Instabilities in Planetary Rings

Henrik Nils Latter

Trinity College

November 2006

A dissertation submitted for the degree of Doctor of Philosophy
Abstract

Saturn’s rings are among the most familiar, beautiful, and puzzling objects in the Solar system, if not all of Space. Their complex, striated structure, much like the grooves carved in a vinyl record, inspires equal degrees of aesthetic pleasure and theoretical agitation. My thesis takes this radial stratification as its theme, and examines the physical mechanisms which generate and sustain it. Specifically, I explore structure formation on the finest scales we have observed, those of about 100 m, where recent images show abundant and irregular patterns. These are thought to be the product of a pulsational instability associated with the viscous properties of the system.

In order to model the instability I attend to the subtle collective dynamics of a ‘gas’ of icy particles — dynamics that the usual tools of fluid mechanics neglect but which in this context are essential. This level of attention can only be supplied by kinetic theoretical models, which have generally been thought too mathematically involved to deploy in detailed dynamical studies. My thesis, however, presents a kinetic formalism that is both accessible and permits me to undertake the analyses necessary to understand the sophisticated behaviour of a ring of particles.

My thesis first develops the linear theory of a dilute ring, which, though not directly applicable in the Saturnian context, permits us to put in place a general framework for the later chapters. It also lets us isolate analytically the interesting effects of anisotropy and non-Newtonian stress. Once this is accomplished I outline a dense gas kinetics based on the work of Araki & Tremaine (1986) but which is much simpler and more general. The formalism is then put into use explaining the onset of the viscous overstability, where its predictions agree well with both Cassini observations and N-body sim-
ulations. In addition, some work is presented which examines its nonlinear saturation in the simple case of an isothermal two-dimensional disk.

Finally, I study the role of the viscous overstability in the excitation, or decay, of eccentricity in gaseous accretion disks. Because of the relative thickness of such disks, the behaviour of the overstability can be quite different to that in a planetary ring and its full three-dimensional character must be included. Until now, this aspect of the problem has received little detailed attention. My dissertation provides the first fully consistent linear stability analysis of the overstability in a three-dimensional disk.
# Contents

1 Introduction to Saturn’s rings 6

1.1 General properties of astrophysical disks 6

1.2 Observations 10

1.2.1 Brief history 10

1.2.2 Observational data 12

1.2.3 Experiments 23

1.3 Theories 27

1.3.1 Early history 27

1.3.2 Recent ideas 28

1.4 Modelling a particulate disk 36

1.4.1 Formalisms 37

1.4.2 Kinetic theory 40

1.5 Summary 45
Bibliographical abbreviations


Chapter 1

Introduction to Saturn’s rings

1.1 General properties of astrophysical disks

Picture a cloud of gas and rubble enveloping, and swirling about, a massive spherical compact body. The cloud will experience a flattening tendency issuing from the opposite action of two basic forces — the gravitational pull of the central body and the centrifugal force of the rotation. The central body will attract the cloud radially, while the cloud’s centrifugal force, if sufficiently strong, will block its compression perpendicular to the rotation axis. As a consequence, rotating matter that escapes falling onto the central mass will collapse onto the orbital plane. Acting against the flattening tendency will be the cloud’s pressure, or velocity dispersion, which can support the cloud against gravity in all directions. Systems in which flattening dominates pressure constitute the family of astrophysical disks and include galactic disks, protoplanetary disks, and planetary rings. Systems dominated by pressure include elliptical galaxies and stars.

Dissipative ensembles, such as planetary rings, are more likely to form disks, as interactions between cloud particles remove energy and as a consequence diminish the pressure while conserving the cloud’s total angular momentum. A ‘cooling’ flattening system will come to equilibrium when the axial gravity abates sufficiently for the (weakening) pressure to cancel it. Thus such disks are governed by three balances: centrifugal force ver-
sus the central mass’s radial gravitational force, pressure against the axial gravitational force, and heating versus cooling.

It is the last of these balances, the thermal balance, that controls the thickness of the disk, as it implicitly determines the equilibrium pressure. A strongly dissipative system will usually establish a cold thermal equilibrium, i.e. one in which the velocity dispersion is significantly less than the rotational velocity. Consequently, the disk finds its axial balance when it is very thin, because only then is the axial gravity of the central body sufficiently small to cancel the weak pressure. Planetary rings, being an aggregate of regularly colliding inelastic particles, are highly dissipative and thus both cold and thin. In stark contrast, the equilibrium shapes of certain hot accretion disks are nearer to tori.

Now suppose that the central body’s gravitational potential is slightly aspherical. In fact, imagine that, like many planets, it is oblate. In such a potential field interactions between elements of orbiting matter only conserve the component of angular momentum perpendicular to the central body’s equatorial plane. The result is a disk whose orbital plane coincides with the equatorial plane, a fact that may be observed in the planetary rings of Jupiter, Saturn, Uranus, and Neptune.

A non-spherical gravitational potential may also influence the horizontal character of the disk’s orbit in important ways. But for the moment let us assume that the central body is sufficiently spherical for this effect to be negligible compared to other processes. Subsequently, conservation of angular momentum ensures that the disk moves according to Kepler’s third law, i.e. a layer of disk at radius $r$ orbits with an angular velocity proportional to $r^{-3/2}$. Particles closer to the planet possess faster orbital velocities than those further out; hence Keplerian disks exhibit significant shear.

Disks also possess appreciable viscous stresses, though often the nature of these is difficult to pin down. The conjunction of stress and shear draws orbital energy from the mean flow and transforms it into ‘heat’ (the random motions of the disk particles with respect to their mean motion). The thermal power generated by the Saturnian rings is estimated to be about 100kW, which is quite small considering its size; yet, if there were no means to dissi-
pate this energy, the random motions, and hence pressure, would grow until the disk ‘explodes’. Planetary rings dissipate this energy via particle collisions, which are inelastic\(^1\). However, collisions do not only remove energy, they play a key role in its injection by enabling the viscous stress. They effect this in two ways. First, collisions scatter particles, which ensures their motions possess a small deviation from their circular orbits; the aggregate of these random motions may transfer appreciable momentum across the background shear flow. This mode of momentum transport occurs between collisions, and is often referred to as ‘translational’ or ‘local’. Interparticle gravitational encounters (‘gravitational scattering’) perform the same function and can be important if the velocity dispersion is comparable to the particles’ escape velocity. Second, momentum is transferred \textit{during} collisions — from the centre of one particle to the centre of the other at the speed of the particle material’s sound speed. This collisional, or ‘nonlocal’, mode of transfer is comparable to the ‘local’ kind if particle sizes are not negligible\(^2\).

So, the thermal balance for a particulate ring consists of a budget of heat injected by the stress and heat dissipated by inelastic collisions. It therefore determines a characteristic relation between collisional properties (such as the degree of inelasticity) on one hand and collision frequency on the other. This is because the collision frequency controls both the amount of energy dissipated and the magnitude of the viscous stress induced by both scattering and collisions.

The reservoir of orbital energy that is tapped by the viscous stress is not infinite, and the energy balance we sketched above holds only in a localised region and on time-scales much shorter than the time of appreciable change in the orbital energy. In fact, if this separation of scales did not exist there would be no steady thermal equilibrium. A planetary ring is, of course, related to an accretion flow and cannot last forever\(^3\). Collisions will transfer mass inward and angular momentum outward until all the disk mass has

---

\(^1\)A fraction of the kinetic energy of two colliding particles is transformed into heat via acoustic waves within them. The heat energy radiates away in the infra-red.

\(^2\)This is discussed in detail in Section 1.4.

\(^3\)It is not however a ‘free’ accretion disk, in that the gravitational torques of the near moons prevent matter from spreading in certain regions.
fallen onto the planet and all the angular momentum has escaped into space (carried by none of the mass, as it happens). The time-scale of this process is just the diffusion time, rough estimates of which can approach $10^{12}$ yr (Esposito, 1986).\(^4\)

In contrast, the orbital period of a ring particle is on the order of a day. Thus one can regard planetary rings as exceptionally ‘long-lived’, if we were to measure ‘disk life’ in terms of orbits. This is valid, in some sense, as an orbit is the time-scale of much dynamical behaviour, not least those involved in the thermal balance. In fact, planetary rings possess a large collection of such rapid interwoven processes. As Fridman & Gorkavyi (1999) point out, a thin disk is, necessarily, a complicated dynamical system. This is a result of the relative enormity of the main radial forces (arising from the planet’s gravity and the disk’s rotation) whose mutual cancellation releases into the ‘dynamic arena’ of the linearised equations a large and varied array of subdominant processes. These include viscosity, electromagnetism, and self-gravity. The interplay of these weaker mechanisms furnish complex dynamical behaviour upon a very large range of lengths and times (owing to the relative thinness of the disk). Such processes are mostly unaware of the disk’s large scale slow accretion, and their analysis does not require its inclusion.

