthal and radial trapping simultaneously would lead to stronger concentrations of dust in pressure traps.

Part III
Gaps in Protoplanetary Discs

Chapter 7

Disc-Planet Interaction

The dynamical evolution of dust grains in protoplanetary discs hosting planets has become a topic of renewed interest in the last years. Recent high-resolution ALMA and SPHERE observations have provided evidence that a large variety of ring-like structures exist in protoplanetary disc, regardless of their age (ALMA Partnership et al. 2015; Nomura et al. 2016; Andrews et al. 2016; Canovas et al. 2016; van der Plas et al. 2016; de Boer et al. 2016; Ginski et al. 2016; van Boekel et al. 2016). Despite the large number of dynamical and non-dynamical mechanisms that could potentially induce gap-like structures in protoplanetary discs, the planetary hypothesis is the widely invoked mechanism to explain their formation.

In this Chapter we provide an overview of the tidal interaction between a planet and the surrounding disc material with a focus on the mechanism of gap opening in the gas density structure and the links between the morphological features of gaps and both the planet's properties and the local gas conditions. The basic framework described in this Chapter was firstly introduced in the seminal papers of Goldreich & Tremaine (1979); Lin & Papaloizou (1979) and recently reviewed by Kley & Nelson (2012) and Baruteau et al. (2014), to which we refer for a more detailed analysis.

7.1 Planetary torque

Planets embedded in a protoplanetary disc perturb the motion of the surrounding fluid elements by deflecting their orbit due to its gravity. The perturbation induced by the planet leads to oscillations of fluid particles around their circular unperturbed orbit, exciting waves which propagate across the disc far from the planet. Propagating density waves are sheared out by the Keplerian differential rotation, creating a wave pattern in the form of spirals. As described in Sect. 7.1.2, for small mass planets, the waves can constructively interfer and form a one-armed spiral structure, mostly referred to as the *wake* (see the left panel of Fig. 7.1).

Let us consider a simplified scenario to describe the interaction between the planet and the tidally-induced wake (see a more detailed description in Sect. 7.1.2). The wake overdensity and the perturber interact gravitationally, leading to an exchange of angular momentum between the planet and the wake. The inner part of the wake pulls the planet forward, generating a positive torque on the planet and driving outward migration. Reciprocally, the planet tends to slow down the inner disc, exerting a negative torque. By contrast, the planets tends to pull the outer disc, exerting a positive torque and resulting in an angular momentum gain of the outer disc. Reciprocally, the outer disc slows down the planet, driving inward migration. Hence, the tidal effects of the planet to the inner and outer disc push the outer disc outward and the inner disc inward, leading to the formation of an axisymmetric underdense region around the planet location, usually called as *gap*. The morphology of the gap and the ability of the planet to carve it depend on the disc local properties and on the mass of the planet (see Sect. 7.3 and Chapt. 8 for a detailed description of the gap opening process in the gas and dust, respectively).

It is worth noting that the gap opening process requires not only the generation of density waves, but a damping mechanism responsible for the transfer of the angular momentum from

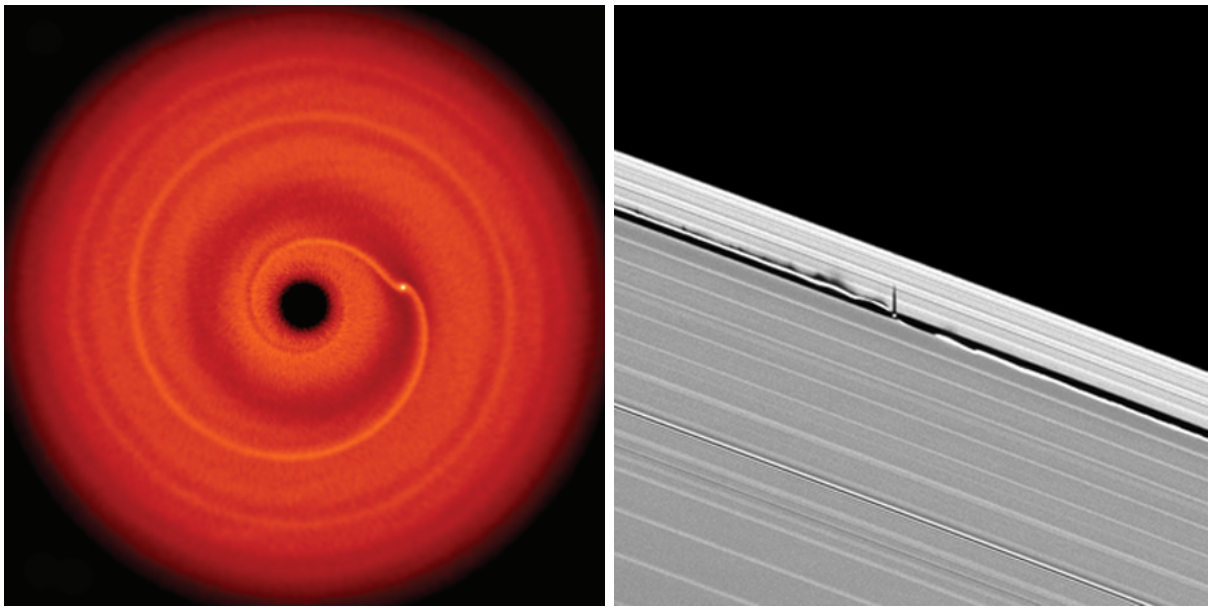

Figure 7.1: (left) Surface density structure of a gas disc with an embedded protoplanet of mass $M_\mathrm{p} = 0.5 M_\mathrm{J}$ computed by means of three dimensional SPH simulations. The disc model is described in Chapter 8. (right) Density waves in Saturn's rings formed by the gravitational pull of the tiny moon Daphnis. Image captured by NASA's Cassini spacecraft (`https://saturn.jpl.nasa.gov/`).

the wake to local fluid elements. In other words, the planet's torque affects the structure of the disc at the locations where the waves are dissipated. The idea of non-local deposition of angular momentum was first introduced by Goldreich & Tremaine (1978) to explain the wide gap in the Cassini division of Saturn's rings and then confirmed by the direct observations of density waves excited by Saturn's moons firstly captured by Voyager I (Cuzzi et al. 1981) and recently by the NASA's Cassini spacecraft (see the right panel of Fig. 7.1). It is therefore necessary to distinguish between the amount of angular momentum added to the density waves, the so-called *excitation* torque, and the angular momentum transferred from the waves to the disc, mostly known as *deposition* torque, which controls the gap opening.

In order to gain insights on the excitation torque in the linear regime, i.e. the amount of angular momentum added to the density waves due to the tidal interaction with a low mass planet, two different approaches can be followed:

- Particle-based approach: the angular momentum is exchanged by individual particles passing by the planet undergoing small deflections on short timescale. This is the so-called *impulse approximation* (Lin & Papaloizou 1979);

- Fluid-based approach: linear analysis of tidally-induced perturbations, calculating the response of the fluid disc adopting linearized hydrodynamic equations. The linear analysis of planet-induced perturbations has been first outlined in the seminal paper of Goldreich & Tremaine (1979) and recently refined by Rafikov & Petrovich (2012).

The impulse approximation provides correct and valuable results, emboding the proper dependences on the various physical quantities. However, the approximation does not lead to the correct dimensionless numerical coefficients of proportionality for the excitation torque and involve some assumptions (such as that the duration of the interaction has to be much shorter than the orbital period) marginally satisfied in protoplanetary disc. However, a full fluid dynamical approach (Goldreich & Tremaine 1979) leads to deeper insight and to more reliable results in the limit of small perturbations.

7.1.1 Impulse approximation

The impulse approximation has been developed by Lin & Papaloizou (1979) and is based on the study of the impulsive exchange of angular momentum between the planet and an approaching particle that undergoes a small deflection over a duration much shorter than the orbital period of the planet. The last approximation is made in order to neglect the orbital motion of the planet around the star. Moreover, the typical random velocity is assumed to be much smaller than the typical orbital velocity v_k. In gaseous discs, since the typical random velocity is $\sim \sqrt{\alpha} c_\mathrm{s}$ (Zhu et al. 2015a), where c_s is the sound speed and α is the Shakura & Sunyaev (1973)'s turbulence parameter, this condition can be rewritten as $H_\mathrm{g}/R \ll \alpha^{-1/2}$, where H_g is the height scale of the disc at vertically hydrostatic equilibrium (see Sect. 1.2.1.1). This condition is typically satisfied in protoplanetary discs. We therefore treat the planet and the approaching particle as moving on a uniform straight motion and consider their interaction as a weak impulsive scattering encounter. The term *impulsive* means that the trajectory remains mostly unperturbed until the particle is at its minimum distance from the satellite or, in other words, the exchange of angular momentum occurs over a duration much shorter than the orbital period of the planet.

Consider a particle of mass $m_\bullet$ on a circular orbit about the star with orbital radius R that approaches the planet orbiting on a fixed circular orbit at radial distance R_p from the central star of mass $M_\star$. The orbital radius of the particle needs to be sufficiently different from R_p to allow it to freely pass by the planet with a small deflection. For simplicity, it is convenient to switch to the frame of reference moving with the planet. In this frame, mostly referred to as *corotating frame*, the planet is motionless and the angular velocity of the particle is $\Omega - \Omega_\mathrm{p}$ where Ω_p is the angular velocity of the planet orbiting around the star measured on the inertial frame centered on the central star (mostly known as *heliocentric frame*). The test particle in the corotating frame of reference is coming from infinity with velocity v_0 and impact parameter $b = R - R_\mathrm{p}$ (see Fig. 7.2). Upon interaction with the planet, the particle is impulsively deflected toward it and gains a velocity on the radial direction (i.e. the direction perpendicular to its unperturbed motion) given by

$$|v_\perp| = \int_{-\infty}^{\infty} \frac{F_\perp}{m_\bullet} dt = \int_{-\infty}^{\infty} \frac{\mathcal{G} M_\mathrm{p}}{d^2} \frac{b}{d} dt, \tag{7.1}$$

where d is the instantaneous separation between the particle and the planet. Adopting $t = 0$ as the moment of the closest approach, the distance along the straight trajectory (we are assuming that the deflection is small) from the point of closest approach is $x = v_0 t$. Therefore, the instantaneous separation is $d^2 = b^2 + v_0^2 t^2$. Changing the variable in the integral from t to $s = v_0 t$, the integral is given by

$$|v_\perp| = \frac{\mathcal{G} M_\mathrm{p}}{v_0} \int_{-\infty}^{\infty} \frac{b}{(b^2 + s^2)^{3/2}} ds = \frac{2\mathcal{G} M_\mathrm{p}}{b v_0}. \tag{7.2}$$

Because of the deflection, the relative particle velocity in the tangential direction decreases from v_0 to $v_\parallel = v_0 \cos\beta$, and gains a radial velocity $|v_\perp| = v_0 \sin\beta \approx v_0 \beta$ expressed in Eq. 7.2, where is assumed that, during the encounter with the planet, the direction of the relative velocity of the particle is changed, but not its value. This can be explained by recalling the conservation of kinetic energy of the particle before and after the interaction with the planet (Armitage 2010). The deflection angle is therefore given by

$$\beta \approx \frac{2\mathcal{G} M_\mathrm{p}}{b v_0^2}. \tag{7.3}$$

During the encounter with the planet, the particle acquires some eccentricity and no longer orbit around a circular trajectory around the star. After the deflection, the particle is assumed to slowly recovers its circular orbit corresponding to the new value of its angular momentum. It is worth remarking that this condition implicitly assumes the presence of some dissipative

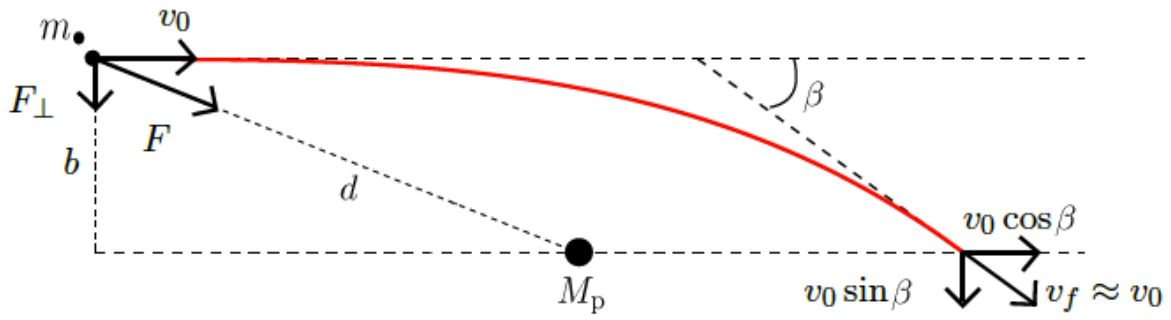

Figure 7.2: Sketch of the impulsive deflection of a particle under the tidal influence of a planet.

mechanism to damp out any eccentric motion (Lodato et al. 2009). Assuming that the particle circularizes its orbit before the next encounter with the planet, which occurs on a synodic period defined as $2\pi/\left|\Omega-\Omega_{\rm p}\right|$, the change in the specific angular momentum of the particle with respect to the star is given by

$$\delta j = Rv_{\parallel} - Rv_0 = Rv_0\left(\cos\beta - 1\right). \tag{7.4}$$

Since the deflection angle is small, $\cos\beta - 1 \approx -\beta^2/2$, yielding

$$\delta j = -Rv_0\frac{\beta^2}{2} = -\frac{2\mathcal{G}^2M_{\rm p}^2R}{b^2v_0^3}, \tag{7.5}$$

the exchange of angular momentum therefore depends on an odd power of the initial relative velocity, expressed by

$$v_0 = R\Omega - R_{\rm p}\Omega_{\rm p} \approx R_{\rm p}\left(\Omega-\Omega_{\rm p}\right) = -R_{\rm p}\left|\Omega-\Omega_{\rm p}\right|\mathrm{sgn}\left(b\right), \tag{7.6}$$

where $R/R_{\rm p} = (b+R_{\rm p})/R_{\rm p} \approx 1$, since $b \ll R_{\rm p}$. In this limit,

$$\left|\Omega-\Omega_{\rm p}\right| \simeq \left.\frac{d\Omega}{dR}\right|_{R_{\rm p}} b = \frac{3\Omega_{\rm p}}{2R_{\rm p}}b. \tag{7.7}$$

The exchange of angular momentum expressed by Eq. 7.5 represents the loss of the specific angular momentum during a single encounter with the planet. Since such an encounter occurs every synodic period, the specific torque acting on the particle, i.e. the rate of loss of specific angular momentum, assuming $R/R_{\rm p} \approx 1$, is given by

$$\frac{\delta j}{\delta t} = \delta j\left(\frac{2\pi}{\left|\Omega-\Omega_{\rm p}\right|}\right)^{-1} = \mathrm{sgn}\left(b\right)\frac{\mathcal{G}^2M_{\rm p}^2}{\pi R_{\rm p}^2b^2\left(\Omega-\Omega_{\rm p}\right)^2} = \mathrm{sgn}\left(b\right)\frac{4}{9\pi}\left(\frac{\mathcal{G}M_{\rm p}}{\Omega_{\rm p}}\right)^2\frac{1}{b^4}, \tag{7.8}$$

which can be restated, in terms of $M_{\rm p}/M_\star$ and as a function of the radial location of the particle, as

$$\partial_m\Gamma^{\rm T}(R) = \mathrm{sgn}\left(R-R_{\rm p}\right)\frac{4}{9\pi}\left(\frac{M_{\rm p}}{M_\star}\right)^2\Omega_{\rm p}^2R_{\rm p}^2\frac{R_{\rm p}^4}{\left(R-R_{\rm p}\right)^4}. \tag{7.9}$$

This result represents the tidal torque per unit mass acting on an elementary ring of particles located at radius R. This formula is not strictly exact because of the approximation $R \approx R_{\rm p}$ made in the calculus, which is not valid when R tends to infinity. However, since $\left(R-R_{\rm p}\right)^{-4}$ vanishes when $R \to \infty$, the error made using this approximation can be neglected. The total torque on the disc can be evaluated by integrating the single particle torque over all the mass in the disc. The mass of an annulus close to the planet position between b and $b+db$ exterior to the planet is given by

$$dm \approx 2\pi R_{\rm p}\Sigma_{\rm g}db, \tag{7.10}$$

where Σ_g is the surface density of the disc. Assuming that Σ_g is constant[1], the total torque acting on the outer disc, mostly referred to as *one-sided tidal torque*, is given by

$$\Gamma^T = \frac{8}{9}\left(\frac{M_p}{M_\star}\right)^2 \Sigma_g \Omega_p^2 R_p^7 \int_{b_{min}}^{\infty} \frac{db}{b^4} = \frac{8}{27}\left(\frac{M_p}{M_\star}\right)^2 \Sigma_g \Omega_p^2 R_p^4 \left(\frac{R_p}{b_{min}}\right)^3, \tag{7.11}$$

where b_{min} denotes the minimum impact parameter below which the encounter would be not fast and weak enough to justify the impulse approximation. The condition for small deflection requires that the impact parameter $|b|$ needs to be larger than the Hill radius R_H, i.e. the radius of the region around the perturber inside which the gravitational force of the perturber on fluid particles dominates over stellar and centrifugal forces, defined as

$$R_H = R_p \left(\frac{M_p}{3M_\star}\right)^{1/3}. \tag{7.12}$$

Inside the Hill sphere the gravity of the planet dominates over that of the star and particles experience heavily deflected trajectories (e.g. Bate et al. 2003; D'Angelo & Lubow 2008).

Up to now, the results shown can be applied regardless of the nature and composition of the disc material. It can be applied to a fluid disc, collisionless gas or a collection of dust particles. In order to give an estimate of the minimum impact parameter, let us consider a disc made of gas. In this case, it is necessary to take into account also the possibility that the gas particle is aware of the presence of the planet due to pressure effects. Using simple arguments, if the relative velocity between the particle and the planet is less than the sound speed, the particles will be aware of the presence of the planet. In this case, the condition of impulsive interaction is no longer valid since the particle could interact with the planet before the closest approach on longer timescale. The minimum impact parameter is therefore given by

$$v_0 \approx R_p \left|\Omega - \Omega_p\right|_{b_{min}} \approx \frac{3\Omega_p}{2} b_{min} = c_s \quad \rightarrow \quad b_{min} = \frac{2}{3} H_g, \tag{7.13}$$

where Eq. 7.6 have been used to express the relative velocity and the vertical hydrostatic equilibrium condition for the gas scale height at the planet location H_g ($c_s = H_g \Omega_p$, see Sect. 1.2.1.1). To include this pressure effect and the requirement that the impact parameter needs to be greater than the Hill radius, the specific excitation torque is often calculated from the expression given by Eq. 7.11, but with the following modified b_{min}

$$b_{min} = \max\left(\frac{2}{3} H_g, R_H\right). \tag{7.14}$$

Therefore, the tidal torque exerted by a fixed planet on an elementary gas ring per unit mass is given by

$$\partial_m \Gamma^T(R) = \operatorname{sgn}(R - R_p)\, \frac{4}{9\pi}\left(\frac{M_p}{M_\star}\right)^2 \Omega_p^2 R_p^2 \frac{R_p^4}{\Delta_g^4} \quad \text{with} \quad \Delta_g = \max\left(|R - R_p|, \frac{2}{3} H_g, R_H\right). \tag{7.15}$$

In a pressureless fluid such as a collection of dust particles (see Chapter 2), the value of Δ is therefore simply given by $\Delta_d = \max(|R - R_p|, R_H)$. The net effect of the angular momentum transfer between particle and the planet is that the gas is repelled from the orbit of the planet and that the planet itself is subject to a migration. Due to the conservation of the total angular momentum, the planet will change its specific angular momentum by $\delta j_p = -\delta j$. Therefore, as mentioned above, the planet moves outward (inward) under the influence of the gas inside (outside) its orbit (see Sect. 7.2), while the gas is pushed away from the planet, resulting in the formation of a gap (see Sect. 7.3).

[1]It is worth remarking that, since $(R - R_p)^{-4}$ is the most rapidly varying quantity, it can be safely assumed that this analysis is valid also with slightly not constant Σ_g.

7.1.2 Lindblad Torque

The linear theory of disc-planet interaction has been first introduced by Goldreich & Tremaine (1979, 1980) and is essentially based on evaluating the linear response of the flow to small tidally-induced perturbations. This analysis is valid for low-mass perturbed, since, in this limit, the morphological features and angular momentum content of the waves excited by the planet can be described in the framework of linear analysis of the perturbed fluid quantities. It can be shown that the linear perturbation analysis on the response of the fluid to the perturber is valid for planets with masses satisfying (Goodman & Rafikov 2001; Rafikov 2002)

$$M_\mathrm{p} \ll M_\mathrm{th} = \frac{c_\mathrm{s}^3}{\mathcal{G}\Omega_\mathrm{k}} = M_\star \left(\frac{H_\mathrm{g}}{R_\mathrm{p}}\right)^3, \tag{7.16}$$

where H_g is the disc height scale at the planet location. This limit mass, mostly known as thermal mass M_th, is the mass of a planet at which the Hill radius (Eq. 7.12) is comparable to the disc scale height H_g.

First of all, let us describe the effect of small perturbations on a pressureless disc characterised by differential rotation $\Omega(R)$. It can be shown that if the flow is weakly perturbed, a fluid particle oscillates around its orbit at equilibrium at a frequency given by the epicyclic frequency κ defined in Eq. 4.37 (Binney & Tremaine 1987). The dispersion relation of these oscillations is simply given by $\omega^2 = \kappa^2$, where it can be easily noticed the stabilising effect of the differential rotation when $\partial\left(\Omega R^2\right)/\partial R > 0$[2] (see Eq. 4.37). In the general case, whenever the frequency of any perturbed component of the hydrodynamic quantities as seen by a fluid particle matches the epicyclic frequency, a resonance occurs, inducing a strong disc response to the perturbations. The resonances induced by a perturbing gravitational potential in the disc are known as *Lindblad resonances* (Lindblad 1927).

Let us assume a flat two-dimensional non-self-gravitating axisymmetric pressureless Keplerian ($\kappa = \Omega_\mathrm{k}$) disc with an embedded protoplanet of mass $M_\mathrm{p}(\ll M_\mathrm{th})$ orbiting on a fixed circular and coplanar orbit (i.e. with zero eccentricity and inclination) with angular velocity $\Omega_\mathrm{p} = \sqrt{\mathcal{G}M_\star/R_\mathrm{p}^3}$. As previously assumed in Sect. 7.1.1, the typical random velocity (which, for the time being, it is denoted by c) has to be much smaller than the typical orbital velocity, in order to consider particles on circular orbit around the central star. Moreover, in order to simplify the wave equations, we adopt the WKB approximation: $m \ll \Omega_\mathrm{k} R/c$, where m is the azimuthal wavenumber of the perturbation and Ω_k/c is a measure of the minimum radial wavenumber of the perturbations that can be described according to the WKB approximation (corresponding to a maximum wavelength equal to the scale height $H = c/\Omega$). The gravitational potential of the planet in polar coordinates centered on the star is $\Phi_\mathrm{p}(R,\phi) = -\mathcal{G}M_\mathrm{p}/d$ where $d = \sqrt{R^2 + R_\mathrm{p}^2 - 2RR_\mathrm{p}\cos\left(\phi - \phi_\mathrm{p}\right)}$ is the distance from the planet. The potential is clearly a 2π-periodic function of the relative azimuthal position $\phi - \phi_\mathrm{p}$ where $\phi_\mathrm{p} = \Omega_\mathrm{p} t$ is the azimuth of the planet. Therefore, the perturbing potential can be expanded in a Fourier series of the variable $\phi - \phi_\mathrm{p}$, as

$$\Phi^\mathrm{p}(R,\theta,t) = \Re\left(2\sum_{m=0}^{\infty}\Phi_m^\mathrm{p}(R)\exp\left[im\left(\phi(t) - \Omega_\mathrm{p}t\right)\right]\right) = 2\sum_{m=0}^{\infty}\Phi_m^\mathrm{p}(R)\cos\left[im\left(\phi(t) - \Omega_\mathrm{p}t\right)\right], \tag{7.17}$$

[2]It can be clearly noticed from Eq. 4.37 that the epicyclic frequency is related to the radial variation of the density of angular momentum of the fluid. The epicyclic motion is therefore connected to the conservation of angular momentum. To simplify the analysis, let us describe the motion in the heliocentric frame. Assuming that the fluid particle is displaced towards larger values of R, the orbital velocity v_ϕ of the fluid element will decrease in order to conserve the angular momentum. The decrease v_ϕ result in a radial inward motion of the particle, since the centrifugal acceleration is not high enough to counterbalance the gravity. By contrast, if the particle is displaced towards smaller values of R, v_ϕ increases and the particles moves outward, restoring the initial position at equlibrium on a typical stabilising time of κ^{-1}. In Keplerian discs ($\kappa = \Omega_\mathrm{k}$) the rotation of the guiding center around the star and the epicyclic oscillations have the same frequency. Thus, the epicyclic trajectory is a closed ellipse in the heliocentric frame.

where the factor 2 corresponds to the full expansion of the Fourier series over $[-\infty, \infty]$ in m and $\Phi_m^{\rm p}(R)$ denotes the potential coefficient for each m, expressed by (Goldreich & Tremaine 1980)

$$\Phi_m^{\rm p}(R) = -\frac{2\mathcal{G}M_{\rm p}}{\pi R_{\rm p}} \int_0^{\pi} \left[1 + \left(\frac{R}{R_{\rm p}}\right)^2 - 2\left(\frac{R}{R_{\rm p}}\right)\cos\phi\right]^{-1/2} \cos\left(m\phi\right) d\phi. \tag{7.18}$$

For each m, a tidal torque arises due to the effects of the potential perturbation component of order m. As mentioned above, when the frequency of an individual potential component as seen by a fluid particle in the disc matches the epicyclic frequency $\Omega_{\rm k}$, the Lindblad resonance condition is satisfied. The frequency of each potential component, i.e. the time derivative of the argument of the cosine function in equation 7.17 along a particle's path $\phi(t)$, is given by $m\left(\Omega_{\rm k}(R) - \Omega_{\rm p}\right)$. The Lindblad resonances therefore correspond to disc radial locations $R_{\rm m}$ satisfying

$$m\left(\Omega_{\rm k}(R_{\rm m}) - \Omega_{\rm p}\right) = \pm\Omega_{\rm k}(R_{\rm m}) \quad \rightarrow \quad \Omega_{\rm k}(R_{\rm m}) = \frac{m}{m\pm 1}\Omega_{\rm p} \quad \rightarrow \quad R_{\rm m} = \left(\frac{m\pm 1}{m}\right)^{2/3} R_{\rm p}. \tag{7.19}$$

Lindblad resonances represent the so-called mean motion resonances of particles: while particles at the resonance locations describes m orbit around the central star, the planet describes $m \pm 1$ orbits. For $m \geq 2$, there are two Lindblad resonance locations for each azimuthal mode, a inner Lindblad resonance (ILR) occurring inside the planetary orbit for which the minus sign is taken in Eq. 7.19, and the outer Lindblad resonance (OLR) occurring outside the planetary orbit, corresponding to the plus sign in Eq. 7.19. For $m = 1$ there is only one outer Lindblad resonance. Moreover, another resonance occurs for $m\left(\Omega_{\rm k}(R) - \Omega_{\rm p}\right) = 0$. This resonance, mostly known as *corotation resonance* occurs at the planet location, where the local disc and planet orbital speeds are equal, $\Omega_{\rm k}(R) = \Omega_{\rm p}$. We will focus on the corotation torque in Sect. 7.1.2.3.

The Lindblad torque arising from the m-component of the potential is given by (Goldreich & Tremaine 1979; Meyer-Vernet & Sicardy 1987)

$$\Gamma_m^{\rm L} = {\rm sgn}\left(R - R_{\rm p}\right) \frac{\pi^2\Sigma}{3\Omega_{\rm k}\Omega_{\rm p}} \left(R\frac{d\Phi_m^{\rm p}}{dR} + \frac{2m^2\left(\Omega_{\rm k} - \Omega_{\rm p}\right)}{\Omega_{\rm k}}\Phi_m^{\rm p}\right), \tag{7.20}$$

which has to be computed at the m-resonance location $R = R_{\rm m}$. At each resonance location $R_{\rm m}$, m-fold density waves are launched, producing an angular momentum flux towards outer radii as they propagate away from the planet.

Importantly, in gaseous discs, it can be shown that density waves with different m can constructively interfer, leading to the formation of a one-armed spiral (Ogilvie & Lubow 2002, see Fig. 7.3), with a pitch angle given by

$$i = \tan^{-1}\left(\frac{m}{kR}\right) = \tan^{-1}\left(\frac{c_{\rm s}}{R\left|\Omega - \Omega_{\rm p}\right|}\right). \tag{7.21}$$

The shape of the wake in polar coordinates far from the planet position $(R_{\rm p}, \phi_{\rm p})$ is given by (Rafikov 2002; Muto et al. 2012)

$$\begin{aligned} \phi(R) &= \phi_{\rm p} - {\rm sgn}\left(R - R_{\rm p}\right)\frac{R}{H_{\rm g}}\left\{\left(\frac{R}{R_{\rm p}}\right)^{1+\beta}\left[\frac{1}{1+\beta} - \frac{1}{1-\alpha+\beta}\left(\frac{R}{R_{\rm p}}\right)^{-\alpha}\right]\right. \\ &- \left.\left(\frac{1}{1+\beta} - \frac{1}{1-\alpha+\beta}\right)\right\}, \end{aligned} \tag{7.22}$$

where α and β are power law exponents of the disc rotation angular frequency $\Omega(R) \propto R^{-\alpha}$ and sound speed $c_{\rm s} \propto R^{-\beta}$. This formula has been used to fit the shape of the observed spiral structure in order to infer the local disc conditions. As an example, by measuring the pitch angle of the spiral structure, one can gain insights on the local disc temperature (see Eq. 7.21) and,

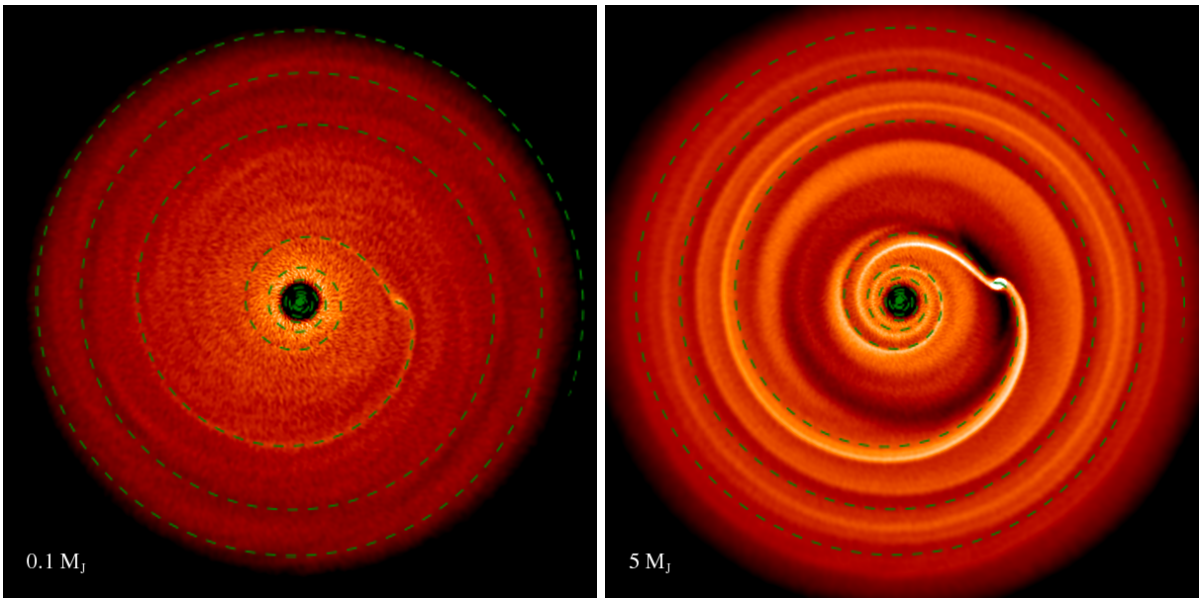

Figure 7.3: Surface density structure of disc computed by means of three dimensional SPH simulations of a disc with an embedded protoplanet with mass 0.1 (left) and 5 M_J (right) where M_J is the Jupiter mass. The disc model is described in Chapter 8. The green dashed curves are the spiral wakes from linear theory (Eq. 7.22). The two images are on different color scale. The one-armed spiral density structure is formed due to the constructive interference of different azimuthal modes (Ogilvie & Lubow 2002). Deviation from the linear theory are expected for $M_p > M_{th}$ (Zhu et al. 2015a; Fung & Dong 2015) where $M_{th} \approx 0.5\, M_J$ for the adopted disc model.

consequently, on the local disc aspect ratio (Muto et al. 2012; Benisty et al. 2015, see Sects. 4.3.2.1 and 4.3.2.2). Numerical simulations have shown that the shape of the wake induced by low mass planet, i.e. $M_p \ll M_{th}$, is well approximated by Eq. 7.22 (see the left panel of Fig. 7.3, Zhu et al. 2015a). For more massive planets, i.e. $M_p \gg M_{th}$, the density waves generated by the planet are strongly nonlinear as soon as they are produced and steepen to shocks very close to the planet. In this limit, the spiral shock will expand in both azimuthal directions away from the trajectory predicted by Eq. 7.22, resulting in a larger pitch angle of the spiral wake. Moreover, a secondary inner spiral arm appears with some azimuthal shift from the primary arm (see the right panel of Fig. 7.3, Zhu et al. 2015a). For high mass planets, the azimuthal separation between the primary and secondary spiral arm increases with the planetary to star mass ratio. In this context, Fung & Dong (2015) found that the azimuthal separation is $\propto (M_p/M_\star)^{0.2}$ up to $M_p/M_\star = 0.016$, beyond which the separation remains approximately constant at 180°.

Importantly, the waves excited at the outer (inner) Lindblad resonances carry an excess (deficit) amount of angular momentum compared with the gas, resulting in a gain (loss) of the angular momentum of the fluid when the waves are damped, forcing the gas outside (inside) the planetary orbit to move radially outward (inward) from the planet. As a result, a gap is formed under the action of the tidal torque. The schematic illustration of the gap formation and the position of the Lindblad resonances are shown in Fig. 7.4.

7.1.2.1 Torque cut-off

Eq. 7.19 shows that for $m \to \infty$ the Lindblad resonances locations accumulate at the corotation radius $R_m \to R_p$. However, this limit cannot be included in this analysis since the WKB approximation, i.e. $m \ll m_{crit} = \Omega R/c$, is no longer valid. By numerical computing the non-WKB equation for the wave equations, Goldreich & Tremaine (1980) found that for $m \gtrsim m_{crit}$, the total torque (see Fig. 3 and Eq. 92 in Goldreich & Tremaine 1980) decreases rapidly with m. For gaseous discs, the critical azimuthal mode number is simply given by $m_{crit} \sim R_p/H_g$ where H_g is the vertical scale height, $H_g = c_s/\Omega_k$. The suppression of all the resonances closer than

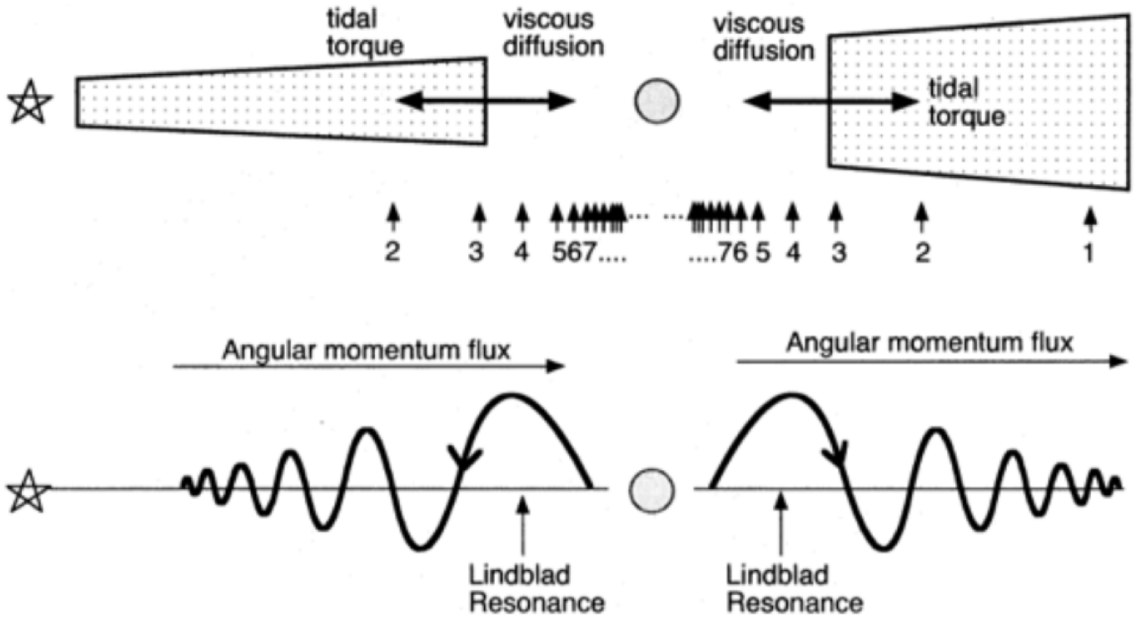

Figure 7.4: Sketch of the gap formation process under the action of the tidal and viscous forces. Arrows indicates the location of the Lindblad resonances at each azimuthal mode. Image taken from Takeuchi et al. (1996).

$\sim H_{\rm g}$ from the planet (i.e. those with azimuthal mode numbers $m > R_{\rm p}/H_{\rm g}$) suggests that the torque density maximum occurs at a radial distance from the planet given by $|R - R_{\rm p}| \approx H_{\rm g}$. Moreover, in gaseous discs, the resonance conditions is modified in order to account for the pressure effect, since the dispersion relation for linear perturbation in the WKB limit is given by (see the derivation for self-gravitating discs in Sect. 4.2.1)

$$m^2 \left[\Omega_{\rm k}(R) - \Omega_{\rm p}(R)\right]^2 = c_{\rm s}^2\, k^2 + \kappa^2, \tag{7.23}$$

where k is the radial wavenumber and κ is the epicyclic frequency defined in Eq. 4.37. The new term is related to dispersion relation describing the propagation of pressure supported waves: $c_{\rm s}^2 k^2$. It can be shown that the pressure effects shift the locations of Lindblad resonances from the nominal values expressed in Eq. 7.19 into new values satisfying

$$m\left(\Omega_{\rm k}(R_{\rm m}) - \Omega_{\rm p}\right) = \mp\Omega_{\rm k}(R_{\rm m})\sqrt{1 + m^2\left(\frac{H_{\rm g}}{R_{\rm p}}\right)^2}. \tag{7.24}$$

For $m \to \infty$ the radial location of the Lindblad resonances in the gas disc pile up at

$$\lim_{m\to\infty} R_{\rm m} = R_{\rm p} \pm \frac{2}{3} H_{\rm g}. \tag{7.25}$$

This prevents the divergence of the torque at the planet location, giving rise to the phenomenon known as *torque cut-off* (Goldreich & Tremaine 1980; Artymowicz 1993). However, this calculation breaks down if some Lindblad resonances falls into the corotation region, i.e. $|R - R_{\rm p}| \sim R_{\rm H}$ (D'Angelo & Lubow 2008). Moreover, in order to take into account the possible overlap of the coorbital region ($|R - R_{\rm p}| < R_{\rm H}$) and the region where the Lindblad resonances are suppressed ($|R - R_{\rm p}| < 2H_{\rm g}/3$), it is expected that the torque density maximum occurs at a radial distance from the planet given by $|R - R_{\rm p}| \approx \max(2H_{\rm g}/3, R_{\rm H})$, which has been already obtained by simple arguments in Section 7.1.1.

The pressure effect needs to be taken into account in the evaluation of the Lindblad torque acting on the disc (Eq. 7.20). The Lindblad torque in a gaseous disc is therefore given by (Artymowicz 1993; Ward 1997)

$$\Gamma_m^{\rm L} = {\rm sgn}\left(R - R_{\rm p}\right) \frac{\pi^2 \Sigma}{3\Omega_{\rm k}\Omega_{\rm p}} \frac{1}{\sqrt{1+\xi^2}\,(1+4\xi^2)} \left(R\frac{d\Phi_m^{\rm p}}{dR} + \frac{2m^2\left(\Omega_{\rm k} - \Omega_{\rm p}\right)}{\Omega_{\rm k}} \Phi_m^{\rm p}, \right), \tag{7.26}$$

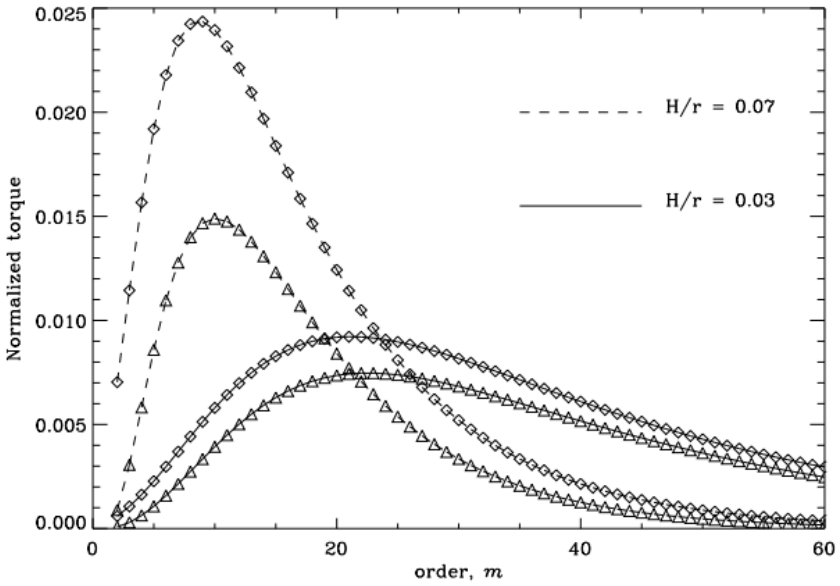

Figure 7.5: Inner (triangles) and outer (diamonds) Lindblad torque expressed in Eq. 7.26 for two different aspect ratios, as a function of m. The torques are normalized to $(M_\mathrm{p}/M_\star)^2 \Sigma_\mathrm{g} \Omega_\mathrm{p}^2 R_\mathrm{p}^4 (H_\mathrm{g}/R_\mathrm{p})^{-3}$. The ticker the disc, the smaller is the value over which the tidal torque decreases. Image taken from Masset & Kley (2006).

where $\xi = mH_\mathrm{g}/R_\mathrm{p}$. Fig. 7.5 shows the one-sided Lindblad torques (in absolute value) expressed in Eq. 7.26 for two different values of the local disc aspect ratio $H_\mathrm{g}/R_\mathrm{p}$. It can be clearly noticed that the tidal torque decreases as m increases for large $m \gtrsim R_\mathrm{p}/H_\mathrm{g}$ and, importantly, the torque cut-off occurs at larger m in thinner disc compared to ticker disc. Since higher m resonances are located closer to the planet, the excitation of density waves occurs close to the planet for thinner discs compared to ticker disc, as shown in Eq. 7.25. Moreover, the differences of the peaks of the inner and outer Lindblad torques appears to increase linearly with the disc aspect ratio. Therefore, while the one sided tidal torque scales with $(H_\mathrm{g}/R_\mathrm{p})^{-3}$ (the tidal torques in Fig. 7.5 are normalized to $(M_\mathrm{p}/M_\star)^2 \Sigma_\mathrm{g} \Omega_\mathrm{p}^2 R_\mathrm{p}^4 (H_\mathrm{g}/R_\mathrm{p})^{-3}$), the sum of the torques is proportional to $(H_\mathrm{g}/R_\mathrm{p})^{-2}$, as described in Sect. 7.2.1.

7.1.2.2 Torque density

The total Lindblad torque in the linear regime can be evaluated by summing over all individual components at each Lindblad resonances: $\Gamma = \sum_{m=1}^{\infty} \Gamma_m^\mathrm{L}$. Assuming that total tidal torque is a smooth function of R, the specific torque, i.e. the amount of torque that is exerted by a planet on the disc per unit radial distance and unit mass density can be simply evaluated. This can be done in gaseous discs since the excitation torque density produced by the planet is in principle non-zero everywhere in the disc, thanks to the finite radial extent of forcing at the Lindblad resonances (Rafikov & Petrovich 2012). The excitation tidal torque exerted by a fixed planet on an elementary gas ring per unit mass is given by

$$\partial_m \Gamma^\mathrm{L} = \mathrm{sgn}(R - R_\mathrm{p}) \frac{16 \widetilde{K}}{81 \pi} \left(\frac{M_\mathrm{p}}{M_\star} \right)^2 \Omega^2 R^2 \left(\frac{R}{\Delta_\mathrm{g}} \right)^4, \tag{7.27}$$

where

$$\widetilde{K} = \left[2K_0 \left(\frac{2}{3} \right) + K_1 \left(\frac{2}{3} \right) \right]^2 \approx 6.35, \tag{7.28}$$

where K_0 and K_1 are modified Bessel functions. As previously discussed (see Sect. 7.1.1), the value of Δ_g is $\sim \max\left(|R - R_\mathrm{p}|, 2H_\mathrm{g}/3, R_\mathrm{H}\right)$. This prescription for Δ_g has been extensively used in a variety of contexts, such as to model disc-planet interactions in protoplanetary discs

(e.g. Bryden et al. 1999; Lodato & Rice 2004; Crida et al. 2006) or simulate the evolution of supermassive black hole binaries embedded in a disc (e.g. Lodato et al. 2009).

It is worth pointing out that the expression of the torque density shown here differs from the one computed adopting the impulse approximation (Eq. 7.15) by a numerical factor of $4\widetilde{K}/9 \approx 2.83$ and the approximation $R = R_\mathrm{p}$. The total one-sided excitation torque can be evaluated by integrating Eq. 7.27 multiplied by $dm \approx 2\pi R \Sigma_\mathrm{g} d\Delta_\mathrm{g}$ in Δ_g from $\Delta_\mathrm{g,min}$ to ∞. Assuming $R = R_\mathrm{p}$ and $\Sigma_\mathrm{g} = \mathrm{const}$, the one-sided torque acting on the outer disc is given by

$$\Gamma^\mathrm{L} = \frac{32\widetilde{K}}{243}\left(\frac{M_\mathrm{p}}{M_\star}\right)^2 \Sigma_\mathrm{g}\Omega_\mathrm{p}^2 R_\mathrm{p}^4 \left(\frac{R_\mathrm{p}}{\Delta_\mathrm{g,min}}\right)^3 \quad \text{with} \quad \Delta_\mathrm{g,min} = \max\left(\frac{2}{3}H_\mathrm{g}, R_\mathrm{H}\right), \tag{7.29}$$

which recovers the same scaling relations with the fluid and planet properties as in Eq. 7.11.

