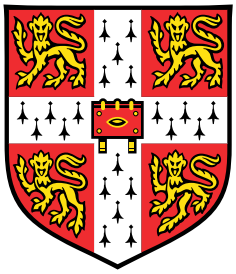
On the Dynamics of Noncircular Accretion Discs



Elliot Matthew Lynch

Supervisor: Prof. Gordon Ogilvie

Department of Applied Mathematics and Theoretical Physics
University of Cambridge

This dissertation is submitted for the degree of
Doctor of Philosophy

Clare College

October 2020

Dedicated to my parents Anne and Philip Lynch

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification at the University of Cambridge, or any other University or similar institution except as declared here and specified in the text. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and declared below. Finally it does not exceed the (unlimited) word limit for the Mathematics Degree Committee.

- **Chapter 5** is an alternative derivation of the Hamiltonian eccentric disc theory developed by Professor Gordon Ogilvie, and published in the Monthly Notices of the Royal Astronomical Society as Ogilvie & Lynch (2019), on which I was a co-author. I contributed to that paper through suggesting some of the formalisms in that paper, such as the introduction of the "geometrical" part of the Hamiltonian.
- **Chapter 6** describes research conducted in collaboration with Professor Gordon Ogilvie, and published in the Monthly Notices of the Royal Astronomical Society as Lynch & Ogilvie (2019).
- **Chapter 7** describes research conducted in collaboration with Professor Gordon Ogilvie, and submitted for publication in the Monthly Notices of the Royal Astronomical Society
- **Chapter 8** describes unpublished research conducted in collaboration with Professor Gordon Ogilvie.
- **Chapter 9** describes unpublished research conducted in collaboration with Professor Gordon Ogilvie.
- **Chapter 10** describes research conducted in collaboration with Professor Gordon Ogilvie, and submitted for publication in the Monthly Notices of the Royal Astronomical Society

- **Chapter 11** describes unpublished research conducted in collaboration with Professor Gordon Ogilvie

Elliot Matthew Lynch
October 2020

Acknowledgements

I am incredibly fortunate to be supervised by Professor Gordon Ogilvie, to whom I owe a great deal. Thank you for you patience, your instruction, and guiding me towards useful avenues of research, without which my PhD would have long ago derailed onto some mathematical tangent with little astrophysical relevance. Hopefully, with your help, I have managed to make a useful contribution to this field. Thanks also go to my adviser Dr Roman Rafikov for advice, for getting the group down the Maypole regularly and for encouraging all of us to interact with the wide astrophysical community, we've all been better off for it.

I would like to express my gratitude to everyone at the AFD research group, our discussions were always interesting and insightful. Thanks in particular to Jansoz Dewberry and Loren Held for your help with simulations and your willingness to be dragged down to Happy Hour, and Antranik Sefilian for insightful discussions about dynamical systems. In the wider astrophysical community, I acknowledge very fruitful discussions with Dr John Zanazzi, Dr Henrik Latter, Dr Robert Izzard, Prof Mark Wyatt, Dr Adam Jermyn and many others.

Thank you to the members of the IoA for allowing me to gate-crash your meetings and putting up with my tendency to throw hypothetical spanners at everything, hopefully a couple of them were useful. Thanks in particular to Tom Comerford for you friendship and mutual support since part III.

Monetarily, I am grateful to the Science and Technology Facilities Council (STFC) for funding this work through a STFC studentship, and for extending my studentship after the bit of bother in 2020. On a technical note this thesis template is derived from a template produced by Krishna Kumar.

Finally thanks to my parents, Anne and Philip Lynch, for always being supportive and helping me through tough times.

Abstract

The classical picture of an accretion disc is of a geometrically thin, nearly axisymmetric, fluid flow undergoing supersonic circular motion and slowly accreting due to angular momentum transport by the disc turbulence. There are, however, strong theoretical and observational grounds for considering non-circular motion in accretion discs. In this thesis I study the dynamics of such non-circular accretion discs using a mix of analytical and semi-analytical methods. One such example of a non-circular accretion disc is an eccentric disc, where the dominant fluid motion consists of slowly evolving, nested, confocal, Keplerian ellipses. I focus on the dynamics of these eccentric discs, along with the dynamical non-axisymmetric vertical structure that they set up.

In the first part of this thesis I consider eccentric waves in ideal fluid discs. I present a secular Hamiltonian theory describing the evolution of the disc orbits due to pressure gradients and show that it can be used to calculate the eccentric standing wave patterns in the disc (the eccentric modes). I derive a "short wavelength" theory for nearly circular orbits that nevertheless can have substantially non-axisymmetric surface density/pressure due to nonlinear eccentricity gradients and disc twist. I use this to show that the pressure profile and precessional forces (such as general relativistic apsidal precession) can focus eccentric waves, causing them to become highly nonlinear in the inner disc.

In the second part of this thesis I consider the action of non-ideal terms on the eccentric disc, such as excitation and damping by viscosity. I derive the ordinary differential equations describing simple horizontally invariant "laminar flows" in a local model of an eccentric disc, which can be used to study the disc's dynamical vertical structure. I also move beyond the purely hydrodynamic models of the preceding chapters and consider the action of magnetic fields in a local model.

Finally I apply the theory presented in this thesis to the highly eccentric discs expected from the tidal disruption of a star by a supermassive black hole. It is currently an open question how, and indeed if, the discs in tidal disruption events circularise. As a step towards understanding the evolution of the disc orbits, I calculate the dynamical vertical structure of highly eccentric discs, emphasising the role of radiation pressure

and thermal stability, and showing that magnetic fields may be important where the disc is highly compressed near the periapsis.

Table of contents

List of figures

Part I

Introduction

Chapter 1

Astrophysical Accretion Discs

1.1 Introduction

This thesis will mostly be concerned with the theory of non-standard accretion discs, where the fluid motion is strongly non-axisymmetric. Such non-axisymmetric structure is important for discs in many astrophysical context. Before considering these non-circular discs it is worth reviewing classical disc theory, where the dominant motion in the disc is assumed to be circular. We do so in this Chapter, alongside reviewing the varied astrophysical contexts where accretion discs are found.

1.2 Classical accretion disc theory

In this section we shall cover some of the results of classical accretion disc theory. Much of the theory covered in this section follows derivations given in Pringle (1981) and González Martínez-País et al. (2014), along with the Part III Dynamics of Astrophysical Discs lecture notes (Ogilvie, 2005, http://www.damtp.cam.ac.uk/user/gio10/accretion.html; Ogilvie, 2020, http://www.damtp.cam.ac.uk/user/gio10/dad.html).

1.2.1 Classical viscous accretion disc model

Consider a fluid flow in a spherically symmetric potential $\Phi(R)$ where R is the spherical polar radius. The continuity equation is

$$\partial_t \rho + \nabla \cdot \rho u = 0 \quad , \tag{1.1}$$

along with the momentum equation,

$$\rho D_t u_i = -\nabla_i p - \rho \nabla_i \Phi + \nabla_j T^j{}_i \quad , \tag{1.2}$$

where ρ is the fluid density, u the fluid velocity, p the fluid pressure and $T^j{}_i$ is the stress tensor, assumed to include the contributions from the turbulent Reynolds and Maxwell stresses. $D_t = \partial_t + u \cdot \nabla$ is the Lagrangian time derivative. These equations must be supplemented with a suitable equation of state which we will assume to be barotropic,

$$p = p(\rho), \tag{1.3}$$

with $p(0) = 0$.

Adopting cylindrical polar coordinates (r, ϕ, z), with the spherical radius given by $R^2 = r^2 + z^2$. For the azimuthal boundary condition, assume 2π periodicity so that $X(t, r, \phi, z) = X(t, r, \phi + 2\pi, z)$ for all quantities X. In the vertical direction we assume symmetry about the midplane so that scalar quantities satisfy $X(t, r, \phi, z) = X(t, r, \phi, -z)$, while u^z, $T^z{}_r$, $T^z{}_\phi$ are antisymmetric. We define z_s to be the height of the disc surface (which could extend to infinity) and require that the density vanish on the disc boundaries so that $\rho(z_s) = \rho(-z_s) = 0$. Above the disc sits a tenuous photosphere, containing negligible mass, which transports heat away from the disc radiatively. For a thin disc we can expand Φ about the midplane as follows:

$$\Phi(R) \approx \Phi(r) + \frac{1}{2}\psi(r) z^2 \quad , \tag{1.4}$$

where we have defined $\psi := \frac{1}{r}\Phi'(r)$, where the prime denotes a partial derivative with respect to r. The linear term in z vanishes by symmetry. Strictly for this expansion to be valid requires $z_s \ll r$, although this should approximately hold near the midplane for large z_s (as is the case for an isothermal disc) if the mass is suitably concentrated near the midplane (i.e. as long as the vertical scaleheight H satisfies $H \ll r$).

Defining a suitable azimuthal average and an integration over the vertical extent of the disc,

$$\langle \cdot \rangle := \frac{1}{2\pi} \int_0^{2\pi} \cdot \, d\phi, \qquad \bar{\cdot} := \int_{-z_s}^{z_s} \cdot \, dz \quad , \tag{1.5}$$

which, with the assumed boundary conditions, means $\langle \partial_\phi X \rangle = 0$ and $\overline{\partial_z(\rho X)} = 0$, for some quantity X. Azimuthally averaging and vertically integrating the continuity equation gives

$$2\pi \frac{\partial \Sigma}{\partial t} + \frac{1}{r}\frac{\partial \mathcal{F}}{\partial r} = 0 \quad , \tag{1.6}$$

where we have introduced the azimuthally averaged surface density Σ and mass flux $\mathcal{F}$,

$$\Sigma := \overline{\langle\rho\rangle}, \quad \mathcal{F} := 2\pi r\overline{\langle\rho u^r\rangle} \quad . \tag{1.7}$$

Making use of the following identity for tensors in cylindrical coordinates,

$$\nabla_i A^i{}_j = \frac{1}{r}\partial_i(rA^i{}_j) - r^{-1}A^\phi{}_\phi\delta^r_j \tag{1.8}$$

allows the momentum equation to be written in the form,

$$\partial_t(\rho u_j) + \frac{1}{r}\partial_i\left[r(\rho u^i u_j - T^i{}_j) - r^{-1}(\rho u^\phi u_\phi - T^\phi{}_\phi)\delta^r_j\right] = -\partial_j p - \rho\partial_j\Phi. \tag{1.9}$$

Multiply by $2\pi r$, azimuthally averaging and vertically integrating to find

$$\begin{aligned}\partial_t\overline{\langle 2\pi r(\rho u_j)\rangle} + \partial_r\overline{\langle 2\pi r(\rho u^r u_j - T^r{}_j)\rangle} - \overline{\langle 2\pi(\rho u^\phi u_\phi - T^\phi{}_\phi)\rangle}\delta^r_j \\ = -2\pi r\delta^r_j(\partial_r P - \Sigma\partial_r\Phi) - 2\pi r\delta^z_j\overline{\langle\rho\rangle\psi z},\end{aligned} \tag{1.10}$$

where we have introduced the vertically integrated pressure $P := \overline{\langle p\rangle}$.

We propose that the dominant fluid motion is fixed circular orbits with angular speed Ω so that the mean velocity can be written as a sum of the circular background and a perturbation v^i,

$$u^i = \Omega(r)e^i_\phi + \epsilon v^i \tag{1.11}$$

where e^i_ϕ is the unit vector in the ϕ direction and ϵ is a book keeping parameter to keep track of the perturbation to the background circular motion. Assume that $T^i{}_j = O(\epsilon)$ and $\partial_t = O(\epsilon)$. While the mean velocity v^r responsible for the accretion flow is $O(\epsilon)$ the velocity associated with the turbulent fluctuations responsible for the accretion is larger. As $T^i{}_j = O(\epsilon)$ this implies that the magnitude of the turbulent velocity fluctuations $u_{\rm turb} = O(\epsilon^{1/2})$. So the picture to have in mind is an $O(1)$ circular orbital motion, on top of which sit $O(\epsilon^{1/2})$ turbulent fluctuations which drive an $O(\epsilon)$ mean flow responsible for disc accretion.

With these assumptions, the components of the momentum equation are

$$-\overline{\langle 2\pi\rho r^2\Omega^2\rangle} = 2\pi r(\partial_r P - \Sigma\partial_r\Phi) + O(\epsilon) \quad , \tag{1.12}$$

$$\epsilon\partial_t\overline{\langle 2\pi r\rho r^2\Omega\rangle} + \epsilon\partial_r\overline{\langle 2\pi r(\rho r^2 v^r\Omega - T^r{}_\phi)\rangle} = O(\epsilon^2) \quad . \tag{1.13}$$

Finally the z component of the momentum equation implies $\overline{\langle\rho\rangle z} \approx 0$, which is satisfied exactly by symmetry about the midplane. More usefully, multiplying the z component of Equation 1.9 by z, azimuthally averaging and vertically integrating we find, at leading order,

$$P = \psi\overline{\langle\rho z^2\rangle}. \tag{1.14}$$

Defining the scale height of the disc to be the vertical standard deviation of the mass distribution, such that $H^2 = \overline{\langle\rho z^2\rangle}/\overline{\langle\rho\rangle}$, this allows us to obtain a relation between the Σ, P and H,

$$P = \psi\Sigma H^2. \tag{1.15}$$

Equation 1.12 can be rearranged to obtain the background angular velocity,

$$\Omega^2 = \frac{1}{r}\frac{d\Phi}{dr} - \frac{1}{r\Sigma}\frac{dP}{dr} \tag{1.16}$$

while equation 1.13 can be written as

$$2\pi r\partial_t(\Sigma h) + \partial_r(h\mathcal{F} + \mathcal{G}) = 0 \tag{1.17}$$

where we have made use of the previously defined mass flux $\mathcal{F}$ and have defined $\mathcal{G}$, the torque due to the $T^r{}_\phi$ stress component,

$$\mathcal{G} := -2\pi r\overline{\langle T^r{}_\phi\rangle} \quad , \tag{1.18}$$

and $h := r^2\Omega$ the specific orbital angular momentum. Using Equation 1.6, Equation 1.17 simplifies to:

$$\mathcal{F}\frac{\partial h}{\partial r} + \frac{\partial\mathcal{G}}{\partial r} = 0 \quad . \tag{1.19}$$

Combining equation 1.19 with equation 1.6 we obtain an equation linking angular momentum transport to mass accretion, which causes evolution of the surface density profile,

$$\frac{\partial\Sigma}{\partial t} - \frac{1}{2\pi r}\frac{\partial}{\partial r}\left[\left(\frac{\partial h}{\partial r}\right)^{-1}\frac{\partial\mathcal{G}}{\partial r}\right] = 0 \quad . \tag{1.20}$$

Assuming that the stress responsible for the angular momentum transport can be modelled as a viscous stress then $\mathcal{G}$ can be written as

$$\mathcal{G} = -2\pi\overline{\nu}\Sigma r^3 \frac{d\Omega}{dr} \quad , \tag{1.21}$$

where $\overline{\nu} = \overline{\nu}(r, t, \Sigma)$ is a vertically averaged viscosity. From this we can obtain a (nonlinear) diffusion type equation for the surface density,

$$\frac{\partial \Sigma}{\partial t} + \frac{1}{r}\frac{\partial}{\partial r}\left[\left(\frac{\partial h}{\partial r}\right)^{-1}\frac{\partial}{\partial r}\left(\overline{\nu}\Sigma r^3 \frac{d\Omega}{dr}\right)\right] = 0 \quad . \tag{1.22}$$

The steady state solution to equation 1.22 is straightforward to obtain, even for the nonlinear case where $\bar{\nu}$ depends on Σ, by setting $\frac{\partial \Sigma}{\partial t}$ to zero, multiplying by r and directly integrating to obtain

$$\frac{\mathcal{F}}{2\pi} = \left(\frac{\partial h}{\partial r}\right)^{-1}\frac{\partial}{\partial r}\left(\overline{\nu}\Sigma r^3 \frac{d\Omega}{dr}\right) = -\frac{\dot{M}}{2\pi} \quad , \tag{1.23}$$

where $\mathcal{F}$ is the mass flux across the boundary. This can be expressed in terms of a constant mass accretion rate with $\mathcal{F} = -\dot{M}$. We rearrange and integrate again to obtain

$$0 = \frac{\mathcal{G}_{\text{in}}}{2\pi} + \frac{1}{2\pi}\dot{M}h + \overline{\nu}\Sigma r^3 \frac{d\Omega}{dr} \quad , \tag{1.24}$$

where $\mathcal{G}_{\text{in}}$ is a constant torque applied to the inner boundary.

Finding the time dependent solution to Equation 1.22, for general $\overline{\nu}$, must be done numerically. However when $\overline{\nu}$ (and Ω) is a fixed function of r the equation is linear and can be solved via Green's method. For a Keplerian disc with $\Omega \propto R^{-3/2}$, constant $\overline{\nu} = \nu_0$ and no torque on the inner boundary the Green's function for equation 1.22 was given in Pringle (1981):

$$G(x, \tau) \propto \tau^{-1} x^{-1/4} \exp\left(-(1 + x^2)/\tau\right) I_{1/4}(2x/\tau) \quad , \tag{1.25}$$

where $x := R/R_0$ and $\tau := 12\nu_0 t R_0^{-2}$ are dimensionless radius and time. I_p is the modified Bessel function. From the initial delta spike this solution viscously spreads, with the peak of the distribution moving inwards until it reaches the inner boundary (See Figure 1.1). The long term behaviour of this solution is for the mass to accrete across the inner boundary (formally at r=0), while the angular momentum is carried off to infinity by a vanishingly small fraction of the initial mass.

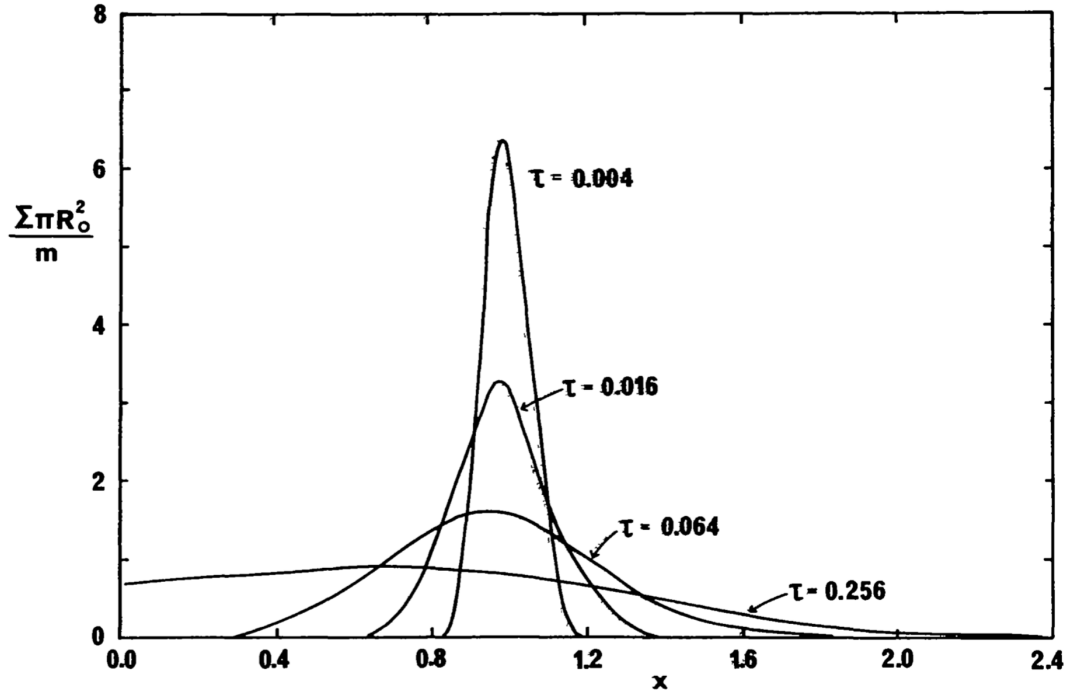


Fig. 1.1 Spreading viscous ring from Pringle (1981). $x := R/R_0$ and $\tau := 12\nu_0 t R_0^{-2}$ are a dimensionless radius and time.

1.2.2 Disc vertical structure

To obtain the vertical structure of a disc in hydrostatic and thermal balance requires knowledge of the heat transport and generation mechanisms within the disc. If, as assumed in the model considered in Section 1.2.1, the disc is truncated by a photosphere, then heat is lost radiatively from the disc surface and must be replenished by some source of heat within the disc. For the classical viscous disc model disc heating is due to viscous dissipation of the orbital energy.

The simplest form of viscosity is that of a Newtonian fluid, where the viscous stress is proportional to the rate of strain. Assuming the viscosity is also isotropic then the viscous stress tensor is given by

$$T^{ij} = \mu S^{ij} + \left(\mu_b - \frac{2}{3}\mu\right) \nabla_k u^k g^{ij} \quad . \tag{1.26}$$

where S^{ij} is the rate of strain tensor, g^{ij} the (inverse) metric tensor, μ, μ_b the dynamic shear and bulk viscosity respectively. The associated viscous heating rate per unit volume is given by

$$\mathcal{H} = T^{ij} S_{ij} \quad . \tag{1.27}$$

For an axisymmetric disc where $u^i = \Omega(r)\delta^i_\phi$, $\nabla_k u^k = 0$ and rate of strain tensor is

$$S_{ij} = \delta^r_{(i}\delta^\phi_{j)} r^2 \frac{d\Omega}{dr} \quad , \tag{1.28}$$

which yields a heating rate per unit volume

$$\mathcal{H} = \mu \left(r \frac{d\Omega}{dr} \right)^2 \quad . \tag{1.29}$$

If the dominant heat transport mechanism is from radiation then the cooling rate per unit volume can be written as the divergence of the radiative heat flux F^i,

$$\mathcal{C} = \nabla_i F^i \tag{1.30}$$

The radiative flux is given by

$$F_i = -\frac{16\sigma T^3}{3\kappa\rho} \nabla_i T \tag{1.31}$$

where σ is the Stefan-Boltzmann constant and κ is the Rosseland mean opacity.

Under the thin disc assumptions it is assumed that the vertical derivatives are much greater than the horizontal derivatives: $\partial_z \gg \partial_r, \partial_\phi$. This means $F_z \gg F_r, F_\phi$ and the radiation predominantly escapes the disc vertically meaning the volumetric cooling rate is given by

$$\mathcal{C} \approx -\frac{\partial}{\partial z} \left(\frac{16\sigma T^3}{3\kappa\rho} \frac{\partial T}{\partial z} \right) \quad . \tag{1.32}$$

The thermal energy equation is written in terms of temperature as

$$D_t T = (\Gamma_3 - 1) \frac{T}{\rho} D_t \rho + \frac{1}{C_v} \left(\mathcal{H} - \mathcal{C} \right) , \tag{1.33}$$

where Γ_3 is the third adiabatic exponent and C_v is the specific heat capacity at constant volume. Thermal balance requires $\mathcal{H} = \mathcal{C}$.

With these details for the heating and cooling in the disc, we are able to obtain equations for the vertical structure of a viscously heated steady axisymmetric thin disc in hydrostatic and thermal equilibrium where heat is transported radiatively,

$$\frac{dp}{dz} = -\rho\psi z \quad , \tag{1.34}$$

$$\frac{dF^z}{dz} = \mu \left(r \frac{d\Omega}{dr} \right)^2 \quad , \tag{1.35}$$

$$\frac{dT}{dz} = -\frac{3\kappa\rho}{16\sigma T^3} F^z \quad . \tag{1.36}$$

These equations can be solved for a given opacity and viscosity law with the requirement that $p(z_s) = \rho(z_s) = 0$ where z_s is the height of the disc surface which acts like an eigenvalue of the problem and must be determined by the solution. Additionally we have $F^z(0) = 0$ by symmetry. These equations are similar to the equations of stellar structure but in a plane parallel atmosphere. They can be solved in a similar way via a shooting method, or by nondimensionalising the solution and numerically integrating from the midplane for arbitrary initial conditions and rescaling the solutions to the physical quantities in the disc.

1.2.3 $\alpha-$viscosity prescription for disc turbulence

In reality almost no astrophysical disc has appreciable viscosity. Instead the viscosity is thought of as a model for turbulent angular momentum transport. This classic disc model is not thought to be a good model when the dominant angular momentum transport mechanism is not turbulence, such as discs which accrete by losing angular momentum in outflows or disc winds.

Shakura & Sunyaev (1973) presented an argument that the turbulent stress in the disc should be proportional to the pressure, which has become known as the $\alpha-$viscosity prescription for disc turbulence. This model has found extensive use in many applications of accretion disc theory. In this model the viscosity $\overline{\nu}$ is sourced by turbulent eddies with typical lengthscale $l_{\rm turb}$ and velocity $u_{\rm turb}$ so that

$$\overline{\nu} \approx l_{\rm turb} u_{\rm turb}. \tag{1.37}$$

The viscosity is assumed to be dominated by the largest and fastest eddies. The eddies must fit within the disc so the characteristic lengthscale must be smaller than the scale height H so $l_{\rm turb} \lesssim H$. Additionally any eddy which exceeds the sound speed c_s should shock and dissipate so that $u_{\rm turb} \lesssim c_s$. An alternative argument (and the original one given in Shakura & Sunyaev (1973)) is that supersonic turbulence would result in heating which raises the soundspeed sufficiently so that $u_{\rm turb} \lesssim c_s$. Thus we can write

$$\overline{\nu} = \alpha H c_s, \tag{1.38}$$

where $\alpha < 1$. This, after using hydrostatic equilibrium, leads to

$$\overline{\nu}\Sigma = \alpha \frac{P}{\Omega}. \tag{1.39}$$

For a steady disc (equation 1.24) this leads to a solution, for the vertically integrated pressure,

$$P = \frac{1}{2\pi\alpha r^3} \left(\frac{-1}{\Omega} \frac{d\Omega}{dr} \right)^{-1} \left(\mathcal{G}_{\rm in} + \dot{M} h \right) \quad , \tag{1.40}$$

from which the disc surface density and temperature can be obtained using the equation of state (equation 1.3). If the disc radiates as a blackbody this gives a prediction for the spectrum of the disc. This corresponds to a sum of the individual black body spectra of each radius, and is approximately the blackbody spectrum of the innermost and outermost orbits connected by a powerlaw with a slope which depends on the steady state surface density distribution in the disc (Lynden-Bell, 1969; Pringle, 1981).

One question raised by the classical disc model is the source of the accretion disc turbulence, responsible for the angular momentum transport. For a circular hydrodynamic disc with a Keplerian rotational profile this runs into the issue that the disc is stable to the development of hydodynamical turbulence by the Rayleigh stability criterion:

$$\frac{\partial h^2}{\partial r} > 0 \quad , \tag{1.41}$$

as $h \propto r^{1/2}$ in a Keplerian disc. The rapid rotation rate of Keplerian discs acts to stabilise them against many hydrodynamic instabilities, such as convection. This stability is a problem for self-consistently modelling disc accretion with the framework of hydrodynamics.

1.2.4 Magnetorotational Instability, the MHD origin of disc turbulence

Even in the classic Shakura & Sunyaev (1973) $\alpha-$disc model it was realised that the disc turbulence was likely magnetohydrodynamic in origin. However it wasn't until the

rediscovery of the magnetorotational instability (MRI) [1] by Balbus & Hawley (1991) and Hawley & Balbus (1991) that there was a plausible mechanism for disc turbulence. In the MRI the presence of a weak magnetic field can destabilise a rotational shear flow by allowing angular momentum exchange between neighbouring annuli. For ideal MHD the condition for the disc to be unstable to the MRI is

$$\frac{\partial \Omega^2}{\partial r} < 0 \quad . \tag{1.42}$$

In other words a disc is unstable to the MRI if the magnitude of the angular velocity decreases outwards, which is precisely the situation found in a (circular[2]) Keplerian disc (Balbus & Hawley, 1991). These authors provided a toy model to understand the destabilising action of the magnetic field in a Keplerian disc. Consider two neighbouring fluid elements elastically tethered by a magnetic field, which can be useful to envisage as two masses within the fluid connected by a spring. If one of these fluid elements is displaced outwards then, in addition to trying to return the fluid element to its original position (which is obviously a stabilising effect), the magnetic field will attempt to enforce rigid rotation with respect to the undisplaced fluid element. This causes the displaced fluid element to rotate faster than the circular rotation rate of its current radial position, causing it to move further out. Similarly displacing a fluid element inwards causes it to fall further in as it is rotating too slowly for its new radial position, thus leading to an instability.

The picture considered in the toy model of the MRI, where the action of the magnetic field within the fluid is envisaged as a pair of masses within the fluid connected by a spring, has proven useful in connecting the dynamics of MHD fluids to other areas of fluid dynamics. In particular, considering the stress associated with pairs of masses connected by springs embedded within the fluid is one way of deriving the Oldroyd-B model for a viscoelastic fluid. An Oldroyd-B fluid is a model of a polymeric fluid, where elastic polymers within the fluid give the model a viscoelastic behaviour which is analogous to the presence of a magnetic field. In fact this connection is more than just an analogy, as it was shown by Ogilvie & Proctor (2003) that there is an exact asymptotic correspondence between an MHD fluid in the limit of large magnetic Reynolds number and an Oldroyd-B fluid in the limit of large relaxation time. This

[1] The MRI was first discovered by Velikhov (1959) and Chandrasekhar (1960) who noted that the Rayleigh stability criterion was not recovered in the zero magnetic field limit of ideal MHD. Because typical stellar rotation profiles are stable to the MRI this instability did not receive much attention and went essentially unnoticed until its rediscovery by Balbus & Hawley (1991) and Hawley & Balbus (1991) in the disc context.

[2] For an eccentric disc this may not be satisfied around the entire orbit

correspondence leads to the suggestion that there should be an elastic analogue of the MRI (Ogilvie & Potter, 2008), which has subsequently been experimentally verified (Boldyrev et al., 2009; Bai et al., 2015).

1.2.5 The thermal instability

Another important area of classical accretion disc theory for this thesis is the thermal instability. The thermal instability is where the thermal equilibrium considered in subsection 1.2.2 is unstable to perturbations to the temperature. The origin of a thermal stability criterion in astrophysics is due to Field (1965) who was interested in the condensation of structures out of a diffuse medium when self gravity is unimportant (e.g. solar prominences, planetary nebulae and interstellar clouds). The Field criterion takes the form

$$\left(\frac{\partial(\mathcal{H}-\mathcal{C})}{\partial T}\right)_A > 0 \quad , \tag{1.43}$$

where, as with many thermodynamic relations, this differential is carried out keeping a thermodynamic variable (A) constant. It is not immediately apparent which thermodynamic quantity should be held fixed. In the original application to planetary nebula Field (1965) suggested that this derivative should be carried out at constant pressure.

For a classical accretion disc Condition 1.43 with $A = p$ isn't appropriate. To see this we must consider the relevant timescales in the disc. The dynamic time is set by the orbital motion with $t_{\rm orb} = \Omega^{-1}$. Hydrostatic equilibrium is set up on the vertical sound crossing time $\sim H/c_s \sim \Omega^{-1}$, comparable with the dynamical time. This is typically much shorter than the thermal time $t_{\rm therm} \sim P/(\bar{\nu}\Sigma\Omega^2) \sim \alpha^{-1}\Omega^{-1}$, which is the timescale on which the thermal content of the disc adjusts, and thus the timescale of any thermal instability. Thus for accretion discs the constraint $A = p$ is not appropriate as the derivative should be carried out preserving hydrostatic equilibrium.

Perturbing the temperature and density away from the thermal equilibrium gives the linearised thermal energy equation,

$$D\delta T = (\Gamma_3 - 1)\frac{T}{\rho}D\delta\rho + \frac{1}{C_v}\left(\frac{\partial\mathcal{H}}{\partial T} - \frac{\partial\mathcal{C}}{\partial T}\right)\delta T \quad . \tag{1.44}$$

If the accretion timescale of the disc is long compared to the thermal time then Σ is constant (which, with hydrostatic equilibrium, implies that p/H is constant) and $\delta\rho = -\frac{\rho}{H}\delta H$. We can now rearrange to find

$$\frac{D\delta T}{\delta T} = \frac{\left(\frac{\partial \mathcal{H}}{\partial T}\right)_{p/H} - \left(\frac{\partial \mathcal{C}}{\partial T}\right)_{p/H}}{C_v \left[1 + (\Gamma_3 - 1)\left(\frac{\partial \ln H}{\partial \ln T}\right)_{p/H}\right]} \tag{1.45}$$

Thermal instability occurs when $\frac{D\delta T}{\delta T} > 0$. Assuming that $1 + (\Gamma_3 - 1)\left(\frac{\partial \ln H}{\partial \ln T}\right)_{p/H}$ is positive (which is usually the case) the condition for thermal instability is:

$$\left(\frac{\partial \mathcal{H}}{\partial T}\right)_{p/H} > \left(\frac{\partial \mathcal{C}}{\partial T}\right)_{p/H} \quad . \tag{1.46}$$

The thermal stability of a radiation pressure dominated disc is currently an open question. For a circular disc with an alpha model (Shakura & Sunyaev, 1973) of the viscous stress a radiation pressure dominated disc is thermally unstable when the stress scales with the total (radiation + gas) pressure (Shakura & Sunyaev, 1976; Pringle, 1976; Piran, 1978), but thermally stable when stress scales with gas pressure (Meier, 1979; Sakimoto & Coroniti, 1981).

To see this consider an α disc, with constant opacity, subject to radiative cooling. Assuming that $t_d \ll t_t \ll t_v$, where t_d is the dynamical time (on which hydrostatic equilibrium is established), t_t is the thermal time and t_v is the viscous time (on which the surface density evolves), we can assume Σ is constant and the disc remains in hydrostatic equilibrium. We also assume the disc is initially in thermal equilibrium, with $\mathcal{H} = \mathcal{C}$. As the opacity κ and Σ are constant the optical depth τ is constant. Thus the cooling rate is

$$\mathcal{C} \sim \sigma_B T_s^4 \sim \frac{\sigma_B T_m^4}{\tau} \quad , \tag{1.47}$$

where T_s is the temperature at the disc surface and T_m is the midplane temperature. As the disc satisfies hydrostatic equilibrium $p \sim p_r \sim T_m^4 \sim H$. When stress scales with total pressure ($T^{R,\phi} \sim \alpha p \sim \alpha p_r$) the heating rate is given by $\mathcal{H} = \frac{\alpha p H}{\Omega} \sim T^8$, and thus we have

$$\frac{\partial \mathcal{H}}{\partial T} \sim \frac{8\mathcal{H}}{T} \sim \frac{8\mathcal{C}}{T} > \frac{4\mathcal{C}}{T} \sim \frac{\partial \mathcal{C}}{\partial T} \quad , \tag{1.48}$$

implying thermal instability. While for stress scaling with gas pressure ($T^{R,\phi} \sim \alpha p_g$) the heating rate is given by $\mathcal{H} = \frac{\alpha p_g H}{\Omega} \sim T$, and thus we have

$$\frac{\partial \mathcal{C}}{\partial T} \sim \frac{\mathcal{H}}{T} \sim \frac{\mathcal{C}}{T} < \frac{4\mathcal{C}}{T} \sim \frac{\partial \mathcal{C}}{\partial T} \quad , \tag{1.49}$$

implying thermal stability.

The $\alpha-$viscosity prescription is a model of the turbulent stress which arises from the action of the MRI (Balbus & Hawley, 1991; Hawley & Balbus, 1991). While early work on the MRI tended to support thermal stability (Hirose et al., 2009), more recent work (Jiang et al., 2013; Mishra et al., 2016; Ross et al., 2017), with large enough box size to allow the disc to expand upon heating, supports stress scaling with total pressure and thermal instability in radiation pressure dominated discs. The thermal instability in a MRI turbulent circular disc either results in thermal runaway (Jiang et al., 2013; Ross et al., 2017) or collapse to a thin gas pressure dominated disc (Jiang et al., 2013; Mishra et al., 2016).

1.3 Diversity of observed accretion discs

Owing to the scale free nature of gravity and the equations of ideal magneto-hydrodynamics, accretion discs occur in a wide array of astrophysical contexts spanning many orders of magnitudes in spatial and temporal scales. The diversity in accretion disc phenomena often come about due to properties of the disc which break the scale free nature of the underlying equations. Radiation physics, non-ideal MHD, chemistry and kinetic effects all break the scale free nature of accretion discs and modify both the dynamics of the underlying fluid flow and the observed emission from the disc.

Orbital companions, whether comparative in mass to the central object or much smaller with negligible gravitational influence on the central object, act to shape the observed diversity in spatially resolved discs. Orbital companions can clear gas from regions of the disc, either truncating the disc entirely or clearing a narrow gap in the disc. They can act to misalign the disc and deform the disc midplane. They can also launch a a host of different waves into the disc. Additionally companions may be the source of the disc material itself through winds or Roche Lobe overflow.

While the main influence of the central object on the disc is through its "point mass" potential the nature of the central object can also strongly affect the inner region of the disc. Torque from central binaries, radiation pressure from young stars and strong magnetic fields can all act to truncate the inner disc. Additionally the orbital stability of Keplerian orbits in Newtonian gravity in a point mass potential, which allows the disc to be long lived, can break down when additional forces modify the potential. Both strong GR effects around black holes and the departure from Keplerian motion close to central binaries can destabilise the disc by modifying the angular momentum profile in the disc. This can set up nearly radial inflows interior to the disc.

While the basic disc physics remains the same across the wide diversity of astrophysical discs the details will be strongly influenced by the astrophysical context. Arguably the main factor controlling what additional physics are important in a given disc is the disc temperature. This is of course of observational interest as well as it controls what wavelengths the disc predominantly emits in. In what follows we run through the diversity of accretion discs discovered in nature, loosely organised in terms of temperature and highlight some of the key additional physical processes important in the disc along with some of the key observational phenomena.

Around 50 % of stars exist as binaries (De Marco & Izzard, 2017). At the start of their lives tight binaries will be surrounded by a circumbinary disc. Additionally towards the end of their lives binaries can also be surrounded by a circumbinary disc, the remnants of a stellar envelope which was ejected during a phase of common envelope evolution. In both cases the circumbinary disc typically has a high density and is very cold, due to its large extent. This prevents MRI turbulence from operating, although the disc could be gravitationally unstable. The inner region of the disc is strongly perturbed by the binary, with the binary clearing a common cavity within the disc (Artymowicz & Lubow, 1994). The binary often excites eccentricity in the disc, and in turn the disc can excite eccentricity in the binary (Dermine et al., 2013; Miranda et al., 2017; Thun et al., 2017; Thun & Kley, 2018; Izzard & Jermyn, 2018; Muñoz et al., 2019; Pierens et al., 2020; Muñoz & Lithwick, 2020).

Protoplanetary discs represent the cold end of accretion disc phenomena. Their low temperature and high density imply low levels of ionisation in the disc leading to the so called "dead zone" where the MRI cannot operate (Gammie, 1996; Jin, 1996; Hawley & Stone, 1998; Fleming et al., 2000; Kunz & Balbus, 2004; Kunz & Lesur, 2013; Riols et al., 2015; Béthune et al., 2016b,a). While the inner disc close to the star and the upper surface of the disc are sufficiently well ionised for MRI to transport angular momentum, the angular momentum transport mechanism in the bulk of the disc remains an open question. Due to the low temperature and poor ionisation non-ideal MHD effects are important in protoplanetary discs and a non-ideal MHD disc wind is a promising candidate for driving disc accretion in such discs (Blandford & Payne, 1982; Wardle & Koenigl, 1993; Béthune et al., 2017). The outer regions of protoplanetary discs may contain enough mass to undergo gravito-turbulence from the disc self gravity (Gammie, 2001; Durisen et al., 2007; Kratter & Lodato, 2016) although it can be difficult to drive turbulence as strong as the MRI without fragmenting the disc (Gammie, 2001; Riols & Latter, 2018, 2019; Deng et al., 2020). When combined

with nonideal MHD the gravitational instability in the outer disc appears to produce strong dynamo action (Riols & Latter, 2018, 2019; Deng et al., 2020).

Chemistry in protoplanetary discs is important for the formation of planetary material. Additionally it dictates how effectively solids can form in the disc which can modify the dynamics. This solid material has allowed for detailed studies of the disc substructure through millimetre/submillimetre imaging of the disc dust with ALMA (e.g. the famous ring of HL Tau observed by ALMA and the asymmetric transition discs observed by Pérez et al. (2014)). Scattered light observations from VLT/SPHERE have revealed a wealth of dust substructure, including spiral arms (Benisty et al., 2015; Gratton et al., 2019), rings (de Boer et al., 2016; Gratton et al., 2019) and shadows thought to be cast by an inner misaligned disc (Benisty et al., 2018).

Similarly cold, but much lower density, discs are found in older planetary systems which possess massive debris discs. Occasionally such debris discs can produce gaseous discs through the sublimation of dust or during collisions of planetesimals (Kral et al., 2016). The much lower densities of these discs allow them to be well ionised so that MRI can operate (Kóspál et al., 2013; Lieman-Sifry et al., 2016; Marino et al., 2016; Hughes et al., 2017; Greaves et al., 2016; Moór et al., 2017; Matrà et al., 2017b,a; Kral et al., 2019). Photo-chemistry in such discs is important in setting their composition. Additionally the high concentration of dust in these discs may be important for the dynamics. Some of these discs have been found around white dwarfs and may be important in supplying metals to white dwarf atmospheres responsible for white dwarf pollution (Manser et al., 2016a; Cauley et al., 2018; Manser et al., 2020).

Associated with protoplanetary discs is the likely existence of circumplanetary discs around planets massive enough to open a gap in their host protoplanetary discs. These have been found in some higher resolution simulations of gap opening in discs (Lubow et al., 1999; D'Angelo et al., 2002). They represent a promising target for imaging planets forming in protoplanetary discs due to the increased area of emitting material associated to the planet (Zhu, 2015).

Next in our zoo of accretion discs are the discs in symbiotic systems. These are discs around white dwarfs formed by accreting from a red giant companion (Luna et al., 2014; Luna, 2019). These discs are typically fairly large ($\lesssim 1$ Au, e.g. Lee et al. (2019b)) and, like protoplanetary discs, should be poorly ionised in all but their innermost regions. Hence MRI may not operate in these discs leading to similar questions about how angular momentum transport operates. The inner regions of these discs are well ionised and UV bright. Depending on how magnetised the white dwarf is the disc should either accrete onto the white dwarf through a narrow X-ray bright boundary

layer or be truncated by the magnetic field and follow the field lines to the white dwarf poles.

Closely related to (and better studied than) these discs are the cataclysmic variables. In many ways these are the prototypical accretion discs and were amongst the first to be discovered. Much of classic accretion discs theory was developed to explain the observed phenomena in these discs. The discs are well ionised, so angular momentum transport occurs through the MRI (although it has been suggested that non-ideal effects may kill the MRI in quiescence (Gammie & Menou, 1998; Coleman et al., 2016)), and their temperature means they typically emit in the UV. One characteristic feature of these discs is their variability, undergoing outburst cycles where the luminosity increases by several orders of magnitude (Warner, 2003). The standard explanation for this behaviour has been the thermal-viscous instability with the hot high accretion state corresponding to outburst (Pringle, 1981; Meyer & Meyer-Hofmeister, 1981; Smak, 1982; Cannizzo et al., 1982; Faulkner et al., 1983; Mineshige & Osaki, 1983). This instability is associated with the ionisation of hydrogen at ~ 5000K. In the cold state the disc temperature slowly rises as mass builds up in the disc. The ionisation of hydrogen causes a sharp increase in the disc opacity from electron scattering that results in a drop in the disc cooling rate so that the disc experiences runaway heating and transitions to the hot state. The disc then cools as mass accretes in the hot state, until hydrogen recombines resulting in a transition back to the cold state. The details of this mechanism are still fairly uncertain as typically cataclysmic variable outbursts are far more abrupt than what is predicted by the model which predicts an unobserved rising phase prior to outburst due to the slope of the $\Sigma - T$ relation (Smak, 2000; Hameury, 2020). Additionally the details of how the thermal instability operates when the heating is sourced by MRI turbulence, rather than modelled as viscous heating, has only recently been addressed (Hirose et al., 2009; Latter & Papaloizou, 2012; Scepi et al., 2018a,b).

Next on our list are the Active Galactic Nuclei (AGN) or quasars. These are the most energetic phenomena in our universe typically outshining their host galaxies. Like protoplanetary discs the outer region of AGN discs can be cold and poorly ionised, but dense enough to be gravitationally unstable causing gravito-turbulence and potentially clumping/undergoing star formation (Paczynski, 1978a,b; Kozlowski et al., 1979; Kolykhalov & Syunyaev, 1980; Goodman, 2003). The inner region near the SMBH is hot and emits in the X-rays. Due to there large size the timescales in these discs are long with disc variability operating on timescales longer than current observational timescales (Komberg, 2007), although in principle it might be possible to detect state

changes in AGN with observations spanning several generations. Many AGN discs launch highly relativistic collimated jets which extend to scales much larger than their host galaxy and jet launching mechanisms have received much attention (Blandford & Znajek, 1977; Blandford & Payne, 1982; Event Horizon Telescope Collaboration et al., 2019; White et al., 2019; Liska et al., 2020). Another extensively studied area is mechanisms by which these discs can accrete at super-Eddington rates in order to grow the central SMBH to their current size in less than the Hubble time. This has motivated the study of various thick disc prescriptions where the standard thin disc assumptions are relaxed (Nityananda & Narayan, 1984; Abramowicz et al., 1988; Tchekhovskoy et al., 2011; McKinney et al., 2012).

Also around SMBHs are the short lived discs formed by the tidal disruption of stars. In a tidal disruption event (TDE) a star, on a nearly parabolic orbit, approaches close enough to the black hole for the black hole's tidal forces to overcome the star's self-gravity. Typically occurring around black holes with lower mass than those in AGN (and potentially occurring around intermediate mass black holes (Clausen & Eracleous, 2011; Mezcua, 2017; Chen & Shen, 2018; Maguire et al., 2020)) these discs are correspondingly hotter and more compact than AGN. The violent nature of their formation leads to an expectation that they should circularise rapidly and have a comparatively high luminosity; however this isn't necessarily borne out by observations (Gezari et al., 2006, 2008; van Velzen et al., 2011; Wang et al., 2011, 2012; Arcavi et al., 2014; Liu et al., 2017; Cao et al., 2018). Whether they in fact circularise is the subject of current research and will be discussed further in this thesis.