This dissertation is concerned primarily with the rapid processes of a planetary ring, specifically those associated with viscosity and self-gravity. These may combine in such a way as to draw energy from the Keplerian shear flow and redirect it into fuelling the growth of small disturbances that can form observable structures in Saturn’s A and B-rings. The linear theory of these instabilities have previously been analysed with simple hydrodynamic models, which, though mathematically convenient, are insensitive to important aspects of the granular flow. A more sophisticated analysis should model in more detail the crucial role of the (relatively) infrequent inelastic collisions

\(^4\)For a time the viscous time-scale was incorrectly calculated as only a few million years, because the ring thickness was greatly overestimated. As a result, theoreticians really struggled with the idea that the Saturnian system was a young, transient phenomenon (Trulsen, 1972a). However, the estimate of $10^{12}$ yr is rough, as it omits angular momentum exchange with nearby satellites and the role of various instabilities in stirring up motions which may enhance viscosity.
and the (resulting) non-Newtonian behaviour of the stress, in addition to the
anisotropy of the particle velocity distribution. Through appropriate kinetic
formalisms, we provide such an analysis. This research is therefore of funda-
mental value in describing the subtle and non-trivial behaviour of the viscous
stress within the Saturnian system.

But before we get into all that we shall present a brief review of the
observational and theoretical history of Saturn’s rings. This will set the
groundwork and context for our later analyses.

1.2 Observations

1.2.1 Brief history

Saturn’s rings were discovered by Galileo in 1610, though on account of the
poor quality of his telescope he believed them to be two large satellites (or
‘branch stars’) located symmetrically on either side of the planet (GT82).
These ‘attendants’ to the aged Saturn were observed by a number of astron-
omers over the next 45 years, but no one could correctly explain them
despite clues such as the periodicity in the phases of their visibility\(^5\), as first
reported by Gevelius (FG99).

It was a young Huygens in the winter of 1655/6 who posited that the ob-
ject was in fact a symmetric ring — a thick solid structure. Though Huygens
argued that his conclusion proceeded from the superiority of his telescope,
it is now considered unlikely his instrument was much better than those
of his contemporaries (Van Helden, 1984). Rather, it is probable that re-
cent Cartesian ideas played the crucial role in the \textit{interpretation} of his data.
Particularly important was Descartes’ hypothesis that space was filled with
vortices, or disks, and that the Solar system was but one of many vortices
containing other stars; planets orbited upon stellar vortices, and in turn cre-
ated smaller vortices about themselves (Van Helden, 1984). It was not long,

\(^5\)This occurs because the angle that Saturn’s ring plane makes with Earth varies
throughout a Saturnian year; at times we view the ring from below, other times we view
it from above, and for a period in between we are edge-on and the rings ‘disappear’.
however, before the solid ring hypothesis was questioned; in fact, upon publication of Huygens’ results, the Medici court conducted a formal examination of the theory where it was postulated that the disk was composed of a multitude of small moons or ‘stars of ice’ (Van Helden, 1973). In 1675, Giovanni Domenico Cassini argued the same, following his discovery of the division (which now bears his name) between the two main Saturnian rings, the A and B (Alexander, 1980). This observation showed that the rings were not a single, rigid, opaque body, as previously conjectured. Nevertheless, the solid disk model persisted into the 19th Century, not least because of the favour granted it by the prestigious astronomer Frederick William Herschel. It was not until Keeler’s measurement of the rings’ differential rotation, in 1895, that their particulate structure was observationally confirmed (FG99), though by then the theoretical arguments of Roche, in 1848, and especially Maxwell, in 1857, had settled the issue (Van Helden, 1984).

In the intervening centuries, astronomers had revealed additional features, such as the dark interior C-ring (William and George Bond, 1850), the Encke gap, located in the A-ring (Johann-Franz Encke, 1837), and axisymmetric variations in ring brightness (William Dawes, 1851; William and George Bond, 1855). Furthermore, in the century that followed, and before the Pioneer and Voyager missions, astronomers concluded important facts about particle sizes, composition, and density from the rings’ infra-red spectrum and surface brightness properties, specifically the ‘opposition effect’ (Bobrov, 1970; Kuiper, Cruickshank and Fink, 1970), and the interesting azimuthal variability of brightness in the A-ring (Camichel, 1958).

However, the data harvested from Pioneer 11 and Voyagers 1 and 2 (which visited Saturn in 1979, 1980, and 1981 respectively) eclipsed in abundance and detail all that ground-based astronomy had accumulated to that point. In particular, these spacecraft sent back startling reports of complicated radial structure at which earlier terrestrial observations had barely hinted. The rings were certainly not broad and homogeneous as many expected (Esposito, 1986). Moreover, they discovered four new rings: the tenuous D-ring that extends to near the planet’s surface, the diffuse E-ring, which spreads itself outside the main system, the narrow and knotted F-ring, 3600 km be-
yond the A-ring, and the G-ring, a faint, narrow structure near the orbit of the moon, Mimas. A number of new moonlets were also found, nestled amidst the ring structures: Atlas, Pan, Prometheus, and Pandora, the most prominent. And the B-ring was seen mottled by shortlived ‘spokes’, which astronomers deduced were clusters of magnetized dust moving in co-rotation with the planet’s magnetic field.

This plethora of new discoveries stimulated a rush of analysis and theoretical activity in the 1980s and 1990s that saw the rapid construction, and sometimes rapid deconstruction, of a number of theoretical models. Currently we are receiving another massive dispatch of information, this time from the Cassini spacecraft, which entered Saturn’s orbit in July 2004. So it appears the imbalance between the surfeit of observation on one hand and its theoretical digestion on the other seems set to continue for the foreseeable future.

1.2.2 Observational data

The principal measurements of Saturn’s rings are of surface brightness, infrared spectrum and optical depth, which are determined from UV and radio occultation experiments and direct imaging. From these, scientists armed with quantitative theoretical models can derive estimates for the intrinsic properties of the system, such as its mass, particle composition and size distribution, velocity dispersion, collision frequency, and thickness.

Large-scale properties

The main rings (A to D) extend from a radius of roughly 67,000 km to 140,000 km, though they are estimated to be only a few tens of metres thick. They are hence razor thin — proportionally a sheet of paper is thicker (GT82). This said, the observed edge-on thickness will be larger due to tidal perturbations from Saturn and its moons, which can warp the disk or set up vertical oscillations (bending waves) with non-negligible amplitudes. Furthermore,  

---

6See Section 1.3 in FG99 for an entertaining account
Figure 1.1: A schematic diagram of Saturn’s rings and satellites and the ring crossings of the Pioneer and Voyager space probes (from www.seds.org)
the perceived thickness will be artificially inflated by the presence of a small number of large bodies (GT82).

The total mass of the rings is about $5 \times 10^{-8}$ times the mass of Saturn, which is similar to the nearby moon, Mimas. This is suggestive, and encourages theories of ring origin based on the cataclysmic destruction of a pre-existing moon. In addition to normal optical depth, we can determine the surface mass density $\sigma(r)$ in certain parts of the ring by analysing the properties of density waves (Spilker et al., 2004). Most of these are located in the A-ring, but few exist in the B, and so our understanding of the rings’ mass distribution is incomplete.

The rings are accompanied by a number of moons and moonlets. Some of these are ensconced within the narrow gaps existing in the A-ring, namely Pan, in the Encke gap, and Daphnis, in the Keeler gap (Fig. 1.7). Shepherd-ing the F-ring on either side are Prometheus and Pandora, while floating near the A-ring edge is Atlas. The former two are about 50 km in radius, while Atlas, Pan, and Daphnis are much smaller. Of the many satellites outside the F-ring, the most important to the ring dynamics are Epimetheus, Janus, Mimas, Enceladus, and Titan at radii of 151,420 km, 151,470 km, 185,000 km, 237,948 km, and 1,221,850 km respectively. Epimetheus and Janus are the ‘co-orbital satellites’ and their orbits are an example of a ‘horseshoe resonance’. Titan, though very distant, is the most massive and therefore plays some part in the ring dynamics.

**Detailed structure**

The conventional measure of surface density, at least when astronomers discuss ring structure, is *normal optical depth* ($\tau$). It is a dimensionless number that quantifies the opaqueness of a medium to radiation. In the context of particulate rings we define it through

$$\tau = \pi a^2 \int_{-\infty}^{\infty} n \, dz,$$

(1.1)
Figure 1.2: An image taken by Cassini of the dark interior C- and D-rings from beneath (from www.ciclops.lpl.arizona.edu).

where $a$ is the average radius of a particle, $n$ is the volumetric number density, and $z$ is the height above the disk midplane. Prosaically, $\tau$ is just the number of particles one would expect to find in a cylinder of cross-section $\pi a^2$ (the area shadowed by a typical particle) if it were to be plunged through the disk. Thus an optical depth of 1 corresponds to a ring with a surface number density of 1. Alternatively, $\tau$ can be defined as the total cross-section of the particles divided by the area of the ring.