It is worth pointing out that, using Eq. 7.27 (or 7.15) to mimic a torque acting on discrete locations might lead to mistakes if applied in regions where the Lindblad resonance locations are not enough spatially dense. Importantly, outside the $m = 2$ inner and the $m = 1$ outer resonances, located at $R_\mathrm{ILR} = 2^{-2/3}R_\mathrm{p}$ and $R_\mathrm{OLR} = 2^{2/3}R_\mathrm{p}$ respectively (see Eq. 7.19 and Fig. 7.4), the disc exerts no torque on the planet and vice-versa. Since $\partial_m \Gamma^\mathrm{L} \propto (M_\mathrm{p}/M_\star)^2$, for high mass planet the use of the tidal torque density expression shown in Eq. 7.27 (or 7.15) might lead to a non negligible contribution outside the resonances. In order to take into account the suppression of tidal forces outside the outermost resonances, gaussian cut-offs are commonly used (Lodato et al. 2009; Tazzari & Lodato 2015),

$$\partial_m \Gamma^\mathrm{L}(R) = \begin{cases} \frac{16\widetilde{K}}{81\pi}\left(\frac{M_\mathrm{p}}{M_\star}\right)^2 \Omega^2 R^2 \left(\frac{R}{\Delta_\mathrm{g}}\right)^4 \exp\left[-\left(\frac{R-R_\mathrm{OLR}}{w_\mathrm{OLR}}\right)^2\right] & \text{for } R \geq R_\mathrm{OLR}, \\ -\frac{16\widetilde{K}}{81\pi}\left(\frac{M_\mathrm{p}}{M_\star}\right)^2 \Omega^2 R^2 \left(\frac{R}{\Delta_\mathrm{g}}\right)^4 \exp\left[-\left(\frac{R-R_\mathrm{ILR}}{w_\mathrm{ILR}}\right)^2\right] & \text{for } R \leq R_\mathrm{ILR}, \end{cases} \tag{7.30}$$

where w_ILR and w_OLR determines the widths of the cut-offs which can be fine tuned in order to reproduce numerical results (e.g. Artymowicz & Lubow 1994; Tazzari & Lodato 2015).

7.1.2.3 Corotation Torque

Fluid particles experience a strong interaction at small separations from the planet. A fluid element initially on an orbit $|R - R_\mathrm{p}| < x_\mathrm{s}$ around the central star will experience horseshoe orbits, similar to those occurring in the restricted three body problem. Here, x_s denotes the so-called *separatrix* limit of the horseshoe region and its value is of the order of $\sim R_\mathrm{H}$ (Masset 2002; Masset et al. 2006). A fluid element initially on a orbit $R = R_\mathrm{p} + b$ with $0 < b < x_\mathrm{s}$ will be shifted from its initial location to $R = R_\mathrm{p} - b$. On this orbit, the particle angular velocity will be lower than the planet's and it moves away from the planet. After a revolution on this new orbit, on a timescale $\sim |\Omega(R) - \Omega_\mathrm{p}|$, the fluid element encounters the planet again and is shifted on its initial orbit at $R = R_\mathrm{p} + b$. This motion describes a trajectory that, in the corotating frame, has the shape of a horseshoe (e.g. Bate et al. 2003). Therefore, the particles approach the planet at the two ends of the horseshoe orbit leading to a periodic exchange of angular momentum between coorbital disc material and the planet. If the disc material does not interact with the rest of the disc (for example, in pressureless flows), the exchange of angular momentum averages out to 0 and there is no net torque exerted on the planet. However, in typical gaseous discs, an asymmetry in the local gas conditions exists between the two sides of the horseshoe region, leading to an angular momentum exchange and a net torque (known as corotation torque) acting on the planet given by

$$\Gamma_\mathrm{C}^\mathrm{L} = \frac{3}{4} x_\mathrm{s}^4 \Omega_\mathrm{p}^2 \Sigma_\mathrm{g,p} \frac{d\log\left(\Sigma_\mathrm{g}/B\right)}{d\log R}, \tag{7.31}$$

where $\Sigma_\mathrm{g,p}$ is the gas surface density at the planet location and $B = (1/2R)d(R^2\Omega)/dR = \kappa^2/(4\Omega)$ is the second Oort constant, physically half of the vertical component of the flow vorticity. Eq. 7.31 shows that the corotation torque depends on the radial gradient of specific vorticity across the horseshoe region.

7.1.3 Coupled evolution for the disc-planet system

Let us analyze the evolution of a gaseous thin viscous accretion disc in the presence of an embedded planet, in order to extend the previous analysis of the gas density and angular momentum evolution described in Section 1.2.1.3.

The presence of the planet in the disc produce an additional torque on the surrounding gas. In this Section, we refer to the tidal torque acting on the elementary gas ring per unit mass expressed in the two forms $\partial_m \Gamma^{\rm T}$ (Eq. 7.15) or $\partial_m \Gamma^{\rm L}$ (Eq. 7.27) as Λ. Since the tidal force can be thought as a force in the direction ϕ acting on a surface whose normal is in direction R (similarly to viscous forces in circular shear flows), the radial and the vertical components of Navier-Stokes remain unchanged, while the azimuthal component integrated in the vertical direction becomes (see Eq. 1.21)

$$\Sigma_{\rm g}\left[\frac{\partial v_{{\rm g},\phi}}{\partial t}+v_{{\rm g},R}\left(\frac{v_{{\rm g},\phi}}{R}+\frac{\partial v_{{\rm g},\phi}}{\partial R}\right)\right]=\nabla\cdot T|_{\phi}+\frac{\Sigma_{\rm g}\Lambda}{R}, \tag{7.32}$$

where the last term in the right hand side represent the tidal force per unit area. Similarly to what shown in Sect. 1.2.1.3, the equation of the conservation of angular momentum for a viscous fluid disc with an embedded planet is (see Eq. 1.22)

$$\frac{\partial}{\partial t}\left(\Sigma_{\rm g}Rv_{{\rm g},\phi}\right)+\frac{1}{R}\frac{\partial}{\partial R}(Rv_{{\rm g},R}\Sigma_{\rm g}Rv_{{\rm g},\phi})=\frac{1}{R}\frac{\partial}{\partial R}\left(\nu\Sigma_{\rm g}R^3\frac{\partial\Omega}{\partial R}\right)+\Sigma_{\rm g}\Lambda. \tag{7.33}$$

The steady state radial gas velocity can be obtained by multiplying each term of Eq. 7.32 by R and putting $\partial v_{{\rm g},\phi}/\partial t=0$, yielding

$$\begin{aligned} v_{{\rm g},R} &= \frac{1}{R\Sigma_{\rm g}\left(R^2\Omega\right)'}\left[\frac{\partial}{\partial R}\left(\nu\Sigma_{\rm g}R^3\frac{\partial\Omega}{\partial R}\right)+R\Sigma_{\rm g}\Lambda\right] &(7.34)\\ &= -\frac{3}{\Sigma_{\rm g}R^{1/2}}\frac{\partial}{\partial R}\left(\nu\Sigma_{\rm g}R^{1/2}\right)+\frac{2\Lambda}{\Omega_{\rm k}R}, &(7.35)\end{aligned}$$

where the second equality holds for Keplerian discs. This equation shows, as expected, that the planet repels the gas from its orbit, leading to the formation of a gap. Combining the angular momentum equation (Eq. 7.33) with the continuity equation (Eq. 1.17), it is possible to obtain the evolution equation for the gas surface density,

$$\begin{aligned} \frac{\partial\Sigma_{\rm g}}{\partial t} &= -\frac{1}{R}\frac{\partial}{\partial R}\left\{\frac{1}{\left(R^2\Omega\right)'}\left[\frac{\partial}{\partial R}\left(\nu\Sigma_{\rm g}R^3\frac{\partial\Omega}{\partial R}\right)+R\Sigma_{\rm g}\Lambda\right]\right\} &(7.36)\\ &= \frac{3}{R}\frac{\partial}{\partial R}\left[R^{1/2}\frac{\partial}{\partial R}\left(\nu\Sigma_{\rm g}R^{1/2}\right)\right]-\frac{2}{R}\frac{\partial}{\partial R}\left(\frac{\Sigma_{\rm g}\Lambda}{\Omega_{\rm k}}\right), &(7.37)\end{aligned}$$

where the second equality is valid for Keplerian discs. It can be noticed that, while the first term in the right hand side leads to a diffusive evolution of the gas surface density due to viscous torques (assuming that the viscosity is constant, see Sect. 1.2.1.3), the second term indicates an advection of the gas due to the tidal torque acting on the disc which repels the gas from the planet's orbit.

7.2 Planet migration

The combined effect of tidal torques acting on the planet leads to an exchange of angular momentum between the disc and planet, driving planetary migration. Assuming a planet embedded in a gaseous disc, there are two different regimes of planet migration, depending on the planet mass and the local gas properties. It is worth remarking that a third migration regime, still under debate, occurs for partially opened gap (Masset & Papaloizou 2003). However, the analysis shown here will be focused on the two well know migration regimes.

7.2.1 Type I migration

Low mass planets in highly viscous disc are not able to significantly affect the local disc structure and, as a result, cannot produce gaps (see Sect. 7.3). Due to the presence of the gas close to the planet and since the potential components $\Phi_m^{\rm p}$ decreases simmetrycally away from the planet, the exchange of angular momentum between the planet and the density waves excited at the closer resonances is expected to be highly efficient, driving planet migration. The outcome of this interaction is mostly known as *Type I migration*.

As already remarked, the disc response to outer (inner) Lindblad resonances leads to a lose (gain) of angular momentum of the planet, driving inward (outward) migration. The net torque acting on the planet by the surrounding gas produces a migration of the planet, with a sense of migration that depend upon which of the effect dominates. Importantly, the tidal interaction between the excited density waves at the $m^{\rm th}$ Lindblad resonance and the planet appears to be stronger at the outer resonances compared to the inner ones (see Fig. 7.5). First, from the definition of $R_{\rm m}$ (Eq. 7.19), it is evident that the outer resonances are closer to the planet orbit than the inner one, due to the decreasing radial profile of $\Omega_{\rm k}(R)$. Second, since $\Gamma_m^{\rm L} \propto 1/\Omega_{\rm k}(R_{\rm m})$, the disc response is larger for outer resonances compared to the outer ones, since the angular velocity decreases with radius. Third, the term in parentheses in Eq. 7.20 (or Eq. 7.26) increases with R and is therefore larger for outer resonances. Therefore, the net torque acting on the planet, mostly known as *differential Lindblad torque*, is negative, driving inward migration (Ward 1997).

A number of analytical and computational investigations have been carried out in the past in order to correctly compute the tidal torque onto the planet and infer the timescale of its migration (Goldreich & Tremaine 1979; Lin & Papaloizou 1979; Ward 1997; Tanaka et al. 2002). The analytical investigations are strictly based on the linear approximation, which is mostly valid for low mass planet and in the region close to the planet, where non linear effect can be neglected (see Sect. 7.1.2). Due to the high complexity in computing the tidal torque acting on the planet taking into account temperature, surface density effects and damping mechanisms, linear studies have been supplemented in recent years by many numerical studies in two and three dimensions based on simplified disc model. The most up-to-date estimate of the differential Lindblad torque has been computed by D'Angelo & Lubow (2010) by means of 3D nested-grid hydrodynamic simulations of planet-disc interaction. This result extends the previous 2D formulation of Paardekooper et al. (2010) and the 3D formulation of Tanaka et al. (2002), strictly valid for isothermal disc models. Assuming a disc model with a power law surface density and temperature radial profile, i.e. $\Sigma_{\rm g} \propto R^{-p}$ and $T \propto R^{-q}$, they computed a differential Lindblad torque acting on the planet in the form

$$\Gamma_{\rm dLR} = -\frac{1}{\gamma}\left(1.36 + 0.62p + 0.43q\right)\Gamma_0 \quad \text{with} \quad \Gamma_0 = \left(\frac{M_{\rm p}}{M_\star}\right)^2 \left(\frac{H_{\rm g}}{R_{\rm p}}\right)^{-2} \Sigma_{\rm p} R_{\rm p}^4 \Omega_{\rm p}^2, \tag{7.38}$$

where γ is the ratio of specific heats, $H_{\rm g}$ and $\Sigma_{\rm p}$ denotes the disc height scale and gas surface density at the planet location. The migration timescale can be simply calculated by evaluating the variation of the angular momentum of the planet $J_{\rm p} = M_{\rm p} R_{\rm p}^2 \Omega_{\rm p}$ under the action of the tidal torque $\Gamma_{\rm dLR} = dJ_{\rm p}/dt$. Assuming that the planet mass is stationary, the Type I migration timescale is therefore given by

$$t_{\rm migr,I} = \frac{R_{\rm p}}{|dR_{\rm p}/dt|} = \frac{J_{\rm p}}{2\,|\Gamma_{\rm dLR}|} \sim \left(\frac{H_{\rm g}}{R_{\rm p}}\right)^2 \frac{1}{M_{\rm p}\Omega_{\rm p}} \frac{M_\star^2}{\Sigma_{\rm p} R_{\rm p}^2}. \tag{7.39}$$

It can be noticed that the Type I migration increases with increasing temperature (or, equivalently, disc aspect ratio) and decreasing planet mass. Therefore, a large mass planet in cold disc region will quickly migrate inwards. The typical migration timescale for a $10 M_\oplus$ orbiting at 5 au in a Minimum Mass Solar Nebula (Weidenschilling 1977b) with $H_{\rm g}/R = 0.05$ is ~ 0.2 Myr, shorter than the disc lifetime. Using plausible disc models, it can be shown that the typical

migration timescales are shorted than the disc lifetime, challenging the planet formation theories. The fast Type I migration is still under debate and several theoretical and numerical works has been recently carried out, mostly focused on how to stop or slow down Type I migration. For example, the influence of the corotation resonance, magnetic effect and the dynamical and thermodynamical influence of the accreting material can slow down the inward Type I migration (see Sect. 3.1.4 of Baruteau et al. 2016 for a detailed summary of recent results).

7.2.2 Type II migration

If the planet is massive enough and if the angular momentum can be effectively deposited locally in the disc, the material inside (outside) the planet loses (gains) angular momentum and recedes from the planet, leading to the formation of a gap. The formation of the gap reduces the amount of gas close to the planetary orbit. This would weaken the exchange of angular momentum between the gas and the planet, reducing the migration efficiency. Contrarily to Type I migration, the planet appears to be locked in the gap. If the planet approaches the inner (outer) disc, the gas will induce a positive (negative) torque, which repels the planet outward (inward). In this regime, assuming that most of the angular momentum content is in the gas rather than in the planet, the viscous evolution of the disc drives the planet migration[3]. This regime of migration, mostly controlled by the viscous evolution of the planet-disc system is mostly known as *Type II migration*. The typical timescale of this migration is of the order of the viscous timescale (Eq. 1.28), with a typical speed in steady state of (Eq. 1.39)

$$v_{\rm migr,II} \approx -\frac{3\nu}{2R_{\rm p}} = -\frac{3}{2}\alpha\left(\frac{H_{\rm g}}{R_{\rm p}}\right)^2 v_{\rm k}, \tag{7.40}$$

where ν is the kinematical viscosity at the planet's location using the α-prescription (Eq. 1.47, Shakura & Sunyaev 1973). Contrarily to Type I migration, the speed of Type II migration depends on a few parameters and, importantly, appears to be independent on the mass of the planet. However, inertial effects of very massive planets would resist the viscous driving, halting the migration. In order to take into account the angular momentum content of both the planet and the surrounding gas, it is necessary to evaluate the angular momentum conservation (see Eq. 1.22) of the disc-planet coupled system. In this analysis, the tidal torque is assumed to be concentrated in a narrow annulus around the planet's orbit at radius $R_{\rm p}$. Consider the evolution of the angular momentum content of the annulus lying outside the planet's orbiting at local Keplerian angular velocity. Due to the tidal interaction with the planet, the annulus receives an amount of angular momentum from the satellite, producing a variation of its angular momentum given by

$$\left.\frac{\partial J_{\rm p}}{\partial t}\right|_{R=R_{\rm p}} = \frac{d}{dt}\left(M_{\rm p}R_{\rm p}^2\Omega_{\rm p}\right) = \frac{1}{2}M_{\rm p}\Omega_{\rm p}R_{\rm p}\dot{R}_{\rm p}. \tag{7.41}$$

Starting from the conservation of the angular momentum per unit area expressed in Eq. 1.22, the migration timescale of the planet can be evaluated by restating the angular momentum conservation of the annulus as (Ivanov et al. 1999)

$$\frac{\partial J_{\rm p}}{\partial t} + 2\pi R\dot{R}\Sigma_{\rm p}R^2\Omega{\rm k}\Big|_{R=R_{\rm p}} = 3\pi\nu\Sigma_{\rm g}R^2\Omega_{\rm k}\big|_{R=R_{\rm p}}, \tag{7.42}$$

that has been evaluated by multiplying Eq. 1.22 by $2\pi R_{\rm p}dR$. After a little algebra, the Type II migration timescale is given by (Syer & Clarke 1995; Ivanov et al. 1999)

$$t_{\rm migr,II} = \frac{M_{\rm p}+4\pi R_{\rm p}^2\Sigma_{\rm p}}{4\pi R_{\rm p}^2\Sigma_{\rm p}}\frac{2}{3}\frac{R_{\rm p}^2}{\nu(R_{\rm p})} = \frac{M_{\rm p}+M_{\rm g}(R_{\rm p})}{M_{\rm g}(R_{\rm p})}t_{\rm visc}(R_{\rm p}) = \frac{1+B}{B}t_{\rm visc}(R_{\rm p}), \tag{7.43}$$

[3]It is worth remarking that the viscous gas evolution can also point outward, especially at the outer disc's edge. If a planet forms near the outer edge of the gaseous disc, the viscous expansion of the gas will drive outward planetary migration.

where

$$B = \frac{M_{\rm g}\left(R_{\rm p}\right)}{M_{\rm p}}, \tag{7.44}$$

where $M_{\rm g}\left(R_{\rm p}\right) = 4\pi R_{\rm p}^2 \Sigma_{\rm p}$ is a measure of the local gas disc mass where $\Sigma_{\rm p}$ is the surface density of the unperturbed disc at the planet location. It can be clearly distinguished two limiting cases:

- **B $\ll$ 1: Planet-dominated regime**: the planet's inertia halts its migration and acts as a dam for the downstream fluid material, resulting in a strong accumulation of gas at the gap edge. For extremely massive planet, i.e. $B \to 0$, the timescale is $t_{\rm migr,II} = t_{\rm visc}/B \gg t_{\rm visc}$,
- **B $\gg$ 1: Disc-dominated regime**: the planet is less massive than the disc and tends to follow its viscous driven radial evolution. In this case the timescale approaches the viscous timescale $t_{\rm migr,II} \to t_{\rm visc}$.

Importantly, since B generally increases with radius, it is therefore much easier for a planet to inertially halt its migration in the inner disc region than the outer ones. Moreover, contrary to the Type I migration, the planet migration timescale does not depend on the functional form of the tidal torque.

These results on planetary migration may suggest a possible scenario to explain the observed population of hot Jupiters, a class of exoplanets that are physically similar to Jupiter and are located very close to the central star, i.e. $R_{\rm p} < 0.1$ au (Armitage et al. 2002). According to the widely accepted scenarios of planet formation, it appears rather difficult to form *in situ* such massive objects (Lin et al. 1996; Bodenheimer et al. 2000). Therefore, planetary migration might be a valuable mechanism to allow mass planet, which are typically formed at larger radii (see Sect. 4.2.6), to populate the inner disc regions. However, the details on planetary migration is still under debate in literature and further analysis needs to be carried out in order to fully understand how the local gas condition, planet accretion and the presence of additional massive planets in the system can lead to different migration regimes (Kley & Nelson 2012).

7.3 Gap opening in gas discs

As previously mentioned, the planet moves outward (inward) under the influence of the gas inside (outside) its orbit, while the gas is pushed away from the planet, resulting in the formation of a gap structure. This Section is focused on the gap opening process in gaseous disc with a particular focus on the link between the gap morphology with the planet and local fluid properties.

7.3.1 Gap opening criteria

The question of criteria for gap formation in gas discs has been explored in a number of work (Ward & Hourigan 1989; Lin & Papaloizou 1993; Crida et al. 2006; Crida & Morbidelli 2007; Duffell & MacFadyen 2013). The first criterion for gap opening in gaseous disc is based on the requirement that a strong shock forms within a scale height of the planet's orbit (Lin & Papaloizou 1993),

$$R_{\rm H} \gtrsim H_{\rm g} \quad \text{or, equivalently,} \quad \frac{M_{\rm p}}{M_\star} \gtrsim \left(\frac{M_{\rm p}}{M_\star}\right)_{\rm th} \sim 3\left(\frac{H_{\rm g}}{R_{\rm p}}\right)^3. \tag{7.45}$$

This first criterion is mostly referred to as "thermal criterion" or "strong shock limit" and states that the planet's Hill radius needs to be larger than the disc scale height. The limit mass according to this criterion corresponds to, as described in Sect. 7.1.2, a planetary mass greater that three times the thermal mass introduced in Eq. 7.16 which represent the threshold mass over which the density waves shock instantaneously where they are excited. Therefore, if

the planet satisfies this criterion, it is reasonably expected that the excitation tidal torque is comparable to the deposition torque due to the strong shocks that efficiently damp the density waves excited at Lindblad resonance locations. Moreover, this condition allows to treat the planet-disc interaction two-dimensionally.

A second criterion is based on the requirement that gap opening should be faster than the viscous refilling (Lin & Papaloizou 1993). This criterion assumes that the first criterion is satisfied by the planet, since it is based on the fact the tidal torque is efficiently deposited in the disc. It can be evaluated by comparing the time required to evacuate all the gas contained in the gap with the timescale for gap closing driven by viscosity (Armitage 2010). It is reasonably expected that, taking into account the location of the maximum of the tidal torque (see Eq. 7.29), a gap of half width $\Delta_{\rm g,min} = \max\left(2H_{\rm g}/3, R_{\rm H}\right)$ is roughly the smallest gap that can be opened. Since the validity of this criterion requires that the planet satisfy the strong shock criterion, $R_{\rm H} \gtrsim H_{\rm g}$, the gap width in this analysis can be taken equal to $R_{\rm H}$, consistent with the reasonable expectation that the width of the gap for high mass planets is dependent on the mass. The timescale to evacuate all the angular momentum between $R_{\rm p}$ and $R_{\rm p} + R_{\rm H}$ is

$$\Delta J = 2\pi R_{\rm p} R_{\rm H} \Sigma_{\rm g} \left.\frac{\mathrm{dj}}{\mathrm{dR}}\right|_{R_{\rm p}} R_{\rm H} = \pi R_{\rm p} R_{\rm H}^2 \Sigma_{\rm g} v_{\rm k}(R_{\rm p}), \tag{7.46}$$

where $j(R) \equiv R v_{\rm k}(R)$ denotes the specific angular momentum. The typical time to evacuate all the gas in this region is

$$t_{\rm open} = \frac{\Delta J}{|\mathrm{dJ/dt}|}, \tag{7.47}$$

where $|\mathrm{dJ/dt}|$ is the one-sided total torque expressed in Eq. 7.11 or 7.29. The gap closing timescale driven by the viscous-refilling effect is given by (see Eq. 1.28)

$$t_{\rm close} = \frac{R_{\rm H}^2}{\nu}. \tag{7.48}$$

Using the formulation of the tidal torque given, for example, by the impulse approximation (Eq. 7.11), the criterion is given by

$$\frac{M_{\rm p}}{M_\star} \gtrsim \left(\frac{M_{\rm p}}{M_\star}\right)_{\rm visc, R_H} = \frac{3}{8}\pi \left(\frac{H_{\rm g}}{R_{\rm p}}\right)^2 \alpha, \tag{7.49}$$

where we have used the α-prescription (see Eq. 1.47, Shakura & Sunyaev 1973).

Another important gap opening criterion is the so-called "inertial limit", which take into account the planet migration (Li et al. 2009; Yu et al. 2010). In this case, the gap opening rate must be faster than the migration rate across its own horseshoe region. Assuming a power law profile for the unperturbed gas density $\Sigma_{\rm g} \propto R^{-p}$, Ward (1997) calculated this limit as

$$\frac{M_{\rm p}}{M_\star} \gtrsim \left(\frac{M_{\rm p}}{M_\star}\right)_{\rm migr} \sim \frac{7}{8} p^2 \frac{H_{\rm g}}{R_{\rm p}} \left(\frac{\pi H_{\rm g}^2 \Sigma_{\rm g}(R_{\rm p})}{M_\star}\right). \tag{7.50}$$

Comparing the strong shock limit with the inertial one, it can be shown that for a disc with profile $p = 3/2$ the strong shock limit is stricter if $M_\star = 0.7\pi R_{\rm g}^2 \Sigma_{\rm g}(R_{\rm p})$, which is satisfied in typical circumstellar discs. Thus, planets satisfying the strong shock criterion will automatically satisfy the Ward (1997)'s criterion.

Since the previously mentioned criteria are based on the assumption that the exchange of angular momentum between the waves and the gas is instantaneous, these limits for the mass might be considered sufficient conditions but not necessary (Duffell & MacFadyen 2013). Generally, the spiral density wave generated by the planet can affect the structure of the disc once the angular momentum that they carry is transferred to the disc material. The transfer of angular momentum from the wave to local fluid elements can be accomplished by damping

mechanisms, such as viscosity and/or shocks (e.g. Goldreich & Nicholson 1989; Goodman & Rafikov 2001; Rafikov 2002; Kanagawa et al. 2015b). Therefore, the density waves excited by planets radially propagate across the gaseous disc, leading to a non-local deposition of the tidal torque. As a result, the tidal interaction between the gaseous disc and the planet would be weaker at the resonances locations because the angular momentum is carried away by density waves and deposited in a wider region of the disc. The portion of the tidal torque carried away by the density waves, introduced by Crida et al. (2006) and mostly known as "pressure torque", depends on the relative radial gradient of the azimuthally averaged gas surface density. Starting from this evidence, Crida et al. (2006) derived a semi-analytical criterion based on the balance between pressure, gravitational and viscous torques for a planet on a fixed circular orbit. This criterion, defining a gap as a reduction of 10% of the unperturbed density at the planet's location, is given by

$$\frac{3}{4}\frac{H_\mathrm{g}}{R_\mathrm{p}} + \frac{50\alpha M_\star}{M_\mathrm{p}}\left(\frac{H_\mathrm{g}}{R_\mathrm{p}}\right)^2 \lesssim 1. \tag{7.51}$$

However, due to the nonlinear wave steepening, even a low mass planet can open gaps in inviscid discs, in contradiction to the "thermal criterion" and the Crida's criterion for gap opening (Duffell & MacFadyen 2012, 2013). In fact, recent analytical and numerical results have shown that non-migrating planets can violate these criteria and open gaps on longer timescale (Dong et al. 2011; Duffell & MacFadyen 2012, 2013; Zhu et al. 2013). Duffell & MacFadyen (2013) found an empirical criterion that can be evaluated starting from the same timescale arguments used above (Eq. 7.49) but taking into account the case where the planet's mass is low enough that the planet's Hill radius is smaller than the scale height. In this case, the condition $t_\mathrm{open} = t_\mathrm{close}$ with $\Delta_\mathrm{g,min} = 2H_\mathrm{g}/3$ leads to the following criterion

$$\frac{M_\mathrm{p}}{M_\star} \gtrsim \left(\frac{M_\mathrm{p}}{M_\star}\right)_{\mathrm{visc},H_\mathrm{g}} = \pi^{1/2}\alpha^{1/2}\left(\frac{H_\mathrm{g}}{R_\mathrm{p}}\right)^{5/2}. \tag{7.52}$$

However, numerical simulations have shown reasonable agreement with Crida et al. (2006) for large viscosity and planet mass (Duffell & MacFadyen 2013). This can be explained by the fact that the transfer of angular momentum for waves excited by high mass planet or by low mass planets in highly viscous disc occurs nearly at the same positions where the waves are generated. Therefore, in highly viscous discs, the criterion based on the equilibrium between tidal and viscous torque are approximately valid (Eqs. 7.49 and 7.51).

Moreover, the assumption that the planet remains on a fixed orbit is used in most of works in this area, finding usually good agreement with theoretical predictions. However, Malik et al. (2015) have shown that the migration may affect the ability of the planet to carve gaps, and the critical mass given by all these criteria might be underestimated. Importantly, the timescale for gap opening is longer for smaller planets (Zhu et al. 2013), implying that migration would play an important role in the gap opening process. This suggests that the criterion for gap opening in the gas should be based on a competition between the timescale for gap opening, the viscous timescale and the migration timescale, resulting in a stricter condition for gap opening in the gaseous discs compared to previously mentioned criteria.

Although significative efforts have been made in the last decades, the gap opening criterion in gaseous discs is still under debate. From a numerical point of view, because gap opening is still possible when the Hill radius is smaller than the scale height (i.e. violating the strong shock limit, Eq. 7.45), three-dimensional calculations are necessary to fully capture the disc-planet interaction in the low-mass regime.

7.3.2 Gap shape

The gap shape of the gas density structure affected by a fixed and non accreting planet can be easily computed by comparing the tidal and viscous torque for an elementary ring of gas (Varnière et al. 2004). Let us assume that the excitation torque introduced in Eq. 7.27 corresponds to

the effective exchange of angular momentum between the density waves and the fluid at the points of waves generation. Assuming steady state, the viscous torque must be equal to the tidal torque on every elementary ring of gas. In Keplerian disc, the net viscous torque acting on a elementary ring at distance R from the central object is given by (see Sect. 1.2.1.3)

$$\partial_R \Gamma^\nu(R) = 2\pi \frac{\partial}{\partial R} T_{R\phi} = 2\pi \frac{\partial}{\partial R}\left(\nu \Sigma_\mathrm{g} R^3 \frac{\partial \Omega_\mathrm{k}}{\partial R}\right) = -3\pi R \Sigma_\mathrm{g} \nu \Omega_\mathrm{k}\left[\frac{R}{\Sigma_\mathrm{g}}\frac{\partial \Sigma_\mathrm{g}}{\partial R} + \frac{1}{2}\right], \tag{7.53}$$

where it is assumed that ν is constant with R. The tidal torque on the elementary ring is given by

$$\partial_R \Gamma^\mathrm{T}(R) = 2\mathrm{sgn}(R - R_\mathrm{p}) f \left(\frac{M_\mathrm{p}}{M_\star}\right)^2 \Sigma_\mathrm{g} \Omega_\mathrm{p}^2 R R_\mathrm{p}^2 \left(\frac{R_\mathrm{p}}{\Delta_\mathrm{g}}\right)^4, \tag{7.54}$$

which has been obtained by multiplying Eq. 7.27 (adopting $R = R_\mathrm{p}$) with $2\pi R \Sigma_\mathrm{g}$ and where $\Delta_\mathrm{g} = \max\left(|R - R_\mathrm{p}|, 2H_\mathrm{g}/3, R_\mathrm{H}\right)$. Here, $f = 4/9\pi$ or $16\widetilde{K}/81\pi$ depending on the use of the functional form of the torque density expressed in Eq. 7.15 or 7.27, respectively. Matching the tidal and viscous torque, we obtain a differential equation in Σ_g (Varnière et al. 2004; Crida et al. 2006) given by

$$\frac{1}{\Sigma_\mathrm{g}}\frac{\partial \Sigma_\mathrm{g}}{\partial R} = \frac{\partial_R \Gamma^\mathrm{T}(R)}{3\pi\alpha R^2 H_\mathrm{g}^2 \Omega_\mathrm{k}^2 \Sigma_\mathrm{g}} - \frac{1}{2R}, \tag{7.55}$$

using the Shakura & Sunyaev (1973)'s α-prescription (Eq. 1.47).

The gap shape can be analytically inferred by assuming that, on the local scales of interest, Σ_g and Δ_g^{-4} are the most rapidly varying quantities, compared to R and Ω_k. Therefore, in Eq. 7.53, the second term in the brackets is neglected and all the terms in Eq. 7.54 are assumed to be constant with R except for the term $\propto \Delta_\mathrm{g}^{-4}$. The resulting equation can be integrated to obtain the gap profile (Lubow & D'Angelo 2006; Kanagawa et al. 2015b)

$$\Sigma_\mathrm{g}(R) = \Sigma_{\mathrm{g},0} \exp\left[-\frac{2}{9} f \left(\frac{M_\mathrm{p}}{M_\star}\right)^2 \alpha^{-1} \left(\frac{H_\mathrm{g}}{R_\mathrm{p}}\right)^{-2} \left(\frac{R_\mathrm{p}}{\Delta_\mathrm{g}}\right)^3\right], \tag{7.56}$$

where $\Sigma_{\mathrm{g},0}$ is the unperturbed disc surface density at $R = R_\mathrm{p}$.

The width of the gap Δ_g can be simply estimated by comparing the opening effect due to the tidal torque with the closing one induced by viscous forces. In Keplerian discs, the specific viscous torque acting on a annulus at distance R and width Δ_g is given by (see Sect. 1.2.1.3)

$$\left|\frac{1}{R\Sigma_{\mathrm{g},0}\Delta_\mathrm{g}}\frac{\partial}{\partial R} T_{R\phi}\right| = \frac{3}{2}\frac{R\Omega_\mathrm{k}}{\Delta_\mathrm{g}}\nu, \tag{7.57}$$

where $\Sigma_{\mathrm{g},0}$ is the unperturbed surface density profile. Equating the specific tidal torque adopting $R - R_\mathrm{p} = \Delta_\mathrm{g}$ with the viscous torque, the width of the gap is given by

$$\left(\frac{\Delta_\mathrm{g}}{R}\right)^3 = \frac{2}{3} f \left(\frac{M_\mathrm{p}}{M_\star}\right)^2 \frac{\Omega_\mathrm{k} R^2}{\nu} = f\left(\frac{M_\mathrm{p}}{M_\star}\right)^2 Re, \tag{7.58}$$

where Re is the Reynolds number of the flow defined as (see Eq. 1.29)

$$Re = \frac{2}{3}\frac{R^2\Omega_\mathrm{k}}{\nu} \sim \alpha^{-1}\left(\frac{H_\mathrm{g}}{R_\mathrm{p}}\right)^{-2}. \tag{7.59}$$

It can be noticed that, due to the dependence on viscosity, a low mass planet might not manage to overwhelm the intrinsic viscous forces, while for low viscosity, the gap is easier to be carved and its width is larger. Note that the same result could be inferred by analyzing the shape of the surface density profile across the gap expressed in Eq. 7.56. As previously mentioned, it is reasonably expected that a gap of half width $\Delta_{\mathrm{g,min}} = \max\left(2H_\mathrm{g}/3, R_\mathrm{H}\right)$ is roughly the smallest

gap that can be opened. By comparing the gap size estimate above with $\Delta_{\rm g,min}$, two additional criterion for gap opening can be obtained,

$$\Delta_{\rm g} \gtrsim R_{\rm H} \quad \rightarrow \quad \frac{M_{\rm p}}{M_\star} \gtrsim \frac{1}{3fRe} \sim \alpha \left(\frac{H_{\rm g}}{R_{\rm p}}\right)^2, \tag{7.60}$$

$$\Delta_{\rm g} \gtrsim \frac{2}{3} H_{\rm g} \quad \rightarrow \quad \frac{M_{\rm p}}{M_\star} \gtrsim \frac{2}{3} \left(\frac{H_{\rm g}}{R_{\rm p}}\right)^{3/2} \left(\frac{1}{fRe}\right)^{1/2} \sim \left(\frac{H_{\rm g}}{R_{\rm p}}\right)^{5/2} \alpha^{1/2}. \tag{7.61}$$

However, these two criteria are embodied in the criteria expressed in Eqs. 7.49 and 7.52, since in both cases the ability of the planet to carve a gap is analyzed in terms of tidal versus viscous torque for a gap width equal to $\Delta_{\rm g,min}$.

Taking into account the non-local deposition of the tidal torque due to the effect of the propagation of the pressure supported density waves, Crida et al. (2006) derived a radial profile of the surface density across the gap carved by a non-migrating planet. The equilibrium between all the torques acting on the fluid yields a surface density profile satisfying

$$\frac{R_{\rm H}}{\Sigma_{\rm g}} \frac{\partial \Sigma_{\rm g}}{\partial R} = \frac{\Gamma_{\rm crida} - \frac{3}{4}\alpha c_{\rm s}^2}{\left(\frac{H_{\rm g}}{R_{\rm p}}\right)^2 R R_{\rm p} \Omega_{\rm p}^2 a'' + \frac{3}{2}\alpha c_{\rm s}^2 \frac{R}{R_{\rm H}}}, \tag{7.62}$$

where $\Gamma_{\rm crida}$ is the tidal torque on an elementary ring directly measured in their simulation and a'' is a dimensionless function. However, the semi-analytic approach performed by Crida et al. (2006) does not take into account the nonlinear wave steepening of the density waves and the consequent shocks which, in inviscid flows, can occur far from the planet. However, the transfer of angular momentum from the planet to the disc by shocks is an inherently non-local process, since waves propagate some distance before shocks occur. This suggests that if the non locality of the tidal torque deposition is not taken into account properly, the depth (width) of the gap predicted by these studies would be overestimated (underestimated).

Starting from the non-local analysis of the tidal torque deposition performed by Goodman & Rafikov (2001) and Rafikov (2002), Duffell (2015) derived a simple analytical model for the gap shape. Assuming a Keplerian disc in steady state (see Sect. 1.2.3), the profile of the gap is given by (see further details in Duffell 2015)

$$\Sigma_{\rm g}(R) = \Sigma_{\rm g,0}(R) \left(1 - \frac{f(R) K/3\pi}{1 + f_0 K/3\pi} \sqrt{\frac{R_{\rm p}}{R}}\right) \quad \text{with} \quad K \equiv \left(\frac{M_{\rm p}}{M_\star}\right)^2 \alpha^{-1} \left(\frac{H_{\rm p}}{R_{\rm p}}\right)^{-5}, \tag{7.63}$$

where $\Sigma_{\rm g,0}(R)$ is the surface density of the unperturbed state and $f(R)$ is a function related to the deposition of tidal torque in the regions where the waves are dissipated (expressed in Eq. 15 of Duffell 2015 and derived in Goodman & Rafikov 2001) with $f_0 = f(R_{\rm p})$, whose value has been derived to best match the outcome of numerical simulations. By comparing the predicted gap shape with the gap computed by 2D numerical simulations, Duffell (2015) found that the shape of the gap is reasonably well-captured by the analytical model. However, once the gap becomes deep enough, numerical simulations show a clear deviation from the analytical model, probably due to the fact that $\Phi(R)$ was computed assuming that the excitation and propagation of waves occurs on an underlying unpertubed profile of the surface density. Importantly, the scaling of the gap depth ($\Sigma_{\rm g}(R = R_{\rm p}) \propto K^{-1}$) with parameters has been tested by using numerical calculations, showing a good consistency to empirical results, especially for low Neptune-like masses (Duffell & MacFadyen 2013; Fung et al. 2014; Kanagawa et al. 2015a). Moreover, for Jupiter-like planets, Fung et al. (2014) found $\Sigma_{\rm gap} \propto (M_{\rm p}/M_\star)^{-2.2}\, \alpha^{1.4}\, (H_{\rm g}/R)^{6.6}$, while for brown dwarfs $\Sigma_{\rm gap} \propto (M_{\rm p}/M_\star)^{-1}\, \alpha^{1.3}\, (H_{\rm g}/R)^{6.1}$.

Fig. 7.6 shows the comparison between the gap profiles given by the solutions of Eq. 7.55, Eq. 7.62 and the one derived by Duffell (2015) expressed in Eq. 7.63. First of all, the depth is significantly lower in the two cases of gap profiles evaluated taking into account the non-local deposition of the tidal torque compared (dashed line, Eq. 7.62 and dotted line, Eq. 7.63) to the

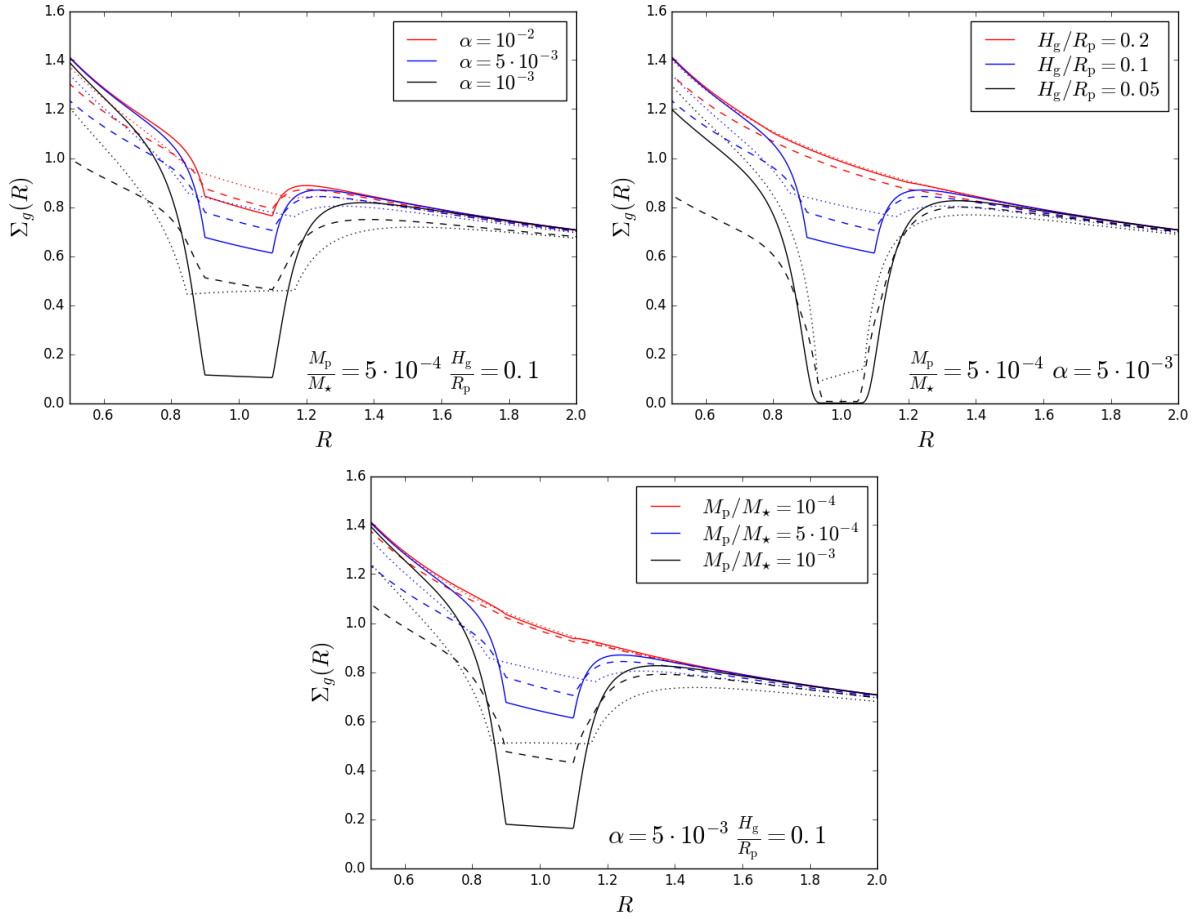

Figure 7.6: Radial profile of the surface density evaluated by solving (solid) Eq. 7.55 with $f = 16\widetilde{K}/81\pi$, (dashed) Eq. 7.62 and (dotted) Eq. 7.63, adopting different values of α (upper-left), H_g/R_p (upper-right) and $M_p/M_\star$ (bottom). The tidal torque is assumed to be null in the region $|R - R_p| < \Delta_{g,min} = H_g$. The disc model adopted here is characterised by a constant aspect ratio radial profile and the boundary condition is imposed to be $\Sigma_g(R = 2) = 1/\sqrt{2}$ in normalized units. The unperturbed radial profile, i.e. $\partial_R \Gamma^T(R) = 0$, is $\propto R^{-1/2}$ (see Eq. 7.55).

one derived by assuming instantaneous wave damping (solid line, Eq. 7.55). In some cases, the width of the gap computed according to Duffell (2015)'s analysis is slightly larger with respect to the other cases. This is due to the fact that the non local deposition of the tidal torque, weakens the intensity of the tidal torque, resultin in a shallower and wider gap, especially for low viscosity and low planet masses. Therefore, the tidal torque is less effective in carving the gap but, at the same time, act on a wider region around the planet (for more details see Crida et al. 2006 and Duffell 2015).

However, the problem of the gap morphology is still an active area of research and further numerical modelling is needed to comprehensively figure out the gap opening process and give predictive links between the gap shape and the local fluid and planet properties.

Chapter 8

Two Mechanisms for Dust Gap Opening in Discs: an Opening Criterion

Based on the papers by ***Giovanni Dipierro****, Guillaume Laibe, Daniel J. Price and Giuseppe Lodato: "Two mechanisms for dust gap opening in protoplanetary discs" published in Monthly Notices of the Royal Astronomical Society Letters,* ***459****, L1-L5 (2016) and by* ***Giovanni Dipierro*** *and Guillaume Laibe: "An opening criterion for dust gaps in protoplanetary discs" submitted to Monthly Notices of the Royal Astronomical Society.*

Recent high-resolution ALMA and SPHERE observations have provided evidence that a large variety of ring-like structures exist in protoplanetary disc, regardless of their age (see Chapter 3, ALMA Partnership et al. 2015; Nomura et al. 2016; Andrews et al. 2016; Canovas et al. 2016; van der Plas et al. 2016; de Boer et al. 2016; Ginski et al. 2016; van Boekel et al. 2016). Most of the theory on planet-induced gaps in discs has considered purely gaseous discs (see Sect. 7.3.1), with dust responding to the gas distribution mainly via drag. As a result, there is one physical mechanism discussed in the literature for explaining the origin of planet-induced gaps and ring-like structures in dusty discs: High mass planets exert a tidal torque that overpowers the local viscous torque and forms a gap in the gas (Papaloizou & Lin 1984; Bryden et al. 1999). Trapping of dust at the pressure maxima (see Sect. 2.5) prevents the large grains from being accreted onto the planet and produces a deep structure in the dust (Paardekooper & Mellema 2004, 2006; Fouchet et al. 2007, 2010), easily detectable by ALMA (Pinilla et al. 2012b, 2015; Gonzalez et al. 2012, 2015).