Around low mass black holes and neutron stars are X-ray binary discs, which are very hot and are strong X-ray sources. Typically these discs emit predominantly in the soft X-rays and there is an additional non-thermal hard X-ray source from the disc corona, a hot magnetised optically thin plasma with somewhat unknown geometry (typically having whatever geometry people need it to have at a given moment) around the inner disc. Like cataclysmic variables XRB discs are also variable, on long (month-year) timescales these discs transition between quiescent and outburst as a result of the thermal-viscous instability in the outer disc (Lasota, 2001; Remillard & McClintock, 2006). During outburst the disc cycles between a high luminosity, disc dominated, soft state and a low luminosity, corona dominated, hard state (Remillard & McClintock, 2006). The hard/soft transition occurs at temperatures where there is no sharp opacity change able to trigger the thermal instability, so the mechanism responsible for the soft/hard transition is not yet agreed upon (Narayan & Yi, 1995; Meyer et al., 2000; Potter & Balbus, 2014, 2017). Another feature of XRBs is the presence of quasi-

periodic oscillations (also present in some cataclysmic variables (Warner, 1986; Warner & Woudt, 2005, 2008)) - coherent peaks in the power spectra thought to probe key disc frequencies (Remillard & McClintock, 2006).

Finally our last disc is the densest and hottest of them all and is that formed during the merger of a neutron star and a black hole or second neutron star. This disc has recently received a lot of attention due to the discovery of GW 170817 and the associated gamma-ray burst GRB 170817A, and optical transient AT 2017gfo, which matched that expected for the merger of two neutron stars (Abbot et al., 2017).

1.4 Plan of the thesis

This thesis is divided up into four parts and an epilogue. These parts are separated broadly by the physical processes considered in the disc. Part I is a review of the theoretical and observational material relevant to non-circular discs. Part II considers ideal, conservative theories of eccentric discs. Part III considers the role of non-ideal effects. Finally Part IV deals with the application of eccentric disc theory to TDE discs.

Chapter 2 reviews the existing theoretical work which has been developed to describe the dynamics of non-circular discs, in addition to providing the theoretical motivation for considering non-axisymmetric discs.

Chapter 3 reviews the observational evidence for eccentric and tidally distorted discs. There is now a large body of both indirect evidence for eccentric discs and directly measured or imaged eccentricities in discs.

This thesis will make use of certain mathematical formalisms which are not commonly used in astrophysics. Chapter 4 will review the necessary mathematical background which will be used in the main body of the thesis.

In Chapter 5 a Hamiltonian theory for an ideal eccentric disc will be formulated. In this theory the fluid motion consists of nested, confocal, Keplerian ellipses, which slowly evolve according to Hamilton's equations, owing to pressure gradient and external perturbations.

Chapter 6 considers a nonlinear limit of the theory developed in Chapter 5, where the eccentricity is small, but with potentially nonlinear eccentricity gradients or twists. We shall show that there are conditions in which an eccentric wave can be focused by the pressure profile and precessional forces in the disc, causes it to become nonlinear.

In Chapter 7 we study a local model of an eccentric disc which includes nonideal effects, such as viscous heating and radiative cooling, and consider a more general

equation of state consisting of a mixture of a perfect gas and radiation. We derive the simplest horizontally invariant "laminar flow" solution to the flow in the eccentric disc.

In Chapter 8 we derive dissipative terms in the short-wavelength limit and show how they add sources/sinks to the conservation laws found in Chapter 6.

Chapter 9 looks at magnetic fields in eccentric discs and derives the "locally axisymmetric" solutions to the ideal induction equation along with presenting a new constitutive model of the MRI.

Chapter 10 is an introduction to TDEs which will be the focus of the remaining part of the thesis. Chapter 11 applies the model derived in Chapter 7 to the highly eccentric discs expected in TDEs, while Chapter 12 generalises the result to MHD using the results of Chapter 9.

Finally Chapter 13 is a summary with suggestions for future work.

Chapter 2

Development of the Theory of Distorted Discs

2.1 Theory of eccentric discs

While the discs of classical disc theory are axisymmetric by construction, there are many theoretical reasons for considering non-axisymmetric disc structures. Early work on accretion discs mostly ignored the possibility of discs possessing significant asymmetries under the (as it would turn out highly questionable) assumption that the disc would rapidly circularise any asymmetries in the disc. In classical disc theory, discs must be able to efficiently lose energy (via radiative losses through the disc surface) but not angular momentum (which is redistributed on the viscous timescale), in order to form an extended geometrically thin disc (Lynden-Bell & Pringle, 1974; Pringle, 1981). For a fixed angular momentum, a circular orbit is the lowest energy configuration for a single Keplerian orbit. Thus the expectation is that, as a disc efficiently loses energy at fixed angular momentum, the disc should settle into a minimum energy configuration; which, in classical disc theory, was taken to mean the disc should become circular (Lynden-Bell & Pringle, 1974). That the situation may be more complicated than this should be apparent from the fact that the minimum energy configuration of a disc (at fixed angular momentum) is a single particle at infinity carrying the entire angular momentum budget, while the rest of the mass is located at the origin (Lynden-Bell & Pringle, 1974). The path by which a disc gets there need not involve a circular disc.

2.1.1 Celestial mechanics

Consider the orbit of a test particle around Newtonian point mass with mass M_1. Let (r, ϕ) be polar coordinates in the orbital plane. The polar equation for an elliptical Keplerian orbit of semimajor axis a, eccentricity e and longitude of pericentre ϖ is

$$r = \frac{a(1-e^2)}{1+e\cos f} \quad , \tag{2.1}$$

where $f = \phi - \varpi$ is the true anomaly. Often it is useful to introduce an alternative angular coordinate to the true anomaly, particularly when orbital averaging needs to be carried out. One such angular coordinate is the eccentric anomaly E (see Figure 2.1) related to the true anomaly by

$$\cos f = \frac{\cos E - e}{1 - e\cos E}, \quad \sin f = \frac{\sqrt{1-e^2}\sin E}{1-e\cos E} \quad , \tag{2.2}$$

with the radius given by

$$r = a(1 - e\cos E) \quad . \tag{2.3}$$

The time-dependence of the particle's location around the orbit is expressed through Kepler's equation,

$$M = E - e\sin E, \tag{2.4}$$

where $M = n(t-\tau)$ is the mean anomaly, $n = (GM_1/a^3)^{1/2}$ is the mean motion and τ is the time of pericentre passage. The mean anomaly corresponds to the true anomaly of the equivalent circular orbit with the same semimajor axis.

Another useful concept from celestial mechanics is the complex eccentricity,

$$\mathcal{E} = e\exp(i\varpi) \quad , \tag{2.5}$$

which defines a vector in the orbital plane which points towards the pericentre with length equal to the eccentricity.

We need two additional quantities to define an orbit in 3D space. A Keplerian orbit is confined to a single orbital plane. Given some reference plane, and a reference direction within the reference plane, the orbital plane can be specified in terms of the orbital inclination ι and longitude of ascending node Ω. The orbital inclination is the angle between the reference plane and the orbital plane, while the longitude of ascending node is the angle between the reference direction and the line of nodes in

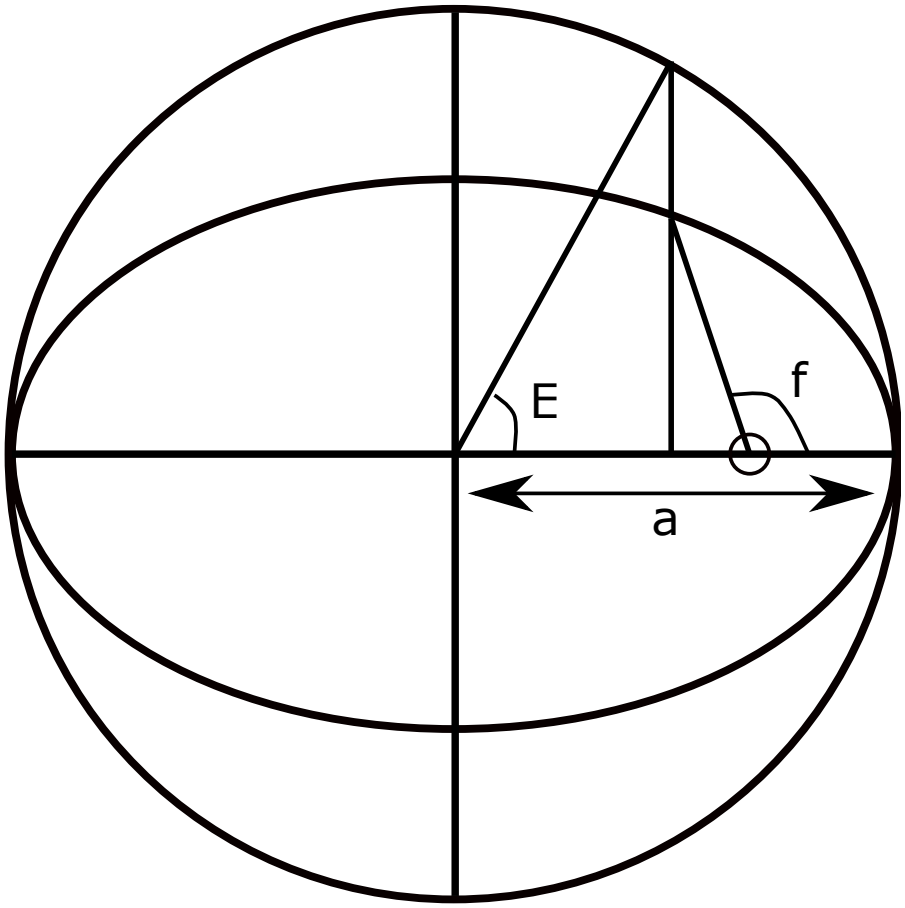


Fig. 2.1 Geometric construction of the eccentric anomaly in an eccentric orbit. The large circle is the circular orbit with the same semi-major axis as the eccentric orbit.

the direction of the ascending node. The line of nodes is the intersection between the orbital plane and the reference plane, with the ascending node the point on the orbit where a particle crosses the reference plane in the positive direction (i.e. from below to above).

Before being studied in relation to gaseous discs, the dynamics of eccentric orbits was extensively studied in celestial mechanics. That orbits in Newtonian gravity can be eccentric, predates Newtonian gravity itself and originates in Kepler's empirical laws for the planetary motion. In the Newtonian 2-body problem the Laplace-Runge-Lenz vector (the vector pointing towards pericentre with length equal to the eccentricity) is an invariant of the system. The application of perturbation theory to the problem of a small force applied to a 2-body orbit led to the development of osculating orbital elements by Lagrange (de Lagrange, 1853). Here the main motion of the particle is

that of a Keplerian orbit which is described by 5 orbital elements (and orbital phase). These orbital elements evolve slowly over many orbits due to the small force leading to the slow evolution of the Keplerian orbit.

The development of Hamiltonian mechanics and the theory of nearly integrable systems provided a natural framework for the osculating orbital elements and perturbations to Keplerian motion. The osculating orbital elements then have a natural description in terms of the action-angle coordinates of the Hamiltonian, with the conserved actions associated with the invariance of the Hamiltonian under continuous symmetries (for instance global rotations and time translation). The Newtonian 2-body problem is unusual among integrable systems in that the orbits are always periodic - rather than the quasiperiodicity seen in most integrable systems. Perturbations to Keplerian orbits come from adding a small potentially non-integrable term to the Hamiltonian,

$$H = H_{\rm Kepler}(\overrightarrow{J}) + \epsilon H_1(\overrightarrow{J}, \overrightarrow{\Theta}) \quad . \tag{2.6}$$

Here $H_{\rm Kepler}$ is the Hamiltonian of the Newtonian 2-body system which has conserved actions $\overrightarrow{J}$, and associated angles $\overrightarrow{\Theta}$. $H_1(\overrightarrow{J}, \overrightarrow{\Theta})$ is a nonintegrable Hamiltonian perturbation which depends on both the action and angle coordinates and $\epsilon \ll 1$.

In general orbits of an integrable Hamiltonian are tori in phase space. Adding a small integrable term to the Hamiltonian deforms the tori. What is the effect of adding a small non-integrable term to the Hamiltonian? Do the tori deform as with the addition of an integrable term? Or does the breaking of the symmetry of the Hamiltonian result in the destruction of the tori? The Kolmogorov–Arnold–Moser theorem guarantees that for a small enough non-integrable terms ($\epsilon \ll 1$) the majority of the tori survive and are deformed (Arnold, 1978).

Not all the tori survive the addition of a non-integrable perturbation, with more tori destroyed the larger the non-integrable term is. The destruction of the stable tori is associated with overlapping resonances, with the stable orbits of one resonance 'punching holes' in the torus resulting in a structure known as a Cantorus. Eventually when the non-integrable term is large enough ($\epsilon \sim 1$) the disruption of the quasi-periodic tori of the integrable system results in a chaotic system.

Interestingly for Keplerian orbits, perturbations of the Hamiltonian (integrable or not) change the character of the orbit and destroys the periodic orbits of the $2-$body problem, which become quasi-periodic tori. This quasi-periodicity which is induced by the perturbation is the origin of the slow evolution of the orbital elements, which

manifest in libration and precession of the orbit. It can also induce periodic variation of e, ι and a.

There are various action-angle coordinates for the Keplerian 2-body problem, many having coordinate singularities for certain orbital geometries (normally for zero eccentricity or zero inclination), due to one (or more) of the angle coordinates becoming undefined. In this Thesis we shall make use of the modified Delaunay variables with,

$$\Lambda = \sqrt{GM_1 a}, \qquad \lambda = M + \varpi, \tag{2.7}$$

$$\Gamma = \Lambda\left(1 - \sqrt{1-e^2}\right), \qquad \gamma = -\varpi, \tag{2.8}$$

$$\mathfrak{H} = 2\Lambda\sqrt{1-e^2}\sin^2\frac{\iota}{2}, \qquad \eta = -\Omega, \tag{2.9}$$

Here (Λ, λ), (Γ, γ), $(\mathfrak{H}, \eta)$ are canonically conjugate action-angle pairs. Λ is related to the energy of the orbit while Γ is a positive definite measure of the orbit's eccentricity, being proportional to e^2 when $e^2 \ll 1$. Γ is also closely related to the angular momentum deficit (AMD), a quantity of great importance in eccentric discs, which is the difference between the angular momentum of the orbit and that of a circular orbit with the same energy. For a single orbit the AMD equals $\mathcal{C} = m\Gamma$, where m is the mass on the orbit. $\mathfrak{H}$ and η are the action-angle coordinates describing the orientation of the orbital plane.

One question, when studying eccentric orbits, is whether there exist mechanisms to excite the orbital eccentricity. An important result of celestial mechanics is the discovery that if a particle orbit is perturbed by an inclined external perturber then the particle can exchange inclination and eccentricity cyclically. This cycle between a highly inclined nearly circular orbit and a moderately inclined highly eccentric orbit is called a Lidov-Kozai cycle (Lidov, 1962; Kozai, 1962)[1].

Two constants of motion in the hierarchical 3-body problem are

$$\sqrt{m_i a_i (1-e_i^2)}\cos\iota_i = \text{const}, \tag{2.10}$$

where the subscript i refers to either the inner or outer orbit. They are related to the existence of an invariable plane which is the plane perpendicular to the conserved angular momentum vector of the system. In a secular theory the orbital elements (with the exception of λ_i) are approximately constant on the orbital timescale. Averaging the Hamiltonian over the orbital time removes its dependence on λ_i meaning the

[1]It now appears that a version of this mechanism was originally discovered by Hugo von Zeipel in 1910 (von Zeipel, 1910; Ito & Ohtsuka, 2019).

conjugate momenta Λ_i (and thus a_i) become integrals of motion. So in the secular 3-body problem,

$$\sqrt{1-e_i^2}\cos\iota_i = \text{const.} \tag{2.11}$$

So changes in the inner orbit's inclination due to the external perturber necessarily lead to changes in its inclination. For this effect to occur the inclination must exceed a critical value $\approx 39^\circ$ (von Zeipel, 1910; Lidov, 1962; Kozai, 1962). Recent work has demonstrated that this mechanism can operate in fluid discs (Martin et al., 2014; Fu et al., 2015a,b, 2017; Lubow & Ogilvie, 2017; Zanazzi & Lai, 2017).

Another aspect of celestial mechanics which is important in eccentric disc theory are the mean motion resonances. These are locations at which secular theory is invalid as periodic terms interact in such a way to cause changes in the osculating elements on an orbital timescale. As discussed above interaction of resonances is responsible for the onset of chaos. However the response of a fluid disc to a mean-motion resonance is different from that of non-interacting particles so we shall defer a discussion of mean motion resonances to later sections.

2.1.2 Galactic dynamics

The exploration of non-axisymmetric structures in discs began in galactic dynamics. Galactic discs are large and were resolved long before gaseous discs and many have non-axisymmetric structures, notably the presence of spiral arms. This led to the development of density wave theory by Lin & Shu (1964, 1966) - which explains the spiral arms as resulting from a concentration of stars which occurs when several non-axisymmetric orbits bunch up (Julian & Toomre (1966); Simkin (1970); Hunter (1973), See Figure 2.2) . The structure propagates in the disc of stars due to pressure (or velocity dispersion) and their collective self gravity.

Spiral density wave theory was created to solve the winding problem. The winding problem occurs in a differentially rotating fluid with a material spiral. The differential rotation will cause a material spiral to become increasingly tightly wound until it breaks apart due to phase mixing. To see this consider a differentially rotating fluid with velocity

$$u^i = \Omega(r)\delta^i_\phi \quad , \tag{2.12}$$

and density ρ. As $\nabla_i u^i = 0$ the continuity equation is

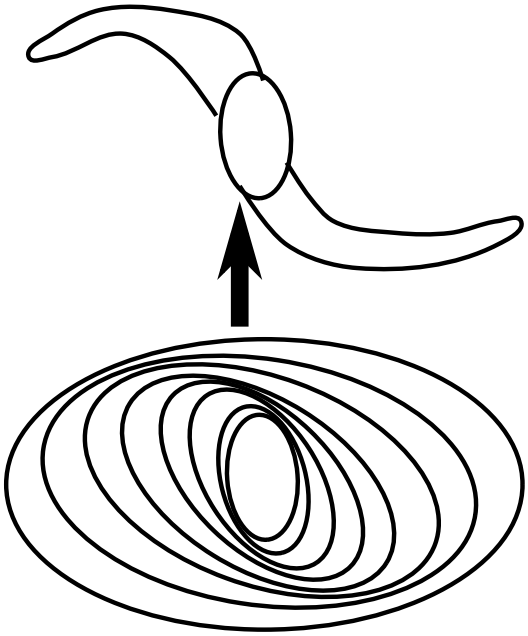


Fig. 2.2 Cartoon showing how concentric elliptical orbits can create a spiral structure in a galactic disc. The bunching up of the ellipses due to their differing orientations causes high density regions which are responsible for the spiral patterns of galaxies.

$$D\rho = \dot{\rho} + \Omega(r)\partial_\phi \rho = 0 \quad . \tag{2.13}$$

The general solution to this PDE is $\rho = \rho_0(r)f(\varphi)$, where $\varphi = \phi - \Omega(r)t$. As an example $f(\varphi) = \cos(m\varphi)$ represents an m-armed structure which takes the form of an increasingly wound spiral for $t > 0$ if Ω has a radial dependence, ultimately resulting in the spiral feature having such a short radial wavelength that it becomes unobservable (and will ultimately be destroyed by stellar dispersion).

A second area of galactic dynamics which was important in the development of eccentric disc theory comes from eccentric nuclear discs. Images of the centres of several galaxies, notably M 31 (the Andromeda galaxy) have two separate components; one component (P1 in the classification used in Tremaine (1995)) is centred on the nuclear core, consistent with an isotropic star cluster on close orbits around the central black hole, while the second (P2) is offset from the nuclear core. The second is too close to the central black hole to be another isotropic cluster around a second SMBH, as tidal forces from the black hole at the centre of P1 would disrupt the cluster. An explanation was put forward by Tremaine (1995), where the second cluster of stars is identified with the apocentre of an aligned eccentric cluster of stars , and he presented

a mathematical model of an aligned eccentric disc to demonstrate this can produce the desired second peak in the surface brightness consistent with P2.

To elucidate the model, Tremaine (1995) presented a simple argument for an eccentric disc having a bright point at apocentre. According to this "apocentre hang time" argument, the lower velocities of the stars at apocentre causes them to bunch up, creating a bright point. This argument is overly simplistic and does not take into account the fact that stars are more spread out at apocentre. For a small number of particles/stars, that don't adequately sample the orbital phase, it is possible this mechanism works. However, as the number of particles/stars gets larger, Kepler's laws guarantee the greater number of particles/stars at apocentre due to their low velocity is precisely balanced out by the greater area over which the stars are spread. In the large-N limit a disc of particles/stars with constant eccentricity will have a constant surface brightness around the orbit - a fact pointed out in Statler (1999). In fact the original eccentric nuclear disc models (Tremaine, 1995; Statler et al., 1999; Statler, 1999; Salow & Statler, 2001; Peiris & Tremaine, 2003) required the eccentricity to decrease outwards in the disc in order that the stars bunch up at apocentre to produce the observed surface brightness maxima. For instance, from Peiris & Tremaine (2003), the ratio of the surface brightness at apocentre Σ_a to that at pericentre Σ_p is

$$\frac{\Sigma_a(a)}{\Sigma_p(a)} = \frac{1 - ae'/(1-e)}{1 + ae'/(1+e)} \quad , \tag{2.14}$$

where $e' = \frac{de}{da}$. To quote Peiris & Tremaine (2003): "To meet the condition that the disk have a substantially higher surface brightness at apoapsis we require $e' < 0$ and either e or $|ae'|$ of order unity". The same conclusion was already noted in Tremaine (1995), Statler et al. (1999) and Statler (1999). So it appears that, while the detailed eccentric disc models constructed in Tremaine (1995), Statler (1999) and Peiris & Tremaine (2003) work as an explanation for the eccentric nuclear discs, the simple "apocentre hang time" argument used originally in Tremaine (1995) to clarify the model is wrong. Unfortunately it appears more people are familiar with the incorrect 'apocentre hang time' argument of Tremaine (1995) than the mathematics and detailed discussion of the eccentric disc model presented in that paper which remain essentially correct.

This "apocentre hang time" argument has had unfortunate influence on the debris disc community where it is has led to the erroneous assertion that debris discs with a constant eccentricity have a surface density maximum at apocentre and thus should be brighter at apocentre at long wavelengths (e.g. Pan et al. (2016)).

2.1.3 Planetary rings

Next density wave theory was adapted to explain non-axisymmetric structures in planetary rings (Goldreich & Tremaine, 1978a,b; Shu, 1984). Again this was motivated by the existence of resolved substructures, in this case in the planetary rings present in our solar system, particularly those of Saturn.

Quoting Shu (1984):

"Superficially, apart from a difference in scale of a factor of a trillion, the disks of spiral galaxies and of Saturn's rings would seem to have many similarities."

Of course Shu (1984) then goes on to point out the chief difference between the two systems, namely the difference in the strength of the collective effects. The mass of a planetary ring system, and gaseous discs, is dominated by the central object and its gravity is the dominant force acting on the ring/disc (much more so than in galaxies).

An important development of the planetary ring literature, which has subsequently been applied to gaseous discs, is the idea that density waves can be launched in the ring by an orbital companion at one of the Lindblad resonances (Goldreich & Tremaine, 1978b). Spiral density waves carry angular-momentum and can transport it far from their origin before damping (Goldreich & Tremaine, 1978a,b). The angular momentum transported and dissipated by these waves can be responsible for clearing gaps in the disc (Goldreich & Tremaine, 1978a). If a spiral density wave, excited at a Lindblad resonance, carries angular momentum outwards and damps in the disc then the angular momentum gets transferred to the disc particles, in the form of an accretion flow, causing them to move inwards towards the resonance (Shu, 1984). If this effect is strong enough it can carve a gap in the disc. This effect has been used to explain some of the features in Saturn's rings, notably the Cassini division, which is cleared by the $2:1$ resonance with the moon Mimas (Goldreich & Tremaine, 1978b; Cuzzi et al., 1981); the outer edge of the A ring which is maintained by a $7:6$ resonance with the moons Janus and Epimetheus (Lissauer et al., 1985); and the Encke division which is cleared by the moon Pan via tidal truncation (Cuzzi & Scargle, 1985; Borderies et al., 1989; Showalter, 1991), similar to the gaps caved by planets in protoplanetary discs (Lin & Papaloizou, 1979; Lin & Papaloizou, 1986; Borderies et al., 1989).

The evolution of eccentricity in planetary rings, subject to the planetary quadrupole, self gravity, collisional viscosity and excitation by companions, was derived by Borderies et al. (1983) who found it was possible to obtain stable coherently precessing eccentric rings. This was motivated by the measured eccentricity (and inferred eccentricity gradients) of the rings of Uranus, in particular the ϵ ring (Nicholson et al., 1982), along

with the Titan and Huygens ringlets of Saturn (Esposito et al., 1983; Porco, 1983; Porco et al., 1984a). The amplitude of the eccentricity was found to be set by a balance between the viscous damping and excitation by companion moons. In many of these planetary rings the eccentricity of the ring is very low, but the eccentricity gradients can be high (e.g. the ϵ ring of Uranus has $e \approx 8 \times 10^{-3}$ but $a\frac{de}{da} \approx 0.6$ (Nicholson et al., 1982)). Equivalent results would later be found for eccentric gas discs.

Several planetary rings are eccentric despite lacking an accompanying satellite capable of exciting the observed eccentricities (e.g. the α and β rings of Uranus (Elliot, 1984), and the Maxwell and Colombo ringlets of Saturn (Porco, 1990)). This lead to a search for other mechanisms to excite eccentricity in rings. One mechanism, which would become relevant to eccentric gaseous discs, is excitation by the viscous overstability (Borderies et al., 1985, 1986). Depending on the rheological model of the ring material, density wave can be excited, rather than damped, by viscosity (Borderies et al., 1985). Thus it was suggested that the eccentricity in these rings is an eccentric standing wave excited to observable amplitudes by this viscous overstability (Borderies et al., 1985).

2.1.4 Early work on eccentric gas discs

In gaseous discs the restoring force responsible for density waves is the disc pressure instead of (or as well as) the disc self gravity. Linear theory of non-axisymmetric disturbances in a vertically averaged self-gravitating ring/disc including gas pressure was carried out for a global radially narrow ring (Papaloizou & Lin, 1989) and for a radially extended disc in Papaloizou & Savonije (1991). In the WKB limit, the dispersion relation for a linear m-armed wave propagating in a disc with self gravity and pressure, that is stationary in a frame rotating with angular frequency ω (or equivalently has angular pattern speed ω in the inertial frame), is (Goldreich & Tremaine, 1978b)

$$(\Omega - \omega)^2 m^2 - \kappa^2 = c_0^2 k^2 - 2\pi G\Sigma|k| \quad , \tag{2.15}$$

where Ω is the local angular velocity, κ is the epicyclic frequency, c_0 the sound speed, k the radial wavenumber. In the inertial frame the wave has angular frequency $m\omega$, while the angular frequency is $m(\omega - \Omega)$ in a frame that rotates with the local angular velocity of the disc. For a slowly precessing eccentric disc $m = 1$ and $\omega \ll \Omega$ and we can obtain an expression for the precession rate (which is the same as the wave frequency) of a linear eccentric wave in the WKB limit,

$$\omega = \frac{1}{2\Omega}\left(\Omega^2 - \kappa^2\right) - \frac{1}{2}H^2k^2\Omega + \frac{\pi G\Sigma}{\Omega}|k| \quad , \tag{2.16}$$

where we have used $c_0 = H\Omega$, with H the scale height. While of historical interest this expression, particularly the pressure term $-\frac{1}{2}H^2k^2\Omega$, has caused a couple of misconceptions that have persisted to the present day. Notably the pressure term causes retrograde precession which, while true in this (short-wavelength) limit, is sometimes overgeneralised to being the contribution pressure makes to precession in general. Later work has subsequently shown that, taking into account 3D effects, pressure can cause prograde precession (Ogilvie, 2001, 2008).

The second is in the estimates for the precession frequency and phase velocity of a eccentric Keplerian disc subject to pressure forces. Typically assuming $k \sim r^{-1}$ gives a precession frequency $\omega \sim (H/r)^2\Omega$ and group velocity $v_g \sim (H/r)c_0$, implying in a thin disc an eccentric wave will propagate much slower than the sound speed. In reality however if $k \sim r^{-1}$ then the WKB limit and equation 2.16 aren't strictly valid, so this estimate is more of a lower bound. Additionally, as indicated by equation 2.16, eccentric waves are dispersive and do not propagate at a single velocity. More generally we expect $r^{-1} \leq k < H^{-1}$ (the upper limit being required for the thin disc assumption to hold). The precession frequency can thus significantly exceed the simple estimate (and of course can switch sign), the group velocity of the eccentric wave will be $(H/r)c_0 \lesssim v_g \lesssim c_0$. As in general an eccentric disturbance will contain contributions from all wavenumbers the leading edge of the eccentric disturbance should propagate through the disc at roughly the sound speed. This shouldn't be a surprise as the limiting precession rate of a spiral density wave should be such that the wave propagation velocity, relative to the gas, exceeds the local sound speed, causing the wave to shock.

In addition to the pressure term equation 2.16 contains a contribution from self gravity. Neglecting the pressure and forced precession frequency term the self gravity causes the eccentric wave to propagate at a single velocity regardless of wavelength $v_{\rm sg} = \mathrm{sgn}(k)\frac{\pi G\Sigma}{\Omega}$. Written in terms of the Toomre Q parameter, then (for a Keplerian disc where $\kappa = \Omega$) $v_{\rm sg} = \mathrm{sgn}(k)Q^{-1}H\Omega$. Comparing this against the velocity of an eccentric wave mediated by pressure $v_{\rm pre} = H^2k\Omega$, self gravity becomes dominant when $QHk \ll 1$. So for a sufficiently thin disc, long wavelength disturbances will be communicated by self gravity rather than pressure even when $Q \gg 1$.

Early work on eccentric gas discs was motivated by the superhump phenomenon, a periodic feature in dwarf novae lightcurves, which was proposed to be caused by a precessing eccentric disc (Vogt, 1982; Osaki, 1985). Based on SPH simulations of a disc

perturbed by a companion Whitehurst (1988) suggested that eccentricity was excited by an orbital resonance with the companion. Following this, along with previous work on exciting eccentricity in planetary rings, Lubow (1991a) showed that eccentricity can be excited at eccentric Lindblad resonances, with the 3:1 eccentric Lindblad resonance being particularly important for exciting eccentricity in dwarf novae. The excitation of the eccentricity in the disc occurs through wave coupling: the eccentric wave and the tidal disturbance couple to generate an intermediate wave, and this intermediate wave can couple with the tidal disturbance to generate an eccentric wave, amplifying the initial eccentric disturbance.

In general a periodic time varying gravitational potential can be decomposed into a double Fourier series in the mean longitudes of the companion and of a particle in the disc. This leads to a sum of terms of the form $R_{m\tilde{m}}e^{i\phi_{m\tilde{m}}}$ where

$$\phi_{m\tilde{m}} = m\lambda + \tilde{m}\tilde{\lambda} \tag{2.17}$$

is a potentially resonant angle with m,$\tilde{m}$ integers, λ the mean longitude of the particle with mean motion n and $\tilde{\lambda}$ the phase of the time dependent potential with angular frequency $\tilde{n}$. Close to the resonance each term in the resonant expansion causes a forced precession of the orbital elements with precession frequency $f = \dot{\phi}^{-1}_{m,\tilde{m}}$.

When $mn + \tilde{m}\tilde{n} = 0$ we encounter a mean motion resonance. One complication of dealing with resonances in discs is that for test particles a mean motion resonance causes the precession frequency to diverge as $\dot{\phi}_{m,\tilde{m}} = 0$. For test particles this resonant singularity is resolved by libration, with the resonant angle $\dot{\phi}_{m\tilde{m}}$ oscillating around an equilibrium value rather than vanishing at the resonance. For a continuous disc this behaviour may no longer be possible if it causes orbital intersection. Instead collective effects are expected to resolve the resonant singularity. The simplest approach to resolving the resonant singularity is to displace the resonant pole away from the real axis,

$$f = \frac{1}{\dot{\phi}_{m,\tilde{m}} - is} \quad , \tag{2.18}$$

which was the approach adopted by Lubow (1991a). This results in both a finite precession rate along with excitation/damping of eccentricity. A more complete treatment of resonances in eccentric discs was carried out by Ogilvie (2007) who included the effects of dissipation. This leads both to the growth/decay of eccentricity and to the resonance having a finite width, with a peak centred on the location of the test particle resonance.

The first forays into nonlinear eccentric disc theory were carried out in Syer & Clarke (1992, 1993); Lyubarskij et al. (1994); Ogilvie (2001) and Statler (2001). Statler (2001) extended his work on self-gravitating eccentric nuclear discs (Statler, 1999; Salow & Statler, 2001) to 2D eccentric fluid discs with gas pressure, to calculate long lived eccentricity profiles in an isolated Keplerian disc. It was assumed in Statler (2001) that, in order to be long lived, an eccentric fluid disc should not do $P\,dV$ work around the orbit, under the expectation that this should cause the orbits to evolve and damp e on the thermal time. This turns out to be a very restrictive requirement and, for a 2D disc, requires the disc to be untwisted and have a constant eccentricity (except perhaps for a small region close to the inner boundary). This also places quite strict constraints on the allowed surface density profiles in the disc, resulting in a single surface density profile for a given equation of state and inner boundary location.

Statler (2001) provided a neat explanation for why the expectation from classical disc theory that the circular disc is the lowest energy state, for a fixed angular momentum, is wrong (see Figure 2.3). For a single Keplerian orbit the lowest energy state, at fixed angular momentum, is the circular orbit. However, in a disc where angular momentum and energy can be exchanged between orbits the picture is more complicated. Following the argument in Statler (2001), consider a system which conserves total angular momentum, but not energy, consisting of two, initially circular, narrow rings which can exchange energy and angular momentum. The inner ring can move to a more tightly bound circular orbit by lowering its energy by ΔE_1 and transferring ΔL angular momentum to the outer ring, so that angular momentum is conserved. Some amount of energy ΔE_2 that is lost from the inner ring is transferred to the outer ring, with $|\Delta E_2| < |\Delta E_1|$. Because of the shape of the locus of circular orbits, in angular momentum-energy space, the outer ring can become larger and more eccentric while the total energy of the system is still reduced. To quote Statler (2001): "Thus for any circular disc, there is an eccentric disc with the same angular momentum and lower energy". While one could lower the energy in the orbit still further by circularising the eccentric orbit, one runs into the problem that we can repeat this process until one orbit reaches the origin, and the other infinity. In other words, when studying accretion discs, it isn't particularly useful to ask what the lowest energy configuration for a given angular momentum is, as the answer is to accrete the entire disc (barring some negligible mass at infinity) onto the central object, at which point we no longer have a disc. Thus one cannot use the "lowest energy configuration" to argue for a circular disc. The disc proceeds from its initial state to the lowest energy

state along some path to be determined, which may or may not include an eccentric disc.

This argument of Statler (2001) is relevant to the action of viscosity in an eccentric disc and shows that the naive expectation from classical theory that viscosity acts to damp out eccentricity is overly simple. Lyubarskij et al. (1994) derived the evolutionary equations for non-stationary viscous accretion discs and showed that it is possible to obtain eccentric discs which remain on eccentric orbits as they accrete, for a range of viscosity laws. Additionally Lyubarskij et al. (1994) showed that, for an alpha viscosity, a circular disc is unstable to the development of eccentricity. A similar result was obtained by Syer & Clarke (1992, 1993) using Gauss's perturbation equations of celestial mechanics, along with a "sticky particle" simulation where a grid based pseudo-viscous force is applied to test particles moving in a Keplerian potential. This effect is a form of viscous overstability, originally studied, in the case of axisymmetric perturbations, by Kato (1978). The explanation for the viscous overstability can be framed in terms of the argument in Statler (2001). If neighbouring rings can exchange energy and angular momentum then outer disc can become eccentric by extracting energy and angular momentum from the inner disc, which causes the inner disc to accrete faster. A complementary viewpoint is that the eccentric wave can extract AMD from the accretion flow; this causes the eccentric wave to increase in amplitude and drives faster accretion in the disc.

A more complete theoretical framework for studying nonlinear eccentric fluid discs was provided by Ogilvie (2001). In this paper Ogilvie (2001) derived the equations for the evolution of the mass, angular momentum and eccentricity vector of a thin disc whose dominant motion consists of elliptical Keplerian orbits, by means of an asymptotic analysis of the fluid-dynamical equation in three dimensions. Ogilvie (2001) formulated an orbital coordinate system, adapted from the orbital elements widely used in celestial mechanics; this coordinate system is both non-orthogonal and slowly evolves due to the evolution of the orbits. Ogilvie (2001) derived the evolutionary equations for the disc orbits and surface density of a thin disc subject to weak perturbations by pressure, external perturbations and a Maxwellian viscoelastic model of the turbulent stresses. A viscoelastic stress model equipped with a relaxation time τ behaves elastically for deformations occurring on timescales much shorter than τ, but behaves viscously on timescales much longer than τ. This is intended as a model of the MRI where, under rapid deformations, the stress behaves like a magnetic field which obeys the ideal induction equation, while for slow deformations the turbulent stress model behaves like an alpha viscosity. Ogilvie (2001) confirmed the results

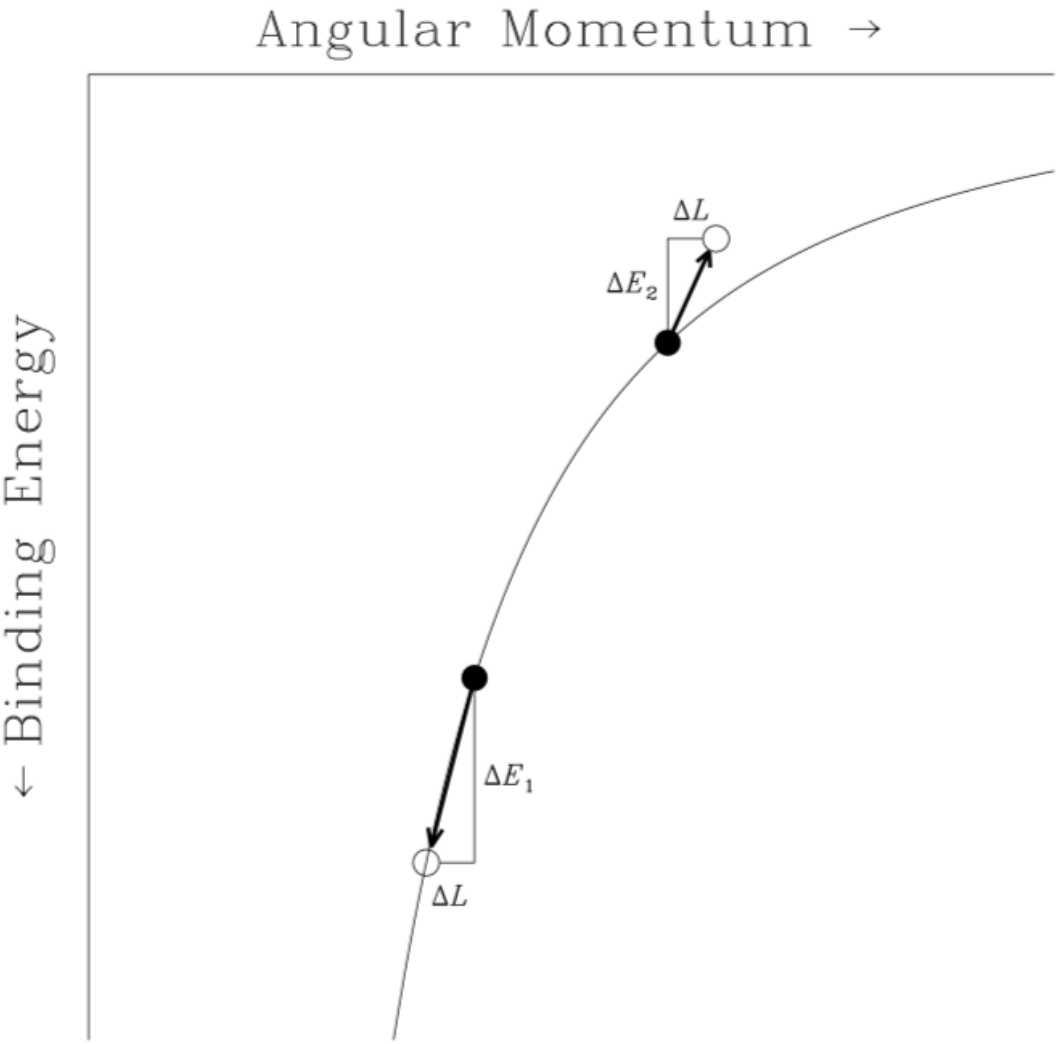


Fig. 2.3 Lindblad diagram from Statler (2001) showing that the lowest energy state in a continuous Keplerian disc need not be circular. A pair of initially circular rings (closed circles) can move to a lower energy state (open circles), while conserving angular momentum, if angular momentum can be exchanged between the rings (say by fluid forces). In the resulting state the inner ring moves to smaller radii (thus reducing the energy of the configuration), while the outer ring moves outward to conserve angular momentum, and can become eccentric. The smooth curve shows the locus of circular orbits.

of Lyubarskij et al. (1994) that a circular disc with an alpha viscosity is unstable to eccentric perturbations. However, the viscous overstability can be stabilised by including either a bulk viscosity or by a non-zero relaxation time in the turbulent stress model which reflects the finite response time of the MRI.

2.1.5 Developments in linear theory

Linearised theory (i.e. working to first order in $\mathcal{E}$ and its gradients, with $e \ll 1$) for a 2D eccentric disc was developed by various authors (Okazaki, 1991; Savonije & Heemskerk, 1993; Okazaki, 1997; Tremaine, 2001; Papaloizou, 2002; Papaloizou & Savonije, 2006; Goodchild & Ogilvie, 2006), who obtain different equations for the dynamics depending on how the vertical averaging was carried out. Linearised theory for 3D eccentric discs was derived by Ogilvie (2008), who found important differences in the dynamics of the eccentricity compared with 2D theory.

Unlike circular discs, the vertical structure of an eccentric disc cannot be decoupled from the horizontal fluid motion. Eccentric discs cannot maintain hydrostatic equilibrium owing to the variation of vertical gravity, and possible lateral compression, around an elliptical orbit. This drives vertical motions in the disc, with the disc scale height varying around the orbit, creating a form of disc breathing mode; this variation modifies the pressure and affects the horizontal dynamics (Ogilvie, 2001, 2008; Ogilvie & Barker, 2014).

As there is no hydrostatic equilibrium for an eccentric disc it is non-trivial to define the scale height of the disc. A definition was provided by Ogilvie & Barker (2014) who used

$$H^2 = \frac{\int \rho z^2 \, dz}{\int \rho \, dz} \quad , \tag{2.19}$$

this is the vertical standard deviation of the mass distribution and corresponds to the scale height of circular theory in the limit of a circular disc. Ogilvie & Barker (2014) derived an expression for the scale height in a linearly eccentric disc,

$$H = H_0 \left[1 + \mathrm{Re} \left(\frac{-3\mathcal{E} + (\gamma - 1 - i\alpha_b)\lambda \mathcal{E}_\lambda}{\gamma - i\alpha_b} e^{-i\phi} \right) \right] \quad , \tag{2.20}$$

where γ is the ratio of specific heats of the gas, α_b the dimensionless alpha associated with a bulk $\alpha-$viscosity and the subscript λ denotes a partial derivative with respect to λ. For an untwisted, inviscid disc this solution is symmetric about pericentre, where the scale height is also smallest. The scale height variation is caused both by the presence of eccentricity and (except when the disc is isothermal) gradients in the

complex eccentricity. Dissipation or disc twist causes the scale height to be asymmetric about pericentre. As we shall see later, these features carry over into the nonlinear solutions.

Written in terms of complex eccentricity $\mathcal{E}$ the linear partial differential equation governing the propagation of eccentricity in a disc, due to pressure forces, is a form of Schrödinger equation (Ogilvie, 2008),

$$i\frac{\partial}{\partial t}|\mathcal{E}\rangle = \mathcal{H}|\mathcal{E}\rangle \quad . \tag{2.21}$$

where $\mathcal{H}$ is a linear operator which corresponds to the Hamiltonian of the Schrödinger equation, and we have introduced the bra-ket notation of Dirac. Associated with this equation is a Hilbert space with an inner product defined as follows,

$$\langle f|g\rangle := \int_{\mathbf{D}} \Sigma a^2 \Omega f^* g \, 2\pi a da \quad , \tag{2.22}$$

which is a weighted integral over the entire domain, with the weight being the angular momentum. With this inner product defined, the angular momentum deficit in the disc is $\frac{1}{2}\langle\mathcal{E}|\mathcal{E}\rangle$. The simplest solutions to equation 2.21 are the eccentric modes and eccentric travelling waves, which correspond to the eigenfunctions of the operator $\mathcal{H}$, and, written in the position basis, take the form

$$\mathcal{E} = \langle \mathbf{r}|\mathcal{E}\rangle = \mathcal{E}_0(a)e^{i\omega t} \quad , \tag{2.23}$$

here $\mathbf{r}$ is just a label, with $|\mathbf{r}\rangle$ the basis vectors of the position basis, i.e. they express the eccentric anomaly as a function of position. For these solutions the disc undergoes rigid precession, with precession frequency ω, and the disc geometry is fixed in the co-rotating frame. These correspond to the states of the Schrödinger equation with energy $-\omega$. In general ω can be complex with $\mathrm{Im}[\omega]$ representing the growth/damping rate of the mode. When E_0 is real and satisfies suitable boundary conditions these solutions are called eccentric modes and correspond to the standing waves in the disc.

For an eccentric disc only subject to perturbations from pressure gradients and secular departures from the Keplerian point mass potential (i.e. in the absence of disc self-gravity, resonances or viscous/turbulent dissipation), $\mathcal{H}$ is a Sturm-Liouville operator and solving for the eccentric modes becomes a Sturm-Liouville Problem (Goodchild & Ogilvie, 2006; Ogilvie, 2008). This means that for a given boundary condition there are a countably infinite number of eccentric modes. The precession rates of the eccentric modes $\{\omega_k\}$ can be ordered such that $\omega_0 > \omega_1 > ... > \omega_n > ...$, so there are (formally) an infinite number of retrograde precessing modes. For 2D

discs (with only pressure and a point mass potential) typically $\omega_0 < 0$ so all the modes precess in a retrograde sense. For 3D discs (or a disc around an object with a deep enough quadrupole potential) ω_0 is typically positive, so there can be a finite number of prograde precessing modes, these modes will also have the smallest number of nodes. Denoting the eigenfunction associated with the $k-$th eigenfrequency ω_k by $|a_k\rangle$. These form an orthonormal basis under the inner product (equation 2.22) with $\langle a_i|a_j\rangle = \delta_{ij}$. This allows any solution to the time dependent eccentric disc equation to be decomposed into independently precessing eccentric modes,

$$|\mathcal{E}\rangle = \sum_k \mathcal{E}_k e^{i\omega_k t}|a_k\rangle \quad , \tag{2.24}$$

with

$$\langle a_k e^{i\omega_k t}|\mathcal{E}\rangle = \mathcal{E}_k \quad . \tag{2.25}$$

This decomposition, along with the inner product structure, allows for the definition of a projection operator which picks out a particular mode from the solution,

$$\mathcal{P}_k := |a_k e^{i\omega_k t}\rangle\langle a_k e^{i\omega_k t}| = |a_k\rangle \int dr\, 2\pi\Sigma r^3 \Omega a_k(r) \quad . \tag{2.26}$$

This operator projects out all the eccentric modes except for $|a_k\rangle$ and allows modes to be considered separately.