The rings may be divided into large-scale structures, which are based primarily on optical depth variation. The D-ring begins just exterior to the limb of the planet, at 66,970 km, and extends to the C-ring, which starts at 74,510 km. Both rings are faint: the D-ring possesses an $\tau$ of about 0.01, and the C-ring between 0.05 and 0.35 (FG99). Nevertheless, both exhibit significant variation in the form of sharp-edged, narrow ringlets, wavelike
structures, and (in the C-ring) low optical depth ‘plateaus’ (Porco et al., 2005). In the C-ring a number of the ringlets are eccentric, a property that results from the aspherical gravitational potential of Saturn, in the case of the Maxwell ringlet, and the gravitational influence of the moonlet, Titan, in the case of the Titan ringlet (Porco et al., 1984). However, not all of the eccentric features, nor the broad plateaus, can be explained (Porco & Nicholson, 1987). Interestingly, recent Cassini images have revealed that the D-ring has evolved considerably since Voyager 2 (Hedman et al., 2005).

A puzzling feature that has been observed since the 19th Century is the sharp division between the optically thin C-ring and the optically thick B-ring (at 92,000 km). This edge is preceded by a ‘ramp’ of linearly increasing optical depth and leads into a large ‘hump’ of high optical depth on the B-ring side (see Fig. 1.2). Whilst it is true that a low order orbital resonance with a Saturnian moon could maintain such a structure none exists at this radius and so its provenance is somewhat mysterious. The B-ring itself possesses large variations in optical depth: the lowest measurements are roughly 0.5, but much of the ring is opaque. Mirroring these features are radial surface brightness variations which range over similarly vast scales — from the limits of Cassini’s resolution, about 100 m, to hundreds of kilometres (Horn & Cuzzi, 1996, Porco et al., 2005, Fig.’s 1.3 and 1.4). A small proportion of this remarkably rich and irregular structure coincides with gravitational resonances (Thiessenhusen et al., 1995), and some originates in non-dynamical effects, such as variations in albedo and phase function (see Cuzzi & Estrada, 1996, for details), but the majority does not, and is likely to be the result of collective processes (instabilities, etc). Cassini has recently shown that the densest regions exhibit disordered striations on the shortest scales (0.1 km to 1 km), while lower optical depth regions lack this fine-scale structure, though they sometimes support smooth undulations in brightness of roughly 100 km wavelength (Porco et al., 2005). Evidently B-ring structure is very sensitive to the background optical depth.

The Cassini division, like the C-ring, is a region of low optical depth with irregular structure: broad featureless plateaus and narrow ringlets specifically (Flynn & Cuzzi, 1989). Amongst, or at times superimposed upon, these are
Figure 1.3: Normal optical depth versus radius at the inner edge of the B-ring obtained from the Voyager ISS occultation data. The radial resolution is 12 km and \( R_s \) is Saturn’s radius. Significant features have been pointed out (from Durisen et al., 1992).

Figure 1.4: Profile of the optical depth of (a) the inner B-ring and (b) the outer parts of the B-ring (from Cuzzi et al. 1984). In the inner part of the ring one can discern a roughly 50 km structure, while in the outer structure there appears variation on a greater range of scales.
Figure 1.5: An image taken by Cassini of the C, B, and A-rings with the Cassini and Encke divisions (from www.ciclops.lpl.arizona.edu).
a number of spiral waves, which are presumed the result of perturbations from the nearby Prometheus, Atlas, Iapetus, and Pan moonlets (Porco et al., 2005; FG99). Mimas is responsible for maintaining the Cassini gap with a strong 1:2 orbital resonance near the outer B-ring edge.

The A-ring, extending from 122,170 km to 136,780 km, supports a great variety of structure caused by assorted gravitational interactions with a clutch of nearby moons and moonlets. Low order resonances, with Mimas and Janus especially, exert torques on ring matter which excite spiral wave trains of wavelengths up to tens of kilometres (FG99). These are categorized as either density waves or bending waves; the former oscillate in, and the latter oscillate out of, the ring plane (Fig. 1.6). Other phenomena include ‘corduroy’ patterns, or ‘wakes’, caused by the gravitation of the moonlet Pan (which is orbiting within the nearby Encke gap), as well as kilometre sized, quasi-parallel variations in brightness labeled ‘straw’, and slightly larger ‘ropy’ features. These manifest themselves in the troughs of strong density waves, or wakes, and may be caused by an enhancement of self-gravitation in the crests of these oscillations. In addition, Cassini has resolved ‘mottled’ structures on the A-ring edge which have yet to find an explanation (Porco et al., 2005).

Interior of the A-ring by some 250 km sits the narrow (∼42 km) Keeler gap, which possesses similar properties to the Encke gap (which lies closer to Saturn at a radius of 133,600 km). These gaps are maintained against viscous spreading by the moonlets Daphnis and Pan, respectively. The passage of these moonlets cause the scalloped inner ring edges, and also draw out bright streamers, or spikes, of ring material into the gap (Fig. 1.7).

The structure of the eccentric F-ring is remarkably complex (see Fig. 1.8). Consisting of three narrow ‘braided’ strands of ring material and a diffuse 700 km wide sheath, it exhibits knots, travelling kinks and other local and short-term fluctuations (Porco et al., 2005). This behaviour is tied closely to the neighbouring shepherd moon, Prometheus, which draws matter away at periodic intervals (roughly when the moonlet reaches apoapse). These form trailing evanescent ‘drapes’. However the gravitational role of Prometheus cannot explain all the complicated structure and dynamic behaviour exhib-
Figure 1.6: An image taken by Cassini of the Prometheus 12:11 density wave (lower left of the image) and the Mimas 5:3 bending wave (the middle of the image). The pixel scale of this image is about 290 metres/pixel (from www.ciclops.lpl.arizona.edu).

Figure 1.7: An image taken by Cassini of Daphnis in the Keeler gap. The wavy (scalloped) edges it induces can be clearly seen (from www.ciclops.lpl.arizona.edu).
Beyond the main sequence of rings lie two extremely tenuous dusty belts: the G and E-rings. The former is located between the orbits of the co-rotational satellites, Janus and Epimetheus, and the orbit of Mimas and is only some 8000 km in width. The E-ring in contrast extends from Mimas all the way to Rhea, gradually growing to a thickness of some tens of thousands of kilometres, though it is densest at the orbit of Enceladus. E-ring dust is roughly between 0.3 to 3 µm in size and is thought to mainly consist of material thrown up by the impact of micrometeroids upon the icy satellites it encompasses — Enceladus, Tethys, Dione, and Rhea (see Fig. 1.1). However, Cassini has recently detected evidence of immense geisers on the South pole of Enceladus which provides an additional source of icy dust (Spahn et al., 2006). The geophysical mechanism for this activity is unclear but could result from the sublimation of water ice beneath the surface (Spencer et al., 2006; Kargel, 2006).

Quite recently, in September 2006, a new extremely faint ring was resolved by Cassini at the orbits of Janus and Epimetheus. Like the E and G-rings it is postulated to consist of icy dust kicked off the nearby moons by micrometeoroids.

Particle composition and size distribution

The Saturnian system comprises trillions of icy grains ranging in radii from a few centimetres to a few metres according to a power-law distribution. Particle composition has been determined from high-frequency measurements of their infra-red spectrum (Kuiper, Cruikshank, and Fink, 1970; Cuzzi et al., 1984); and later spectrometric and photopolarimetric data show that they are coated in a thin frost (Pollack, 1978; Steigmann, 1984). However, reflectivity data suggest, and microwave data confirm, that a small proportion (1–10%) of the ring mass is non-icy in nature, and this explains their faint ‘salmon’ colour. Moreover, ring colour appears to vary on large scales — in particular, lower optical depth regions, such as the C-ring and Cassini gap, are ‘less red’

---

than the A and B-rings (Estrada & Cuzzi, 1996, and references therein). Additionally, matter in lower optical depth regions is consistently darker. It has been postulated that this variation is the manifestation of pollution by interplanetary dust: less opaque areas are more susceptible to pollution and will have reduced particle albedos (Cuzzi & Estrada, 1998).

Size distributions can be deduced from occultation experiments (Marouf et al., 1983; Zebker et al., 1985; Showalter & Nicholson, 1990; French & Nicholson, 2000). As with compositions, these distributions vary throughout the rings, though a power law fit can be applied successfully in each region. The exponent of the power law is approximately $-3$, with a cut-off at approximately 5-10 m. The distribution of particle sizes above this limit is poorly constrained, though recent observations of ‘propellors’, presumably caused by the gravity of ‘skyscraper’-sized objects, reveal that there exist a not-inconsiderable population of particles between metre and moonlet sizes (Tiscareno et al., 2006).
Generally, however, the power law indicates that there are many small bodies and few large ones, but with most of the ring mass in the latter. Particles inhabiting the A-ring are larger, followed by the B-ring, while the C-ring is comprised of significantly smaller particles. In each case, the effective particle radius (the equivalent radius if the rings were monodisperse) is roughly 10 m, 7 m, and 2 m according to Showalter & Nicholson (1990). Esposito (1986) points out that particle sizes are smaller in regions more vigorously ‘stirred’ by moonlet gravity (for example, the eccentric ringlets in the C-ring and the spiral wave trains in the outer A-ring). He concludes that ring particles are brittle: a small increase in velocity dispersion (and hence average impact velocity) corresponds to a large decrease in particle sizes, i.e. particles are very vulnerable to fracturing in collisions. If correct, such a view is in tune with the idea of particles as loose and transient agglomerations of ice (‘dynamic ephemeral bodies’), rather than rigid, hard blocks (Weidenreich et al., 1984). Certainly this may be the case in the A-ring where tidal forces are weaker and gravitationally bound aggregates more likely to be stable. That said, it is difficult to know exactly the nature of Saturn’s ring particles, none ever having been directly observed.