In this Chapter we study the tidal interaction between the planet and the dust component of a protoplanetary disc. We describe two distinct physical mechanisms for dust gap opening by embedded planets in protoplanetary discs based on the symmetry of the drag-induced motion around the planet: I) A mechanism where low mass planets, that do not disturb the gas, open gaps in dust by tidal torques assisted by drag in the inner disc, but resisted by drag in the outer disc; and II) The usual, drag assisted, mechanism where higher mass planets create pressure maxima in the gas disc which the drag torque then acts to evacuate further in the dust. The first mechanism produces gaps in dust but not gas, while the second produces partial or total gas gaps which are deeper in the dust phase. *Dust gaps do not necessarily indicate gas gaps.*

Moreover, we focus on the gap opening process by Mechanism I by deeply analyzing the dust dynamics in disc with embedded protoplanets, in order to derive a grain size-dependent criterion for opening dust gaps in viscous protoplanetary discs. Surprisingly, the minimum planet mass required to open a gap in the *dust* disc has not been clearly identified yet. In this Chapter we

revisit the theory of dust drift developed by Nakagawa et al. (1986) to include disc-planet tidal interactions and viscous forces. We apply this formalism to discuss the differences in the gap opening process between dust and gas as a function of the size of grains. We provide a necessary and sufficient condition for the minimum planet mass able to carve a dust gap, comparing the opening and closing effects of tides and gas drag. We also provide an estimate of the location of the outer edge of the dust gap, which can be useful to estimate the mass of the planet from high-resolution millimetre observations. We also provide an estimate of the minimum Stokes number above which low mass planets are able to carve gaps only in the dust. The criterion is benchmarked against 3D Smoothed Particle Hydrodynamics simulations of dusty discs hosting an embedded planet. The minimum planet mass to carve dust gaps and the position of the outer gap edge predicted by our analysis can be useful to estimate the mass of the planets responsible for the gap-like structures observed in protoplanetary discs.

8.1 Methodology

8.1.1 Dust and gas simulations

We use the PHANTOM SPH code (Price & Federrath 2010; Lodato & Price 2010; Price 2012; Nixon, King, & Price 2013) to perform global 3D simulations of dusty discs containing one or more embedded planets. The three processes involved in the gap-opening process — gas viscosity, gravity and drag — are computed consistently. Artificial viscosity is applied, calibrated to mimic the viscous transport of gas described with a Prandtl-like model of turbulence (Lodato & Price 2010). The central star and the planets are modelled using sink particles. Sink particles are allowed to both accrete and migrate as a consequence of their interactions with the disc (Bate, Bonnell, & Price 1995) and their gravitational interaction. In particular the gravity between the sink particles and the gas and dust particles in the disc is modelled.

Drag between gas and dust is computed using the two-fluid algorithm described in Laibe & Price (2012b). The algorithm has been extensively benchmarked on test problems with known analytic solutions (Laibe & Price 2011, 2012b; Price & Laibe 2015). We additionally perform simulations where gravity from the planets is artificially turned off in the dust. This allows us to contrast our results with a model where the dust is assumed to respond to the gas only via the drag torque.

8.1.2 Initial conditions

We setup a disc as in Lodato & Price (2010). We assume a central star of mass $1.3\,M_\odot$ surrounded by a gas disc made of 5×10^5 gas particles and a dust disc made of 3×10^5 dust particles. The two discs extend from $r_{\rm in}=1$ au to $r_{\rm out}=120$ au. We model the initial surface density profiles of the discs using power-laws of the form $\Sigma(R)=\Sigma_{\rm in}(R/R_{\rm in})^{-p}$. We adopt $p=0.1$ and set $\Sigma_{\rm in}$ such that the total gas mass contained between $R_{\rm in}$ and $R_{\rm out}$ is 0.0002 M$_\odot$. We assume 1 mm dust grains with an intrinsic grain density of 1 g/cm^3, corresponding Stokes number (the ratio between the stopping time and the orbital timescale), St ~ 10. The initial dust-to-gas ratio is 0.01 and St $\propto 1/\Sigma_{\rm g}\sim r^{0.1}$ in the disc. We simulate only the inner part of the disc since this is what can be observed with ALMA e.g. in HL Tau. If the gas phase were to extend to $r_{\rm out}=1000$ au, the total mass of the system is $\simeq 0.01$ M$_\odot$. We assume a vertically isothermal equation of state $P=c_{\rm s}^2\rho$ with $c_{\rm s}(r)=c_{\rm s,in}(R/R_{\rm in})^{-0.35}$ and an aspect ratio of the disc that is 0.05 at 1 au. We set an SPH viscosity parameter $\alpha_{\rm AV}=0.1$ giving an effective Shakura & Sunyaev (1973) viscosity $\alpha_{\rm SS}\approx 0.005$. We setup a planet located at 40 au and evolve the simulations over 40 planetary orbits. This is sufficient to study the physics of dust gap opening with our assumed grain size, though we caution that further evolution occurs over longer timescales. We vary the planet mass in the range $[0.05, 0.1, 0.5, 1]\,M_{\rm J}$ to evaluate the relative contributions of the tidal and drag torques.

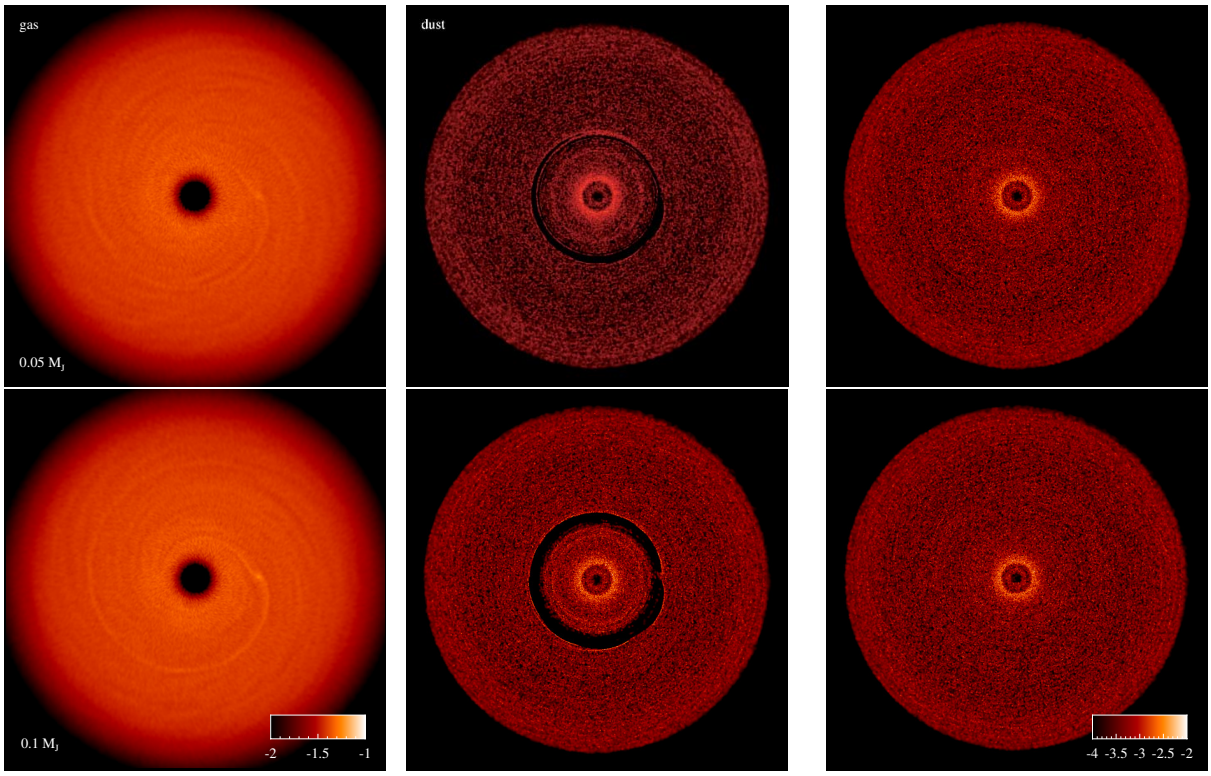

Figure 8.1: Gap opening via Mechanism I, where low mass planets carve a gap in the dust but not the gas. Plots show gas (left) and millimetre dust grain (centre) surface densities in a dusty disc hosting planets of mass 0.05 M_J (top row) and 0.1 M_J (bottom row). While the 0.05 M_J creates a depletion of dust at the planet location, the 0.1 M_J planet is able to carve a gap in the dust. Neglecting the gravity between the planet and the dust (right panels) shows that the gap is opened by the tidal torque. The drag torque acts to close the gap due to the radial migration of dust particles from the outer disc (top centre panel).

8.2 Results

Gap formation is a competition between torques. In a gas disc the competition is between the tidal torque from the planet trying to open a gap and the viscous torque trying to close it (see Sect. 7.3.1). Dust, by contrast, is pressureless and inviscid, and the competition is between the tidal torque and the aerodynamic drag torque (see Chapt. 2).

Dust efficiently settles to the midplane in our simulations, forming a stable dust layer with dust to gas scale height ratio of $\sim \sqrt{\alpha_\mathrm{SS}/\mathrm{St}} \sim 0.02$ (see Sect. 2.4.2), consistent with the Dubrulle et al. (1995) model and other SPH simulations of dusty discs (e.g. Laibe et al. 2008). Settling of grains is expected to slightly reinforce the contribution from the tidal torque by local geometric effects.

In the following sections we describe qualitatively the two mechanism of gap opening. In Sect. 8.3 we will refine our interpretation of the Mechanism I by analytically investigating the dynamics of dust particles.

8.2.1 Mechanism I — low mass planets

Fig. 8.1 demonstrates gap-opening when the planet is not massive enough to carve a gap in the gas disc. The gas shows only a weak one-armed spiral density wake supported by pressure, as predicted by linear density wave theory (Ogilvie & Lubow 2002).

In a disc undisturbed by a planet the radial pressure gradient and the gas viscous velocity are, in general, negative. Hence the drag torque on the dust phase is negative, pushing dust particles inwards. The planet exerts a tidal torque on the dust which is positive outside the planet orbit and negative inside, forcing grains away from the planet. This torque increases as

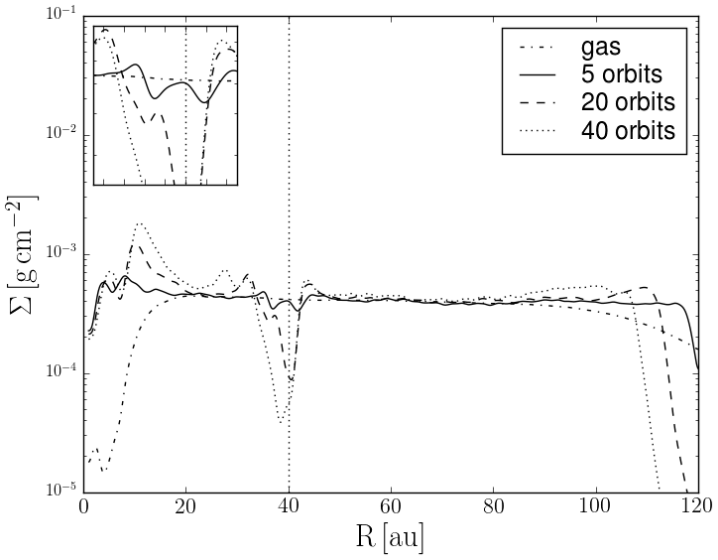

Figure 8.2: Evolution of the azimuthally averaged dust surface density profile for the disc hosting a 0.1 M_J planet corresponding to the bottom centre panel of Fig. 8.1. The dotted-dashed line provides the gas surface density scaled by a factor of 0.01 at the end of the simulation, for direct comparison with the dust phase. The dotted vertical line indicates the planet orbit. An asymmetric gap profile is seen in the dust, but not in the gas.

the planet mass increases. The balance between these two torques determines the evolution of the dust: A gap opening process resisted by drag from the outer disc.

At radii smaller than the orbital radius of the planet, both contributions add to make the grains drift inwards. At larger radii the tidal torque resists the drag torque by repelling the particles from the planet. For any planet mass, the tidal torque therefore decelerates dust particles outside the planetary orbit. A strong enough tidal torque in the outer disc prevents the replenishment of the dust population in the inner disc due to the action of the drag torque. Gap closing (opening) induced by drag outside (inside) the planetary orbit is maximal for grains with St = 1, decreases linearly with Stokes number for St $\ll$ 1 and is inversely proportional to Stokes number for St $\gg$ 1 (see Sect. 2.3.1, Takeuchi & Artymowicz 2001). However, small grains follow the viscous evolution of the gas. We therefore expect a less effective replenishment of larger particles from outside the planetary orbit due to the reduced drag torque, resulting in a more symmetric gap around the planet location.

Whether or not a gap can be opened in the dust depends on the planet mass. For a planet of very low mass (top row in Fig. 8.1; 0.05 M_J), we only see a partial depletion of dust at the planet location, while the regions of the disc inside the planetary orbit are constantly replenished by an inflow of particles from the outer disc. Replenishment occurs because the drag torque produces a constant inward drift of dust particles, while the action of the the tidal torque is localised at the planet position. If the time taken by dust particles to refill the gap is shorter than the orbital timescale, the gap (and inner disc) is refilled by the action of the drag torque. However, if the tidal torque from the planet on the inner disc is strong enough to overcome this refilling, an inner cavity will slowly open.

The lower panel in Fig. 8.1 shows that a sufficiently massive planet (here 0.1 M_J) clears a complete orbit in the dust. This occurs once the tidal torque is large enough to overcome the drag torque outside of the planetary orbit. Fig. 8.2 shows that the gap is not centred around the planet orbit, but is shifted towards the inner disc. Initially the gap is symmetric around the planet location, but becomes more and more asymmetric as the dust in the inner part of the disc drifts inwards, away from the planets orbit. As a result, no dust population is maintained in the corotation region. We confirm our physical picture above by noting that our 0.05 M_J simulation shows the same gap structure after a further 20 orbits, while the gap in the 0.1 M_J

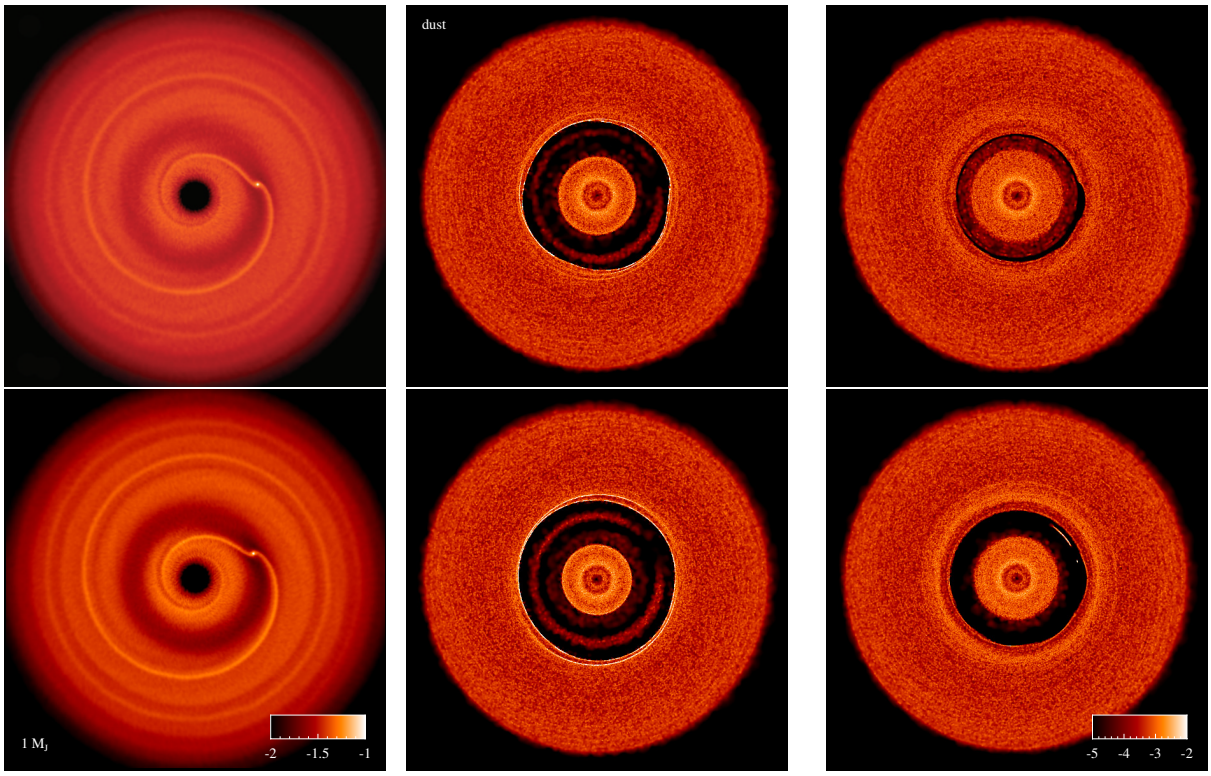

Figure 8.3: Gap opening via Mechanism II, where high mass planets carve a partial or total gap in the gas, and the dust is evacuated from the gap via drag and tidal torques. Plots are as in Fig. 8.1, but with planet masses 0.5 M_J (top) and 1 M_J (bottom). Although the tidal torque modifies the structure of the gap and stabilises the corotation region, the structure in the dust phase is dominated by the drag torque (comparing centre and right panels).

simulation grows more asymmetric with time.

The right panel of Fig. 8.1 demonstrates the role played by the tidal torque in the gap-opening mechanism, showing the dust dynamics computed with the gravitational force from the planet on the dust phase switched off. In this case, solid particles are unaffected by the surrounding gas and simply migrate towards the central star.

8.2.2 Mechanism II — high mass planets

Previous authors have identified two regimes for dust gap formation: high-mass planets opening gas and dust gaps (e.g. Crida et al. 2006) and low-mass planets creating mild radial pressure gradients in the gas but deeper gaps in the dust (e.g. Paardekooper & Mellema 2004). Both of these regimes involve the same, drag assisted, gap opening mechanism (our Mechanism II), but with different intensities. For sufficiently massive planets (0.5 M_J and 1 M_J, top and bottom rows, respectively), the tidal torque dominates the viscous torque at the planet location, expelling the gas. This results in a stable gap in the gas surface density. A pressure maximum is always created at the outer edge of the gap. At the inner edge, a pressure maximum forms only if the perturbation of the pressure profile induced by the tidal torque exceeds the background pressure gradient. This is more likely to occur for flat pressure profiles and a planet located close to the star (Crida & Morbidelli 2007; Fouchet et al. 2010). The drag torque on the dust phase is influenced by the formation of these pressure maxima at the gap edges. Dust accumulates at the location of the pressure maxima, forming a deeper dust gap than in the gas.

As expected, the width and depth of the gaps increase as the planet mass increases (top and bottom rows in Fig. 8.3). Compared to the low mass planet case, the additional pile-up induced by the presence of pressure maxima concentrates dust more efficiently at the outer edges of the gap. Fig. 8.4 shows that the gaps in the dust disc are W-shaped and asymmetric around the

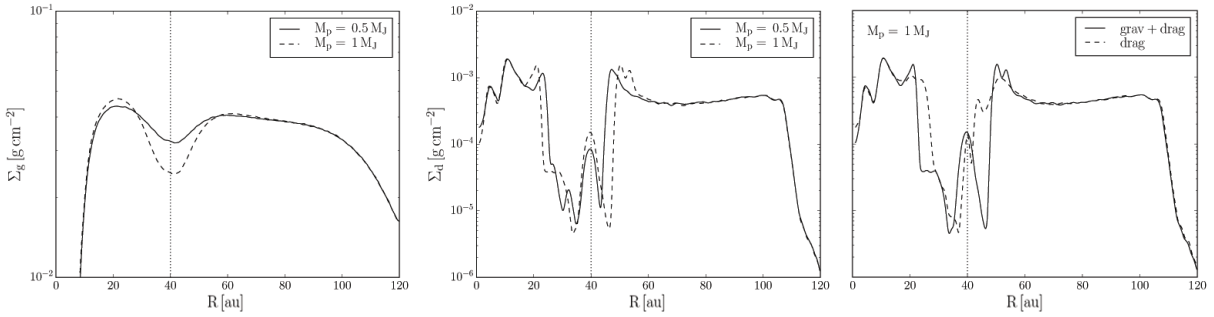

Figure 8.4: Azimuthally averaged surface density of the gas corresponding to the simulations of Fig 8.3. The dotted vertical line indicates the planet's location. Gaps are created in both the gas and the dust phases.

location of the planet. Dust grains located initially between the inner edge of the gap in the gas and the corotation radius are depleted towards the inner pressure maximum by the drag torque. In contrast to Mechanism I, particles with St = 1 concentrate fastest on the gap edges. It is therefore easier for marginally coupled particles to open gaps in the dust (Weidenschilling 1977a).

A large and stable population of dust grains is observed in the corotation region where the tidal torque provides a stabilising effect. The efficiency of the drag torque decreases locally as the gas surface density decreases (Laibe & Price 2012a) and the background pressure gradient is significantly reduced (see left panel of Fig. 8.4). Interestingly, excitation of particles eccentricities at the outer edge of the gap forms narrow ridges just outside the orbit of the planet (Ayliffe et al. 2012; Picogna & Kley 2015). This effect does not occur inside of the orbital radius of the planet because the drag torque is strong enough to efficiently damp any resonances that develop (Fouchet et al. 2007; Ayliffe et al. 2012), similar to what occurs in the whole disc in Mechanism I. For our 1 M_J planet, the outer edge of the dust gap is close to the 3:2 resonance ($R \sim 52$ au), inducing a double peaked outer edge in the dust density profile (see Fig. 8.4).

The right panel of Figs. 8.3 and 8.4 shows that for high mass planets, the formation of a gap in the dust can be recovered simply by considering drag effects and neglecting the action of the gravitational potential of the planet. However, the detailed structure of the gap is still different when the tidal torque is included: the gap is wider, deeper, with a corotation region, sharper edges and more asymmetries due to external resonances.

8.3 Interpretation of Mechanism I: Dynamics in a viscous disc with a planet

In this section we analyze the Mechanism of gap opening via Mechanism I (see Sect. 8.2.1) by investigating the motion of dust particles and the ability of the planet to carve a gap in the dust structure without affecting the local gas structure.

8.3.1 Disc-planet tidal interaction

As described in Chapter 7, a planet transfers angular momentum to the surrounding disc by driving material into mean-motion resonances at specific locations referred to as Lindblad resonances. Goldreich & Tremaine (1979, 1980) derived an analytic expression for the tidal torque per unit mass Λ (the *excitation* torque density) exerted by a fixed planet on an elementary ring of a pressureless disc given in Eq. 7.27 and repeated here in a general form:

$$\Lambda(r) = \mathrm{sgn}(R - R_\mathrm{p}) f \frac{(\mathcal{G}M_\mathrm{p})^2}{\Omega_\mathrm{p}^2} \frac{1}{\Delta^4}, \tag{8.1}$$

where M_p is the mass of the planet, $\mathcal{G}$ is the gravitational constant, R_p is the planet location, $\Omega_\mathrm{p} = (\mathcal{G}M_\star/R_\mathrm{p}^3)^{1/2}$ is the Keplerian angular velocity at r_p where $M_\star$ is the mass of the central

star, $\Delta \equiv |R - R_\mathrm{p}|$, and f is a constant of order unity. While the key features of the tidal torque equation, such as the Δ^4 dependence, have been widely accepted (Lin & Papaloizou 1986; Bryden et al. 1999; Bate et al. 2003; Varnière et al. 2004; D'Angelo & Lubow 2008), the proportionality coefficient f is mostly inferred by analyzing the shape of the gap carved by the planet (Armitage et al. 2002; Tazzari et al. 2015). In particular, Armitage et al. (2002) found that the gap extension carved by a planet is consistent with $f \sim 10^{-2}$. The factor in front of the right-hand side of Eq. 8.1 shows that a planet tends to push material outside of its orbit.

8.3.1.1 Gas disc

As described in Chapter 7, in a disc with non-zero pressure, the exchange of torque between the planet and the disc is a two-step mechanism. The planet first excites density waves in the gas. The amount of initial torque density stored in these density waves is the one given by Eq. 8.1. Then, a fraction of this torque (the *deposition* torque) is transferred from the waves to the disc by a damping processes such as viscosity or shocks (Goldreich & Nicholson 1989; Goodman & Rafikov 2001; Rafikov 2002), whereas the remaining part (the *pressure* torque, Crida et al. 2006) is carried away by the waves until being deposited in a more distant region of the disc. This portion of the tidal torque depends on the relative radial gradient of the azimuthally averaged gas surface density. Therefore, in the specific case of a planet not able to carve a gap in the gas density structure, we can safely ignore the details of wave damping and assume that the deposition torque is equal to the excitation torque. This pressure-supported transport modifies also the effective location of the Lindblad resonance. An important consequence is that the deposition torque is essentially zero in the vicinity of the planet, since high order Lindblad resonances are shifted away from the planet by a typical length that is the scale height of the disc H_g (see Sect. 7.1.2.1, Goldreich & Tremaine 1980; Artymowicz 1993). To include the torque cut-off due to the pressure and the corotation region, the deposition toque is often calculated from the expression given by Eq. 8.1, but with the Δ_g given by $\Delta_\mathrm{g} = \max(|R - R_\mathrm{p}|, H_\mathrm{g}, R_\mathrm{H})$. It is worth pointing out that a self-consistent calculation of the torque density should include the non-linear propagation of the density waves induced by the planet and the details of the wave damping processes.

8.3.1.2 Dust disc

From the previous discussion, we expect the tidal torque to be larger in the dust than in the gas, and this for three reasons:

1. Dust is a pressureless fluid, where density waves cannot propagate. The deposition torque equals therefore the excitation torque.

2. For the same reason, there is no torque cut-off in the dust, and the effective Δ_d is equal to $\Delta_\mathrm{d} = \max(|R - R_\mathrm{p}|, R_\mathrm{H})$. Close to a low mass planet (i.e. when $R_\mathrm{H} < H_\mathrm{g}$), angular momentum can be deposited in the dust, but not in the gas. Hence, the ratio between the maximum torque density in the gas and in the dust is of order $\sim (R_\mathrm{H}/H_\mathrm{g})^4 = (R_\mathrm{p}/H_\mathrm{g})^4 (M_\mathrm{p}/3M_\star)^{4/3}$.

3. For large grains, the thickness of the dust layer H_d is smaller than H_g, as a result from the competition between settling and turbulent stirring (see Sect. 2.4.2, Dubrulle et al. 1995). For dust layers with $H_\mathrm{d} \geq R_\mathrm{H}$ this results in an enhancement of the tidal torque due to local geometrical effects.

8.3.2 Dust-gas aerodynamical coupling

8.3.2.1 Equations of motion for low-mass planet

The motion of dust particles in protoplanetary discs is affected by the aerodynamical interaction with the gas and *vice-versa* (see Chapter 2). The differential motion between the two phases

gives rise to a drag force that tends to damp this velocity difference. In discs hosting planets, an additional velocity difference can be produced since the interactions of the planet with the gas differs from the ones with the dust (Sect. 8.3.1).

We assume a thin, non-magnetic, non-self-graviting, viscous and vertically isothermal protoplanetary disc hosting a non-migrating planet, and treat the dust phase as a continuous pressureless and viscousless fluid. The equations of motion for the gas and the dust are given by Eqs. 2.22–2.23 with an additional term related to the tidal torque

$$\frac{\partial \mathbf{v}_\mathrm{g}}{\partial t} + (\mathbf{v}_\mathrm{g}\cdot\nabla)\,\mathbf{v}_\mathrm{g} = \frac{K^V}{\rho_\mathrm{g}}(\mathbf{v}_\mathrm{d}-\mathbf{v}_\mathrm{g}) - \nabla(\Phi+\Phi_\mathrm{p}) - \frac{1}{\rho_\mathrm{g}}(\nabla P - \nabla\cdot\sigma), \tag{8.2}$$

$$\frac{\partial \mathbf{v}_\mathrm{d}}{\partial t} + (\mathbf{v}_\mathrm{d}\cdot\nabla)\,\mathbf{v}_\mathrm{d} = -\frac{K^V}{\rho_\mathrm{d}}(\mathbf{v}_\mathrm{d}-\mathbf{v}_\mathrm{g}) - \nabla(\Phi+\Phi_\mathrm{p}), \tag{8.3}$$

where the indices g and d refer to the gas and the dust phases, v and ρ denote the velocities and the densities, Φ and Φ_p denote the gravitational potentials of the star and the planet. P and σ denote the pressure and the viscous tensor of the gas. K^V denotes the drag coefficient, whose expression is given in Eq. 2.25 and depends on the the stopping time of the mixture or, analogously, on the Stokes number expressed in Eq. 2.37.

To reduce Eqs. 8.2 – 8.3 to a system of equations that describes a steady steady solution for the gas and the dust we adopting the same approach described in Sect. 2.3.1. The equations of motion for the perturbed velocities in cylindrical coordinates are given by

$$\frac{K^V}{\rho_\mathrm{g}}(v_{\mathrm{d},R}-v_{\mathrm{g},R}) - \frac{1}{\rho_\mathrm{g}}\frac{\partial P}{\partial R} + 2\Omega_\mathrm{k} v_{\mathrm{g},\theta} = 0, \tag{8.4}$$

$$\frac{K^V}{\rho_\mathrm{g}}(v_{\mathrm{d},\theta}-v_{\mathrm{g},\theta}) - \frac{\Omega_\mathrm{k}}{2}v_{\mathrm{g},R} + \frac{1}{\rho_\mathrm{g}}\nabla\cdot\sigma|_\theta + \frac{\Lambda_\mathrm{g}}{R} = 0, \tag{8.5}$$

$$\frac{K^V}{\rho_\mathrm{d}}(v_{\mathrm{d},R}-v_{\mathrm{g},R}) - 2\Omega_\mathrm{k} v_{\mathrm{d},\theta} = 0, \tag{8.6}$$

$$\frac{K^V}{\rho_\mathrm{d}}(v_{\mathrm{d},\theta}-v_{\mathrm{g},\theta}) + \frac{\Omega_\mathrm{k}}{2}v_{\mathrm{d},R} - \frac{\Lambda_\mathrm{d}}{R} = 0. \tag{8.7}$$

8.3.2.2 Steady-state solutions

Solving the linear system Eqs. 8.4 – 8.7, we find

$$\begin{aligned} v_{\mathrm{g},R} =& -\frac{1}{1+\epsilon}\left\{\frac{\epsilon\,\Delta v}{\mathrm{St}+\mathrm{St}^{-1}} - \left(1+\epsilon\frac{\mathrm{St}^2}{1+\mathrm{St}^2}\right)v_\mathrm{visc}\right\} \\ &+ \frac{2\Lambda_\mathrm{d}}{v_\mathrm{k}}\frac{\epsilon}{(1+\epsilon)(1+\mathrm{St}^2)} + \frac{2\Lambda_\mathrm{g}}{v_\mathrm{k}}\frac{1+\mathrm{St}^2(1+\epsilon)}{(1+\epsilon)(1+\mathrm{St}^2)}, \end{aligned} \tag{8.8}$$

$$\begin{aligned} v_{\mathrm{d},R} =& \frac{1}{1+\epsilon}\left\{\frac{\Delta v}{\mathrm{St}+\mathrm{St}^{-1}} + \frac{v_\mathrm{visc}}{1+\mathrm{St}^2}\right\} \\ &+ \frac{2\Lambda_\mathrm{d}}{v_\mathrm{k}}\frac{\epsilon\left(1+\mathrm{St}^2\right)+\mathrm{St}^2}{(1+\epsilon)\left(1+\mathrm{St}^2\right)} + \frac{2\Lambda_\mathrm{g}}{v_\mathrm{k}(1+\epsilon)(1+\mathrm{St}^2)}, \end{aligned} \tag{8.9}$$

$$\begin{aligned} v_{\mathrm{g},\theta} =& \frac{1}{2(1+\epsilon)}\left\{\left(1+\epsilon\frac{\mathrm{St}^2}{1+\mathrm{St}^2}\right)\Delta v + \frac{\epsilon}{\mathrm{St}+\mathrm{St}^{-1}}v_\mathrm{visc}\right\} \\ &+ \frac{\epsilon(\Lambda_\mathrm{g}-\Lambda_\mathrm{d})}{v_\mathrm{k}(1+\epsilon)(\mathrm{St}+\mathrm{St}^{-1})}, \end{aligned} \tag{8.10}$$

$$\begin{aligned} v_{\mathrm{d},\theta} =& \frac{1}{2(1+\epsilon)}\left\{\frac{\Delta v}{1+\mathrm{St}^2} - \frac{v_\mathrm{visc}}{\mathrm{St}+\mathrm{St}^{-1}}\right\} \\ &- \frac{\Lambda_\mathrm{g}-\Lambda_\mathrm{d}}{v_\mathrm{k}(1+\epsilon)(\mathrm{St}+\mathrm{St}^{-1})}, \end{aligned} \tag{8.11}$$

where $\epsilon = \rho_\mathrm{d}/\rho_\mathrm{g}$ is the dust-to-gas density ratio, Δv and v_visc are expressed in Eqs. 2.42 and 2.43, respectively. We point out that we use the value of the viscous gas velocity obtained in a disc without an embedded protoplanet. Our treatment in justified by the fact that, since we are studying the Mecahanism I for dust gap opening process for low mass planets not able to carve gaps in the gas, surface density gradients develop over large scales. From Eqs. 8.10 – 8.11, the differential azimuthal velocity between the gas and dust is

$$v_{\mathrm{d},\theta} - v_{\mathrm{g},\theta} = -\frac{\Delta v}{2(1+\mathrm{St}^{-2})} - \frac{v_\mathrm{visc}}{2(\mathrm{St}+\mathrm{St}^{-1})} - \frac{\Lambda_\mathrm{g}-\Lambda_\mathrm{d}}{v_\mathrm{k}(\mathrm{St}+\mathrm{St}^{-1})},$$

which is independent on the dust-to-gas ratio. The specific drag torque exerted by the gas phase on an elementary dust ring si therefore

$$\Lambda_{\mathrm{g}\to\mathrm{d}} = -R\frac{K^V}{\rho_\mathrm{d}}\left(v_{\mathrm{d},\theta} - v_{\mathrm{g},\theta}\right) = \frac{v_\mathrm{k}}{2(1+\epsilon)}\left\{\frac{\Delta v}{\mathrm{St}+\mathrm{St}^{-1}} + \frac{v_\mathrm{visc}}{1+\mathrm{St}^2}\right\} - \frac{\Lambda_\mathrm{g}-\Lambda_\mathrm{d}}{(1+\epsilon)(1+\mathrm{St}^2)}, \quad (8.12)$$

while the back reaction drag torque from the dust to the gas is

$$\Lambda_{\mathrm{d}\to\mathrm{g}} = R\frac{K^V}{\rho_\mathrm{g}}\left(v_{\mathrm{d},\theta} - v_{\mathrm{g},\theta}\right) = -\epsilon\Lambda_{\mathrm{g}\to\mathrm{d}}. \quad (8.13)$$

8.3.2.3 Physical interpretation

In the limit of the dust being perfectly decoupled from the dust ($\mathrm{St} = +\infty$), grains orbit the star with Keplerian velocity ($v_{\mathrm{d},\theta} = 0$) and are pushed outside of the planet orbit by the tidal torque at the constant velocity $v_{\mathrm{d},R} = 2\Lambda_\mathrm{d}/v_\mathrm{k}$. The gas orbits the star at the sub-Keplerian velocity $\Delta v/2 < 0$, the gas radial velocity being $v_\mathrm{visc} + 2\Lambda_\mathrm{g}/v_\mathrm{k}$. When drag couples the two phases, the dust motion is dominated by the gas when $\epsilon \ll 1$ and $1/(1+\epsilon) \sim 1$. When $\epsilon \gg 1$ the gas motion is dominated by the dust and $\epsilon/(1+\epsilon) \sim 1$. Both the gas and the dust strongly feel the other phase when the dust-to-gas ratio is of order unity (see Sect. 2.3.1).

As it is known in the case without a planet ($\Lambda_\mathrm{g} = \Lambda_\mathrm{d} = 0$, differential dynamics due to gas pressure and viscosity are communicated from one phase to the other via drag. See Sect. 2.3.1 for further details.

In the general case, the differential motion between the gas and the dust induced by the different tides only affect the motion of smaller grains, i.e. $\mathrm{St} \lesssim 1$. This concerns particularly the region $R_\mathrm{H} < |r - R_\mathrm{p}| < H_\mathrm{g}$, where $\Lambda_\mathrm{d} \neq \Lambda_\mathrm{g} = 0$. In this region, for a gas-dominated dynamics with $\epsilon \ll 1$, only large dust grains, i.e. $\mathrm{St} \gtrsim 1$, are pushed away from the planet orbit, since they experience a lower drag related to the different tidal torque between the two phases (see the third term in the right hand side of Eq. 8.12). In other words, for smaller grains, the motion induced by the tidal torque in this region is damped by the drag torque that tends to reduce the velocity difference with the unperturbed gas flow. As a result, small grains are forced to stick to the fixed gas. On the other hand, it can be noticed that the gas is not affected by the tides in this region if $\epsilon \ll 1$ (see the third and fourth terms in the right hand side of Eq. 8.8). Sufficiently far away from the planet, where $|r - R_\mathrm{p}| > H_\mathrm{g}$, $\Lambda_\mathrm{g} = \Lambda_\mathrm{d}$, there is no differential tidal torque, and the tidal barycentric velocity is spread over the two phases proportionally to the respective density of each phase.

From this analysis, it is reasonably expected that planets which are not able to carve gaps in the gas might be able to open gaps in the dust structure only for larger grains, i.e. $\mathrm{St} \gtrsim 1$.

8.4 Criterion for gap opening in dusty discs

8.4.1 Orders of magnitude

We compare the orders of magnitude of the different velocity terms in Eq. 8.9, related to the radial pressure gradient, the viscous and the tidal contribution. Until specified, we shall not

restrict our analysis to the case of an unperturbed gas density profile. We obtain

$$\left|-\frac{1}{\rho_{\rm g}\Omega_{\rm k}}\frac{\partial P}{\partial r}\right| \sim \left(\frac{R}{l}\right)\left(\frac{H_{\rm g}}{R}\right)^2 v_{\rm k}, \tag{8.14}$$

$$\left|\frac{1}{\rho_{\rm g}\Omega_{\rm k}}\nabla\cdot\sigma|_\theta\right| \sim \left(\frac{R}{l}\right)\alpha\left(\frac{H_{\rm g}}{R}\right)^2 v_{\rm k}. \tag{8.15}$$

where l denotes the typical length over which the gas surface density varies and where we have used the seminal Prandtl-like turbulent viscosity $\nu = \alpha c_{\rm s} H_{\rm g}$ (Shakura & Sunyaev 1973). The prefactor in Eq. 8.15 originates from the second derivative of the Keplerian deviation (Eq. 2.27). For an unperturbed gas profile, $l = r$ and the pre-factors in Eqs. 8.14 – 8.15 equal unity. The maximal tidal contributions, obtained at the cut-off locations, are of order

$$\left|\frac{\Lambda}{v_{\rm k}}\right|_{\rm max} \sim \frac{R_{\rm H}^6}{R^2\Delta^4} v_{\rm k}. \tag{8.16}$$

This gives for the gas and the dust respectively

$$\left|\frac{\Lambda_{\rm g}}{v_{\rm k}}\right|_{\rm max} \sim \left(\frac{R_{\rm H}}{H_{\rm g}}\right)^6\left(\frac{H_{\rm g}}{R}\right)^2 v_{\rm k}, \tag{8.17}$$

$$\left|\frac{\Lambda_{\rm d}}{v_{\rm k}}\right|_{\rm max} \sim \left(\frac{R_{\rm H}}{H_{\rm g}}\right)^2\left(\frac{H_{\rm g}}{R}\right)^2 v_{\rm k}. \tag{8.18}$$

For the sake of clarity, we now assume $\epsilon \ll 1$ and limit our analysis to large grains which are most affected by the tidal torque without being decelerated by the drag torque arising due to the differential tidal torque between the two phases. For larger grains, i.e. ${\rm St} \gtrsim 1$, the terms in Eq. 8.9 proportional to $1/(1+{\rm St}^2)$ are negligible. The terms in Eq. 8.9 are given by

$$v_{{\rm d},R}|_{\Delta v} \sim \frac{\rm St}{1+{\rm St}^2}\left(\frac{H_{\rm g}}{R}\right)^2 v_{\rm k}, \tag{8.19}$$

$$v_{{\rm d},R}|_{\Lambda_{\rm d}} \sim \frac{{\rm St}^2}{1+{\rm St}^2}\left(\frac{R_{\rm H}}{H_{\rm g}}\right)^2\left(\frac{H_{\rm g}}{R}\right)^2 v_{\rm k}. \tag{8.20}$$

The term ${\rm St}^2/\left(1+{\rm St}^2\right)$ in Eq. 8.20 expresses that only large grains are entrained by the tides since small grains stick to the gas. For these grains the drag torque is dominated by the contributions from the pressure gradient and the tides. The scaling for the minimum mass for a planet to open a gap in the dust is obtained by equating Eqs. 8.19 and 8.20 with the definition given in Eq. 7.12:

$$\frac{M_{\rm p}}{M_\star} \sim {\rm St}^{-3/2}\left(\frac{H_{\rm g}}{R_{\rm p}}\right)^3. \tag{8.21}$$

Eq. 8.21 shows that the critical mass required to open a gap in the dust is lower for large Stokes numbers and large aspect ratios, when the drag torque becomes smaller. More interestingly, the exponent $-3/2$ implies a sharp transition between small and large grains at ${\rm St} \simeq 1$. Fixing $H_{\rm g}/R_{\rm p}$ and decreasing St in Eq. 8.21 shows that $M_{\rm p}$ increases efficiently, up to reach the critical mass required to open a gap in the gas as well, i.e. $M_{\rm p}/M_\star \sim (H_{\rm g}/R_{\rm p})^3$, corresponding to the thermal criterion introduced in Sect. 7.3.1 (Eq. 7.45, Lin & Papaloizou 1993). This suggests, as expected, that the condition ${\rm St} \gtrsim 1$ should be fulfilled for a gap to be carved in the dust only (see Sect. 8.4.4).

8.4.2 Dust velocities

8.4.2.1 Dust gap width

The tidal torque acts to clean its orbit, whereas drag makes the grains drift inwards. In particular, the flux of solids coming from the outside of the planet orbit tends to refill the dust depletion

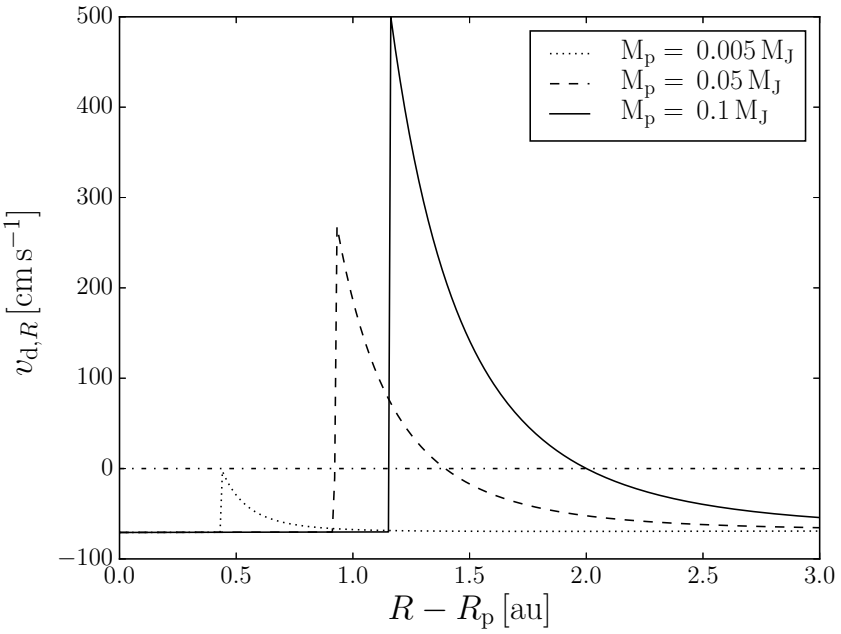

Figure 8.5: Total radial dust velocity of millimetre grains outside the orbit of the planet for different planet masses: 0.005, 0.05 and 0.1 M_{J} adopting the model described in Sect. 8.1.2 and assuming the proportionality constant in front of the torque density prescription expressed in Eq. 8.1 equal to the nominal value of 0.4 introduced in Goldreich & Tremaine (1979). The velocity peaks at $R - R_{\mathrm{p}} = R_{\mathrm{H}}$ and decreases with increasing distance from the planet. For planets with masses $\lesssim 0.005 M_{\mathrm{J}}$, the tidal torque is not strong enough to halt the radial inflow induced by the drag torque ($v_{\mathrm{d},R}(r) < 0$).

locally induced by the tides. In the presence of a low mass planet, i.e. not able to affect the local pressure structure, the drag torque can be derived using the unperturbed pressure profile of the gas (Eq. 8.12). The azimuthally averaged radial dust velocity of Eq. 8.9 can be written

$$v_{\mathrm{d},R} = \frac{\zeta \mathrm{St} - (6+3\zeta)\,\alpha}{(1+\epsilon)(1+\mathrm{St}^2)} \frac{c_{\mathrm{s}}^2}{v_{\mathrm{k}}} + \frac{2\Lambda_{\mathrm{d}}}{v_{\mathrm{k}}} + \frac{2(\Lambda_{\mathrm{g}} - \Lambda_{\mathrm{d}})}{v_{\mathrm{k}}(1+\mathrm{St}^2)(1+\epsilon)}. \tag{8.22}$$

where ζ is expressed in Eq. 1.16. As an example, Fig. 8.5 shows the radial velocity of millimetre-in-size grains near the orbit of the planet for the disc model described in Sect. 8.1.2, assuming f to be equal to the nominal value of 0.4 introduced in Goldreich & Tremaine (1979). In the presence of a planet of very low-mass, the tidal torque is not strong enough to halt the inward radial flow triggered by the drag torque (e.g. the case $\lesssim 0.005\,M_{\mathrm{J}}$, for which $v_{\mathrm{d},R} < 0$). Larger planet masses empower stronger tidal toques in the disc, and the balance between the tidal and the drag torques leads to an outward drift (Fig. 8.5 shows regions where $v_{\mathrm{d},R} > 0$ for $M_{\mathrm{p}} = 0.05\,M_{\mathrm{J}}$ and $M_{\mathrm{p}} = 0.1\,M_{\mathrm{J}}$). The location of the outer edge of the gap R_{gap} can be estimated simply, by evaluating the distance to the planet where $v_{\mathrm{d},R}(R_{\mathrm{gap}}) = 0$, i.e. where the drift induced by the tides balances steadily the drift induced by the drag. We use $\Delta_{\mathrm{gap,d}}$ to denote the width of the dust gap outside the planetary orbit $R_{\mathrm{gap}} - R_{\mathrm{p}}$, and obtain

$$\frac{\Delta_{\mathrm{gap,d}}}{R_{\mathrm{p}}} \simeq (2f)^{1/4}\, z^{-1/4}\, \mathrm{St}^{1/4} \left(\frac{\mathrm{H_g}}{\mathrm{R_p}}\right)^{-1/2} \left(\frac{\mathrm{M_p}}{\mathrm{M_\star}}\right)^{1/2}, \tag{8.23}$$

where

$$z(\mathrm{St}, \epsilon, \alpha) = \begin{cases} \dfrac{-\zeta + (6+3\zeta)\,\alpha/\mathrm{St}}{(1+\epsilon) + \epsilon/\mathrm{St}^2}, & R_{\mathrm{H}} \leq \Delta_{\mathrm{gap,d}} < H_{\mathrm{g}}, \\ \dfrac{-\zeta + (6+3\zeta)\,\alpha/\mathrm{St}}{(1+\epsilon) + (1+\epsilon)/\mathrm{St}^2}, & \Delta_{\mathrm{gap,d}} \geq H_{\mathrm{g}}. \end{cases} \tag{8.24}$$

Note that the difference between the two expressions in Eq. 8.24 has no sensible effect on the value of z in practice for larger grains. This suggests, as expected, that the different tidal torque

experienced by the two phases does not cover the key role in the gap opening process. For $\mathrm{St} \gtrsim \alpha$, the second term of the numerator vanishes and for $\mathrm{St} \gtrsim 1+\epsilon$, Eq. 8.24 reduces to

$$z \simeq \frac{-\zeta}{1+\epsilon}. \tag{8.25}$$

Eq. 8.23 is consistent with the fact that infinitely small grains at low dust densities follow the gas, in which no gap forms ($\Delta_{\mathrm{gap,d}} = 0$ for $\mathrm{St} = 0$ and $\epsilon = 0$). The location of the outer edge of the dust gap is an increasing function of the Stokes number, since it corresponds to a weaker drag torque. For large grains, $\Delta_{\mathrm{gap,d}} \propto \mathrm{St}^{1/4}$, an expression only weakly sensitive to exact value of St. Moreover, this analysis shows that, since large grains experience a lower drag related to the differential tidal torque between the two phases, the criterion is mostly based on the balance between the tidal torque Λ_{d} and the drag torque related to pressure forces. The differential tidal torque between the two phases mostly influence the dynamics of small grains, forcing them to follow the unperturbed gas flow.