Consider now the situation where $\mathcal{H}$ can be expressed as the sum of a Sturm-Liouville operator $\mathcal{L}$, which contains the contributions from pressure and secular departures from the Keplerian point mass potential, and a perturbation from a linear operator $\epsilon\mathcal{D}$, with $\epsilon \ll 1$, that includes contributions from the disc self gravity and dissipative processes. Then equation 2.21 becomes

$$i\frac{\partial}{\partial t}|\mathcal{E}\rangle = \mathcal{L}|\mathcal{E}\rangle + \epsilon\mathcal{D}|\mathcal{E}\rangle \quad . \tag{2.27}$$

If ω_k, $|a_k\rangle$ and $\mathcal{P}_k$ are the k-th eigenvalue, eigenfunction and projection operator for $\mathcal{L}$ then applying the projection operator to equation 2.27 we obtain

$$i\frac{\partial}{\partial t}\left(\mathcal{E}_n e^{i\omega_n t}\right)|a_n\rangle = -\omega_n \mathcal{E}_n e^{i\omega_n t}|a_n\rangle + \epsilon|a_n\rangle \sum_k \langle a_n|\mathcal{D}\mathcal{E}_k e^{i\omega_k t}|a_k\rangle \quad , \tag{2.28}$$

which can be rearranged to obtain an equation for the slowly evolving coefficients $\mathcal{E}_n$,

$$\frac{\partial \mathcal{E}_n}{\partial t} = -i\epsilon \sum_k \mathcal{E}_k e^{i(\omega_k - \omega_n)t} \langle a_n | \mathcal{D} | a_k \rangle \quad . \tag{2.29}$$

Averaging over the "fast" precession time then, as in general the different eigenfrequencies aren't resonant (see for instance the expressions for ω_k in Ogilvie (2008)), this simplifies to

$$\frac{\partial \mathcal{E}_n}{\partial t} = -i\epsilon \langle a_n | \mathcal{D} | a_n \rangle \mathcal{E}_n \quad . \tag{2.30}$$

So the main effect of the forces encapsulated by the operator $\mathcal{D}$ is to modify the (complex) precession frequency of the mode and, at lowest order, the eigenfunctions remain independent of each other. Through a similar line of reasoning it should be possible to show that for a disc and system of planets the general disc state can be decomposed into a (countably) infinite number of coherently precessing disc-planet modes (analogous to those found in Teyssandier & Ogilvie (2016, 2017)). For a disc with complex eccentricity $\mathcal{E}(a)$ and a system of n planets with complex eccentricities $\mathcal{E}^1$,...,$\mathcal{E}^n$ then, assuming that the dynamics are dominated by the disc pressure forces and a fixed departure from a Keplerian point mass potential, the disc state has the form

$$|\mathcal{E}, \mathcal{E}^1, ..., \mathcal{E}^n\rangle = \sum_{k=0}^{\infty} A_k e^{i\omega_k t - i\langle a_k, p_k^1, ..., p_k^n | \mathcal{D} | a_k, p_k^1, ..., p_k^n \rangle t} |a_k, p_k^1, ..., p_k^n\rangle, \tag{2.31}$$

where $|a_k, p_k^1, ..., p_k^n\rangle$ is the eigenfunction for the k-th mode. This property, which is a general properly of the Schrödinger equation, was implicitly used by the simplified disc planet interaction model of Teyssandier & Lai (2019), who proposed integrating the eccentric disc equations with $\int dr\, \Sigma r^3 \Omega f(r)$, where $f(r)$ is an ansatz for the radial profile of the dominant eccentric mode and the integration acts like an approximate projection operator.

Concretely, the linear equation for the evolution of complex eccentricity in a 2D adiabatic disc, including dissipation through a bulk viscosity term and precessional forces (such as from GR or from an orbital companion) through $\omega_f := \frac{1}{2n}(\Omega^2 - \kappa^2)$, where Ω is the angular frequency, κ is the epicyclic frequency and n the mean motion of the Keplerian orbit, was given by Teyssandier & Ogilvie (2016) :

$$-\Sigma r^2 n \left(\omega_f \mathcal{E} + i\dot{\mathcal{E}}\right) = \frac{1}{r}\frac{\partial}{\partial r}\left(\frac{1}{2}(\gamma - i\alpha_b) P r^3 \frac{\partial \mathcal{E}}{\partial r}\right) + \frac{r}{2}\frac{dP}{dr}\mathcal{E} \quad , \tag{2.32}$$

along with a similar equation for a 3D disc,

$$-\Sigma r^2 n\left(\omega_f \mathcal{E}+i\dot{\mathcal{E}}\right)=\frac{1}{r}\frac{\partial}{\partial r}\left[\frac{1}{2}\left(2-\frac{1}{\gamma}-i\alpha_b\right)Pr^3\frac{\partial \mathcal{E}}{\partial r}\right]+\frac{1}{2}\left(4-\frac{3}{\gamma}\right)r\frac{dP}{dr}\mathcal{E} + \frac{3}{2}\left(1+\frac{1}{\gamma}\right)P\mathcal{E} \quad . \tag{2.33}$$

This highlights the effect of the 3D structure as in addition to slightly modifying some of the coefficients the presence of the breathing mode in the disc adds an additional term $\propto E$ which is the origin of the prograde precession in the 3D discs. Similar differences occur between the linear equation for complex eccentricity in 2D and 3D isothermal discs.

The WKB approximate solution to the linearised eccentric disc equations was obtained by Ferreira & Ogilvie (2008), who applied it to eccentric waves propagating in the inner regions of a black hole disc. One important result of this analysis was to show that the eccentric waves become highly oscillatory as they approach the marginally stable orbit with both the eccentricity and eccentricity gradients diverging (see Figure 2.4), so that the linear theory predicts its own breakdown. This motivates the development of a nonlinear shortwavelength approximation for the eccentric waves to explore what happens when the linear theory breaks down, which will form a part of this thesis.

Following Ferreira & Ogilvie (2008) the WKB solution to equation 2.33 is

$$\mathcal{E}=\mathcal{E}_0\left(P\Sigma r^6(\omega_f-\omega)\right)^{-1/4}\exp\left[\pm i\int\sqrt{\frac{2\Sigma(\omega_f-\omega)}{\left(2-\frac{1}{\gamma}-i\alpha_b\right)P}}dR\right] \quad . \tag{2.34}$$

This solution is useful in highlighting some of the properties of linear eccentric waves. The first is that in order to be wavelike we must have $\omega_f-\omega>0$; this can occur either due to the disc globally precessing in a retrograde sense, or due to the presence of non-Keplerian forces which cause disc precession, such as GR. The eccentric wave in the WKB limit is a slowly varying dispersive wave, whose amplitude and wavelength changes on the scale of the variation of pressure, surface density and precessional forces in the disc. The presence of a bulk alpha viscosity $\alpha_b\neq 0$ causes an exponential damping of the wave along with a shift in the wave phase.

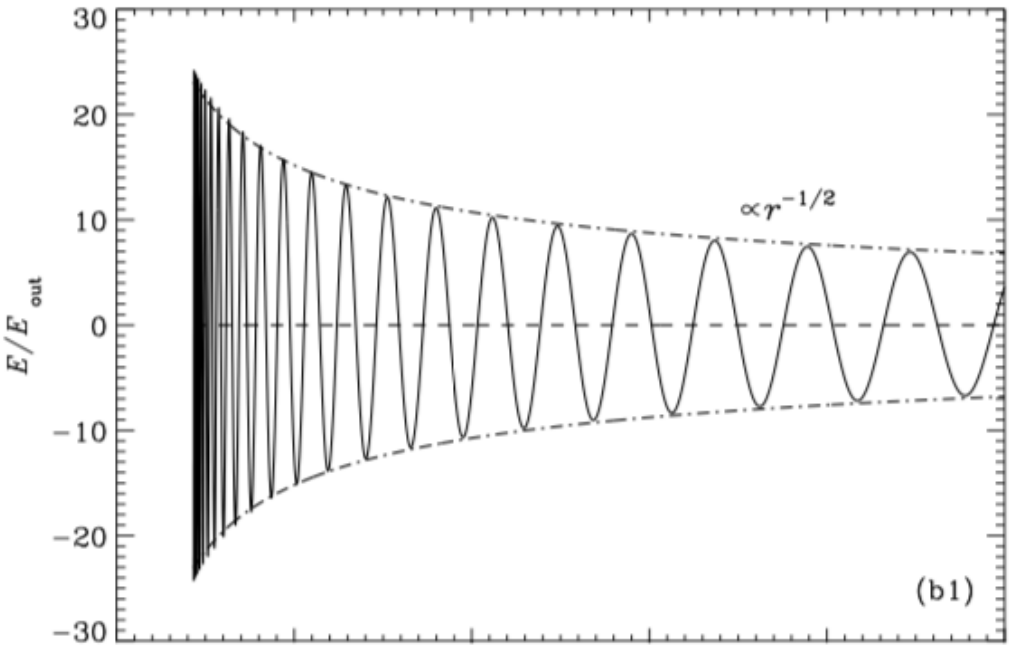


Fig. 2.4 Eccentricity in the interior of a black hole disc from Ferreira & Ogilvie (2008), with envelope predicted by WKB theory. The solution becomes highly oscillatory as it approaches the marginally stable orbit, resulting in a breakdown of linear theory.

2.1.6 Later developments in nonlinear theory

A local linear stability analysis of a 3D unstratified fluid disc with eccentric streamlines was carried out by Papaloizou (2005a) who found that the periodic variation around the streamline leads to a parametric instability which excites inertial waves. As the spectrum of inertial-gravity waves is dense in a nearly circular orbit there will always be a mode which is resonant with the orbital motion, allowing for the excitation. Generically, the resonant inertial-gravity wave will have non-zero vertical wavenumber and will not be excited in a 2D eccentric disc.

Ogilvie & Barker (2014) rederived the nonlinear eccentric disc theory of Ogilvie (2001) from an alternative viewpoint, by deriving a continuum analogue to the Gauss perturbation equations of celestial mechanics. By use of the orbital coordinate system and applying a multiple timescale analysis the fast orbital timescale can be separated from the timescale over which the orbits evolve due to fluid forces and secular forcing (e.g. from an orbital companion). This allows the equation for the evolution of complex eccentricity to be expressed in terms of orbit averaged stress gradients. Using this formalism Ogilvie & Barker (2014) derived a local model for an eccentric disc, which generalised the commonly used Cartesian shearing box (Goldreich & Lynden-Bell, 1965; Hawley et al., 1995; Latter & Papaloizou, 2017) to eccentric orbits (see Figure 2.5). In this model the domain consists of a box centred on a reference orbit which moves with the orbital motion; unlike the circular shearing box, an eccentric shearing box deforms as it moves around the orbit due to the non-uniformity of the orbital motion (See Figure 2.5). By use of a shearing coordinate system the geometry and background quantities can be made to depend on the Lagrangian time coordinate only, making the box horizontally homogeneous. This is advantageous for setting boundary conditions as this allow for periodic/shear periodic boundary conditions. A generalisation to the MHD equations, which will be important for the latter part of this thesis, was included in their Appendix C.

Ogilvie & Barker (2014) derived the simplest solutions to their local model, being a horizontally invariant laminar flow which undergoes vertical expansion and contraction due to the breathing mode. An important result of this calculation was in showing that for low ratio of specific heats $\gamma < 1.5$ the variation of the scale height around the orbit becomes extreme as $e \to 1$. Similarly for high $\gamma > 1.5$ extreme compression occurs when the eccentricity gradients are large. A linear stability analysis of this solution was carried out in Barker & Ogilvie (2014) who found that the disc was unstable to the parametric excitation of small scale inertial waves (the parametric instability of Papaloizou (2005a,b), but with significantly higher growth rates as the instability can

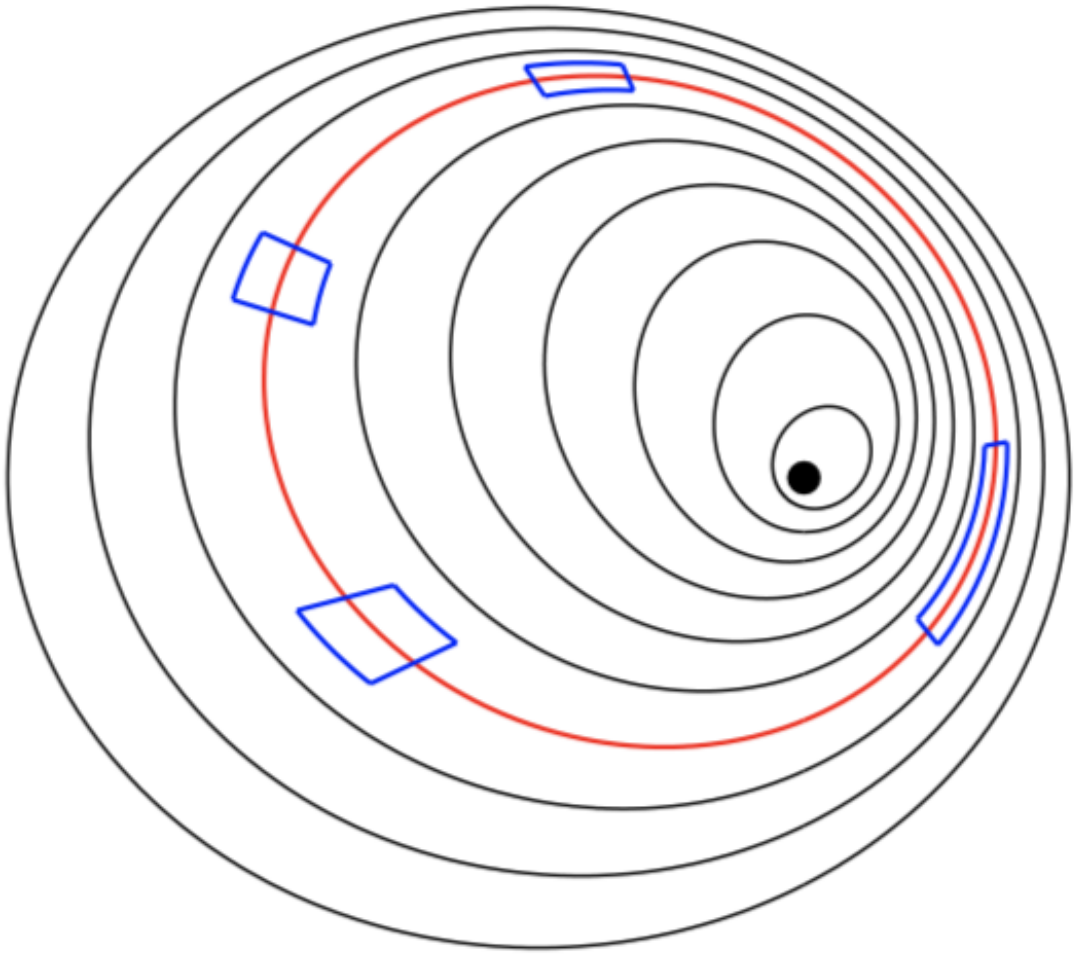


Fig. 2.5 Eccentric shearing box model from Ogilvie & Barker (2014).

feed off the breathing mode). Wienkers & Ogilvie (2018) used the eccentric shearing box framework to study the parametric instability, while Chan et al. (2018) used the eccentric shearing box to study the MRI in an eccentric disc, although the authors of that paper looked at an incompressible unstratified box which, due to the presence of the breathing mode, is only likely to be applicable to weakly eccentric discs.

2.2 Simulations of eccentric discs

In addition to the analytical work reviewed in the previous section, eccentric discs have been explored through numerical simulations. Disc eccentricity is particularly prevalent in simulations of discs with orbital companions (internal and external) due to the excitation of eccentricity at the 3:1 Lindblad resonances (Lubow, 1991a).

One challenge of studying eccentric discs in simulations is the significant damping of the eccentric wave which can occur due to artificially strong fluid effects present in most numerical simulations. For reasons of numerical stability and computational efficiency, simulated discs are typically thicker and more viscous than their astrophysical counterparts. Both of these factors lead to lower amplitude eccentric waves, due to the greater dissipation. Additionally thicker discs typically have stronger pressure forces, which require stronger forcing to excite/maintain an eccentric wave of a given amplitude.

This means that certain numerical techniques, notably SPH, are particularly poorly suited to studying eccentric discs due to the relatively thick discs and the requirement that artificial bulk viscosity be included for numerical stability. It has also presented challenges for studying realistically thin 3D eccentric discs. Currently the state of the art numerical simulations of eccentric discs are mostly 2D in order to attain realistically low viscosity and thin discs conducive to the propagation of eccentric waves. However, as discussed in the previous section, several important dynamical effects are missing in a 2D disc which are important for 3D eccentric discs. The relatively large numerical or artificial viscosities in (global) 3D disc simulations can be viewed as a stand in for unmodelled disc turbulence, in hydrodynamical simulations of disc where the MRI is expected to operate, although such simulated discs have historically been relatively thick. The extent to which such viscous models capture the effects of the MRI is debatable, particularly when interacting with an eccentric disc. Only recently have we had the computational resources necessary to begin the study of 3D eccentric discs with realistic parameters for a MRI dead-zone.

Due to the historical interest in explaining the superhump phenomena, early simulation efforts of eccentric discs focused on validating the excitation of eccentricity in a disc by the 3:1 resonance of an orbital companion. Despite its drawbacks, early simulations of the superhump phenomenon where carried out using SPH, starting with Whitehurst (1988). 2D SPH simulations (Whitehurst, 1988; Lubow, 1991b; Kunze et al., 1997; Murray, 1998) found evidence for the excitation of an eccentric deformation by the 3:1 resonance, superimposed on an $m = 2$ tidal deformation. A 3D SPH simulation of the superhump phenomenon was carried out by Smith et al. (2007), who similarly found a superposition of the eccentric and $m = 2$ deformation. Unfortunately the expectation in Smith et al. (2007) that pressure forces cause retrograde precession of the disc meant that the prediction of 3D eccentric disc theory, that the breathing mode contributes to a prograde precession from pressure forces, was not tested by this work.

The grid based simulation of Kley et al. (2008) studied the dynamics of a 2D eccentric disc around a white dwarf in a CV system and found that the disc could be excited to high eccentricities by the companion - to the extent that the outer part of the disc can periodically spill over the white dwarf Roche lobe (see Figure 2.6). This simulation confirmed the existence of a coherently precessing eccentric disc with a precession period consistent with the requirements for the superhump phenomenon. Kley et al. (2008) found their solution appears to strongly depend on what is assumed for the inner boundary conditions, with the disc eccentricity, precession rate and even precession direction dependent on the boundary condition used.

The results of Kley et al. (2008) highlight two features that are common across many simulations of eccentric discs. This first is that thin 2D eccentric discs excited by orbital companions are often far more eccentric than people expect from classical/linear theory. The second is that the discs truncate at greater distances from the companion than would be expected from either the torque balancing (Goldreich & Tremaine, 1980; Papaloizou & Lin, 1984; Lin & Papaloizou, 1986; Artymowicz & Lubow, 1994) or resonance overlap Wisdom (1980) theories. The equivalent results for circumbinary discs were shown by Thun et al. (2017) and Thun & Kley (2018) (see Figure 2.7). The presence of both these effects is likely not a coincidence as a thinner, less viscous disc allows the disc to attain much higher eccentricities, which in turn strengthens the resonances responsible for disc truncation. In particular the standard resonant expansion used to calculate the truncation radius will not be accurate for a disc attaining an eccentricity $e \approx 1/2$. It is likely that higher order resonances that are negligible for nearly circular discs become important for highly eccentric discs; as these

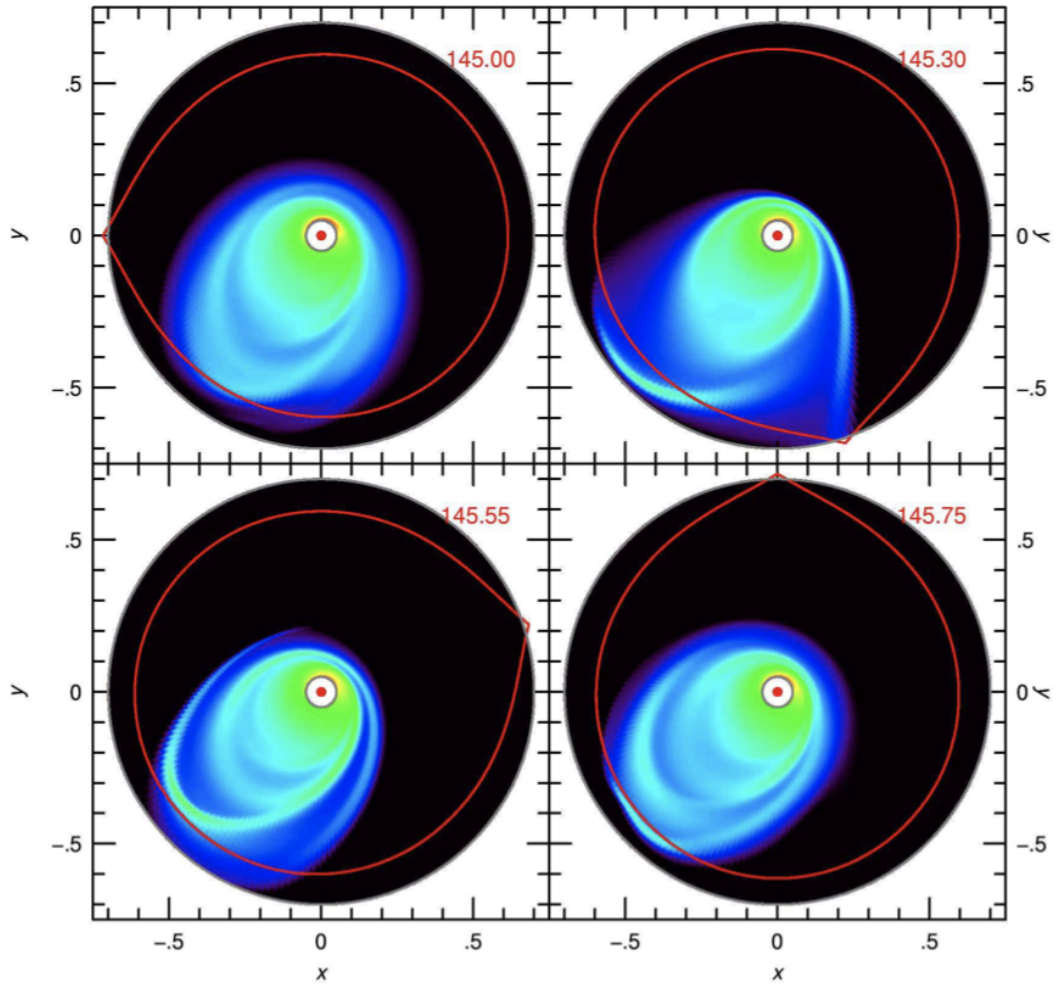


Fig. 2.6 Simulations of an eccentric disc around one component of a binary from Kley et al. (2008). The eccentricity is excited by the 3:1 Lindblad resonance and results in a precessing eccentric disc.

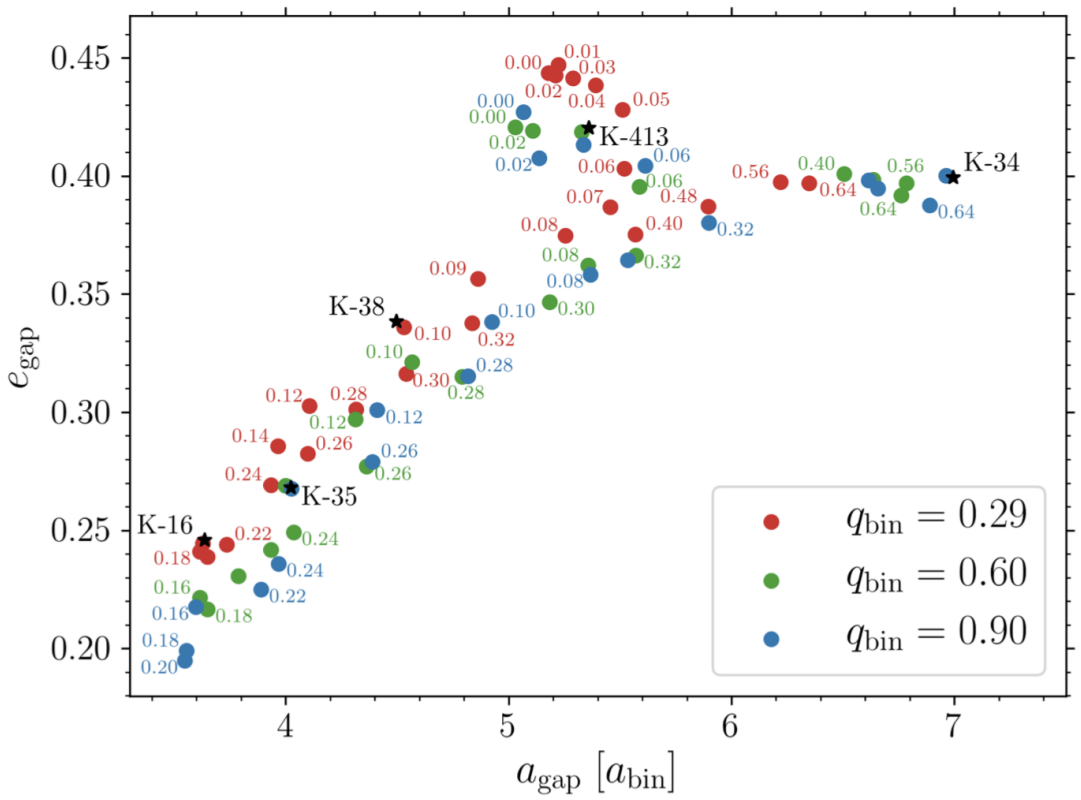


Fig. 2.7 Properties of the eccentric cavity cleared by a binary from Thun & Kley (2018).

are located further from the companion this may explain the small disc/large cavity sizes seen in the simulations.

Interestingly the recent simulations by Miranda et al. (2017) and Muñoz et al. (2019) of circumbinary discs differ from Thun et al. (2017) and Thun & Kley (2018) in that they do not find enlarged cavities, although they still find the binary can excite significant eccentricities in the disc (e.g. see Figure 2.8). The Miranda et al. (2017) and Muñoz et al. (2019) results are mostly in agreement with existing theory. The main difference between the simulations of Miranda et al. (2017) and Muñoz et al. (2019) and those of Thun et al. (2017) and Thun & Kley (2018) are the strength of the fluid forces in the disc. With the Miranda et al. (2017) and Muñoz et al. (2019) discs being (slightly) thicker and significantly more viscous (more appropriate to some MRI active disc) compared to the discs of Thun et al. (2017) and Thun & Kley (2018) (which are more applicable to a MRI dead zone). This suggests there may be a difference between MRI active circumbinary discs, where the existing theory of Artymowicz & Lubow (1994) appears to work well, and the nearly inviscid protoplanetary/post-AGB circumbinary discs where we may need to revisit the theory of disc truncation. Thun

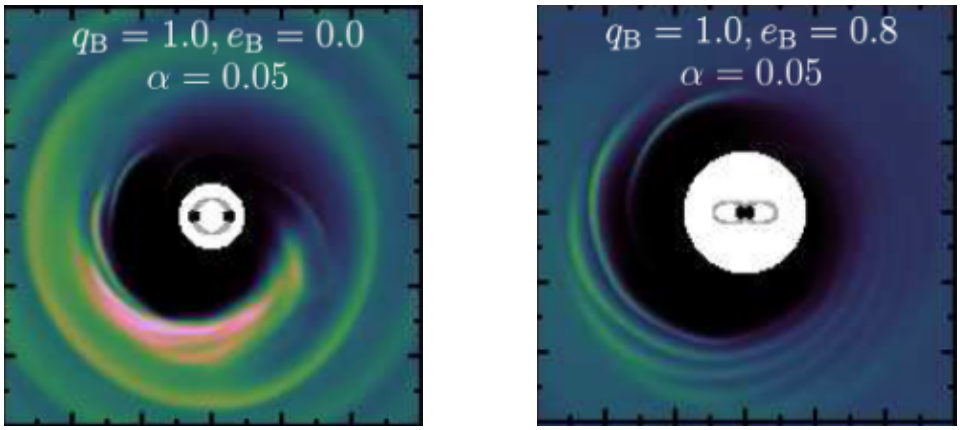


Fig. 2.8 Eccentric cavities excited in a circumbinary disc from the simulations of Miranda et al. (2017)

& Kley (2018) and Kley et al. (2019) demonstrated that, depending on their mass, planets migrating in these discs can end up on eccentric orbits close to the cavity cleared in the disc.

Eccentric discs have also been simulated in the regime where the orbital companion has a low mass relative to the central object. This is of particular interest for understanding the origin of planetary eccentricity and possible formation channels for Hot Jupiters. Kley & Dirksen (2006) simulated this situation and found that for a massive enough planet eccentricity was excited by the 3:1 resonance. Subsequent numerical simulations have been able to test the prediction of Teyssandier & Ogilvie (2016) and Teyssandier & Ogilvie (2017) that disc-planet interactions should lead to the excitation of a coherent disc-planet mode that uniformly precesses as a single system (Ragusa et al. (2017), see Figures 2.9 and 2.10). Additionally the simulations found evidence for discs switching between two different disc modes, with an associated change in the precession rate, owing to the different growth rates, a result also expected from the linear theory of Teyssandier & Ogilvie (2016) and Teyssandier & Ogilvie (2017).

In the absence of excitation and dissipation mechanisms secular eccentric disc theory predicts the existence of coherently precessing nonlinear eccentric modes. Barker & Ogilvie (2016) demonstrated the existence of long lived coherently precessing eccentric mode in a 2D hydrodynamic simulation of a disc in a Newtonian point mass potential, with the precession rates consistent with that predicted by secular theory. Barker & Ogilvie (2016) found that, in the simulation, the eccentric mode was damped owing to the presence of numerical dissipation and the presence of spiral density waves excited by the Papaloizou-Pringle Instability (Papaloizou & Lin, 1984; Papaloizou & Pringle, 1985, 1987) as a consequence of using closed boundaries.

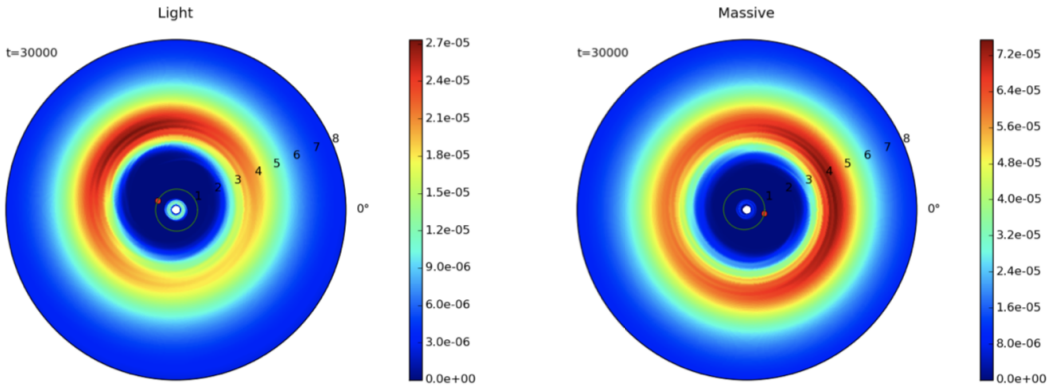


Fig. 2.9 Simulations of disc planet interaction from Ragusa et al. (2017) showing the excitation of eccentricity in the disc.

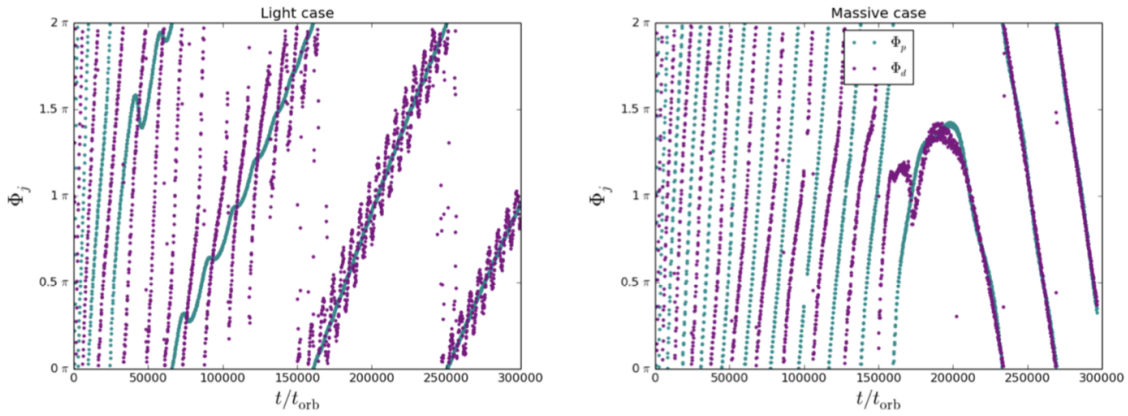


Fig. 2.10 Longitude of pericentre for the disc and planet in the simulations of Ragusa et al. (2017) showing how the disc and planet precess as a coherent combined mode as predicted by Teyssandier & Ogilvie (2016). The reversal in direction in the massive disc is likely due to a switch in the dominant mode.

Recently computational resources have allowed for the simulation of 3D eccentric discs. As yet, there is no confirmation from simulations of the dynamical vertical structure of an eccentric disc and its effect on the eccentricity dynamics - in particular the change from retrograde to prograde precession from pressure forces when going from a 2D to a 3D disc. This is because, at present, most 3D models of eccentric discs are either local or unstratified. Instead simulation efforts have focused on the other prediction of 3D eccentric disc theory, namely demonstration of the parametric instability. The earliest demonstration of the parametric instability in a simulation was in Papaloizou (2005b) who studied the evolution of a global $m = 1$ mode in an unstratified 3D disc and found that, while the eccentricity in the disc was long lived it was damped by the turbulence produced by the parametric instability. Wienkers & Ogilvie (2018) looked at the turbulence generated by the parametric instability in an eccentric shearing box, and found that the turbulence was regulated by a predator-prey dynamic between the turbulence excited by the eccentric mode and zonal flows excited by the turbulence which inhibit the parametric instability. Dewberry et al. (2020a) studied the parametric instability in a global simulation of a pseudo-Newtonian black hole disc and confirmed that the inertial waves can be confined to a narrow region close to the marginally stable orbit, a requirement of the mechanism proposed by Kato (2008) to explain one of the high frequency quasi-periodic oscillations (HFQPOs, a feature of X-ray binary spectra). Dewberry et al. (2020b) showed that this trapping could remain effective in the presence of a magnetic field, as it has been suggested that this would allow the inertial wave to escape across the marginally stable orbit (Fu & Lai, 2009; Fu & Lai, 2013). Pierens et al. (2020) performed a 3D simulation of the eccentric circumbinary discs considered by Miranda et al. (2017), Thun et al. (2017), Thun & Kley (2018) and Muñoz et al. (2019) both confirming the existence of the eccentric cavities in the 3D discs and demonstrating that the disc is unstable to the parametric instability, which drives turbulence in the inner eccentric region around the cavity cleared by the binary.

While primarily motivated as a study of excitation and trapping of inertial waves, the simulations of Dewberry et al. (2020a) and Dewberry et al. (2020b) considered the propagation of eccentric waves in Newtonian and pseudo-Newtonian discs for hydrodynamic (Dewberry et al., 2020a) and magnetohydrodynamic (Dewberry et al., 2020b) discs. The simulations are a mix of 2D and 3D unstratified domains, the latter being unstable to the parametric instability but not exhibiting the breathing mode present in a fully 3D eccentric disc. The results of these papers give valuable

information about the propagation of eccentric waves in discs, and the interaction of the eccentric wave with MRI, and will be discussed further in the body of this thesis.

Eccentric disc simulations have consistently encountered the problem of what boundary conditions should be implemented to reduce the influence of the numerical setup. Circumbinary discs, Newtonian discs around single objects and relativistic/pseudo-Newtonian discs possessing a marginally stable orbit all have their own unique challenges when it comes to the correct inner boundary conditions. The outer boundary condition can also be a challenge, particularly for simulations designed to study a coherent eccentric disc by exciting an eccentric wave at the outer disc boundary. We now review some of the challenges encountered with the boundary condition for eccentric discs in various contexts. This discussion will mostly focus on boundary conditions in 2D simulations; however, many of the issues raised are likely to be relevant to 3D simulations of eccentric discs.

The inner boundaries of most simulations have either open/outflow or closed boundary conditions. An open/outflow boundary sets the gradients of the fluid variables to zero on the boundary. To avoid inflow the radial velocity on the boundary is only allowed to exit the domain, and is set to zero in regions where it is directed back into the domain. For a closed boundary, like the open boundary, the gradients of the fluid variables are zero. However the radial velocity is set to zero. Often the angular velocity of the closed boundary is set to that required for an axisymmetric radial force balance.

A third type of boundary condition that is sometimes used is a viscous outflow condition (Pierens & Nelson, 2008; Kley et al., 2008; Mutter et al., 2017). For this boundary condition the material in the ghost zone is prescribed a radial velocity proportional to the viscous drift rate. This boundary typically ends up acting like either a closed or an open boundary depending on the viscosity (Kley et al., 2008; Thun et al., 2017).

For a simulation of a Newtonian disc around single point mass, some relevant boundary conditions and related issues are:

1. **Open/Outflow Boundary:** In this setup open boundaries appear to be ill-posed (or at least numerically poorly behaved) and can often cause the spontaneous generation of eccentricity in the disc (Kley et al., 2008). This happens if an $m = 1$ deformation occurs close to the boundary as material will flow out of the domain, but cannot return due to the diode boundary condition. This causes the formation of an asymmetric cavity on the inner disc boundary which grows due to the lack of inflow on the boundary and makes the disc eccentric. An example

where this boundary condition leads to questionable results is Teyssandier & Lai (2020), who were looking at the excitation of eccentricity in a disc with an embedded hot Jupiter, as it is not clear how much of the eccentric mode is excited by the planet and how much is excited/amplified by the boundary condition.

2. **Closed Boundary:** In this setup closed boundaries are numerically well behaved. This setup will be unstable to the Papaloizou-Pringle Instability (a form of co-rotation instability), which produces high azimuthal order spiral density waves that can damp the eccentricity in the disc by nonlinear interactions (Barker & Ogilvie, 2016). The extent to which it is relevant to a real astrophysical disc depends on how good a model of the boundary between the disc and central object a closed boundary is expected to be. One drawback of this boundary condition, for eccentric discs, is that it forces the disc to be circular on the boundary.

3. **Viscous Outflow Boundary:** As mentioned above the viscous boundary typically behaves like an open or a closed boundary depending on the viscosity. This means that viscous outflow boundaries are typically well behaved when they have a sufficiently low viscosity so that the inflow is slow and the boundary remains impervious to eccentric/density waves. As an advantage over the closed boundary conditions this boundary condition allows for a viscous steady state to be achieved.

For a simulation of a pseudo-Newtonian/relativistic disc or a disc with a marginally stable orbit, the inner edge of the disc transitions to a supersonic inflow. Some relevant boundary conditions and important considerations are:

1. **Open/Outflow Boundary:** In this setup open/outflow boundaries are numerically well behaved if they are located sufficiently far within the marginally stable orbit (Dewberry et al., 2020a). If the boundary is located too close to the marginally stable orbit then a supersonic inflow does not have the space to develop. This leads to excitation of spiral density waves via the co-rotation instability (Fu & Lai, 2013; Dewberry et al., 2020a). One issue that can be encountered is that in inviscid discs or discs with closed outer boundaries the disc cannot settle down into a steady state. This can be a problem for long term simulations of the disc, although may be less of an issue for relatively short simulations focusing on the propagation of eccentric waves on a disc background.

2. **Closed Boundary:** In this setup the closed boundary condition is ill-posed, unless it is located outside the marginally stable orbit (in which case it behaves much like it does for the Newtonian point mass disc). This is because the inflow on the inner disc boundary causes mass to build up on the boundary and causes the surface density profile to be non-convergent (Thun et al., 2017), with the shape of the azimuthally averaged surface density profile, and the size of the central cavity, dependent on numerical resolution and the code used.

3. **Viscous Outflow Boundary:** Again this behaves like either an open or a closed boundary depending on viscosity. When the viscous speed is low this boundary the surface density profile becomes non-converged as in the closed boundary case (Thun et al., 2017).

Simulations of circumbinary discs have much in common numerically with the pseudo-Newtonian/relativistic simulations. For binaries with comparable masses, the disc possesses a marginally stable orbit so when the binary's orbit is located outside the domain the simulation boundaries have the behaviour discussed above. Alternatively the binary orbit can be located within the domain. This can be achieved either through the use of moving unstructured meshes (Muñoz & Lai, 2016; Muñoz et al., 2019, 2020) or by excising a small circular region at the origin (Thun et al., 2017), with either open or closed boundaries.

It would be beneficial to be able to study eccentric waves in accretion discs in isolation from other effects, in particular in isolation from the effects of an orbital companion normally responsible for exciting the wave. One way of doing this is to use boundary conditions which excite an eccentric wave within the domain to model an eccentric wave propagating into the domain from a different part of the disc. Such a boundary condition was developed in Dewberry et al. (2020a). This boundary is closed (zero radial velocity) with a circular Keplerian angular velocity. The eccentric wave is excited by imposing a density on the ghost cells taken from the nonlinear mode calculations of Barker & Ogilvie (2016), with the density pattern precessing with precessional frequency equal to the eigenfrequency of the mode, such that $\rho(\phi, t) = \rho(\phi - \omega t)$. This leads to an eccentric wave being launched into the disc. In principle there is no reason to require the precessional frequency to equal the eigenfrequency as any frequency will set up some form of eccentricity profile in the disc. However an arbitrary frequency may lead to the solution becoming nonlinear close to the boundary which can cause issues for the stability of the hydrocode, potentially resulting in the code terminating unsuccessfully.

As discussed in section 2.1 it is now understood that the 3D structure is important for several processes in eccentric discs. Which of these processes occur in simulations of eccentric discs depends on the treatment of the vertical structure. For the strictly 2D simulations the parametric instability, breathing mode and the prograde contribution to precession made by 3D pressure forces are absent. The local 3D simulation, within the eccentric shearing box framework, of Wienkers & Ogilvie (2018) possessed both the breathing mode and the parametric instability. Unstratified global simulations of eccentric discs (Dewberry et al., 2020a,b) are unstable to the parametric instability, but do not posses a breathing mode or associated prograde precession. The choice of vertical boundary condition influences the vertical structure, in particular choosing periodic vertical boundaries leads to the formation of large scale 'elevator flows' (Dewberry et al., 2020a), a steady column of fluid with constant vertical velocity. It was suggested by Dewberry et al. (2020a) that these 'elevator flows' might be due to the formation of large scale circulation which is not adequately captured by an unstratified simulation. There have been relatively few full 3D simulations of eccentric discs carried out at reasonable resolution. Fully 3D SPH simulations (Smith et al., 2007) do not appear to be unstable to the parametric instability, likely due to artificially high numerical viscosity and insufficient resolution. The fully 3D grid code simulation of a circumbinary disc by Pierens et al. (2020) did find the parametric instability. So far there have been no reported demonstrations, within hydrodynamical simulations, of the breathing mode and associated prograde precession from pressure forces; both these effects could have been in Smith et al. (2007) but may have gone unnoticed as this prediction of eccentric disc theory was not fully appreciated at the time and any prograde precession due to pressure would have been masked by that induced by the companion. A careful look at simulation snapshots in Pierens et al. (2020) strongly suggests the breathing mode is present in their simulations in the form of a large scale anti-symmetric pattern in the meridional velocity (most noticeable in their Figure 3). The amplitude of this pattern appears comparable to the small scale turbulence attributed to the parametric instability. This suggests that a significant contribution to the vertical kinetic energy, used by Pierens et al. (2020) as a measure of the strength of the parametrically excited turbulence, comes from the breathing mode.

2.3 Theory of tidally distorted discs

Closely related to, and developed alongside, the theory of eccentric discs is the theory of tidally distorted discs, where the disc is distorted by the gravity of the orbital companion.

For low to moderate mass companions this is a common source of eccentricity in eccentric discs, however for massive enough companions there can be a strong departure from a Newtonian point mass potential which requires separate theoretical consideration from the slowly evolving Keplerian orbits of eccentric disc theory. Except for very massive discs the correct theoretical framework for understanding tidally distorted discs is in perturbations to the orbits of the restricted 3-body problem. Even for test-particles the (restricted) 3-body problem remains an unsolved problem, in so far as a general solution is not known (or would not necessarily be particularly useful), and is the subject of current research. Hence the theory of tidally distorted discs is much less advanced than that of eccentric discs; even so a lot of progress in this area has been made over the last few decades.

Broadly speaking a thin disc theory is predicated on finding an asymptotic expansion in H/R about some set of test particle orbits. To be able to make such an expansion requires that there exist closed (or nearly closed), nested, non-intersecting orbits. Paczynski (1977) was the first to obtain a set of periodic nested non-intersecting orbits around a star with an orbital companion of comparable mass on a circular orbit (see Figure 2.11 for an example taken from Ogilvie (2002a)). These orbits are initially circular close to the star but become increasingly distorted by the companion's gravity at larger radii. An important feature of these orbits is the existence of an outermost orbit, beyond which the periodic orbits intersect. If a disc extends beyond the outermost orbit fluid streamlines that follow the test particle orbits will intersect, resulting in shocks that cannot be treated within the framework of a thin disc theory. The self intersection of the fluid streamlines is potentially responsible for truncating the disc. Alternatively the disc may heat up due to shock heating so that the streamlines have significant pressure support and depart strongly from the test particle motion, and thus avoid intersecting other streamlines. Lin & Papaloizou (1993) found that low viscosity circumcomponent discs extend slightly beyond the last nonintersecting Paczynski (1977) orbit, with gas pressure preventing the crossing of the outermost streamlines. For large viscosities the disc can be much larger than the test particle orbits, and even extend beyond the components Roche lobe.