Ring particles collide inelastically at rates which are of the same order of magnitude as the orbital frequency (Stewart et al., 1984). In very dense regions, such as the high optical depth regions of the B-ring, the collision frequency per orbit may be much higher. Particle collisions are very gentle; an upper bound on their rms speed is 0.2 cm s\(^{-1}\) (Weidenschilling et al., 1984). This velocity scale is of the same magnitude as a typical particle’s escape velocity, and also the difference in mean velocity across a particle diameter; thus gravitational encounters and nonlocal viscosity effects should be important. In addition, the adhesive effect of frost may play a role.

### 1.2.3 Experiments

Parallel to the interpretation of Voyager data, a number of laboratories, those of Santa Cruz principally, have conducted experiments in order to ascertain the collisional properties of ice spheres in Saturnian conditions. They have
focused especially on the relationship between the normal coefficient of restitution $\varepsilon$ and normal impact velocity $v_n$ (see Bridges et al., 1984; Hatzes et al., 1988; Supulver et al., 1995; Dilley & Crawford, 1996). The normal coefficient of restitution is a simplistic but convenient measure of the inelasticity of particle collisions, being the ratio of the normal relative speed after and before a collision. It is generally a function of normal impact velocity $v_n$ and possibly other parameters like ambient temperature and pressure, particle size and mass (Hatzes et al., 1988; Dilley, 1993).

These experiments usually consist of a cryostat in which a block of stationary ice is struck by a small ice sphere attached to a pendulum. Velocities immediately before and after the collision are measured by various means: laser beams, capacitive displacement on the metal pendulum, or video camera (Bridges et al., 1984, Hatzes et al., 1988, and Dilley & Crawford, 1996, respectively).

Excepting Dilley, these studies successfully fit a step-wise power law to their data for collisions sufficiently gentle and/or surfaces sufficiently frosted:

$$\varepsilon(v_n) = \begin{cases} (v_n/v_c)^{-p}, & \text{for } v_n > v_c, \\ 1, & \text{for } v_n \leq v_c, \end{cases} \quad (1.2)$$

where $v_c$ and $p$ are parameters contingent on the material properties of the ice balls and their environment. This is plotted in Fig. 1.9a. Bridges’ data admit $p = 0.234$ and $v_c = 0.0077 \text{ cm s}^{-1}$ for frosted particles of radius 2.5 cm at atmospheric pressure and at a temperature of 210 K (significantly higher than the appropriate conditions). Hatzes finds $p = 0.20$ and $v_c = 0.025 \text{ cm s}^{-1}$ for the case of frosted particles at 123 K at pressures as low as $10^{-5}$ torr (however, for the case of smoother particles an exponential law provides a better fit). At slightly lower temperatures ($\approx 100$ K) and at atmospheric pressure Supulver obtains $p = 0.19$ and $v_c = 0.029 \text{ cm s}^{-1}$ with a fixed torsional pendulum. Hatzes reports that there is little change in $\varepsilon$ with pressure.

All these studies reveal that very gentle ice collisions can be remarkably dissipative, which has profound implications for ring energetics. Addition-
ally, they find that the coefficient of restitution varies considerably as the physical condition of the contact surface is more or less frosty or sublimated. Hatzes and coworkers showed that after a few very dissipative collisions (of constant $v_n$) the coefficient of restitution approached an asymptotic value, corresponding to an ice surface of compacted frost. This suggests that the compactification of regolith at typical impact velocities should be important. A layer of compacted frost may mitigate collisional erosion because it can buffer the icy nucleus of the particles (Weidenschilling et al., 1984). There also exists some experimental research on glancing collisions and the functional form of the tangential coefficient of restitution (Supulver, Bridges & Lin, 1995).

A number of theoretical studies have obtained an expression for $\varepsilon$ by modelling an ice particle as a viscoelastic solid (Gorkavyi, 1985; Dilley, 1993; Spahn, Hertzsch & Brilliantov, 1995). The predictions of these models all seem consistent with the experimental data, though notable is the fact that Dilley’s $\varepsilon$ depends substantially on particle mass and size, and Spahn et al.’s model also determines the tangential coefficient of restitution.

There may be two more collisional regimes which are relevant: that of fracturing (if $v_n$ exceeds a critical value) and that of adhesion (if $v_n$ is less than a critical value and there exists sufficient surface regolith). In the first case the kinetic energy of the collision is so large that the nucleus of the particle deforms or shatters irreversibly. In the second, the kinetic energy is used up compressing the loose surface frost. More specifically, freshly frosted surfaces consist of a jagged interlace of micrometre ice ‘whiskers’ and gentle collisions between two such surfaces may allow this complicated structure to mesh, much like ‘Velcro’ (Hatzes et al., 1991). Experimental studies of the process of frost formation and the adhesive properties of ice have been undertaken by Hatzes et al. (1991) and Supulver et al. (1997). High velocity ice impacts have been experimentally analysed by Higa, Arakawa & Maeno (1996, 1998).

Some, but not much, theoretical work exists which deals with adhesion effects at low impact velocities, and how this modifies the restitution coefficient (see Brilliantov & Pöschel, 2004b, and references therein). Heuristically,
Figure 1.9: The coefficient of restitution $\varepsilon$ as a function of impact velocity $v_n$ for (a) the piecewise power law of (1.2) and (b) adhesive (frost coated) ice particles.

for an adhesive particle, there must exist a characteristic velocity ($v_1$) below which $\varepsilon = 0$, and another velocity ($v_2$) at which $\varepsilon$ possesses a turning point. For $v_n > v_2$ we might expect $\varepsilon$ to behave similarly to the power law sketched in Eq. (1.2). This behaviour is plotted schematically in Fig. 1.9b.

A number of theoretical estimates have been proposed for the mass erosion rate of ice collisions (GT78; Gorkavyi, 1985; Borderies et al., 1984; Longaretti, 1989), though not all of these appear to be consistent with experiments (Hartmann, 1978, 1985; FG99). Furthermore, Longaretti (1989) and Weidenschilling et al. (1984) have built dynamical models for the mass erosion and re-accretion processes, and have obtained equilibrium mass distributions roughly consistent with those observed. Their models incorporate either tidal or collisional erosion on one hand and gravitational accretion on the other; omitted effects include particle sticking, collisional mass transfer, and regolith compactification, all of which the later Santa Cruz experiments show to be important.
1.3 Theories

1.3.1 Early history

The first theoretical investigations of ring dynamics were concerned chiefly with large-scale structure and stability. Laplace in the late 18th Century was the first to seriously tackle the problem. He showed that the tidal forces exerted on a rigid solid ring (the model suggested by Huygens a century earlier) were too extreme for known materials to withstand, and thus suggested that the rings were composed of a sequence of concentric, solid ringlets, each too narrow to be disrupted by Saturn’s tide. But Laplace also showed that such an arrangement was unstable because the potential energy of each ringlet possesses a maximum when it is centred on the planet (FG99, GT82). His solution was to claim that the ringlets were solid but of an unknown character which could render them immune to his stability analysis.

It was Maxwell in 1859 who offered detailed analyses of incompressible fluid and particulate disks, as well as of Laplace’s concentric solid ringlets (Laplace never considered a particulate disk). He showed that a solid ringlet could only remain stable if its mass distribution was very nonuniform, specifically if all its mass was focused in one small region, a situation which seems implausible. He also claimed a disk composed of incompressible fluid is unstable, though his discussion is not correct. His conclusion that the rings were particulate is right, of course, though perhaps not completely justified by his analysis (for further details see Cook & Franklin, 1964).

It was Jeffreys in 1947 who dispensed with the gaseous and liquid models. He argued that a liquid disk would reflect the planet (which it does not) and a gaseous disk would be far thicker than the observations show it to be (GT82). Jeffreys also proved that small satellites with an appreciable tensile strength could withstand tidal forces within the Roche lobe, and thus strengthened the argument for a particulate ring.
1.3.2 Recent ideas

The Voyager data, with their dramatic and unexpected panorama of irregular and varied phenomena, have excited and confounded theoreticians from the 1980s until the present day. While studies predating the Voyager mission focused on broad-brush questions, theoretical work since has concentrated on the the new fine-scale details it has revealed — not that these are divorced from the larger questions. For all their (assumed) age, Saturn’s rings exhibit, in their complexity and irregularity, what seems to be significant dynamical immaturity, and appear (naively) to have not yet settled into a steady state (Esposito, 1986). If nothing else, viscous diffusion should smooth away structure on fine scales ($\Delta r$) in a time $t_\nu = (\Delta r)^2/\nu$, which should be very short for $\Delta r \ll r$ (where $r$ is radius and $\nu$ is viscosity). This is important, because if we claim that the Saturnian system is old, then we have to find processes that could generate and maintain on long time-scales the complicated phenomena we see.