Eq. 8.23 also shows that, since $\mathrm{St} \propto 1/(1+\epsilon)$, the gap width does not depend on the dust-to-gas ratio when large grains are considered, i.e. when Eq. 8.25 is valid.

For large Stokes number, i.e. $\mathrm{St} > \alpha$, the increase of the dust-to-gas ratio at the midplane affect the motion of dust and gas, as described in Sect. 8.3.2. The eventual modification of the gas surface density profile is not included in our model (see discussion in Sect. 8.6.2).

8.4.2.2 Necessary condition for dust gap opening

Eq. 8.23 provides a simple way to estimate the minimum planet mass M_{p} able to halt the inward radial drift induced by the drag, by noting that the minimum radius of the outer edge of the dust gap is the Hill radius. M_{p} is the planet mass for which the radial dust velocity is zero at $R = R_{\mathrm{p}} + R_{\mathrm{H}}$. With the reasonable assumption that the temperature and surface density of the gas are uniform over the domain $|R - R_{\mathrm{p}}| \sim R_{\mathrm{H}}$, we obtain

$$\frac{M_{\mathrm{p}}}{M_\star} \geq \xi \left(\frac{z}{\mathrm{St}}\right)^{3/2} \left(\frac{H_{\mathrm{g}}}{R_{\mathrm{p}}}\right)^3, \tag{8.26}$$

with z given by the first expression in Eq. 8.24 and with

$$\xi = \frac{1}{9}(2f)^{-3/2}. \tag{8.27}$$

Eq. 8.26 provides the minimum mass for a planet to produce a density depletion in the dust outside the planet orbit. Adopting the same model described in Sect. 8.1.2 and assuming $f = 0.4$ (see Eq. 7.27, Goldreich & Tremaine 1979), the minimum mass mass has a value of $\sim 0.005\, M_{\mathrm{J}}$ for millimetre-in-size grains, as expected from the previous analysis about the radial velocity (see dotted line in Fig. 8.5). Eq. 8.26 is in agreement with the orders-of-magnitude estimate (see Eq. 8.21).

8.4.3 Evacuation of angular momentum

8.4.3.1 Opening and closing timescales

Eq. 8.26 gives the minimum mass of a planet able to halt the radial inward drift of dust particles. However, for a gap to form, dust must not refill dust depletions as it explores different azimuths. Since refilling is a *non-axisymmetric* process, its effects is not picked up by the previous analysis based on averaged axisymmetric torques. An alternative approach based on timescales estimates is therefore developed hereafter to take dust refilling into account (see Sect. 7.3.1).

The gap opening timescale is the time required to evacuate all the dust contained between R_{p} and $R_{\mathrm{p}} + R_{\mathrm{H}}$. As previously mentioned, a gap of half width R_{H} is roughly the smallest gap that can be opened in the dust, since the Lindblad resonances are most effective at this distance

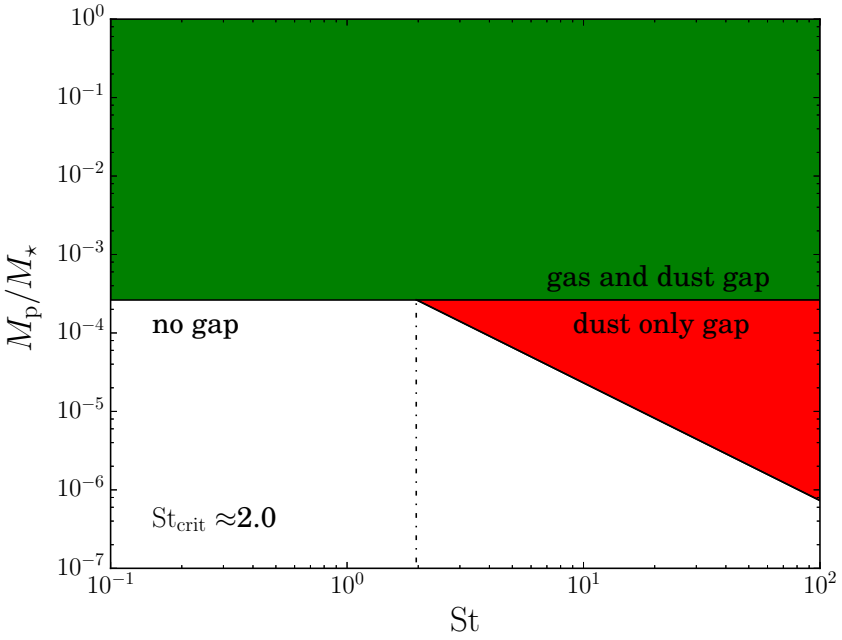

Figure 8.6: Sufficient condition for dust gap opening adopting the model described in Sect. 8.1.2for different Stokes numbers. The shaded areas indicates the range of planet masses and Stokes number for which a gap is carved in both phases (green) and only in the dust (red). The range of masses able to carve a gap in the dust only increases with the Stokes number.

from the planet. Focussing our analysis on the dust in the outer planetary orbit, the angular momentum that must be removed to open the gap between R_p and $R_\mathrm{p} + R_\mathrm{H}$ is

$$\Delta J = 2\pi R_\mathrm{p} R_\mathrm{H} \Sigma_\mathrm{d} \left.\frac{\mathrm{dj}}{\mathrm{dR}}\right|_{R_\mathrm{p}} R_\mathrm{H} = \pi R_\mathrm{p} R_\mathrm{H}^2 \Sigma_\mathrm{d} v_\mathrm{k}, \tag{8.28}$$

where j denotes the specific angular momentum. The typical time to evacuate all the dust in this region is

$$t_\mathrm{open} = \frac{\Delta J}{|\mathrm{dJ/dt}|}, \tag{8.29}$$

where $|\mathrm{dJ/dt}|$ is the one-sided total torque on the planet due to its interaction with dust outside the orbit. This total torque is the integral of the torque density given by Eq. 8.1 over the entire outer disc, i.e.

$$\frac{\mathrm{dJ}}{\mathrm{dt}} = \int_{R_\mathrm{H}}^{\infty} 2\pi r_\mathrm{p} \Sigma_\mathrm{d} \Lambda_\mathrm{d} (R - R_\mathrm{p})\, \mathrm{d(R - R_p)}. \tag{8.30}$$

The gap opening timescale is therefore

$$t_\mathrm{open} = (3\xi)^{2/3} \left(\frac{M_\mathrm{p}}{M_\star}\right)^{-1/3} \Omega_\mathrm{k}^{-1}. \tag{8.31}$$

Setting $\Lambda_\mathrm{g} = \Lambda_\mathrm{d} = 0$ in Eq. 8.22 and assuming $v_{\mathrm{d},R}$ constant over the domain $|R - R_\mathrm{p}| \sim R_\mathrm{H}$, the closing time $t_\mathrm{close} = R_\mathrm{H}/v_{\mathrm{d},R}$ is

$$t_\mathrm{close} = \frac{(1+\epsilon)(1+\mathrm{St}^2)}{-\zeta \mathrm{St} + (6+3\zeta)\,\alpha\, c_\mathrm{s}^2} \frac{v_\mathrm{k}}{} R_\mathrm{H}. \tag{8.32}$$

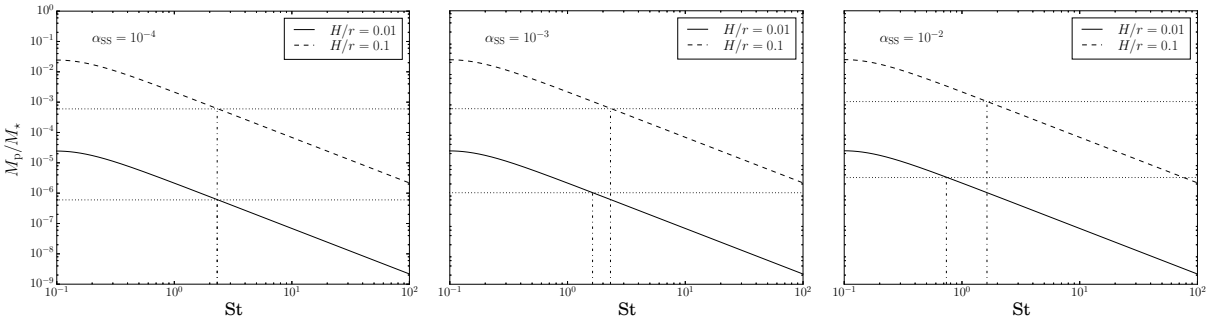

Figure 8.7: Sufficient condition of gap opening assuming different values of the viscosity α_{SS} and aspect ratio at the planet location. In the viscous force dominates the gap closing mechanism, the value of the critical Stokes number decrease with increasing viscosity and decreasing aspect ratio (see Eq. 8.36).

8.4.3.2 Sufficient condition for dust gap opening

The critical mass ratio $M_p/M_\star$ above which a planet sustains its gap in a dust disc is obtained by equating the opening and the closing timescale, which gives

$$\frac{M_p}{M_\star} \geq 3^{3/2}\xi \left(\frac{z}{St}\right)^{3/2} \left(\frac{H_g}{R_p}\right)^3, \tag{8.33}$$

with z given by the second expression in Eq. 8.24. As already mentioned, the two expressions in Eq. 8.24 have the same value for our aims, i.e. for $St \gtrsim \alpha$ and for $St \gtrsim \epsilon$. Thus, Eq. 8.33 equals to $3^{3/2} \approx 5.2$ times the critical mass derived in Eq. 8.26. Assuming $f = 0.4$ (Goldreich & Tremaine 1979), it gives a typical mass of $\sim 0.026\, M_J$ to open a dust disc for the disc model described in Sect. 8.1.2. The critical Stokes number over which the planet carve gaps only in the dust is close to unity. For large grains, Eq. 8.33 reduces to

$$\frac{M_p}{M_\star} \geq \xi \left(\frac{-3\zeta}{1+\epsilon}\right)^{3/2} St^{-3/2} \left(\frac{H_g}{R_p}\right)^3, \tag{8.34}$$

This criterion provides a good estimator for the minimum mass to open a gap in dusty disc, although we will show hereafter that the constant in front can be fine-tuned against numerical simulations.

The criteria derived above remain valid as long as the gas density profile remains unperturbed. In Sect. 7.3.1 we provide an overview of all the gap opening criteria developed in the past. In this work we consider that the minimum mass able to carve a gap in gas disc is given by the maximum of all the minimum masses predicted by the previously mentioned criteria (see Sect. 7.3.1)

$$M_{p,lim} = \max(0.2\, M_{p,th}, M_{p,visc,H_g}). \tag{8.35}$$

where the first term is $\propto M_{p,th}$ is the thermal criterion (Eq. 7.45) with a corrective term to account for recent results of gas simulations of discs hosting planets which have shown that planets with mass $\gtrsim 0.2\, M_{p,th}$ are able to create a pressure maximum at the gap outer edge (Lambrechts et al. 2014; Rosotti et al. 2016). The second term is related to the gap filling effect of the viscosity and is given by Eq. 7.52 multiplied by a factor of $(27/8)^{1/2}$, as adopted in Armitage (2010).

8.4.4 Critical Stokes number

Fig. 8.6 shows the minimum planet mass to open a dust gap given by Eq. 8.33 as a function of the Stokes number adopting the model described in Sect. 8.1.2 and assuming $f = 0.4$ (Goldreich & Tremaine 1979). The red shaded area indicates the range of planet masses and Stokes numbers for which a gap is carved only in the dust. The green area shows the domain for which the

planet carves a gap in the gas as well, i.e. for $M_\mathrm{p} \geq M_\mathrm{p,lim}$ (Eq. 8.35). As expected, the range of masses increases with increasing Stokes numbers, due to the reduced replenishment from the outer disc induced by the drag torque. For planets with masses inside the green area, the drag assist the gap opening in the dust, leading to a well-defined dusty gaps with a shape closely related to the Stokes number. Fig. 8.7 shows the same condition applied for various Shakura & Sunyaev (1973) viscosity α_SS and different aspect ratios of the disc at the location of the planet.

An important feature mentioned in Sect. 8.4.1 appears clearly: low mass planets can carve gaps in the dust only if the Stokes number is above a critical value of order unity. The value of the critical Stokes number is obtained by comparing the gap opening criterion for the gas and dust. Since the value of the critical Stokes number is expected to be higher than unity, we assume $\mathrm{St} \gtrsim \alpha$ and $\mathrm{St} \gtrsim \epsilon$ and obtain

$$\mathrm{St}_\mathrm{crit} = \begin{cases} 3z \left(\dfrac{\xi}{0.6}\right)^{2/3}, & M_\mathrm{p,lim} = 0.2\, M_\mathrm{p,th}, \\ 2z \left(\dfrac{\xi^2}{\pi}\right)^{1/3} \alpha^{-1/3} \left(\dfrac{H_\mathrm{g}}{R_\mathrm{p}}\right)^{1/3}, & M_\mathrm{p,lim} = M_\mathrm{p,visc}, \end{cases} \tag{8.36}$$

where z is given by Eq. 8.25. The small exponents $\pm 1/3$ involved make the the value of the critical Stokes number is very close to unity. When the gap opening criterion for the gas is dominated by pressure forces, i.e. $M_\mathrm{p,lim} = 0.2\, M_\mathrm{p,th}$, the value of the critical Stokes number does not depend on $H_\mathrm{g}/R_\mathrm{p}$ anymore (left panel of Fig. 8.7) since the gap opening criterion for the gas and the dust scales equally with the aspect ratio. When the gap opening criterion for the gas is dominated by viscosity, the critical Stokes number decreases with decreasing aspect ratio (center and right panels of Fig. 8.7). Planet migration would also increase of the planet mass required to carve the gap in the gas and hence, decrease slightly the value of the critical Stokes number.

8.5 Testing the criterion: Numerical simulations

Adopting the model described in Sect. 8.1.2 we study the evolution of the dust density resulting from the tides of one embedded planet located at a distance of 40 au from the central star. We perform a series of simulations varying the planet mass, the aspect ratio of the disc and the exponent p of the gas surface density profile in order to test our criterion over a wide range of disc models.

8.5.1 Results

8.5.1.1 Planet mass

We first use the disc model described in Sect. 8.1.2, and vary the planet mass around the value predicted by Eq. 8.33. The power law exponent of the gas surface density profile is $p = 0.1$, and the aspect ratio of the disc equal to 0.07 at the planet location. We simulate the evolution of 1 mm dust grains that have an initial Stokes number of $\mathrm{St} \sim 10$. As described in Sect. 8.2, those large grains settle efficiently to the midplane of the disc in a stable dust layer with dust-to-gas scale height ratio of $\sim \sqrt{\alpha_\mathrm{SS}/\mathrm{St}} \sim 0.02$, consistent with the Dubrulle et al. (1995) model and SPH simulations of dusty discs (Laibe et al. 2008). A dust-to-gas ratio in density ϵ of order unity is achieved in the midplane of the disc. This effect reduces the Stokes number compared to a case where the dust mass would be neglected (Eq. 2.37), and reinforces the gap-closing (opening) effect induced by the drag torque outside (inside) the orbit of the planet.

Our analytic criterion (Eq. 8.33) predicts for the minimum mass able to carve a gap in the dust disc to be $M_\mathrm{p} \simeq 0.026\, M_\mathrm{J}$ assuming the nominal value for $f = 0.4$ (Goldreich & Tremaine 1979). Above this mass, we expect to see gaps only in the dust for planets up to $M_\mathrm{p} \simeq 0.35\, M_\mathrm{J}$, according to our estimation of minimum mass for the gas gap (Eq. 8.35). We therefore vary the planet mass in the range $[0.01, 0.05, 0.075, 0.085, 0.095, 0.1]\, M_\mathrm{J}$ and look at the eventual structure

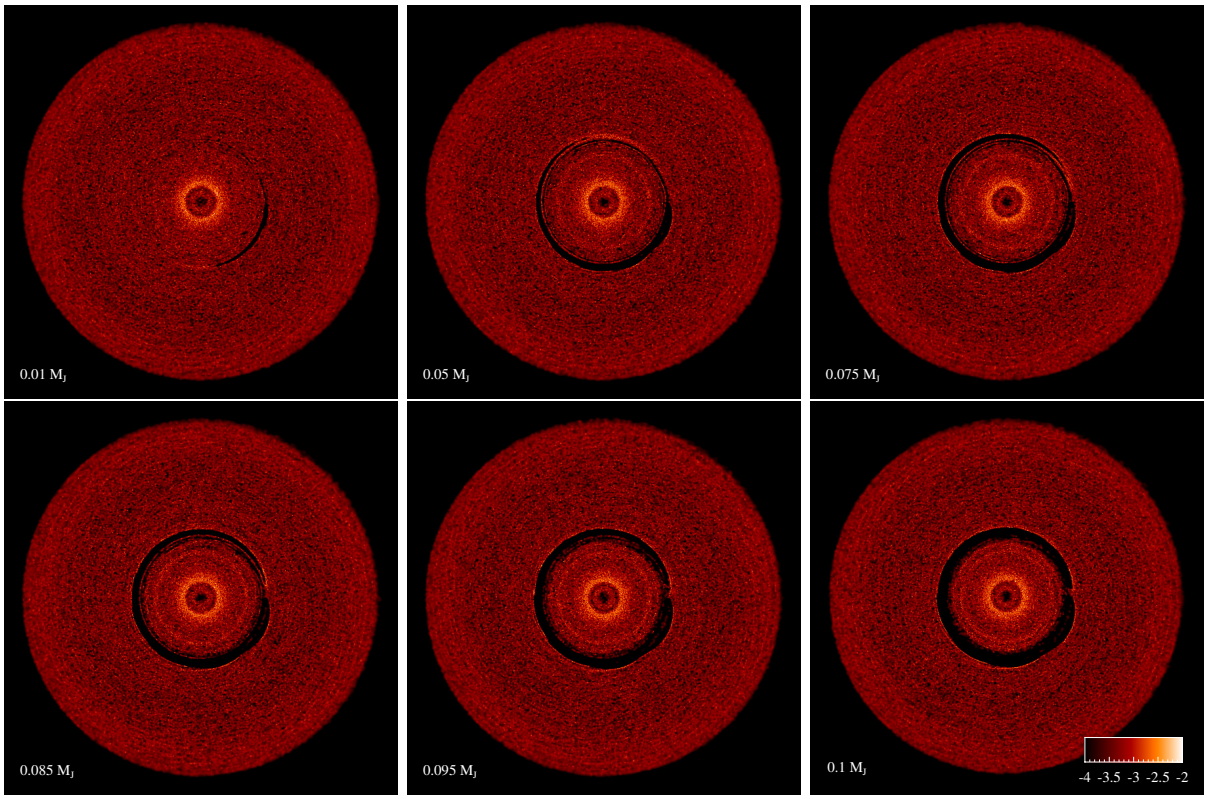

Figure 8.8: Rendered images of the steady-state dust surface density for the disc model described in Sect. 8.1.2 hosting embedded planets of mass 0.01 (top-left), 0.05 (top-center), 0.075 (top-right), 0.085 (bottom-left), 0.095 (bottom-center) and 0.1 (bottom-right) $M_{\rm J}$ initially located at 40 au after 40 planetary orbits. A full-cleared gap is carved by a planet with a minimum mass of $\sim 0.09\,M_{\rm J}$, consistent with our analytical expectations adopting $f \sim 0.2$.

of dust gaps. We evolve the simulations over 40 planetary orbits, which leaves enough time for the dust gap to form for the grain sizes considered. We find planets with mass $\lesssim 0.09\,M_{\rm J}$ produce a local depletion that is constantly replenished by the drag after the planet transit, leading to the formation of a non-axisimmetric depletion of dust. According to the criterion of gap opening expressed in Eq. 8.33, the value of the minimum planet mass to carve a gap in the dust inferred by numerical simulations is consistent with a value of the proportionality constant in front of the tidal torque expression (Eq. 8.1) given by $f \sim 0.2$. Moreover, it can be shown that the position of the outer gap edge is consistent with the value predicted by Eq. 8.23. Importantly, we have tested the no gas gap condition by running a simulation where the gravity is switched off in the dust, in order to figure out which mechanism of gap opening is in action. We have already found that for the range of planet mass presented here, the planet is not massive enough to affect the pressure structure and the gap is carved effectively by the mechanism explained in this paper (see lower right panel in Fig. 8.1).

8.5.1.2 Aspect ratio

The disc aspect ratio covers a crucial role for the gap closing process in the dust phase, since the drag induced dust motion that tends to close the gap is strictly dependent on the value of the disc aspect ratio. We therefore test our criterion using a disc model with $H/R = 0.035$ at 40 au, two times lower than the value previously adopted. According to Eq. 8.33, the minimum mass of the planet able to carve a gas gap in this disc model is 0.07 $M_{\rm J}$ (Eq. 8.35). We model the disc using a higher resolution in order to ensure the same level of viscosity of the previous model. Fig. 8.9 shows that indeed, a planet with mass $\gtrsim 0.01\,M_{\rm J}$ ($3\,M_{\oplus}$) is able to carve dust gaps, again consistent with our predictions using $f \sim 0.2$. By running a simulation with no

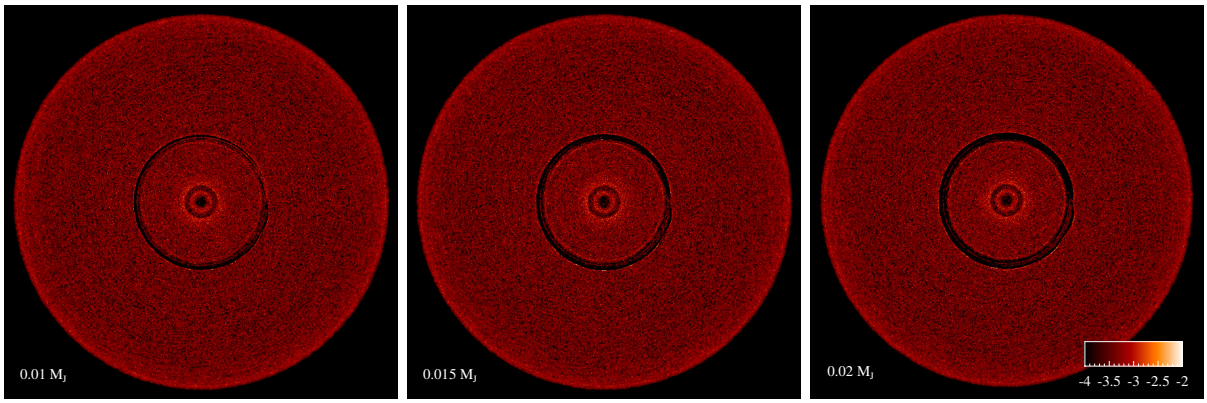

Figure 8.9: Rendered images of dust surface density for the disc model described in Sect. 8.1.2 with an aspect ratio $H/R = 0.02$ at 1 au hosting a planet with mass 0.01 (left), 0.015 (center) and 0.02 (right) M_J initially located at 40 au after 40 planetary orbits. A planet with mass $\gtrsim 0.01\,M_\mathrm{J}$ ($3\,M_\oplus$) is able to carve dust gaps, again consistent with $f \sim 0.2$

tides in the dust, we observe that such low massive planets does not affect the local pressure structure as expected.

8.5.1.3 Gas surface density profile

As mentioned in Sect. 8.4.3, the minimum mass to carve a gap in dusty discs increases with the steepening of the surface density profile. In order to test our criterion we run a simulation with a steeper surface density profile $p = 0.5$ with the same disc mass and using the same disc parameters described in Sect. 8.1.2. From Eq. 8.33, assuming $f = 0.2$, the mass of the planet able to carve a gap in dusty discs without producing pressure bumps in the gas structure is in the range $0.13\,M_\mathrm{J} \lesssim M_\mathrm{p} \lesssim 0.35\,M_\mathrm{J}$. Fig. 8.10 shows that planet with mass $0.1\,M_\mathrm{J}$ does not produce a gap in the dust, in contrast with the previous case where the same planet is able to carve a gap in a disc with a shallower surface density profile (see the bottom-right panel of Fig. 8.8). The left panel of Fig. 8.10 shows, as expected, that a slightly massive planet ($0.125\,M_\mathrm{J}$) is able to carve gap in the dust. This value of the minimum mass is consistent with the mass given by the criterion, assuming $f \sim 0.2$. Additionally, for steeper surface density, the drag helps the gap opening from inside the planetary orbit more effectively. Thus, although the gap closing effect is more enhanced outside the planetary orbit, the gap is more easily carved compared to the case of shallower density profiles and it results much wider and asymmetric around the planet location. Moreover, in order to ensure that the planet does not affect the local structure of the gas, we switched off the tidal interaction between the planet and dusty disc. We found that the drag alone cannot produce any gap in the dust, since the planet does not affect the local pressure structure.

8.5.2 Summary

We have considered a disc hosting a low-mass planet, so that no gap is carved in the gas phase. We obtained two analytic criteria for the minimum planet masses required to i) stop the inflow of dust particles (axisymmetric mechanism) and ii) ensure that drift can not refill the inner regions of the disc in dust (non-axisymmetric mechanism). These two criteria represent the necessary and sufficient conditions for dust gap opening, respectively. The exact value of the minimum masses predicted by these conditions depends on the proportionality constant in front of Eq. 8.1.

By combining our various numerical tests, for $\mathrm{St} > \alpha$, our final sufficient condition for gap opening in dusty discs is given by

$$\frac{M_\mathrm{p}}{M_\star} \gtrsim 2.2 \left(\frac{-\zeta}{1+\epsilon}\right)^{3/2} \mathrm{St}^{-3/2} \left(\frac{\mathrm{H_g}}{\mathrm{R_p}}\right)^3 . \tag{8.37}$$

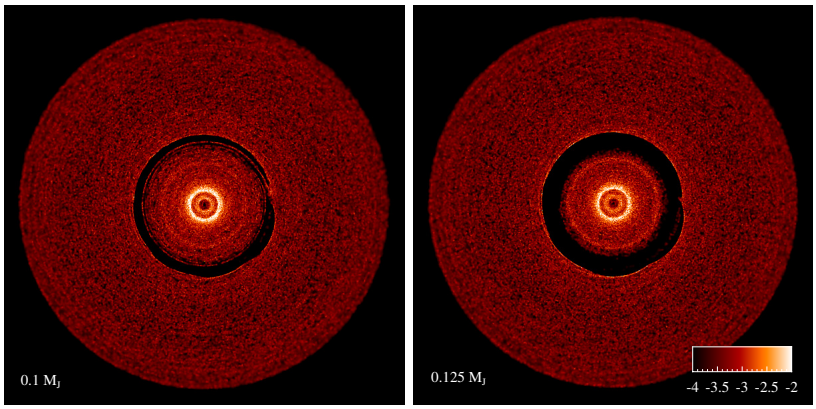

Figure 8.10: Rendered images of dust surface density for a disc with power law index of the gas surface density $p = 0.5$ hosting a planet with mass 0.1 (left) and 0.125 (right) M_J initially located at 40 au after 40 planetary orbits. A planet with mass $\gtrsim 0.125\,M_\mathrm{J}$ is able to carve gap in the dust, as predicted by the criterion adopting $f \sim 0.2$.

We predict the outer edge of the dust gap to be located at a distance $\Delta_\mathrm{gap,d}$ from the planet, where

$$\frac{\Delta_\mathrm{gap,d}}{R_\mathrm{p}} \simeq 0.8 \left(\frac{-\zeta}{1+\epsilon}\right)^{-1/4} \mathrm{St}^{1/4} \left(\frac{\mathrm{H_g}}{\mathrm{R_p}}\right)^{-1/2} \left(\frac{\mathrm{M_p}}{\mathrm{M_\star}}\right)^{1/2}. \tag{8.38}$$

Planets with masses larger than the limit given by Eq. 8.37 but lower than the one given by Eq. 8.35 shall carve a deep gap in the dust without affecting the gas structure. A dust gap can only form if $\mathrm{St} \geq \mathrm{St_{crit}} = \mathcal{O}(1)$. However, the constant in front of the criterion and the outer edge location are not accurate, since we used approximations for our model to be analytically tractable. As an example, we applied the tidal torque formula derived for a gas phase in linear regime (Eq. 8.1), which might need corrections for a gas and dust mixture in non-linear regime. As an example, modifying the value of Δ by 20% modifies the constant by a factor of 2 due to the fourth power in its expression.

8.6 Using the criterion

8.6.1 Interpreting observations of gaps

8.6.1.1 Gap detectability

Relating the morphology of dust gaps to the properties of the planet and the disc gives insights about the planet formation process. Ideally, multiple-wavelengths observations should be combined to infer density distributions of grains experiencing different aerodynamical regimes (see Chapt. 3). Scattered light emission at optical and near-infrared frequencies trace small dust grains at the surface of the disc ($\simeq 0.1 - 10\ \mu$m), where stellar photons are absorbed or scattered (Watson et al. 2007). The scattering emission intensity probes the gas structure at the surface of the disc (see Sect. 3.2.1), since these grains are efficiently coupled with the gas ($\mathrm{St} \lesssim \alpha$). Emission at (sub)-millimetre wavelengths probes surface density of large grains in the mid-plane of the disc ($\simeq 0.1 - 10$ mm), since discs are usually optically thin at these wavelengths in the vertical direction (see Sect. 3.3, Dullemond et al. 2007; Williams & Cieza 2011). A narrow beam is required to resolve the gap, together with a large signal-to-noise ratio to discriminate its weak emission. The gap depth can then be extrapolated, assuming that the weak emission in the gap is solely due to a low dust surface density.

8.6.1.2 Estimating the Stokes number

Eq. 8.37 involves the Stokes number of the grains. St can be estimated directly if the gas surface density is known. Unfortunately, hydrogen density is a quantity which is not directly measurable

in a disc. Gas masses are therefore usually estimated by processing the (sub-)millimetre continuum or line measurements of CO isotopologues (see Sect. 3.3.3). This requires to model fractions of isotopologues abundances, dust grains opacities and local gas-to-dust mass ratios (Williams & Cieza 2011; Williams & Best 2014; Miotello et al. 2014a). An estimate of the value of the Stokes number St can also be inferred indirectly via the ratio St/α coming from the thickness of the dust layer (e.g. Dubrulle et al. 1995), assuming a fixed value for the turbulence parameter α. Similarly, assuming a fixed value of the Stokes number, it is possible to infer the level of gas turbulence by analyzing the dust settling. As an example, Pinte et al. (2016) measured $H_\mathrm{d} \sim 1$ au at $R \sim 100$ au in the disc around HL Tauri. Assuming coupled grains (e.g. $\mathrm{St} \simeq 0.01$), this thickness implies $\alpha \simeq 10^{-4}$. However, i) this value is $1-2$ orders of magnitude smaller than the value consistent with typical accretion rates of protostars, ii) such low St implies typical Minimum Mass Solar Nebula gas density, which would make the planet migrate very quickly onto the star, and iii) when planets interact with discs of such low viscosity, vortices develop via the Rossby wave instabilities, trap grains and produce non-axisymmetric structures detectable by ALMA which are not observed (Lyra & Lin 2013). Assuming $\mathrm{St} \simeq 1$ and $\alpha \simeq 10^{-2}$ is more consistent with current knowledge about the planet formation process.

8.6.1.3 Gap in the gas

The mass of the hypothetical planet and the local properties of the disc can be estimated when observing a gap in optical or near-infrared scattered light emission, given a degeneracy over the ratio M_p/α (Fung et al. 2014; Kanagawa et al. 2016; Rosotti et al. 2016). If the signal-to-noise ratio is high enough, the depth of the gap (Fung et al. 2014; Kanagawa et al. 2015a) or its shape (Kanagawa et al. 2015a, 2016) can also be used to the same purpose. These methods suffer large systematic errors, mostly due to uncertainties on the local disc geometries which heavily affect the surface brightness around the gap (Jang-Condell & Turner 2012). Scattering emission may additionally reveal spiral structures, whose morphologies may be related to the mass of the planet and the aspect ratio of the disc (Zhu et al. 2015a; Juhász et al. 2015; Dong et al. 2015b).

8.6.1.4 Applying the criterion

A good use of the criterion starts with two preliminary remarks. Firstly, Fig. 8.6 shows that $M_\mathrm{p} \gtrsim M_\mathrm{p,lim}$ is the only condition required for planets to open gaps in both the gas and the dust and this, independently of the Stokes number. Thus, if a gap is detected in near-infrared scattering *and* thermal-mm emission, no information can be extracted on St with the criterion derived in this study. Indeed, grains of all sizes tend to drift towards the pressure maximum at the outer edge of the gas gap, as long as $\mathrm{St} > \alpha$. Further analysis of the morphological details of the gap should be conducted to infer the properties of the system. Secondly, the absence of *any* gap in both gas and dust does not necessarily reflect the absence of any gravitational body in the disc. Grains may replenish the orbit of the planet as they drift inwards if $\mathrm{St} \lesssim 1$. However, if any additional detection limit is given for the maximum planet mass a disc can embed, Eq. 8.37 provides a condition on the minimum Stokes number compatible with the eventual existence of the planet.

We now focus on the non-trivial case, i.e. a gap detected only at millimetre wavelengths. This work shows that any low mass planet can create this structure as long as the Stokes number is large enough. As an example, for $\mathrm{St} \simeq 20$, the necessary mass of the planet required to open the dust gap is $\simeq 100$ lower than the one required to open a gap in the gas due to the factor $\mathrm{St}^{3/2}$ in Eq. 8.37. We therefore expect observations of dust only gaps to be more frequent in the outer disc, where grains of a given sizes have Stokes numbers much larger than unity. At least, $\mathrm{St} \gtrsim 1$, which constrains the *maximum* local density of the gas. Eq. 8.38 gives the expression of the distance between the planet mass and the outer edge of the dust gap. Even if the Stokes number is only roughly approximated, the weak sensitivity brought by the factor $\mathrm{St}^{1/4}$ allows to determine the planet mass relatively precisely (assuming that the aspect ratio of the disc is

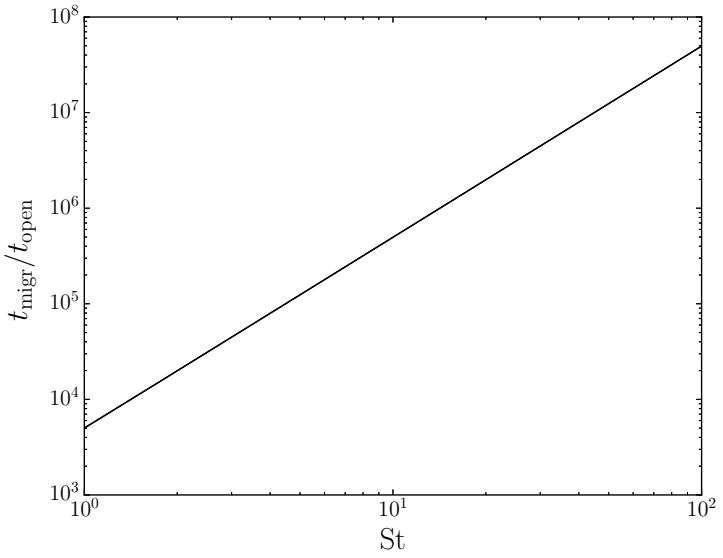

Figure 8.11: Ratio between the type I migration time scale and the gap opening time in dusty discs with an embedded planet with mass 0.1 $M_{\rm J}$. The gap is carved in the dusty disc faster than the planet migration.

known). The absence of any gap in the gas additionally implies that $M_{\rm p} < M_{\rm p,lim}$. Combined with the roughly known value of $M_{\rm p}$, this condition provides a minimum value for α in the disc. Comparing this value and the degree of dust settling may give a way to infer if the seminal diffusive description of turbulence in disc is relevant or not. Interestingly, if several dust only gaps are detected in the same disc, the degeneracy over the constant in front of the criterion can be broken, helping to determine the masses of the planets more precisely.

Finally, a sufficiently massive planet may slightly affect the local pressure structure leading to a deceleration of marginally coupled dust particles, i.e. St ~ 1. This alternative mechanism has been suggested to produce a shallow variation of the dust surface density of order unity that can be detected with ALMA (Rosotti et al. 2016). This mechanism has not been considered in the analysis presented in this work, since our main assumption is that the pressure structure is not affected by the presence of the planet.

8.6.2 Limitations

In this study, we have restricted our analysis to planets on fixed orbits. Migration may strongly affect the ability of a planet to carve gaps in the gas phase (Malik et al. 2015). On the contrary, Fig. 8.11 shows that the dust gap opening time scale obtained from Eq. 8.31 is much shorter than the migration time scale, estimated from the differential Lindblad torque derived in Tanaka et al. (2002) (low mass planets) and the disc model described in Sect. 8.1.2 with an embedded planet with mass 0.1 $M_{\rm J}$. Hence, planet migration is not expected to affect the ability of a planet to carve gaps in our mechanism. Note that the ratio between both timescales depends indirectly on the Stokes number through the local gas density.

We have also assumed dust grains of constant sizes. It is known that dust coagulation or fragmentation may strongly affect the dust dynamics (e.g. Laibe et al. 2008; Birnstiel et al. 2010). However, we can safely neglect the grains size evolution over the small time required to open the dust gap. Finally, we have assumed that back-reaction is weak enough and does not affect the gas surface density significantly as grains are repelled outside of the planet orbit. Whether back-reaction can trigger the formation of gaps in the gas would be worth investigating, but remains beyond the scope of this study. Hence, the criterion has not been proven to work at dust-to-gas ratio larger than unity.

8.7 Conclusion

We have identified two physical mechanisms for dust gap opening by embedded planets in dusty protoplanetary discs. Our conclusions are:

i) Gap formation in the gas is not a necessary condition for opening a gap in the dust.

ii) For low mass planets that do not produce pressure maxima in the gas, drag resists (assists) the tidal torque outside (inside) the planetary orbit, forming an asymmetric gap around the planet orbit.

iii) For high mass planets that create pressure maxima in the gas, solid particles are prevented from accreting onto the planet. Here drag assists the tidal torque, leading to the formation of a deep dust gap with a stable population of grains at the corotation region.

We derived an analytic criterion that predicts the minimum mass required for a planet to open a gap in the dust phase of a viscous protoplanetary disc in the case where the planet do not perturb the local pressure profile of the disc. In this regime, a gap opens in the dust if the tidal torque overpowers the drag torque outside the planet orbit. We generalised the approach of NSH86 to include the disc-planet tidal interaction and the viscous forces in the equations of motion. Gas and dust velocities in steady state were obtained analytically (Eq. 8.8-8.11). From there, assuming that the planet is not able to affect the local pressure structure, we derived a relation between the minimum mass required to open a gap in the dust, and the key parameters of the dust motion: the Stokes number, the aspect ratio of the disc and the dust-to-gas ratio (Eq. 8.34). We benchmarked the value of the scaling constant in front of the tidal torque density formula using 3D dust-and-gas SPH simulations of various discs.

Our final opening criterion for dust gaps is given by Eq. 8.37. We found that low mass planets are able to carve dust gaps when the Stokes number $\mathrm{St} \geq \mathrm{St}_{\mathrm{crit}} \simeq 1$ (Eq. 8.36). We also derived an analytic formula for the radial extension of the outer dust gap edge (Eq. 8.38).

Observations in near-infrared scattering or millimetre thermal dust emission might reveal gaps, in both phases or in the dust only. This criterion can be used to constrain the mass of the planet embedded in the disc, or the Stokes number of the grains.

Chapter 9

On Planet Formation in HL Tau

Based on the paper by ***Giovanni Dipierro****, Daniel J. Price, Guillaume Laibe, Kieran Hirsh, Alice Cerioli and Giuseppe Lodato: "On planet formation in HL Tau" published in Monthly Notices of the Royal Astronomical Society Letters,* ***453****, L73-L77 (2015).*

In this Chapter we apply our study on the gap opening processes in dusty discs (see Chapter 8) on the HL Tau protoplanetary disc, recently imaged in long-baseline observations by ALMA. The image of the disc revealed a striking series of concentric, axisymmetric gaps, thought to indicate the presence of massive protoplanetary bodies caught in the act of forming planets (ALMA Partnership et al. 2015). While gap carving by planets embedded in discs has been predicted theoretically since at least Lin & Papaloizou (1986), the axisymmetry of the rings in HL Tau is puzzling. First, prominent spiral waves induced by the planet are seen in all theoretical simulations of planets carving gaps in gas discs whenever the planet is massive enough to open a gap (e.g. de Val-Borro et al. 2006). Second, prior observational estimates of the relative disc mass in HL Tau suggested a disc of 0.03–0.14 $M_\odot$ around a 0.5-1.3 $M_\odot$ star (e.g. Robitaille et al. 2007; Kwon et al. 2011, 2015; the latter authors estimated ≈ 0.1 $M_\odot$ based on fitting disc models to the previous best observations at mm wavelengths). Massive discs should almost certainly show gravitationally instability and hence prominent and large scale spiral arms (see Part II, Lodato & Rice 2004; Dipierro et al. 2014, 2015a), as seen in the theoretical predictions for HL Tau made by Greaves et al. (2008). Third, HL Tau is in an early phase of star formation (Class I evolving to Class II), emitting a 100-200 km/s highly collimated jet (Mundt et al. 1990). The ongoing accretion from the envelope means that the disc is likely relatively thick, massive and hot, yet the clarity of the gaps are more reminiscent of those found in simulations of thin, cold discs.

In this Chapter we explain the axisymmetric gaps seen in the ALMA image as being due to the different response of gas and dust to embedded planets in protoplanetary discs. We perform global, three dimensional dusty smoothed particle hydrodynamics calculations of multiple planets embedded in dust/gas discs which successfully reproduce most of the structures seen in the ALMA image. We find a best match to the observations using three embedded planets with masses of 0.2, 0.27 and 0.55 M_J in the three main gaps observed by ALMA, though there remain uncertainties in the exact planet masses from the disc model.

9.1 Methods

9.1.1 Dust/gas simulations

We perform 3D global simulations of dust/gas discs with embedded protoplanets using the PHANTOM smoothed particle hydrodynamics (SPH) code written by Daniel Price (Price & Federrath 2010; Lodato & Price 2010; Price 2012; Nixon et al. 2013).

We assume a single grain size per calculation. For grains of mm-size and larger we employ the two fluid algorithm described in Laibe & Price (2012b,a). For calculations with smaller grains we employ the new single fluid algorithm based on the terminal velocity approximation (Laibe & Price 2014; Price & Laibe 2015). As described in Sect. 8.1, we assume a physical drag prescription where the drag regime is chosen automatically depending on the ratio of the grain size to the gas mean free path (Kwok 1975; Paardekooper & Mellema 2006; Laibe & Price 2012b). All of the grain sizes used in this paper fall into the Epstein regime, meaning that the stopping time is given by Eq. 2.37 assuming an intrinsic grain density of 1 g/cm^3.

9.1.1.1 Sink particles

We place three planets, initially located at 13.2, 32.3 and 68.8 au, in the three main gaps observed in HL Tau (ALMA Partnership et al. 2015). We represent the protoplanetary bodies and the central star using sink particles (Bate, Bonnell, & Price 1995) with accretion radii of 0.25, 0.25 and 0.75 au, respectively. Dust and gas can be accreted onto the sinks provided that the material is bound to the sink and the velocity divergence is negative. The planets are free to migrate due to planet-disc interactions and the viscous disc evolution. We found migration to be negligible on the timescales simulated.

9.1.2 Initial conditions

9.1.2.1 Gas

We adopt the same initial considtions described in Sect. 8.1.2 and repeated here for clarity and justified for the HL Tau disc modelling. The system comprises a central star with mass $1.3M_\odot$ surrounded by a gaseous disc of 10^6 SPH particles which extends from $R_{\rm in} = 1$ to $R_{\rm out} = 120$ au. We assume a power-law surface density profile given by $\Sigma_{\rm g}(R) = \Sigma_{\rm g,in}(R/R_{\rm in})^{-p}$, where $\Sigma_{\rm g,in}$ is the surface density at the inner edge. We assume a radially isothermal equation of state $P = c_{\rm s}^2(R)\rho$, where $c_{\rm s} = c_{\rm s,in}(R/R_{\rm in})^{-q}$, such that the gas temperature power law is $T(R) \propto (R/R_{\rm in})^{-2q}$ and where $c_{\rm s,in}$ is the sound speed at $R_{\rm in}$. The aspect ratio in the gas is thus prescribed by the p and q indices according to $H_{\rm g}/R = (H_{\rm g}/R)_{\rm in}\,(R/R_{\rm in})^{\frac{1}{2}-q}$. We set the SPH $\alpha_{\rm AV}$ viscosity parameter to 0.1 giving an effective Shakura & Sunyaev (1973) viscosity $\alpha_{SS} \approx 0.005$.

9.1.2.2 Constraining the disc model from HL Tau

While Kwon et al. (2011) fit various disc profiles to their previous low-resolution observations of HL Tau to constrain the surface density and temperature profiles, we use the gaps observed in the disc (ALMA Partnership et al. 2015) to directly constrain the aspect ratio.

In particular, assuming that the gap is carved both in the gas and in the dust, while the gap is deeper in the dust (see Chapter 8, Fouchet et al. 2007), the *width* is the same. Since a gap cannot be opened with a width less than the gas disc scale height the three prominent gaps observed in HL Tau provide an upper limit on $H_{\rm g}$ as a function of radius. Combining this with the dust temperature profile observed by ALMA, which, assuming that the dust and gas are thermally coupled, provides a lower limit on the gas temperature, we can directly constrain the sound speed profile. Figure 9.1 shows the constraint provided by the gap widths together with the lower limit and our best fitting power law $q = 0.35$ ($T \propto R^{-0.7}$) with $(H_{\rm g}/R)_{\rm in} = 0.04$. This gives $(H_{\rm g}/R)_{\rm out} \approx 0.08$.