The exploration of test particle orbits in the restricted 3-body problem relevant to circumbinary and circumcomponent discs was further developed by Pichardo et al. (2005) and Pichardo et al. (2008) who advocated searching for "invariant loops". An "invariant loop" is a closed orbit which changes shape periodically with the binary orbital phase. Pichardo et al. (2005) and Pichardo et al. (2008) searched for such invariant loops by use of a shooting method. For the circumcomponent discs, starting from

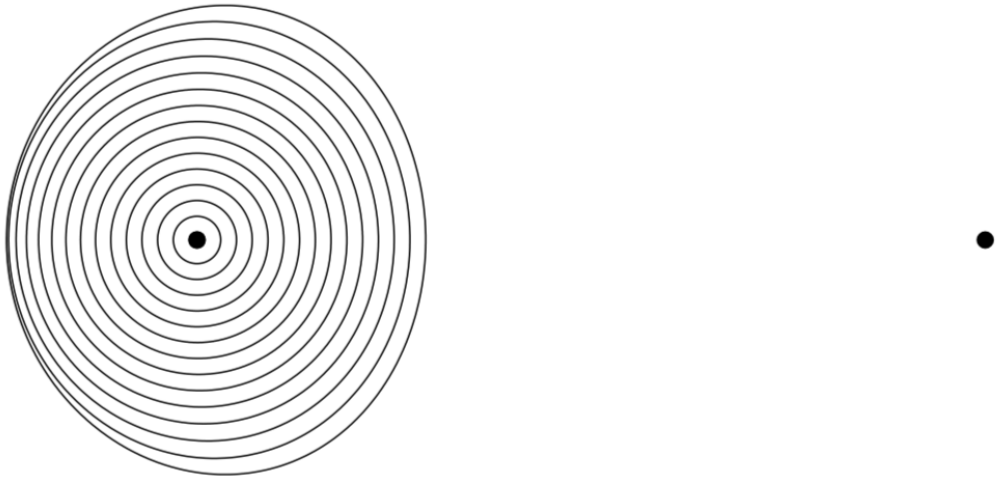


Fig. 2.11 Paczyński orbits of a prograde disc around a component of an equal mass ratio binary taken from Ogilvie (2002a). Extending the disc further out results in an orbital intersection as the periodic orbits at large radii intersect, providing a mechanism for disc truncation by the companion.

initially circular orbits near the star and proceeding outwards until the neighbouring orbits intersect leads to a generalisation of the Paczynski (1977) orbits to eccentric binaries. Similar orbits can be obtained for a circumbinary disc by starting at large radii and proceeding inwards until again neighbouring orbits intersect, or the orbits become unstable.

A thin disc model based on the Paczynski (1977) orbits was developed by Ogilvie (2002a). In a manner similar to Ogilvie (2001) for eccentric discs, Ogilvie (2002a) developed an asymptotic expansion in H/R to determine how the fluid properties vary around the Paczynski (1977) orbits. As in an eccentric disc there is no hydrostatic equilibrium, as a nonlinear forced breathing mode is set up by the variation of gravity, and the divergence of the orbital velocity field, around the orbit. The height of the disc surface varies around the orbit and Ogilvie (2002a) suggests that if the disc reprocesses light from the central object, thicker part of the disc will intercept, and reprocess, more light producing a non-axisymmetric pattern in Doppler tomograms.

In addition to distorting the shape of the disc orbits, as described above, orbital companions can influence the orientation of the orbital plane of the disc. Early work on misaligned discs around binaries showed that a disc which is marginally misaligned from the binary orbital plan can be aligned by torques from the binary (Papaloizou & Pringle, 1983; Kumar, 1987, 1990; Ivanov et al., 1999). A similar effect occurs around spinning black hole though the Bardeen-Patterson effect (Bardeen & Petterson, 1975;

Papaloizou & Pringle, 1983; Kumar & Pringle, 1985; Pringle, 1992). However SPH (and recently grid code Liska et al. (2020)) simulations of strongly misaligned discs around binary or spinning black holes have demonstrated disc breakage or tearing with the inner disc aligning to the binary angular momentum vector/black hole spin axis, while the outer disc remains misaligned (Nixon et al., 2012, 2013; Nealon et al., 2015; Doğan et al., 2015; Nealon et al., 2016).

In exploring these highly misaligned discs a second effect was found. For eccentric binaries, in addition to being aligned/anti-aligned to the angular momentum vector, there is another stable configuration where the disc orbital angular momentum vector is aligned to the binary eccentricity vector (Martin & Lubow, 2017; Zanazzi & Lai, 2018; Lubow & Martin, 2018), perpendicular to the orbital plane of the binary. For viscous discs it has been shown that the disc can align to either the angular momentum vector of the binary or eccentricity vector depending on the initial configuration, with the size of the basin of attraction for alignment to the eccentricity vector increasing with larger binary eccentricity (Zanazzi & Lai, 2018). Initial work on the truncation of these polar discs indicates that they truncate closer in than their coplanar counterparts (Miranda & Lai, 2015; Franchini et al., 2019). Additionally it has been shown that more massive discs have a greater deviation from alignment for both the polar and coplanar configurations (Martin & Lubow, 2019).

Alignment of discs with the binary eccentricity vector due to fluid forces has been demonstrated in SPH simulations (Martin & Lubow, 2017, 2018; Franchini et al., 2019; Martin & Lubow, 2019). These have also backed up the expectation that polar discs are truncated at smaller radii than their coplanar counterparts (Franchini et al., 2019). Recently there has been an observational confirmation of the existence of the polar configuration by Kennedy et al. (2019), who found that the debris disc in the HD 98800 system is perpendicular to the central binary.

As discussed in Section 2.1.1 an orbit that is sufficiently inclined with respect to an orbiting companion can undergo Lidov-Kozai oscillations. For a fluid disc this results in the excitation of eccentric waves in the disc (Martin et al., 2014; Fu et al., 2015a,b, 2017; Lubow & Ogilvie, 2017; Zanazzi & Lai, 2017). Fluid discs and orbital companions (such as planets) respond differently to the binary potential (primarily due to the dissipation in the disc). Alignment of the disc with the binaries angular momentum or eccentricity vector, due to dissipative effects, may be one way of ending up with a large mutual inclination with an external planetary companion. This companion could then excite large eccentricities in the disc.

Chapter 3

Astrophysical Context of Distorted Discs

3.1 Observational evidence for eccentric discs

3.1.1 Indirectly inferred disc eccentricity

Early indirect evidence for eccentric discs came from the superhump phenomenon in lightcurves of cataclysmic variables (CVs) during superoutburst (e.g. see Figure 3.1). The superhump is a periodic modulation of the CV lightcurve with a period close to, but not equal to, the binary orbital period (Patterson, 1979; Vogt, 1980). The superhump phenomena appear in CVs systems with small mass ratio (< 0.25) (Whitehurst, 1988).

An initial attempt to explain the superhump phenomena using a precessing eccentric disc was due to Osaki (1985), who suggested that the mass accretion rate onto the disc varies due to variable irradiation of the donor by the precessing eccentric disc. The current precessing eccentric disc model, where the outer parts of the accretion disc experience a periodic variation in the tidal torque from the companion was developed by Whitehurst (1988), Osaki (1989), Hirose & Osaki (1990) and Lubow (1991a). In this model the varying tidal torque is responsible for removing angular momentum from the outer parts of the disc and thus drives a variable accretion rate (and thus luminosity) which depends on the orientation of the eccentric disc, with the maxima in the luminosity occurring as the disc apocentre passes the donor resulting in the greatest tidal torque. The model explains the mismatch between the binary period and the superhump period; the superhump period equals the precession period of the disc in the frame co-rotating with the binary. The precession frequency of a thin disc in the

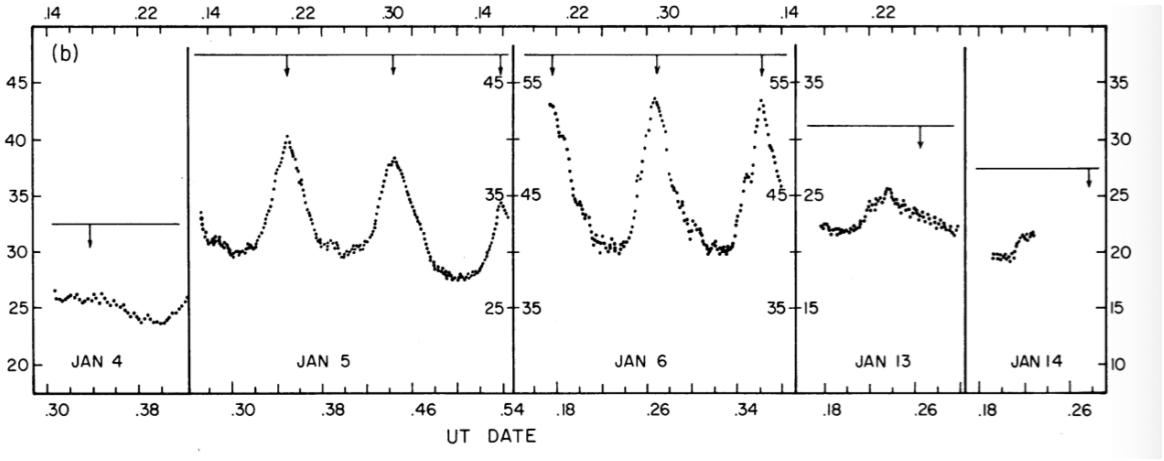


Fig. 3.1 Observations of YZ Cnc during superoutburst, clearly exhibiting the superhump phenomena with periodic variations in the lightcurve. From Patterson (1979).

inertial frame is slow compared to the orbital frequency, which results in the precession frequency in the co-rotating frame having a small departure from the binary orbital frequency. The model also explains why superhumps were only found in systems with small mass ratios, as it's only in these systems that the disc can grow large enough to encounter the 3 : 1 Lindblad resonance responsible for exciting the eccentric disc (Whitehurst, 1988; Lubow, 1991a).

Superhumps are only observed in superoutbursts and not regular outbursts (Lasota, 2001). Within the framework of the disc instability model regular outbursts occur due to the thermal-viscous instability, with the disc too small to reach the 3:1 resonance explaining the lack of superhump. Each outburst only accretes a small fraction of the mass built up during quiescence, so that the disc size increases with each successive outburst. Eventually the disc reaches the 3 : 1 resonance triggering the superoutburst and the excitation of the superhump, along with the accretion of most of the material built up over multiple outburst cycles (Osaki, 1989; Ichikawa et al., 1993; Osaki, 1996; Lasota, 2001). The superhump could be responsible for the superoutburst through the enhanced tidal torque, or the superoutburst maybe trigger separately and is the only outburst where the disc is large enough to reach the 3 : 1 resonance (Smak, 1996, 2000; Lasota, 2001).

Further indirect observational evidence for eccentric discs comes from the discovery of a broad eccentricity distribution amongst observed exoplanets (Butler et al., 2006) (see Figure 3.2, see also Alei et al. (2020), Figure 3.3, for a more recent catalogue). This contrasts with the orbits of the solar system planets which are all close to circular. There are many competing explanations for the origin of the exoplanetary

eccentricities, either involving the interaction of the planet with the protoplanetary disc (Kley & Nelson, 2012; Teyssandier & Ogilvie, 2017; Teyssandier & Lai, 2019; Ragusa et al., 2017) with the excitation of a coherent disc-planet eccentric mode, or through gravitational interaction with a companion which can excite eccentricity by the Lidov-Kozai mechanism (Naoz, 2016), planet-planet scattering (Rasio & Ford, 1996; Papaloizou & Terquem, 2001; Ford & Rasio, 2008; Jurić & Tremaine, 2008; Mustill et al., 2017) or secular chaos (Wu & Lithwick, 2011; Lithwick & Wu, 2011; Hamers & Lai, 2017; Deibert et al., 2018; Petrovich et al., 2019; Teyssandier et al., 2019). The recent discovery of an eccentric Hot Jupiter (CI Tau b), with $e = 0.28 \pm 0.16$, around the disc bearing star CI Tau (Johns-Krull et al., 2016) has provided a strong motivation for studying eccentricity excitation in disc planet interactions (e.g. Rosotti et al. (2017); Ragusa et al. (2017)).

Regardless of whether the disc is involved in the excitation of the eccentricity or not, if the excitation mechanism operates during the disc lifetime (probable given the aforementioned eccentric Hot Jupiter, CI Tau b), the planet can impose a forced eccentricity on the disc. Additionally, if the planetary eccentricity is due to the Lidov-Kozai mechanism this has the potential to excite eccentricity in the disc (Martin et al., 2014; Fu et al., 2015a,b, 2017; Lubow & Ogilvie, 2017). As debris discs, the remnants of the protoplanetary disc, have observed eccentricities (e.g. Fomalhaut, HD 202628 Kennedy (2020), see Figure 3.12), which in some instances are expected to originate in the forced eccentricity from a planetary companion (Dermott et al., 1998; Wyatt et al., 1999), there is good evidence that protoplanetary discs can be eccentric over at least part of their lifetime.

Van Eylen et al. (2019) found evidence that systems with singly transiting planets have generally higher eccentricities than systems with multiply transiting planets. This makes sense if the planetary eccentricities originated from the Lidov-Kozai mechanism as planet chains are more resistant to the Lidov-Kozai effect than single planets (Pu & Lai, 2018). According to Van Eylen et al. (2019) this result is also consistent with planet-planet scattering as the excitation of eccentricities by gravitational self stirring leads to wider planet spacings and mutual inclinations making it harder to detect multiple planets (Lai & Pu, 2017). This result would be somewhat surprising if the planetary eccentricity originated in secular chaos, as secular chaos requires at least 3 planets to operate (Wu & Lithwick, 2011; Lithwick & Wu, 2011; Petrovich et al., 2019) (although hierarchical 2 planet systems can exhibit other forms of chaos (Wu & Lithwick, 2011)). At present little work has been done on multiple planets in eccentric discs (although the case of a disc with two planets was considered in Teyssandier &

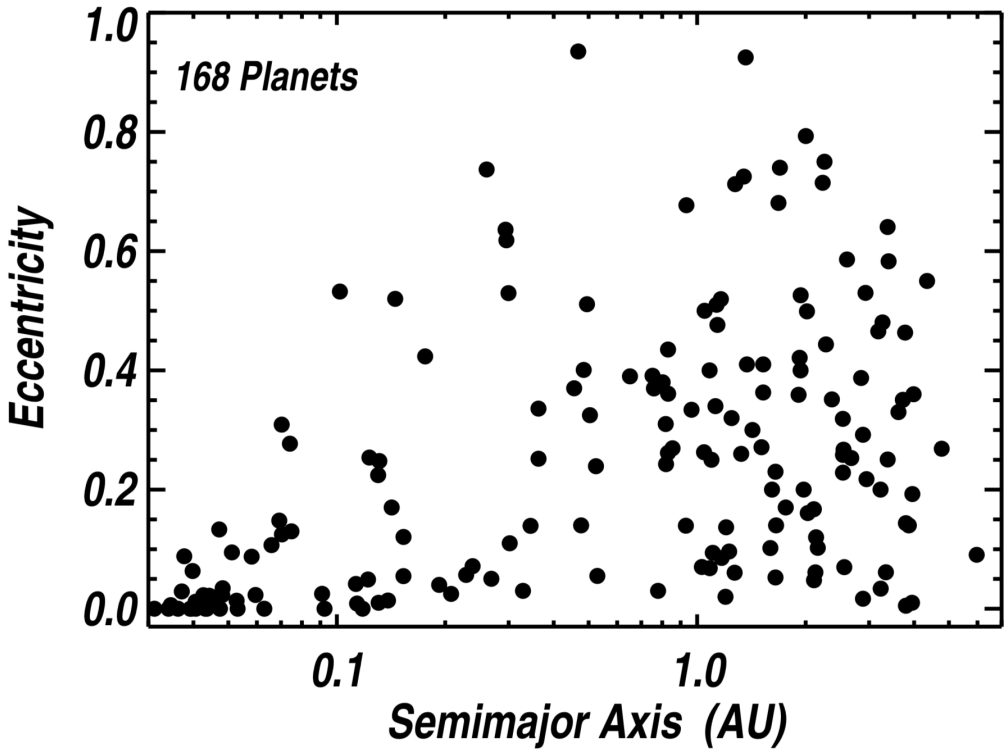


Fig. 3.2 Eccentricity against semimajor axis for nearby exoplanets from Butler et al. (2006). Planets with small semimajor axis have low eccentricities due to tidal circularisation.

Lai (2019)), so it is unclear if these observations are consistent with the excitation of planetary eccentricities by disc planet interactions.

In addition to planets the eccentricity distribution in stellar binaries is indirect evidence for the existence of eccentric discs. Notably many stellar populations feature an eccentricity distribution where short period binaries are nearly circular (due to tidal circularisation), while long period binaries have large eccentricities. This eccentricity distribution is seen in pre-main sequence/young main sequence (pre-MS/MS) stars (Mathieu et al., 1989; Mathieu, 1994), barium stars (Jorissen et al., 1998; Izzard et al., 2010) and post-AGB (asymptotic giant branch) stars (Van Winckel et al., 1997; Frankowski & Jorissen, 2007; Izzard et al., 2010). This occurs even in systems, such as the post-AGB stars (Van Winckel et al., 1997; Frankowski & Jorissen, 2007; Izzard et al., 2010), where the system is expected to have been circularised during a period of common envelope evolution. In pre-MS/MS stars and post-AGB stars a leading explanation for this distribution is eccentricity pumping via interaction

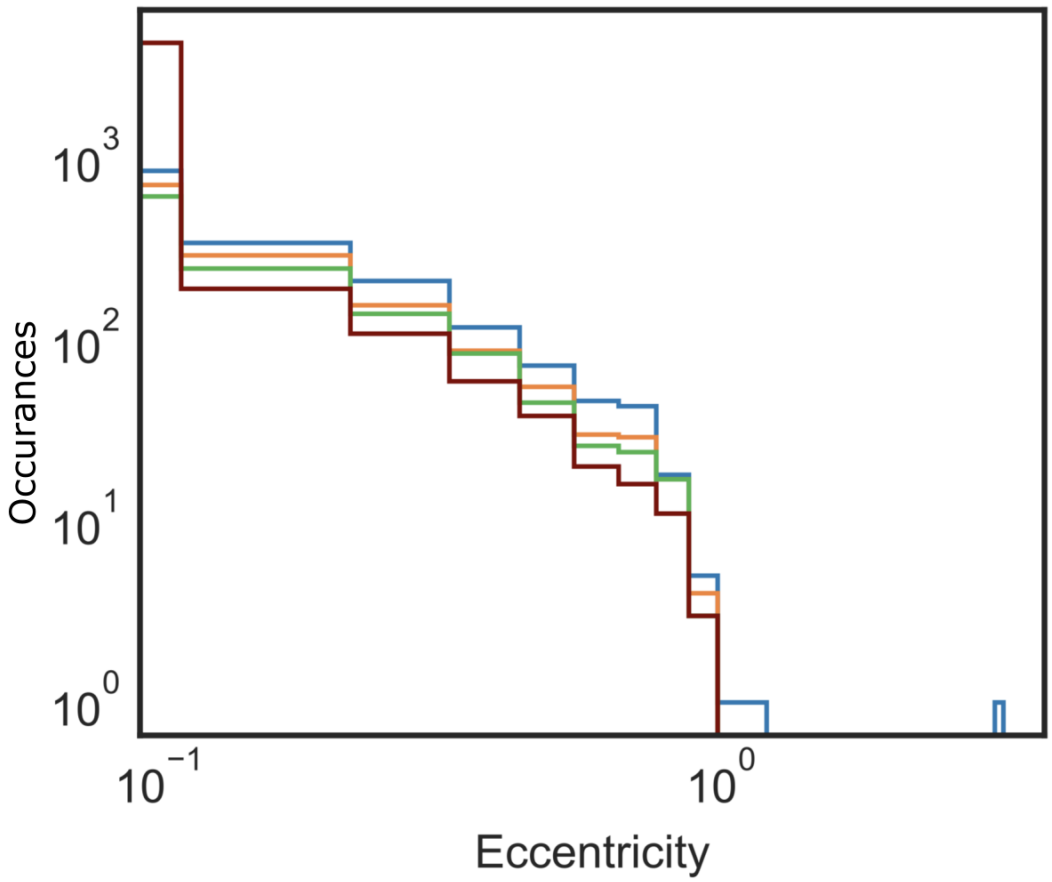


Fig. 3.3 Occurrences for exoplanets of a given eccentricity from 4 different exoplanet catalogues collated by Alei et al. (2020). While this may suggest that the majority of planets have low eccentricities, the surveys are biased towards short period planets where tidal circularisation is effective. The extremely high $e \geq 1$ eccentricities for the blue line (Exoplanet Encyclopaedia Schneider et al. (2011)) are possibly "rogue planet" candidates which are included in that database.

with a circumbinary disc. For pre-MS/MS this would involve interaction with the natal protoplanetary disc prior to its dispersal (Artymowicz et al., 1991; Artymowicz & Lubow, 1994), while for Post-AGB stars this will be with the circumbinary disc formed from an ejected stellar envelope after a phase of common-envelope evolution (Frankowski & Jorissen, 2007; Dermine et al., 2013; Izzard & Jermyn, 2018).

Other indirect evidence for eccentric discs comes from high frequency quasiperiodic oscillations (HFQPOs). These are peaks in the power spectrum of the X-ray binaries which are centred on frequencies which correspond to disc frequencies close to the marginally stable orbit. A leading explanation for HFQPOs is that they are inertial modes trapped close to the marginally stable orbit (Kato & Fukue, 1980; Kato, 2004, 2008). These trapped inertial waves are excited parametrically by $m = 1$ patterns in the disc as are formed by warps, eccentric waves and the breathing mode (Kato, 2008; Ferreira & Ogilvie, 2008). The details of the HFQPO mechanism is the subject of ongoing research (Fu & Lai, 2009; Fu & Lai, 2013; Dewberry et al., 2018, 2019, 2020a,b) so, at present, the correspondence between the appearance of HFQPOs and eccentricity in the disc is tentative. In the future HFQPOs might provide a probe of the geometry of the inner accretion disc. The occasional suggestion in the literature that HFQPOs might be useful in searching for physics beyond GR is somewhat naive; while the GR of an isolated black hole is relatively simple and well understood the fluid dynamic of the disc is substantially more complicated and any inconsistency between the expected and observed HFQPOs are likely a result of the (currently) poorly understood HFQPO mechanism and not an indication of an issue with GR.

Eccentric discs are also implicated in the relativistic precession QPO model (Stella & Vietri, 1998, 1999). In this model HFQPOs are associated with the three main disc frequencies, the orbital frequency Ω, and the nodal and apsidal precession frequencies. In this model the higher frequency HFQPO is associated with the orbital frequency, while the lower frequency HFQPO are associated with the apsidal precession frequency (Stella & Vietri, 1999; Zhang & Dolgov, 2002; Ingram & Motta, 2014), and is associated with a rapidly precessing inner eccentric disc. One advantage of this model is this naturally explains the correlation between the higher and lower frequency HFQPOs, as these are just the angular and apsidal precession frequencies of a given orbit. A major issue with this model is it suggests the inner eccentric disc has become decoupled from the rest of the disc, as a coherently precessing eccentric mode (even a strongly confined one) should precess slower than the apsidal precession frequency at the inner edge of the disc (likely at a suitably weighted average of the apsidal precession frequency over the inner disc). Zhang & Dolgov (2002) proposed associating the 15-60 Hz QPO to

the slower apsidal precession in the outer disc; however, given that the precession of the outer disc implies some form of global eccentric mode, mediated by pressure, it is highly unlikely the precession of the disc is directly related to the epicyclic frequency and would typically be expected to precess much slower. While certainly promising, it is clear that significant theoretical uncertainties remain with QPO mechanisms that invoke eccentric discs.

3.1.2 Directly measured disc eccentricity

The main observational evidence for eccentric discs has come from the modification to the Doppler shift of spectral features by the Keplerian motion. A characteristic feature of spectral lines in circular discs is that they are split by the rotational motion of the disc into approaching (blueshifted) and receding (redshifted) peaks (See Figure 3.4). For a circular disc these double peaked lines are symmetric. For an eccentric disc the approaching and receding sides of the disc will have different velocities that depend on the orientation of the disc. This breaks the symmetry of the line profile causing an asymmetric peak with the degree of asymmetry dependent on the eccentricity and disc orientation (e.g. see Figure 3.5, along with a real life example from Hanuschik (1988), Figure 3.6). Slow global precession of the disc as expected by eccentric disc theory manifests in a slow, periodic variation of the line asymmetry as the apocentre/pericentre moves from the receding to the approaching side. Decretion discs around Be stars often exhibit asymmetric Balmer lines which vary on $\sim$ 7 yr timescales (Copeland & Heard, 1963) (long compared with the 1 day orbital period of the disc (Okazaki, 1991)), which is consistent with a trapped prograde eccentric mode (Okazaki, 1991, 1997; Ogilvie, 2008). Recently Zamanov et al. (2020) observed asymmetric Hα, Hβ and HeI 6678 emission lines in the Be star disc in X Persei over the period December 2017 - January 2020. As the lines occur at different disc temperatures they are thought to probe different radial locations in the disc. Zamanov et al. (2020) have claimed their observations show evidence that the disc becomes eccentric initially in the inner disc before the eccentric wave travels outwards and is observed in the lines which probe the outer disc. This method could provide a powerful method of observing the evolution of the eccentricity profile in eccentric discs.

Another system where this technique has been used is in the gaseous discs around polluted white dwarfs; here three Ca II emission lines are observed to have a slowly varying double peaked structure which is attributed to the presence of an eccentric gas disc (Manser et al. (2016b,a); Cauley et al. (2018), see 3.8). A polluted white dwarf is a white dwarf with trace metals detected in its atmosphere (Koester et al., 1997;

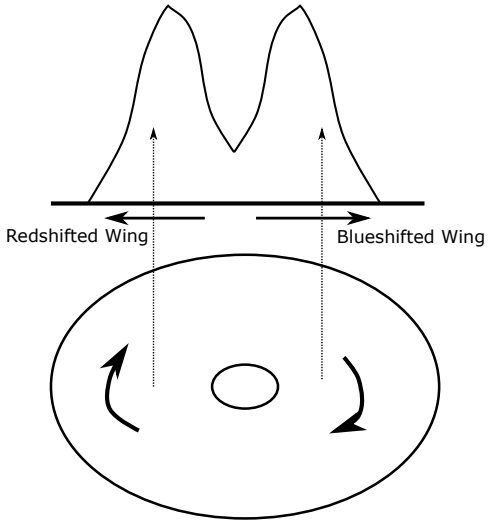


Fig. 3.4 Diagram showing how the redshifted and blueshifted wing of an emission line originate in the receding and approaching sides of a circular disc.

Zuckerman et al., 2007; Jura & Young, 2014), these trace metals have a relatively short sinking timescale and should sink out of the atmosphere if material isn't periodically resupplied to the white dwarf surface (Koester et al., 1997; Zuckerman et al., 2007; Jura & Young, 2014). These polluted white dwarfs are often accompanied by a debris disc (Manser et al., 2016a). The leading theory regarding the origin of both the debris disc and the pollution in the white dwarf atmosphere is that it originates in a remnant planetary system (Zuckerman et al., 2007; Jura & Young, 2014) and is thrown inwards by perturbations from an unseen gas/ice giant in a similar manner to how Neptune creates comets by close encounters with Kuiper belt objects (Debes & Sigurdsson, 2002; Jura, 2003). The origin of the gas disc is somewhat uncertain with various proposals for how gas can form from the dusty debris disc (Jura, 2008; Rafikov, 2011b,a; Metzger et al., 2012). Regardless of how the gas forms, the requirement that there exist some orbital companion, and the fact that any material originating in the remnant planetary system will arrive on highly eccentric orbits (Debes & Sigurdsson, 2002; Jura, 2003; Malamud & Perets, 2019b,a; Nixon et al., 2020), means it should not be a surprise that the gaseous discs have some level of eccentricity.

This technique has begun to be used to constrain the highly uncertain geometry of discs formed during Tidal Disruption Events (TDEs). Observational evidence for highly eccentric disc in TDEs comes from fits to the Hα emission made by Liu et al. (2017), Cao et al. (2018) and Holoien et al. (2019). In these papers the authors proposed that the Hα emission originates from the reflection of the X-ray corona off a highly eccentric

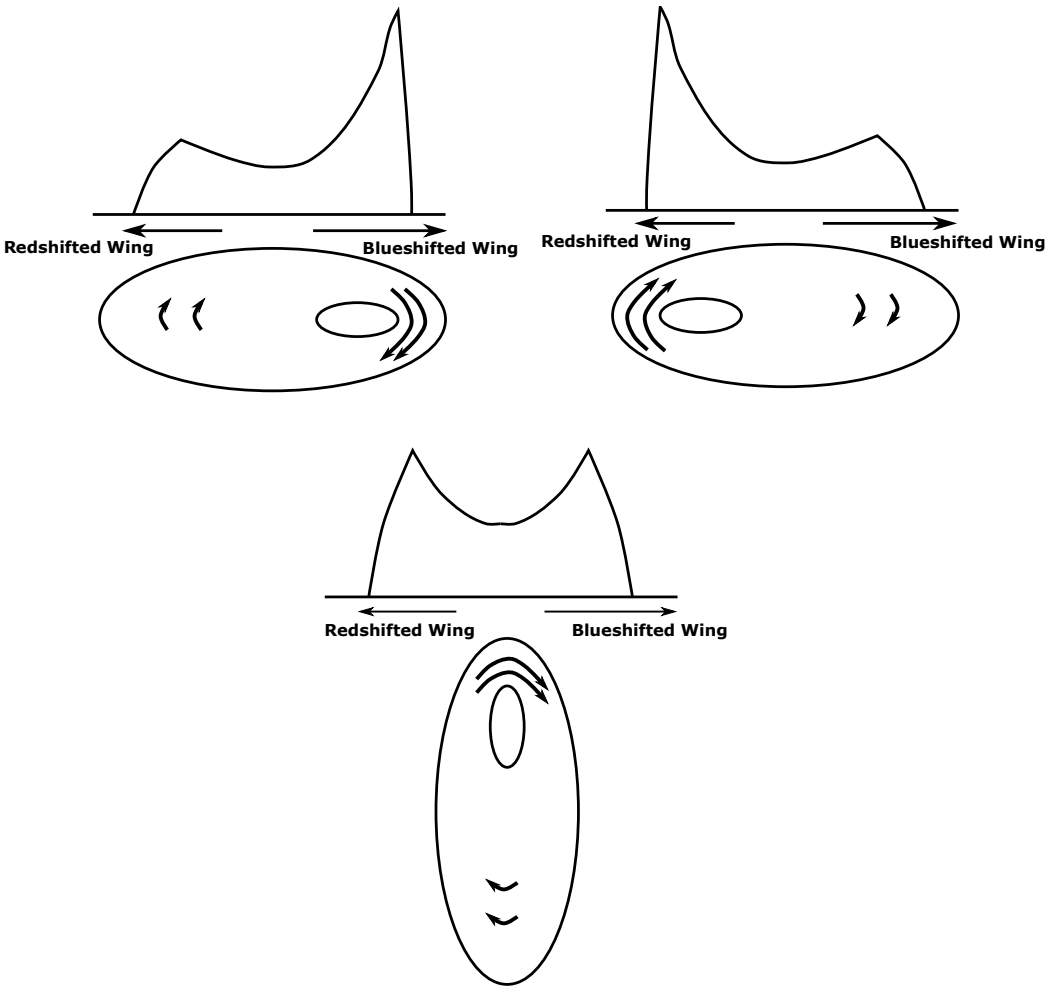


Fig. 3.5 Diagram showing how the asymmetry an eccentric disc modifies the emission line with the faster material on the pericentric side of the disc producing a stronger signal, while the apocentric peak is smeared out. The orientation of the eccentric disc affects the shape of the line, so the strong peak associated with pericentre alternates between the redshifted and blue shifted wings (Manser et al., 2016a).

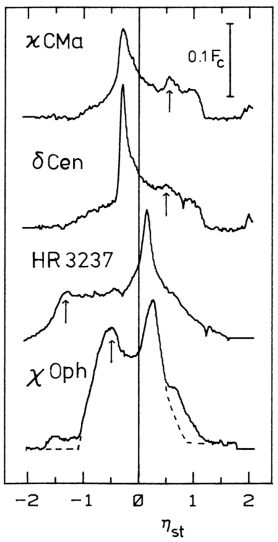


Fig. 3.6 Asymmetric Fe II emission line profiles for 4 Be star discs, taken from Hanuschik (1988). Data first presented in Hanuschik (1987).

($e \approx 0.97$) disc, in order to fit the double peaked structure of the Hα line. Holoien et al. (2019) also showed that the Hα emission of PS18kh indicated non-axisymmetric structures in the TDE disc, consistent with either an eccentric disc or spiral arm. Recently Hung et al. (2020) have claimed a much lower eccentricity of $e \approx 0.1$ for the TDE AT 2018hyz.

For discs around black holes the iron Kα line is an important probe of the conditions of the inner disc, close to the marginally stable orbit (Fabian et al., 2000). In principle the iron Kα line is a double peaked line like the Balmer and Ca II lines. In practice the line is smeared out by the effects of special relativistic beaming (as the disc velocity is close to the speed of light) and gravitational redshift (as the line originates deep within the black hole's potential well), along with gravitational lensing and frame dragging effects (Fabian et al., 2000). These effects are summarised in Figure 3.7, taken from Fabian et al. (2000). This line should also be sensitive to any eccentricity in the disc, which, under certain conditions, can propagate to the inner disc (Ferreira & Ogilvie, 2009). When combined with relativistic beaming, this could strongly modify the shape of the line. This possibility, and the effect of nonaxisymmetric structures in general, has received very little attention (Karas et al. (2001); Hartnoll & Blackman (2002); Fukumura & Tsuruta (2004) are notable exceptions), most likely due to the high degree of degeneracy that exists between the effects of the disc geometry and the effects of the space-time geometry.

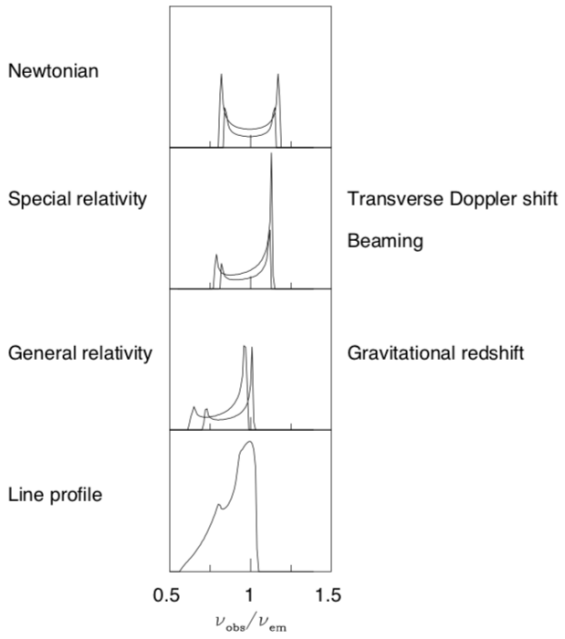


Fig. 3.7 Effects of relativistic beaming and gravitational redshift on the iron Kα line from Fabian et al. (2000). Note the similarity of the asymmetric relativistically corrected lines in a circular disc, to the non-relativistic Fe II emission lines in the eccentric Be star discs (Figure 3.6).

With sufficiently good spectrometry, Doppler tomograms of the emission lines can be constructed, allowing for a more detailed determination of the velocities in the disc. This technique has the advantage that it gives an indication of where in the disc the eccentricity is as the inner and outer disc are separated in velocity space, due to their differing speeds. This technique also captures information about the orbital phase dependence of the velocity. Doppler tomography has been used for both the gas discs around polluted white dwarfs (Manser et al. (2016a), see Figure 3.9) and various cataclysmic variables (Steeghs et al., 1996; Richards, 2007).

In addition to observing the Doppler shifts of emission lines, another common technique for determining eccentricity in discs is through eclipse timing. This was important in one of the first applications of eccentric disc theory, which was as an explanation of the superhump phenomenon in cataclysmic variables. A superhump is a periodic variation of the disc lightcurve with a frequency close to the orbital period of the binary. The modulation of the lightcurve comes from a precessing eccentric disc excited at the 3:1 Lindblad resonance. Evidence for this model comes from the variations in the eclipse of the hot spot by the donor (Hessman et al., 1992; Rolfe et al., 2000). The "hot spot" in a superhump system is the point where the accretion

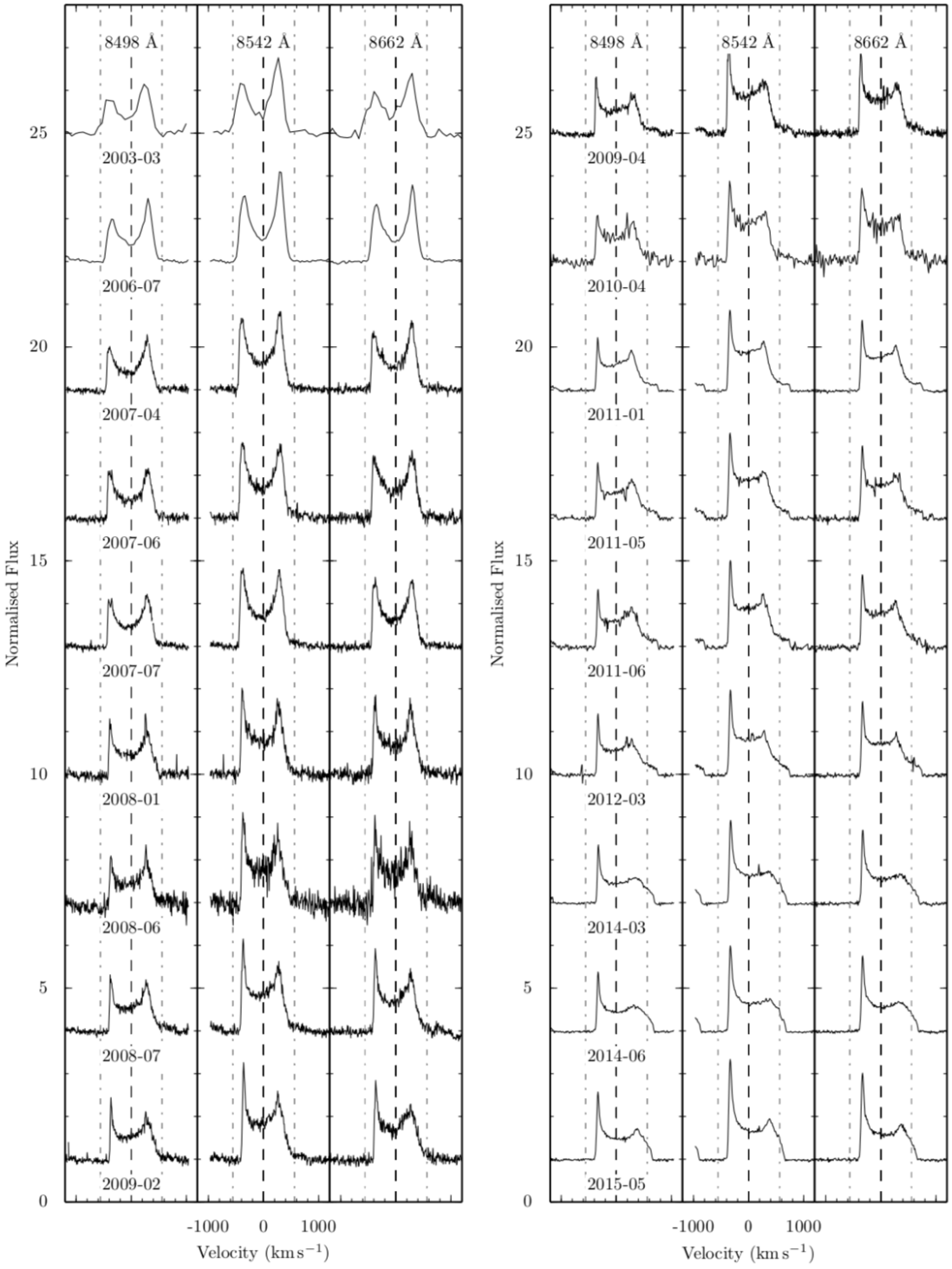


Fig. 3.8 Secular variation of asymmetric line profiles from the disc around SDSS J122859.93+104032.9 from Manser et al. (2016a).

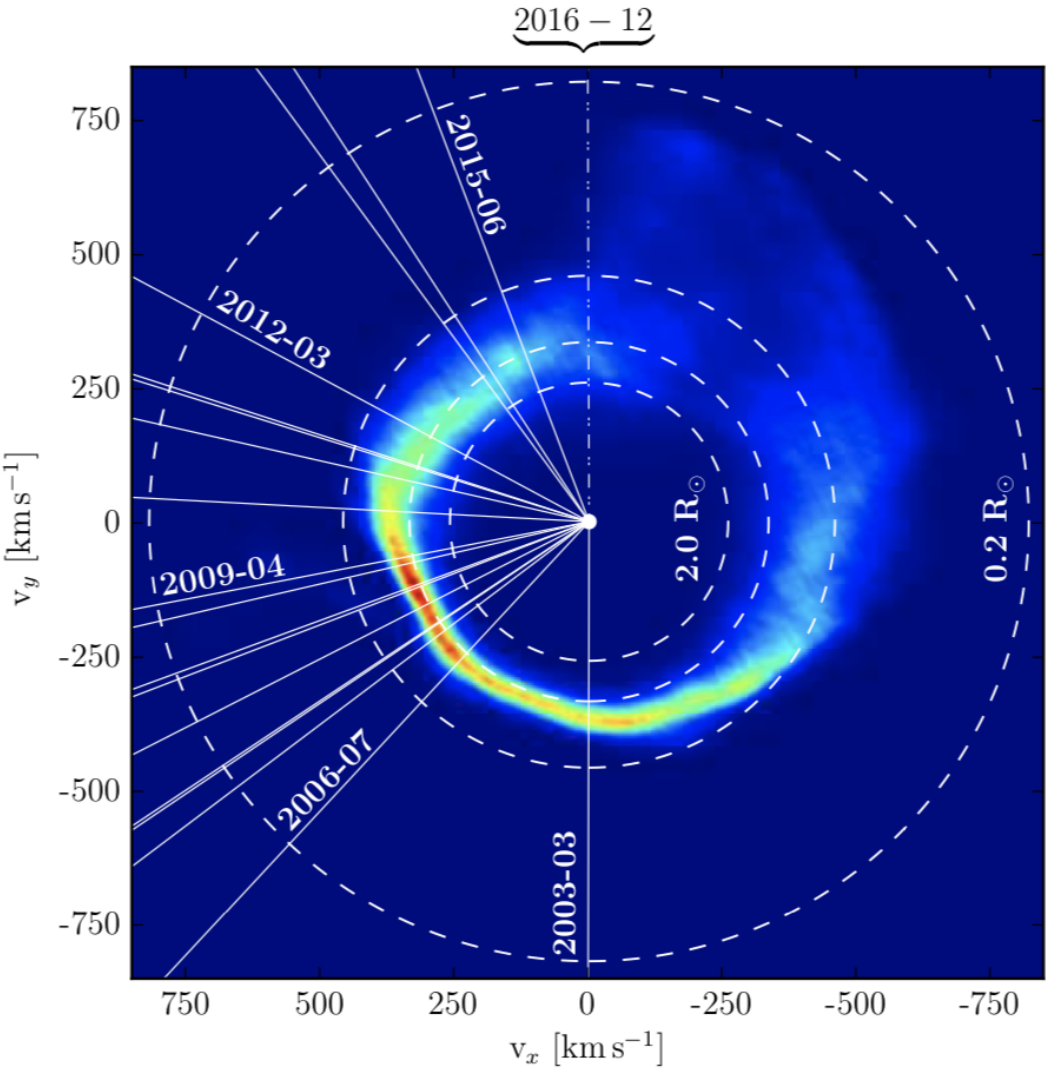


Fig. 3.9 Doppler tomography of an eccentric gaseous disc around the white dwarf SDSS J122859.93+104032.9 from Manser et al. (2016a).

stream from the donor star impacts the disc resulting in shock heating. This spot is periodically eclipsed by the donor star. The orbital phase where the hotspot is first eclipsed varies between orbits, with a period consistent with the precession period of the disc (as determined by the orbital and superhump periods). This has been interpreted to be due to the varying of the position of the hotspot due to the change in the radial location of the outer disc edge as an eccentric disc precessed (Hessman et al., 1992) (see Figure 3.10). As the hotspot location changes the orbital phase where its first blocked by the donor star changes also producing the observed variability. Additionally superhump binaries have long been known to possess the asymmetric double peaked Hα line profiles which have become a characteristic feature of eccentric discs (Hessman et al., 1992).

Spatially resolved observations of eccentric discs are uncommon and are predominantly infrared observations of dust emission. Non-axisymmetric features in dust in protoplanetary discs have been observed with ALMA. How the features in the dust relate to features in the gas is uncertain and is the subject of current research. The prevailing explanation for large scale asymmetries/horseshoe shaped features (such as those seen in ALMA observations of SAO 206462 and SR 21; Figure 3.11 taken from Pérez et al. (2014)) in transition discs is they originate from dust trapped in large vortices excited by the Rossby wave instability at the edge of a gap cleared by a planet embedded in the disc. One issue with this explanation is the stability of the vortex (Chang & Oishi, 2010; Railton & Papaloizou, 2014), particularly in 3D where they are unstable to the elliptical instability (Lithwick, 2009; Lesur & Papaloizou, 2009) (a parametric instability which occurs when a multiple of the vortex turnover frequency matches an inertial wave frequency). An alternative explanation, which has received less attention, is the possibility that these features are due to an eccentric dust disc (Ataiee et al., 2013; Ragusa et al., 2017, 2020). The lack of attention is perhaps surprising given that similar, if less pronounced, features are apparent in debris discs which have measured eccentricities from the offset between the central star and the disc maxima. Observationally it should be possible to distinguish an eccentric deformation from a vortex, given a long enough time between observations. This is because a vortex will circulate at roughly the orbital frequency and can move by a detectable amount on $10 - 15$ yr timescales, whereas the precession frequency of an eccentric mode is so slow that it will remain fixed between observations.