It is now believed that a host of physical mechanisms are responsible for ring structure, each operating in different regions and/or on different scales. Also, it is agreed that much of the irregular structure (especially in the B-ring) stems from a collective dynamics — from the rings’ intrinsic self organisation. The rings are not, as some previously held, merely a passive accumulation of rubble sculpted by the gravitational torques of moons (FG99; Tremaine, 2003). In particular, axisymmetric instabilities are regarded as the most likely culprits of irregular stratification, their nonlinear evolution presumably leading to a saturated state exhibiting the same radial fine structure as that observed.

The following subsections present a brief survey of the principal mechanisms and theories that have been advanced in the quest to explain structure in Saturn’s rings. We will choose however to concentrate on those dealing with the B-ring, given that is the focus of the dissertation. Also, we will not comment here on the (massive) subject of moonlet-ring interactions, as we do not draw on it in later chapters. If readers are interested in this topic we refer them to Goldreich & Tremaine 1978b, 1979, 1980; and Shu et al., 1983,
Jetstreams and inelastic collapse

In the 1970s, the Scandinavian theorists, Alfvén, Arrhenius, and Trulsen, argued that a collection of orbiting bodies undergoing inelastic collisions should collapse into a narrow stream with correlated orbital elements (Alfvén, 1970; Alfvén & Arrhenius 1976; Trulsen 1972b). Inelasticity is essential to the mechanism they propose. In an elastic collision energy is directed from one dimension to others by a change in relative momentum, a process that leads to a ‘spreading’ of the distribution in phase space. But an inelastic collision decreases the relative momentum, and locally the distribution function contracts. Put more concretely, an inelastic collision diminishes the normal component of the relative velocity but conserves the tangential component, and hence the motion of two particles which have recently collided are aligned. This permits velocity correlations to develop on small length-scales, and thus, particles can drift into similar orbits.

Numerical simulations, and some analytical work, have indeed exhibited the formation of jetstreams (Trulsen, 1972b; Baxter & Thompson, 1973), but these required collisions to be unrealistically dissipative (Stewart et al., 1984). Moreover, the coefficient of restitution (or its surrogate) was assumed constant, which is a poor approximation especially as an ε dependent on impact velocity should mitigate this effect (by analogy with the force-free case, BP04). In addition, when clustering has reached a stage when random velocities are small, we would expect the Keplerian shear across a particle diameter to scatter colliding particles in such a way to counteract the clustering.

The jetstream mechanism is no longer considered a realistic generator of fine structure in Saturn’s rings, but it is worth mentioning because of its relationship to inelastic collapse and the clustering instabilities which manifest in dissipative granular gases, phenomena which have enjoyed much attention of late (see BP04, and references therein).

In fact, this idea was presaged crudely by Kant 220 years earlier, who was the first to predict the existence of multiple ringlets (FG99).
Meteoric bombardment

The particles that make up the rings of Saturn are subject to continuous bombardment by interplanetary projectiles which issue from comets, the Oort cloud, and possibly other sources (Cook & Franklin, 1970; Morfill et al., 1983; Ip, 1983; Durisen, 1984; Cuzzi & Durisen, 1990). Saturn can gravitationally focus a significantly larger stream of meteroids than would otherwise be the case. The mass flux incident on the rings has been estimated to be as high as $2.2 \times 10^6$ g s$^{-1}$ (Morfill et al., 1983) and at such a rate significant dynamical consequences follow.

The most drastic of these is ring erosion by vaporisation or the loss of ejecta dispersed by the impacts. Initial conservative estimates put the total mass eroded per unit area per year as $5 \times 10^{-7}$ g cm$^{-2}$ yr$^{-1}$. Ring mass density is of the order of tens of grams per square centimetre (Cuzzi et al., 1984) and so the rings should be eroded away on a time-scale much shorter than that of the Solar system. To avoid this conclusion it has been postulated that a substantial amount of ejecta is recycled back onto the ring, and that a sizable portion of meteoroid mass remains in the disk (Cook & Franklin, 1970; Cuzzi & Durisen, 1990).

But this opens the door to a number of other effects, which we can characterise as either ‘direct’ or due to ‘ballistic transport’ (Durisen et al., 1996). Direct effects include changes to surface density and specific angular momentum due to the deposition of meteoroid mass. The deposition of angular momentum is by far the more important effect (especially if meteoroids are absorbed asymmetrically) and gives rise to radial secular drifts. These global motions have been calculated and suggest that much of Saturn’s rings will fall into the planet in less than a Solar system lifetime. In particular, the drift time-scale of the C-ring is only of order $10^8$ yr (Durisen et al., 1996). This problem has yet to be resolved, and we may conclude that some ring regions (particularly the C-ring) are young (Esposito, 1986).

Ballistic transport refers to the carriage of mass and angular momentum between different radii by the ejecta thrown up by meteoroid impacts. This matter, post-collision, will remain in orbit for some characteristic time before
reaccreting at a different location, usually at a different radius. The net effect of this transport mechanism can be significant, especially in regions where there are pre-existing optical depth variations on length-scales comparable to the radial distances traveled by the ejecta (Durisen, 1990). Durisen and coworkers have shown that the ballistic transport mechanism can maintain sharp inner edges against viscous spreading, such as that between the B and C-rings, and also build up the ‘ramp’ and ‘hump’ structures on either side of the edge, as we observe in Fig. 1.3 (Durisen et al., 1992). It also can create large undulatory structures on length-scales of order 100 km near a ring edge that may be linearly unstable. This is because high density regions tend to absorb more of the ejecta than neighbouring less dense regions. The unstable oscillations grow slowly, on time-scales of millions of years (Durisen, 1995).

**Electromagnetic effects**

Amongst the ejecta thrown up by an impacting meteoroid may be a quantity of charged, sub-micrometre dust. A substantial proportion of this, like the larger ejecta, will eventually fall back upon the ring. But, unlike larger particles, they will be subject to electromagnetic forces induced by their motion across the planet’s dipolar magnetic field. The magnitude of these forces will be substantial, because the dust particles have a large charge to mass ratio, and they will tend to force particles into co-rotation with the planetary field (Goertz et al., 1986). When these particles are reabsorbed the ring will feel a torque proportional to the dust flux: matter located outside the co-rotation radius will drift outward and matter located inside will drift inward. The time-scale for this global drift has been estimated to be roughly $5 \times 10^9$ yr (Goertz et al., 1986).

Angular momentum coupling between the disk and the planet’s magnetic field can also excite growing waves. The mechanism of instability is analogous to that of ballistic transport (Goertz & Morfill, 1988), and, similarly, may lead to growing fluctuations with wavelengths of several hundred kilometres, though this depends closely on the distribution of radial distances the dust particles travel. The basic idea is that charged dust elevated above
an ‘overdense’ region will move a characteristics radial distance $\Delta$ because of Saturn’s magnetic field. If this dust settles onto an ‘underdense’ region then the resulting torque will gently push matter in this region radially. But dust ejected from underdense regions settling on overdense regions will produce a smaller radial motion, because of the region’s greater mass and also because it is assumed that dilute areas will eject less dust than denser areas. Thus it is possible that material may move from the rarefied regions to dense regions. Working against this tendency are viscous diffusion and ‘gardening’, i.e. the fact that dust transport itself smooths out gradients by removing matter preferentially from dense to dilute regions. Goertz & Morfill (1988) show that sufficiently long waves are rendered unstable, though the fastest growing wavelengths are (naturally) of the order of $\Delta$. Growth rates are relatively large and are proportional to $(k\Delta)^2 \times 10^{-4} \text{ yr}^{-1}$, where $k$ is radial wavenumber. A simple model for particle ‘hopping’ gives $\Delta \sim 10,000 \text{ km}$ though the actual distance depends closely on the charge to mass ration $Q/m$ of the dust (Goertz et al., 1986). It is quite possible that irregular large-scale structure in the B-ring represents the nonlinear saturation of this instability (Shan & Goertz, 1991).

**Phase changes**

Tremaine has suggested that particles in dense sections of the B-ring are so closely packed that the cohesive forces between particles will locally ‘freeze’ the ring into a ‘solid’ (Araki & Tremaine, 1986; Tremaine 2003). If true, then we should think of dense portions of the ring as divided into ringlets of ‘solid’ and ‘fluid’, and thus of rigid body and differential rotation. Observed structure in this case results from variations in shear, not variations in surface density.

The hypothesis relies on the stability of such aggregates to the disruptions caused by tidal forces and impacting particles. Acting against these effects is the yield stress of the assembly which arises primarily from sticking forces between particles. However, no rigorous stability proof yet has been offered. (Maxwell’s analysis, which disposed of Laplace’s ringlets, may fail
for a weakly-bound ‘rubble pile’.) Moreover, it is unclear whether the weak cohesive forces between particles could ever be sufficiently strong, or indeed whether very dense ensembles in Keplerian rotation could ever be susceptible to the required ‘freezing’ instability.