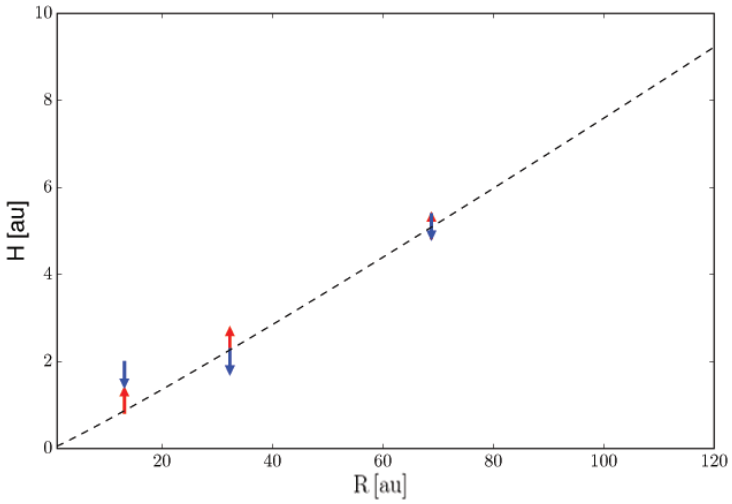

Figure 9.1: Constraint on the disc scale height profile obtained from the gap widths observed in the ALMA image (upper limits; blue points) combined with the lower limit obtained from the dust temperature as a function of radius (red points) along with our best fit $T \propto R^{-0.7}$ (dashed line).

We choose a shallow surface density profile $p = 0.1$ for computational efficiency since it avoids expensive computation of many particles on the shortest orbits. This reproduces the inner disc in our simulations, but gives an outer disc that is over-luminous compared to the inner disc in the simulated observations (Fig. 9.4). While more work is needed on the disc model, our main aim in the present paper is to reproduce the axisymmetry and size of the gaps.

9.1.2.3 Gas disc mass and dust-to-gas ratio

Hayashi et al. (1993) found a H_2 gas mass of $0.03 M_\odot$ contained within 1400 au in HL Tau based on the observed ^{13}CO emission. We assume this as our starting point, noting that the gas mass contained within a particular radius depends on the density profile according to

$$M_\mathrm{g}(< R) = \frac{2\pi}{2-p} \Sigma_\mathrm{g,in} R_\mathrm{in}^2 \left[\left(\frac{R}{R_\mathrm{in}} \right)^{2-p} - 1 \right]. \tag{9.1}$$

We assume the dust was initially co-located with the gas but has migrated to a radius ≈ 10 times smaller than the initial radius to give the 120 au dust disc visible with ALMA. Assuming an initial dust-to-gas ratio of 0.01 in the undifferentiated material at 1400 au, this gives $\Sigma_\mathrm{g} \approx \Sigma_\mathrm{d}$ within the smaller radius given our assumed p index of 0.1. Hence we assume a dust-to-gas ratio of unity within 120 au and set $\Sigma_\mathrm{g,in}$ such that $M_\mathrm{g} = 0.03 M_\odot$ at 1400 au. The total gas mass within 120 au is thus 0.0002 $M_\odot$, but this is low mainly due to our assumed shallow surface density profile.

A steeper surface density profile, e.g. $p = 1$ would give a more realistic gas mass of $0.003 M_\odot$. While even this may seem in conflict with earlier fits to much higher disc masses (e.g. Robitaille et al. 2007; Kwon et al. 2011, 2015), these assume that the dust is co-located with the gas, whereas we expect settling and radial inward migration to produce a more compact dust disc, consistent with observations in other systems (see Sect. 3.3.2, Andrews et al. 2012; de Gregorio-Monsalvo et al. 2013). Moreover, Tamayo et al. (2015) found that such a massive gas disc within 120 au would smear out the eccentric features observed in HL Tau since the precession periods of pericentres would be faster than the timescale on which planets carve gaps. Our lower disc mass also ensures that the mm grains are close to St = 1 in order to guarantee an efficient migration towards the pressure maxima in the disc induced by the embedded protoplanets. Again, a steeper surface density profile would produce less coupling of the mm grains in the outer disc, giving dynamics similar to those we find with cm grains. Indeed, as we shall see, the spiral

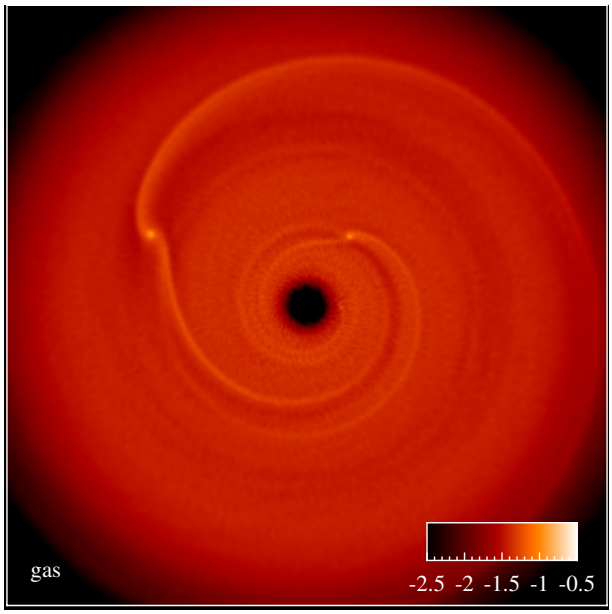

Figure 9.2: Rendered image of gas surface density for a disc containing three embedded protoplanets of mass 0.2, 0.27 and 0.55 $M_{\rm J}$ initially located at the same distance as the gaps detected in HLTau, at 13.2, 32.3 and 68.8 au.

waves induced in larger grains by the outer planet do appear to be visible at mm wavelengths in HL Tau.

9.1.2.4 Dust

The dust disc is setup assuming an initially constant dust-to-gas ratio (equal to unity, see above). Since we simulate one grain size at a time we assume a power-law grain size distribution as expressed in Eq. 3.8,

$$n(a) \propto a^{-m} \quad \text{for} \quad a_{\rm min} < a < a_{\rm max}, \tag{9.2}$$

where we assume $a_{\rm min} = 1\,\mu$m, $a_{\rm max} = 10$ cm and $m = 3.5$ and perform six different simulations using $s = 1\,\mu$m, $10\,\mu$m, $100\,\mu$m, 1 mm, 1 cm and 10 cm, respectively (Fig. 9.3). The grain size distribution is appropriate to interstellar dust and particles growing by agglomeration in protoplanetary discs (Draine 2006). While approximate, this is more realistic than assuming a constant dust-to-gas ratio for each grain size, capturing the main idea that small grains are more abundant than large grains.

In our two fluid simulations we set up the dust disc as a separate set of dust particles. We use one quarter of the number of dust particles compared to gas to prevent dust from becoming artificially trapped below the gas resolution (Laibe & Price 2012b; Ayliffe et al. 2012). For the calculations employing the single fluid algorithm (Price & Laibe 2015) there is only one set of particles describing the mixture, set up as described above for the gas but with a dust fraction ϵ set for each particle to provide an initially constant dust-to-gas ratio.

9.1.3 Simulated ALMA observations

We perform simulated observations of our model using the RADMC-3D Monte Carlo radiative transfer code (Dullemond 2012) together with the CASA ALMA simulator (version 4.1), focussing on ALMA band 6 (continuum emission at 233 GHz). The source is located in the position of HL Tau, adopting the disc inclination and position angle given by ALMA Partnership et al. (2015). The procedure is as in Chapter 6 except that the dust distribution is used directly from our simulation data rather than being prescribed. The dust model consists of spherical

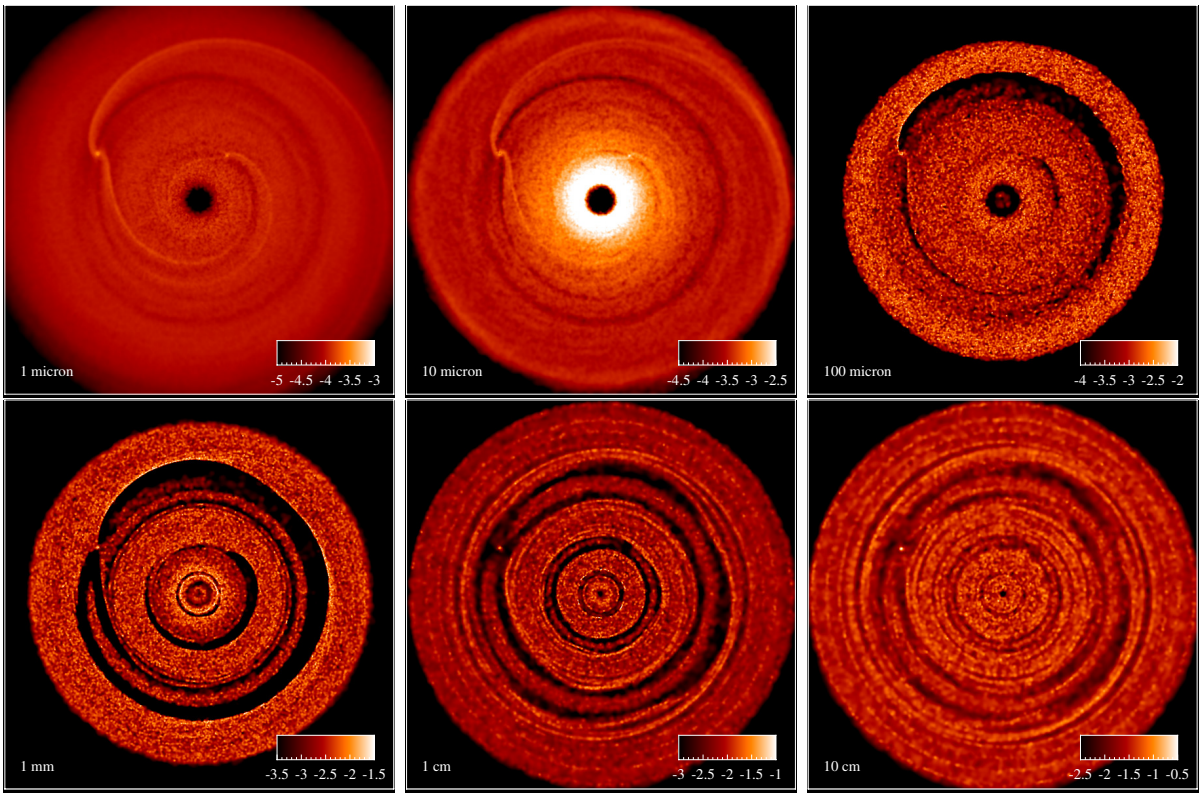

Figure 9.3: Rendered images of dust surface density for a disc containing three embedded protoplanets of mass 0.2, 0.27 and 0.55 M_J initially located at the same distance as the gaps detected in HL Tau. Each panel shows the simulation with gas plus grains of a particular size (as indicated).

silicate grains with optical constant for magnesium-iron grains taken from the Jena database[1]. Full-resolution images produced by RADMC-3D simulations are used as input sky models to simulate realistic ALMA observations accounting for thermal noise from the receivers and the atmosphere and assuming a perfect calibration of the visibility measurements. A transit duration of 8 hours is used to reach an optimal signal-to-noise ratio. The observation parameters are chosen to match the spatial resolution of the observations (see caption of Fig. 9.4).

9.2 Results

Fig. 9.2 shows the gas surface density, while Fig. 9.3 shows rendered images of the dust surface density from the disc model with three embedded protoplanets of masses of 0.2, 0.27 and 0.55 M_J, respectively, shown after ∼10 orbits of the outer planet. None of the planets produce clean gaps in the gas disc (Fig. 9.2), but as expected, the dust density perturbations induced by protoplanets depend strongly on the grain size (e.g. Fouchet et al. 2007; Ayliffe et al. 2012), or more specifically on the Stokes number. The density distribution of μm-sized particles (top left and centre) are similar to the gas distribution due to the stronger coupling. The micrometer sized grains (top left) capture the spiral density wave launched by the protoplanets. Millimetre-sized particles (bottom left) are most affected by the density waves induced by the protoplanets, exhibiting the largest migration towards the gap edges. Interestingly, the dust density distributions of 10 mm and cm-sized particles (bottom centre and right) show axisymmetric waves launched by the planets propagating across the whole disc. The gaps carved by protoplanets in cm grains (St ≈ 10) show the formation of horseshoe regions. As expected, the formation of the horseshoe region in planetary gaps depends on the dust-gas coupling, since poorly coupled large particles tend to have frequent close encounters with the planet. We stress that the important parameter

[1] `http://www.astro.uni-jena.de/Laboratory/Database/databases.html`

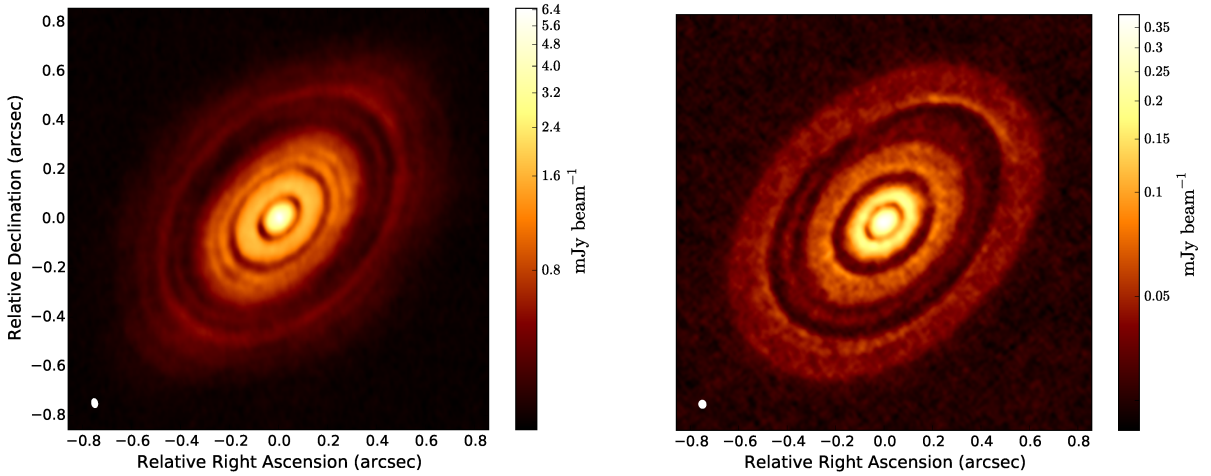

Figure 9.4: Comparison between the ALMA image of HL Tau (left) with simulated observations of our disc model (right) at band 6 (continuum emission at 233 GHz). Note that the color bars are different. The white colour in the filled ellipse in the lower left corner indicates the size of the half-power contour of the synthesized beam: (left) 0.035 arcsec × 0.022 arcsec, P.A. 11°; (right) 0.032 arcsec × 0.027 arcsec, P.A. 12°.

is the Stokes number rather than the actual grain size, so equivalent dynamics can be produced in different grains by changing the gas density or with different assumptions about the grain composition (see Eq. 6.1).

An intriguing possibility is to infer the disc viscosity from the morphology of the gaps. Since our results show that the planets are not massive enough to open gaps in the gas or, equivalently, the viscosity is high enough to effectively smooth the density profile, we can only evaluate a lower limit for α_{SS} below which planets are able to carve gaps in the disc model adopted here. Using the criterion in Crida et al. (2006) we infer $\alpha_{SS} \gtrsim 0.0002$.

Fig. 9.4 compares the ALMA simulated observation of our disc model at band 6 (right) with the publicly released image of HL Tau (left). The pattern of bright and dark rings is readily detectable with ALMA using a selection of observation parameters that ensure a resolution close to the one reached in the real ALMA image. As expected, the emission probes the mid-plane disc surface density in large grains (0.1-10 mm), due to their high opacity at these wavelength (Dullemond et al. 2007; Williams & Cieza 2011). However, the intensity in the ALMA simulated observation of this model is lower than in the HL Tau image. This difference is likely due to the low dust mass, suggesting that the higher disc mass from a steeper surface density profile would produce better signal-to-noise ratio.

9.3 Summary

Inspired by the recent ALMA image of HL Tau, we have performed 3D dust/gas SPH simulations of a disc hosting multiple planets aimed at reproducing the peculiar morphology of alternating dark and bright rings shown in the observation. We used 3D radiative transfer modelling to produce synthetic ALMA observations of our disc model in order to test whether dust features induced by embedded protoplanets can explain those detected. We find:

1. The absence of spiral structure observed in HL Tau is due to the different response of the dust compared to the gas to embedded planets. Axisymmetric rings such as those observed are a natural product of embedded planets in the disc when observing the dust. Spiral structure may still be present in the gas. Hence dust/gas modelling is crucial for interpreting these systems.

2. In agreement with previous authors and our analysis described in Chapter 8, we find it

is easier to carve gaps in dust than in gas. Hence gaps observed in the continuum at mm-wavelengths do not necessarily indicate gaps in gas.

3. We can reproduce all the major features of HL Tau observations with three embedded planets at 13.2, 32.3 and 68.8 au with planet masses of 0.2, 0.27 and 0.55 $M_{\rm J}$, respectively, though there is uncertainty in the exact masses from the assumed disc model. This supports the conclusion that embedded protoplanets are responsible. Our masses are similar to the 0.2 $M_{\rm J}$ suggested for all three planets in Dong et al. (2015c) and consistent with a $10M_{\rm J}$ upper limit for the third planet given by Testi et al. (2015). The lower drag makes these waves more visible since the drag effect on the radial dust motion is negligible. We therefore suggest that the axisymmetric perturbations in the ALMA images in the outer part of the second and third planet orbits are generated by the lower coupling of mm-particles. Further simulations with a steeper surface density profile should be able to reproduce this.

There are many caveats to our results, and these are preliminary investigations. Nevertheless they illustrate that the ALMA observations of HL Tau can be understood from the interaction between dust, gas and planets in the disc.

Part IV
Horseshoes in Transitional Discs

Chapter 10

Transitional Discs

Transitional discs are some of the most intriguing and studied circumstellar discs in the star and planet formation community. Despite their small fraction, ∼10% of the total of the spatially resolved protoplanetary disc sample (Williams & Cieza 2011; Koepferl et al. 2013), transitional discs have received special attention in circumstellar discs studies. The distinct shape of their SEDs in different spectral regions and the wide range of (non-)axisymmetric structures detected in near-infrared scattering and thermal dust emission at (sub-)millimetre wavelengths have lead many scientists to conclude that these objects are being caught in an important phase of disc evolution. Transitional discs are thought to be excellent candidates to study and test the planet formation theories and to investigate the mechanism that drives the disc dispersal. The wide range of structures detected in transitional discs encompasses all the disc substructures explored in this thesis: spiral arms (see Part II, Muto et al. 2012; Grady et al. 2013; Garufi et al. 2013; Christiaens et al. 2014; Benisty et al. 2015; Wagner et al. 2015; Garufi et al. 2016; Stolker et al. 2016), dust and gas depleted cavities/gaps (see Part III, Brown et al. 2009; Andrews et al. 2011; Kraus & Ireland 2012; van der Marel et al. 2016b; van der Plas et al. 2016; de Boer et al. 2016; Ginski et al. 2016; van Boekel et al. 2016) and azimuthal asymmetries (Brown et al. 2009; Casassus et al. 2012, 2013; Isella et al. 2013; van der Marel et al. 2013; Pérez et al. 2014; Casassus et al. 2015b; van der Marel et al. 2015a). In this Chapter we provide a brief overview of transitional discs with particular focus on explaining the main features observed to date: cavities and horseshoe (see Chapter 11). A comprehensive description of the observational results and theoretical investigations on this class of discs can be found in Espaillat et al. (2014) and Owen (2015), to which we refer for a more detailed analysis.

10.1 Insights from observations

The first detection of this new class of discs was performed by Strom et al. (1989), who revealed that, in a small fraction of disc in their sample, the disc emission shows a little or no excess at near-infrared (NIR) and/or mid-infrared (MIR) wavelengths, but a significant far-infrared (FIR) excess. From the earliest evidences, these objects were called *transitional discs*. The term "transitional" was used to indicate that these objects can be in a transition phase from optically thick gas/dust-rich discs extending inward to the stellar surface to objects where the disc has been dispersed. Although there is no agreed definition of a transitional disc (e.g. Currie et al. 2009; Sicilia-Aguilar et al. 2011), the general idea to identify a transitional disc based on a large depletion in opacity in the NIR appears to be well accepted by the community.

The initial interpretation of transitional discs was based on the analysis of their SEDs, which, as described in Sect. 3.1, trace different radial regions of the disc at different wavelengths. The lack of NIR emission might indicate that their inner regions are depleted of small dust grains by several orders of magnitude, while the flux emission at longer wavelengths resembles the typical emission coming from a dust-rich object in the outer regions. These hypothesis were confirmed by recent IRS/Spitzer observations of a large number of objects that have shown

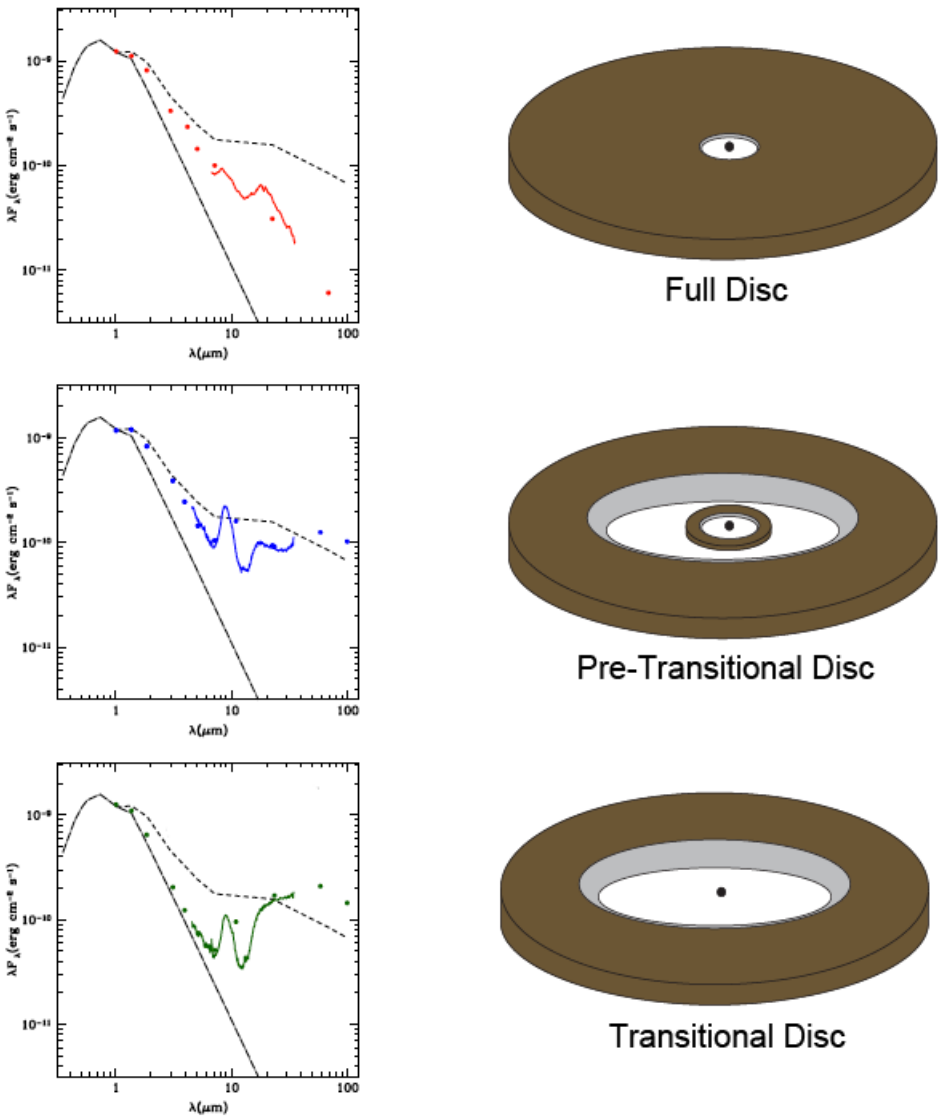

Figure 10.1: (left) SEDs of (top) full disc (Ingleby et al. 2011), (middle) pre-transitional disc (Espaillat et al. 2007) and (bottom) transitional disc (Calvet et al. 2005). The short-dashed line is a typical SEDs of a typical young disc in Taurus (D'Alessio et al. 1999) and the long-dashed line represent the stellar photosphere emission model. (right) Schematic of the expected small grains dust structure in (top) full, (middle) pre-transitional and (bottom) transitional disc. Image adapted from Espaillat et al. (2014).

SEDs with dips in the NIR and/or MIR, consistent with a (partial) depletion of small grains in the inner regions (e.g. Calvet et al. 2005; D'Alessio et al. 2005). Moreover, discs with dips only in the MIR part of the spectra (roughly $\sim 5 - 20\,\mu$m) were classified as *pre-transitional discs*, consistent with the presence of an optically thick annular ring in the inner regions that is thought to be dispersed and then form the typical cavity of a transitional disc. The left panels of Fig. 10.1 show the comparison among SEDs of a typical evolved disc (that shows less emission at all wavelengths), pre-transitional (only MIR deficit, except for the $10\,\mu$m silicate emission feature) and transitional disc (NIR and MIR deficit). The right panels of Fig. 10.1 shows the widely believed interpretation of this kind of discs based on a simple SED-based fitting model (Espaillat et al. 2014). Importantly, recent observations have revealed that some SEDs shows a stronger FIR excess compared to a typical SED. This can be explained by the fact that the stellar irradiation that would normally be absorbed in the inner disc at NIR wavelengths of a primordial disc is instead absorbed and re-irradiated by the outer disc. However, as described in Sect. 3.1, modelling the radial dusty disc structure by SED-fitting is highly degenerate.

The idea of dust hole or cavities to explain the SEDs of this new class of discs gained substantial success when transitional discs were directly imaged by high-resolution observations at (sub-)millimetre wavelengths, showing clearly dust cavities close to the star (e.g. Brown et al. 2009; Isella et al. 2010a; Andrews et al. 2011; Pérez et al. 2014). The detected cavities were thought as a clear evidence that the NIR deficit can be related to a lack of dust emitting temperatures from ∼1000 K to a few hundred K in the inner disc regions. Observations at different wavelengths reveal a wavelength dependency of the cavity size (e.g. Pinilla et al. 2015) and a radial dependence of the spectral index α (see Sect. 3.3.1, van der Marel et al. 2015a; Wright et al. 2015), indicating that the larger dust grains are usually more radially concentrated in the density maxima at the cavity edge, consistent with the expectation of dust trapping of large grains (see Sect. 2.5). This feature of transitional discs can be interpreted by taking into account the different motion of dust with different sizes (see Chapt. 2). While the radial distribution of micron-sized particles is expected to be similar to the gas distribution, a large radial separation between the gas and dust is expected for marginally-coupled (typical $\sim$ millimetre) particles (see Chapter 2).

In some cases, the SEDs of transitional discs show strong 10 μm silicate emission features (e.g. Calvet et al. 2005; Espaillat et al. 2010). As described in Sect. 3.2.2, this emission feature is related to the optically thin emission of $\sim$ micron sized dust grains close to the star, probably inside the dust hole. This suggests that there is a small amount of optically thin dust in the hole and consequently, since small grain size are tightly coupled with the gas, it is highly expected that the dust cavity is partially filled with gas. This conclusion has been recently confirmed by ALMA high-resolution observations of ^{13}CO and C^{18}O molecular lines, which have revealed gas inside the dust cavity (e.g. Pontoppidan et al. 2008; Bruderer 2013; Carmona et al. 2014; Banzatti & Pontoppidan 2015; van der Marel et al. 2016b; Carmona et al. 2016). Interestingly, in some cases the ^{12}CO becomes optically thin inside the dust cavity, allowing to infer the level of gas content inside the cavity. Remarkably, it was found that the gas density drops inside the dust cavity. The gas cavity extent is smaller than the dust cavity, with a value of up to twice as small in radius (Bruderer et al. 2014). In some cases, physical-chemical modeling of transitional discs reveals that the gas surface density drops inside the cavity by at least a factor 10, while the dust density drops by at least a factor 1000 (Bruderer et al. 2012; Bruderer 2013; van der Marel et al. 2015b).

Furthermore, one of the most intriguing aspect of transitional discs is that the majority of this objects also exhibit gas accretion rates comparable to (or slightly smaller than) T-Tauri disc accretion rates, $\sim 10^{-7} - 10^{-8}\,\mathrm{M_\odot\,yr^{-1}}$ (Muzerolle et al. 2000; Calvet et al. 2002, 2005; Espaillat et al. 2007, 2008). These finding were rather unexpected due to the depletion of gas inside the cavity, suggesting that a significant mass reservoir of gas in the inner disc is present.

In summary, there are three observational features to define (pre-)transitional discs:

- a wide dust cavity optically thin at NIR wavelengths with a sharp transition at the cavity edge,
- a gas cavity smaller and shallower than the dust cavity,
- a moderate gas accretion rate.

The most important conflict in transitional disc modelling is between the moderate disc accretion rate and the NIR deficit. These features are related to the small grains in the cavity, which are expected to be tightly coupled with the gas (see Chapter 2), putting strong constraints on any theoretical attempts to explain these objects.

A sample of transitional discs have been imaged with high-resolution polarized scattering images at NIR wavelengths. These images reveals a wide range of (non-)axisymmetric features, such as gaps and spirals (see Sect. 4.3.2, e.g. Stolker et al. 2016; van Boekel et al. 2016), further challenging the modelling of the disc emission. Importantly, one of the most spectacular first results with ALMA was the observation of a horseshoe-like structure in the dust continuum

emission in Oph IRS 48 (van der Marel et al. 2013), with subsequent observations revealing asymmetric structures in several other transitional discs (Casassus et al. 2013; Isella et al. 2013; Pérez et al. 2014; van der Marel et al. 2016b). Such features are most commonly interpreted as vortices, for example arising from the Rossby Wave Instability (RWI) at the edge of the gap formed by a young planet (Lovelace et al. 1999; Lyra et al. 2009; Lyra & Lin 2013; Zhu & Stone 2014) or by a dead zone (Regály et al. 2012, 2013), or as a result of internal dynamical processes associated with the presence of a weak magnetic field (Ruge et al. 2016). RWI arises in sufficiently inviscid discs, with equivalent α parameters (Shakura & Sunyaev 1973) of the order of $\alpha \lesssim 10^{-4}$. Such a low viscosity allows the planet-induced vortex to survive for thousands and up to 10^4 planetary orbits (de Val-Borro et al. 2007; Ataiee et al. 2013; Zhu & Stone 2014; Fu et al. 2014b). Vortices can effectively trap dust particles, leading to a more azimuthally and radially concentrated dust density distribution of larger grains at the center of the vortex (Barge & Sommeria 1995; Birnstiel et al. 2013; Lyra & Lin 2013; Ruge et al. 2016). However, the combined effect of the dust settling and trapping inside the vortex produces an enhanced dust-to-gas mass ratio in the vortex, leading to an increase of the dust back-reaction. This produces an alteration of the coherent vorticity pattern and destroys the vortex (Johansen et al. 2004; Fu et al. 2014a).

Lopsided discs have been identified in high-resolution observations at (sub-)millimetre wavelengths. The most extreme examples are Oph IRS 48 (van der Marel et al. 2013) and HD 142527 (Casassus et al. 2015b,a) with contrasts > 30. Minor azimuthal asymmetries with contrasts of $\lesssim 10$ appear in DoAr 44, SR 21 and HD 135344B (Pinilla et al. 2015; van der Marel et al. 2016b). While in the case of Oph IRS 48 (van der Marel et al. 2013) millimetre grains appear to be more concentrated in the horseshoe region compared to smaller sizes, in other cases (SR 21 and HD 135344B; Pinilla et al. 2015), dust trapping is not observed, further challenging the vortex scenario. HD 142527 shows a large horseshoe in millimetre continuum emission but whether or not dust trapping occurs is more controversial (Pérez et al. 2015; Muto et al. 2015; Casassus et al. 2015b). We will focus on an alternative mechanism to explain horseshoes in Chapter 11.

10.2 Theoretical models for the gas/dust cavities

The intriguing observational features of transitional discs have led to the development of many models to describe them. Most of the theoretical models developed to date were focused on explaining the sharp cavity in the dust and gas phase, without finding a consistent mechanism to induce a moderate gas accretion.

One of the first invoked mechanism is related to the grain growth processing to explain the low flux level in the inner cavity (Dullemond & Dominik 2005). Observations indicate a decreased optical depth in the inner regions of transitional discs, which do not necessarily indicate a low level of dust density (see Sect. 3.3). Grain growth processing might produce a decrease of the optical depth, since the absorption opacity of large grains is very low, leading to a less efficient emission (Draine 2006). However, Birnstiel et al. (2012) have shown that grain growth produce dips in the SED consistent with those detected in many transitional discs but they fail to reproduce the cavities in the (sub-)millimetre images. The widely accepted mechanisms to explain the cavity-like features of transitional disc are photoevaporation (Clarke 2011; Alexander et al. 2014; Gorti et al. 2015, 2016) and planet-disc interactions (Zhu et al. 2012; de Juan Ovelar et al. 2013).

10.2.0.1 Photoevaporation

By comparing the SEDs of transitional disc with primordial and evolved circumstellar discs, transitional discs might be interpreted as disc in a transition phase from optically thick gas/dust-rich discs to disc-less objects (Owen 2015). Since the majority of transitional discs shows a lack of NIR excess but show a FIR excess, there is a strong indication that disc dispersal proceeds

from inside out (Owen 2015). The idea of disc dispersal from the inside out have been introduced to explain some recent observational results. It was found that the IR excess typically declines on a timescale of $\lesssim 3-5$ Myr (Haisch et al. 2001; Hernández et al. 2007; Mamajek 2009; Ribas et al. 2014). Importantly, the accretion rates approximately decline with age on nearly the same timescale (Fedele et al. 2010; Mendigutía et al. 2012), as expected from viscous accretion theory (Pringle 1981; Hartmann 2009). A recent survey by the Herschel Space Observatory shows that the disc dispersal timescale at outer radii might be longer that what inferred by the star accretion rate on the central star. Dent et al. (2013) found that almost all the gas in the region between $\sim 5-200$ au in the disc sample are dispersed by $\sim 10-20$ Myr, while a recent survey by the Spitzer Science Telescope derived a dispersal timescale of the order of about $5-30$ Myr at radii $\sim 1-40$ au, based on non-detections of H_2 (Pascucci et al. 2006). Since the typical dispersal of gas at larger radii is roughly longer than what expected by the star accretion rate, i.e. $\lesssim 3-5$ Myr, the hypothesis of disc dispersal from inside out can be considered a consistent scenario. For what concern the dust, Ribas et al. (2014) found that the mid-IR excesses are systematically more frequent at longer wavelengths, deriving a timescale of $\sim 4-6$ Myr at $22-24\,\mu$m and $\sim 2-3$ Myr at $3-12\,\mu$m, consistent with inside out dust disc clearing (Alexander et al. 2014; Gorti et al. 2015, 2016). Therefore, rather than homogeneously dispersing by pure viscosity evolution (Lynden-Bell & Pringle 1974), discs seem to have a final stage of clearing from the inside out (Ercolano et al. 2011; Koepferl et al. 2013).

Photoevaporation is one of the most investigated mechanism to trigger disc dispersal from inside out. The mechanism is based on the interaction between the high-energy irradiation coming from the central star and the gas in the disc inner regions. If the irradiation energy and flux are high enough to increase the thermal energy of the gas over the binding energy to the tidal potential of the star, the gas will be removed through a pressure supported wind, carrying micron sized dust grains (e.g. Hollenbach et al. 1994; Hutchison et al. 2016). This mechanism is likely to affect the gas structure at the intermediate and late stage of disc evolution, when the infall rate driven by viscous torque decreases under the photoevaporative wind mass loss rate. Importantly, Clarke et al. (2001) have shown that the combined effect of the viscous driven infall flows and photoevaporative wind eventually lead to the formation of a inner region clear of gas and small dust grains with a radius dependent on the temperature and the star mass (Alexander et al. 2006a,b; Owen 2015). The photoevaporation mechanism is therefore a threshold process: once the mass infall rate induced by viscous torque decreases under the photoevaporative rate value (numerical simulations shows that this occurs after ~ 3.5 yr, Owen et al. 2011a), the photoevaporative flow removes gas and small dust grains in the region below a critical radius (Alexander et al. 2006a,b), leading to the formation of a dust and gas cavity. However, although photoevaporation can explain the dust and gas cavity in the inner regions, it fails in reproducing the high accretion rate of transitional discs, since the typical accretion rate in disc models with photoevaporation is $\lesssim 10^{-8}\,M_\odot\,yr^{-1}$ (Owen et al. 2011b).

10.2.0.2 Planetary or sub-stellar companion

The alternative widely accepted mechanism to carve a transitional disc-like cavity in both dust and gas is related to the presence of an (or multiple) embedded protoplanet(s) or sub-stellar companion(s) orbiting around the central protostar. As described in Chapter 7, the tidal interaction between the planet and the gas/dusty disc leads, if circumstances allow, to the formation of a gap around the planetary orbit.

Importantly, several planet candidates have been found in transitional discs, such as CoKu Tau/4 (Ireland & Kraus 2008), T Cha (Huélamo et al. 2011), LkCa 15 (Kraus & Ireland 2012), HD 100546 (Quanz et al. 2013, 2015), HD 142527 (Biller et al. 2012; Lacour et al. 2016, see also Sect. 11.3.1.5) and HD 169142 (Reggiani et al. 2014). In some cases, it is possible to infer constraints on the separation and the mass of the companion (see Sect. 11.3.1). Given the challenges in detecting young planets or sub-stellar companion in discs due to their low brightness compared to their host stars and their proximity to the star, it is convenient to

theoretically analyze the observational signatures of embedded perturber in discs and verify if they are consistent with what has been observed to date in transitional discs.

First of all, the gap formation in the gas structure leads to the creation of a gas surface density maximum outside the orbit of the planet, at a radial location where the viscous inflow of material is approximately balanced by the expulsion of material due to the planet's tidal torque (see Sect. 7.3). Thus, at the gap outer edge, large dust particles are expected to be trapped in the pressure maximum on a timescale dependent on their grain sizes and local gas properties (see Chapter 2). An additional property of transitional discs that might credit the planetary hypothesis is the fact that they are characterised by an accretion rate onto the central star similar to what inferred in typical T-Tauri discs. In this context, planet are expected to heavily accrete the surrounding gas (Artymowicz & Lubow 1994; Lubow et al. 1999; Bryden et al. 1999; Bate et al. 2003; Papaloizou & Nelson 2005; Lubow & D'Angelo 2006) with a rate depending on the planet mass and disc properties. Lubow & D'Angelo (2006) found that the accretion efficiency of a non-migrating low mass ($\lesssim 1 M_{\rm J}$) planet, i.e. the ratio between the accretion rate onto the planet and the one onto the star, varies from a low value of 3 for the highest and lowest mass planets considered, to a high value of 12 at an intermediate mass. Therefore, although the majority of accretion matter occurs onto the planet, around 10% of of the accretion flow continues past the orbit of the planet and onto the central star, leading to a decrease of the accretion rate on the central star by a order of magnitude, consistent with the measured accretion rate in transitional discs. However, the exact value of the reduction in accretion rate is closely dependent on the mass of the planet and the local gas disc conditions such as viscosity and disc aspect ratio (Papaloizou & Nelson 2005; Lubow & D'Angelo 2006; Ragusa et al. 2016).

As previously mentioned, the high accretion rate measured in transitional discs indicates the presence of gas inside the dust hole. Contrary to large grains which are expected to be trapped at the gap outer edge, smaller grains will be carried with the gas across the gap (Paardekooper & Mellema 2004, 2006; Rice et al. 2006a; Zhu et al. 2012). This spatially-sorted distribution is expected to produce observational signatures in good agreement with the observations of transitional discs over a wide range of disc wavelengths (Dong et al. 2012; Pinilla et al. 2012a; de Juan Ovelar et al. 2013; Follette et al. 2013). However, since small particles dominate the opacity at near-infrared wavelengths, if they are able to follow the gas through the gap and populate the dust cavity, the disc emission would show a strong NIR-excess, resulting in a SED consistent with a primordial disc (Zhu et al. 2012; Owen 2014). One possible way to explain the lack of NIR emission and, at the same time, to allow gas and small dust particles to penetrate into the inner disc is to consider the possibility that multiple planets are present in the inner (from 0.1 au to tens of au) disc regions (Dodson-Robinson & Salyk 2011; Zhu et al. 2011; Duffell & Dong 2015; Dong & Dawson 2016). However, multiple planets produce much shallower and wider cavities than single planets, as recently numerically found by Duffell & Dong (2015). Moreover, the accretion rate onto the star is significantly reduced in this multiple planetary systems by a least two-orders of magnitude with respect to the single-planet case (Zhu et al. 2011). Therefore, very massive planets are necessary to explain large density depletion inside the cavity and an alternative mechanism needs to be invoked to explain the high accretion rate.

Furthermore, the mechanisms of photoevaporation and planet formation might occur simultaneously. In this context, Rosotti et al. (2013) found that the giant planet formation could accelerate disc dispersal and the interplay with the X-ray photoevaporation may be responsible for the origin of transitional discs with large holes and mass accretion rates (see also Rosotti et al. 2015). Detecting the amount of gas accretion and/or gas within dust inner holes can help to distinguish between inner holes related to photoevaporation and those opened by the dynamical clearing driven by embedded protoplanets.

Chapter 11

On the Origin of Horseshoes in Transitional Discs

Based on the paper by Enrico Ragusa, ***Giovanni Dipierro****, Giuseppe Lodato, Guillaume Laibe and Daniel J. Price: "On the origin of horseshoes in transitional discs" published in Monthly Notices of the Royal Astronomical Society,* ***464****, 1449-1455 (2017).*

In this Chapter, we investigate an alternative explanation for the development of non-axisymmetric gas and dust structures in transitional discs (see Chapter 10). Recent numerical studies of discs around black hole binaries (Farris et al. 2014; D'Orazio et al. 2016; Ragusa et al. 2016) have shown that, for mass ratios $q \gtrsim 0.04$, the wide cavity around the primary object carved by the companion becomes eccentric and develops a strong overdensity at the cavity edge, orbiting at the local Keplerian frequency. This arises naturally even in relatively viscous discs in the presence of a sufficiently massive companion. In the protostellar case, Ataiee et al. (2013) showed with 2D hydrodynamic simulations that, for lower mass ratios ($\sim 10^{-3}$), the asymmetries at the cavity edge are weaker than in the vortex scenario, resulting in ring-like rather than horseshoe morphologies. It is therefore timely to explore the case with higher mass ratio to determine whether the horseshoe-like density features revealed by ALMA observations might be explained by the presence of a massive companion inside the cavity. We explore this hypothesis using global, 3D smoothed particle hydrodynamics (SPH) simulations of gas and dust evolution in a circumbinary disc, where the binary consists of a young star and either a massive planet or low mass stellar companion. We demonstrate the formation of crescent-like structures with emissivity contrast up to ~ 10, sufficient to explain many of the 'dust horseshoes' and other non-axisymmetric features observed in transitional discs.

The idea of transitional discs as circumbinary discs has been recently explored by Ruíz-Rodríguez et al. (2016). They found that the spectral energy distributions of $\sim 40\%$ of transitional discs in their sample can be explained as being produced by the flux emission of discs orbiting around binary systems.

11.1 Methods

We perform a set of 3D gas and dust numerical simulations of a circumbinary disc surrounding a binary object (to be interpreted both as star-star or star-planet system), using the PHANTOM SPH code (Lodato & Price 2010; Price & Federrath 2010; Price 2012). The binary is represented by two sink particles (e.g. Bate et al. 1995; Nixon et al. 2013) that exert the gravitational force on

each other and on the gas particles. The sink particles are free to move under the backreaction force of the gas on them, allowing the overall conservation of the binary-disc angular momentum, migration and eccentricity evolution of the binary. The sink particles are also allowed to accrete gas particles when they cross the sink radius and satisfy several dynamical conditions assuring that they are not able to escape from the gravitational field (Bate et al. 1995). We model the gas-dust interaction using the one fluid model SPH formulation developed by Laibe & Price (2014) and Price & Laibe (2015), assuming small grains (i.e. a Stokes number St < 1) such that the terminal velocity approximation is valid. We set the dust grain size to be $s = 1\,\mathrm{mm}$ and vary the mass ratio of the binary. Pressure is computed using a locally isothermal equation of state assuming a radial power-law temperature profile. We exploit the SPH artificial viscosity to model the physical processes responsible for the angular momentum transfer throughout the disc. With reference to the notation used in Lodato & Price (2010), we set the artificial viscosity parameter $\alpha^{\mathrm{AV}} = 0.1$ that corresponds, given our initial conditions, to a Shakura & Sunyaev (1973) α_{SS} parameter ranging between $0.01 \lesssim \alpha_{\mathrm{SS}} \lesssim 0.04$ across the disc. To prevent particle interpenetration we set the parameter $\beta = 2$ as prescribed in Price (2012). SPH artificial viscosity provides also a natural way to reproduce turbulent diffusion of the gas (Arena & Gonzalez 2013) which is transmitted to the dust by the drag. Each simulation is evolved for 140 binary orbits, corresponding to a physical time of $\sim 5600\,\mathrm{yr}$, which is long enough to allow the dust to settle from its initial displacement and to reach quasi-stationarity in the disc shape.

11.1.1 Initial conditions

Our initial setup consists of a sink particle binary surrounded by a disc of 2×10^6 SPH particles. The binary has a total mass $M_{\mathrm{tot}} = 2.2\,M_\odot$ (note, however, that the dynamics are only sensitive to the mass ratio, not the absolute mass), a binary separation $a = 15\,\mathrm{au}$ and an orbital eccentricity $e = 0$. We performed a set of four simulations varying the mass ratio $q = \{0.01; 0.05; 0.10; 0.20\}$. The disc extends between an inner radius $R_{\mathrm{in}} = 18\,\mathrm{au}$ and an outer radius $R_{\mathrm{out}} = 100\,\mathrm{au}$, centred on the centre of mass of the binary. The surface density distribution is $\Sigma = \Sigma_0 R^{-p}$, where R is the radial coordinate in the disc, $p = 0.5$ and Σ_0 determines the total disc mass, $M_{\mathrm{disc}} = (1+\epsilon)M_{\mathrm{g,disc}}$, where $\epsilon = 10^{-3}$ is the millimetre dust-to-gas ratio (corresponding to a total dust-to-gas ratio of 0.01, for our assumed grain size distribution, see Section 11.1.2), and $M_{\mathrm{g,disc}} = 0.05\,M_\odot$ is the gas disc mass. Particles are distributed vertically according to a Gaussian distribution with thickness $H = c_{\mathrm{s}}/\Omega_{\mathrm{k}}$, where c_{s} is the gas sound speed and $\Omega_{\mathrm{k}} = \sqrt{GM_{\mathrm{tot}}/R^3}$. We assume that $H/R \propto R^{0.25}$ and that $H/R = 0.05$ at $R = 18$ au. The dust-to-gas ratio is $\epsilon = 10^{-3}$ throughout the entire disc, implying that the dust has initially the same vertical structure as the gas. After a few orbits of the secondary, the dust has settled from its initial displacement forming a layer with thickness $H_{\mathrm{d}} = H\sqrt{\alpha_{\mathrm{SS}}/\mathrm{St}} \sim 0.7\,H$, consistently with the Dubrulle et al. (1995) model. The velocity of each particle follows a Keplerian profile centered on the binary centre of mass, with orbital velocities corrected to take into account of the radial pressure gradient. The average vertical resolution of this setup can be expressed as $\langle h/H \rangle \sim 0.2$, where h is the SPH smoothing length Price (2012). Since, for our parameter choice, $H_{\mathrm{d}} \sim H$, the disc remains vertically well resolved both in the gas and in the dust.