Spatially resolved debris discs can have asymmetric surface brightnesses, as found in protoplanetary discs, along with measurable offsets from the central star owing to the disc's eccentricity (see Figure 3.12 for the particularly clear example provided by

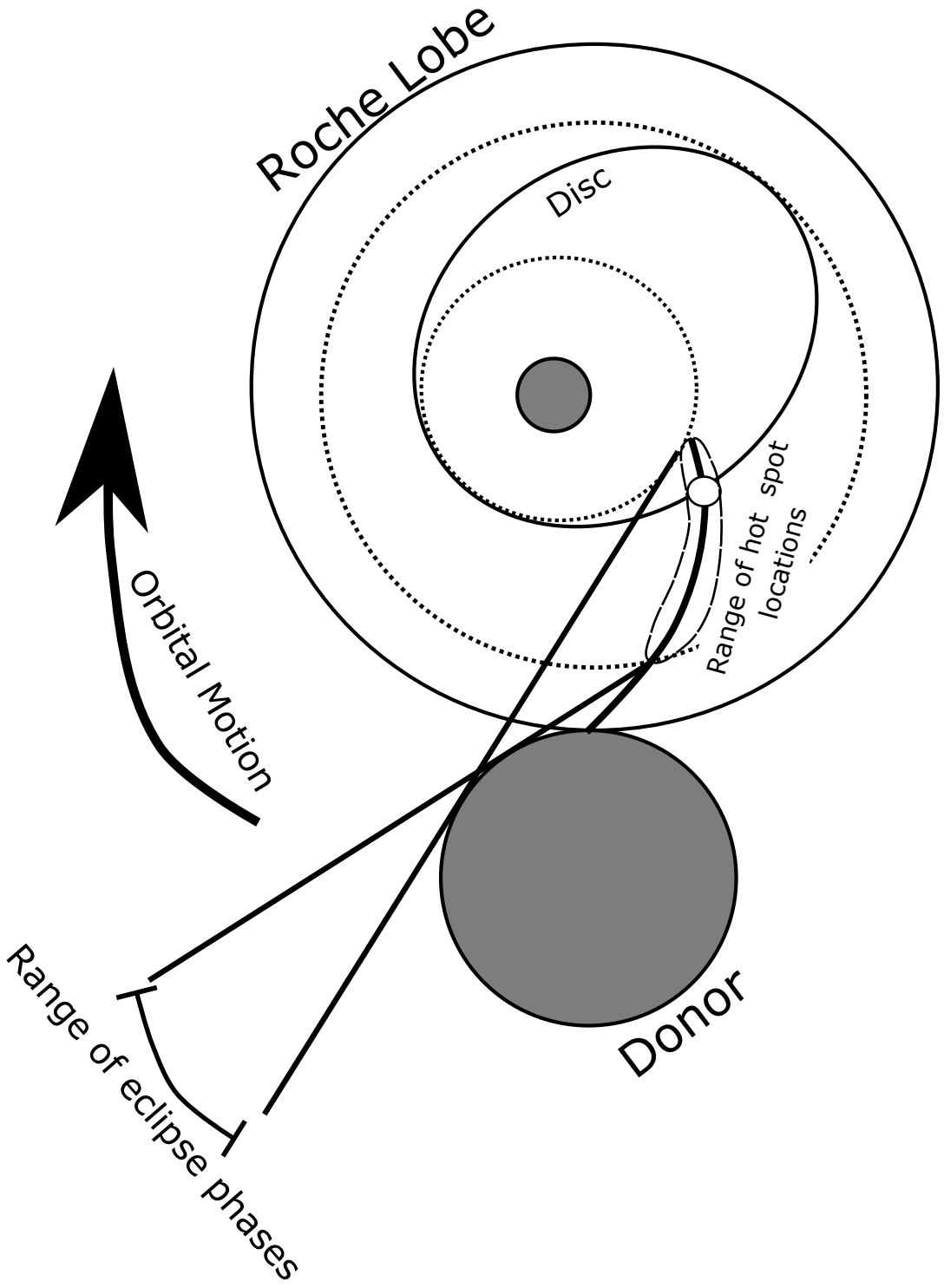


Fig. 3.10 Depiction of the eclipse timing variation due to an eccentric disc, based off the figures and explanation given in Hessman et al. (1992). The location where the accretion stream hits the disc can change depending on the orientation of the eccentric disc. As the disc precesses the hot spot (located at the disc stream intersection) moves along the ballistic arc traced by the stream. The orbital phase where the hot spot is first eclipsed by the donor depends on the location of the hotspot.

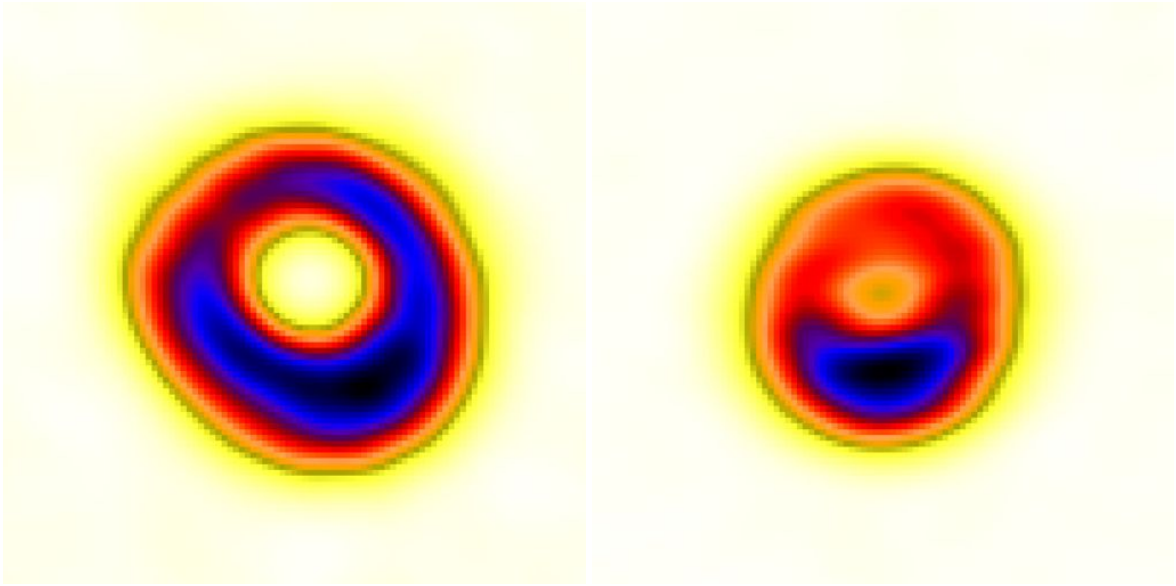


Fig. 3.11 Dust emission observations with ALMA of SAO 206462 and SR 21 showing significant asymmetries from Pérez et al. (2014)

Fomalhaut). At short wavelengths the observed asymmetries in eccentric debris disc are well understood as being caused by the hotter dust at pericentre glowing brighter than those at apocentre (Wyatt et al., 1999). At longer wavelengths the situation is complicated by any variation of the dust surface density around the orbit and is poorly understood. As mentioned in 2.1 the false "apocentre hang time" argument has had unfortunate influence on the study of eccentric debris discs at long wavelengths, where there can be the expectation that debris discs with constant eccentricity should be denser (and thus brighter at long wavelengths) at apocentre.

The "apocentre hang time" argument may be the reason the eccentric disc hypothesis for the asymmetries in protoplanetary discs has not received much attention. For instance Ataiee et al. (2013) explained the overdensity at apocentre seen in their eccentric disc simulations as being due to the lower velocities at apocentre as per the "apocentre hang time" argument, rather than the correct explanation that the orbits are bunched up at apocentre due to the eccentricity falling off with distance from the planetary perturber. In principle a steep enough eccentricity gradient could lead to an arbitrarily strong asymmetric feature without the disc becoming noticeably eccentric. Teasing out the influence of dust surface density and dust temperature on the dust emission is non-trivial, with both influenced by eccentricity in the disc.

In our solar system system several planetary rings have eccentricities which can be directly imaged. Study of these rings is aided greatly by our ability to send space probes

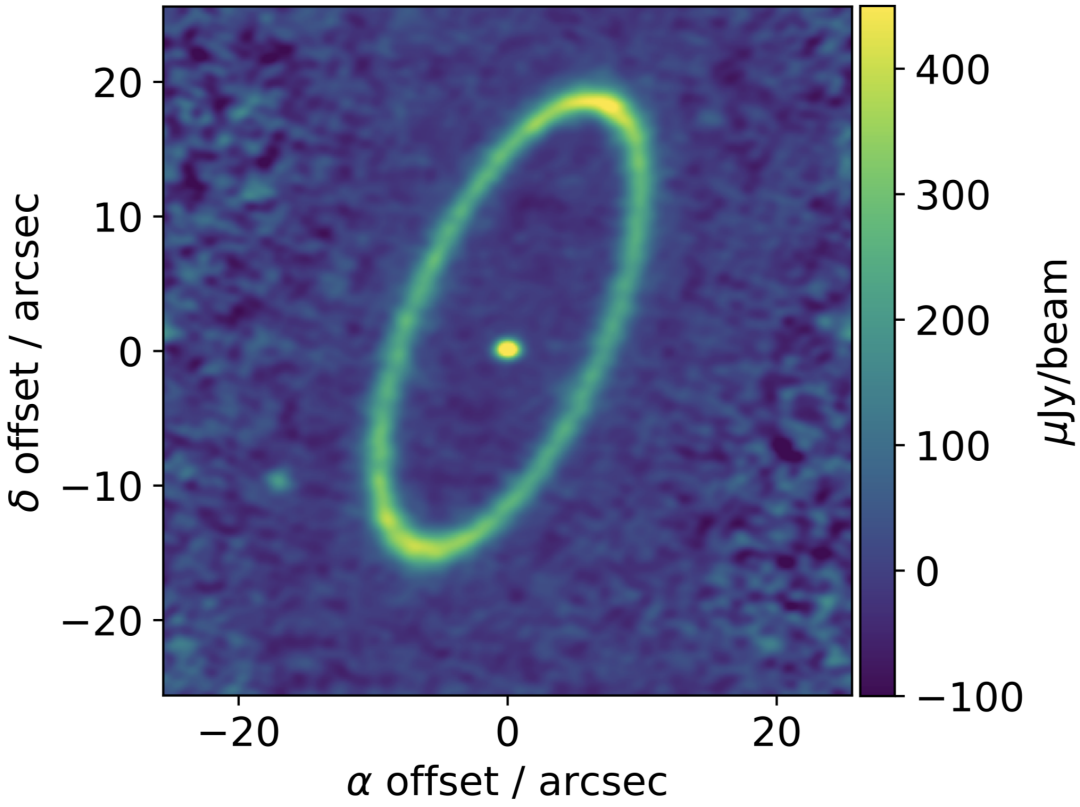


Fig. 3.12 ALMA image of Fomalhaut from Kennedy (2020) showing the star with significant offset from the centre of the ring, due to the debris disc's eccentricity.

to the outer planets allowing planetary rings to be imaged in great detail. This led to the discovery of the eccentric Titan and Huygens ringlets of Saturn (Esposito et al., 1983; Porco, 1983; Porco et al., 1984a) by the Voyager missions. Additionally eccentricity has been measured in Saturn's Maxwell ringlet, despite having no neighbouring moon responsible for the observed eccentricity (Porco et al., 1984b). Several rings of Uranus are also eccentric, in particular the ϵ ring with an eccentricity $\approx 8 \times 10^{-3}$ (Nicholson et al., 1982). Precision measurements of the eccentricities of planetary rings can be obtained by the stellar occultation method. In this method a distant star is occluded by the ring, by combining the time the star spends occluded by the ring with the known geocentric velocity of Uranus an accurate measurement of the (sky projected) width and position of the ring can be obtained. This gives both an accurate determination of the eccentricity and allows an eccentricity gradient to be inferred based on the eccentricities of the inner and outer edges of the ring. In the case of Uranus' ϵ ring the inferred eccentricity gradient across the ring is $a\frac{de}{da} \approx 0.6$ (Nicholson et al., 1982). The Adam's ring of Neptune also has measured eccentricity, and is in a $42:43$ outer Lindblad resonance with the moon Galatea (Miner et al., 2007). This ring exhibits additional non-axisymmetric structure (Neptune so called "ring arcs") the origin and azimuthal confinement of which is debated, but are typically attributed to interaction with Galatea.

In addition to the eccentric rings found around the gas giant, several minor bodies in the solar system also posses eccentric ringlets. The Centaur's Chariklo (Sicardy et al., 2014; Braga-Ribas et al., 2014) and Chiron (Ortiz et al., 2015) both posses eccentric rings. The trans-Neptunian dwarf planet Haumea also possess rings (Ortiz et al., 2017) with potentially complex azimuthal structure.

Another dust feature, that is a possible indicator of an eccentric disc, are the butterfly like features seen in some nearly face on dust discs (Hines et al. (2007); Maness et al. (2009); Ricarte et al. (2013), see Figure 3.13). These features are interpreted as being due to aligned highly eccentric dust orbits (Lee & Chiang, 2016). Theoretically differential precession should disrupt this configuration and a gaseous eccentric disc has been suggest as a mechanism by which this structure can be preserved (Lee & Chiang, 2016; Lin & Chiang, 2019).

3.2 Observational evidence for tidally distorted discs

Like the eccentric disc, much of the observational evidence for tidally distorted discs comes from Doppler tomography. Doppler tomography of discs in binary systems has

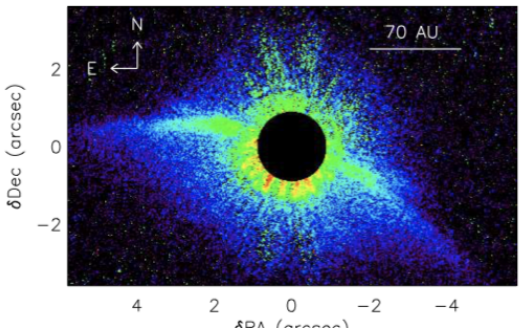


Fig. 3.13 HST image of dust in HD 61005 showing swept butterfly like morphology thought to be caused by highly eccentric dust grains. Observation by Maness et al. (2009).

provided evidence for an $m = 2$ distortion to the disc. For instance in several dwarf nova binaries a two armed structure is seen in their Doppler tomograms (see Figure 3.14, along with Steeghs et al. (1996, 1997); Harlaftis et al. (1999); Morales-Rueda et al. (2000); Joergens et al. (2000); Groot (2001); Steeghs (2001)). This two armed structure is persistent during the outburst (Morales-Rueda et al., 2000; Groot, 2001), but its morphology evolves with time (Groot, 2001). The two arms can have differing intensities and shapes (Groot, 2001; Smak, 2001), indicating an asymmetry in the underlying $m = 2$ structure (potentially a superposed odd m structure, that could be a result of disc eccentricity).

Circumbinary discs represent another class of tidally distorted discs, where the companion is located interior to the inner disc edge. Hence the greatest tidal distortion is to the inner disc. The existence of such discs has been observationally confirmed by both the direct detection of circumbinary discs (Ireland & Kraus, 2008; Piétu et al., 2011; Yang et al., 2017; Keppler et al., 2020) and of circumbinary planets (Doyle et al., 2011; Welsh et al., 2012; Orosz et al., 2012a,b; Schwamb et al., 2013; Kostov et al., 2014; Welsh et al., 2015; Kostov et al., 2016). Recently a lot of observational work has gone into attempting to image the common inner cavity of the circumbinary disc, predicted to exist if the mass ratio is close to 1 (as opposed to a cleared gap for a lower mass companion). This has been carried out with an interest in how material can overcome the tidal torque responsible for truncating the disc to accrete onto the central binary. Notably near-IR scattered light imaging of GG Tau A, carried out by Yang et al. (2017), revealed a wealth of substructure which has analogues in simulations of circumbinary discs that include the central region (See Figure 3.15). The observations of Yang et al. (2017) revealed evidence for an accretion stream from the circumbinary disc onto one of the binary components (a feature also seen in the continuum observations of Piétu et al.

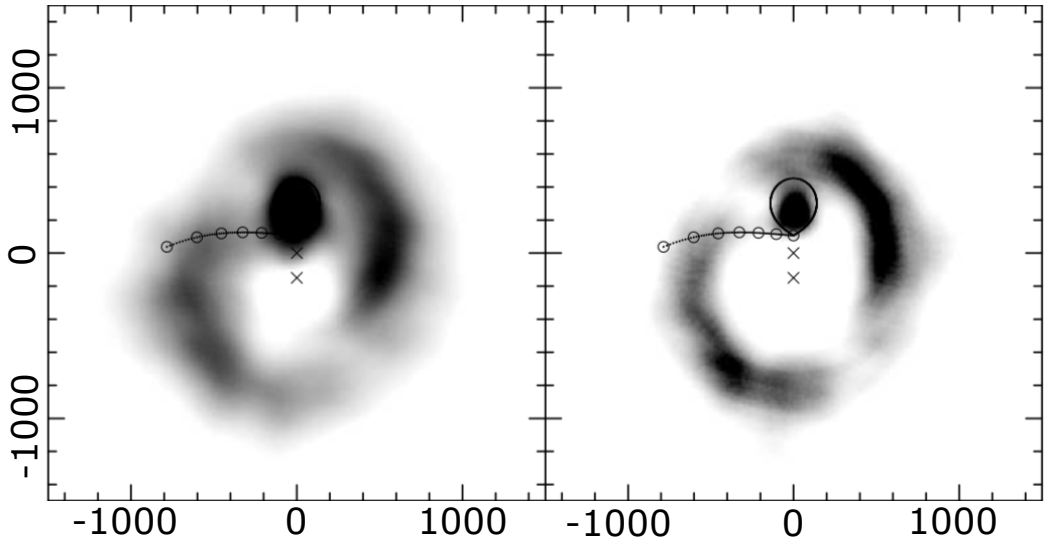


Fig. 3.14 Doppler tomography of the dwarf nova binary IP Pegasi from Steeghs et al. (1997) showing a two armed structure indicating some form of $m = 2$ distortion in the underlying disc. Left is Hα, right is HeI(6678), axes are velocity in km s^{-1}.

(2011)), along with tentatively identified mini-discs around the stars and a possible connecting stream between them. Additionally Yang et al. (2017) found that the inner edge of the disc could be fitted with an ellipse with eccentricity $e = 0.2$. Scattered light observations of GG Tau A by Keppler et al. (2020) also find similar features to Yang et al. (2017).

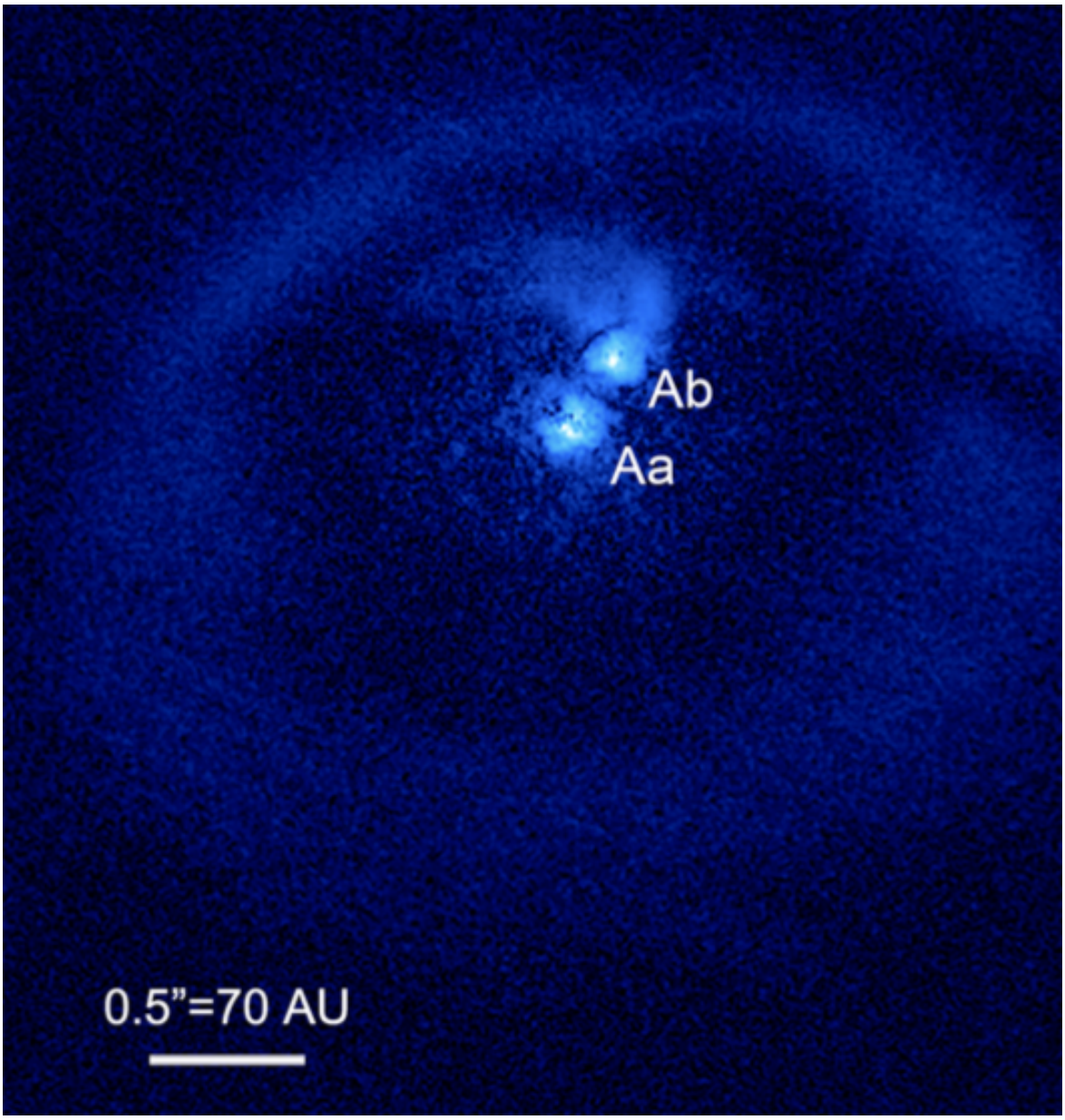


Fig. 3.15 Near-IR scattered light imaging of GG Tau A of Yang et al. (2017) showing possible substructure in the disc cavity. Note that the non circular shape and variable width of the ring, which cannot be produced by a flat inclined circular disc, provides evidence for an eccentric cavity.

Chapter 4

Underlying Mathematics

4.1 Introduction

This thesis will make use of several mathematical formalisms not commonly used in astrophysics. As such, in this chapter I will introduce and explain the less well known mathematics that I will make use of throughout this thesis. The main part of this chapter will be a discussion of Lagrangian and Hamiltonian dynamics of continuous (or infinite dimensional) systems, and how to develop multiple time/length-scale analysis of the dynamics. We shall also cover advection of tensors/tensor densities by a fluid flow, important in constitutive models of the MRI, and introduce some of the common notational conventions used in this thesis. Readers who are familiar with the mathematical formalisms reviewed here, or who are primarily interested in the astrophysical consequences of this work, can safely skip this chapter. Derivations presented in the rest of this thesis will reference results reviewed here.

This thesis will make use of several notational conventions from differential geometry. Namely the use of "upstairs" and "downstairs" indices to denote a contravariant and covariant indices respectively; Einstein summation convention where if a covariant and contravariant index are labelled with the same letter they are summed over and the notation for symmetrising indices,

$$X^{(ij)} := \frac{1}{2}\left(X^{ij} + X^{ji}\right) \quad . \tag{4.1}$$

Sometimes we shall make use of a subscript to denote a partial derivative so that $X_a := \frac{\partial X}{\partial a}$, and very occasionally as a label. This of course clashes with the notation for a covariant index, which it is should be clear in context. Where this needs clarifying

we shall use a comma to denote a partial derivative and a semi-colon for a covariant derivative. So for a covariant vector X_a we have $X_{a,b} = \frac{\partial X_a}{\partial x_b}$ and $X_{a;b} = \nabla_b X_a$.

4.2 Continuous media as dynamical systems

In this section we shall discus how continuum mechanics can be formulated as an infinite dimensional dynamical system, and (if conservative) can thus be treated within the framework of Lagrangian or Hamiltonian mechanics. Much of this discussion can be found in textbooks on variational calculus, such as Gelfand et al. (2000).

4.2.1 Functional calculus

Throughout this thesis I will make extensive use of Functional Calculus, particularly functional derivatives. These are particularly important in Lagrangian or Hamiltonian formalisms describing the dynamics of continuous media. A functional $F(u)$ is a map from a space of functions into the real numbers. Variational calculus can be defined in terms of a functional derivative defined as followed,

$$\int \frac{\delta F[u]}{\delta u}\phi(x)dx = \lim_{\epsilon\to 0}\int \frac{F[u(x)+\epsilon\phi(x)] - F[u(x)]}{\epsilon}dx \quad , \tag{4.2}$$

where $\phi(x)$ is a 'test function'. The functional derivatives has a number of useful properties, and obeys both a chain rule and a Leibniz rule. An important example of the use of a Functional derivative is in the derivation of the Euler-Lagrange equation from the Hamilton's principle. For instance, if a functional F has the form,

$$F = \int F_x(u, u_x, x)dx \tag{4.3}$$

then the functional derivative with respect to u is given by

$$\frac{\delta F}{\delta u} = \frac{\partial F_x}{\partial u} - \frac{\partial}{\partial x}\frac{\partial F_x}{\partial u_x}, \tag{4.4}$$

which takes the form of an Euler-Lagrange equation for u when set to zero. Indeed if F is the action then setting the above equal to zero corresponds to making the action stationary with respect to u.

Another important case is the Functional derivative of a function,

$$\int \frac{\delta f(u(x'))}{\delta u(x)}\phi(x)dx = \frac{\partial f}{\partial u}\phi(x') \quad , \tag{4.5}$$

which, in applied mathematics is usually written in terms of a Dirac delta,

$$\frac{\delta f(u(x'))}{\delta u(x)} = \frac{\partial f}{\partial u}\delta(x' - x) \quad . \tag{4.6}$$

4.2.2 Hamiltonian dynamics of continuous media

In Section 4.2.1 we saw how the functional derivative of a functional $F[u] = \int F_x[u, u_x, x]dx$ lead to an Euler-Lagrange equation. Here we take that further and review how this can be used to formulate Lagrangian and Hamiltonian Dynamics for continuous media. Consider a Lagrangian, a functional of u, a function of time and one spacial dimension, of the form,

$$L[u] := \int L_x[u_t, u_x, u, x, t]dx \quad , \tag{4.7}$$

with associated Action $S[u] = \int L[u]dt$. Then by Hamilton's principle we can obtain the equations of motion by taking the functional derivative of $S[u]$ with respect to u and setting equal to zero,

$$\frac{\delta S[u]}{\delta u} = \frac{\partial L_x}{\partial u} - \frac{\partial}{\partial t}\frac{\partial L_x}{\partial u_t} - \frac{\partial}{\partial x}\frac{\partial L_x}{\partial u_x} = 0 \quad . \tag{4.8}$$

This Euler-Lagrange equation is a PDE which is second order in space and time, and thus describes a form of nonlinear wave equation. We can reformulate this equation by introducing a momentum canonically conjugate to u,

$$m_x \pi_u := \frac{\partial L_x}{\partial u_t} \quad , \tag{4.9}$$

where we have introduced appropriate mass weighting with $m_x = m_x(x)$ the mass per unit x. We can define a Hamiltonian, another functional of u, via a Legendre transform of Equation 4.7,

$$H[u] := \int m_x \pi_u u_t(\pi_u, u_x, u, x, t) - L_x[u_t(\pi_u, u_x, u, x, t), u_x, u, x, t]dx \quad , \tag{4.10}$$

where we have inverted equation 4.9 to obtain u_t as a fixed functions of π_u, u_x, u, x and t. The form of Hamilton's equations which agree with the Euler-Lagrange equations are

$$m_x \frac{\partial \pi_u}{\partial t} = -\frac{\delta H[u]}{\delta u}, \tag{4.11}$$

$$m_x \frac{\partial u}{\partial t} = \frac{\delta H[u]}{\delta \pi_u}. \tag{4.12}$$

To see this carry out the respective functional derivatives, for that in Equation 4.11,

$$\begin{aligned} -\frac{\delta H[u]}{\delta u} &= -m_x \pi_u \frac{\partial u_t}{\partial u} + \frac{\partial}{\partial x}\left(m_x \pi_u \frac{\partial u_t}{\partial u_x}\right) + \frac{\partial L_x}{\partial u_t}\frac{\partial u_t}{\partial u} - \frac{\partial}{\partial x}\left(\frac{\partial L_x}{\partial u_t}\frac{\partial u_t}{\partial u_x}\right) + \left(\frac{\partial L_x}{\partial u} - \frac{\partial}{\partial x}\frac{\partial L_x}{\partial u_x}\right), \\ &= \frac{\partial L_x}{\partial u} - \frac{\partial}{\partial x}\frac{\partial L_x}{\partial u_x}. \end{aligned} \tag{4.13}$$

Substituting this expression, and the definition of the canonical momentum (Equation 4.9), into Equation 4.11 confirms that this equation agrees with the Euler Lagrange equation (Equation 4.8). For the functional derivative in Equation 4.12,

$$\begin{aligned} \frac{\delta H[u]}{\delta \pi_u} &= m_x u_t + m_x \pi_u \frac{\partial u_t}{\partial \pi_u} - \frac{\partial L_x}{\partial u_t}, \frac{\partial u_t}{\partial \pi_x}, \\ &= m_x u_t, \end{aligned} \tag{4.14}$$

which, upon substituting into Equation 4.12, gives the trivially true statement $\frac{\partial u}{\partial t} = u_t(\pi_u, u_x, u, x, t)$.

Associated with Hamilton's Equations we can define a Poisson bracket such that for some functional F, it's time derivative is given by

$$\frac{\partial F}{\partial t} = \{F, H\} + \frac{\eth F}{\eth t} \quad . \tag{4.15}$$

Here we must distinguish between the partial derivative acting on (x, t) space and that acting on (π_u, u, x, t) phase space. For the partial time derivative $\frac{\partial}{\partial t}$, F is treated as a function of (x, t), describing its spatial distribution within the disc and its evolution in time. For the partial time derivative $\frac{\eth}{\eth t}$, F is a function of the dynamical parameters $(\pi_u, \pi_{u,x}, u, u_x, x, t)$ each of which is itself a function of (x, t).

The Poisson bracket is given by

$$\{F, G\} = \int \left(\frac{\delta F}{\delta u}\frac{\delta G}{\delta \pi_u} - \frac{\delta F}{\delta \pi_u}\frac{\delta G}{\delta u}\right)\frac{dx}{m_x} \quad . \tag{4.16}$$

By antisymmetry $\{F, F\} = 0$. Additionally for any canonically conjugate coordinate, momenta pair (q, π_q) we have,

$$\{q(x), \pi_q(x')\} = \frac{1}{m_x}\delta(x - x') \quad . \tag{4.17}$$

4.2.3 Lagrangian and Hamiltonian formulation of ideal fluid dynamics

One can obtain a Lagrangian and Hamiltonian formulation for ideal fluid dynamics; a more complete discussion of the Lagrangian formalism of ideal fluids, including magnetic fields, is given in Ogilvie (2016). To do this one formulates the dynamics as a time-dependent map from a set of Lagrangian coordinates, that are conserved by the flow, into the fluid properties. The most obvious map is the flow map ϕ^t, which maps the initial conditions $\mathbf{x}_0$ into the current fluid state,

$$\mathbf{x}_0 \mapsto \mathbf{x}(\mathbf{x}_0, t) = \phi^t \mathbf{x}_0, \tag{4.18}$$

with $\mathbf{x}(\mathbf{x}_0, 0) = \mathbf{x}_0$. For simplicity, in this section, we shall only consider Cartesian $\mathbf{x}_0$ and $\mathbf{x}$. Fluid elements conserve the position vector of their initial state, $\mathbf{x}_0$, which can be used as a Lagrangian variable. The fluid velocity is obtained by taking the Lagrangian time derivative (i.e. at fixed $\mathbf{x}_0$) of the position vector,

$$\mathbf{u} = \frac{D\mathbf{x}}{Dt} = \left(\frac{\partial \mathbf{x}}{\partial t}\right)_{\mathbf{x}_0}. \tag{4.19}$$

The Jacobian of the transform (also known as the deformation tensor of the flow), along with it's inverse are

$$J_{ij} = \frac{\partial x_i}{\partial x_0^j}, \quad J_{ij}^{-1} = \frac{\partial x_{0i}}{\partial x^j}, \tag{4.20}$$

with J the Jacobian determinant. At $t = 0$, $J_{ij} = \delta_{ij}$. Finally the continuity equation is solved by a density of the form

$$\rho = \rho_0(\mathbf{x}_0) J^{-1}, \tag{4.21}$$

while entropy conservation in an ideal fluid means that the entropy is $s = s_0(\mathbf{x}_0)$.

With these expressions we can write the Lagrangian for an ideal, non-self-gravitating fluid,

$$L = \int \rho_0 \left(\frac{1}{2}\mathbf{u}^2 - \Phi(\mathbf{x}) - e(J^{-1}\rho_0, s0) \right) d^3\mathbf{x}_0, \tag{4.22}$$

where e is the internal energy which, assuming an equation of state of the form $e = e(\rho, s)$, has been expressed in terms of the known entropy and density. The variational equation for this Lagrangian is

$$\frac{\delta L}{\delta x^i} = -\frac{D}{Dt}\frac{\partial L}{\partial u^i} - \frac{\partial}{\partial x_0^j}\frac{\partial L}{\partial J^{ij}} + \frac{\partial L}{\partial x^i} = 0, \tag{4.23}$$

and after using $de = -pd\rho^{-1} = -p\rho_0^{-1}dJ$ and $\frac{\partial J}{\partial J_{ij}} = JJ_{ji}^{-1}$ this leads to the momentum equation,

$$\frac{\delta L}{\delta x^i} = -\rho_0 \frac{Du^i}{Dt} - JJ_{ji}^{-1}\frac{\partial p}{\partial a^j} - \rho_0 \frac{\partial \Phi}{\partial x^i} \tag{4.24}$$

$$= J\left(-\rho \frac{Du^i}{Dt} - \frac{\partial p}{\partial x^i} - \rho \frac{\partial \Phi}{\partial x^i} \right) = 0. \tag{4.25}$$

The fluid momentum can be obtained in the normal manor from the Lagrangian,

$$p_i = \frac{\partial L}{\partial u^i} = \rho_0 u_i \tag{4.26}$$

and we can obtain the Hamiltonian via a Legendre transform,

$$H = \int \rho_0 \left(\frac{1}{2\rho_0^2}\mathbf{p}^2 + \Phi(\mathbf{x}) + e(J^{-1}\rho_0, s0) \right) d^3\mathbf{x}_0. \tag{4.27}$$

The appropriate Poisson bracket describing the dynamics of the fluid is

$$\{F, G\} = \int \left(\frac{\delta F}{\delta x^i}\frac{\delta G}{\delta p^i} - \frac{\delta F}{\delta p^i}\frac{\delta G}{\delta x^i} \right) d^3\mathbf{x}_0 \quad . \tag{4.28}$$

The Hamiltonian generates time translations in the fluid

$$\hat{H} = \{H, \cdot\} = -\frac{\partial \cdot}{\partial t}, \tag{4.29}$$

similarly the momenta are the generators of spatial translations in the fluid

$$\hat{p} = \{\mathbf{p}, \cdot\} = -\frac{\delta \cdot}{\delta \mathbf{x}}. \tag{4.30}$$

4.2.4 The affine disc model of Ogilvie (2018)

One use of the Lagrangian formulation of ideal fluid dynamics is the affine disc model of Ogilvie (2018), from which we shall derive the ideal theory of eccentric discs in the next two chapters. This model is analogous to the shallow-water equations that have found use in geophysical fluid dynamics (Miles & Salmon, 1985). In this model the disc consists of infinitesimal, stably stratified fluid columns that undergo translation, rotation and rescaling (i.e. affine transformations), and interact via the pressure gradients in the disc.

The affine disc starts with a circular, coplanar, reference disc with coordinates $\mathbf{x}_0$. Under the affine transforms this reference disc is mapped into the dynamical state, with the position vector of a point in the dynamical state given by

$$\mathbf{x} = \mathbf{X}(\bar{\mathbf{x}}_0, t) + \mathbf{H}(\bar{\mathbf{x}}_0, t)\zeta, \tag{4.31}$$

where $\mathbf{X}$ is the position vector of the centre of the column in the dynamical state, $\mathbf{H}$ is a scale vector which describes the height and orientation of the column and ζ is a dimensionless label which denotes the height in the fluid column. The fluid velocity is then given by $\mathbf{u} = \mathbf{v} + \mathbf{w}\zeta$, where

$$\mathbf{v} = \frac{D\mathbf{X}}{Dt}, \quad \mathbf{w} = \frac{D\mathbf{H}}{Dt}. \tag{4.32}$$

Within a single column the fluid is stably stratified, but can undergo homogeneous expansion or contraction. The fluid in the reference disc obeys hydrostatic equilibrium and the vertical structure of the physical disc is obtained via the affine transform of this structure. The affine transformations do not affect the stratification of the disc, instead the hydrostatic vertical structure of the reference disc is rotated and rescaled to produce the vertical structure of the physical disc. This vertical structure is often referred to as the generalised equations of hydrostatic equilibrium (e.g. Ogilvie (2000), Ogilvie (2001), Ogilvie & Barker (2014) and Ogilvie (2016)).

The Lagrangian for the affine disc model can be obtained from the Lagrangian for an ideal fluid (Equation 4.22) by imposing the homogeneous vertical structure/generalised hydrostatic equilibrium assumption and adopting the affine transform as the map from the reference to dynamical state (as opposed to the flow map considered in the previous section). Additionally the disc is assumed to be thin so that the gravitational potential and the Jacobian matrices describing the affine transform are expanded to second order in height above the midplane. An explicit derivation for this can be found in Ogilvie (2016); here we will simply write the resulting Lagrangian,

$$L = \int \Sigma_0 \left[\frac{1}{2} \left(u^2 + w^2 \right) - \Phi(\mathbf{x}) - \frac{1}{2} \mathbf{HH} : \nabla\nabla\Phi(\mathbf{x}) - \bar{\varepsilon} \right] d^2\bar{\mathbf{x}}_0 \quad , \tag{4.33}$$

where $J_3 = JH$ is the Jacobian determinant of the 3D "affine transform", J is the Jacobian determinant of the 2D orbital coordinate system, $H = \mathbf{H} \cdot \mathbf{n}$, where $\mathbf{n}$ is the unit normal to the disc midplane, and $w = \frac{D\mathbf{H}}{Dt}$. $\bar{\varepsilon}$ is the mass weighted vertically integrated internal energy defined by

$$\bar{\varepsilon} := \frac{\int \varepsilon \rho dz}{\int \rho dz} = \frac{\int \varepsilon \rho_0 dz}{\Sigma_0} \quad . \tag{4.34}$$

To summarise the assumptions underlying the affine disc model are

1. **The disc is geometrically thin:** The gravitational potential and Jacobian of the affine transform are expanded to quadratic order in the height above the midplane.

2. **Homogeneous vertical structure:** The affine disc has a homogeneous vertical structure. The vertical structure within a single column remains fixed up to rescaling by a constant factor. The fluid within a column obeys the generalised equations for hydrostatic equilibrium.

3. **Ideal theory:** The affine disc model is an ideal theory, based on a Lagrangian formalism. Thus it contains no viscous dissipation or radiative cooling. In principle a non-ideal affine disc theory could be developed using a pseudo-Lagrangian prescription, however this is yet to be attempted.

These assumptions will be implicit in the theories developed in the next two chapters.

4.3 Orbital coordinates

4.3.1 Global orbital coordinates for an continuous eccentric disc

In Section 2.1.1 we described the orbital elements $(a, E, e, \varpi, i, \Omega)$ of a Keplerian orbit, which describe the orientation and geometry of a single orbit. We also introduced the Poincaré variables (also known as modified Delauny Variables), action-angle coordinates for the 2-body problem in Newtonian gravity,

$$\begin{aligned}
\Lambda &= \sqrt{GMa}, \quad \lambda = M + \varpi, \\
\Gamma &= \Lambda(1 - \sqrt{1-e^2}), \quad \gamma = -\varpi, \\
\mathfrak{H} &= \Lambda\sqrt{1-e^2}(1-\cos i), \quad \eta = \Omega.
\end{aligned} \tag{4.35}$$

In this section we adapt these coordinates to a continuous disc. A similar (non-canonical) coordinate system was developed by Ogilvie (2001) and Ogilvie & Barker (2014).

A planar eccentric disc involves a continuous set of nested elliptical orbits. We can describe the shape of the disc by considering e and ϖ to be functions of a. The derivatives of these functions are written as $e_a = de/da$ and $\varpi_a = d\varpi/da$. Equivalently when using modified Delaunay variables, we can consider Γ and γ to be functions of Λ, with derivatives $\Gamma_\Lambda = d\Gamma/d\Lambda$ and $\gamma_\Lambda = d\gamma/d\Lambda$. Thus e_a (or $\Gamma_\Lambda - \Gamma/\Lambda$) is a measure of the eccentricity gradient, while ϖ_a (or γ_Λ) is a measure of the twist. For a warped disc (not considered here) $\mathfrak{H}$ and η would also be functions of Λ.

We can use (Λ, λ) as a canonical orbital coordinate system covering the disc. The first coordinate Λ labels the elliptical orbits, while around each orbit λ ranges from 0 to 2π. This coordinate system covers $\mathbb{R}^2$ (or at least a local patch of $\mathbb{R}^2$, such as an elliptical annulus) and there exists a map between the (Λ, λ) coordinate system and the Cartesian (x, y) coordinate system of the disc midplane. In this map Λ determines which orbit a point lies upon, while the functions $\Gamma(\Lambda)$ and $\gamma(\Lambda)$ determine the shape of the orbit. λ then determines the point on the orbit. The polar (r, ϕ) and Cartesian (x, y) coordinates can be obtained from the equations for a Keplerian ellipse (Equations 2.2-2.4).

The coordinate system can be extended to cover $\mathbb{R}^3$ (or some local patch) by introducing the cylindrical height z. Then we have a coordinate system (Λ, λ, z) which determines a point in 3D space by first specifying a point on the orbital plane with (Λ, λ) and then translating to a point of height z above(/below) the midplane. For many applications it is useful to introduce a stretched coordinate system $(\Lambda, \lambda, \tilde{z})$, where $\tilde{z} = \frac{z}{H}$ is a stretched vertical coordinate where the cylindrical z coordinate has been rescaled by some characteristic disc thickness H. Typically H is taken to be the scale height of the disc, which varies around the orbit for an eccentric disc. The precise meaning of H is that it is the vertical standard deviation of the density density distribution defined as follows,

$$H^2 = \int \rho z^2 dz / \int \rho dz \quad , \tag{4.36}$$

with the surface density $\Sigma = \int \rho dz$.

The Jacobian of the orbital coordinate system (Λ, λ) is

$$J = \frac{\partial(x, y)}{\partial(\Lambda, \lambda)} = J^{\circ}(a) j(E), \tag{4.37}$$

where

$$J^{\circ} = \frac{2}{n}, \tag{4.38}$$

is the Jacobian for a circular disc and,

$$\begin{aligned} j &= 1 - \frac{\Gamma}{2\Lambda} - \frac{\Gamma_{\Lambda}}{2} + \frac{(\Gamma - \Lambda\Gamma_{\Lambda})\cos E}{2\sqrt{\Gamma(2\Lambda - \Gamma)}} + \frac{\sqrt{\Gamma(2\Lambda - \Gamma)}\gamma_{\Lambda} \sin E}{2} \\ &= \frac{1 - e(e + ae_a)}{\sqrt{1 - e^2}} - \frac{ae_a}{\sqrt{1 - e^2}} \cos E - ae\varpi_a \sin E, \end{aligned} \tag{4.39}$$

is dimensionless and purely related to the elliptical geometry of the disc. The Jacobian of the 3D (Λ, λ, z) coordinate system is just equal to J, while the Jacobian for the stretched coordinate system $(\Lambda, \lambda, \tilde{z})$ is

$$J_3 = JH \quad . \tag{4.40}$$

It is often useful to write the dimensionless part of the Jacobian j as

$$j = \frac{1 - e(e + ae_a)}{\sqrt{1 - e^2}} (1 - q \cos(E - E_0)) \quad , \tag{4.41}$$

where q is an orbital intersection parameter given by

$$q^2 = \frac{(ae_a)^2 + (1 - e^2)(ae\varpi_a)^2}{[1 - e(e + ae_a)]^2} \quad , \tag{4.42}$$

where we require $|q| < 1$ to avoid an orbital intersection. The angle E_0, which determines the location of maximum horizontal compression around the orbit, is determined by the relative contribution of the eccentricity gradient and twist to q,

$$\frac{ae_a}{1 - e(e + ae_a)} = q \cos E_0, \quad \frac{\sqrt{1 - e^2} ae\varpi_a}{1 - e(e + ae_a)} = q \sin E_0 \quad . \tag{4.43}$$

When $|q| > 1$ we encounter a coordinate singularity caused by the intersection of the disc orbits. This coordinate singularity arises because a (or Λ) becomes multi-valued at the orbital intersection. In a real disc this would likely result in a shock, with a discontinuous jump in orbital elements as one past the shock. Obviously such a situation requires the orbital elements to vary on the orbital timescale and cannot be treated within the framework of a thin disc theory. It can be shown that the condition for orbital intersection $|q| > 1$ is equivalent to similar criteria given by Statler (2001), $(e + ae_a)^2 + (ae\varpi_a)^2 > 1$, and Ogilvie (2001), $(e - \lambda e_\lambda)^2 + (\lambda e\varpi_\lambda)^2 > 1$.

The orbital coordinate system discussed here is not orthogonal. Conservative eccentric disc theories can normally be formulated in terms of scalar quantities, thus avoiding the need for tensor calculus. Where tensor calculus is required it is normally more convenient to switch back to the orbital coordinate system of Ogilvie (2001) and Ogilvie & Barker (2014) as the metric tensor and Christoffel symbol components are simpler in that coordinate system. In the (Λ, λ) coordinate system, the contravariant components of the orbital velocity are particularly simple and are given by $u^\Lambda = 0$ and $u^\lambda = n = (GM_1)^2/\Lambda^3$, which depends only on Λ.