**Self-gravity**

An inviscid fluid disk possesses an instability analogous to the classical Jeans instability, through which a gas cloud collapses under the action of self-gravity. In the case of a disk, unstable axisymmetric modes can grow on a band of intermediate wavelengths, with rotation and pressure stabilising the long and short scales respectively (Toomre, 1964; Julian & Toomre, 1966). The criterion for instability is $Q < 1$ in which appears the ‘Toomre’ parameter $Q$, a quantity that measures the relative strength of disk pressure and rotation on one hand against gravitational attraction on the other. The denser and colder the system, and the slower it orbits, the lower the $Q$.

The Toomre parameter varies throughout Saturn’s rings and its value is uncertain, primarily because the variously sized populations probably possess different velocity dispersions. It may be the case that a stable subpopulation of particles can stabilize those populations which are Toomre unstable. However, the C-ring at least should be stable given its rapid rotation. In other parts of the ring $Q$ is estimated a little above unity, which may suggest a self-regulating mechanism is in place analogous to that operating in galactic disks. In this scenario the disk’s velocity dispersion is maintained just on the margins of instability. If it ever falls below, gravitational instability will intervene and heat the disk until $Q > 1$.

A *viscous* disk, however, paints a slightly more complex picture. Then it is better to think of self-gravity as ‘extending’ the viscous instabilities, which we discuss later, into larger areas of parameter space. In these ‘extensions’ the viscous instabilities grow only in a certain confined range of intermediate wavelengths, unlike the non-self-gravitating cases in which the longest wavelengths are the first to become unstable.

Self-gravitating $N$-body simulations reveal transient, non-axisymmetric
density wakes (Salo, 1992a; Daisaka & Ida, 1999; Tanaka et al., 2003). These structures form on length-scales of a few hundred metres for parameter regimes associated with the A and B-rings, and are generally thought to be analogous to the local trailing wavelets investigated by Julian & Toomre (1966) in galactic disks. Observations of azimuthal brightness variations in the A-ring may be attributable to such formations (Salo, 1992a). Recently Griv and coworkers (Griv et al., 2000; Griv & Gedalin, 2003) have presented a kinetic theory which interprets the wake structures as global, tightly wound spiral waves, but it is unclear if this conclusion is justified, as their model neglects crucial dense gas effects, and \( N \)-body simulations have yet to verify the theory.

Simulations have also described the collapse of wakes into particle aggregates as suggested by Weidenschilling et al. (1984), but these appear only in the outer portions of the A-ring, near the Roche radius (Salo, 1992a). Here tidal forces are weaker and more is possible in the way of gravitational accretion of particles. Because these aggregates are full of voids their density is low and after an initial growth phase additional particles can no longer accumulate.

**Viscous instabilities**

Finally we turn to the local instabilities of viscous fluid disks upon which this dissertation concentrates. These instabilities function by tapping the energy the viscous stress draws from the differential rotation: if the stress varies in an appropriate way with surface density, then a small amplitude wave or inhomogeneity may be excited and grow.

The ‘viscous instability’ was first put forward as a cause of fluctuations in the radiation of accretion disks (Lightman & Eardley, 1974). But in the early 80s it stimulated enormous interest in planetary ring circles and quickly became a popular candidate for the newly discovered fine structure in Saturn’s rings (Lin & Bodenheimer, 1981; Ward, 1981; Lunkari, 1981). Essentially a monotonic ‘clumping’, the viscous instability is associated with an outward angular momentum flux which decreases with surface density: \( d(\nu \sigma)/d\sigma < 0 \)
(where $\nu$ is kinematic viscosity and $\sigma$ surface mass density). Physically, a small localised increase in density leads to a decrease in radial angular momentum flux in that area. Consequently, angular momentum will build up on the inside border of the density clump and decrease on the outside border. The material with greater angular momentum will move outward and the matter with less will move inward; thus mass will accumulate in the higher density region and deplete the lower density regions, and the gradient will be exacerbated. The longest length-scales are the most susceptible to this runaway process, though they will grow slowly because the growth rate of the clumping is proportional to $k^2$, where $k$ is radial wavenumber. For sufficiently small wavelengths, pressure extinguishes the instability, and so there is a preferential intermediate scale on which the viscous instability grows most vigorously.

A dilute ring’s viscosity depends on surface density in a manner that promises the existence of the instability (GT78, SS85), but Saturn’s rings are most likely ‘dense’, and theoretical and numerical $N$-body studies have since revealed that such rings do not manifest the appropriate viscosity for this instability to develop (Araki & Tremaine, 1986; Wisdom & Tremaine, 1988). As a result the viscous instability has been all but abandoned as a generator of ring structure.

Like the viscous instability, the ‘viscous overstability’ was first examined in the context of accretion disks (Kato, 1978), and, as the name suggests, originates in an overcompensation by the system’s restoring forces: the stress oscillation which accompanies the epicyclic response in an acoustic-inertial wave will force the system back to equilibrium so strongly that it will ‘overshoot’. The mechanism relies on:

a) the synchronisation of the viscous stress’s oscillations with those of density,

b) the viscous stress increasing sufficiently in the compressed phase.

In hydrodynamics only the latter consideration is relevant, which furnishes the criterion for overstability: $\beta \equiv (d \ln \nu / d \ln \sigma) > \beta^*$, where $\beta^*$ is a number dependent on the thermal properties of the ring (Schmit & Tscharnuter,
1995). More generally, a non-Newtonian stress may oscillate out of phase with the shear, which would check the mechanism.

The axisymmetric viscous overstability has been a favoured explanation for smaller-scale B-ring structure in recent years, a status stemming, primarily, from the viscosity profiles computed by Araki & Tremaine's (1986) dense gas model and Wisdom & Tremaine's (1988) particle simulations. Both appear to satisfy the above criterion. Consequently, the linear behaviour of the instability has been thoroughly examined, though only within a hydrodynamic framework (Schmit & Tscharnuter, 1995; Spahn et al., 2000; Schmidt et al., 2001). In addition, Schmidt & Salo (2003) have constructed a weakly nonlinear theory, and the overstability’s long-term, nonlinear behaviour has been numerically studied by Schmit & Tscharnuter (1999). An isothermal model was adopted in both cases. Recently $N$-body simulations have exhibited the viscous overstability on length-scales of 100-200m (Salo et al., 2001). These results would suggest that the finest scale structure observed in Saturn’s B-ring is possibly caused by this instability.

The viscous overstability has also enjoyed much attention in its non-axisymmetric guise. It is then associated with the evolution of the $m = 1$ global mode which controls the growth or decay of eccentricity in dense narrow rings (Borderies, Goldreich & Tremaine, 1985; Papaloizou & Lin, 1988; Longaretti & Rappaport, 1995). This mechanism may cause some of the eccentric features observed not only in the D and C-rings and the Cassini gap, but also in the Uranian system.

### 1.4 Modelling a particulate disk

Having discussed observations and theories generally, we shall begin to concentrate on the particular phenomena this dissertation hopes to tackle and the theoretical tack it shall take.

We seek to describe the collective dynamics of a particulate ring in the absence of external forces (other than Saturn’s gravity, of course), such as moonlets, meteoroids, and magnetic fields. In particular, we shall examine in some detail the properties of the viscous instabilities discussed in the
preceding section. These benefit from a treatment more attuned to the particularities of granular flow, and so we have mainly employed a kinetic gas theory over hydrodynamics, but not gone as far as $N$-body simulations. Such a formalism permits us to examine in a direct and more accurate way important processes such as collisions, their inelasticity, the viscous stress induced by the particles' random motion, and the viscous stress arising from collisional momentum transport. These processes have important implications for the development of the viscous instabilities which should not be omitted.

1.4.1 Formalisms

We briefly describe the approaches which have been employed in the modelling of planetary rings, of which there are quite a few. The principal ones include $N$-body simulations, continuum models (hydrodynamic and kinetic), and generalisations of celestial mechanics. This plurality is indicative, I believe, of how the field of planetary ring dynamics falls uncomfortably between the more familiar frameworks of classical physics. It is also provides striking evidence of the great variety of dynamical behaviour planetary rings exhibit.