11.1.2 Simulated ALMA observations

We performed mock ALMA observations of our models using the RADMC-3D Monte Carlo radiative transfer code (Dullemond 2012) together with the Common Astronomy Software Application (CASA) ALMA simulator (version 4.5.3), focusing on ALMA band 7 (continuum emission at 345 GHz). The source of radiation is assumed to be the central star, located at the centre of the coordinate system, with $M_\star = 2M_\odot$, $T_{\mathrm{eff}} = 5500$ K and $R_\star = 2R_\odot$. Dust opacities were produced using the routine[1] developed by Woitke et al. (2016) adopting the dust model from Min et al. (2016). Since ALMA band 7 images essentially trace millimetre particles with

[1] `https://dianaproject.wp.st-andrews.ac.uk/data-results-downloads/fortran-package`

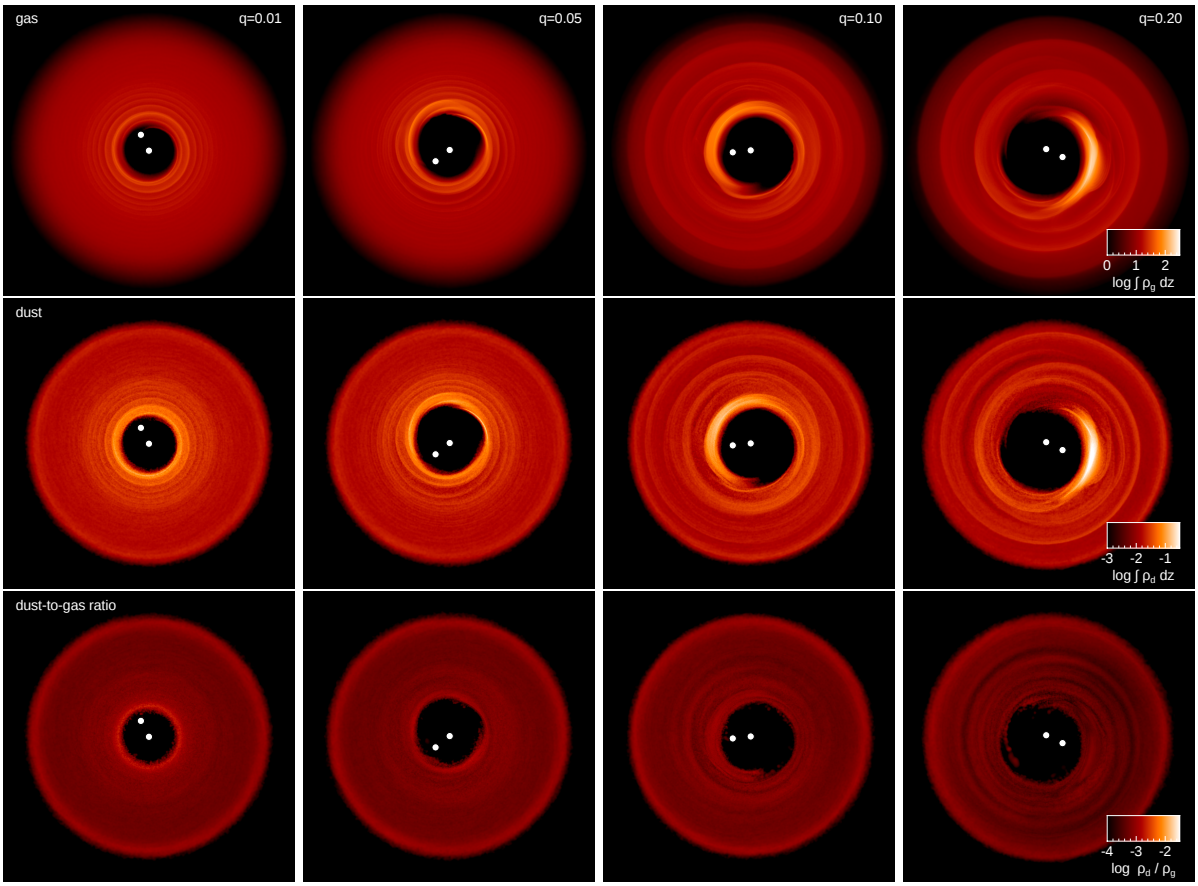

Figure 11.1: Gas (top row) and dust (middle row) surface density in units of g/cm^2 in logarithmic scale after 140 binary orbits for four different binary mass ratios: $q = \{0.01, 0.05, 0.1, 0.2\}$ (left to right, respectively). High mass ratio binaries drive the formation of a large eccentric cavity leading to non-axisymmetric overdensities in both gas and dust ($q \gtrsim 0.05$; right columns). Low binary mass ratios, by contrast, produce more axisymmetric overdensities around a smaller central cavity ($q \lesssim 0.05$; left columns). The bottom row shows the column averaged dust-to-gas ratio in logarithmic scale for the different mass ratios. Note that, for the millimetre size particles we simulate, no dust trapping occurs in the overdense region. Simulated observations of these calculations are shown in Figure 11.2.

a maximum size of $\sim 3\lambda$ (Draine 2006, see Chapter 3), we assumed a dust population with a power-law grain size distribution given by $n(s) \propto s^{-m}$ between $s_{\rm min} = 0.1$ mm to $s_{\rm max} = 3$ mm, with $m = 3.5$. Starting from the 3D density distribution of millimetre grains of our model, we computed the spatial densities of grains in this size range by scaling the dust mass for each grain sizes according to the assumed size distribution, with a total dust mass in the size range $[0.1\,\mu{\rm m}, 10\,{\rm cm}]$ equal to 0.01 of the gas mass.

We computed full-resolution images using 10^8 photon packages. These images were then used as input sky models to simulate realistic ALMA observations taking into account the thermal noise from the receivers and the atmosphere and assuming a perfect calibration of the visibility measurements. We assumed that all the sources were located in Ophiucus star-forming region ($d \sim$130 pc), observed with a transit duration of 3 minutes. We assumed Cycle 4 ALMA capabilities adopting an antenna configuration that provides a beam of 0.12×0.1 arcsec ($\sim 16 \times 13$ au).

11.2 Results

Fig. 11.1 shows the surface density after 140 binary orbits in the gas (top panels) and dust (middle panels) for four different disc models with increasing binary mass ratio ($q = 0.01, 0.05, 0.1$ and 0.2; left to right, respectively). As expected, the cavity size increases with the mass ratio

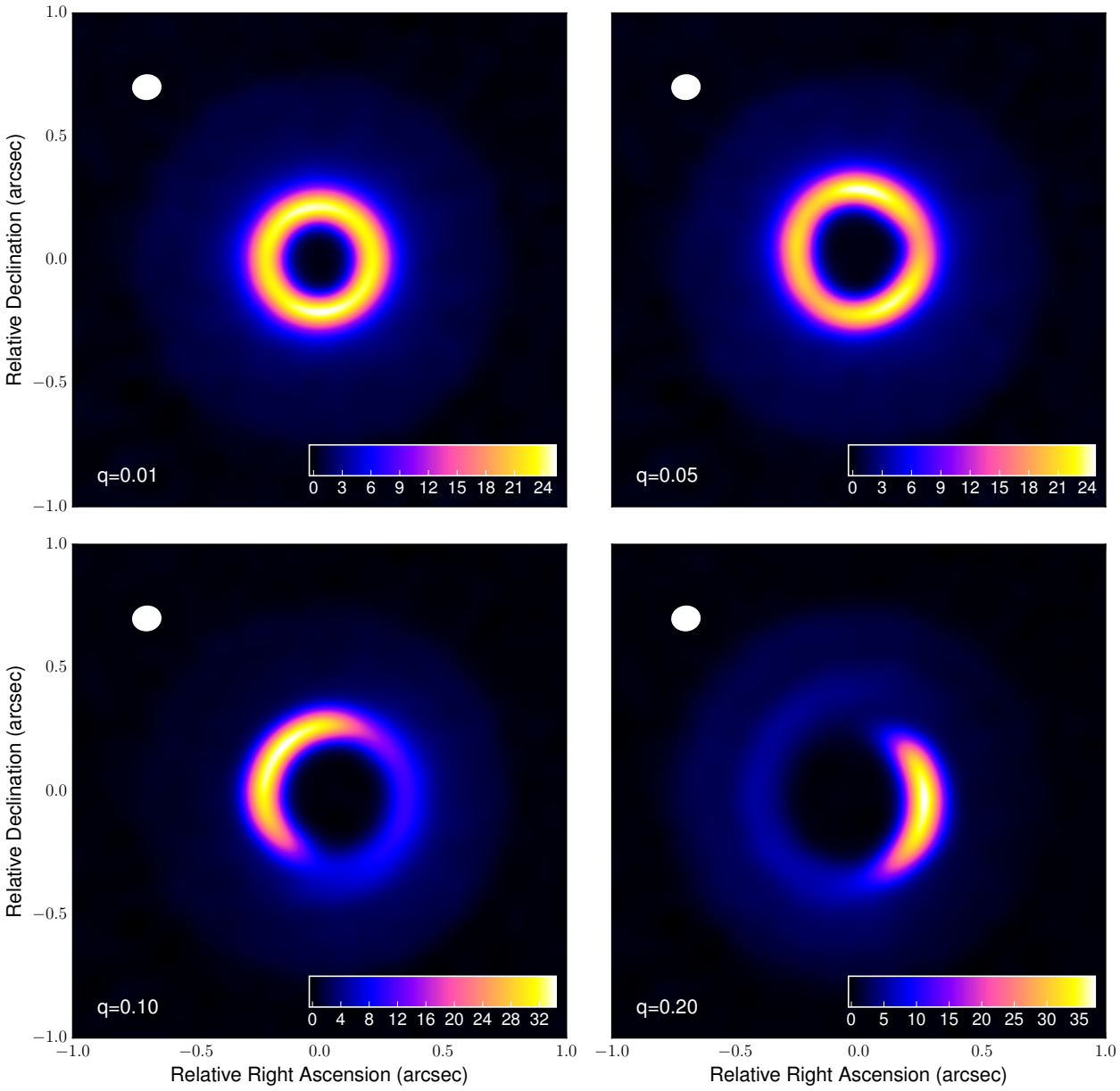

Figure 11.2: Comparison of ALMA simulated observations at 345 GHz of disc models with a mass ratio $q = 0.01$ (upper left), $q = 0.05$ (upper right), $q = 0.1$ (bottom left) and $q = 0.2$ (bottom right). Intensities are in mJy beam^{-1}. The white colour in the filled ellipse in the upper left corner indicates the size of the half-power contour of the synthesized beam: 0.12×0.1 arcsec ($\sim 16 \times 13$ au at 130 pc.).

(e.g. Artymowicz & Lubow 1994). The orbital eccentricity of the gas at the cavity edge also increases with q, reaching $e = 0.1$ for our highest mass ratio case ($q = 0.2$; last right column).

For lower mass ratios ($q = 0.01$ and $q = 0.05$; left two columns), the dust and gas density distribution are more axisymmetric, showing a ring-like overdensity at the cavity edge. For $q \gtrsim 0.05$, an asymmetric crescent-shaped overdensity develops at the cavity edge, with surface densities up to a factor ~ 10 denser than the surrounding gas, consistent with previous numerical simulations in the context of black hole binaries (D'Orazio et al. 2013; Farris et al. 2014; Shi et al. 2012; Ragusa et al. 2016). The overdensity is a Lagrangian feature that rotates with the local orbital frequency. For fixed mass ratio, the level of contrast in the surface density across the crescent-shaped region is similar in both in the gas and in the dust. This is due to the fact that the high gas density in the lump produces a strong aerodynamical coupling between the gas and the dust in the disc. Interestingly, the sharpness of the region increases with increasing mass ratio.

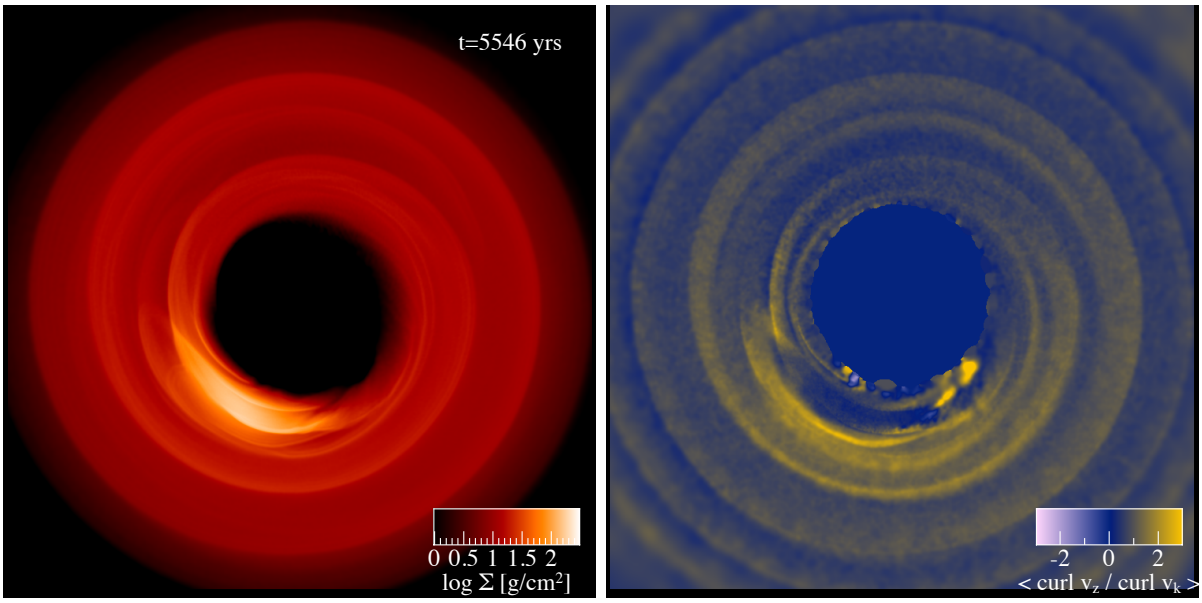

Figure 11.3: Map of the vorticity $\boldsymbol{\omega} = \nabla \times \mathbf{v}$ (right panel) in the region of the overdensity, scaled to the Keplerian value $\boldsymbol{\omega}_\mathrm{K} = \nabla \times \mathbf{v}_\mathrm{K}$, for $q = 0.2$. The left panel shows the gas density structure. There is no evidence of vortices associated with the overdense region (cf. Owen & Kollmeier 2016).

Fig. 11.2 shows mock ALMA images of our disc models at band 7 for the four different mass ratios. The simulated ALMA images reflect the density structures observed in Fig. 11.1. In particular, a crescent or 'dust horseshoe' is evident for $q > 0.05$, with the contrast increasing with increasing mass ratio: for $q = 0.1$ the typical contrast is ≈ 5, while for $q = 0.2$ we obtain a contrast ≈ 7. For $q = 0.05$, the ALMA image shows a double-lobed feature with a low contrast ~ 1.5, similar to those observed in SR21 or DoAr 44 (van der Marel et al. 2016b). For $q = 0.01$ a ring-like structure can be observed, as observed e.g. in Sz 91 (Canovas et al. 2016).

The right panel of Fig. 11.3 shows a snapshot of the vorticity $\boldsymbol{\omega} = \nabla \times \mathbf{v}$, scaled to the Keplerian value $\boldsymbol{\omega}_\mathrm{K} = \nabla \times \mathbf{v}_\mathrm{K}$, where v_K is the Keplerian velocity field. The flow is close to Keplerian in the outer regions of the disc, while in the overdense region the value of the vorticity is $0 \lesssim \boldsymbol{\omega}/\boldsymbol{\omega}_\mathrm{K} < 1$. The extended region outside the overdense crescent where $\boldsymbol{\omega} > \boldsymbol{\omega}_\mathrm{K}$ is due to the steeper than Keplerian gradient of the azimuthal velocity. Vortices induced by the Rossby wave instability typically result in much higher vorticities, with anti-cyclonic vortices reaching $|\boldsymbol{\omega}/\boldsymbol{\omega}_\mathrm{K}| \sim 2$ (Owen & Kollmeier 2016).

11.3 Discussion

The idea that large scale asymmetries might be due to a planetary companion was explored by Ataiee et al. (2013), who concluded that planetary mass objects only produce ring-like features in the disc, in contrast to the observed horseshoe. However, we have shown the dynamics induced in the disc by low and high mass companions is markedly different. It is known that low-mass companions, with $q \sim 10^{-3}$ can produce eccentric cavities, that precess slowly around the star-planet system (Papaloizou et al. 2001; Kley & Dirksen 2006). In contrast, more massive companions, with $q \gtrsim 0.04$ (Shi et al. 2012; D'Orazio et al. 2016) produce strong non-axisymmetric lumps that orbit at the local Keplerian frequency. We have explored the latter case in this paper. For sufficiently massive companions (binary mass ratio $q = 0.2$) we obtain an azimuthal contrast of the order of ~ 10 in mm-wave map, with the contrast an increasing function of the binary mass ratio.

The mechanism causing the formation of the gas overdensity at the cavity edge is still unclear (Shi et al. 2012), but is thought to be related to shocks in the gas at the cavity edge, arising from the intersection of gas flows within the cavity. Thus, it might be expected that the chemistry

would be affected by shocks. Processes such as desorption of various chemical species from the surface of disc dust grains and gas-phase chemical reactions due to shocks occurring in the cavity wall, produce clear chemical signatures of the disc dynamics (see e.g. Ilee et al. 2011 in the case of shocks induced by gravitational instabilities) which may be detected using ALMA. Additionally, shocks might induce the emission of forbidden lines, the detection of which would confirm eccentric cavities as the origin of these structures.

As previously mentioned, horseshoes in transitional discs are often assumed to be due to a vortex induced by a low-mass companion in the cavity (see Sect. 10.1). In this case, a variation of the azimuthal extents of the horseshoe at different wavelengths is expected. Indeed, models of dust trapping by a vortex predict that larger grains would be more azimuthally concentrated in the centre of the vortex. However, in some cases (SR21 and HD 135344B Pinilla et al. 2015) smaller grains appear to be more trapped than larger grains. Since we compute the dynamics of a single species of dust, we cannot predict if our model would reproduce this scenario.

11.3.1 Comparison with observed systems

Non-axisymmetric features have been observed in a handful of transitional discs (see Table 11.1). We report here below a brief summary about the cavity features and the current evidence for the presence of massive companions in these systems.

For what concern the upper-limits on the mass of putative companions, the most accurate results have been obtained applying the aperture masking interferometric observations and speckle imaging in the near-IR waveband. It is worth noticing that detecting planets through imaging is challenging due to the proximity of planets to the central star and their low contrast ratio in emission compared to the brightness of their host star. Additionally, massive companions might have eccentric or misaligned orbits with respect to the disc. This implies that the size of the cavity they are able to carve can be much larger than ~ 2 times the separation at which they are located in the imaging due to projection effects or orbital phase (a planet might not be resolved at the pericentre of its orbit, while resolved and thus detectable at the apocentre).

11.3.1.1 HD 135344B

Recent observations of line and continuum emission from HD 135344B (also known as SAO 206462) evidenced the presence of a cavity both in the gas ($\sim$ 30 au) and in the dust ($\sim$ 40 au) (van der Marel et al. 2016b). The continuum emission shows also a well defined crescent shaped overdense feature at the cavity edge with a mild contrast (van der Marel et al. 2016b). A spiral structure has also been detected in the near-IR scattered light, constituting a strong indication of the presence of a massive companion (see Sect. 4.3.2.1, Garufi et al. 2013). Using the "locally optimized combination of images" (LOCI) technique, in order to be able to possibly resolve and locate the exact position of the companion, Vicente et al. (2011) put an upper-limit of $M_c \sim 230 M_J$ at separations $a \lesssim 14$ au and $M_c \sim 85\, M_J$ at $a \lesssim 37$ au.

Given the central star estimated mass $M_\star \sim 1.7 M_\odot$, the upper-limits on the secondary mass imply mass ratios $q \lesssim 0.05$ at $a \lesssim 37$ au and up to $q \sim 0.13$ for separations $a \lesssim 14$ au. This is consistent with our models, since the crescent-like feature with contrast $\lesssim 10$ observed in HD 135344B (van der Marel et al. 2016a) is similar to what we obtain for our $q = 0.1$ case (bottom left panel of Fig. 11.2).

11.3.1.2 SR 21

The continuum emission from this system shows different asymmetric features at different wavelengths: a crescent shaped overdense feature at 690 GHz (Pérez et al. 2014), and a double-lobed structure at 345 GHz (van der Marel et al. 2016b). In both cases the contrast is mild ($\lesssim 10$). Modeling the dust emission van der Marel et al. (2016b) inferred a cavity edge in the dust at $\sim$ 25 au, while the gas cavity appears to be much smaller ($\sim$ 7 au, Pontoppidan et al. 2008).

The presence of a warm companion surrounded by a cloud of accreting gas in this system was invoked by Eisner et al. (2009) to explain an excess in the near-IR and mid-IR SEDs, which could be explained by an additional warm ($\sim 700\,\mathrm{K}$) black body emission from an extended region of $40\,R_\odot$. The total luminosity produced by the companion in this framework appears to be consistent with a T-Tauri star with mass $M_\mathrm{c} \sim 0.2\,M_\odot$ enveloped in a gaseous cloud (Follette et al. 2013). Using the angular differential imaging technique to increase the image resolution, Follette et al. (2013) was able to rule out the presence of secondary stellar object for separations $a \gtrsim 18$ au; based on the contrast sensitivity achieved by Follette et al. (2013), Wright et al. (2015) constrained the upper-limit on the companion mass to $\sim 40 - 60\,M_\mathrm{J}$ at separations $a \gtrsim 18$ au, implying that such a stellar source needs to be located at separations $a \lesssim 18$ au. This implies possible secondary-to-primary mass ratios of $q \sim 0.1$ for $a \lesssim 18$ au and $q \sim 0.03$ for $a \gtrsim 18$ au. The double-lobed structures detected in the continuum emission in SR 21 (e.g. Pinilla et al. 2015) at 345 GHz are consistent with our models with $q \sim 0.05$, which is in agreement with the detection limits reported in literature.

11.3.1.3 IRS 48

With an azimuthal contrast of $\gtrsim 130$ of the peak emission compared to the background disc, IRS 48 represents the source with the strongest crescent-shaped dust structure in our sample. This dramatic azimuthal range at the cavity wall observed in IRS 48 is best described by the segregation of millimetre grain sizes induced by an azimuthal bump in pressure (van der Marel et al. 2013). The dust cavity was found to be extended 60 au from the central star (Bruderer et al. 2014). The continuum asymmetry has been modeled as a major dust trap, triggered by the presence of a substellar companion with a mass of $9M_\mathrm{J}$ (Zhu & Stone 2014). This would seem very unlikely to correspond to an equally sharp gas distribution. Alternatively, Wright et al. (2015), based on the detection limit reported in Ratzka et al. (2005), rule out a potential companion with a mass $\gtrsim 100\,M_\mathrm{J}$, which correspond to a mass ratio of $q \sim 0.05$ at a radial separation of 19 au. As expected, this result does not appear consistent with the mechanism presented in this paper.

11.3.1.4 DoAr 44

DoAr 44 (also known as ROX 44 and Haro 1-16) is the source with the mildest contrast in (sub-) mm continuum emission in our sample (van der Marel et al. 2016b). This source has been classified as a pre-transitional disc with a dust cavity between 2 and 32 au (Espaillat et al. 2010; van der Marel et al. 2016b). Based on the companion detection limits reported in Ratzka et al. (2005), Wright et al. (2015) derived that a potential companion should have a mass $\lesssim 80\,M_\mathrm{J}$ at a separation $\lesssim 12.5$ au which correspond to a mass ratio of $\lesssim 0.06$, adopting the star properties reported in Espaillat et al. (2010). The values of the mass ratio inferred from observations appear to be consistent with our scenario: the double-lobed structures observed in DoAr 44 can be explained for $q \sim 0.05$ according to our model.

11.3.1.5 HD 142527

HD 142527 harbours a disc with a wide dust cavity extending from 10 au to 120 au. At the cavity inner wall, the (sub-) millimetre dust continuum emission show a bright horseshoe with contrast ∼30. The existence of a massive close companion with $q \approx 0.16$ has been established for HD 142527 using the sparse masking aperture technique (Biller et al. 2012; Lacour et al. 2016). This is particularly interesting since the contrast in HD 142527 of ~ 30 is within a factor of 3 of the contrast we find in our highest mass ratio. However, this companion is also inclined by $\sim 70^\circ$ with respect to the disc (Lacour et al. 2016) and this case might be further complicated by a strong warp (Casassus et al. 2015a). Intriguingly, Casassus et al. (2015b) comment that "*the large sub-mm crescent [in HD 142527] mostly reflects the gas background, with relatively inefficient trapping, so that the observed contrast ratio of ∼ 30 is accounted for with a contrast*

Name	Contrast	Dust trapping	Companion	Consistency
HD 135344B	$\lesssim 10$	No	Strong indication	Yes
SR 21	$\lesssim 10$	No	Indication	Yes
DoAr 44	$\lesssim 10$	?	?	Yes
IRS 48	$\gtrsim 100$	Yes	?	No
HD 142527	~ 30	cm grains?	Yes	Yes
Lk Hα 330	$\lesssim 10$	?	Indication	Yes

Table 11.1: Summary of transition discs displaying horseshoe or other non-axisymmetric features. For each source, we indicate the observed contrast in mm images, whether there is evidence for dust trapping in the crescent, and whether the system is known to host a massive companion. The last column indicates whether the observed structures are consistent with our model, given the upper-limits on the companion mass as reported in the literature.

of 20 in the gas", consistent with our model. Seemingly this conflicts with Muto et al. (2015) who found variations of $\sim$ 10-30 in the dust-to-gas ratio. This difference may be explained by uncertainties in grain surface chemistry, in particular whether or not a fraction of CO is depleted on dust grains (Casassus 2016).

11.3.1.6 Lk Hα 330

Lk Hα 330 is characterised by a millimetre dust cavity with a size of about 40 au and an azimuthal intensity variation of a factor of two. Recent observations performed by Willson et al. (2016), using the sparse aperture masking technique in the K' near infrared band, on Lk Hα 330 revealed the presence of a possible massive companion characterized by a value of $M_{\rm c}\dot{M}_{\rm c} \sim 10^{-3}\, M_{\rm J}^2\,{\rm yr}^{-1}$ orbiting at a separation $a \sim 37\,$au from the central star. Assuming an accretion rate on the secondary object of $\dot{M}_{\rm c} \lesssim 10^{-8}\,{\rm M}_\odot\,{\rm yr}^{-1}$ implies an mass $M_{\rm c} \gtrsim 100\,M_{\rm J}$ and a mass ratio $q \gtrsim 0.05$. It should be noted that previous works (Brown et al. 2009; Andrews et al. 2011; Isella et al. 2013) had reported private communications that, based on near-IR observations, ruled out the presence of secondary objects with masses $M_{\rm c} \gtrsim 50\,M_{\rm J}$ ($q \gtrsim 0.025$) at separations $M_{\rm c} \gtrsim 10\,$au, indicating how elusive these objects might be. In any event, these estimates for the companion mass appear to be consistent with the low sharpness of the crescent-shaped structure predicted by our model (see the upper panels of Fig. 11.2).

11.4 Conclusions

We performed 3D SPH gas and dust simulations of circumbinary discs surrounding a protostar and either a low mass stellar companion or massive protoplanet. We showed that the companion carves a wide, eccentric cavity in the disc, resulting in a non-axisymmetric gas overdensity at the cavity edge. For sufficiently large binary mass ratios this feature appears as a 'horseshoe' in millimetre wavelength dust continuum images, as observed in several transition discs.

Our model makes testable predictions that can be used to observationally distinguish the eccentric cavity model from the more commonly assumed 'gap edge vortex' model. We identify the following main features differentiating the two processes, which can be used as the basis for observational tests of our hypothesis:

1. Dust and gas kinematics. In our model, the fluid velocity is close to Keplerian, and does not show the large vorticity expected in the vortex model (see Fig. 11.3).

2. Our mechanism applies both in high and in low viscosity discs, while vortices only arise for $\alpha \lesssim 10^{-4}$.

3. Dust horseshoes arise from eccentric cavities only for relatively large mass ratios $q \gtrsim 0.04$, while in principle vortices can arise for lower mass planets. Clearly, establishing whether

a relatively massive companion is present within the cavity is a key observational test of our model.

4. The structures described in this paper only occur at the edge of the central cavity, while vortices can occur in principle at any location within the disc.

5. Less massive companions should produce more axisymmetric structures, potentially explaining the 'dust rings' seen in (e.g.) Sz 91 and DoAr 44. We predict that a higher degree of non-axisymmetry around larger central cavities.

6. Our model does not require azimuthal dust trapping and the observed contrast largely reflects the gas density contrast.

In summary, cavities opened by massive companions are a promising mechanism for explaining rings, lopsided features and horseshoes seen in transition discs.

Chapter 12

Conclusions

In this Thesis I have investigated a problem of great interest in the present star and planet formation research field: the characterization of substructures recently detected in protoplanetary discs. The dramatic advances in observational techniques of protoplanetary discs in the past few years are giving unprecedented insights on some of the complex phenomena occurring in them. Substructures such as gaps, spirals and horseshoes will be especially important in developing a comprehensive model for the evolution of planetary system. At the same time, the general framework for interpreting protoplanetary disc observations developed in the past appears to be crude and over-simplified when applied to modern high resolution observations. Standard assumptions, such as adopting axisymmetric disc models or ignoring the dynamical evolution of solids and their impact on the radiative transfer, can be considered simplistic.

In this Thesis, I have developed a state-of-the-art framework to investigate the observational predictions of a variety of disc models in order to compare the expected disc emission computed by my analysis with realistic observations. The final goal is to understand what physical processes induce the peculiar disc substructures observed to date. The key of my analysis is to investigate the size-sorted 3D density distribution of dust grains in protoplanetary discs with (non-)axisymmetric gas density structures such as spirals, gaps and horseshoes. The numerical approach developed in this Thesis can be summarizes into three main steps (see Fig. 12.1) as follows:

1. The dynamics of dust and gas in protoplanetary discs is computed by means of 3D dust/gas hydrodynamic simulations taking into account the aerodynamical coupling between the gas and dust phase. In this Thesis I consider three different models: gravitationally unstable discs (see Part II), discs with embedded protoplanets (see Part III) and circumbinary discs (see Part IV);

2. Starting from the the 3D spatial distributions of dust grains with different sizes, the expected full-resolution disc emission over a wide range of wavelengths is computed by means of 3D radiative transfer simulations taking into account a realistic set of absorption and scattering dust opacities;

3. I perform simulated observations of the disc emission models taking into account the atmospheric and instrumental noises in order to compare the expected (sub-)millimetre (and, in one case study, near infrared scattering) observations of my disc models with real observations.

Our analysis shows that the size-dependent dynamics of dust grains in the variety of disc models taken into consideration in this Thesis can reproduce the substructures detected in protoplanetary discs. I can summarize my findings as follows:

- In **Part II** I investigate if the occurrence of local pressure maxima induced by gravitational instabilities can trap dust particles and explore the detectability of these inhomogeneities

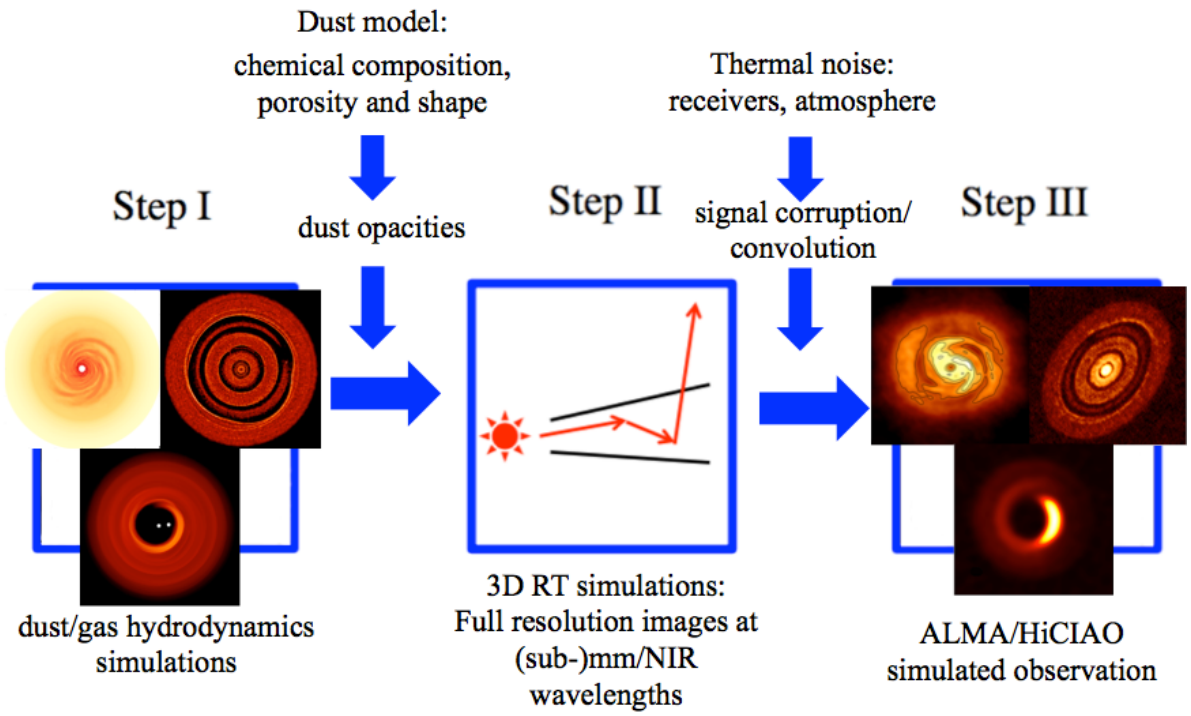

Figure 12.1: A simple schematic of the general workflow of this Thesis.

at near-infrared and (sub-)millimetre wavelengths. The observational predictions of the resulting models show that the resolution capabilities and sensitivity of ALMA and HiCIAO are sufficient to spatially resolve the peculiar spiral structure of gravitationally unstable discs. This work therefore demonstrates that the development of gravitational instability can be considered a valid theoretical explanation of the spiral structure observed to date in protostellar discs.

- In **Part III** I explore the dynamical clearing mechanism induced by forming protoplanets embedded in discs with the aim, on the one hand, of understanding the physics of dust gap opening and, on the other hand, of reproducing the gap-like structures observed in the disc around HL Tau. I provide a necessary and sufficient condition for the minimum planet mass able to carve a dust gap, as well as an estimate of the location of the outer edge of the dust gap, which can be useful to estimate the mass of the planet from high-resolution (sub-)millimetre observations. Importantly, I identify a new mechanism to carve dusty gaps in protoplanetary disc by low massive planets, without necessarily requiring the presence of gas gaps. I apply my findings to the case of HL Tau, showing that the three gaps detected by ALMA can be described by the presence of sub-Jupiter mass planets in the three main gaps observed by ALMA orbiting around the central star.

- Finally, in **Part IV**, I investigate if the horseshoes observed at (sub-)millimetre wavelengths in transitional discs can be explained by the dynamics of gas and dust at the edge of the cavity carved by a binary object. I found that the cavities carved by binaries with large mass ratio becomes eccentric, leading to a horseshoe-like feature in the gas density and, due to the tight coupling between millimetre sized grains and the gas, also in the dust surface density. my simulated observations reveal that low contrast non-axisymmetric structures detected in transitional discs can be explained by the presence of massive companions in these system.

Although a variety of disc models can potentially reproduce the disc substructures observed to date in protoplanetary discs (e.g. Casassus 2016), the analysis performed in this Thesis shows that my models can accurately reproduce the observational results and can be useful for shedding

light on the mechanism responsible for the disc substructures.

12.1 Limitations of the model

The greatest challenge in modern protoplanetary disc field is to develop a single model capable of reproducing high resolution, multi-tracer and multi-wavelength observations of a given protoplanetary disc. In principle, the perfect disc simulation should simultaneously compute the three dimensional global dynamics of the gas and multiple growing and fragmenting dust grain species, coupled to a self-consistent modelling of the disc chemical evolution, as well as the radiation and magnetic fields (Haworth et al. 2016). In this Thesis, I focus on a limited part of the full disc modelling, based on computing the dynamics of gas and a pressureless dust fluid (see Sect. 2.2.1) made of non-growing spherical grains in non-magnetized discs and a post-processing evaluation of the temperature structure and disc emission due to the irradiation coming from the central star. Although I used state of the art dynamical and radiative transfer modelling of protoplanetary discs, my approach has some uncertainties:

- The primary source of uncertainty comes from the crucial assumption at the basis of my approach: I assume that the disc substructures observed at (sub-)millimetre wavelengths are related to variations of the dust density structure across the disc. As described in Section 3.3 any structure detected in optically thin disc emission can be also related to the disc temperature variations. By running isothermal simulations the radiative effect induced by the central star, viscous heating and cooling onto the gas and dust dynamics is not fully captured. This limitation might affect the observational predictions of the disc models (e.g. Isella & Turner 2016).

- A second limitation of the model developed in this Thesis is based on the computation of the evolution of a single species of dust grains. Performing multiple grain species simulations allows to fully capture the aerodynamical coupling between the dust and gas and self-consistently compute the spatial distribution of solids spanning an entire grain size distribution.

- Thirdly, in my simulations and analysis of discs with embedded planets and circumbinary discs, I assume that the global evolution of the gas structure is governed by a viscous-like mechanism of transport of angular momentum and energy across the disc. Although this approach is common in many theoretical studies of discs, in real protoplanetary discs the transport process is believed to be driven by some kind of turbulence phenomena, induced by, for example, magnetorotational and gravitational instabilities (see Sect. 1.3). However, due to the small evolution time scale of my simulations, the details of viscous gas evolution are not expected to heavily affects my results.

- For what concerns the radiative transfer modelling, the main assumption of my approach is to consider a uniform choice for the shape, porosity and chemical composition of dust across all the discs. Although this approach is mostly adopted in literature, a self consistent model of the structural and chemical composition of the dust is required to better compute the observational predictions of my disc models.

- Moreover, the only source of heating in my models is assumed to be due to the irradiation from the central star without taking into account the envelope contribution and the viscous heating. However, neglecting the viscous heating would significantly affect the temperature structure in the inner disc regions, while in the outer disc it can be reasonably assumed that the temperature is governed by the irradiation coming from the central star (see Sect. 1.4).

12.2 Future perspectives

With the completion of ALMA and SPHERE, several survey projects aimed at resolving protoplanetary discs on scales of a few au are underway. In the near future, with the data available from such surveys, there are therefore many opportunities to address some fundamental questions in the field. For example, are the observed structures really related to sub-stellar/planet formation and/or disc instabilities? If the gap-like structures are related to the presence of embedded protoplanets, are the observed structure a cause or a consequence? Alternatively, non-dynamical clearing mechanism be efficient enough to carve sharp gaps/cavities in the spatial distribution of large grains? Addressing these questions requires a consistent numerical and observational approach in investigating the disc substructures and comparing with theoretical models. Taking the limitations of the approach developed in this Thesis into account, my research can be extended into the following directions:

- Theoretical analysis of the physics of gap opening process induced by an embedded protoplanet in the dust density distribution. As recently discussed in Chapter 8, it is reasonably expected that it is easier to carve and sustain a gap in the dust than in the gas. In the light of this, it would be interesting to analytically evaluate a grain size-dependent shape of the dust density radial profile across the gap carved by an embedded planet. This analysis can then be tested by performing numerical simulations which may reproduce the wavelength-dependent extension of dusty gaps revealed in recent observational results.

- Analysis of the influence of the back reaction from the dust to the gas density structure in a variety of disc models. The back reaction from the dust to the gas might play a non-negligible role in affecting the gas distribution across the gap/cavity carved by a (sub-)stellar companion/planet. Moreover, it would be interesting to include its influence in current theoretical models on the gap opening process in the gas and dust distribution that can be tested by performing numerical simulations. Moreover, disc clearing mechanisms such as photoevaporation would be inevitably affected by the back reaction from large dust grains, leading to a significant deviation of the gas surface density profile close to the central star and, consequently, of the mass accretion rate compared to the current models of inside-out disc disc dispersal (Gorti et al. 2016).

- Develop an accurate prescription for grain growth and fragmentation processes and incorporate it into 3D dynamical models of dust and gas evolution, taking into account the transfer of angular momentum between the gas and the multi species growing dust grains. Collisions between grains provide a means of altering the size distribution of grains, both through coagulation and destructive collisions, which are expected to affect the appearance of the disc (Birnstiel et al. 2016).

12.2.1 Towards a global evolutionary model of gas-dust viscous discs

Among the wide range of improvements of the work carried out in my Ph.D., the most intriguing development of my study is to incorporate the global evolution of the gas and dust mixture described in Chapter 2 into a 1D analytical model on the evolution of the gas and dust surface density $\Sigma_{\rm g}(R)$ and $\Sigma_{\rm d}(R)$ as a function of disc cylindrical radius R for a viscous disc, including consistently the exchange of angular momentum between the two phases of the mixture. Assuming a dust-gas mixture in a thin, axisymmetric, non-magnetic, non-self-gravitating, laminar viscous and vertically isothermal disc, the evolution of the surface density of the two phases can be expressed in cylindrical coordinates as a system of two coupled equations of continuity of mass

$$\frac{\partial \Sigma_{\rm g}}{\partial t} = -\frac{1}{R}\frac{\partial}{\partial R}(R\,\Sigma_{\rm g}\,v_{{\rm g},R}), \tag{12.1}$$

$$\frac{\partial \Sigma_{\rm d}}{\partial t} = -\frac{1}{R}\frac{\partial}{\partial R}\left\{\Sigma_{\rm d}\left[R\,v_{{\rm d},R} - D\,\frac{\partial \log\left(\Sigma_{\rm d}/\Sigma_{\rm g}\right)}{\partial \log R}\right]\right\}, \tag{12.2}$$

where the gas and dust radial velocity are given by

$$v_{\mathrm{g},R} = -\frac{1}{1+\epsilon}\left[\frac{\epsilon\,\Delta v}{\mathrm{St}+\mathrm{St}^{-1}} - \left(1+\epsilon\frac{\mathrm{St}^2}{1+\mathrm{St}^2}\right)v_{\mathrm{visc}}\right], \tag{12.3}$$

$$v_{\mathrm{d},R} = \frac{1}{1+\epsilon}\left[\frac{\Delta v}{\mathrm{St}+\mathrm{St}^{-1}} + \frac{v_{\mathrm{visc}}}{1+\mathrm{St}^2}\right], \tag{12.4}$$

with

$$\epsilon = \frac{\Sigma_{\mathrm{d}}}{\Sigma_{\mathrm{g}}}, \qquad \mathrm{St} = \frac{\rho_{\mathrm{p}} a}{\Sigma_{\mathrm{g}}}\frac{\pi}{2}, \qquad D = \frac{\nu}{1+\mathrm{St}^2}, \tag{12.5}$$

and

$$\Delta v \equiv \frac{1}{\Sigma_{\mathrm{g}}\Omega_{\mathrm{k}}}\frac{\partial P}{\partial R}, \qquad v_{\mathrm{visc}} = -\frac{3}{\Sigma_{\mathrm{g}} R^{1/2}}\frac{\partial}{\partial R}\left(\nu\Sigma_{\mathrm{g}} R^{1/2}\right), \tag{12.6}$$

where I have assumed classical shear viscosity (see further details in Chapt. 1 and Sect. 2.3.1). The solutions of Eqs. 12.1 – 12.2 needs to be simultaneously computed in order to account for variation of the aerodynamical coupling regimes and the dust feedback across the disc. However, it is worth remarking that these equations are valid if the dust phase is treated as a passive trace species, i.e. the particles have no influence on the gas motion and spatially coexist with the gas (Morfill 1985; Clarke & Pringle 1988).

These equation might also include the effect of photoevaporation by high energy photons (by adding a mass loss term in Eq. 12.1, Alexander & Armitage 2007; Gorti et al. 2015) in order to figure out the inside out disc dispersal due to the combined effect of photoevaporation and dust back reaction. As described in Chapter 2, for standard disc geometries, i.e. $\partial P/\partial R < 0$, the dust back reaction forces the gas to move towards the regions of low pressure (i.e. outer disc regions), creating a gas cavity around the central star. In addition, the photoevaporative wind opens a cavity in the gas disc, carrying away only small dust grains which are tightly coupled to gas (Alexander & Armitage 2007; Hutchison et al. 2016). Importantly, the decrease of the gas density close to the central star produces an increase of the Stokes number of small particles, leading to an enhancement of the dust population with Stokes number close to unity, further powering the dust back reaction. As recently confirmed in Gorti et al. (2015), photoevaporation removes gas and negligible amounts of dust mass, increasing the dust to gas mass ratio in the disc and affecting further dust evolution (Alexander & Armitage 2007). Therefore, the combined effect of photoevaporation and dust back reaction are expected to act both to promote the inside out disc dispersal at the beginning of the disc evolution. Once the cavity is formed, the dynamics of both phases are expected to be very different. Large grains are expected to drift toward the edge of the gas cavity, leading to an increase of the local dust to gas ratio (Alexander & Armitage 2007). In this case, the dust back reaction tends to supply with gas the cavity by pulling the gas at the edge of the cavity towards the central star, leading to a further photoevaporation. Importantly it is reasonably expected that the accretion rate onto the star will be strongly affected by the dust back reaction. It is therefore timely to figure out the combined impact of the dust back reaction and photoevaporation on the disc dispersal (Alexander et al. 2014; Gorti et al. 2015, 2016).