Assuming the density, in an inertial frame, does not evolve on an orbital timescale the continuity equation is given by,

$$\nabla_i \rho u^i = \frac{1}{J_3}\partial_\lambda \left(J_3 \rho u^\lambda \right) = 0 \quad , \tag{4.44}$$

The mass of the disc is given by the integral,

$$\int \rho dV = \iiint \rho J_3 d\Lambda d\lambda d\tilde{z} = 2\pi \int J\Sigma d\Lambda, \tag{4.45}$$

we see that

$$2\pi J\Sigma = M_\Lambda, \tag{4.46}$$

where $M(\Lambda)$ is the mass of the disc contained within the orbit labelled by Λ, and $M_\Lambda = dM/d\Lambda$ is the one-dimensional mass density with respect to Λ .

For several applications it can be useful to rewrite the time derivative, following the orbital motion, in terms of the eccentric anomaly,

$$\frac{\partial}{\partial t} = \frac{n}{(1 - e\cos E)}\frac{\partial}{\partial E}. \tag{4.47}$$

4.3.2 The eccentric shearing coordinates of Ogilvie & Barker (2014)

Ogilvie & Barker (2014) developed a local model of an eccentric disc by use of a non-orthogonal eccentric shearing coordinate system based on a local expansion around a reference orbit. To derive this Ogilvie & Barker (2014) started from the (λ, ϕ) orbital coordinates of Ogilvie (2001). These are related to the polar coordinates (r, ϕ) by the polar equation of an ellipse,

$$r = R(\lambda, \phi) = \frac{\lambda}{1 + e(\lambda)\cos(\phi - \varpi(\lambda))}. \tag{4.48}$$

The semi-latus rectum is a orbital labelling coordinate, similar to a in the (a, E) coordinate system described in Section 4.3.1, while ϕ is the usual azimuthal angular coordinate. Ogilvie (2001) derived a number of geometrical quantities necessary for preforming vector and tensor calculus with this coordinate system.

Following Ogilvie & Barker (2014), to obtain the eccentric shearing coordinate system, one must first specify a reference orbit. Selecting a reference orbit with $\lambda = \lambda_0$ and angular velocity $\Omega(\lambda_0, \phi)$, and a reference point that follows this orbit, starting from pericentre $\phi = \varpi_0$ at $t = 0$. Let $\varphi(t)$ be a solution of $d\varphi(t)/dt = \Omega(\lambda_0, \varphi(t))$ subject to the initial condition $\varphi(0) = \varpi_0$. Then the orbital coordinates of the reference point are $(\lambda, \phi, z) = (\lambda, \varphi(t), 0)$.

We examine the neighbourhood of the reference point by letting,

$$\lambda = \lambda_0 + \xi, \quad \phi = \varphi(t) + \eta, \quad z = \zeta, \quad t = \tau, \tag{4.49}$$

where (ξ, η) are (non-orthogonal) local coordinates in the orbital plane, while ζ and τ are introduced only for notational uniformity. As we are interested in a region with an extent comparable to H in each dimension around the reference point ξ, η and ζ are small.

In the local model of Ogilvie & Barker (2014) the geometrical coefficients are approximated by their values at the reference point $(\lambda, \phi) = (\lambda, \varphi(t), 0)$, making them periodic function of τ only. In the (ξ, η, ζ) coordinate system there is a background shear flow due to the differing velocities of orbits with different ξ. As in the shearing box, used to study circular discs, we can absorb this background shear flow into the coordinate system by adopting a shearing (and oscillating) local coordinates $(\xi', \eta', \zeta', \tau')$ that are Lagrangian with respect to the local orbital motion and are defined by,

$$\xi' = \xi, \quad \eta' = \aleph(\tau)\eta + \beth(\tau)\xi, \quad \zeta' = \zeta, \quad \tau' = \tau. \tag{4.50}$$

$\aleph$ and $\beth$ correspond to a time dependent stretching of the azimuthal coordinate ($\aleph$) and a shearing of the coordinate system ($\beth$). This model is "horizontally homogeneous", with background/geometrical quantities dependent only on τ' provided that,

$$\dot{\aleph} = -\aleph\Omega_\phi, \quad \dot{\beth} = -\aleph\Omega_\lambda. \tag{4.51}$$

The Lagrangian time derivative in this coordinate system is,

$$D = \partial_{\tau'} + v^\xi(\partial_{\xi'} + \beth\partial_{\eta'}) + \aleph v^\eta\partial_{\eta'} + v^\zeta\partial_{\zeta'}, \tag{4.52}$$

where v^i is the velocity relative to the orbital motion. Unlike the Lagrangian time derivative of the (ξ, η, ζ) coordinate system, the Lagrangian time derivative of the (ξ', η', ζ') coordinate system is horizontally invariant.

Figure 2.5 of Chapter 2 illustrates the setup of the eccentric shearing box, where a nonorthogonal box is deformed as it moves around the eccentric orbit due variations between neighbouring orbits.

4.4 Averaged Lagrangians and pseudo-Lagrangians

An asymptotic method we shall make extensive use of in this thesis is the averaged Lagrangian theory of Whitham (1965) and it's extension to pseudo-Lagrangians, describing non-conservative dynamics by Jimenez & Whitham (1976). This method describes a (nonlinear) wave propagating on a slowly varying background, making use of the separation of scales between the wavelength of the wave and the lengthscale on which the background varies.

4.4.1 Averaged Lagrangian theory of Whitham (1965)

In the method of Whitham (1965) the dynamics of a (nonlinear) wave propagating on a slowly varying background can be separated into a (nonlinear) oscillator describing how the wave varies over one wavecycle, where the background can be regarded as fixed, and the modulation of this wave by the slowly varying background over many wave cycles. This long-length/time-scale dynamics is obtained by varying the Lagrangian averaged over one wave-phase. The long-lengthscale dynamics obtained describe a set of

conserved fluxes which the wave carries through the background medium. Conservation of these fluxes as the background varies results in variations in the wave amplitude.

The method of Whitham (1965) can be viewed as a method of obtaining the wave action and wave action fluxes for a nonlinear wave on a slowly varying background. The averaged-Lagrangian will be independent of wave phase, which is averaged over. Thus the conserved fluxes are the Noether currents generated by the (approximate) symmetry of the Lagrangian under a change of phase.

Consider a Lagrangian $L = L(q, q_x, x)$, the Lagrangian depends on x due to the slowly varying background. This form of Lagrangian describes the dynamics of a time independent wave, such as a standing wave, in one spatial dimension. Assume that this wave can be described by a slowly (spatially) varying wavetrain, $q = q(x, \varphi)$ where φ is the wave phase. If the wavelength of the wave is much shorter than the lengthscale l on which the the background varies then $\varphi_x \gg l^{-1}$ and we have $q_x \approx \varphi_x q_\varphi$. The Lagrangian can then be approximated by $L \approx \int L_x(q, \varphi_x q_\varphi, x)dx$. The short lengthscale dynamics is obtained by holding x fixed and varying with respect to q,

$$\frac{\delta L}{\delta q} = \frac{\partial L_x}{\partial q} - \frac{\partial}{\partial \varphi}\frac{\delta L_x}{\delta q_\varphi} = 0. \tag{4.53}$$

This is an ODE describing the dynamics of the wave on the short lengthscale as a form of nonlinear oscillator. Associated with this Lagrangian is a Hamiltonian $H = H(\pi_q, q, x)$, where $\pi_q = \frac{\partial L_x}{\partial q_\varphi}$ is the canonical momenta, which can also be used describe the short lengthscale dynamics. As the Hamiltonian does not depend explicitly on wave phase, it is constant on the short lengthscale. Setting $H = A$ where A is some constant (a measure of the wave amplitude), this provided a convenient way of expressing the phase averaged Lagrangian,

$$\langle L_x \rangle = \frac{1}{2\pi}\oint \pi_q(q, A, \varphi_x, x)dq - A \quad , \tag{4.54}$$

where $\langle\cdot\rangle = \frac{1}{2\pi}\int \cdot d\varphi$. π_q can be obtained by inverting $H(\pi_q, q, x) = A$, which is easy to do if the Hamiltonian is expressible as a sum of kinetic and potential energies. In the above expression q is just a parameter so that the integral $\oint \pi_q(q, A, \varphi_x, x)dq$ is a well defined function of A, φ_x and x only.

As an example consider the standing waves on a string with a position dependent string tension $m(x)$. For this system the Lagrangian can be written as

$$\begin{aligned} L_x &= \frac{1}{2}q_x^2 - \frac{1}{2}m(x)q^2 \\ &\approx \frac{1}{2}\varphi_x^2 q_\varphi^2 - \frac{1}{2}m(x)q^2. \end{aligned} \tag{4.55}$$

Associated with this Lagrangian we have the following Hamiltonian which describes the short lengthscale dynamics,

$$H = \frac{1}{2}\frac{\pi_q^2}{\varphi_x^2} + \frac{1}{2}m(x)q^2 \quad . \tag{4.56}$$

Setting $H = A$ and rearranging for π_q we obtain the following for the averaged Lagrangian,

$$\begin{aligned} \langle L_x \rangle &= \frac{1}{2\pi}\oint \pi_q dq - A \\ &= \frac{2\sqrt{2A}}{\pi}\varphi_x \int_0^{q_{\max}} \left(1 - \frac{m(x)}{\sqrt{2A}}q^2\right)^{1/2} dq - A \\ &= A\left(\frac{\varphi_x}{\sqrt{m(x)}} - 1\right). \end{aligned} \tag{4.57}$$

The dispersion relation for the wave can be obtained by varying the averaged Lagrangian $\langle L \rangle$ with respect to the wave amplitude A,

$$\frac{\delta\langle L\rangle}{\delta A} = \frac{1}{2\pi}\oint \frac{\partial \pi_q(q, A, \varphi_x, x)}{\partial A} dq - 1 = 0 \quad , \tag{4.58}$$

which in the above example gives $\varphi_x = \sqrt{m(x)}$. This is the lowest order dynamics of the wave and captures the same information as Equation 4.53. To obtain an equation describing how the wave is modulated by the slowly varying background we must vary L with respect to φ. The averaged Lagrangian obviously doesn't depend explicitly on φ, but it does depend on it's derivative,

$$\frac{\delta\langle L\rangle}{\delta\varphi} = -\partial_x\left(\frac{1}{2\pi}\oint \frac{\partial \pi_q(q, A, \varphi_x, x)}{\partial \varphi_x} dq\right) = 0 \quad . \tag{4.59}$$

We recognise this as a form of flux conservation law where the flux $\mathcal{J} = \frac{1}{2\pi}\oint \frac{\partial \pi_q(q,A,\varphi_x,x)}{\partial \varphi_x} dq$ is carried by the wave as it propagates. The requirement that $\mathcal{J}$ remain fixed as the wave propagates across a varying background causes a modulation of the wave ampli-

tude A, which causes it to vary with position. In the simple example of a wave on a string with variable string tension this yields $A \propto \sqrt{m(x)}$.

The averaged Lagrangian theory of Whitham (1965) is very powerful and it, along with some of it's generalisations, underpin a substantial part of this thesis.

4.4.2 Whitham's method extended to more variables

The generalisation of the method to multiple independent variables is straightforward and simply results in additional flux directions when the averaged Lagrangian is varied with respect to φ resulting in a flux conservation law of the form

$$\partial_i j^i = -\partial_i \left(\frac{\partial \langle L \rangle}{\partial \varphi_i} \right) \quad , \tag{4.60}$$

where φ_i is the derivative of the phase with respect to the i'th dimension. The short lengthscale behaviour is essentially unchanged by the inclusion of more independent variables as this is still an ODE in the phase variable. The original formulation of Whitham (1965) described a 1 dimensional travelling wave including 1 spacial and 1 time dimension, which is all we will need in this thesis.

The inclusion of additional dependent variables to the theory of Whitham (1965) is more involved and will (mostly) not be needed in this thesis. Thus we omit a general discussion of how to include additional dependent variables, although the next section will deal with a special case which is important to this work.

4.4.3 Whitham's method for a nearly integrable system

While the above describes how to use Whitham's method to deal with systems where the short lengthscale dynamics are integrable, it is useful to have a version which works for a nearly integrable system. This is particularly true when considering the dynamics of accretion discs as we can often separate the Lagrangian into a fully integrable term describing Keplerian test particle motion and a small non-integrable perturbation (e.g. due to pressure forces or perturbations from a planetary companion).

Consider a Lagrangian of the form

$$L = \sum_i p_i \dot{q}_i - H_0, \tag{4.61}$$

where H_0 is a fully integrable Hamiltonian with action-angle coordinates $\overrightarrow{J}$, $\overrightarrow{\theta_0}$, specifically the action-angle coordinates resulting from a solution to the Hamilton-Jacobi equation. Then we can rewrite the Lagrangian as

$$L = \sum_i \overrightarrow{J} \cdot \frac{\partial \overrightarrow{\theta_0}}{\partial t} \quad , \tag{4.62}$$

which results in the trivial equations of motion $\overrightarrow{J} = \text{const}$, $\overrightarrow{\theta_0} = \text{const}$. Note that, being solutions to the Hamilton-Jacobi equation (i.e. canonical coordinate/momenta which set the Hamiltonian to zero), $(\overrightarrow{J}, \overrightarrow{\theta_0})$ are subtly different from the usual action-angle coordinates $(\overrightarrow{J}, \overrightarrow{\theta})$ with the angle coordinates related by $\overrightarrow{\theta_0} = \overrightarrow{\theta} - \overrightarrow{\omega} t$, where $\overrightarrow{\omega}$ is a constant vector. The action coordinates are the same for both coordinate systems.. So while the normal angle coordinate $\overrightarrow{\theta}$ increases(/decreases) linearly with time, $\overrightarrow{\theta_0}$ always remains fixed.

To deal with this equation within the Averaged Lagrangian framework then, in addition to a phase variable φ, we must introduce a second phase like function τ with $\tau_t = \gamma = O(\epsilon)$ where $\epsilon \ll 1$. Unlike φ we do not average over the second phase function. The Action $\overrightarrow{J} = \overrightarrow{J}(\varphi, \tau)$ and angle $\overrightarrow{\theta_0} = \overrightarrow{\theta_0}(\varphi, \tau)$ and we obtain the Lagrangian,

$$L = \varphi_t \overrightarrow{J} \cdot \overrightarrow{\theta_0}_{,\varphi} + \gamma \overrightarrow{J} \cdot \overrightarrow{\theta_0}_{,\tau} \quad . \tag{4.63}$$

Averaging over the phase φ, and noting that as they are action angle coordinates $\overrightarrow{J}$ and $\overrightarrow{\theta}_0$ are independent of φ then the averaged Lagrangian becomes,

$$L = \gamma \overrightarrow{J} \cdot \overrightarrow{\theta_0}_{,\tau} \quad . \tag{4.64}$$

As expected this has no substantive difference from Equation 4.62, beyond a rescaling of the time coordinate, with $\overrightarrow{J}$ and $\overrightarrow{\theta_0}$ constant on the long timescale τ. Strictly speaking the long timescale τ is not properly defined, as there is no perturbation to the short timescale dynamics which would set the length of the long timescale.

Now consider a Lagrangian of the form

$$L = \sum_i p_i \dot{q}_i - H_0 + \epsilon H_1, \tag{4.65}$$

where H_0 is the same integrable Hamiltonian as before, while H_1 is an additional non-integrable term, and ϵ is a measure of the size of the nonintegrable terms. When $\epsilon \ll 1$, this is a form of 'nearly integrable' system. Introducing φ and τ as before we can rewrite this Lagrangian as

$$L = \varphi_t \overrightarrow{J} \cdot \overrightarrow{\theta_0}_{,\varphi} + \gamma \overrightarrow{J} \cdot \overrightarrow{\theta_0}_{,\tau} - \epsilon H_1(\overrightarrow{J}, \overrightarrow{\theta_0}, \varphi) \quad . \tag{4.66}$$

The short lengthscale dynamics are still given by $\overrightarrow{J}_{,\varphi} = \overrightarrow{\theta_0}_{,\varphi} = 0$, as the nonintegrable term only contributes on the long timescale τ, so this can be straightforwardly averaged over φ to give

$$L = \gamma \overrightarrow{J} \cdot \overrightarrow{\theta}_{0,\tau} - \epsilon \langle H_1(\overrightarrow{J}, \overrightarrow{\theta_0}, \varphi) \rangle \quad . \tag{4.67}$$

Varying this Lagrangian with respect to $\overrightarrow{J}$ and $\overrightarrow{\theta_0}$ we obtain a set of equations,

$$\frac{\partial \overrightarrow{\theta_0}}{\partial \tau} = \frac{\delta \langle H_1(\overrightarrow{J}, \overrightarrow{\theta_0}, \varphi) \rangle}{\delta \overrightarrow{J}}, \tag{4.68}$$

$$\frac{\partial \overrightarrow{J}}{\partial \tau} = -\frac{\delta \langle H_1(\overrightarrow{J}, \overrightarrow{\theta_0}, \varphi) \rangle}{\delta \overrightarrow{\theta_0}}, \tag{4.69}$$

where, we have set $\gamma = \epsilon$ (which just sets the slow timescale τ). These are just Hamilton's equations for the evolution of the (formerly) action-angle coordinate on the slow-timescale due to the non-integrable terms. Varying with respect to τ we obtain,

$$\frac{\delta L}{\delta \varphi} = -\partial_t \left(\frac{\partial L}{\partial \gamma} \right) = -\partial_t \left(\overrightarrow{J} \cdot \overrightarrow{\theta_0} \right) = 0 \quad . \tag{4.70}$$

So we have preserved one conserved quantity $\overrightarrow{J} \cdot \overrightarrow{\theta_0}$ from the original fully integrable system. This is because the averaging process removes one of the degrees of freedom. In the case of the secular theory of perturbed Keplerian orbits, $\overrightarrow{J} \cdot \overrightarrow{\theta_0}$ is related to the orbital energy. Hence, Equation 4.70 corresponds to the conservation of the orbital energy in secular theory as the secular Hamiltonian has lost any dependence on the orbital phase that the non-secular Hamiltonian might have had.

4.4.4 Pseudo-Lagrangian and weakly non-conservative terms in Whitham's averaging method

A Pseudo-Lagrangian prescription is an extension of the Lagrangian to allow for the derivation of dissipative terms from a variational principle. It works by doubling up the coordinates, with the equations of motion obtained by varying with respect to one copy of the coordinate system followed by a restriction of the second copy of the coordinates to be equal to the first.

As an example if we have a set of generalised coordinates $\overrightarrow{q}$ and duplicated coordinates $\overrightarrow{q}'$ then we can consider a Pseudo-Lagrangian of the form

$$F(\overrightarrow{q}, \overrightarrow{q}') = L(\overrightarrow{q}) + \epsilon I(\overrightarrow{q}, \overrightarrow{q}'), \tag{4.71}$$

(ϵ needn't be small), where L is a functional of $\overrightarrow{q}$ which is the Lagrangian for the conservative terms and I is a functional of $\overrightarrow{q}$ and $\overrightarrow{q}'$ from which we obtain the non-conservative terms. The equation of motion are obtained by varying F with respect to $\overrightarrow{q}$ followed by setting $\overrightarrow{q}' = \overrightarrow{q}$:

$$\frac{\delta L(\overrightarrow{q})}{\delta \overrightarrow{q}} = -\epsilon \frac{\delta I(\overrightarrow{q}, \overrightarrow{q}')}{\delta \overrightarrow{q}}\bigg|_{\overrightarrow{q}'=\overrightarrow{q}}, \tag{4.72}$$

with the terms on the right being the non-conservative terms. The order of this procedure matters as setting $\overrightarrow{q}' = \overrightarrow{q}$ before varying with respect to $\overrightarrow{q}$ would only yield conservative terms.

As an example consider adding a position dependent damping to the string considered as a toy model in Section 4.4.1. The dynamics of the steady damped wave on this string can be obtained from the following pseudo-Lagrangian density,

$$F_x = \frac{1}{2}q_x^2 - \frac{1}{2}m(x)q^2 - D(x)qq_x'. \tag{4.73}$$

Varying the pseudo-Lagrangian with respect to q we obtain

$$q_{xx} = -m(x)q - D(x)q_x, \tag{4.74}$$

which can be recognised as a form of damped harmonic motion.

Jimenez & Whitham (1976) generalised the formalism of Whitham (1965) to work with Pseudo-Lagrangian of the form (4.71) when $\epsilon \ll 1$. In this case the non-conservative terms only have an effect over many wave cycles and as such the short lengthscale dynamics are determined by L. The non-conservative terms from I appear as source/sink terms in the flux conservation equation derived for the long lengthscale dynamics.

The short lengthscale dynamics are obtained in the same manner as Whitham (1965) and are obtained from

$$\frac{\delta L}{\delta \overrightarrow{q}} = 0. \tag{4.75}$$

The variables $\overrightarrow{q}$ and $\overrightarrow{q}'$ are independent of each other and, in the short wavelength limit, are both assumed to take the form of slowly modulating waves: $\overrightarrow{q} = \overrightarrow{q}(x, \varphi(x))$, $\overrightarrow{q}' = \overrightarrow{q}'(x, \varphi'(x))$. As these waves are independent they have independent phase variables φ and φ', which are both functions of the (same) spatial variable x.

To obtain the long lengthscale dynamics Jimenez & Whitham (1976) proposed expanding the complimentary phase variable around the value of the phase variable so that we have $\overrightarrow{q} = \overrightarrow{q}' + (\varphi - \varphi')\overrightarrow{q}'_\varphi$ and

$$I(\overrightarrow{q}, \overrightarrow{q}') \approx I(\overrightarrow{q}', \overrightarrow{q}') + (\varphi - \varphi')\overrightarrow{q}'_\varphi I_q(\overrightarrow{q}', \overrightarrow{q}'), \tag{4.76}$$

with the long lengthscale dynamics obtained in a similar manner to Whitham (1965) by varying the phase averaged F with respect to the phase φ. This works because when the variables and complementary variables are equated at the end of the procedure the $O((\varphi - \varphi')^2)$ terms are set to zero. Defining $\langle\cdot\rangle := \frac{1}{2\pi}\int \cdot d\varphi$ the long lengthscale dynamics is

$$\begin{aligned}\frac{\delta\langle F\rangle}{\delta\varphi} &= \frac{\delta\langle L\rangle}{\delta\varphi} + \frac{\delta\langle I\rangle}{\delta\varphi} \\ &= -\partial_i j^i + \frac{\partial}{\partial\varphi}\left[\langle I\rangle + \langle(\varphi - \varphi')\overrightarrow{q}'_\varphi I_q(\overrightarrow{q}', \overrightarrow{q}')\rangle\right] \\ &= -\partial_i j^i + \langle\overrightarrow{q}'_\varphi I_q(\overrightarrow{q}', \overrightarrow{q}')\rangle\end{aligned} \tag{4.77}$$

when the first term is the same as the flux conservation equations, obtained with Whitham (1965)'s method, and the second term are sources/sinks generated by the non-conservative terms. This method is reliant on the non-conservative terms being weak; non-conservative terms strong enough to act on a single wave cycle break the proposed asymptotic scaling.

For our toy model of a standing wave on a string with variable string tension and (weak) damping (given by the pseudo-Lagrangian density 4.73), the relevant pseudo-Lagrangian density for the short wavelength limit is

$$F_x = \frac{1}{2}\varphi_x^2 q_\varphi^2 - \frac{1}{2}m(x)q^2 - D(x)\varphi'_x q q'_{\varphi'}, \tag{4.78}$$

where we require $D(x)|\varphi_x q_\varphi| \ll m(x)|q|$ so that the damping term is weak. As the damping is weak, it only occurs over many wave cycles meaning the short lengthscale dynamics are identical to the ideal theory described in Section 4.4.1. We can thus write the average pseudo-Lagrangian density as

$$\begin{aligned}\langle F_x\rangle &= \frac{1}{2\pi}\oint \pi_q(q,A,\varphi_x,x)dq - A - D(x)\varphi'_x\langle qq'_{\varphi'}\rangle \\ &\approx \frac{1}{2\pi}\oint \pi_q(q,A,\varphi_x,x)dq - A - D(x)\varphi'_x\langle q'q'_{\varphi'}\rangle - (\varphi-\varphi')D(x)\varphi'_x\langle (q'_{\varphi'})^2\rangle \\ &\approx \frac{1}{2\pi}\oint \pi_q(q,A,\varphi_x,x)dq - A - D(x)\varphi'_x\langle q'q'_{\varphi'}\rangle - (\varphi-\varphi')\frac{D(x)}{\varphi'_x}\frac{1}{2\pi}\oint \pi'_q dq'.\end{aligned} \tag{4.79}$$

The $\langle q'q'_{\varphi'}\rangle$ term can be dropped as it has no influence on the dynamics. Substituting in the expressions for π_q, which are obtained from the Hamiltonian describing the short lengthscale dynamics (Equation 4.56) and carrying out the averaging we obtain

$$\langle F_x\rangle = A\left(\frac{\varphi_x}{\sqrt{m(x)}} - 1\right) - (\varphi-\varphi')\frac{D(x)}{\sqrt{m(x)}}A'. \tag{4.80}$$

Varying the pseudo-Lagrangian with respect to A we obtain the same dispersion relation as the ideal theory,

$$\frac{\delta\langle F\rangle}{\delta A} = \frac{\varphi_x}{\sqrt{m(x)}} - 1 = 0. \tag{4.81}$$

To obtain the modulation of the wave on the long lengthscale we vary with respect to φ to obtain

$$\left.\frac{\delta\langle F\rangle}{\delta\varphi}\right|_{.'=.} = -\frac{\partial}{\partial x}\left(\frac{A}{\sqrt{m(x)}}\right) - D(x)\frac{A}{\sqrt{m(x)}}, \tag{4.82}$$

we see that the damping term now acts as a sink for the flux $\mathcal{J} = \frac{A}{\sqrt{m(x)}}$, which was conserved in the ideal theory, with $\partial_x\mathcal{J} = -D(x)\mathcal{J}$.

4.5 Advection operators for tensors and tensor densities

In the latter part of this thesis we shall discuss various closure models for the turbulent MRI stress. In order to do this we shall need to consider how tensors and tensor like quantities are advected by the fluid flow. For a scalar this is straight forward, and a scalar quantity ϕ passively advected by the flow obeys

$$D\phi = \frac{\partial \phi}{\partial t} + u^i \nabla_i \phi = 0 \quad , \tag{4.83}$$

where D is the Lagrangian time derivative. To obtain the equation for a tensor field A passively advected by a vector field u we require that it's Lie derivative, along the flow lines, vanishes,

$$\mathcal{L}_U A = 0 \quad , \tag{4.84}$$

where $U = \partial_t + u$, in three dimensional space plus time. For a rank (2,0) tensor the Lie derivative is given by

$$\mathcal{L}_U A^{ij} = \overset{\triangledown}{A}{}^{ij} := DA^{ij} - 2A^{k(i}\nabla_k u^{j)} \quad , \tag{4.85}$$

while for a rank (0,2) tensor this is given by

$$\mathcal{L}_U A_{ij} = \overset{\triangle}{A}_{ij} := DA_{ij} + 2A_{k(i}\nabla^k u_{j)} \quad . \tag{4.86}$$

Here D is the usual Lagrangian time derivative and $\overset{\triangle}{\cdot}$ and $\overset{\triangledown}{\cdot}$ are the Upper and Lower convective time derivatives respectively. The convention of referring to these operators as convective time derivatives is somewhat unfortunate as they are more closely related to the physical process of advection than they are to convection. However, as this is the standard naming convention, we shall make use of it here.

In Ogilvie (2003) it was shown that there is an exact asymptotic correspondence between an incompressible Oldroyd-B fluid (in the limit of large Deborah number) and an incompressible MHD fluid (in the limit of large magnetic Reynolds number). In both cases the Maxwell stress and the polymeric stress are frozen into the fluid. It is reasonable to ask what happens to this correspondence for compressible fluids.

As astrophysical plasma's have negligible molecular viscosity, we can set the solvent viscosity in the Oldroyd-B fluid to zero and consider the slightly simpler case of a compressible Maxwell fluid. There is currently a great deal of debate about the correct formulation of the compressible Maxwell model (e.g. see Bollada & Phillips (2012) for a review). For our purposes we shall use the following form:

$$\overset{\triangledown}{\mathbf{T}} + (\nabla \cdot u)\mathbf{T} = -\frac{1}{\tau}\left(T_p - \frac{\mu_p}{\tau}\mathbf{1}\right) \quad . \tag{4.87}$$

This formulation of the compressible Maxwell model is meant as the simplest generalisation of Ogilvie (2003) and is not meant to be rigorous. In general the left hand side of this equation is widely agreed upon, while the right hand side is not (Bollada

& Phillips, 2012). As the right hand side vanishes in the asymptotic correspondence of Ogilvie (2003), for our purposes, the correct formulation of the right hand side is irrelevant. In the asymptotic correspondence of Ogilvie (2003) this stress tensor should correspond to a modified Maxwell tensor,

$$\mathrm{M} = \frac{1}{\mu_0}\mathbf{BB}, \tag{4.88}$$

where $\mathbf{B}$ obey the ideal induction equation, in the limit $\tau \to \infty$. Substituting Equation 4.88 into Equation 4.87, taking the limit $\tau \to \infty$ and rearranging to give an equation for $\mathbf{B}$ we obtain

$$DB^i - B^k \nabla_k u^i + \frac{1}{2}\nabla \cdot u B^i = 0 \quad , \tag{4.89}$$

which is certainly not the ideal induction equation! Despite the fact that both the Maxwell stress and the polymeric stress are frozen into the fluid in the limit considered. So what has gone wrong? The answer comes from the fact that neither the Maxwell stress 'tensor' or the polymeric stress 'tensor' are strictly tensors and do not transform like tensors under change of volume. In fact both the Maxwell and polymeric stresses transform like tensor densities.

A tensor density of weight W transforms like a tensor, but picks up a factor of the Jacobian raised to the power W. If $\mathfrak{A}$ is a tensor density of weight W then $\mathfrak{A}J^{-W}$ transforms tensorally. This is concept is particularly useful in compressive flows as the density transforms like J^{-1} and many quantities are expressible as a tensor multiplied by the density raised to some power.

For a contravariant(/covariant) tensor density $\mathfrak{A}$ of rank W conservative advection of the associated tensoral quantity $\mathfrak{A}J^{-W}$ is given by setting it's Lie derivative, along the flow lines, to zero,

$$\mathcal{L}_U \mathfrak{A} J^{-W} = J^{-W}\mathcal{L}_U \mathfrak{A} - W J^{-W}\mathfrak{A}\nabla \cdot u = 0 \quad . \tag{4.90}$$

For contravariant $\mathfrak{A}$ this is given by

$$\mathcal{L}_U \mathfrak{A}^{ij} J^{-W} = J^{-W} D\mathfrak{A}^{ij} - 2J^{-W}\mathfrak{A}^{k(i}\nabla_k u^{j)} - W J^{-W}\mathfrak{A}^{ij}\nabla \cdot u \quad , \tag{4.91}$$

and for covariant $\mathfrak{A}$,

$$\mathcal{L}_U \mathfrak{A}_{ij} J^{-W} = J^{-W} D\mathfrak{A}_{ij} + 2J^{-W}\mathfrak{A}_{k(i}\nabla^k u_{j)} - W J^{-W}\mathfrak{A}_{ij}\nabla \cdot u \quad , \tag{4.92}$$

It is useful to introduce an advective derivative $\mathcal{D}$ which acts on a tensor densities (with weight W) so that $\mathcal{D}\mathfrak{A} = 0$ passively advects the tensoral quantity $\mathfrak{A}^{ij}J^{-W}$. We define it using the Lie derivative so that

$$\mathcal{D}\mathfrak{A} := J^W \mathcal{L}_U \mathfrak{A} J^{-W}. \tag{4.93}$$

This operator obeys all the usual properties associated with a derivative, such as linearity,

$$\mathcal{D}(a\mathfrak{g} + b\mathfrak{f}) = a\mathcal{D}\mathfrak{g} + b\mathcal{D}\mathfrak{f}, \tag{4.94}$$

where $\mathfrak{g}$ and $\mathfrak{f}$ are tensor densities with the same weight and a, b are scalar constants. This follows straightforwardly from the properties of the Lie derivative. $\mathcal{D}$ also obeys a Leibniz rule. For $\mathfrak{g}$ a tensor density of weight n and $\mathfrak{f}$ a tensor density of weight n,

$$\begin{aligned}
\mathcal{D}\mathfrak{g}\mathfrak{f} &= J^{(n+m)}\mathcal{L}_U J^{-(n+m)}\mathfrak{g}\mathfrak{f} \\
&= J^{(n+m)}\mathcal{L}_U(J^{-n}\mathfrak{g}J^{-m}\mathfrak{f}) \\
&= J^n\mathcal{L}_U(J^{-n}\mathfrak{g})\mathfrak{f} + \mathfrak{g}J^m\mathcal{L}_U(J^{-m}\mathfrak{f}) \\
&= \mathcal{D}(\mathfrak{g})\mathfrak{f} + \mathfrak{g}\mathcal{D}\mathfrak{f},
\end{aligned} \tag{4.95}$$

where we have made use of the Leibniz rule for the Lie derivative, along with the fact that $J^{-n}\mathfrak{g}$ and $J^{-m}\mathfrak{f}$ are tensoral quantities.

The Maxwell stress M^{ij} is a tensor density with weight -2 and the ideal induction equation is given by $\mathcal{D}M^{ij} = 0$, explicitly,

$$\mathcal{D}M^{ij} = (\partial_t + u^k\nabla_k)M^{ij} - 2M^{k(i}\nabla_k u^{j)} + 2\nabla\cdot u M^{ij} = 0 \quad , \tag{4.96}$$

while the elastic stress T^{ij} for an elastic fluid is a tensor density with weight -1 (Bollada & Phillips, 2012). The elastic stress is advected by the fluid and thus evolves according to

$$\mathcal{D}T^{ij} = (\partial_t + u^k\nabla_k)T^{ij} - 2T^{k(i}\nabla_k u^{j)} + \nabla\cdot u T^{ij} = 0 \quad . \tag{4.97}$$

This highlights an important difference between viscoelastic fluids and MHD plasmas, namely the Maxwell stress and viscoelastic stresses transform differently under changes of volume.

The different weights of the Maxwell and polymeric stresses can be understood by considering the underlying quantities advected by the flow. For the Maxwell stress, freezing in of the magnetic field means that the quantity $\frac{B^i}{\rho}$ is passively advected by

the fluid flow. This is equivalent to the advection of the tensoral quantity $\frac{B^i B^j}{\rho^2} \propto \frac{M^{ij}}{\rho^2}$, hence M^{ij} has weight -2. In the case of a viscoelastic stress, it is possible to derive the viscoelastic stress model by considering a fluid containing a large number of "polymers", which are modelled as two masses connected by a spring. The two ends of the "polymer" are advected by the fluid, with the spring resisting fluid shearing that would separate the two ends of the polymer, resulting in a stress. The magnitude of this stress is proportional to the number of these polymers in a given volume. Therefore the polymeric stress is proportional to the fluid density, and thus has weight -1.

Part II

Ideal Theory of Eccentric Discs

Chapter 5

Hamiltonian Theory of Eccentric Discs

5.1 Introduction

Eccentric gaseous discs, in which the dominant motion of the fluid consists of elliptical Keplerian orbits around a central mass, have numerous applications in astrophysics. In many of these applications the eccentricities and their gradients are not small enough to be well described by the linear theories discussed by Papaloizou (2002), Goodchild & Ogilvie (2006), Ogilvie (2008) and Teyssandier & Ogilvie (2016). A nonlinear theory of eccentric discs was developed by Ogilvie (2001) and rederived from an alternative viewpoint by Ogilvie & Barker (2014). Evolutionary equations in one spatial dimension were obtained for the shape and mass distribution of the disc; however, these equations are complicated because they contain nonlinear functions of the eccentricity, eccentricity gradient and twist that must be (pre-)computed numerically by solving a set of nonlinear ordinary differential equations around each orbit and evaluating a number of integrals of stress components combined with geometrical factors. Perhaps as a result of this complexity, the nonlinear aspects of the theory have not been investigated widely in the literature, although inviscid nonlinear modes have been computed and tested in 2D simulations by Barker & Ogilvie (2016).

The purpose of this chapter is to present an alternative, and in some ways much simpler, approach to the nonlinear dynamics of eccentric discs. We focus on the special case in which dissipation can be neglected; it is then possible to apply a Hamiltonian formulation, which has attractive mathematical properties and assists in the physical interpretation of the solutions. As in the case of the classical dynamics of systems of

particles or rigid bodies, Hamiltonian methods are useful mainly because they clarify the mathematical structure of the problem and provide new insights.

It is well known from the linear theory that eccentricity propagates through gaseous discs predominantly by means of pressure, in the form of a dispersive wave. Viscous forces or other dissipative effects generally contribute a slow diffusion and/or damping of eccentricity. The ideal fluid theories that we discuss in this chapter describe the nonlinear version of the propagation of eccentricity by means of pressure, and neglect the weaker effects of viscous forces and other dissipative effects. Self-gravity can also make an important contribution to the propagation of eccentricity, e.g. in planetary rings; in principle it is relatively easy to include this within the Hamiltonian framework, however one will run into the issue of the issue of how to regularise the Laplace-Lagrange theory which is singular in the continuous disc limit (normally done by introducing some form of smoothing length). We do not consider self-gravity in this chapter.

5.2 Preliminaries and Assumptions

In this chapter we shall derive a Hamiltonian theory of eccentric discs. The 3D theory will be based of a Whitham average of the affine disc model of Ogilvie (2018). Thus all the assumptions which go into the derivation of the affine disc model apply to the theory derived here. In addition the theory we shall derive is a secular theory, so the orbital geometry (as described by the eccentricity and longitude of pericentre) are constant on orbital timescales. Finally we shall not consider the effects of disc warps and restrict ourselves to discs with a flat midplane.

To summarise the assumptions required for the Hamiltonian theory developed in this chapter (in addition to the obvious assumptions of hydrodynamics and Newtonian mechanics) are

1. **The disc is geometrically thin:** As in the affine disc model the potential is expanded to second order in height above the midplane.

2. **Homogeneous vertical structure:** The affine disc has a homogeneous vertical structure and this carries over to the eccentric disc. One additional consequence of this assumption is the eccentricity and longitude of pericentre will be independent of height.

3. **Ideal theory:** The affine disc, and thus the eccentric disc theory is an ideal fluid theory. This means we neglect viscous dissipation and radiative cooling. This assumption will be partially relaxed in later chapters.

4. **Secular Theory:** The theory developed here is a secular theory, where the dynamics are dominated by Keplerian motion on the orbital timescale. Non-Keplerian forces, such as from pressure gradients, cause the disc orbits to evolve on a timescale long compared with the orbital period. The vertical structure, however, is dynamical and undergoes a periodic oscillation during each orbit.

5. **Coplanar midplane:** Disc warps will be neglected throughout this thesis.

5.3 Derivation of the Hamiltonian for a thin eccentric disc

5.3.1 2D theory

The Lagrangian for 2D ideal hydrodynamics subject to a gravitational potential $\Phi(x)$ is (Ogilvie, 2016)

$$L = \int \Sigma_0 \left[\frac{1}{2}\mathbf{u}^2 - \Phi(\mathbf{x}) - \varepsilon\right] d^2\mathbf{x}_0, \tag{5.1}$$

where $\mathbf{x} = \mathbf{x}(\mathbf{x}_0, t)$ is the position vector of a fluid element in the disc, and the fluid velocity $\mathbf{u} = \frac{D\mathbf{x}}{Dt}$, where $\frac{D}{Dt}$ is the Lagrangian time derivative. $\mathbf{x}_0$ are Lagrangian coordinates which label the fluid element according to their position in some reference state, with $J = |\frac{\partial(\mathbf{x})}{\partial \mathbf{x}_0}|$ the Jacobian determinant of the time dependent map from the reference state to the dynamical state. ε is the specific internal energy of the fluid.

To develop a secular theory for the evolution of the eccentricity we must separate out the $O(1)$ contribution to the Lagrangian L_K, which describes Keplerian motion around a Newtonian point mass potential from the perturbations,

$$L = L_K - \epsilon^2 \int \Sigma_0 \left[V(\mathbf{x}) + \varepsilon\right] d^2\mathbf{x}_0, \tag{5.2}$$

where ϵ is a bookkeeping parameter and $V(\mathbf{x}) = \Phi(\mathbf{x}) - \Phi_K(R)$ is the (small) non-point mass contribution to the potential, with R the cylindrical radius and $\Phi_K = -\frac{GM_1}{R}$ the Keplerian point mass potential.

L_K, the $O(1)$ contribution to the Lagrangian is given by

$$L_K = \int \Sigma_0 \left[\frac{1}{2}\left(\dot{R}^2 + R^2\dot{\phi}^2\right) - \Phi_K(R)\right] d^2\mathbf{x}_0 \quad . \tag{5.3}$$

This is of course completely integrable and can be rewritten in terms of the action angle coordinates:

$$L_K = \int \Sigma_0 \left[\Lambda\dot{\lambda} + \Gamma\dot{\gamma} + \frac{(GM_1)^2}{2\Lambda^2} \right] d^2\mathbf{x}_0 \quad , \tag{5.4}$$

where we have used the (Λ, λ) action angle coordinates which are based on the modified Delauny variables (reviewed in Section 4.3.1). Assuming the ϵ^2 terms are small compared to L_K then the Lagrangian given by Equation 5.2 is in a "nearly integrable form", where the non-integrable $O(\epsilon^2)$ terms cause slow evolution of the action angle coordinates of the integrable $O(1)$ term L_K. This means we can make use of the Whitham's averaging discussed in Chapter 4 (e.g. see equation 4.67). Carrying out this averaging we obtain

$$L = \int M_\Lambda \left[\Lambda\dot{\lambda} + \Gamma\dot{\gamma} + \frac{(GM_1)^2}{2\Lambda^2} - \langle V \rangle - \langle \varepsilon \rangle \right] d\Lambda \quad , \tag{5.5}$$

where $\langle\cdot\rangle$ denote the Whitham average, which, in this instance, corresponds to an orbit average, and we have dropped the ϵ^2 bookkeeping parameters. We have made use of the fact that Λ, Γ, $\dot{\lambda}$ and $\dot{\gamma}$ are constant over one orbit due to the fact they are action-angle coordinates. We have introduced M_Λ the mass per unit Λ and rewritten integration over d^2x_0 by making use of the following,

$$\int \Sigma_0 \cdot d^2\mathbf{x}_0 = \int\int \Sigma_0 J^\circ \cdot d\Lambda d\lambda = \int 2\pi\Sigma_0 J^\circ \langle\cdot\rangle d\Lambda = \int M_\Lambda \langle\cdot\rangle d\Lambda \quad , \tag{5.6}$$

where $J^\circ = 2/n$ is the Jacobian determinant of the orbital coordinate system in the reference circular disc. i.e. J° is the Jacobian determinant of the map from reference state labelled by $\mathbf{x}_0$ to a reference state labelled by the modified Delauny variables. This reference disc is the circular disc with the same mass and entropy on each orbit (labelled by Λ) as the eccentric disc.

For an ideal fluid disc the entropy s is constant on each orbit, $s = s(a)$. We can express the equation of state as a relationship between the specific internal energy, the entropy and the surface density: $\varepsilon = f(s, \Sigma)$, for some function f. As $s = s(a)$ and, by mass conservation, $\Sigma = \Sigma_0(a)/J$ the specific internal energy can be written as a function of the a and J, $\varepsilon(a, J) = f(s(a), \Sigma_0(a)/J)$.

At lowest order the dynamics of Lagrangian (5.5) is just the orbital motion. With the Lagrangian written in this form the orbital motion is particularly simple, with λ advancing linearly with time at fixed Λ. We can separate this part of the Lagrangian describing the orbital motion from that describing the slow evolution of the orbits,

$$L = \int M_\Lambda \left[\Gamma\dot{\gamma} - \langle V \rangle - \langle \varepsilon(a, J) \rangle \right] d\Lambda \quad . \tag{5.7}$$

This is the Lagrangian that describes the dynamics of an eccentric wave in a (2D) disc. Performing a Legendre transform we obtain the Hamiltonian,

$$\mathcal{H} = \int M_\Lambda \left[\langle V\rangle + \langle \varepsilon(a, J)\rangle\right] d\Lambda \quad . \tag{5.8}$$

With Hamilton's equations for an eccentric disc given by

$$M_\Lambda \frac{\partial \Gamma}{\partial t} = -\frac{\delta \mathcal{H}}{\delta \gamma}, \quad M_\Lambda \frac{\partial \gamma}{\partial t} = \frac{\delta \mathcal{H}}{\delta \Gamma}, \tag{5.9}$$

where $\frac{\delta}{\delta\gamma}$ and $\frac{\delta}{\delta\Gamma}$ denote functional derivatives with respect to γ and Γ respectively. The dynamics described by these equations is identical to that given by the Euler-Lagrange equation which can be obtain by independently varying Equation 5.7 with respect to Γ and γ.