A popular and powerful method derived from celestial mechanics is the ‘two-streamline approach’ which monitors the evolution of the parameters that characterise the geometric orbits of fluid elements. Orbits may be described by

$$r = \overline{a}\{1 - \overline{e}\cos(m\theta + m\Delta)\},$$

where $r$ is radius, $m$ an azimuthal wavenumber, $\Delta$ is a phase angle, $\theta$ is longitude, and $\overline{a}$ and $\overline{e}$ are the semi-major axis and eccentricity of the ring fluid particles. The mean orbital elements sketched above will generally change over time as a result of the streamline’s interaction with adjacent streamlines via viscous forces, or with the central planet, a moonlet, or the rest of the ring via gravity. This approach has been utilized primarily to describe the evolution of density spiral waves excited by external gravitational torques (Borderies, Goldreich & Tremaine, 1986; Longaretti & Borderies, 1986), the maintenance of sharp edges by a shepherd moonlet (Borderies, Goldreich & Tremaine, 1982, 1989; Tremaine, Rappaport & Sicardy, 1995), the influence
of an embedded moonlet (Spahn, Scholl & Hertzsch, 1994), and eccentricity growth due to viscous overstability (Borderies, Goldreich & Tremaine, 1985; Longaretti & Rappaport, 1995). More general explications of the method can be found in Borderies, Goldreich & Tremaine (1983), Borderies & Longaretti (1987), and Longaretti & Borderies (1991). The streamline approach is well suited to the task of modelling global non-axisymmetric structures and disk–moon interactions, particularly if the moons travel on elliptical orbits. However, these methods lend themselves less naturally to a detailed examination of the viscous properties of equilibrium states and their stability to axisymmetric perturbations.

\[ -body computations are another useful tool with which one can ‘reproduce’ the behaviour of a small area of a particulate ring. Such simulations usually distribute \( N \) particles in a periodic box and then determine the trajectory of each from the equation of motion: for instance, the \( i \)’th particle moves according to

\[
m_i \frac{d^2 \mathbf{r}_i}{dt^2} = \mathbf{F}_i(\mathbf{r}_1, \mathbf{v}_1, \ldots, \mathbf{r}_N, \mathbf{v}_N, t)
\]

where \( m_i \) is the \( i \)’th particles’ mass, \( \mathbf{r}_i \) its position, \( \mathbf{v}_i \) its velocity, and \( \mathbf{F}_i \) is the force it experiences at a given time due to its Keplerian rotation, its gravitational interactions with moonlets, and its gravitational and collisional interactions with other ring particles. The problem then boils down to integrating numerically \( N \) coupled ODE’s. If self-gravity is excluded the particle trajectories between collisions can be solved analytically, in which case most of the work lies in figuring out the ordering of the collisions and postcollisional velocities. If self-gravity is included the trajectories between collisions must be numerically computed but a number of techniques can be applied which simplify the task (see Salo, 1995).

The \( N \)-body approach provides a useful picture of the properties of an equilibrium state, and some insight into its stability (Brahic, 1977; Wisdom & Tremaine, 1988; Salo, 1991, 1992a, 1995; Mosquiera, 1996; Daisaka & Ida, 1999; Salo et al., 2001), as well as the role of particle spin and size distribution (Salo, 1992b; Morishima & Salo, 2006; Ohtsuki, 2006b). Simulations of
an embedded moonlet predicted the recently observed propeller structures (Spahn, Scholl & Hertzsch, 1994; Hertzsch et al. 1997), while others have simulated the perturbations of a nearby moon and subsequently illuminated the launching of density waves (Hänninen & Salo, 1992, 1994, 1995) and structure formation at ring edges, especially at the F-ring (Hänninen & Salo, 1994; Lewis & Stewart, 2000, 2005; Murray et al., 2005).

However, in order to track structure formation of the kind displayed by Saturn’s rings one needs to simulate the system for thousands or tens of thousands of orbital periods in a box whose size is of the order of kilometres at least, filled with an appropriate number of particles. Also, the longer the length-scale the longer the time required, especially if we seek to resolve the evolution of the viscous instabilities which possess growth rates $\propto 1/\lambda^2$. The largest computations so far have been undertaken in a box of length-scale a few hundred metres for some 3000 orbits (see, for example, Salo et al., 2001; Schmidt & Salo, 2003). Plainly the box sizes employed in most simulations are only large enough to capture structure formation on the smallest of scales, and only for the initial stage of their evolution. Thus computational constraints argue for the continued use of continuum models, particularly in the self-gravitating case.

Of these, theoreticians have used second-order kinetic models solely to solve for the equilibrium state (GT78; SS85; Shukhman, 1984; Araki & Tremaine 1986; Araki, 1988, 1991) and to study density waves (Shu, 1985; Borderies, Goldreich & Tremaine, 1983, 1986). Linear stability calculations are typically left to hydrodynamics, which sometimes employs the transport coefficients computed from the kinetic theoretical steady states (for details, see Lin & Bodenheimer, 1981; Ward, 1981; Stewart et al., 1984) or $N$-body simulations (Schmidt et al., 2001).

However, the adoption of the Navier-Stokes stress model introduces two assumptions that may be inappropriate in the ring context and whose consequences are instructive to investigate. Firstly, the Navier-Stokes model presumes the particles’ velocity dispersion to be nearly isotropic. In the regime of many collisions per orbit this is an acceptable supposition, as collisions scatter particles randomly on the average. However, if the collision rate
$\omega_c$ is of the same order as the orbital frequency $\Omega$ (as it is presumed to be in Saturn’s rings) this need not be true. Secondly, hydrodynamics assumes an ‘instantaneous’ (local in time) relationship between stress and strain and this may not hold when $\omega_c \sim \Omega$. Generally the viscous stress possesses a relaxation time of order $1/\omega_c$, which in this regime will be comparable to the dynamic time-scale. Thus the immediate history of the stress cannot be ignored and must be dynamically determined. Including this physics has the most impact on the stability of oscillating modes, especially the overstability, it depending on the synchronisation of the stress and density oscillations.

A kinetic model can address both issues, accounting for anisotropy within an appropriate collision term and providing a straightforward way, by the taking of moments, to generate dynamical equations for the viscous stress. Another advantage is that a kinetic model lets us explicitly include the microphysics of particle-particle interactions and thence potentially to model a larger set of the physical mechanisms at play (such as collisions, irregular surfaces, spin, size distribution). It also narrows the scope of our simplifying assumptions to the particulars of collisions between spheres of ice, which have been observed in the laboratory (as discussed in Section 1.2.3).

The risk run, of course, is that the formalism becomes so complicated that to obtain a solution we are obliged to enforce assumptions little better than those we criticise. Particularly, the closure of the moment equations causes a significant degree of trouble, as does the simplification of the collision term.

1.4.2 Kinetic theory

Though it is true the formulation of a suitable kinetic theory poses a number of difficulties, I feel the various approximations these require by no means cripple the model. The most fundamental assumptions we make in this dissertation are that our ring is composed exclusively of *hard*, *identical*, and *indestructible* spheres. We now briefly touch on each.

The assumption of ‘hardness’ is equivalent to saying that the time spent during collisions is negligible to the time between them. This permits us to neglect the cumulative effect of ternary, or higher, collisions: if particle
spend so small a time during a collision there is little chance they will be struck by a third or fourth particle. In fact it can be shown that the ratio of ternary to binary collisions scales like \( na^2\xi_{\text{max}} \sim FF\xi_{\text{max}}/a \) where \( n \) is number density, \( a \) is particle radius, \( \xi_{\text{max}} \) is the maximal compression displacement the particles endure in a collision, and \( FF \equiv 4\pi a^3 n/3 \) is the filling factor (or packing fraction), a quantity which denotes the proportion of space occupied by particles (BP04). The assumption of hardness also collapses the details of particle interactions onto a single parameter, the coefficient of restitution, \( \varepsilon \).

Because a collision is effectively instantaneous, the evolution of the system is indifferent to its detailed dynamics. All that matters is the result of the collision, i.e. the postcollision velocities, which proceed from the specification of \( \varepsilon \) and the corresponding collision rule (derived in Section 3.2.1).

The two assumptions of single-size and indestructibility dispel the complexities of size distributions, and erosion and reaccretion processes. However a number of \( N \)-body computations have simulated a polydisperse gas (as mentioned earlier), and some theoretical work exists (Stewart et al., 1984; Salo, 1987; Hämőnen-Anttila & Salo, 1993). Erosion and accretion processes have been modeled in Weidenschilling et al. (1984) and Longaretti (1989). Particle shapes other than spheres present significant theoretical difficulties but the effect of surface irregularities have been approximately accounted for in the detailed formalisms of Salo (1987) and Hämőnen-Anttila & Salo (1993).

In addition we presume that the particles are non-spinning. The inclusion of spin requires knowledge of tangential friction, which at the present moment is poorly determined, even for typical terrestrial materials (BP04; Araki, 1991). Dropping spin also renders the mathematics more convenient, but see Shukhman (1984), Araki & Tremaine (1986), Araki (1988, 1991), Hämőnen-Anttila & Salo (1993), and Ohtsuki (2006a) for interesting assaults on this problem.

Our approach uses a phase space of \( \mathbf{x} \) and \( \mathbf{v} \). An alternative developed by Hämőnen-Anttila is to frame the kinetic theory in terms of orbital elements. His formalism is worked out in a suite of detailed papers (Hämőnen-Anttila, 1975, 1976a, 1976b, 1977, 1978, 1981, 1982, 1984, 1987, 1988; Hämőnen-Anttila & Salo, 1993) but the mathematics is necessarily byzantine. For
simplicity’s sake, we shall stick to a phase space of \( \mathbf{x} \) and \( \mathbf{v} \).