Appendices

Appendix A

The viscous stress tensor

The viscous stress tensor σ indicates the local momentary forces between neighbouring fluid elements as a result of the velocity differences between the elements. The viscous stress tensor is given by:

$$\sigma = \eta \left(\nabla \mathbf{v} + (\nabla \mathbf{v})^T - \frac{2}{3} (\nabla \cdot \mathbf{v}) \mathbf{I} \right) + \zeta (\nabla \cdot \mathbf{v}) \mathbf{I}, \tag{A.1}$$

where $\mathbf{I}$ is the identity matrix, $(\nabla \mathbf{v})^T$ indicates the transpose of the matrix $\nabla \mathbf{v}$ and where η and ζ are respectively the shear and bulk viscosity coefficients. The first term ($\propto \eta$) is associated with momentum transfer in shear flows, making a zero contribution to the diagonal part of the stress tensor. The second term ($\propto \zeta$) is associated with momentum transfer due to bulk compression of the flow, i.e. $\propto \nabla \cdot \mathbf{v}$, which is proportional to the rate of change of density (see Eq. 1.4). Generally, in a component notation with the repeated index summation convention, the force per unit area acting on a surface is given by $F_i = \sigma_{ij} \hat{s}_j$ where $\hat{s}_j$ is the j-component of the unit normal vector to the surface. Thus, the tensor components σ_{ij} indicate the viscous force in the direction i acting on a surface whose normal is in direction j.

Due to the axisymmetric nature of the system (see Sect. 1.2), it is convenient to derive the fluid equations in cylindrical coordinate.

A.1 Stress tensor for a circular shearing flow in cylindrical coordinates

Let us evaluate the stress tensor in cylindrical coordinates. In Cartesian coordinates, the tensor components are given by

$$\sigma_{ij} = \eta \left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} - \frac{2}{3} \delta_{ij} \frac{\partial v_\mathrm{k}}{\partial x_k} \right) + \zeta \delta_{ij} \frac{\partial v_\mathrm{k}}{\partial x_k}. \tag{A.1}$$

Adopting cylindrical coordinates, for a circular shear flow (i.e., a flow in which the angular velocity is a function of radius only) in the azimuthal direction, the only non-zero component of the stress tensor is $\sigma_{\phi R} = \sigma_{R\phi}$ which, as previously mentioned, indicates the force in the direction ϕ acting on a surface whose normal is in direction R. Importantly, the components σ_{RR}, $\sigma_{\phi\phi}$ and σ_{zz} can be considered negligible because they represent normal stress associated to the bulk viscosity[1] of the fluid. In addition, for the component $\sigma_{z\phi}$ and σ_{zR}, the thin-disc approximation ensures us to neglect the shear terms involving the z-component because it is expected that the azimuthal and radial velocity is invariant under translations in z.

Let us calculate the $R\phi$ component of the stress tensor. We firstly notice that the divergence of a vector in two-dimensional cylindrical coordinates is given by

$$\nabla \cdot \mathbf{v} = \frac{1}{R} \frac{\partial}{\partial R} (R v_{\mathrm{g},R}) + \frac{1}{R} \frac{\partial v_{\mathrm{g},\phi}}{\partial \phi} + \frac{\partial v_{\mathrm{g},z}}{\partial z}. \tag{A.2}$$

[1] It is worth pointing out that the bulk viscosity is important in phenomena involving compression and rarefaction such as, for instance, shocks. Anyway, the stress tensor terms related to the bulk viscosity can be neglected in comparison to the shear component $\sigma_{R\phi}$.

According to the basic assumptions to describe a infinitesimally thin accretion discs ($v_{\mathrm{g},R} \ll v_{\mathrm{g},\phi}$ and $v_{\mathrm{g},\phi} = v_{\mathrm{g},\phi}(R)$), the velocity has a vanishing divergence. Moreover, it is expected that the vertical component of the velocity can be neglected. This allow to express the stress tensor as a function of the strain tensor, ϵ which expresses the rate at which the mean velocity changes in the medium. In Cartesian coordinates the stress tensor is given by

$$\sigma = \eta \left(\nabla \mathbf{v} + (\nabla \mathbf{v})^T \right) = 2\eta\epsilon. \tag{A.3}$$

In Cartesian coordinates, the components of the strain tensor are

$$\epsilon_{ij} = \frac{1}{2} \left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} \right). \tag{A.4}$$

It can be shown that, in the case of shear flows, the strain tensor components relative to a set of (n,t,s) directions that are mutually orthogonal are given by

$$\epsilon_{nn} = \mathbf{e}_n \cdot \frac{\partial \mathbf{v}}{\partial x_n}, \quad \epsilon_{tt} = \mathbf{e}_t \cdot \frac{\partial \mathbf{v}}{\partial x_t}, \quad \epsilon_{ss} = \mathbf{e}_s \cdot \frac{\partial \mathbf{v}}{\partial x_s}, \tag{A.5}$$

$$\epsilon_{nt} = \frac{1}{2} \left(\mathbf{e}_n \cdot \frac{\partial \mathbf{v}}{\partial x_t} + \mathbf{e}_t \cdot \frac{\partial \mathbf{v}}{\partial x_n} \right), \tag{A.6}$$

where $\mathbf{e_n}, \mathbf{e_t}, \mathbf{e_s}$ are the unit vectors identifying the directions. Since these results are valid for any directions n, t and s, we can use these expressions to find the strain components in any other coordinate system. In cylindrical coordinates, i.e. $\partial x_R = \partial R, \partial x_\phi = R\partial\phi$ and $\partial x_z = \partial z$, the $R\phi$ component of the strain tensor is given by

$$\epsilon_{R\phi} = \frac{1}{2} \left(\mathbf{e}_R \cdot \frac{1}{R} \frac{\partial \mathbf{v}}{\partial \phi} + \mathbf{e}_\phi \cdot \frac{\partial \mathbf{v}}{\partial \phi} \right) = \frac{1}{2} \left(\frac{1}{R} \frac{\partial v_{\mathrm{g},R}}{\partial \phi} + R \frac{\partial \left(v_{\mathrm{g},\phi}/R \right)}{\partial R} \right) = \frac{1}{2} R \frac{\partial \Omega}{dR}, \tag{A.7}$$

where we have used

$$\frac{\partial \mathbf{e}_R}{\partial \phi} = \mathbf{e}_\phi \qquad \frac{\partial \mathbf{e}_\phi}{\partial \phi} = -\mathbf{e}_R, \tag{A.8}$$

and $\Omega = v_{\mathrm{g},\phi}/R$ is the angular velocity. Therefore, the stress tensor is given by

$$\sigma_{R\phi} = \eta R \frac{\partial \Omega}{dR}. \tag{A.9}$$

Since the Navier-Stokes equation includes the divergence of the stress tensor, it can be shown that the three components of $\nabla \cdot \sigma$ are

$$\nabla \cdot \sigma|_R = 0, \quad \nabla \cdot \sigma|_\phi = \left(\frac{\partial}{\partial R} + \frac{2}{R} \right) \sigma_{R\phi}, \quad \nabla \cdot \sigma|_z = 0. \tag{A.10}$$

Therefore, the viscous forces acts only in the azimuthal direction, leading to a transport of the angular momentum.

Appendix B

Interferometry and Aperture Synthesis

The detection of disc substructures of circumstellar discs plays a crucial role for understanding of their physical and chemical evolution during the initial stage of star formation. To this end, it is necessary to resolve scales of less than a few tens of au for discs located in different star forming regions. The resolution requirements to spatially resolve disc substructures of very young circumstellar discs in the range of distance of the better-known star forming regions, are unachievable for a single dish telescope at the moment.

It is well known that the resolution of a perfect optical system was defined by Lord Rayleigh according to the ability of an optical system to distinguish two separate point sources depending on the overlap of their Airy discs produced by diffraction. The angular resolution of a perfect optical system, ignoring the atmospheric disturbances, is defined to be: $\theta \sim 1.22\,\lambda/D$ where λ is the wavelength and D is the diameter of the optical system. In the astronomical context, it is necessary to account for disturbances induced by atmosphere which perturb the images of stars as seen through a telescope. However, ignoring the atmospheric disturbances, the angular resolution of a telescope is determined solely by its diffraction limit. Therefore, in order to achieve higher resolutions, one has to either decrease the observing wavelength or increase the diameter of the telescope further. The first option has led to a tendency for telescopes to operate at centimetre and millimetre wavelengths. However, it is worth remarking that the choice of the observation waveband is obviously deeply linked to the phenomena one wants to investigate. In addition, another way to increase the resolution is to build telescopes with apertures that are hundreds of kilometers in size. The largest single dish telescope at the moment has a diameter of about 100 meters which correspond in terms of resolution at radio wavelength, e.g. 7 mm, at $\sim 17''$. Therefore, the diffraction limit of this telescope is insufficient to spatially resolve protoplanetary discs which have typical angular scales of few arcseconds even in the nearest star forming region. Thus, it is even more difficult to spatially resolve substructures of a circumstellar disc.

The purpose of astronomical interferometry is to achieve high-resolution observations using a cost-effective cluster of comparatively small telescopes rather than a single very expensive monolithic telescope. This technique is one of the most used for optical, infrared, submillimetre and radio astronomy. It allows to produce, through a complex technique called *aperture synthesis* (see Sect. B.2), images with the same angular resolution as a single telescope with an aperture size comparable to the longest baseline used for interferometry. It is worth pointing out that this technique is well-established in the Radio regime while in infrared waveband significant progresses has been made through the use of several interferometers such as the Very Large Telescope Interferometer (VLTI). As regards the optical and far infrared waveband, this technique is still in the experimental stage.

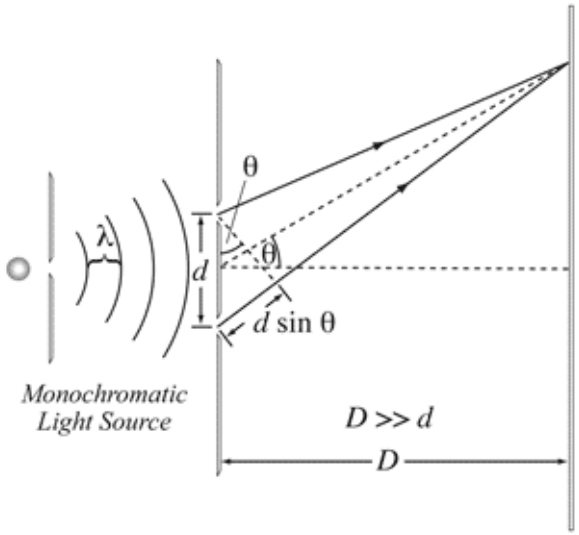

Figure B.1: A simple schematic of the Young's double slit experiment.

B.1 Interferometry

B.1.1 Young's double slit experiment

There is a deep analogy between the Young's double slit experiment with monochromatic light and a two element interferometer. In fact, generically, an interferometer is a device for producing interference between two or more waves. A simple schematic of the experiment is shown in Fig. B.1. It consists in a single point light source which emits radiation in front of a pair of closely spaced narrow slits separated by a distance d. The radiation passing through the slits is observed on a screen located at a distance D.

Let us assume that the point-like source produces a monochromatic (single-wavelength λ) wave. In addition, let us assume that the radiation emitted by the source hits the two slits at the same time. Thus, as the peaks of the wave arrive at the two slits simultaneously, the two waves passing through the slits are then said to be in phase with one another. The wavefronts emerging from the slits overlap and interfere in the region beyond. Assuming that the distance D is large: $D \gg d$, the paths of light from two slits to a point on the measurement screen can be considered nearly parallel. Thus, as it can be seen in Fig. B.1, the light from the two slits arrives at the screen with a relative phase difference $\phi = k\, d \sin\theta$. If this phase difference is an odd multiple of π, the two waves are out of phase and produce destructive interference,

$$k\, d \sin\theta = (2n+1)\pi \quad n = 0, 1, 2, \ldots, . \tag{B.1}$$

Otherwise, the two waves produce constructive interference when the phase is an even multiple of π,

$$k\, d \sin\theta = 2n\pi \quad n = 0, 1, 2, \ldots, . \tag{B.2}$$

Therefore, the interference pattern formed on the measurement screen consists of bright lines, where the light emerging from the two slits constructively interferes, alternating with dark lines where the light destructively interferes. In detail, it is worth pointing out that rather than sharp bright and dark lines, there is a sinusoidal modulation of the light intensity on the screen.

The Young's interference experiment played a crucial role in the origin of the development of the interferometric techniques in astronomy.

B.1.2 The antenna beam pattern

In order to outline a complete description of the response of an interferometer for an extended source, it is necessary to investigate also the antenna performance. Let $I(j, n)$ be the two-

dimensional intensity of the source where (j, n) are generic sky coordinates, with n measured parallel to the aperture spacing vector and n normal to it. Given the positions of the two mirrors, the interferometer is sensitive solely to the intensity projected onto the j direction while in the orthogonal direction the response is simply proportional to the intensity integrated over solid angle. Therefore, the one dimensional profile $I_j\,(j)$ measured by the interferometer is given by

$$I_j\,(j) = \int I\,(j, n)\; dn. \tag{B.3}$$

In order to evaluate the response of a single antenna or an interferometer to a source, we need to investigate how the signal coming from an extended source is captured by the antenna. The geometrical characterization of an antenna is specified by the so-called *radiation pattern*. An antenna *radiation pattern* or *antenna pattern* is defined as a mathematical function or a graphical representation of the radiation properties of the antenna as a function of space coordinates. In the left panel of Fig. B.2 is represented a convenient set of coordinates to represent the three-dimensional spatial distribution of radiated[1] energy. Various parts of a radiation pattern are referred to as lobes, which may be subclassified into major (or main) and minor lobes. The major lobe (also called main beam) is defined as the radiation lobe containing the direction of maximum radiation. Minor lobes are any lobes apart from the major lobe and they represents radiation in undesired directions such as the ground and the telescope itself. The radiation picked up in the antenna minor lobes is commonly referred to as *spillover radiation* and it should be minimized in order to increase the directivity of an antenna.

Associated with the antenna pattern is a parameter designated as *beamwidth*. The beamwidth of a radiation pattern is defined as the angular separation between two identical points on opposite sides of the pattern maximum. The beamwidth is also used to describe the resolution capabilities of the antenna to distinguish between two adjacent radiating sources. The most common resolution criterion states that the resolution capability of an antenna to distinguish between two sources, is equal to the *Full Width Half Maximum* ($FWHM$) defined by the angle between the two directions in which the radiation intensity is half of the value of the patter maximum.

Therefore, the response of a single antenna or an interferometer to a source has to take into account both the intensity distribution of the source and the radiation pattern of the antenna. Mathematically, the response of an interferometer is expressed by a convolution of the last mentioned terms. In the right panel of Fig. B.2 is shown the main lobe of the radiation pattern of an antenna $A\,(\theta)$ and the one-dimensional intensity profile of a source $I_j\,(\theta')$ in which θ' is measured with respect to the center of the source θ_0.

Thus, the total output power from the antenna, omitting the constant factors, is proportional to the convolution of the source intensity with the radiation pattern of the antenna

$$F(\theta) \propto \int_{\text{source}} A\left(\theta' - \theta\right) I_j\left(\theta'\right)\; d\theta' = A\,(\theta) * I_j\,(\theta)\,, \tag{B.4}$$

where it is used the relation $A(\theta' - \theta) = A(\theta - \theta')$ which is generally valid for a radiation pattern.

B.1.3 Two-dimensional source synthesis

The major part of radio astronomy antennas operates at frequencies above 1 GHz because, in this way, it is possible to reach high angular resolutions. The antenna beams is therefore narrower as frequency increases and the source appears extended in the sky. Let us consider how a two element interferometer response to the radiation coming from a bidimensional source.

[1] A fundamental property of antennas states that the receiving pattern of an antenna when used for receiving is identical to the far-field radiation pattern of the antenna when used for transmitting. This is a consequence of the reciprocity theorem of electromagnetism. Therefore the antenna can be viewed as either transmitting or receiving.

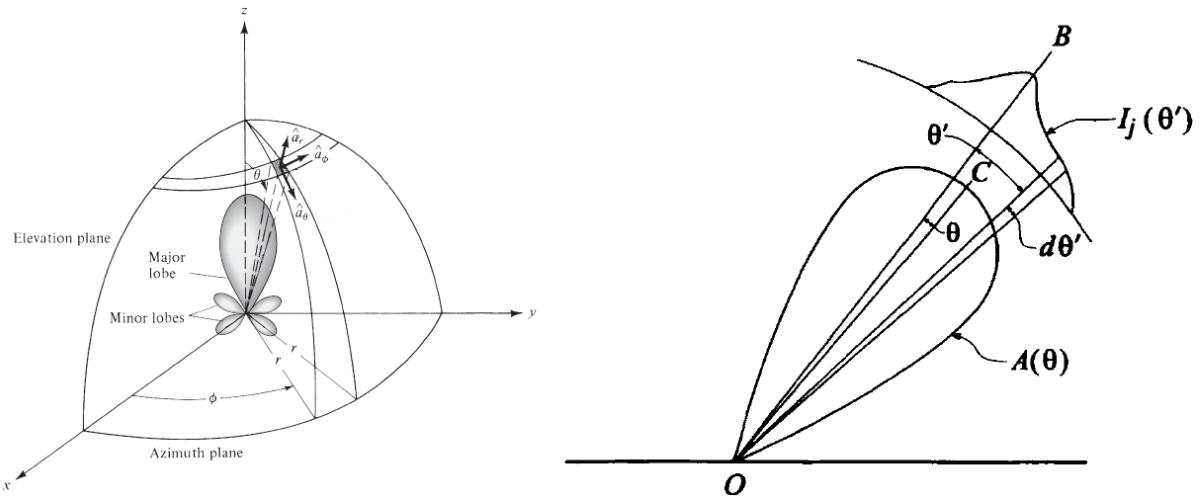

Figure B.2: (left) Coordinate system for antenna analysis. Image adapted from Balanis (2008). (right) The main lobe of the radiation pattern $A(\theta)$ of an antenna and the intensity profile of a source $I_j(\theta')$. The angle θ is measured with respect to the beam center OC and θ' is measured with respect to the direction of the reference position of the source 0B. Image taken from Thompson et al. (2004).

In Fig. B.3 is shown the geometry of the radiation received from a small element of the extended source in the direction $\mathbf{s}$, where $\mathbf{s_0}$ indicates the phase reference position. An element of the source of solid angle $d\Omega$ at position $\mathbf{s} = \mathbf{s_0} + \boldsymbol{\sigma}$ contributes a component of power $\propto A(\boldsymbol{\sigma})I(\boldsymbol{\sigma})\,d\Omega$ where $A(\boldsymbol{\sigma})$ is the pattern of each antenna, $I(\boldsymbol{\sigma})$ is the source intensity distribution (see Eq. B.3). In the two-dimensional source synthesis, the visibility function is defined as

$$V = \int_{4\pi} A_N(\boldsymbol{\sigma})\, I(\boldsymbol{\sigma})\; e^{-i2\pi \mathbf{D}_\lambda \cdot \boldsymbol{\sigma}}\, d\Omega, \tag{B.5}$$

where $\mathbf{D}_\lambda$ is the baseline vector (see Fig. B.3).

At this point, it is convenient to introduce a coordinate system on the sky and for the baseline. The geometry of the coordinate system typically chosen for interferometric observations is illustrated in Fig. B.4. The baseline vector $\mathbf{D}_\lambda$ is measured in wavelengths at the center frequency of the observing band. The two antennas are assumed to be identical, but if they differ, $A_N(\boldsymbol{\sigma})$ can be considered as the geometric mean of the beam patterns of the two antennas. As regards the baseline, let us introduce a right-handed coordinate system with components (u, v, w) where u and v are measured in a plane normal to the direction of the phase reference position. The spacing component v is measured toward the north while u toward the east. The component w is measured in the direction of the phase reference position $\mathbf{s_0}$.

In two-dimensional analysis the components l and m of the coordinate system on the sky, are defined as the cosines of the angles between the direction (l, m) and the u and v axes respectively, as shown in Fig. B.4. With this choice of coordinates, some terms in Eq. B.5 can be rewritten in the following way

$$\mathbf{D}_\lambda \cdot \mathbf{s_0} = w \qquad \mathbf{D}_\lambda \cdot \mathbf{s} = ul + vm + w\sqrt{1 - l^2 - m^2}, \tag{B.6}$$

$$d\Omega = \frac{dl\, dm}{\sqrt{1 - l^2 - m^2}}. \tag{B.7}$$

Noting that $\mathbf{D}_\lambda \cdot \boldsymbol{\sigma} = \mathbf{D}_\lambda \cdot \mathbf{s} - \mathbf{D}_\lambda \cdot \mathbf{s_0}$, the visibility function expressed by the Eq. B.5 can be rewritten in the following manner

$$V(u, v, w) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} A_N(l, m)\, I(l, m)\; e^{-i2\pi\left[ul + vm + w\left(\sqrt{1-l^2-m^2}-1\right)\right]} \frac{dl\, dm}{\sqrt{1 - l^2 - m^2}}. \tag{B.8}$$

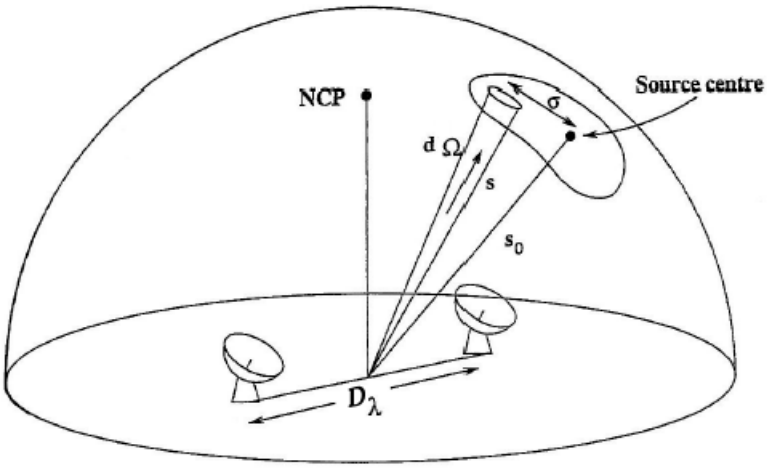

Figure B.3: Geometry of the radiation received from a small element of the extended source subtending a solid angle $d\Omega$ in the direction $\mathbf{s}$ where $\mathbf{s_0}$ indicates the phase reference position.

Moreover, it is possible to rewrite the last equation using an approximation valid so long as the synthesized field is not too large. If l and m are small enough that the term

$$w\left(\sqrt{1-l^2-m^2}-1\right) \simeq -\frac{1}{2}\left(l^2+m^2\right)w, \tag{B.9}$$

can be neglected, the visibility function is given by

$$V\left(u,v,w\right) \simeq V\left(u,v,0\right) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} A_N\left(l,m\right) I\left(l,m\right)\, e^{-i2\pi(ul+vm)} \frac{dl\, dm}{\sqrt{1-l^2-m^2}}. \tag{B.10}$$

With this approximation, it is usual to omit the w dependence and write the visibility as the two-dimensional function $V\left(u,v\right)$. Therefore, for a coplanar array or in the case where a coordinate system can be chosen such as $w=0$, the above relation turns out to be a 2-D Fourier relation. Note that the factor $\sqrt{1-l^2-m^2}$ in Eq. B.10 can be subsumed into the function $A_N\left(l,m\right)$. Equation B.10 is a form of the van Cittert-Zernike theorem which was originally derived in optical context. Therefore, the instantaneous output of the interferometer is proportional to the visibility function which represents the Fourier transform of the source intensity distribution multiplied by the beam pattern.

The coordinates (u,v) represent the couple of spatial frequencies in which the interferometer is sensitive at a given time. In detail, the points at small (u,v) record large scale structure of the source and viceversa. Thus, each instantaneous baselines identifies one point in the system of coordinates (u_0,v_0) where it is possible to measure amplitude and phase of the visibility function $V\left(u_0,v_0\right)$. Thus, for example, one can imagine measuring all required Fourier components with just two antennas, one of which is mobile, by laboriously moving the second antenna from place to place.

However, it is worth pointing out that, as the earth rotates, the baseline projection on the celestial sphere rotates. The variation of the length and direction of the projected baseline is represented by an arc of an ellipse in the u-v plane. The parameters of the ellipse depend on the declination of the source, the length and orientation of the baseline and the observation duration. This technique of using the Earth's rotation to improve the "u-v coverage" was traditionally called *Earth Rotation Synthesis*.

The u-v coverage plays a crucial role for the imaging (and image fidelity) capabilities of an interferometer because it affects the beam size associated to the antenna pattern. In addition,

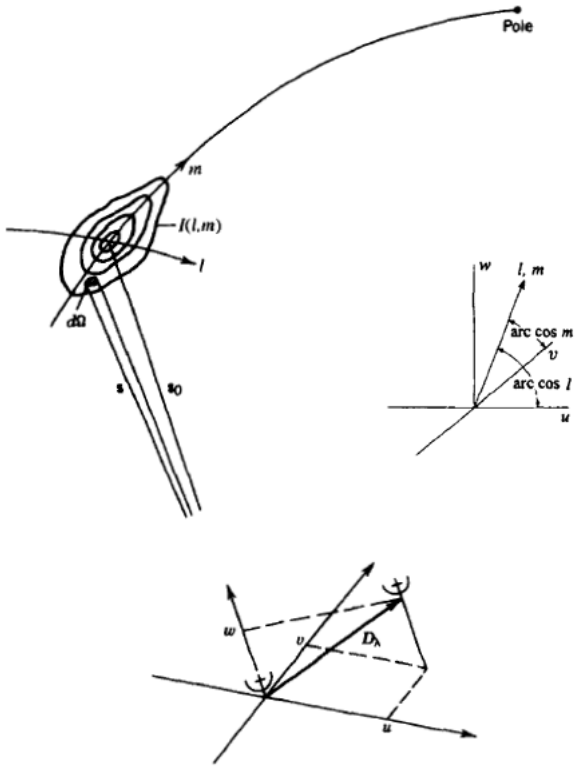

Figure B.4: Geometric relationship between a source under observation $I(l, m)$ and a two-element interferometer. The antenna baseline vector D_λ, measured in wavelengths, has components (u, v, w). Image adapted from Thompson et al. (2004).

as remarked above, the resolution capabilities of an interferometer is determined by the length of the baseline projected onto a plane normal to the direction of the source. Thus, the imaging capabilities of an interferometer for the same antenna configurations and at a given wavelength, can vary with different observation duration and declination of the source. In Fig. B.5 is shown the spatial frequency coverage at 345 GHz for a closed ALMA configuration of 50 antenna with different observation duration for a source located at two different declinations: $\delta = -24°$ and $\delta = +30°$.

Therefore, to summarize, the particular spacing and orientation of all the different baseline[2] at a given time, can be described in terms of the fringe visibilities and the Fourier transform of the brightness distribution in the sky multiplied by the antenna radiation pattern.

B.2 Aperture synthesis

In this section we will outline the techniques aimed at obtaining high-resolution images from interferometric observations of a variety of astronomical objects. In order to obtain a map of the observed source, it is necessary to perform the inverse Fourier transform of the visibility measurements in a proper way. All the complex techniques regarding the creation of the sky image from interferometric data is called *Aperture Synthesis*.

The most straightforward method of obtaining an intensity distribution from measured visibility data is by direct Fourier transformation. Previous results showed that each point in the u-v plane samples one component of the Fourier transform of the brightness distribution multiplied by the beam pattern. From Eq. B.10, it is convenient to define the following function

[2]For a multi-elements interferometer with N antennas, the number of baselines at a given time is $N(N-1)/2$.

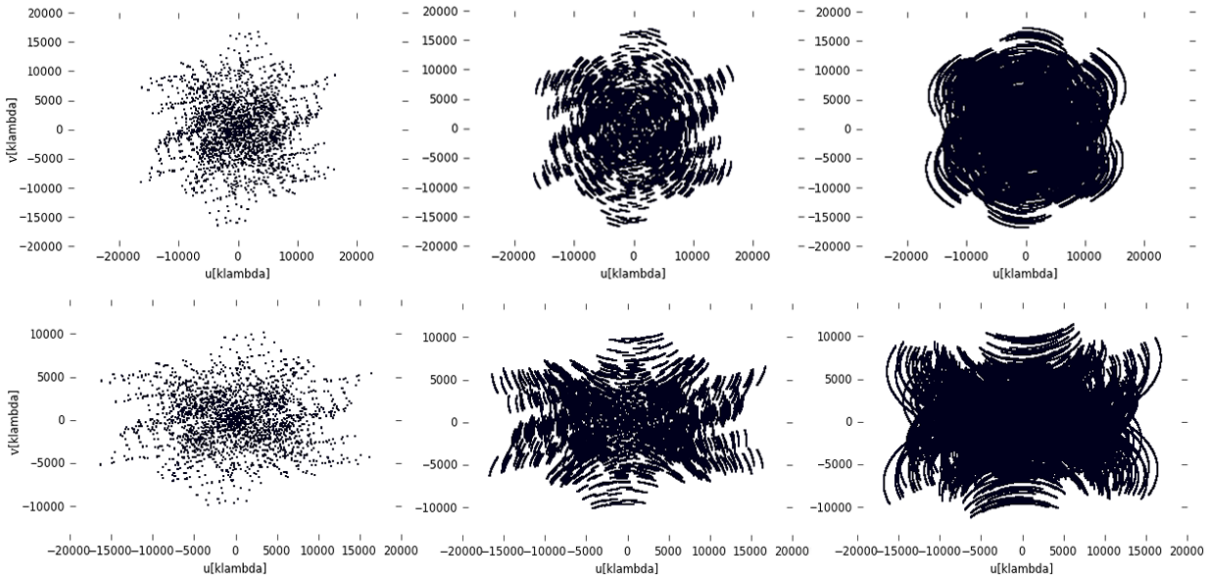

Figure B.5: Spatial frequency coverage at 345 GHz for a closed ALMA configuration of 50 antennas with observation duration of: one minute (left), one hour (centre) and four hours (right) for a source located at two different declinations: (top) $\delta = -24^\circ$ and (bottom) $\delta = +30^\circ$.

$$I^A(l,m) = A_N(l,m)\, I(l,m) = \mathcal{F}^{-1}[V(u,v)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} V(u,v)\; e^{i2\pi(ul+vm)} du\, dv. \quad \text{(B.11)}$$

For a real observation by an interferometer, the completely sampled visibility function is not available. Typically, the sky measurements consist of values of real and imaginary part of the visibility function at discrete locations in the Fourier plane. Moreover, the recorded data are not uniformly distributed across the u-v plane and are mostly concentrated towards the center (see Fig. B.5). Incomplete coverage of the u-v plane gives rise to a loss of information on the source. This means that the number of different images consistent with the data is infinite, since it is possible in principle to fill in the unmeasured parts of the u-v plane in an infinite number of ways. In addition, since the Fourier transform cannot be applied on the raw data points produced by the interferometer, it is necessary to introduce a transfer function which creates a regular gridding in the u-v plane. The gridding is needed to construct a regular matrix of measurements to do the Fourier transform.

Furthermore, in order to reduce effects due to the irregular coverage of the u-v plane, it is convenient to multiply the visibility measurements by a weighting function aimed to reach an effective uniform density of data in the u-v plane and reduce sidelobes caused by the application of the transfer function.

Thus, as already mentioned, the incompleteness of the u-v coverage is represented by introducing a sampling function $S(u,v)$ given by

$$S(u,v) = \sum_{k=1}^{M} w_k\, \delta(u-u_k)\, \delta(v-v_k), \quad \text{(B.12)}$$

where w_k is the weighting applied to the visibility measurements in the u-v plane and $M = N(N-1)/2$ is the number of baselines for N antennas at a given time. Thus, introducing the sampling function $S(u,v)$ into the righthand side of Eq. B.11, the inverse Fourier transform of the ensemble of visibilities is given by

$$I^D(l,m) = \int\int S(u,v)\, V(u,v)\; e^{i2\pi(ul+vm)} du\, dv = \mathcal{F}^{-1}[S(u,v)\, V(u,v)] =$$
$$= \mathcal{F}^{-1}[S(u,v)] * \mathcal{F}^{-1}[V(u,v)] = B(l,m) * I^A(l,m), \quad \text{(B.13)}$$

where $I^D(l,m)$ is called *Dirty Image* and $B(l,m)$ is the *Dirty* (or *Synthesized*) *Beam*. It is also called the *Point Spread Function (PSF)*.

At this point, the next step to solve for $I^A(l,m)$, i.e., the source intensity distribution modified by the beam pattern, is the deconvolution starting from the knowledge of the dirty image $I^D(l,m)$ and the synthesized beam $B(l,m)$ that can be deduced knowing the geometry of the array and the location of the source in the sky.

B.2.1 Cleaning algorithms

The image deconvolution is usually performed using the so-called CLEAN algorithm. This simple algorithm allows synthesis imaging of complex objects even with relatively poor Fourier plane coverage occurring, for instance, with partial earth rotation synthesis or with arrays composed of small numbers of antennas.

One of the most successful deconvolution procedures is the algorithm CLEAN introduced by (Högbom 1974). This is basically a numerical deconvolving process applied in the (l,m) domain. This CLEAN algorithm finds a solution to the convolution equation by positing a model for the true sky brightness composed by a collection of point sources. The principal steps are as follows:

1. Compute the discrete inverse Fourier transforms of visibility and sampling functions to obtain the dirty image I^D and the dirty beam B;

2. Localize the strongest intensity pixel in the dirty image and record the position and amplitude by inserting a delta-function component into a model that will become the cleaned map:
$$I^C = \sum_{q=1}^{n} I_q \, \delta(l - l_q) \, \delta(m - m_q) \, ; \tag{B.14}$$

3. Calculate the values for I_q, l_q and m_q through the search of the peak in the residual image defined by:
$$I^R = I^D - \gamma B I^C, \tag{B.15}$$
where γ is a parameter called *loop gain* which controls the convergence of the method;

4. Update of the model image with the values of amplitudes and peak positions already precomputed;

5. If the strongest feature in the residual image is larger than some threshold, return to step 3 and repeat the procedure iteratively until all significant source structure has been removed from the map. Otherwise, if the strongest feature is below threshold or if the number of iterations $N_{\rm iter}$ is too large, compute the final image I^A. The latter is obtained by convolution of the the final model I^C with a "clean" beam function B^C which is usually chosen to be a Gaussian with a half-amplitude width equal to that of the peak of the original dirty beam B. The output image is given by adding the residuals into the clean-beam map. This step is summarized by the following equation:
$$I^A = I^C * B^C + I^R. \tag{B.16}$$

The key aspect of the Högbom CLEAN algorithm is the way in which it solves iteratively for the positions and strengths of the CLEAN components. However, the Högbom algorithm spends a lot of time searching residuals also in regions where they are negligible. However, performance of the deconvolution can be improved by restricting the search in a "CLEAN window" wherein the image brightness is known to be non-zero.

Beyond the simple Hogbom CLEAN, there are two variations which perform essentially the same algorithm, but do so in a more efficient manner. These are the Clark and the Cotton-Schwab CLEAN algorithm. The major modification introduced by the two last mentioned algorithm is the introduction of the concept of major and minor cycles.

As regards the Clark algorithm, in the minor cycle, the clean components are searched for, and subtracted from the image using the Högbom CLEAN. The difference is that only the central portion of the dirty beam is used for the subtraction. This algorithm is reasonably fast, because of performing the minor cycle in the image plane. The CLEAN components identified during the minor cycles are removed at once by a Fast Fourier Transforms (FFT) during a major cycle.

The Cotton-Schwab CLEAN algorithm is a variant of the Clark algorithm whereby the major cycle subtracts CLEAN components from the un-gridded visibility data. The noise and gridding errors can thus be removed if the inverse Fourier transform of the CLEAN components to each $u-v$ sample is accurate enough. This results in a considerable improvement in the image quality due to reduced sidelobes and the application of the sample function.

B.2.1.1 Weighting

As already mentioned, a critical factor in the cleaning processes is the weighting to be applied to gridded data in order to take into account that the recorded data are not uniformly distributed across the $u - v$ plane. In detail, when the grid is implemented, in some cells (typically in the center) you will have more points than in other ones. Thus, it is necessary to apply weights to the data (w_k, in Eq. B.12) inside each cell in order to assign a proper weight to the whole cell. There are three different methods to evaluate the weight associated to each cell.

The first and the most common possibility is the so-called *natural weighting*. It favors the signal-to-noise ratio using weighting to data measurements inversely as their variances. If the data are obtained with a uniform array of antennas and receivers and the averaging time is the same for all data, then the variances should all be the same. Thus the maximum signal-to-noise ratio is obtained by including all measurements with the same weight. However, it is worth pointing out that the sampling function is more concentrated in the center of the $u - v$ plane. Thus, if the value of the visibility is computed averaging these measurements and assigning a weight that is computed following the correct statistics (natural weighting), then the error associated to the average measurements in the inner regions of the grid will be much smaller than in the outer regions. So the weight of the inner regions will be higher than the outer regions.

In order to face the last mentioned problem, the second option, called *uniform weighting*, applies equal weights to each cell in the grid, without considering that in some cells you have more measurements or that some measurements have lower errors. This choice increases the resolution but decreases the signal-to-noise. This is due to the fact that equal weights are given to cells with few or many measurements.

The third and intermediate option, called *robust weighting*, allows to sweep across these two methods thanks to a parameter that varies setting the best combination of weighting, given the sampling of the source on the *u*-*v* plane.

Bibliography

Adams, F. C. & Shu, F. H. 1986, ApJ, 308, 836

Aikawa, Y., van Zadelhoff, G. J., van Dishoeck, E. F., & Herbst, E. 2002, A&A, 386, 622

Akiyama, E., Muto, T., Kusakabe, N., et al. 2015, ApJ, 802, L17

Alexander, R., Pascucci, I., Andrews, S., Armitage, P., & Cieza, L. 2014, Protostars and Planets VI, 475

Alexander, R. D. & Armitage, P. J. 2007, MNRAS, 375, 500

Alexander, R. D., Clarke, C. J., & Pringle, J. E. 2006a, MNRAS, 369, 216

Alexander, R. D., Clarke, C. J., & Pringle, J. E. 2006b, MNRAS, 369, 229

ALMA Partnership et al. 2015, ApJ, 808, L3

André, P. 2002, in EAS Publications Series, Vol. 3, EAS Publications Series, ed. J. Bouvier & J.-P. Zahn, 1–38

Andrews, S. M. 2015, ArXiv e-prints

Andrews, S. M. & Williams, J. P. 2007a, ApJ, 671, 1800

Andrews, S. M. & Williams, J. P. 2007b, ApJ, 659, 705

Andrews, S. M., Wilner, D. J., Espaillat, C., et al. 2011, ApJ, 732, 42

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., Hughes, A. M., et al. 2012, ApJ, 744, 162

Andrews, S. M., Wilner, D. J., Zhu, Z., et al. 2016, ApJ, 820, L40

Ansdell, M., Williams, J. P., van der Marel, N., et al. 2016, ApJ, 828, 46

Arena, S. E. & Gonzalez, J.-F. 2013, MNRAS, 433, 98

Armitage, P. J. 2010, Astrophysics of Planet Formation, 294

Armitage, P. J. 2015, ArXiv e-prints

Armitage, P. J., Livio, M., Lubow, S. H., & Pringle, J. E. 2002, MNRAS, 334, 248

Armitage, P. J., Livio, M., & Pringle, J. E. 2001, MNRAS, 324, 705

Armstrong, D. J., Osborn, H. P., Brown, D. J. A., et al. 2014, MNRAS, 444, 1873

Artymowicz, P. 1993, ApJ, 419, 166

Artymowicz, P. & Lubow, S. H. 1994, ApJ, 421, 651

Ataiee, S., Pinilla, P., Zsom, A., et al. 2013, A&A, 553, L3

Avenhaus, H., Quanz, S. P., Schmid, H. M., et al. 2014, ApJ, 781, 87

Ayliffe, B. A., Laibe, G., Price, D. J., & Bate, M. R. 2012, MNRAS, 423, 1450

Bai, X.-N. 2011, ApJ, 739, 50

Bai, X.-N. & Stone, J. M. 2010, ApJ, 722, 1437

Bai, X.-N. & Stone, J. M. 2011, ApJ, 736, 144

Balanis, C. A. 2008, Modern Antenna Handbook (Wiley)

Balbus, S. A. & Hawley, J. F. 1991, ApJ, 376, 214

Balbus, S. A. & Hawley, J. F. 1998, Reviews of Modern Physics, 70, 1

Balbus, S. A. & Papaloizou, J. C. B. 1999, ApJ, 521, 650

Banzatti, A. & Pontoppidan, K. M. 2015, ApJ, 809, 167
Banzatti, A., Testi, L., Isella, A., et al. 2011, A&A, 525, A12
Barge, P. & Sommeria, J. 1995, A&A, 295, L1
Baruteau, C., Bai, X., Mordasini, C., & Mollière, P. 2016, Space Sci. Rev.
Baruteau, C., Crida, A., Paardekooper, S.-J., et al. 2014, Protostars and Planets VI, 667
Basri, G. & Bertout, C. 1989, ApJ, 341, 340
Bastien, P. & Menard, F. 1988, ApJ, 326, 334
Bate, M. R., Bonnell, I. A., & Price, N. M. 1995, MNRAS, 277, 362
Bate, M. R., Lubow, S. H., Ogilvie, G. I., & Miller, K. A. 2003, MNRAS, 341, 213
Beckwith, S. V. W., Sargent, A. I., Chini, R. S., & Guesten, R. 1990, AJ, 99, 924
Bell, K. R. & Lin, D. N. C. 1994, ApJ, 427, 987
Benisty, M., Juhasz, A., Boccaletti, A., et al. 2015, A&A, 578, L6
Bergin, E. A., Cleeves, L. I., Gorti, U., et al. 2013, Nature, 493, 644
Bertin, G. 2014, Dynamics of Galaxies
Bertin, G. & Lodato, G. 1999, A&A, 350, 694
Béthune, W., Lesur, G., & Ferreira, J. 2016, A&A, 589, A87
Beuzit, J.-L. 2013, European Planetary Science Congress 2013, held 8-13 September in London, 8, EPSC2013
Biller, B., Lacour, S., Juhász, A., et al. 2012, ApJ, 753, L38
Binney, J. & Tremaine, S. 1987, Galactic dynamics
Birnstiel, T. & Andrews, S. M. 2014, ApJ, 780, 153
Birnstiel, T., Andrews, S. M., & Ercolano, B. 2012, A&A, 544, A79
Birnstiel, T., Dullemond, C. P., & Brauer, F. 2010, A&A, 513, A79
Birnstiel, T., Dullemond, C. P., & Pinilla, P. 2013, A&A, 550, L8
Birnstiel, T., Fang, M., & Johansen, A. 2016, ArXiv e-prints
Blum, J. 2006, Advances in Physics, 55, 881
Blum, J. & Wurm, G. 2008, ARA&A, 46, 21
Boccaletti, A., Pantin, E., Lagrange, A.-M., et al. 2013, A&A, 560, A20
Boccaletti, A., Thalmann, C., Lagrange, A.-M., et al. 2015, Nature, 526, 230
Bodenheimer, P., Hubickyj, O., & Lissauer, J. J. 2000, Icarus, 143, 2
Bohlin, R. C., Savage, B. D., & Drake, J. F. 1978, ApJ, 224, 132
Bohren, C. F. & Huffman, D. R. 1983, Absorption and scattering of light by small particles
Booth, R. A. & Clarke, C. J. 2016, MNRAS, 458, 2676
Borucki, W. J., Koch, D., Basri, G., et al. 2010, Science, 327, 977
Boss, A. P. 2007, ApJ, 661, L73
Boss, A. P. 2011, ApJ, 731, 74
Bouwman, J., Meeus, G., de Koter, A., et al. 2001, A&A, 375, 950
Brandt, T. D., McElwain, M. W., Turner, E. L., et al. 2013, ApJ, 764, 183
Brauer, F., Dullemond, C. P., & Henning, T. 2008, A&A, 480, 859
Brown, J. M., Blake, G. A., Qi, C., et al. 2009, ApJ, 704, 496
Bruderer, S. 2013, A&A, 559, A46
Bruderer, S., van der Marel, N., van Dishoeck, E. F., & van Kempen, T. A. 2014, A&A, 562, A26
Bruderer, S., van Dishoeck, E. F., Doty, S. D., & Herczeg, G. J. 2012, A&A, 541, A91
Bryden, G., Chen, X., Lin, D. N. C., Nelson, R. P., & Papaloizou, J. C. B. 1999, ApJ, 514, 344
Cabrit, S., Edwards, S., Strom, S. E., & Strom, K. M. 1990, ApJ, 354, 687
Calvet, N., D'Alessio, P., Hartmann, L., et al. 2002, ApJ, 568, 1008
Calvet, N., D'Alessio, P., Watson, D. M., et al. 2005, ApJ, 630, L185
Canovas, H., Caceres, C., Schreiber, M. R., et al. 2016, MNRAS, 458, L29