5.3.2 3D theory

Now we generalise the theory to a 3D disc. To do this we shall have to account for the variation in the scale height which arises due to the variation of vertical gravity, and horizontal compression, which occurs around an eccentric orbit. A convenient way of formulating this is with the affine disc model of Ogilvie (2018) (reviewed in Section 4.2.4). The Lagrangian of the affine disc model is given by

$$L = \int \Sigma_0 \left[\frac{1}{2}\left(u^2 + w^2\right) - \Phi(\mathbf{x}) - \frac{1}{2}\mathbf{HH} : \nabla\nabla\Phi(\mathbf{x}) - \bar{\varepsilon}\right] d^2\bar{\mathbf{x}}_0 \quad , \tag{5.10}$$

where $J_3 = JH$ is the Jacobian of the 3D "affine transform", J is the Jacobian of the 2D orbital coordinate system, $H = \mathbf{H}\cdot\mathbf{n}$, where $\mathbf{n}$ is the unit normal to the disc midplane, and $w = \frac{D\mathbf{H}}{Dt}$. For an unwarped disc confined to a single plane $n = \hat{e}_z$ and we can take $\mathbf{H} = H\hat{e}_z$, with $w^2 = \dot{H}^2$. It is also convenient to introduce a Lagrangian vertical coordinate: $\tilde{z} = z/H$. $\bar{\varepsilon}$ is the mass weighted vertically integrated internal energy defined by

$$\bar{\varepsilon} := \frac{\int \varepsilon\rho dz}{\int \rho dz} = \frac{\int \varepsilon\rho_0 dz}{\Sigma_0} \quad . \tag{5.11}$$

In an ideal fluid disc the specific entropy s is a conserved quantity and can be written as a function of the Lagrangian coordinates only, $s = s(\bar{\mathbf{x}}_0, \tilde{z})$. We express the equation of state as a relationship between the specific internal energy, the entropy and the density: $\varepsilon = f(s, \rho)$, for some function f. Thus by conservation of mass and entropy

we can write the vertically integrated internal energy as a function of the Jacobian J_3 and the horizontal Lagrangian coordinates $\bar{\mathbf{x}}_0$:

$$\bar{\varepsilon}(\bar{\mathbf{x}}_0, J_3) = \frac{1}{\Sigma_0}\int f(s(\bar{\mathbf{x}}_0, \tilde{z}), \rho_0(\bar{\mathbf{x}}_0, \tilde{z})J_3^{-1})\rho_0 dz \quad . \tag{5.12}$$

For this unwarped disc the Lagrangian simplifies to

$$L = \int \Sigma_0 \left[\frac{1}{2}\left(u^2 + \dot{H}^2\right) - \Phi(\mathbf{x}) - \frac{1}{2}H^2\Psi(\mathbf{x}) - \bar{\varepsilon}(\bar{\mathbf{x}}_0, J_3)\right] d^2\bar{\mathbf{x}}_0, \tag{5.13}$$

where we have introduced $\Psi(\mathbf{x}) := \frac{\partial^2\Phi(\mathbf{x})}{\partial z^2}|_{z=0}$. As in the 2D theory we separate out the $O(1)$ Keplerian contribution to the Lagrangian L_K (equation 5.3) and a perturbation,

$$L = L_K + \epsilon^2 \int \Sigma_0 \left[\frac{1}{2}\dot{H}^2 - V(\mathbf{x}) - \frac{1}{2}H^2\Psi(R) - \bar{\varepsilon}(\bar{\mathbf{x}}_0, J_3)\right] d^2\bar{\mathbf{x}}_0 \quad . \tag{5.14}$$

We can obtain the dynamics on the short, orbital timescale by varying L with respect to $\mathbf{x}$, from which we obtain

$$\frac{\delta L}{\delta \mathbf{x}} \approx \frac{\delta L_K}{\delta \mathbf{x}} = 0 \quad , \tag{5.15}$$

which as in the 2D disc is solved by Keplerian orbital motion. Additionally we obtain the dynamics of H on the orbital timescale by varying the Lagrangian with respect to H,

$$\begin{aligned} \epsilon^{-2}\frac{\delta L}{\delta H} &= -\ddot{H} - H\Psi(R) - \frac{\partial\bar{\varepsilon}}{\partial H}, \\ &= -\ddot{H} - H\Psi(R) + \frac{P}{\Sigma H} = 0 \quad , \end{aligned} \tag{5.16}$$

where P is the vertically integrated pressure. The last term has been simplified using the thermodynamic relation, valid for adiabatic changes, $d\varepsilon = p\rho^{-2}d\rho$.

As in the 2D theory we can Whitham average the Lagrangian, using the fact that the Lagrangian is nearly integrable and written in the action-angle coordinate of L_K,

$$\begin{aligned} \mathcal{L} = &\int M_\Lambda \left[\Lambda\dot{\lambda} + \Gamma\dot{\gamma} + \frac{(GM_1)^2}{2\Lambda^2}\right] d\Lambda \\ &+ \int M_\Lambda \left\langle \frac{1}{2}\dot{H}^2 - V(\mathbf{x}) - \frac{1}{2}H^2\Psi(R) - \bar{\varepsilon}(\Lambda, J_3)\right\rangle d\Lambda \quad . \end{aligned} \tag{5.17}$$

Here we see in addition to the internal energy the vertical oscillation of the fluid column, described by Equation 5.16, also causes a slow evolution of the disc orbits. We can simplify this Lagrangian by making use of the virial relation for H, obtained by multiplying Equation 5.16 by H and averaging (with integration by parts),

$$\left\langle -\dot{H}^2 + \Psi H^2 - \frac{P}{\Sigma} \right\rangle = 0 \tag{5.18}$$

Using equation 5.18 the Average Lagrangian becomes

$$\mathcal{L} = \int M_\Lambda \left[\Lambda\dot{\lambda} + \Gamma\dot{\gamma} + \frac{(GM_1)^2}{2\Lambda^2} - \langle V(\mathbf{x}) \rangle - \left\langle \bar{\varepsilon}(\Lambda, J_3) + \frac{P}{2\Sigma} \right\rangle \right] d\Lambda \quad . \tag{5.19}$$

Again the dynamics of λ is dominated by the orbital motion, while Λ remains fixed in this secular theory. Therefore we can separate out the terms in the Lagrangian describing the orbital motion from those describing the slow evolution of the disc orbits,

$$\mathcal{L} = \int M_\Lambda \left[\Gamma\dot{\gamma} - \langle V(\mathbf{x}) \rangle - \left\langle \bar{\varepsilon}(\Lambda, J_3) + \frac{P}{2\Sigma} \right\rangle \right] d\Lambda \quad . \tag{5.20}$$

An alternative form of this Lagrangian, written in terms of the non-canonical orbital elements e, ϖ and a, which is useful for many applications, is

$$\mathcal{L} = \int M_a \left[na^2\sqrt{1-e^2}(\dot{\varpi} - \omega_f(a,e)) - \left\langle \bar{\varepsilon}(a, J_3) + \frac{P}{2\Sigma} \right\rangle \right] da. \tag{5.21}$$

Written in this form, the effects of external perturbations/departures from a Newtonian point mass potential, encapsulated by the $M_\Lambda\langle V(\mathbf{x})\rangle$ in Equation 5.20, are captured by the forced precession frequency $\omega_f(a,e)$ defined by

$$\omega_f(a,e) := \frac{1}{na^2\sqrt{1-e^2}}\langle V(\mathbf{x})\rangle \quad . \tag{5.22}$$

An alternative definition, which is useful when $e \ll 1$, is to define ω_f in terms of the radial and azimuthal epicyclic frequencies,

$$\omega_f(a) := \frac{1}{2n}\left(\Omega^2 - \kappa^2\right) \quad , \tag{5.23}$$

where n is still the mean motion of the Keplerian orbit. This form is convenient for studying nearly circular ($e \ll 1$) relativistic discs, where $\Omega \neq \kappa \neq n$, which can never the less be nonlinear as ae_a and $ae\varpi_a$ need not be small. This limit will be considered in detail in the next chapter.

By use of a Legendre transform of Equation 5.20 we obtain the Hamiltonian for the secular evolution of the orbits in a disc around a (nearly Keplerian) gravitational potential subject to pressure forces,

$$\mathcal{H} = \int M_\Lambda \left\langle V(\mathbf{x}) + \bar{\varepsilon}(a, J_3) + \frac{P}{2\Sigma} \right\rangle d\Lambda \quad . \tag{5.24}$$

For a perfect gas $P = K(a)HJ_3^{-\gamma}$, where γ is the ratio of specific heats [1], and the Hamiltonian simplifies to

$$\mathcal{H} = \int M_\Lambda \left\langle V(\mathbf{x}) \right\rangle d\Lambda + \frac{1}{2}\frac{\gamma+1}{\gamma-1} \int 2\pi K J^\circ \left\langle J_3^{-(\gamma-1)} \right\rangle d\Lambda \quad . \tag{5.25}$$

Hamilton's equations are again given by

$$M_\Lambda \frac{\partial \gamma}{\partial t} = \frac{\delta \mathcal{H}}{\delta \Gamma}, \quad M_\Lambda \frac{\partial \Gamma}{\partial t} = -\frac{\delta \mathcal{H}}{\delta \gamma} \quad . \tag{5.26}$$

As in the 2D theory, these match the Euler-Lagrange equations obtained by varying Equation 5.20 with respect to Γ and γ.

We can also obtain Hamilton's equations in terms of the non-canonical orbital elements (a, e, ϖ):

$$M_a \frac{\partial \varpi}{\partial t} = -\frac{\sqrt{1-e^2}}{na^2 e}\frac{\delta \mathcal{H}}{\delta e}, \tag{5.27}$$

$$M_a \frac{\partial e}{\partial t} = \frac{\sqrt{1-e^2}}{na^2 e}\frac{\delta \mathcal{H}}{\delta \varpi} \quad . \tag{5.28}$$

5.4 Symmetries and conserved quantities

It is possible to define the Poisson bracket, in terms of the canonical (Λ, λ) coordinate system, so that for some functional F its evolutionary equation can be obtained by taking the Poisson bracket of F with the Hamiltonian,

$$\frac{\partial F}{\partial t} = \{F, \mathcal{H}\} + \frac{\eth F}{\eth t} \quad . \tag{5.29}$$

As in Chapter 4, we must distinguish between the partial time derivative acting on (Λ, t) space ($\frac{\partial}{\partial t}$) and that acting on $(\Gamma, \gamma, \Lambda, t)$ phase space ($\frac{\eth}{\eth t}$).

The Poisson bracket which agrees with Hamilton's equations (Equations 5.26) is

[1] there is an unfortunate notational clash between the ratio of specific heats and the modified Delauny variables, the distinction should normally be clear in context

$$\{F,G\} := \int \left(\frac{\delta F}{\delta\gamma(\Lambda)} \frac{\delta G}{\delta\Gamma(\Lambda)} - \frac{\delta F}{\delta\Gamma(\Lambda)} \frac{\delta G}{\delta\gamma(\Lambda)} \right) \frac{d\Lambda}{M_\Lambda} \quad . \tag{5.30}$$

In the absence of external perturbations, the Hamiltonian (and Hamiltonian density) do not depend explicitly on time. The Hamiltonian is conserved by antisymmetry of the Poisson bracket $\{\mathcal{H},\mathcal{H}\} = 0$. The conservation of $\mathcal{H}$ is related, by Noether's theorem, to the time translation symmetry of the Hamiltonian

Another symmetry present, in the absence of external perturbations, is global rotational symmetry. This means H_Λ does not depend explicitly on γ, and only depends on the relative orientation of the eccentric orbits through its dependence on γ_Λ. This global rotational symmetry is related to conservation of AMD in the disc, with the AMD of the disc given by

$$C = \int M_\Lambda \Gamma \, d\Lambda \quad . \tag{5.31}$$

Computing its Poisson bracket we get

$$\frac{\partial C}{\partial t} = \{C,\mathcal{H}\} = -\int \frac{\delta\mathcal{H}}{\delta\gamma} d\Lambda = \int \frac{\partial}{\partial\Lambda} \frac{\partial H_\Lambda}{\partial\gamma_\Lambda} d\Lambda = \left. \frac{\partial H_\Lambda}{\partial\gamma_\Lambda} \right|_{\Lambda_{\text{in}}}^{\Lambda_{\text{out}}} \quad . \tag{5.32}$$

This vanishes subject to suitable boundary conditions, such that the AMD flux $-\partial H_\Lambda/\partial\gamma_\Lambda$ entering the domain balances that leaving the domain, i.e.

$$-\left. \frac{\partial H_\Lambda}{\partial\gamma_\Lambda} \right|_{\Lambda_{\text{in}}} = -\left. \frac{\partial H_\Lambda}{\partial\gamma_\Lambda} \right|_{\Lambda_{\text{out}}} . \tag{5.33}$$

The AMD is conserved because angular momentum is conserved and a (or Λ) is a material invariant. The net transport of angular momentum, or of AMD, is related to the twisting of the eccentric disc, because the flux $-\frac{\partial H_\Lambda}{\partial\gamma_\Lambda}$ is found to vanish in the untwisted case $\varpi_a = 0$. For an untwisted eccentric disc, AMD is transported by superimposed ingoing and outgoing eccentric waves, carrying equal and opposite AMD flux.

5.5 Hamiltonian for a disc of perfect gas

In this section we consider the Hamiltonian for a perfect gas in the absence of any external forces. In this case it is convenient to write the Hamiltonian density as

$$H_\Lambda = H^\circ_\Lambda F, \quad \text{or} \quad H_a = H^\circ_a F, \tag{5.34}$$

where F is the "geometric" part of the Hamiltonian density depending only on $e, ae_a, ae\varpi_a$ and the adiabatic exponent γ, while

$$H^\circ_\Lambda = \frac{4\pi P^\circ}{n}, \quad \text{or} \quad H^\circ_a = 2\pi a P^\circ \tag{5.35}$$

is a fixed function of Λ (or a), P° being the (2D, or vertically integrated) pressure of a circular disc with the same distributions of mass and entropy.

For a 2D disc the 'geometric' part of the Hamiltonian density is given by

$$F^{(2D)} = \frac{1}{\gamma - 1} \left\langle j^{-(\gamma-1)} \right\rangle \quad , \tag{5.36}$$

where $j = J/J^\circ$ is a dimensionless Jacobian determinant which has been normalised by the Jacobian determinant of the circular disc. In the isothermal limit of a perfect gas with $\gamma = 1$ we have instead

$$F^{(2D)} = -\langle \ln j \rangle \quad . \tag{5.37}$$

Explicit expressions for $F^{(2D)}$, in the case of a perfect gas, can be obtained in terms of Legendre functions.

In the realistic case of a 3D disc, the geometric part of the Hamiltonian is given by

$$F^{(3D)} = \frac{1}{2}\frac{\gamma + 1}{\gamma - 1} \left\langle (jh)^{-(\gamma-1)} \right\rangle, \tag{5.38}$$

where $h(E) = H/H^\circ$ is a dimensionless vertical scale height, with H° the scale height of the reference circular disc. In general h must be obtained as the solution of a second-order differential equation, which can be obtained by nondimensionalising Equation 5.16. In the isothermal limit we have instead

$$F^{(3D)} = -\langle \ln(jh) \rangle \quad . \tag{5.39}$$

In Figure 5.1 the geometric parts of the 2D and 3D Hamiltonians are compared in the case of untwisted discs. The four panels show how much the dynamics of an eccentric disc is affected by its thermodynamic behaviour and, in particular, by the dynamical vertical structure of a 3D disc.

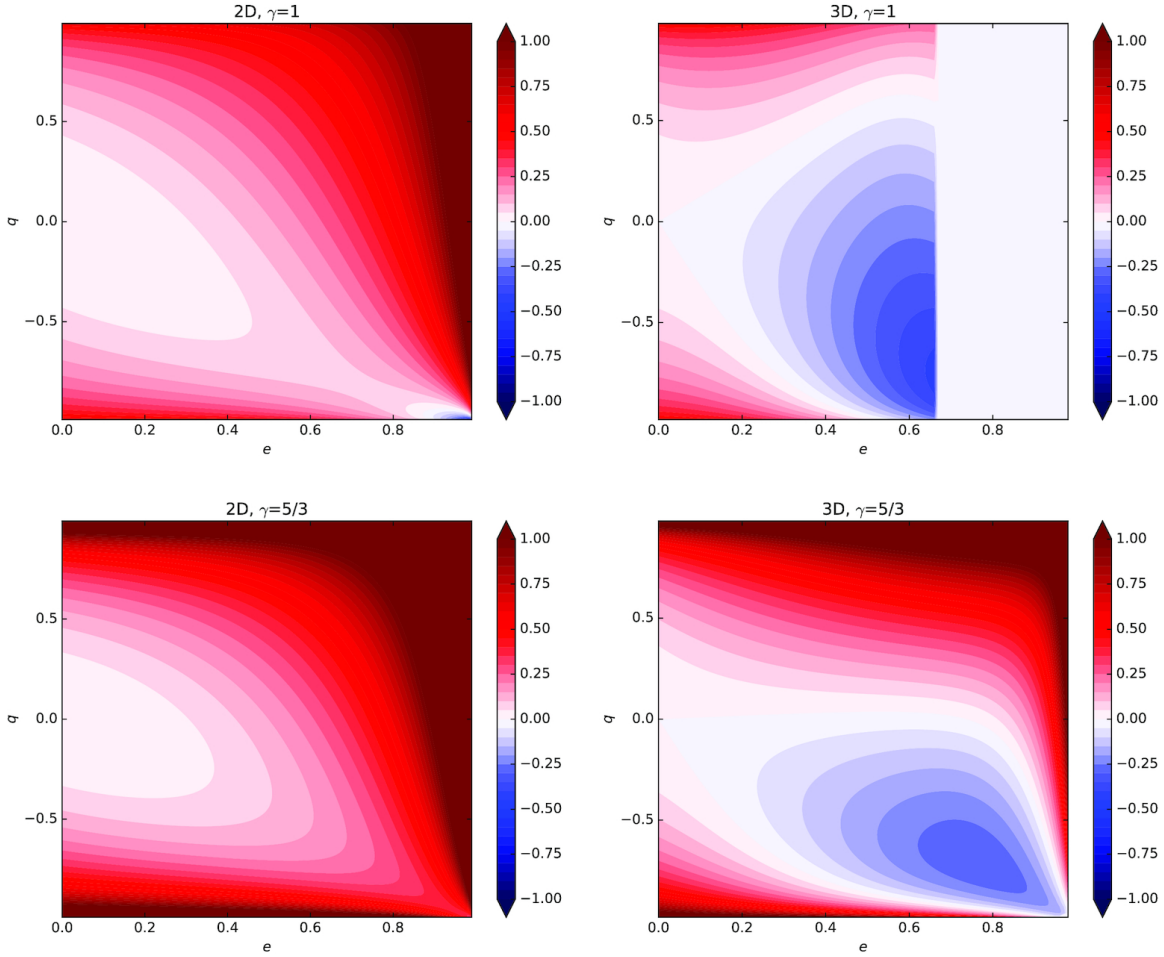


Fig. 5.1 Dimensionless geometric part of the Hamiltonian density for untwisted eccentric discs, as a function of the eccentricity e and the (signed) nonlinearity parameter $q = ae_a/[1 - e(e + ae_a)]$. Top left: 2D case for $\gamma = 1$. Top right: 3D case for $\gamma = 1$. (Values are not plotted for $e > 0.66$ because the solution involves extreme compression at periapsis.) Bottom left: 2D case for $\gamma = 5/3$, with the additive constant term removed. Bottom right: 3D case for $\gamma = 5/3$, with the additive constant term removed.

5.6 Nonlinear eccentric modes

The simplest solutions to Equations 5.27 and 5.28 are obtained by requiring that the complex eccentricity $\mathcal{E} = e \exp i\varpi$ depends on time only through a phase factor so that $\mathcal{E} \propto \exp i\omega t$. This corresponds to a solution that is steady in some rotating frame, i.e. a uniformly precessing eccentric disc. When $\mathcal{E}$ takes this form, the twisted solutions are eccentric travelling waves whilst the untwisted solutions are the standing waves or modal solutions of the disc. The equations for these eccentric waves are

$$0 = \frac{\delta H}{\delta \varpi} \quad , \tag{5.40}$$

$$M_a \omega = -\frac{\sqrt{1-e^2}}{na^2 e} \frac{\delta H}{\delta e} \quad , \tag{5.41}$$

where $\omega = \dot{\varpi}$ is the global precession rate of the disc that acts like a nonlinear eigenvalue of the problem.

The untwisted eccentric modes can be thought of as a superposition of two twisted eccentric waves which reflect off the disc boundaries or are potentially confined away from the disc boundary in a wave cavity (Lee et al., 2019a). In the presence of an excitation mechanism (such as an eccentric Lindblad resonance with an orbital companion, as occurs in superhump binaries, for example) we would expect a reflective disc to develop an untwisted eccentric mode. However, if the ingoing and outgoing travelling waves have different amplitudes, for instance if the boundaries are not perfectly reflecting or if the dissipation is strong enough, then the disc will develop a twist.

For an untwisted eccentric mode equation 5.40 is trivially satisfied. The solutions of equation 5.41 generally involve oscillations of the function $e(a)$, similar to waves on a string. It is conventional in celestial mechanics to regard e as a non-negative quantity and ϖ as being defined modulo 2π wherever $e \neq 0$. In the case of untwisted modes it is convenient to regard e as a signed quantity where $(-e, \varpi)$ is equivalent to $(e, \varpi + \pi)$.

When the Hamiltonian density for an untwisted disc is written as $H_a = H_a^\circ F$, where H_a° depends only on a, and F is a dimensionless function of e and $f = e + ae_a$, the modal equation can be written explicitly as

$$-\frac{(\omega - \omega_f) M_a}{H_a^\circ} \frac{na^2 e}{\sqrt{1-e^2}} = \frac{\partial F}{\partial e} - ae_a \frac{\partial^2 F}{\partial e \partial f} - a(2e_a + ae_{aa}) \frac{\partial^2 F}{\partial f^2} - \frac{d \ln H_a^\circ}{d \ln a} \frac{\partial F}{\partial f} \quad . \tag{5.42}$$

To illustrate this theory consider the case of an isothermal 2D disc with rigid boundaries and a constant surface density, which was studied by Barker & Ogilvie (2016). Constant surface density means that $M_a \propto a$, and along with isothermality means that $H_a^\circ \propto a$. The boundary conditions are $e = 0$ at $a = r_{\rm in}$ and at $a = r_{\rm out} = 2r_{\rm in}$. We can solve ODE 5.42 using a shooting method, and we can set the amplitude of the mode by specifying the value of ae_a at the inner boundary. We obtain a sequence of modes with increasing number of nodes in the eigenfunction $e(a)$ (see Figure 5.2). The variation of the precession rate ω is shown in Figure 5.3. The precession is retrograde and increases with both amplitude and mode number.

As ae_a approaches its maximum possible value of 1, the mode becomes highly nonlinear with f being close to ± 1 for most values of a, so that we have $ae = \pm a + \text{const}$, with sharp transitions in between. This means that the mode structure consists of alternating regions in which either the pericentre or the apocentre are tightly bunched (see Figure 5.4). The limiting form of the lowest-order mode is given by

$$ae = \begin{cases} a - r_{\rm in}, & r_{\rm in} < a < \bar{r}, \\ -a + r_{\rm out}, & \bar{r} < a < r_{\rm out}, \end{cases} \tag{5.43}$$

where $\bar{r} = (r_{\rm in} + r_{\rm out})/2$. It has finite AMD (equal to $0.02318 M l_{\rm in}$, where $l_{\rm in} = \sqrt{GM_1 r_{\rm in}}$), which is the largest AMD that the system can support, and an unbounded Hamiltonian because F diverges as $f \to 1$. The limiting forms of the higher-order modes, which were not considered by Barker & Ogilvie (2016), have similar piecewise linear structures, but the positions of the break points cannot be determined in a straightforward way (except for very high order modes which will be covered in the next chapter).

We next consider a more realistic problem involving an extended 2D disc with $a_{\rm out}/a_{\rm in} = 100$ and free boundaries. We also set $\gamma = 2$ for this illustrative calculation, and in fact consider a polytrope of index 1. In this instance the geometric Hamiltonian of a 2D disc can be obtained in terms of elementary functions for a perfect gas with $\gamma = 1$ or $\gamma = 2$, which are expected to bracket the range of behaviour expected in a real system. We take $M_a/a \propto \varepsilon^\circ \propto T(a)/a$, where $T(a) = \tanh[(a - a_{\rm in})/w_{\rm in}] \tanh[(a_{\rm out} - a)/w_{\rm out}]$ is a tapering function that is close to 1 in most of the disc but declines smoothly to 0 at the inner and outer boundaries over lengthscales $w_{\rm in} = 0.01 a_{\rm in}$ and $w_{\rm out} = 0.01 a_{\rm out}$. Apart from the taper, this model corresponds in the circular limit to a surface density $\propto r^{-1}$ and a sound speed $\propto r^{-1/2}$. The free boundary conditions correspond to regularity conditions at $a = a_{\rm in}$ and $a = a_{\rm out}$, which are singular points of equation 5.42. For this model they take the form

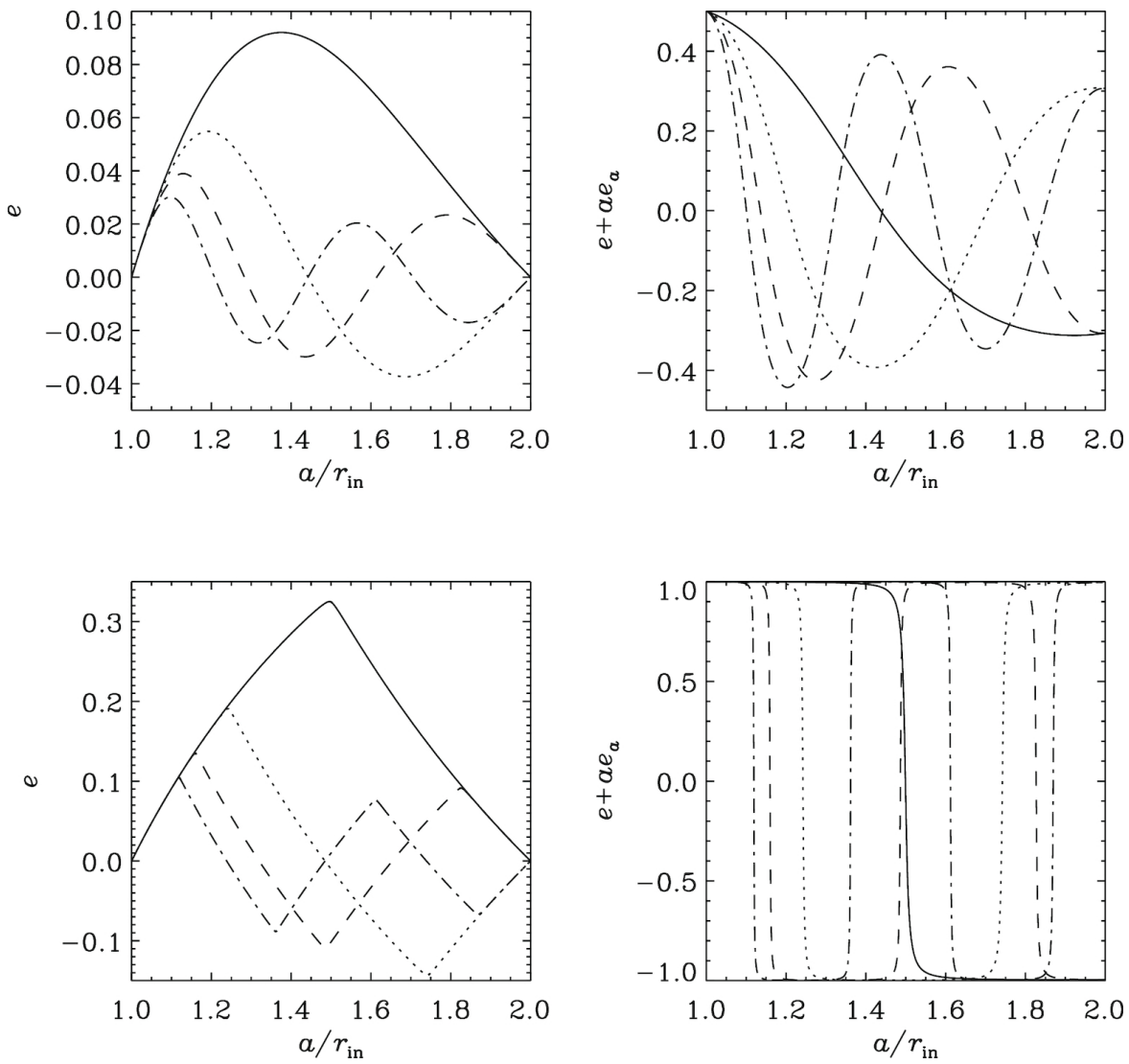


Fig. 5.2 Shape of the nonlinear eccentric mode for the simple model problem considered by Barker & Ogilvie (2016) in which a 2D isothermal disc is contained within rigid circular boundaries with a radius ratio of 2. The upper two panels show the profiles of e and $f = e + ae_a$ for the first four modes with moderately large amplitudes such that $ae_a = 0.5$ at the inner boundary. The lower panels show the same modes with maximal amplitude, such that $ae_a = 0.999$ at the inner boundary. Only the lowest order mode was computed by Barker & Ogilvie (2016).

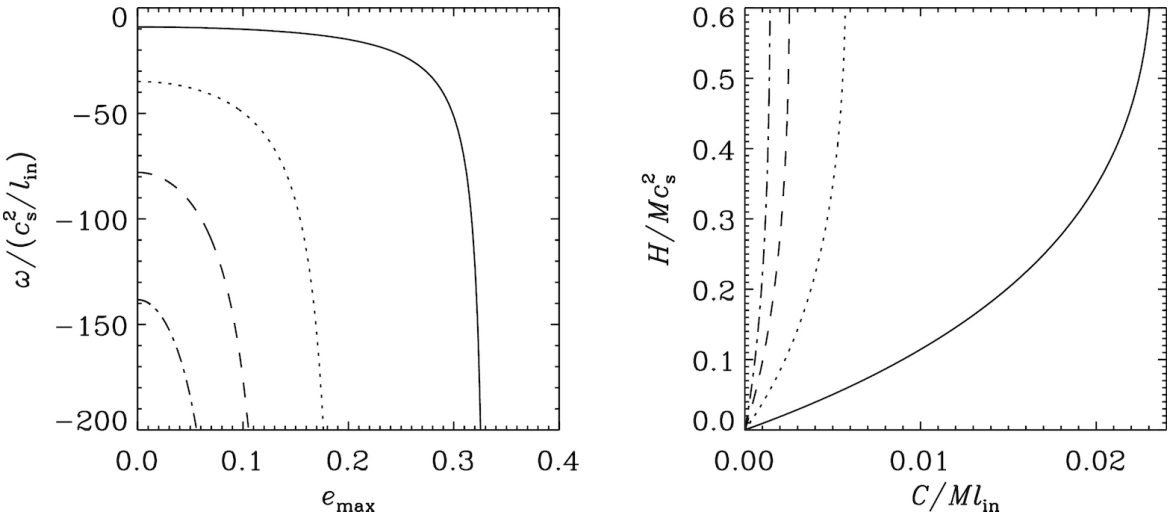


Fig. 5.3 Left: dependence of the precession frequency of the families of nonlinear eccentric modes, considered in Figure 5.2, on the mode amplitude, quantified by $e_{\max} = \max|e|$. Here $l_{\rm in} = \sqrt{GM_1 r_{\rm in}}$ is the specific angular momentum at the inner boundary. Right: Hamiltonian (energy) versus AMD for the same families. Note that each branch has a limiting value of C at which $H \to \infty$ because an orbital intersection is approached.

$$2a_{\rm in}\frac{dT}{da}\frac{\partial F}{\partial f} = \frac{\omega n a^2}{\varepsilon^\circ_{\rm in}}\frac{e}{\sqrt{1-e^2}}, \tag{5.44}$$

where $\varepsilon^\circ_{\rm in}$ is the value of ε° at $a = a_{\rm in}$ neglecting the taper (with an equivalent boundary condition at $a = a_{\rm out}$); this leads to a nonlinear relation between e and f at the boundaries.

The shape of the first four modes are shown in Figure 5.5, where the amplitudes have been chosen such that the maximum eccentricity attained in the disc is 0.5. Again the modes form a sequence in which the number of nodes in $e(a)$ increases by one. These modes are less confined than those of the previous problem; with free boundary conditions e does not need to vanish at the boundaries, and the larger disc means that the eccentricity gradient can avoid the nonlinear limit $f \to 1$. The precession rate again increases with amplitude (see Figure 5.6)and mode number; the numerical value of $-\omega$ in $\epsilon^\circ_{\rm in}/l^\circ_{\rm in}$ are much smaller than that found in the previous problem. This is because (i) the eccentricity gradients are more moderate, which reduces the precession rate; (ii) the average internal energy of the disc is lower than $\epsilon^\circ_{\rm in}$ because of the assumed temperature gradient; and (iii) the majority of the angular momentum is located at large semimajor axis.

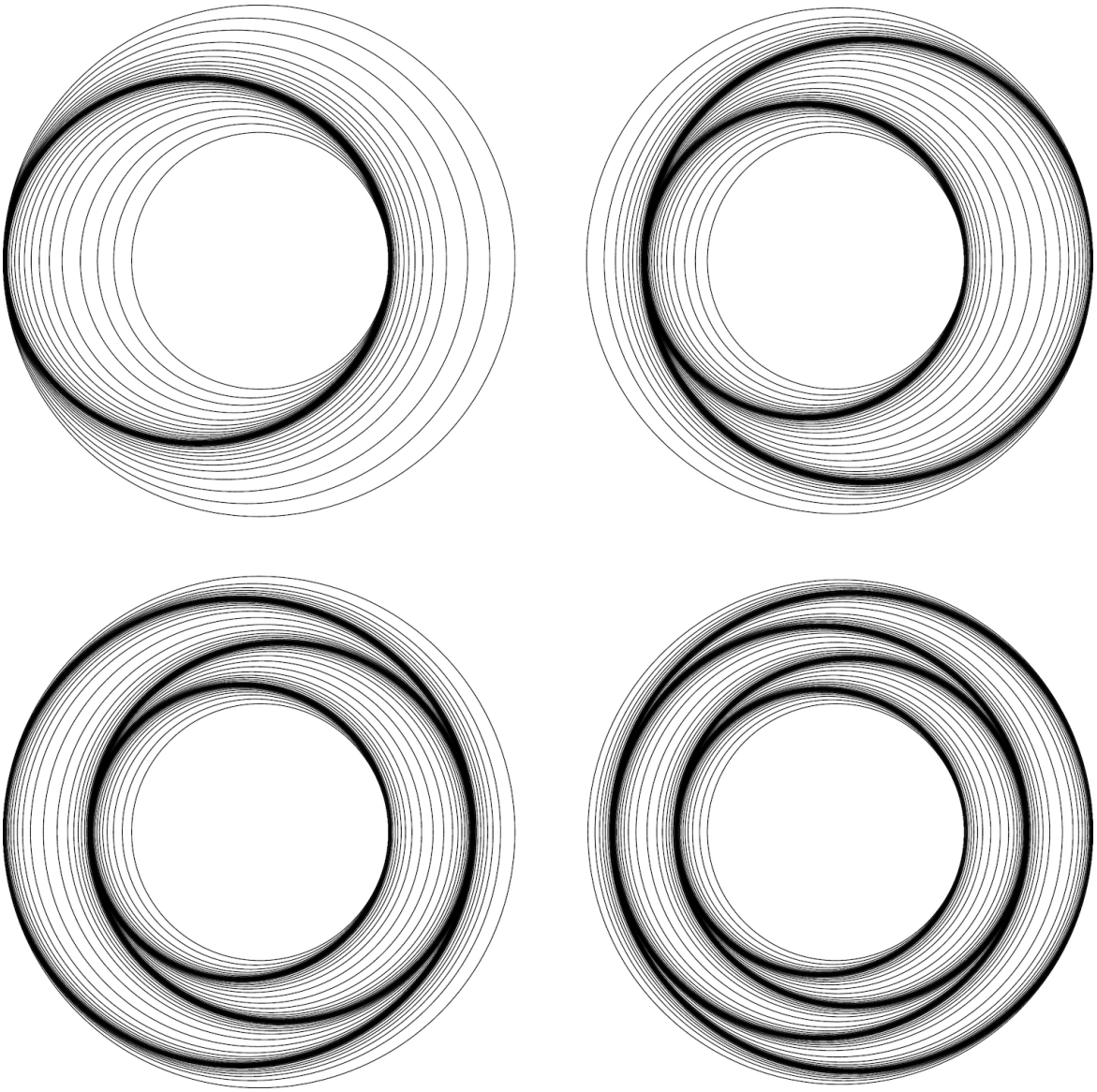


Fig. 5.4 Orbits in the nonlinear eccentric modes considered in Figure 5.3, for the first four modes with nearly maximal amplitudes such that $ae_a = 0.999$ at the inner boundary. The alternate bunching up of orbits at pericentre and apocentre will cause large surface density enhancements. This is the real source of surface density variations around eccentric orbits, and not the "traffic jam" effect postulated in the "apocentre hang time" argument.

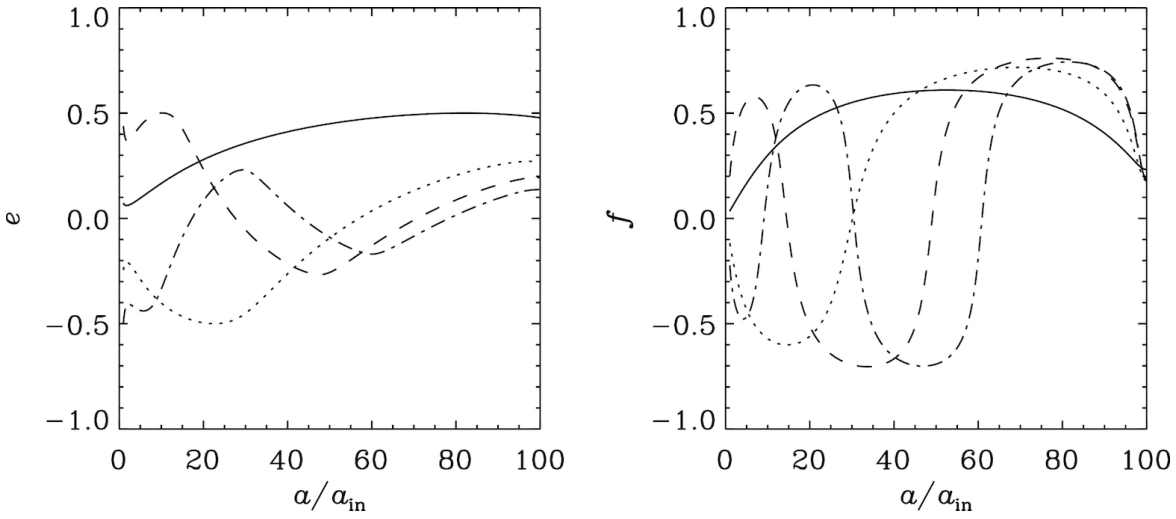


Fig. 5.5 Same as figure 5.2 but for the more realistic problem of an extended disc with $a_{\rm out}/a_{\rm in} = 100$ and free boundaries. The amplitude of the modes are chosen such that the maximum eccentricity attained is $e_{\rm max} = 0.5$.

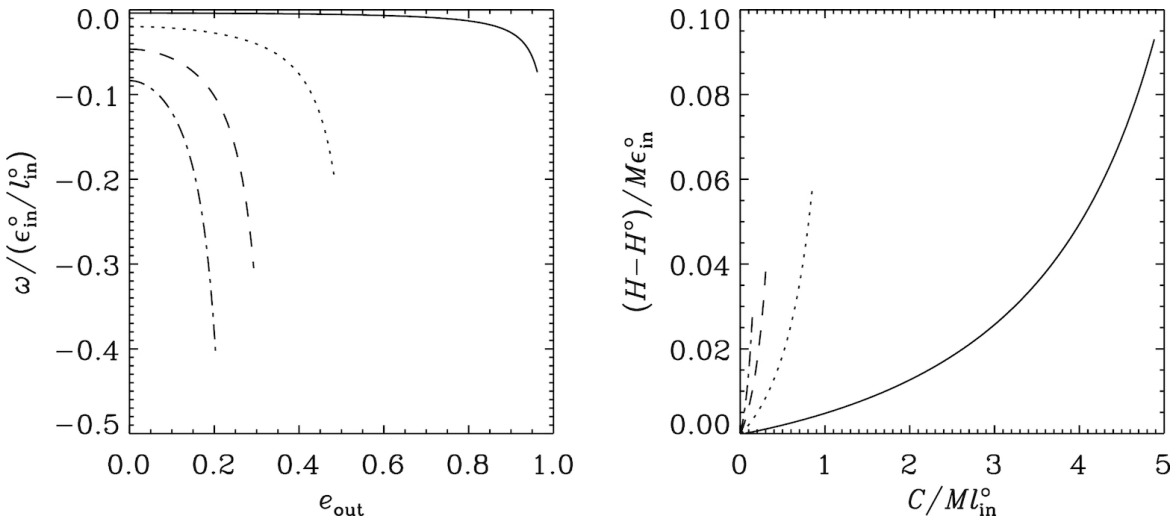


Fig. 5.6 Same as figure 5.3 but for the more realistic problem considered in figure 5.5. Here $\varepsilon^\circ_{\rm in}$ is the specific internal energy of a circular disc at the inner boundary, neglecting the taper, and $l^\circ_{\rm in} = \sqrt{GM_1 a_{\rm in}}$. Apart from the lowest-order mode the branches terminate when the maximum eccentricity attained in the disc $e_{\rm max} \to 1$.

5.7 Conclusion

In this Chapter we have developed a new framework for studying the nonlinear dynamics of thin eccentric astrophysical discs. In the case of an ideal, non-self-gravitating fluid, a thin eccentric disc consists of a continuous set of nested elliptic orbits that evolve slowly as a result of pressure and external perturbations. The semimajor axis of each orbit is conserved, as is the total angular momentum deficit of the disc, while the orbits undergo precession and exchanges of eccentricity. The nonlinear dynamics of this system can be derived by Whitham-averaging the affine disc model of Ogilvie (2018), when the disc orbits are confined to a plane (i.e. in the absence of a disc warp). From this we obtain a Lagrangian, and by a Legendre transform a Hamiltonian, describing the secular dynamics of the disc using either canonical (e.g. modified Delaunay) variables or the more familiar Keplerian orbital elements. In a 2D disc, without external forces, the Hamiltonian is just the orbit averaged internal energy. For the realistic case of a 3D disc, the vertical oscillation of the fluid columns in the affine model provide an additional contribution to the Hamiltonian which modifies the dynamics of the secular theory.

The simplest solutions of the theory are uniformly precessing nonlinear eccentric modes, and we have presented numerical examples of branches of model solutions in illustrative situations up to their maximum possible amplitudes.

In the next chapter we shall explore a nonlinear limit of the theory presented here which is amenable to analytical treatment. The theory presented here is less complete than that presented by Ogilvie (2001) which includes non-ideal effects such as a turbulent stress model and radiative cooling. In principle, however, this can be incorporated into the above theory by starting from a pseudo-Lagrangian of an affine disc model which incorporates these effects. Then the generalisation of the Whitham averaging method by Jimenez & Whitham (1976) to nonideal effects can be used to obtain the secular theory. At present the affine disc model has not been extended to include non-ideal effects so this generalisation must be left to future work.

Chapter 6

The Short Wavelength Limit

6.1 Introduction

WKB theory has been used by various authors to describe the behaviour of short-wavelength linear eccentric modes in the inner regions of discs. Examples of this approach are the work of Tremaine (2001) on galactic "slow modes", Ferreira & Ogilvie (2009) on eccentricity in the inner regions of black hole accretion discs and most recently Lee et al. (2018) who looked at how the relative importance of pressure and self-gravity affects linear eccentric modes in discs. Of particular relevance to this chapter, the theory of Ferreira & Ogilvie (2009) predicts its own breakdown - with eccentric waves in the inner regions of discs around black holes attaining very strong eccentricity gradients highlighting the need for a nonlinear theory to describe this regime. Ferreira & Ogilvie (2008) showed that the eccentricity gradients are important in determining the growth rates of inertial waves via a parametric type instability, which have been proposed as the origin of high frequency quasi-periodic oscillations (QPOs) (Kato, 2008; Ferreira & Ogilvie, 2008, 2009; Dewberry et al., 2018, 2019).

The extension of the short-wavelength limit into the nonlinear regime was first done by Lee & Goodman (1999) who used the theory of Whitham (1965) to derive the dispersion relation for a single-armed spiral density wave in the tight-winding limit, incorporating pressure and self-gravity. They found, for a range of surface density and sound speed profiles, that a tightly wound trailing wave tends to become more nonlinear as it propagates inwards.

In Chapter 5 we showed that the nonlinear equations of eccentric discs can be derived from a variational principle. This nonlinear secular theory describes the slow evolution of nested ellipses due to pressure forces in a thin disc. In the 3D theory the Hamiltonian is proportional to the orbit averaged internal energy of the disc.

In this chapter we apply the theory of Whitham (1965) to determine the behaviour of eccentric waves in the short-wavelength limit. This will be a short wavelength limit of the theory developed in Chapter 5 and thus has all the assumptions/approximations made in deriving that theory (discussed in Section 5.2), in addition to the short wavelength assumption that will be made in this Chapter. This theory has much in common with Lee & Goodman (1999). The main advantage of our approach over theirs is that the short-wavelength limit, derived here, is much less restrictive than their tight winding assumption, allowing for both twisted and untwisted discs. In this chapter we focus on two limiting cases, that of untwisted discs (corresponding to standing waves) and highly twisted discs with a tightly wound eccentric travelling wave. In principle both can be obtained as separate limits of a more general short-wavelength theory, but we make no attempt to derive it in this thesis. Our treatment of the pressure forces is more general than Lee & Goodman (1999), in particular allowing for the inclusion of 3D effects.

6.2 Nonlinear theory of eccentric discs

As shown in Chapter 5 the ideal eccentric disc equations can be obtained from the following Lagrangian:

$$L = \int \left(M_a n a^2 \varpi_t \sqrt{1-e^2} - H_a \right) da. \tag{6.1}$$

This averaged Lagrangian (in the sense of Whitham (1965)) was obtained by averaging the affine disc Lagrangian of Ogilvie (2018) over the orbital timescale. In the following we will be performing a second Whitham (1965) averaging procedure over an intermediate time/lengthscale, which can loosely be thought of as an averaging over the disc precession timescale. It is convenient to split the Lagrangian as follows: $L = L_T + L_M$, where

$$L_T = \int M_a n a^2 \tilde{\varpi}_t \sqrt{1-e^2}\, da, \tag{6.2}$$

$$L_M = \int \left[M_a n a^2 \left(\omega \sqrt{1-e^2} + \int \frac{e \omega_f(a,e)}{\sqrt{1-e^2}}\, de \right) - H_a^\circ F \right] da, \tag{6.3}$$

where we have introduced $\tilde{\varpi} := \varpi - \omega t$. L_M is the Lagrangian for an eccentric wave which is steady in a frame precessing at rate ω. For an untwisted disc L_M is the Lagrangian describing an eccentric mode. L_T contain the terms responsible for time

evolution (in the precessing frame). In this Chapter we shall neglect L_T by setting $\tilde{\varpi}_t = 0$; we thus only consider eccentric waves which are steady in a frame rotating with angular velocity ω.

As in Chapter 5 the precessional forces coming from H_a^f can be expressed in terms of the precession rate $\omega_f(a,e)$ of a test particle,

$$M_a \omega_f(a,e) := -\frac{\sqrt{1-e^2}}{na^2 e}\frac{\partial H_a^f}{\partial e} \quad , \tag{6.4}$$

which can be rearranged for H_a^f,

$$H_a^f = -M_a n a^2 \int \frac{e\omega_f(a,e)}{\sqrt{1-e^2}}\, de \quad . \tag{6.5}$$

6.3 Short-wavelength nonlinear eccentric waves using an averaged Lagrangian

We now consider the situation where the precessional (or test particle) forces are strong relative to the pressure forces. Defining

$$N^2 = \frac{M_a n a^2(\omega_f(a,0)-\omega)}{H_a^\circ} \quad , \tag{6.6}$$

which can be understood as the ratio of these precessional effects (as captured by $\omega_f(a,0)-\omega$) to pressure forces, this situation corresponds to the limit $N^2 \gg 1$. This limit can come about either because of a prograde externally induced precession ($\omega_f(a,0) > 0$, e.g. due to general relativity or an orbital companion), or from a retrograde free precession of the disc ($\omega < 0$). [1]

The method of Whitham (1965) is applicable to nonlinear waves propagating on a background varying on a much longer lengthscale than the wavelength of the wave (it is also applicable to temporal variations but this is not needed here). A separation of scales occurs, as over one cycle, the wave sees a constant background, with the variations in the background medium only having an effect over many wave cycles. The wave amplitude is modulated on the longer lengthscale by this variation in the background medium.