At low densities the behaviour of a gas of particles can be adequately described by the Boltzmann theory adapted to incorporate the inelasticity of collisions. A central assumption of the Boltzmann approximation is the neglect of particle size in all calculations other than that determining the scattering cross-section. However, at higher densities the theory fails because this assumption leads to the neglect of important ‘dense’ processes. In this regime we must turn to the Enskog model (Chapman & Cowling, 1970). This formalism distinguishes two additional processes; the first is associated with a large filling factor and the other arises from the collisional transfer of particle properties like momentum and energy. We describe these in turn.

When particles take up a significant portion of space (i.e. when FF is not small) the volume in which they may move is reduced, and possible colliders may be screened by other particles (Chapman & Cowling, 1970). This means that the statistics of two impacting particles must include the influence of their neighbours and, as a consequence, the evaluation of the collision frequency must take space correlations into account. In some cases velocity correlations will also play a role. Overall this leads to an enhancement of \( \omega_c \) which, in the Enskog theory, is approximately quantified by a factor \( Y(\text{FF}) \) (the ‘Enskog factor’). This quantity cannot be calculated within the bounds of kinetic theory; it must be determined separately from either the equation of state of the particulate gas or from the radial pair correlation function, which usually require data gathered from molecular dynamics experiments (Araki & Tremaine, 1986; Jenkins & Richman, 1984; PB04, and references therein). Dense inelastic systems may differ considerably in this respect to their elastic counterparts because of the postcollisional correlations in particles’ tangential velocity. As we mentioned in the discussion on jetstreams, inelastic collisions diminish the normal component of the relative velocity but conserve the tangential component, and as a result subsequent collisions lead to a statistical alignment of neighbouring particles’ traces (BP04). Other than possibly forming vortices, this effect can reduce the local value of the collision frequency, and so the Enskog factor of an inelastic ensemble may be
smaller than that of a corresponding elastic system (see Pöschel, Brilliantov & Schwager, 2002).

There are two modes of particle property transfer: their free carriage by particles between collisions and their transmission from one particle to another during a collision. In a dilute gas, the former so-called ‘local’, or translational, mode dominates the latter because particles travel relatively long distances between collisions. In a dense gas, the mean free path is much reduced, which can mean that the finite size of the particles is large enough for the exchange of properties between colliding particles to become important. We often find that collisional transfer is at least as effective as translational transfer in a dense gas.

Estimates for the relative magnitude of the two modes are easy to derive. Consider the transport of momentum across a plane by particles of radius \(a\), mass density \(\rho\) and velocity dispersion \(c\). The magnitude of the momentum flux density due to the free carriage of particles is of order \((\rho c)c\). Now consider all the collisions between particles straddling the same surface and suppose the collision rate is \(\omega\). The flux density of momentum carried from the centres of all the colliding particles on one side to all the particles on the other side in a collision is of order \((\rho c)a\omega c\). This implies the scaling:

\[
\frac{|P_{\text{coll}}|}{|P_{\text{tran}}|} \sim \frac{a\omega c}{c} = (\omega c/\Omega)R, \tag{1.3}
\]

where \(P_{\text{coll}}\) and \(P_{\text{tran}}\) refer to the collisional and translational pressure tensors respectively, \(\Omega\) is the local orbital frequency, and \(R\) is defined by

\[
R \equiv \frac{a\Omega}{c}. \tag{1.4}
\]

Elsewhere \(R\) is called the Savage and Jeffrey \(R\)-parameter (Savage & Jeffrey, 1981; Araki, 1991), and it quantifies the ratio of shear motions to velocity dispersion. Note that in a Keplerian disk the shear rate is \(r\Omega/dr = -3\Omega/2\). For a general flow \(R = aU/(cL)\) where \(U\) is a characteristic velocity scale and \(L\) is the length-scale associated with its variation.

Now consider the energy drawn from the mean orbital flow due to its vari-
ation on length-scales of order \( a \). The energetic loss from particle inelasticity is of order \( \omega_c \rho c^2 (1 - \varepsilon^2) \) per unit volume (GT78). The viscous heating rate per unit volume, issuing from the \textit{translational} stress, is \( \sim \Omega \rho c^2 \). However, because of the shear flow, there will be an extra positive contribution when we average over all particle collisions which derives from the \textit{collisional} stress. Particles on one side of the gradient will have less velocity than the other, the magnitude of this difference being approximately \( a \Omega \). On average, the energy density injected into random particle motions by the collisional stress is hence \( \sim (a \Omega)^2 \rho \), and the rate of its production \( \sim \omega_c (a \Omega)^2 \rho \). Consequently, we write
\[
\left| \dot{E}^{\text{coll}} \right| \sim R^2, \quad (1.5)
\]
where \( \dot{E}^{\text{coll}} \) and \( \dot{E}^{\text{inel}} \) denote the rate of energy injected by collisions from the mean flow and the rate of energy lost due to collisional dissipation respectively. It is clear from these Eqs (1.3) and (1.5) that when \( R < 1 \) the collisional transport processes are relatively more significant than collisional production: so if we steadily decrease the velocity dispersion of a gas we will notice collisional effects in the momentum equation before the energy equation.

We subsequently define a \textit{dilute} gas as one in which the effect of the collisional transfer of momentum and the collisional production of energy is unimportant. These two requirements boil down to a single ‘diluteness condition’, \( (\omega_c/\Omega) R \ll 1 \). In fact, let us establish a ‘diluteness parameter’, \( (\omega_c/\Omega) R \) which is effectively zero for a dilute gas.

The effects of large filling factor and collisional transport usually work in tandem, though for very large or very low optical depths there are perverse cases when one can exist without the other. This can be observed in the scaling:
\[
\tau \sim \text{FF}/R. \quad (1.6)
\]
For a substantial discussion on this subject see Araki (1991). In Saturn’s rings \( R \) is probably of order unity (Salo, 1992b), which means that filling factor effects are as important as the optical depth is large. Therefore in low
optical depth regions, such as the C and D-rings and the Cassini division, these are probably negligible. In contrast, collisional transport/production effects will be important throughout the rings on account of their low velocity dispersion \( R \sim 1 \).

1.5 Summary

The dissertation shall be organised as follows. Chapter 2 comprises the exposition and examination of two dilute kinetic theories, those proffered by Shu & Stewart (SS85) and by Goldreich & Tremaine (GT78). These formalisms admit stability criteria for the viscous instabilities we seek to understand, and a comparison shows that their qualitative behaviour is insensitive to the precise details of the collision operators employed. Comparison is also made with an analogous hydrodynamic calculation in which we find significant disagreement, this issuing from the latter’s failure to properly account for the non-Newtonian nature of the stress.

However, a dilute ring is a poor model for the dense Saturnian system, though it provides a useful tool with which to lay out the groundwork for dense calculations. These we undertake in Chapter 3. First a workable formalism is devised which simultaneously generalises and simplifies the model of Araki & Tremaine (1986). This is then put to use on the problem of the onset of viscous overstability. Both particle simulations and observations suggest that viscous overstability occurs in regions with optical depths above a critical value (Salo et al., 2001; Porco et al., 2005). The formalism we derive reproduces this qualitative behaviour and establishes an estimate for the critical optical depth which is quantitatively in good agreement with previous work.

After establishing the theoretical background to the onset of viscous overstability through linear analyses, we study its nonlinear saturation in a simple two-dimensional hydrodynamical model. The results of our simulations are the substance of Chapter 4, and these show that the overstability equilibrates by relaxing into a train of travelling nonlinear waves, as predicted by the weakly nonlinear analysis of Schmidt & Salo (2003). The work we present,
however, is not definitive and is but a preliminary to a more comprehensive computational project we plan for the future.

Finally, in Chapter 5, the full three-dimensional behaviour of the viscous overstability is examined in a gaseous radiative accretion disk. Here the local overstable mode can excite global non-axisymmetric, eccentric modes of modest azimuthal wavenumber, and these can determine important properties of protoplanetary systems (amongst others). We determine the vertical structure of the disk and its modes, treating radiative energy transport in the diffusion approximation. On intermediate scales and low viscosities (for which the 2D theory predicts instability) the three-dimensional system is stable because of the vertical structure of the mode: the horizontal velocity perturbations develop significant vertical shear which induce an increase in viscous dissipation. This behaviour may control the rate of eccentricity decay in protoplanetary disks, and may explain the preferential excitation of large-scale eccentric modes via overstability in thinner disks such as those around Be stars.
Bibliography


55


global one-armed oscillations of equatorial disks. *Journal of the Astro-
nomical Society of Japan*, 43, 75-94.

global one-armed oscillations of equatorial disks. *Astronomy and Astro-
physics*, 210, 369-370.

nonself-gravitating polytropic disk. I - A disk with constant thickness.

in narrowly confined viscous rings. *The Astrophysical Journal*, 331,
838-860.

Angular Momentum Transport Processes. *Annual Review of Astronomy
and Astrophysics*, 33, 505-540.

bility of the accretion disc boundary layer. *Monthly Notices of the Royal


[139] Porco, C. C., Nicholson, P. D., Borderies, N., Danielson, G. E., Gol-
ringlets at 1.29 R(s) and 1.45 R(s). *Icarus*, 60, 1-16.


[141] Pöschel, T., Brilliantov, N. V., Schwager, T., 2002. Violation of Molecu-
v1.