Canovas, H., Ménard, F., Hales, A., et al. 2013, A&A, 556, A123
Carballido, A., Fromang, S., & Papaloizou, J. 2006, MNRAS, 373, 1633
Carmona, A., Pinte, C., Thi, W. F., et al. 2014, A&A, 567, A51
Carmona, A., Thi, W. F., Kamp, I., et al. 2016, ArXiv e-prints
Carmona, A., van den Ancker, M. E., Henning, T., et al. 2008, A&A, 477, 839
Carrasco-Gonzalez, C., Henning, T., Chandler, C. J., et al. 2016
Casassus, S. 2016, ArXiv e-prints
Casassus, S., Marino, S., Pérez, S., et al. 2015a, ApJ, 811, 92
Casassus, S., Perez M., S., Jordán, A., et al. 2012, ApJ, 754, L31
Casassus, S., van der Plas, G., M, S. P., et al. 2013, Nature, 493, 191
Casassus, S., Wright, C. M., Marino, S., et al. 2015b, ApJ, 812, 126
Chandrasekhar, S. 1961, Hydrodynamic and hydromagnetic stability
Chapman, S. & Cowling, T. G. 1970, The mathematical theory of non-uniform gases. an account of the kinetic theory of viscosity, thermal conduction and diffusion in gases
Chatterjee, S. & Tan, J. C. 2014, ApJ, 780, 53
Chen, H.-R. V., Keto, E., Zhang, Q., et al. 2016, ApJ, 823, 125
Chen, X. P., Henning, T., van Boekel, R., & Grady, C. A. 2006, A&A, 445, 331
Chiang, E. I. & Goldreich, P. 1997, ApJ, 490, 368
Chiang, E. I., Joung, M. K., Creech-Eakman, M. J., et al. 2001, ApJ, 547, 1077
Christiaens, V., Casassus, S., Perez, S., van der Plas, G., & Ménard, F. 2014, ApJ, 785, L12
Clarke, C. 2011, The Dispersal of Disks around Young Stars, ed. P. J. V. Garcia, 355–418
Clarke, C. & Carswell, B. 2007, Principles of Astrophysical Fluid Dynamics
Clarke, C. J. 2009, MNRAS, 396, 1066
Clarke, C. J., Gendrin, A., & Sotomayor, M. 2001, MNRAS, 328, 485
Clarke, C. J. & Lodato, G. 2009, MNRAS, 398, L6
Clarke, C. J. & Pringle, J. E. 1988, MNRAS, 235, 365
Collins, K. A., Grady, C. A., Hamaguchi, K., et al. 2009, ApJ, 697, 557
Cossins, P., Lodato, G., & Clarke, C. J. 2009, MNRAS, 393, 1157
Cossins, P., Lodato, G., & Testi, L. 2010, MNRAS, 407, 181
Crida, A. & Morbidelli, A. 2007, MNRAS, 377, 1324
Crida, A., Morbidelli, A., & Masset, F. 2006, Icarus, 181, 587
Cumming, A., Butler, R. P., Marcy, G. W., et al. 2008, PASP, 120, 531
Currie, T., Lada, C. J., Plavchan, P., et al. 2009, ApJ, 698, 1
Cuzzi, J. N., Dobrovolskis, A. R., & Champney, J. M. 1993, Icarus, 106, 102
Cuzzi, J. N. & Hogan, R. C. 2003, Icarus, 164, 127
Cuzzi, J. N., Lissauer, J. J., & Shu, F. H. 1981, Nature, 292, 703
D'Alessio, P., Calvet, N., Hartmann, L., Lizano, S., & Cantó, J. 1999, ApJ, 527, 893
D'Alessio, P., Cantö, J., Calvet, N., & Lizano, S. 1998, ApJ, 500, 411
D'Alessio, P., Hartmann, L., Calvet, N., et al. 2005, ApJ, 621, 461
D'Angelo, G. & Lubow, S. H. 2008, ApJ, 685, 560
D'Angelo, G. & Lubow, S. H. 2010, ApJ, 724, 730
de Boer, J., Salter, G., Benisty, M., et al. 2016, ArXiv e-prints
de Gregorio-Monsalvo, I., Ménard, F., Dent, W., et al. 2013, A&A, 557, A133
de Juan Ovelar, M., Min, M., Dominik, C., et al. 2013, A&A, 560, A111
de Val-Borro, M., Artymowicz, P., D'Angelo, G., & Peplinski, A. 2007, A&A, 471, 1043
de Val-Borro, M., Edgar, R. G., Artymowicz, P., et al. 2006, MNRAS, 370, 529
Debes, J. H., Jang-Condell, H., Weinberger, A. J., Roberge, A., & Schneider, G. 2013, ApJ, 771, 45
Dent, W. R. F., Greaves, J. S., & Coulson, I. M. 2005, MNRAS, 359, 663

Dent, W. R. F., Thi, W. F., Kamp, I., et al. 2013, PASP, 125, 477
Dipierro, G., Laibe, G., Price, D. J., & Lodato, G. 2016, MNRAS, 459, L1
Dipierro, G., Lodato, G., Testi, L., & de Gregorio Monsalvo, I. 2014, MNRAS, 444, 1919
Dipierro, G., Pinilla, P., Lodato, G., & Testi, L. 2015a, MNRAS, 451, 974
Dipierro, G., Price, D., Laibe, G., et al. 2015b, MNRAS, 453, L73
Dodson-Robinson, S. E. & Salyk, C. 2011, ApJ, 738, 131
Dominik, C., Dullemond, C. P., Waters, L. B. F. M., & Walch, S. 2003, A&A, 398, 607
Dong, R. & Dawson, R. 2016, ArXiv e-prints
Dong, R., Hall, C., Rice, K., & Chiang, E. 2015a, ArXiv e-prints
Dong, R., Rafikov, R., Zhu, Z., et al. 2012, ApJ, 750, 161
Dong, R., Rafikov, R. R., Stone, J. M., & Petrovich, C. 2011, ApJ, 741, 56
Dong, R., Vorobyov, E., Pavlyuchenkov, Y., Chiang, E., & Liu, H. B. 2016a, ArXiv e-prints
Dong, R., Zhu, Z., Fung, J., et al. 2016b, ApJ, 816, L12
Dong, R., Zhu, Z., Rafikov, R., & Stone, J. 2015b, ArXiv e-prints
Dong, R., Zhu, Z., & Whitney, B. 2015c, ApJ, 809, 93
D'Orazio, D. J., Haiman, Z., Duffell, P., MacFadyen, A., & Farris, B. 2016, MNRAS, 459, 2379
D'Orazio, D. J., Haiman, Z., & MacFadyen, A. 2013, MNRAS, 436, 2997
Douglas, T. A., Caselli, P., Ilee, J. D., et al. 2013, MNRAS, 433, 2064
Doyle, L. R., Carter, J. A., Fabrycky, D. C., et al. 2011, Science, 333, 1602
Draine, B. T. 2006, ApJ, 636, 1114
Draine, B. T. & Lee, H. M. 1984, ApJ, 285, 89
Dubrulle, B., Morfill, G., & Sterzik, M. 1995, Icarus, 114, 237
Duchêne, G., McCabe, C., Pinte, C., et al. 2010, ApJ, 712, 112
Duffell, P. C. 2015, ApJ, 807, L11
Duffell, P. C. & Dong, R. 2015, ApJ, 802, 42
Duffell, P. C. & MacFadyen, A. I. 2012, ApJ, 755, 7
Duffell, P. C. & MacFadyen, A. I. 2013, ApJ, 769, 41
Dullemond, C. P. 2012, RADMC-3D: A multi-purpose radiative transfer tool, Astrophysics Source Code Library
Dullemond, C. P. & Dominik, C. 2004, A&A, 421, 1075
Dullemond, C. P. & Dominik, C. 2005, A&A, 434, 971
Dullemond, C. P., Dominik, C., & Natta, A. 2001, ApJ, 560, 957
Dullemond, C. P., Hollenbach, D., Kamp, I., & D'Alessio, P. 2007, Protostars and Planets V, 555
Dunkin, S. K., Barlow, M. J., & Ryan, S. G. 1997, MNRAS, 290, 165
Eisner, J. A. 2012, ApJ, 755, 23
Eisner, J. A., Monnier, J. D., Tuthill, P., & Lacour, S. 2009, ApJ, 698, L169
Eisner, J. A., Plambeck, R. L., Carpenter, J. M., et al. 2008, ApJ, 683, 304
Epstein, P. S. 1924, Physical Review, 23, 710
Ercolano, B., Clarke, C. J., & Hall, A. C. 2011, MNRAS, 410, 671
Espaillat, C., Calvet, N., D'Alessio, P., et al. 2007, ApJ, 670, L135
Espaillat, C., Calvet, N., Luhman, K. L., Muzerolle, J., & D'Alessio, P. 2008, ApJ, 682, L125
Espaillat, C., D'Alessio, P., Hernández, J., et al. 2010, ApJ, 717, 441
Espaillat, C., Ingleby, L., Hernández, J., et al. 2012, ApJ, 747, 103
Espaillat, C., Muzerolle, J., Najita, J., et al. 2014, Protostars and Planets VI, 497
Evans, II, N. J., Dunham, M. M., Jørgensen, J. K., et al. 2009, ApJS, 181, 321
Facchini, S., Lodato, G., & Price, D. J. 2013, MNRAS, 433, 2142
Facchini, S., Ricci, L., & Lodato, G. 2014, MNRAS, 442, 3700
Fan, L. S. & Zhu, C. 1988, Principles of Gas-Solid Flows, Cambridge Series in Chemical Engi-

neering (Cambridge University Press)
Farris, B. D., Duffell, P., MacFadyen, A. I., & Haiman, Z. 2014, ApJ, 783, 134
Fedele, D., van den Ancker, M. E., Henning, T., Jayawardhana, R., & Oliveira, J. M. 2010, A&A, 510, A72
Flock, M., Ruge, J. P., Dzyurkevich, N., et al. 2015, A&A, 574, A68
Follette, K. B., Tamura, M., Hashimoto, J., et al. 2013, ApJ, 767, 10
Forgan, D., Rice, K., Cossins, P., & Lodato, G. 2011, MNRAS, 410, 994
Forgan, D. H., Ilee, J. D., Cyganowski, C. J., Brogan, C. L., & Hunter, T. R. 2016, MNRAS, 463, 957
Fouchet, L., Gonzalez, J.-F., & Maddison, S. T. 2010, A&A, 518, A16
Fouchet, L., Maddison, S. T., Gonzalez, J.-F., & Murray, J. R. 2007, A&A, 474, 1037
Frank, J., King, A. R., & Raine, D. J. 1985, Accretion power in astrophysics
Fromang, S. & Nelson, R. P. 2009, A&A, 496, 597
Fu, W., Li, H., Lubow, S., & Li, S. 2014a, ApJ, 788, L41
Fu, W., Li, H., Lubow, S., Li, S., & Liang, E. 2014b, ApJ, 795, L39
Fukagawa, M., Tamura, M., Itoh, Y., et al. 2006, ApJ, 636, L153
Fukagawa, M., Tsukagoshi, T., Momose, M., et al. 2013, PASJ, 65, L14
Fung, J. & Dong, R. 2015, ApJ, 815, L21
Fung, J., Shi, J.-M., & Chiang, E. 2014, ApJ, 782, 88
Gammie, C. F. 1996, ApJ, 457, 355
Garaud, P., Barrière-Fouchet, L., & Lin, D. N. C. 2004, ApJ, 603, 292
Garaud, P. & Lin, D. N. C. 2007, ApJ, 654, 606
Garufi, A., Quanz, S. P., Schmid, H. M., et al. 2016, A&A, 588, A8
Garufi, A. et al. 2013, A&A, 560, A105
Gibbons, P. G., Mamatsashvili, G. R., & Rice, W. K. M. 2014, MNRAS, 442, 361
Gibbons, P. G., Mamatsashvili, G. R., & Rice, W. K. M. 2015, MNRAS, 453, 4232
Gibbons, P. G., Rice, W. K. M., & Mamatsashvili, G. R. 2012, MNRAS, 426, 1444
Ginski, C., Stolker, T., Pinilla, P., et al. 2016, ArXiv e-prints
Girart, J. M., Beltrán, M. T., Zhang, Q., Rao, R., & Estalella, R. 2009, Science, 324, 1408
Glauser, A. M., Ménard, F., Pinte, C., et al. 2008, A&A, 485, 531
Goldreich, P. & Nicholson, P. D. 1989, ApJ, 342, 1075
Goldreich, P. & Tremaine, S. 1979, ApJ, 233, 857
Goldreich, P. & Tremaine, S. 1980, ApJ, 241, 425
Goldreich, P. & Tremaine, S. D. 1978, Icarus, 34, 240
Goldreich, P. & Ward, W. R. 1973, ApJ, 183, 1051
Golimowski, D. A., Ardila, D. R., Krist, J. E., et al. 2006, AJ, 131, 3109
Gonzalez, J.-F., Laibe, G., Maddison, S. T., Pinte, C., & Ménard, F. 2015, MNRAS, 454, L36
Gonzalez, J.-F., Pinte, C., Maddison, S. T., Ménard, F., & Fouchet, L. 2012, A&A, 547, A58
Goodman, A. A., Benson, P. J., Fuller, G. A., & Myers, P. C. 1993, ApJ, 406, 528
Goodman, J. & Rafikov, R. R. 2001, ApJ, 552, 793
Gorti, U., Hollenbach, D., & Dullemond, C. P. 2015, ApJ, 804, 29
Gorti, U., Liseau, R., Sándor, Z., & Clarke, C. 2016, Space Sci. Rev.
Grady, C. A., Muto, T., Hashimoto, J., et al. 2013, ApJ, 762, 48
Grady, C. A., Woodgate, B., Bruhweiler, F. C., et al. 1999, ApJ, 523, L151
Greaves, J. S., Richards, A. M. S., Rice, W. K. M., & Muxlow, T. W. B. 2008, MNRAS, 391, L74
Guidi, G., Tazzari, M., Testi, L., et al. 2016, ArXiv e-prints
Gullbring, E., Hartmann, L., Briceño, C., & Calvet, N. 1998, ApJ, 492, 323
Gundlach, B. & Blum, J. 2015, ApJ, 798, 34

Güttler, C., Blum, J., Zsom, A., Ormel, C. W., & Dullemond, C. P. 2010, A&A, 513, A56
Haghighipour, N. & Boss, A. P. 2003a, ApJ, 598, 1301
Haghighipour, N. & Boss, A. P. 2003b, ApJ, 583, 996
Haisch, Jr., K. E., Lada, E. A., & Lada, C. J. 2001, ApJ, 553, L153
Hall, C., Forgan, D., Rice, K., et al. 2016, ArXiv e-prints
Hartmann, L. 2009, Accretion Processes in Star Formation: Second Edition (Cambridge University Press)
Hartmann, L., Calvet, N., Gullbring, E., & D'Alessio, P. 1998, ApJ, 495, 385
Hatzes, A. P. 2016, Space Sci. Rev.
Haworth, T. J., Ilee, J. D., Forgan, D. H., et al. 2016, ArXiv e-prints
Hayashi, M., Ohashi, N., & Miyama, S. M. 1993, ApJ, 418, L71
Herbig, G. H. 1960, ApJS, 4, 337
Herbig, G. H. 1962, Advances in Astronomy and Astrophysics, 1, 47
Hernández, J., Hartmann, L., Megeath, T., et al. 2007, ApJ, 662, 1067
Hersant, F. 2009, A&A, 502, 385
Hillenbrand, L. A., Carpenter, J. M., Kim, J. S., et al. 2008, ApJ, 677, 630
Högbom, J. A. 1974, A&AS, 15, 417
Hollenbach, D., Johnstone, D., Lizano, S., & Shu, F. 1994, ApJ, 428, 654
Howard, A. W., Marcy, G. W., Johnson, J. A., et al. 2010, Science, 330, 653
Hoyle, F. 1960, QJRAS, 1, 28
Huélamo, N., Lacour, S., Tuthill, P., et al. 2011, A&A, 528, L7
Hull, C. L. H., Plambeck, R. L., Kwon, W., et al. 2014, ApJS, 213, 13
Hunter, T. R., Brogan, C. L., Cyganowski, C. J., & Young, K. H. 2014, ApJ, 788, 187
Hutchison, M. A., Price, D. J., Laibe, G., & Maddison, S. T. 2016, ArXiv e-prints
Ilee, J. D., Boley, A. C., Caselli, P., et al. 2011, MNRAS, 417, 2950
Ilee, J. D., Cyganowski, C. J., Nazari, P., et al. 2016, MNRAS, 462, 4386
Ingleby, L., Calvet, N., Bergin, E., et al. 2011, ApJ, 743, 105
Inoue, A. K., Honda, M., Nakamoto, T., & Oka, A. 2008, PASJ, 60, 557
Ireland, M. J. & Kraus, A. L. 2008, ApJ, 678, L59
Isella, A., Carpenter, J. M., & Sargent, A. I. 2009, ApJ, 701, 260
Isella, A., Carpenter, J. M., & Sargent, A. I. 2010a, ApJ, 714, 1746
Isella, A., Natta, A., Wilner, D., Carpenter, J. M., & Testi, L. 2010b, ApJ, 725, 1735
Isella, A., Pérez, L. M., Carpenter, J. M., et al. 2013, ApJ, 775, 30
Isella, A., Tatulli, E., Natta, A., & Testi, L. 2008, A&A, 483, L13
Isella, A. & Turner, N. 2016, ArXiv e-prints
Ivanov, P. B., Papaloizou, J. C. B., & Polnarev, A. G. 1999, MNRAS, 307, 79
Jacquet, E., Balbus, S., & Latter, H. 2011, MNRAS, 415, 3591
Jang-Condell, H. & Sasselov, D. D. 2003, ApJ, 593, 1116
Jang-Condell, H. & Turner, N. J. 2012, ApJ, 749, 153
Jeans, J. H. 1902, Philosophical Transactions of the Royal Society of London Series A, 199, 1
Jin, S., Li, S., Isella, A., Li, H., & Ji, J. 2016, ArXiv e-prints
Johansen, A., Andersen, A. C., & Brandenburg, A. 2004, A&A, 417, 361
Johansen, A., Henning, T., & Klahr, H. 2006, ApJ, 643, 1219
Johansen, A. & Klahr, H. 2005, ApJ, 634, 1353
Johansen, A., Klahr, H., & Henning, T. 2011, A&A, 529, A62
Johansen, A., Oishi, J. S., Mac Low, M.-M., et al. 2007, Nature, 448, 1022
Johansen, A., Youdin, A., & Klahr, H. 2009a, ApJ, 697, 1269
Johansen, A., Youdin, A., & Mac Low, M.-M. 2009b, ApJ, 704, L75
Johnston, K. G., Robitaille, T. P., Beuther, H., et al. 2015, ApJ, 813, L19

Jones, M. G., Pringle, J. E., & Alexander, R. D. 2012, MNRAS, 419, 925
Joy, A. H. 1945, ApJ, 102, 168
Juhász, A., Benisty, M., Pohl, A., et al. 2015, MNRAS, 451, 1147
Kanagawa, K. D., Muto, T., Tanaka, H., et al. 2015a, ApJ, 806, L15
Kanagawa, K. D., Muto, T., Tanaka, H., et al. 2016, PASJ, 68, 43
Kanagawa, K. D., Tanaka, H., Muto, T., Tanigawa, T., & Takeuchi, T. 2015b, MNRAS, 448, 994
Kant, I. 1755, Allgemeine Naturgeschichte und Theorie des Himmels
Kasper, M., Apai, D., Wagner, K., & Robberto, M. 2015, ApJ, 812, L33
Kataoka, A., Muto, T., Momose, M., Tsukagoshi, T., & Dullemond, C. P. 2015a, ArXiv e-prints
Kataoka, A., Muto, T., Momose, M., et al. 2015b, ApJ, 809, 78
Kenyon, S. J. & Hartmann, L. 1987, ApJ, 323, 714
Kessler-Silacci, J., Augereau, J.-C., Dullemond, C. P., et al. 2006, ApJ, 639, 275
Kley, W. & Dirksen, G. 2006, A&A, 447, 369
Kley, W. & Nelson, R. P. 2012, ARA&A, 50, 211
Koepferl, C. M., Ercolano, B., Dale, J., et al. 2013, MNRAS, 428, 3327
Kratter, K. M. & Lodato, G. 2016, ArXiv e-prints
Kratter, K. M. & Murray-Clay, R. A. 2011, ApJ, 740, 1
Kraus, A. L. & Ireland, M. J. 2012, ApJ, 745, 5
Kudo, T., Tamura, M., Kitamura, Y., et al. 2008, ApJ, 673, L67
Kwok, S. 1975, ApJ, 198, 583
Kwon, W., Looney, L. W., & Mundy, L. G. 2011, ApJ, 741, 3
Kwon, W., Looney, L. W., Mundy, L. G., & Welch, W. J. 2015, ApJ, 808, 102
Lacour, S., Biller, B., Cheetham, A., et al. 2016, A&A, 590, A90
Lada, C. J. 1985, ARA&A, 23, 267
Laibe, G. 2014, MNRAS, 437, 3037
Laibe, G., Gonzalez, J.-F., Fouchet, L., & Maddison, S. T. 2008, A&A, 487, 265
Laibe, G., Gonzalez, J.-F., & Maddison, S. T. 2012, A&A, 537, A61
Laibe, G., Gonzalez, J.-F., Maddison, S. T., & Crespe, E. 2014, MNRAS, 437, 3055
Laibe, G. & Price, D. J. 2011, MNRAS, 418, 1491
Laibe, G. & Price, D. J. 2012a, MNRAS, 420, 2345
Laibe, G. & Price, D. J. 2012b, MNRAS, 420, 2365
Laibe, G. & Price, D. J. 2014, MNRAS, 440, 2147
Lambrechts, M., Johansen, A., & Morbidelli, A. 2014, A&A, 572, A35
Landau, L. D. & Lifshitz, E. M. 1980, Statistical physics. Pt.1, Pt.2
Laplace, P. S. 1796
Lau, Y. Y. & Bertin, G. 1978, ApJ, 226, 508
Lenzen, R., Hartung, M., Brandner, W., et al. 2003, in Proc. SPIE, Vol. 4841, Instrument Design and Performance for Optical/Infrared Ground-based Telescopes, ed. M. Iye & A. F. M. Moorwood, 944–952
Li, H., Lubow, S. H., Li, S., & Lin, D. N. C. 2009, ApJ, 690, L52
Lin, C. C. & Shu, F. H. 1964, ApJ, 140, 646
Lin, D. N. C., Bodenheimer, P., & Richardson, D. C. 1996, Nature, 380, 606
Lin, D. N. C. & Papaloizou, J. 1979, MNRAS, 186, 799
Lin, D. N. C. & Papaloizou, J. 1986, ApJ, 309, 846
Lin, D. N. C. & Papaloizou, J. C. B. 1993, in Protostars and Planets III, ed. E. H. Levy & J. I. Lunine, 749–835
Lin, D. N. C. & Pringle, J. E. 1987, MNRAS, 225, 607
Lindblad, B. 1927, MNRAS, 87, 420

Lissauer, J. J. 1993, ARA&A, 31, 129
Lodato, G. 2007, Nuovo Cimento Rivista Serie, 30
Lodato, G. 2008, New Astronomy Reviews, 52, 21
Lodato, G. & Bertin, G. 2003, A&A, 408, 1015
Lodato, G. & Clarke, C. J. 2011, MNRAS, 413, 2735
Lodato, G., Nayakshin, S., King, A. R., & Pringle, J. E. 2009, MNRAS, 398, 1392
Lodato, G. & Price, D. J. 2010, MNRAS, 405, 1212
Lodato, G. & Pringle, J. E. 2007, MNRAS, 381, 1287
Lodato, G. & Rice, W. K. M. 2004, MNRAS, 351, 630
Lodato, G. & Rice, W. K. M. 2005, MNRAS, 358, 1489
Lorén-Aguilar, P. & Bate, M. R. 2016, MNRAS, 457, L54
Lovelace, R. V. E., Li, H., Colgate, S. A., & Nelson, A. F. 1999, ApJ, 513, 805
Lubow, S. H. & D'Angelo, G. 2006, ApJ, 641, 526
Lubow, S. H., Seibert, M., & Artymowicz, P. 1999, ApJ, 526, 1001
Lynden-Bell, D. & Pringle, J. E. 1974, MNRAS, 168, 603
Lyo, A.-R., Ohashi, N., Qi, C., Wilner, D. J., & Su, Y.-N. 2011, AJ, 142, 151
Lyra, W., Johansen, A., Klahr, H., & Piskunov, N. 2009, A&A, 493, 1125
Lyra, W. & Lin, M.-K. 2013, ApJ, 775, 17
Macintosh, B. A., Graham, J. R., Palmer, D. W., et al. 2008, in Proc. SPIE, Vol. 7015, Adaptive Optics Systems, 701518
Malik, M., Meru, F., Mayer, L., & Meyer, M. 2015, ApJ, 802, 56
Mamajek, E. E. 2008, Astronomische Nachrichten, 329, 10
Mamajek, E. E. 2009, in American Institute of Physics Conference Series, Vol. 1158, American Institute of Physics Conference Series, ed. T. Usuda, M. Tamura, & M. Ishii, 3–10
Manara, C. F., Rosotti, G., Testi, L., et al. 2016, A&A, 591, L3
Marino, S., Perez, S., & Casassus, S. 2015, ApJ, 798, L44
Martin, R. G. & Lubow, S. H. 2011, ApJ, 740, L6
Martin, R. G. & Lubow, S. H. 2014, MNRAS, 437, 682
Masset, F. & Kley, W. 2006, Disk-planet interaction and migration, ed. H. Klahr & W. Brandner (Cambridge University Press), 216
Masset, F. S. 2002, A&A, 387, 605
Masset, F. S., D'Angelo, G., & Kley, W. 2006, ApJ, 652, 730
Masset, F. S. & Papaloizou, J. C. B. 2003, ApJ, 588, 494
Mathis, J. S., Rumpl, W., & Nordsieck, K. H. 1977, ApJ, 217, 425
McCaughrean, M. J. & O'dell, C. R. 1996, AJ, 111, 1977
McClure, M., Bergin, T., Cleeves, I., et al. 2016, ArXiv e-prints
Meeus, G., Montesinos, B., Mendigutía, I., et al. 2012, A&A, 544, A78
Mendigutía, I., Mora, A., Montesinos, B., et al. 2012, A&A, 543, A59
Meru, F. & Bate, M. R. 2011a, MNRAS, 411, L1
Meru, F. & Bate, M. R. 2011b, MNRAS, 410, 559
Meru, F. & Bate, M. R. 2012, MNRAS, 427, 2022
Mestel, L. 1963, MNRAS, 126, 553
Meyer-Vernet, N. & Sicardy, B. 1987, Icarus, 69, 157
Min, M., Dominik, C., Hovenier, J. W., de Koter, A., & Waters, L. B. F. M. 2006, A&A, 445, 1005
Min, M., Rab, C., Woitke, P., Dominik, C., & Ménard, F. 2016, A&A, 585, A13
Miotello, A., Bruderer, S., & van Dishoeck, E. F. 2014a, A&A, 572, A96
Miotello, A., Testi, L., Lodato, G., et al. 2014b, A&A, 567, A32
Mishchenko, M. I., Hovenier, J. W., & Travis, L. D. 2000, Light scattering by nonspherical

particles : theory, measurements, and applications
Miyake, K. & Nakagawa, Y. 1993, Icarus, 106, 20
Monaghan, J. J. 2005, Reports on Progress in Physics, 68, 1703
Morfill, G. E. 1985, in Birth and Infancy of Stars, ed. R. Lucas, A. Omont, & R. Stora, 693–792
Morfill, G. E. & Voelk, H. J. 1984, ApJ, 287, 371
Mulders, G. D., Min, M., Dominik, C., Debes, J. H., & Schneider, G. 2013, A&A, 549, A112
Mundt, R., Buehrke, T., Solf, J., Ray, T. P., & Raga, A. C. 1990, A&A, 232, 37
Muto, T., Grady, C. A., Hashimoto, J., et al. 2012, ApJ, 748, L22
Muto, T., Tsukagoshi, T., Momose, M., et al. 2015, PASJ, 67, 122
Muzerolle, J., Calvet, N., Briceño, C., Hartmann, L., & Hillenbrand, L. 2000, ApJ, 535, L47
Nakagawa, Y., Sekiya, M., & Hayashi, C. 1986, Icarus, 67, 375
Nath, S. K. & Mukhopadhyay, B. 2016, ArXiv e-prints
Natta, A., Prusti, T., Neri, R., et al. 2001, A&A, 371, 186
Natta, A. & Testi, L. 2004, in Astronomical Society of the Pacific Conference Series, Vol. 323, Star Formation in the Interstellar Medium: In Honor of David Hollenbach, ed. D. Johnstone, F. C. Adams, D. N. C. Lin, D. A. Neufeeld, & E. C. Ostriker, 279
Natta, A., Testi, L., Calvet, N., et al. 2007, Protostars and Planets V, 767
Nixon, C., King, A., & Price, D. 2013, MNRAS, 434, 1946
Nomura, H., Tsukagoshi, T., Kawabe, R., et al. 2016, ApJ, 819, L7
O'dell, C. R., Wen, Z., & Hu, X. 1993, ApJ, 410, 696
Ogilvie, G. I. & Lubow, S. H. 2002, MNRAS, 330, 950
Okuzumi, S., Momose, M., Sirono, S.-i., Kobayashi, H., & Tanaka, H. 2015, ArXiv e-prints
Okuzumi, S., Tanaka, H., Takeuchi, T., & Sakagami, M.-a. 2011, ApJ, 731, 96
Ormel, C. W. & Cuzzi, J. N. 2007, A&A, 466, 413
Orosz, J. A., Welsh, W. F., Carter, J. A., et al. 2012a, ApJ, 758, 87
Orosz, J. A. et al. 2012b, Science, 337, 1511
Owen, J. E. 2014, ApJ, 789, 59
Owen, J. E. 2015, ArXiv e-prints
Owen, J. E., Ercolano, B., & Clarke, C. J. 2011a, MNRAS, 412, 13
Owen, J. E., Ercolano, B., & Clarke, C. J. 2011b, MNRAS, 411, 1104
Owen, J. E. & Kollmeier, J. A. 2016, ArXiv e-prints
Paardekooper, S.-J., Baruteau, C., Crida, A., & Kley, W. 2010, MNRAS, 401, 1950
Paardekooper, S.-J. & Mellema, G. 2004, A&A, 425, L9
Paardekooper, S.-J. & Mellema, G. 2006, A&A, 453, 1129
Paczynski, B. 1978, Acta. Astron., 28, 91
Papaloizou, J. & Lin, D. N. C. 1984, ApJ, 285, 818
Papaloizou, J. C. B. & Nelson, R. P. 2005, A&A, 433, 247
Papaloizou, J. C. B., Nelson, R. P., & Masset, F. 2001, A&A, 366, 263
Pardo, J., Cernicharo, J., & Serabyn, E. 2002, in Astronomical Society of the Pacific Conference Series, Vol. 266, Astronomical Site Evaluation in the Visible and Radio Range, ed. J. Vernin, Z. Benkhaldoun, & C. Muñoz-Tuñón, 188
Pascucci, I., Gorti, U., Hollenbach, D., et al. 2006, ApJ, 651, 1177
Pérez, L. M., Carpenter, J. M., Andrews, S. M., et al. 2016, Science, 353, 1519
Pérez, L. M., Carpenter, J. M., Chandler, C. J., et al. 2012, ApJ, 760, L17
Pérez, L. M., Chandler, C. J., Isella, A., et al. 2015, ArXiv e-prints
Pérez, L. M., Isella, A., Carpenter, J. M., & Chandler, C. J. 2014, ApJ, 783, L13
Perez, S., Casassus, S., Ménard, F., et al. 2015, ApJ, 798, 85
Perrin, M. D., Duchêne, G., Kalas, P., & Graham, J. R. 2006, ApJ, 645, 1272
Perrin, M. D., Graham, J. R., Kalas, P., et al. 2004, Science, 303, 1345

Perrin, M. D., Maire, J., Ingraham, P., et al. 2014, in Proc. SPIE, Vol. 9147, Ground-based and Airborne Instrumentation for Astronomy V, 91473J
Picogna, G. & Kley, W. 2015, A&A, 584, A110
Piétu, V., Dutrey, A., Guilloteau, S., Chapillon, E., & Pety, J. 2006, A&A, 460, L43
Pinilla, P., Benisty, M., & Birnstiel, T. 2012a, A&A, 545, A81
Pinilla, P., Birnstiel, T., Ricci, L., et al. 2012b, A&A, 538, A114
Pinilla, P., van der Marel, N., Pérez, L. M., et al. 2015, ArXiv e-prints
Pinte, C., Dent, W. R. F., Ménard, F., et al. 2016, ApJ, 816, 25
Pinte, C., Fouchet, L., Ménard, F., Gonzalez, J.-F., & Duchêne, G. 2007, A&A, 469, 963
Pinte, C., Ménard, F., Duchêne, G., & Bastien, P. 2006, A&A, 459, 797
Pinte, C., Padgett, D. L., Ménard, F., et al. 2008, A&A, 489, 633
Pohl, A., Kataoka, A., Pinilla, P., et al. 2016, A&A, 593, A12
Pohl, A., Pinilla, P., Benisty, M., et al. 2015, MNRAS, 453, 1768
Pollack, J. B., Hollenbach, D., Beckwith, S., et al. 1994, ApJ, 421, 615
Pontoppidan, K. M., Blake, G. A., van Dishoeck, E. F., et al. 2008, ApJ, 684, 1323
Poppe, T., Blum, J., & Henning, T. 1999, Advances in Space Research, 23, 1197
Price, D. J. 2012, Journal of Computational Physics, 231, 759
Price, D. J. & Federrath, C. 2010, MNRAS, 406, 1659
Price, D. J. & Laibe, G. 2015, MNRAS, 451, 813
Pringle, J. E. 1981, ARA&A, 19, 137
Pringle, J. E. & King, A. R. 2007, Astrophysical flows, 199–201
Pringle, J. E., Verbunt, F., & Wade, R. A. 1986, MNRAS, 221, 169
Qi, C., Öberg, K. I., Andrews, S. M., et al. 2015, ApJ, 813, 128
Quanz, S. P., Amara, A., Meyer, M. R., et al. 2015, ApJ, 807, 64
Quanz, S. P., Amara, A., Meyer, M. R., et al. 2013, ApJ, 766, L1
Quanz, S. P., Schmid, H. M., Geissler, K., et al. 2011, ApJ, 738, 23
Quillen, A. C., Varnière, P., Minchev, I., & Frank, A. 2005, AJ, 129, 2481
Rafikov, R. R. 2002, ApJ, 569, 997
Rafikov, R. R. 2005, ApJ, 621, L69
Rafikov, R. R. 2007, ApJ, 662, 642
Rafikov, R. R. 2009, ApJ, 704, 281
Rafikov, R. R. 2015, ApJ, 804, 62
Rafikov, R. R. & Petrovich, C. 2012, ApJ, 747, 24
Ragusa, E., Dipierro, G., Lodato, G., Laibe, G., & Price, D. J. 2017, MNRAS, 464, 1449
Ragusa, E., Lodato, G., & Price, D. J. 2016, MNRAS, 460, 1243
Ratzka, T., Köhler, R., & Leinert, C. 2005, A&A, 437, 611
Regály, Z., Juhász, A., Sándor, Z., & Dullemond, C. P. 2012, MNRAS, 419, 1701
Regály, Z., Sándor, Z., Csomós, P., & Ataiee, S. 2013, MNRAS, 433, 2626
Reggiani, M., Quanz, S. P., Meyer, M. R., et al. 2014, ApJ, 792, L23
Ribas, Á., Merín, B., Bouy, H., & Maud, L. T. 2014, A&A, 561, A54
Ricci, L., Testi, L., Natta, A., et al. 2010, A&A, 512, A15
Ricci, L., Trotta, F., Testi, L., et al. 2012, A&A, 540, A6
Rice, W. K. M., Armitage, P. J., Mamatsashvili, G. R., Lodato, G., & Clarke, C. J. 2011, MNRAS, 418, 1356
Rice, W. K. M., Armitage, P. J., Wood, K., & Lodato, G. 2006a, MNRAS, 373, 1619
Rice, W. K. M., Lodato, G., & Armitage, P. J. 2005, MNRAS, 364, L56
Rice, W. K. M., Lodato, G., Pringle, J. E., Armitage, P. J., & Bonnell, I. A. 2004, MNRAS, 355, 543
Rice, W. K. M., Lodato, G., Pringle, J. E., Armitage, P. J., & Bonnell, I. A. 2006b, MNRAS,

372, L9
Robitaille, T. P., Whitney, B. A., Indebetouw, R., & Wood, K. 2007, ApJS, 169, 328
Rodigas, T. J., Follette, K. B., Weinberger, A., Close, L., & Hines, D. C. 2014, ApJ, 791, L37
Rodríguez, L. F., Loinard, L., D'Alessio, P., Wilner, D. J., & Ho, P. T. P. 2005, ApJ, 621, L133
Rosenfeld, K. A., Chiang, E., & Andrews, S. M. 2014, ApJ, 782, 62
Rosenfeld, K. A., Qi, C., Andrews, S. M., et al. 2012, ApJ, 757, 129
Rosotti, G. P., Ercolano, B., & Owen, J. E. 2015, ArXiv e-prints
Rosotti, G. P., Ercolano, B., Owen, J. E., & Armitage, P. J. 2013, MNRAS, 430, 1392
Rosotti, G. P., Juhasz, A., Booth, R. A., & Clarke, C. J. 2016, MNRAS, 459, 2790
Rousset, G., Lacombe, F., Puget, P., et al. 2003, in Proc. SPIE, Vol. 4839, Adaptive Optical System Technologies II, ed. P. L. Wizinowich & D. Bonaccini, 140–149
Ruge, J. P., Flock, M., Wolf, S., et al. 2016, A&A, 590, A17
Ruge, J. P., Wolf, S., Uribe, A. L., & Klahr, H. H. 2013, A&A, 549, A97
Ruíz-Rodríguez, D., Ireland, M., Cieza, L., & Kraus, A. 2016, MNRAS, 463, 3829
Rybicki, G. B. & Lightman, A. P. 1979, Radiative processes in astrophysics
Safronov, V. S. 1969, Evoliutsiia doplanetnogo oblaka.
Savage, B. D. & Mathis, J. S. 1979, ARA&A, 17, 73
Schräpler, R. & Henning, T. 2004, ApJ, 614, 960
Shakura, N. I. & Sunyaev, R. A. 1973, A&A, 24, 337
Sheehan, P. D. & Eisner, J. A. 2014, ApJ, 791, 19
Shi, J.-M., Krolik, J. H., Lubow, S. H., & Hawley, J. F. 2012, ApJ, 749, 118
Shu, F. H., Adams, F. C., & Lizano, S. 1987, ARA&A, 25, 23
Sicilia-Aguilar, A., Hartmann, L., Calvet, N., et al. 2006, ApJ, 638, 897
Sicilia-Aguilar, A., Hartmann, L. W., Watson, D., et al. 2007, ApJ, 659, 1637
Sicilia-Aguilar, A., Henning, T., Dullemond, C. P., et al. 2011, ApJ, 742, 39
Spitzer, Jr., L. 1942, ApJ, 95, 329
Stepinski, T. F. & Valageas, P. 1996, A&A, 309, 301
Stolker, T., Dominik, C., Avenhaus, H., et al. 2016, ArXiv e-prints
Strom, K. M., Strom, S. E., Edwards, S., Cabrit, S., & Skrutskie, M. F. 1989, AJ, 97, 1451
Supulver, K. D. & Lin, D. N. C. 2000, Icarus, 146, 525
Syer, D. & Clarke, C. J. 1995, MNRAS, 277, 758
Takahashi, S. Z. & Inutsuka, S.-i. 2016, ArXiv e-prints
Takami, M., Hasegawa, Y., Muto, T., et al. 2014, ApJ, 795, 71
Takeuchi, T. & Artymowicz, P. 2001, ApJ, 557, 990
Takeuchi, T. & Lin, D. N. C. 2002, ApJ, 581, 1344
Takeuchi, T., Miyama, S. M., & Lin, D. N. C. 1996, ApJ, 460, 832
Taki, T., Fujimoto, M., & Ida, S. 2016, ArXiv e-prints
Tamayo, D., Triaud, A. H. M. J., Menou, K., & Rein, H. 2015, ApJ, 805, 100
Tamura, M. 2009, in American Institute of Physics Conference Series, Vol. 1158, American Institute of Physics Conference Series, ed. T. Usuda, M. Tamura, & M. Ishii, 11–16
Tanaka, H., Takeuchi, T., & Ward, W. R. 2002, ApJ, 565, 1257
Tazzari, M. & Lodato, G. 2015, MNRAS, 449, 1118
Tazzari, M., Testi, L., Ercolano, B., et al. 2015
Terebey, S., Shu, F. H., & Cassen, P. 1984, ApJ, 286, 529
Testi, L., Birnstiel, T., Ricci, L., et al. 2014, Protostars and Planets VI, 339
Testi, L., Natta, A., Scholz, A., et al. 2016, ArXiv e-prints
Testi, L., Natta, A., Shepherd, D. S., & Wilner, D. J. 2001, ApJ, 554, 1087
Testi, L., Natta, A., Shepherd, D. S., & Wilner, D. J. 2003, A&A, 403, 323
Testi, L., Skemer, A., Henning, T., et al. 2015, ApJ, 812, L38

Thalmann, C., Mulders, G. D., Hodapp, K., et al. 2014, A&A, 566, A51
Thompson, R. A., Moran, M. J., & Swenson, W. G. 2004, Interferometry and Synthesis in Radio Astronomy (Wiley)
Tilling, I., Woitke, P., Meeus, G., et al. 2012, A&A, 538, A20
Tobin, J. J., Hartmann, L., Chiang, H.-F., et al. 2012, Nature, 492, 83
Tobin, J. J., Kratter, K. M., Persson, M. V., et al. 2016, Nature, 538, 483
Toomre, A. 1964, ApJ, 139, 1217
Turner, N. J., Fromang, S., Gammie, C., et al. 2014, Protostars and Planets VI, 411
van Boekel, R., Dullemond, C. P., & Dominik, C. 2005, A&A, 441, 563
van Boekel, R., Henning, T., Menu, J., et al. 2016, ArXiv e-prints
van Boekel, R., Waters, L. B. F. M., Dominik, C., et al. 2003, A&A, 400, L21
van der Marel, N., Cazzoletti, P., Pinilla, P., & Garufi, A. 2016a, ArXiv e-prints
van der Marel, N., Pinilla, P., Tobin, J., et al. 2015a, ApJ, 810, L7
van der Marel, N., van Dishoeck, E. F., Bruderer, S., et al. 2016b, A&A, 585, A58
van der Marel, N., van Dishoeck, E. F., Bruderer, S., et al. 2013, Science, 340, 1199
van der Marel, N., van Dishoeck, E. F., Bruderer, S., Pérez, L., & Isella, A. 2015b, A&A, 579, A106
van der Plas, G., Wright, C. M., Ménard, F., et al. 2016, ArXiv e-prints
van Dishoeck, E. F. & Black, J. H. 1988, ApJ, 334, 771
Varnière, P., Quillen, A. C., & Frank, A. 2004, ApJ, 612, 1152
Verhoeff, A. P., Min, M., Pantin, E., et al. 2011, A&A, 528, A91
Vicente, S., Merín, B., Hartung, M., et al. 2011, A&A, 533, A135
Visser, R., van Dishoeck, E. F., & Black, J. H. 2009, A&A, 503, 323
Voelk, H. J., Jones, F. C., Morfill, G. E., & Roeser, S. 1980, A&A, 85, 316
Vorobyov, E. I. & Basu, S. 2005, ApJ, 633, L137
Vorobyov, E. I. & Basu, S. 2006, ApJ, 650, 956
Vorobyov, E. I. & Basu, S. 2015, ApJ, 805, 115
Vorobyov, E. I., Zakhozhay, O. V., & Dunham, M. M. 2013, MNRAS, 433, 3256
Wada, K., Tanaka, H., Suyama, T., Kimura, H., & Yamamoto, T. 2009, ApJ, 702, 1490
Waelkens, C., Waters, L. B. F. M., de Graauw, M. S., et al. 1996, A&A, 315, L245
Wagner, K., Apai, D., Kasper, M., & Robberto, M. 2015, ArXiv e-prints
Walsh, C., Juhász, A., Pinilla, P., et al. 2014, ApJ, 791, L6
Ward, W. R. 1997, Icarus, 126, 261
Ward, W. R. & Hourigan, K. 1989, ApJ, 347, 490
Ward-Thompson, D., Kirk, J. M., Greaves, J. S., & André, P. 2011, MNRAS, 415, 2812
Watson, A. M., Stapelfeldt, K. R., Wood, K., & Ménard, F. 2007, Protostars and Planets V, 523
Weidenschilling, S. J. 1977a, MNRAS, 180, 57
Weidenschilling, S. J. 1977b, Ap&SS, 51, 153
Weidenschilling, S. J. 1980, Icarus, 44, 172
Whipple, F. L. 1972
Williams, J. P. & Best, W. M. J. 2014, ApJ, 788, 59
Williams, J. P. & Cieza, L. A. 2011, ARA&A, 49, 67
Willson, M., Kraus, S., Kluska, J., et al. 2016, ArXiv e-prints
Winn, J. N. & Fabrycky, D. C. 2015, ARA&A, 53, 409
Woitke, P., Min, M., Pinte, C., et al. 2016, A&A, 586, A103
Wright, C. M., Maddison, S. T., Wilner, D. J., et al. 2015, MNRAS, 453, 414
Youdin, A. & Johansen, A. 2007, ApJ, 662, 613
Youdin, A. N. & Goodman, J. 2005, ApJ, 620, 459

Youdin, A. N. & Lithwick, Y. 2007, Icarus, 192, 588
Youdin, A. N. & Shu, F. H. 2002, ApJ, 580, 494
Young, M. D. & Clarke, C. J. 2015, MNRAS, 451, 3987
Yu, C., Li, H., Li, S., Lubow, S. H., & Lin, D. N. C. 2010, ApJ, 712, 198
Zapata, L. A., Loinard, L., Rodríguez, L. F., et al. 2013, ApJ, 764, L14
Zapata, L. A., Palau, A., Galván-Madrid, R., et al. 2015, MNRAS, 447, 1826
Zhang, K., Blake, G. A., & Bergin, E. A. 2015, ApJ, 806, L7
Zhu, Z., Dong, R., Stone, J. M., & Rafikov, R. R. 2015a, ArXiv e-prints
Zhu, Z., Nelson, R. P., Dong, R., Espaillat, C., & Hartmann, L. 2012, ApJ, 755, 6
Zhu, Z., Nelson, R. P., Hartmann, L., Espaillat, C., & Calvet, N. 2011, ApJ, 729, 47
Zhu, Z. & Stone, J. M. 2014, ApJ, 795, 53
Zhu, Z., Stone, J. M., & Bai, X.-N. 2015b, ApJ, 801, 81
Zhu, Z., Stone, J. M., & Rafikov, R. R. 2013, ApJ, 768, 143
Zsom, A., Ormel, C. W., Güttler, C., Blum, J., & Dullemond, C. P. 2010, A&A, 513, A57
Zuckerman, B., Forveille, T., & Kastner, J. H. 1995, Nature, 373, 494

List of Publications

Refereed publications:

Enrico Ragusa, **Giovanni Dipierro**, Giuseppe Lodato, Guillaume Laibe and Daniel J. Price, *On the origin of horseshoes in transitional discs*, published in Monthly Notices of the Royal Astronomical Society, **464**, 1449, 2017 (**Chapter 11**)

Giovanni Dipierro, Guillaume Laibe, Daniel J. Price and Giuseppe Lodato, *Two mechanisms for dust gap opening in protoplanetary discs*, published in Monthly Notices of the Royal Astronomical Society: Letters, **459**, L1, 2016 (**Chapter 8**)

Giovanni Dipierro, Daniel J. Price, Guillaume Laibe, Kieran Hirsh, Alice Cerioli and Giuseppe Lodato, *On planet formation in HL Tau*, published in Monthly Notices of the Royal Astronomical Society: Letters, **453**, L73, 2015 (**Chapter 9**)

Giovanni Dipierro, Paola Pinilla, Giuseppe Lodato and Leonardo Testi, *Dust trapping by spiral arms in gravitationally unstable protostellar discs*, published in Monthly Notices of the Royal Astronomical Society, **451**, 974, 2015 (**Chapter 6**)

Giovanni Dipierro, Giuseppe Lodato, Leonardo Testi and Itziar de Gregorio Monsalvo, *How to detect the signatures of self-gravitating circumstellar discs with the Atacama Large Millimeter/submillimeter Array*, published in Monthly Notices of the Royal Astronomical Society, **444**, 1919, 2014 (**Chapter 5**)

Publications under review:

Giovanni Dipierro and Guillaume Laibe, *An opening criterion for dust gaps in protoplanetary discs*, submitted to Monthly Notices of the Royal Astronomical Society (**Chapter 8**)

Publications in preparation:

Giovanni Dipierro, Luca Ricci, Giuseppe Lodato, Daniel J. Price and Guillaume Laibe, *On planet formation in Elias 24*