In order to use the theory of Whitham (1965) we introduce a rapidly varying phase variable $\varphi = \varphi(a)$ which satisfies $a\varphi_a = N(a)k(a)$ with $k(a) = O(1)$ a rescaled

[1] The correspondence of $N^2 \gg 1$ to the short wavelength limit can be seen in linear theory. For example, the linear, local dispersion relation for an isothermal disc is $(k_r r)^2 = 2N^2$ where k_r is the radial wavenumber. As we shall show, strongly nonlinear effects can break this correspondence.

dimensionless wavenumber. We can then introduce a rescaled variable $\tilde{e} = O(1)$ such that

$$e = \frac{\tilde{e}(\varphi, a)}{N(a)k(a)} \quad . \tag{6.7}$$

We consider two situations where this limit is applicable. The first is an untwisted disc where the eccentricity oscillates on a short lengthscale; in this case the nonlinearity is

$$\begin{aligned} q &= a\frac{de}{da} \\ &= \tilde{e}_\varphi(\varphi, a) + a\frac{\partial}{\partial a}\left(\frac{\tilde{e}(\varphi, a)}{N(a)k(a)}\right) \\ &\approx \tilde{e}_\varphi(\varphi, a) \quad . \end{aligned} \tag{6.8}$$

The second is a highly twisted disc, with a tightly wound eccentric travelling wave, where the eccentricity is slowly varying and the nonlinearity is

$$\begin{aligned} q &= a\frac{de}{da} + ae\frac{d\varpi}{da} \\ &= \tilde{e}(a)\varpi_\varphi(\varphi, a) + a\frac{\tilde{e}(\varphi, a)}{N(a)k(a)}\frac{\partial\varpi(\varphi, a)}{\partial a} + a\frac{\partial}{\partial a}\left(\frac{\tilde{e}(a)}{N(a)k(a)}\right) \\ &\approx \tilde{e}(a)\varpi_\varphi(\varphi, a) \quad . \end{aligned} \tag{6.9}$$

In both cases the nonlinearity is order unity.

Substituting this into the Lagrangian and keeping only terms of the lowest order in N^{-1}, the Lagrangian becomes

$$\begin{aligned} L &\approx \int H_a^\circ \hat{\mathcal{L}}\, da \\ &= \int H_a^\circ \left(\frac{1}{2}\frac{\tilde{e}^2}{k^2} - F(q)\right) da \quad , \end{aligned} \tag{6.10}$$

where we have neglected a term $(M_a n a^2 \omega)$ which is a function only of a as this has no effect on the dynamics. We have separated L_M into a dimensionless part $\hat{\mathcal{L}}$ which controls the small scale dynamics and a dimensional part H_a° which is important in determining how the solution varies on the length scale of the disc. At lowest order, the geometric part of the Hamiltonian, defined in Chapter 5, $F(e, q, \alpha) \approx F(q)$. The

form of $F(q)$ depends on the ratio of specific heats and whether the disc considered includes the effects of the breathing mode. Example forms for $F(q)$ are

$$F^{(\mathrm{lin})}(q) = \frac{\gamma+1}{2(\gamma-1)} + \frac{(2\gamma-1)}{4\gamma}q^2 \quad , \tag{6.11}$$

$$F^{(\mathrm{2D})}(q) = \frac{1}{(\gamma-1)}\frac{1}{2\pi}\int_0^{2\pi}(1-q\cos E)^{-(\gamma-1)}dE \quad , \tag{6.12}$$

$$F^{(\mathrm{3D})}(q) = \frac{(\gamma+1)}{2(\gamma-1)}\frac{1}{2\pi}\int_0^{2\pi}[h(q,\cos E)(1-q\cos E)]^{-(\gamma-1)}dE \quad , \tag{6.13}$$

$$F^{(\mathrm{iso})}(q) = \ln\left[2q^{-2}(1-\sqrt{1-q^2})\right] \quad , \tag{6.14}$$

where $F^{(\mathrm{lin})}$ is for a 3D adiabatic disc in linear theory, $F^{(\mathrm{2D})}$ and $F^{(\mathrm{3D})}$ are for the 2D and 3D adiabatic discs and $F^{(\mathrm{iso})}$ is for an isothermal disc. Here $h(q,\cos E)$ is the dimensionless scale height (equal to the scale height divided by the orbit averaged scale height); this varies around the orbit owing to the presence of the disc breathing mode which is forced by the lateral compression of the fluid around each orbit (Ogilvie, 2001, 2008; Ogilvie & Barker, 2014). Some examples of $F(q)$ are plotted in Figure 6.1.

At lowest order there is no difference between the 3D and 2D isothermal discs, as the breathing mode is independent of q owing to the vertical specific enthalpy gradient being independent of horizontal compression.

6.4 Short-lengthscale dynamics

With the preliminaries out of the way we now turn to solving the short-lengthscale dynamics of the Lagrangian given by equation 6.10. This determines how $\tilde{e}$, q and ϖ depend on the wave phase φ, which is needed to perform the phase average.

For the tightly wound wave $\tilde{e}$ and q are both integrals of motion so the short lengthscale dynamics are trivial with $\tilde{e}$, q constant on the short lengthscale and $\varpi \propto \varphi$. The Whitham (1965) average of the Lagrangian is trivial and is exactly equation 6.10, where $\tilde{e}$ and $q = q_+$ are regarded as varying only on the long lengthscale and $\varpi_\varphi \tilde{e} = q$,

$$\langle L\rangle = \int H_a^\circ \left(\frac{1}{2}\frac{\tilde{e}^2}{k^2} - F(q_+)\right) da \quad . \tag{6.15}$$

This averaged Lagrangian has the following gauge symmetry:

$$k \to \lambda(a)k \quad , \qquad \tilde{e} \to \lambda(a)\tilde{e}, \tag{6.16}$$

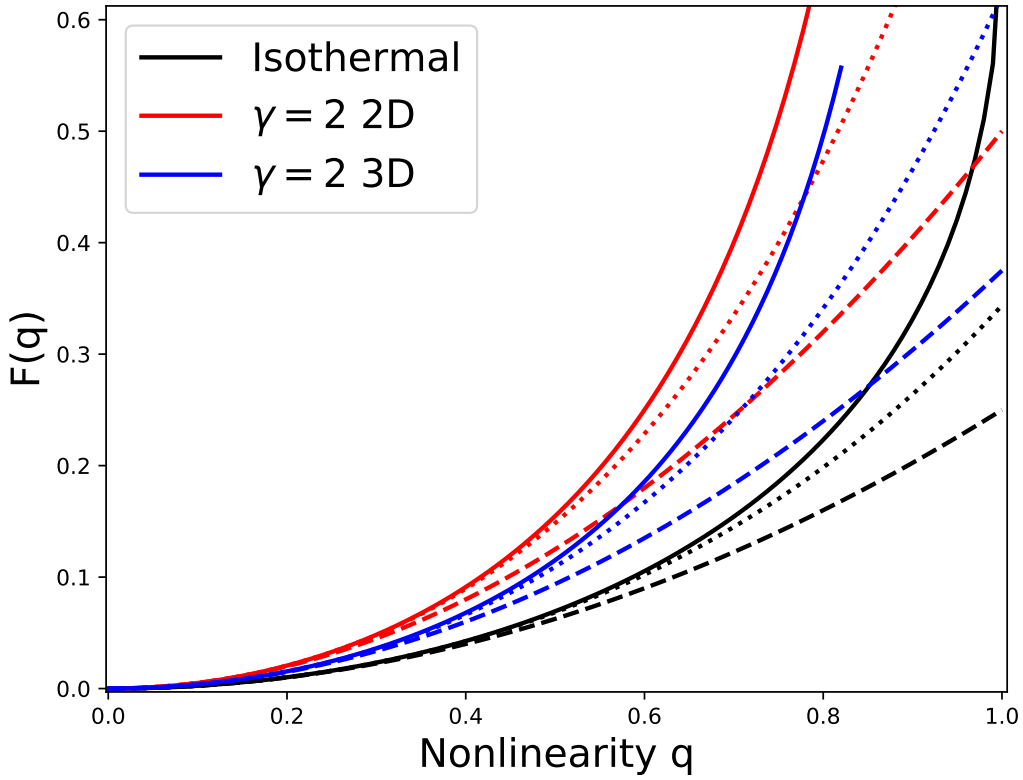


Fig. 6.1 Geometric part of the Hamiltonian in the short-wavelength limit. In each case the unimportant constant term has been subtracted off. Most forms of F of interest are bracketed by the behaviour of these functions. In particular we expect all should have the property of being monotonically increasing even functions which diverge at $q = 1$. The dashed lines are series expansion in q including terms up to $O(q^2)$ (which corresponds to linear theory). The dotted lines include an $O(q^4)$ terms which are the lowest order nonlinear terms.

where $\lambda(a)$ is an arbitrary function of a. This is directly related to $\tilde{e}$ and q not varying on the short lengthscale, meaning the phase φ is arbitrary as the solution is not oscillatory. This introduces a freedom to redefine the wave phase. It is useful to gauge fix out this symmetry by setting $\tilde{e} = q_+$. Upon adopting this gauge the Lagrangian becomes

$$\langle L \rangle = \int H_a^\circ \langle \hat{\mathcal{L}} \rangle \, da \quad , \tag{6.17}$$

with

$$\langle \hat{\mathcal{L}} \rangle = \frac{1}{2} \frac{q_+^2}{k^2} - F(q_+). \tag{6.18}$$

For the untwisted disc it is useful to reformulate the Lagrangian viewing q as the coordinate and obtain a Hamiltonian in terms of q and its conjugate momentum π_q as we expect this to be quadratic in π_q. To do so we vary $\hat{\mathcal{L}}$ with respect to $\tilde{e}$ to obtain the equation of motion:

$$\tilde{e} = -k^2 F''(\tilde{e}_\varphi) \tilde{e}_{\varphi\varphi} \tag{6.19}$$

and thus rewrite the nondimensional Lagrangian as

$$\hat{\mathcal{L}} = \frac{k^2}{2} (F''(q))^2 q_\varphi^2 - F(q), \tag{6.20}$$

where we have made use of the fact that $q = \tilde{e}_\varphi$. From this we obtain π_q:

$$\pi_q = \frac{\partial \hat{\mathcal{L}}}{\partial q_\varphi} = k^2 (F''(q))^2 q_\varphi \quad , \tag{6.21}$$

and thus obtain the Hamiltonian

$$H = \frac{\pi_q^2}{2k^2 (F''(q))^2} + F(q) \quad . \tag{6.22}$$

As we are treating φ as a phase variable governing the short length scale behaviour we require $\tilde{e}$ to be periodic in φ with period 2π. Unlike the tightly wound waves, the phase variable is physically meaningful and cannot be rescaled, although we are free to translate it by an arbitrary function of a. The dynamics on the short length scale can be entirely understood from the contours of constant Hamiltonian in the $(q, \pi_q/k)$ plane (Figure 6.2) where they circulate around the fixed point at the origin.

To solve the short lengthscale dynamics we follow Whitham (1965) in setting $H = A$, where A can be viewed as the amplitude of the nonlinear wave. If F is an

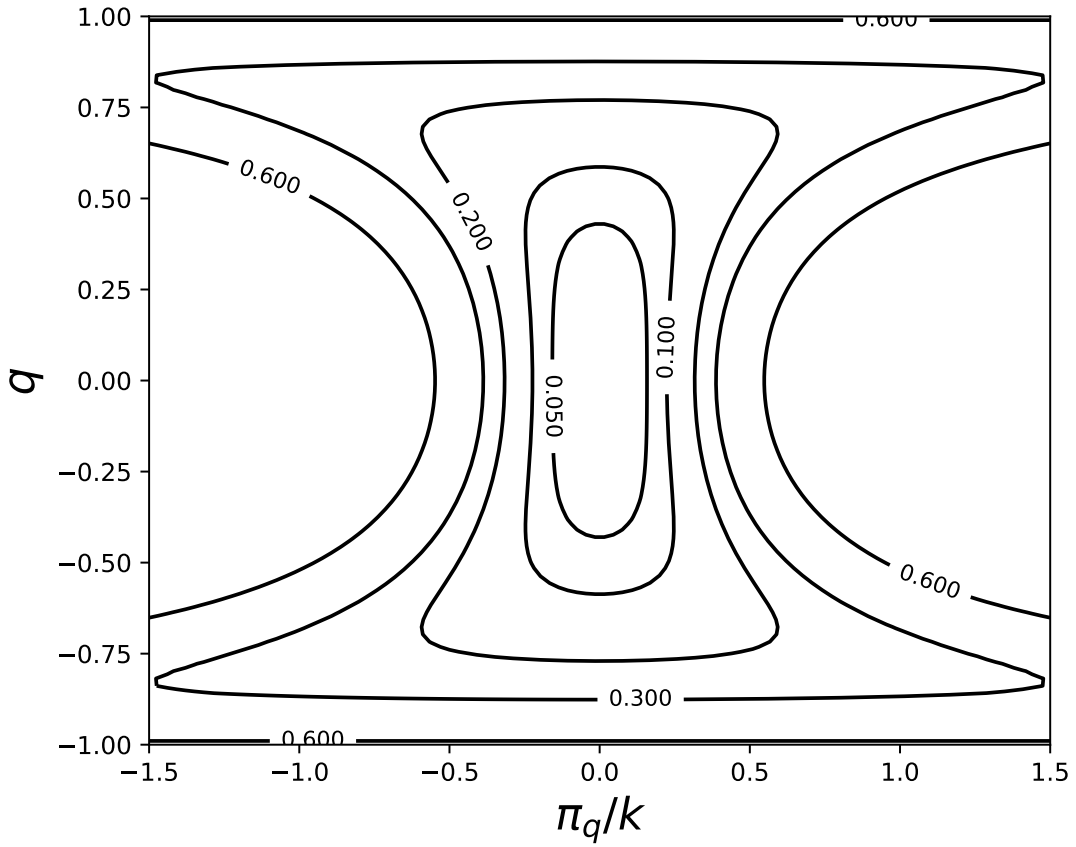


Fig. 6.2 Phase space of the small scale nonlinear oscillator, which occurs in the untwisted case, for $F = F^{(\mathrm{iso})}$ showing the contours of the Hamiltonian in terms of q and π_q/k.

even increasing function of $|q|$ then the maximum value of q is attained when $\pi_q = 0$. Denoting this maximum value q_+ such that $|q| \leq q_+$ then we have $A = F(q_+)$ so that

$$F(q_+) = \frac{\pi_q^2}{2k^2(F''(q))^2} + F(q) \quad . \tag{6.23}$$

This can be rearranged to give

$$q_\varphi = \pm \frac{\sqrt{2}}{k} \frac{[F(q_+) - F(q)]^{1/2}}{F''(q)} \quad , \tag{6.24}$$

and directly integrated to obtain an implicit solution for the short-lengthscale dynamics

$$\pm \frac{\varphi}{\sqrt{2}k} = \int_{-q_+}^{q} \frac{F''(q)}{2\left[F(q_+) - F(q)\right]^{1/2}} \, dq =: I\left(q, q_+\right) \quad , \tag{6.25}$$

where it is useful to define an incomplete integral I (Equation A.1). Here the sign denotes the increasing and decreasing branches of the solution and is necessary as $\varphi(q)$ is multivalued over one period.

k can be determined by imposing 2π periodicity on $q(\varphi)$. With this condition we obtain the nonlinear dispersion relation for the rescaled wavenumber

$$k = \frac{\pi}{\sqrt{2} I\left(q_+, q_+\right)} = \frac{\pi}{\sqrt{2} I\left(q_+\right)} \quad , \tag{6.26}$$

where we have defined the complete integral $I(q_+) = I(q_+, q_+)$, allowing us to write

$$\varphi = \pm \pi \frac{I\left(q, q_+\right)}{I\left(q_+\right)} \quad . \tag{6.27}$$

Figure 6.3 shows the variation of the mode on the short lengthscale for an isothermal disc with different values of q_+. In the linear limit the eccentricity and nonlinearity are sine and cosine waves; as the nonlinear limit is approached they approach a limiting slope resulting in e taking the form of a triangle wave, while q tends to a square wave. This is similar to the nonlinear global modes calculated by Barker & Ogilvie (2016) and in Chapter 5 using the more general theory. The advantage here is that it is easy to obtain the locations of the zeros and peaks of these modes which is not straightforward to obtain in general.

As per Whitham (1965) we obtain an expression for π_q in terms of q and A,

$$\pi_q = \pm \sqrt{2} k F''(q) \left[A - F(q)\right]^{1/2} . \tag{6.28}$$

From this we obtain an expression for the (nondimensional) averaged Lagrangian,

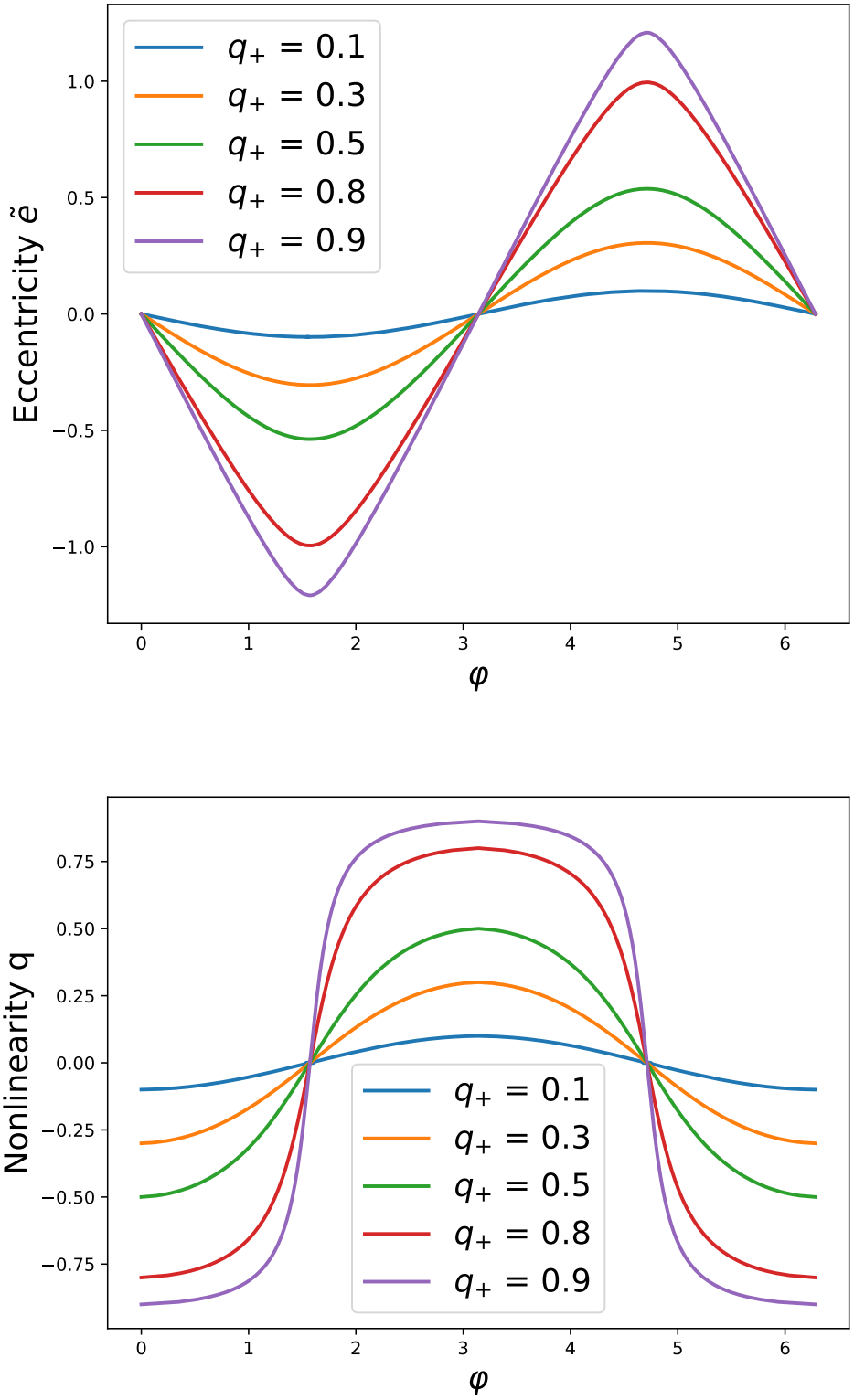


Fig. 6.3 Variation of the eccentric mode on the short lengthscale (φ) for different maximum nonlinearities (q_+) in an isothermal disc. Top: (rescaled) eccentricity, Bottom: nonlinearity q. The eccentricity tends to a triangle wave and the nonlinearity to a square wave in the nonlinear limit $q_+ \to 1$.

$$\langle\hat{\mathcal{L}}\rangle = \frac{1}{2\pi}\sqrt{2}k \oint F''(q)[F(q_+) - F(q)]^{1/2}dq - F(q_+), \tag{6.29}$$

where we have made use of $A = F(q_+)$.

An alternative form of the averaged Lagrangian can be obtained for both the untwisted and tightly wound discs by directly averaging Equation 6.10 (with $\varphi_t = 0$). In this case the (nondimensional) averaged Lagrangian is

$$\langle\hat{\mathcal{L}}\rangle = \frac{1}{2}\frac{\langle\tilde{e}^2\rangle}{k^2} - \langle F(q)\rangle \quad ; \tag{6.30}$$

this form of the averaged Lagrangian is not that useful for determining the dispersion relation as, for the untwisted disc, it's not obvious how to carry out the average $\langle F(q)\rangle$. It is however useful for determining how eccentricity varies on the long lengthscale.

6.5 Modulation of the eccentric wave on the disc length scale

We now consider how the eccentricity and nonlinearity vary on the long lengthscale for a steady disc. At this stage it is useful to introduce certain functionals of $F(q)$ which are independent of the details of the disc model. One of these, $I(q_+)$, was already introduced above for the case of an untwisted disc. In general we shall need two such functionals, $I(q_+)$ and $J(q_+)$, which for an untwisted disc are given by

$$I(q_+) := \int_0^{q_+} F''(q)[F(q_+) - F(q)]^{-1/2}dq \quad , \tag{6.31}$$

$$J(q_+) := \int_0^{q_+} F''(q)[F(q_+) - F(q)]^{1/2}dq \quad . \tag{6.32}$$

For a tightly wound disc these are instead given by

$$I(q_+) := \frac{\pi}{\sqrt{2}}\sqrt{\frac{F'(q_+)}{q_+}} \quad , \tag{6.33}$$

$$J(q_+) := \frac{\pi}{2\sqrt{2}}q_+^{1/2}F'(q_+)^{3/2} \quad . \tag{6.34}$$

In what follows we shall assume that $I(q_+)$, $J(q_+)$ are both increasing functions of q_+ which diverge as $q_+ \to 1$. This assumption is proven for physically reasonable $F(q)$ in Appendix A.

6.5.1 Flux conservation and nonlinear dispersion relation in a steady eccentric disc

For both our untwisted and tightly wound discs we use q_+ as a measure of the wave amplitude. For the tightly wound disc this corresponds to the nonlinearity q of the wave, while for the untwisted disc this is the amplitude of q as it varies on the short lengthscale. To obtain the dispersion relation we vary the averaged Lagrangian with respect to q_+,

$$\frac{\delta\langle L\rangle}{\delta q_+} = H_a^\circ \frac{\partial\langle\hat{\mathcal{L}}\rangle}{\partial q_+} = 0, \tag{6.35}$$

which for the untwisted disc is

$$\frac{\delta\langle L\rangle}{\delta q_+} = H_a^\circ F'(q_+)\left(\frac{\sqrt{2}}{\pi} I(q_+) - 1\right) = 0 \quad , \tag{6.36}$$

while for the tightly wound wave it is

$$\frac{\delta\langle L\rangle}{\delta q_+} = H_a^\circ F'(q_+)\left(\frac{\pi^2}{2k^2 I^2(q_+)} - 1\right) = 0 \quad , \tag{6.37}$$

both of which, after rearranging and rescaling, become

$$a\varphi_a = \frac{\pi N}{\sqrt{2} I(q_+)} \quad . \tag{6.38}$$

For both untwisted and tightly wound waves, varying the averaged Lagrangian with respect to φ gives us a flux conservation law,

$$\frac{\delta\langle L\rangle}{\delta\varphi} = -\frac{\partial}{\partial a}\left(\frac{aH_a^\circ}{N}\frac{\partial\langle\hat{\mathcal{L}}\rangle}{\partial k}\right) = -\frac{\partial}{\partial a}\left(\frac{aH_a^\circ}{N}\frac{2\sqrt{2}}{\pi} J(q_+)\right) = 0, \tag{6.39}$$

which can be rearranged to obtain a conserved flux,

$$\frac{2\sqrt{2} a H_a^\circ}{\pi N} J(q_+) = \text{const} \quad . \tag{6.40}$$

An alternative form of the flux conservation equation can be obtained varying equation 6.30 with respect to φ, from which we obtain an equivalent conserved flux,

$$\frac{aH_a^\circ N^2}{a\varphi_a}\langle e^2\rangle = \text{const} \quad . \tag{6.41}$$

Combining this with the dispersion relation we obtain an equation for the eccentricity envelope,

$$\langle e^2\rangle \propto \frac{1}{aH_a^\circ N I(q_+)} \quad . \tag{6.42}$$

In the linear limit (using $F = F^{\text{lin}}$) we can evaluate I and J and obtain

$$I(q_+) = \frac{\pi}{2}\sqrt{\frac{2\gamma-1}{\gamma}} \quad , \tag{6.43}$$

$$J(q_+) = s\frac{\pi}{16}\left(\frac{2\gamma-1}{\gamma}\right)^{3/2} q_+^2 \quad . \tag{6.44}$$

where $s = 1$ for an untwisted disc and $s = 2$ for the tightly wound wave. This means for a given wave flux J the untwisted eccentric mode will have twice the amplitude of the tightly wound wave. This can be understood by the fact that in the linear limit the untwisted disc is a sum of two tightly wound waves propagating in opposite directions.

In the linear limit the solution to the short lengthscale oscillator has the form $\mathcal{E} = \mathcal{E}_0 \exp(i\varphi)$; this gives a solution in the linear limit of

$$\mathcal{E} \propto (aH_a^\circ N)^{-1/2} \exp\left(\int \sqrt{\frac{2\gamma}{2\gamma-1}}\frac{N}{a}\,da\right) \quad . \tag{6.45}$$

This matches the known WKB solution to the 3D eccentric disc equations (equivalent to equation 2.34), showing that the WKB limit of linear theory and the linear limit of the short wavelength theory for eccentric discs are the same.

6.6 Conditions for wave steepening

We now turn to the question of under what circumstances do nonlinearity q and eccentricity get amplified by the large scale disc properties. In effect, how does the disc behaviour control the long length scale envelope of the solution. Can a small disturbance in the outer region of the disc be focused and become highly nonlinear in the inner region?

We show in Appendix A.1 that the integrals $I(q_+)$, $J(q_+)$ are increasing functions of q_+. Using Equation 6.40 we can obtain a condition on the amplification of q_+ with decreasing a,

$$\frac{\partial q_+}{\partial a} < 0 \quad \text{if and only if} \quad \frac{\partial}{\partial a}\left(\frac{aH_a^\circ}{N}\right) > 0 \tag{6.46}$$

The condition on $\langle e^2 \rangle$ to increase can be obtained from Equation 6.42, from which we obtain the condition

$$\frac{\partial}{\partial a}\left(aH_a^\circ N\langle e^2\rangle\right) < 0 \quad \text{if and only if} \quad \frac{\partial q_+}{\partial a} > 0 \quad . \tag{6.47}$$

Similarly we have a condition on $a\varphi_a$:

$$\frac{\partial}{\partial a}\left(\frac{a\varphi_a}{N}\right) > 0 \quad \text{if and only if} \quad \frac{\partial q_+}{\partial a} > 0 \quad . \tag{6.48}$$

Using Equations 6.38 and 6.42 we can show that

$$\frac{a\varphi_a}{a\varphi_a|_{\text{lin}}} = \frac{I(0)}{I(q_+)} \leq 1 \quad , \tag{6.49}$$

$$\frac{\langle e^2\rangle}{\langle e^2_{\text{lin}}\rangle} = \frac{I(0)}{I(q_+)} \leq 1 \quad , \tag{6.50}$$

where the subscript $_{\text{lin}}$ denotes the prediction from linear theory. So in general we expect the nonlinear eccentric waves to behave like the linear eccentric waves, however amplification of the eccentricity and wavenumber of the eccentric wave due to wave focusing will be suppressed due to nonlinear effects. As such we can conclude that the wavenumber and (absolute) eccentricity of a nonlinear eccentric wave is bounded from above by that obtained from linear theory in the direction in which q_+ is increasing.

Physically this can be understood as the nonlinear effect enhancing the pressure forces from the eccentric wave in the disc so that the eccentric wave does not need to attain as high an amplitude/wavenumber as in linear theory in order to balance the precessional effects captured by $\omega_f - \omega$.

If we now consider a disc with powerlaw profiles for P°, Σ° and $\omega_f - \omega$; with $P^\circ \propto a^r$, $\Sigma^\circ \propto a^\sigma$ and $\omega_f - \omega \propto a^s$. The last of these corresponds to the non-Keplerian forces in the disc, in particular the lowest order contribution from GR is given by $s = -5/2$ and from a quadrupole term (such as that present in discs around binaries or rotating stars) $s = -7/2$.

With these forms for the disc dependent quantities the linear theory predict that $q_+|_{\text{lin}} \propto a^\lambda$ and $\langle e^2_{\text{lin}}\rangle \propto a^\mu$ where the power law exponents are given by

$$\lambda = \frac{1}{4}\left(s + \sigma - 3r - \frac{7}{2}\right) \quad , \tag{6.51}$$

$$\mu = -\frac{1}{2}\left(r + \sigma + s + \frac{9}{2}\right) \quad . \tag{6.52}$$

For the nonlinear theory we obtain the following conditions for the behaviour of the eccentric wave,

$$\frac{\partial q_+}{\partial a} < 0 \quad \text{if and only if} \quad \lambda < 0 \quad , \tag{6.53}$$

which is the same as that from linear WKB theory. The conditions on the behaviour of $\langle e^2 \rangle$ are

$$\lambda > 0 \quad \text{and} \quad \mu < 0 \implies \frac{\partial}{\partial a}\langle e^2 \rangle < 0 \quad , \tag{6.54}$$

$$\lambda < 0 \quad \text{and} \quad \mu > 0 \implies \frac{\partial}{\partial a}\langle e^2 \rangle > 0 \quad . \tag{6.55}$$

These differ subtly from the condition derived from linear theory which is

$$\frac{\partial}{\partial a}\langle e^2 \rangle < 0 \quad \text{if and only if} \quad \mu < 0 \quad . \tag{6.56}$$

This difference comes about because in nonlinear theory the sign of λ controls whether nonlinear effects will compete with or enhance the forcing terms present in linear theory. In addition to conditions (6.54) and (6.55) there are two other possible combinations of signs for λ and μ.

1. $\lambda < 0 \quad \text{and} \quad \mu < 0 \implies \frac{\partial}{\partial a}\langle e^2 \rangle$ bounded from below by linear theory.
2. $\lambda > 0 \quad \text{and} \quad \mu > 0 \implies \frac{\partial}{\partial a}\langle e^2 \rangle$ bounded from above by linear theory.

In both these cases we can make no definite conclusions about the sign of $\frac{\partial}{\partial a}\langle e^2 \rangle$ as the nonlinear effects compete with the linear terms. Instead $\frac{\partial}{\partial a}\langle e^2 \rangle$ has the same sign as linear theory for low q_+ and changes sign as q_+ increases.

In most regions of the disc (apart from the disc inner edge/gap where the disc is truncated) we expect $r \leq 0$ and $\sigma \leq 0$. In addition we expect the steepest precessional frequency that can be dominant in the disc to be that from a quadrupole term so $s \geq -7/2$.

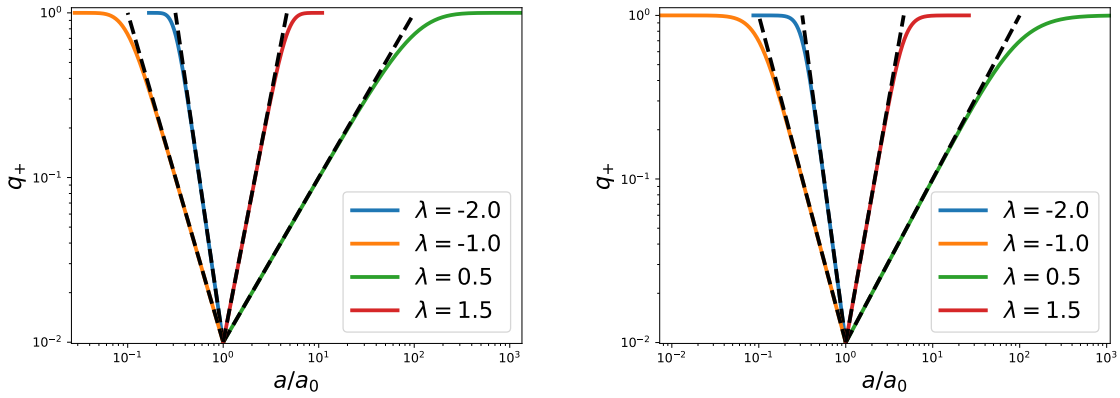


Fig. 6.4 Variation of the maximum nonlinearity attained during a wave cycle (q_+) on the disc lengthscale. Each solution is labelled by λ which corresponds to the power law for q_+ in the linear limit. Dashed lines correspond to power law solutions from linear theory. Left: Untwisted. Right: tightly wound.

6.7 Modulation on the disc lengthscale in power law discs

To illustrate our theory we calculate how the long length scale properties of an eccentric wave vary in a power law disc.

Figure 6.4 shows how q_+ varies on the disc length scale dependent on λ. As expected the solution is a power law at low q_+ where linear theory is valid. As the nonlinearity increases the solution turns over due to the increasingly strong pressure forces which are able to more effectively balance the precessional forces in the nonlinear regime so that q_+ need not be as high as would be predicted in linear theory. The solution asymptotes to $q_+ \to 1$ as the precessional forces diverge which in a power law disc can only occur at $a \to 0$ or $a \to \infty$, dependent on the sign of λ. There is little difference between the variation of the nonlinearity for tightly wound and untwisted discs.

Figure 6.5 similarly show how $\langle e^2 \rangle$ varies on the disc length scale, with its behaviour dependent on both λ and μ. Unlike q the equations are linear in e and as such we are free to rescale our solutions by a constant value. As with q_+ the solution is a power law in the linear regime and turns over at increasing nonlinearities so that it always lies below the power law slope of linear theory. The tightly wound waves are much less effected by nonlinear effects, resulting in much higher eccentricities in the nonlinear limit compared to the eccentric mode.

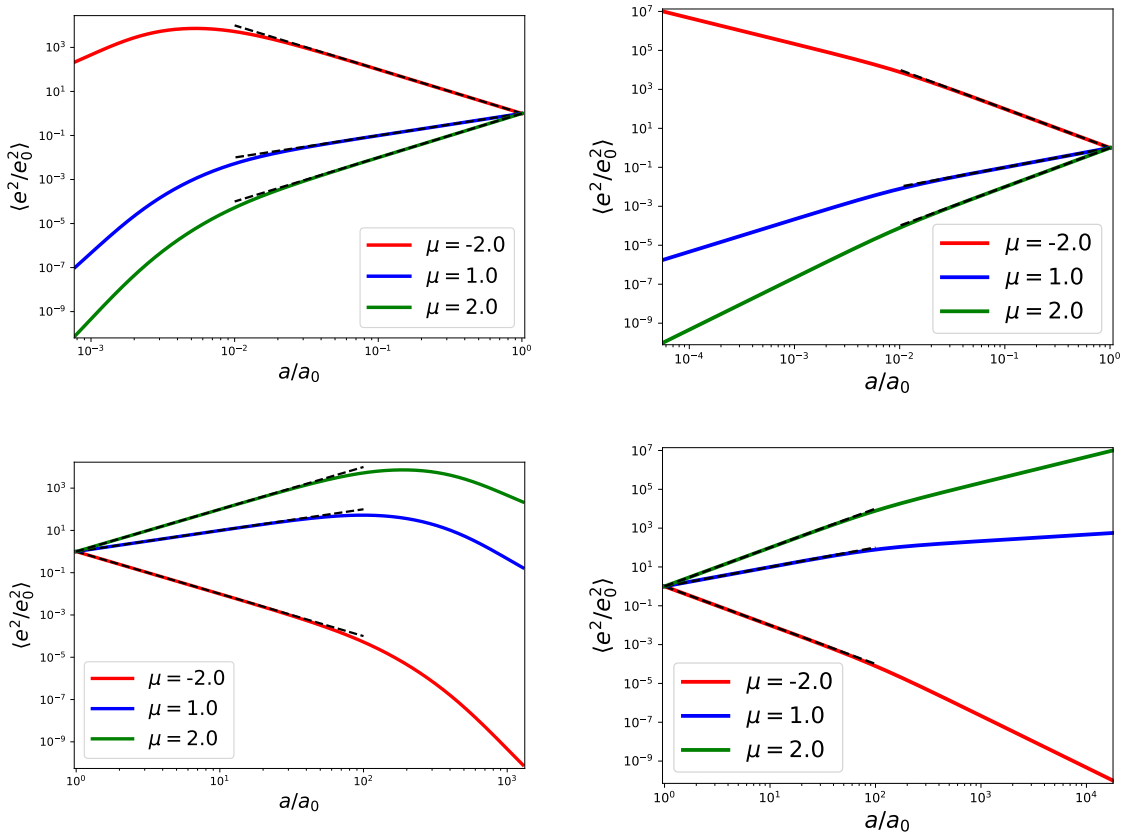


Fig. 6.5 Eccentric amplitude on the disc scale for, as the theory is linear in eccentricity, the absolute value of this is unimportant. Dashed lines correspond to power law solutions from linear theory. Upper left: untwisted $\lambda = -1$. Upper right: tightly wound $\lambda = -1$. Lower left: untwisted $\lambda = 1$. Lower right: tightly wound $\lambda = 1$. The nonlinear effects appear to be weaker for the tightly wound waves.

6.8 Astrophysical consequences

6.8.1 Exploring the parameter space relevant to astrophysical discs

Now we consider some astrophysical consequences of the conditions we have derived for eccentric wave steepening. To do so we will concentrate on the conditions derived for power law discs as these are sufficient to draw conclusions relevant to an astrophysical context. We will assume that the power law profile for the surface density is bounded such that $-2 < \sigma < 0$, which simply states that the surface density must decrease outwards but most of the mass of the disc is at large radii.

The local precessional terms $\omega_f - \omega$ are assumed to be dominated by one precessional effect, and can thus be described by a power law. As above we use $\omega_f - \omega \propto a^s$. Different regions of the disc are likely to be dominated by different precessional terms and the eccentric mode will switch between them as it moves to smaller semimajor axis. In general s should decrease moving from the outer to the inner disc as more centrally concentrated forces become important. Starting from precessional effects that dominate the outer disc and moving inwards, some important precessional effects and their associated values for s are

- Outer quadrupole from an external perturber: $s = 3/2$.
- Global (retrograde) disc precession: $s = 0$.
- Breathing mode: The effect of the breathing mode can be approximated with $s = r - \sigma - 1/2$. This is a special case of the precessional effects as it depends on the pressure profile. More generally the effect of the breathing mode should be included in the functional form of F.
- General relativity: The lowest order GR correction has $s = -5/2$.
- Inner quadrupole term: $s = -7/2$.

Figure 6.6 shows how these different precessional regimes and surface density/pressure profiles affect the behaviour of the eccentric wave. It should be noted that most discs will not occupy a single point on this diagram but will slowly move left to smaller $s + \sigma$ at smaller radii as short range forces become important. We consider vertically integrated pressure profiles $P^\circ \propto a^r$ with $-3 < r < 0$. A steady α-disc should have $r = -1.5$.

The different regions of eccentric wave behaviour are

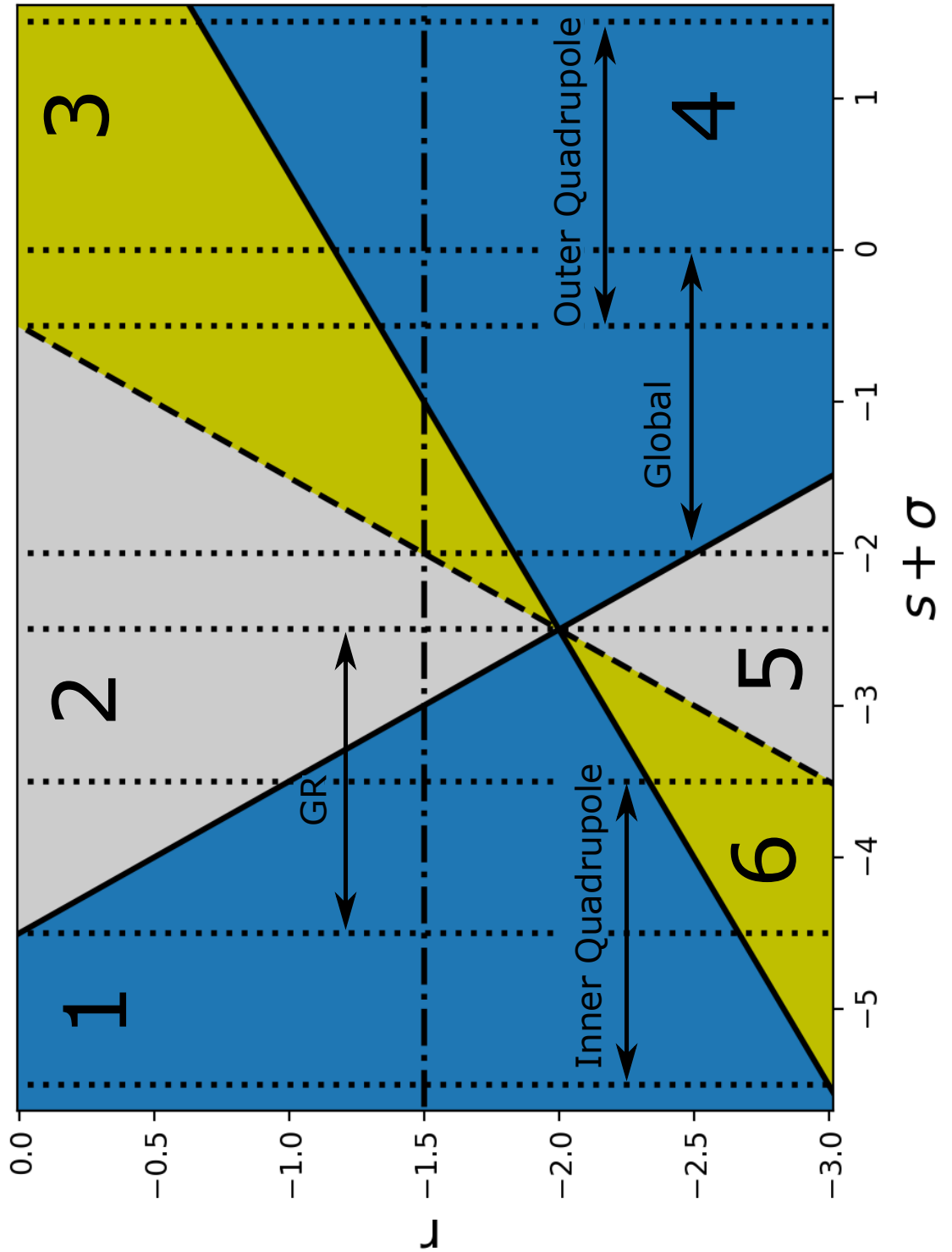


Fig. 6.6 Different regimes (see text) for an eccentric wave in discs with power law profiles. The solid lines separate regions which dictate the behaviour of the eccentric wave. The dashed line corresponds to the precession frequency associated with pressure forces (such as the breathing mode), thus here $\partial_a N = 0$ as the precessional forces are proportional to the pressure forces here. The dotted lines show the region of the parameter space occupied by discs with a given dominant precessional effect assuming the power law for the surface density profile is in the range $-2 < \sigma < 0$. In general the dominant precessional effect changes in a disc and moves towards the left of the diagram at smaller radii, with the dash-dotted line corresponding to the pressure profile of a steady α-disc.

1. $\frac{\partial q_+}{\partial a} < 0$, $\frac{\partial}{\partial a} N^2 < 0$, $\frac{\partial}{\partial a}\langle e^2 \rangle > 0$.
2. $\frac{\partial q_+}{\partial a} < 0$, $\frac{\partial}{\partial a} N^2 < 0$, $\frac{\partial}{\partial a}\langle e^2 \rangle < 0$ for small q_+, $\frac{\partial}{\partial a}\langle e^2 \rangle > 0$ for large q_+.
3. $\frac{\partial q_+}{\partial a} < 0$, $\frac{\partial}{\partial a} N^2 > 0$, $\frac{\partial}{\partial a}\langle e^2 \rangle < 0$ for small q_+, $\frac{\partial}{\partial a}\langle e^2 \rangle > 0$ for large q_+.
4. $\frac{\partial q_+}{\partial a} > 0$, $\frac{\partial}{\partial a} N^2 > 0$, $\frac{\partial}{\partial a}\langle e^2 \rangle < 0$.
5. $\frac{\partial q_+}{\partial a} > 0$, $\frac{\partial}{\partial a} N^2 > 0$, $\frac{\partial}{\partial a}\langle e^2 \rangle > 0$ for small q_+, $\frac{\partial}{\partial a}\langle e^2 \rangle < 0$ for large q_+. This region rarely occurs in real discs as this requires a particularly steep pressure gradient.
6. $\frac{\partial q_+}{\partial a} > 0$, $\frac{\partial}{\partial a} N^2 < 0$, $\frac{\partial}{\partial a}\langle e^2 \rangle > 0$ for small q_+, $\frac{\partial}{\partial a}\langle e^2 \rangle < 0$ for large q_+. As with region 5, this region rarely occurs in real discs.

Regions $1-4$ all feature focusing of the eccentric wave (i.e. an increase in either q_+ or $\langle e^2 \rangle$ inwards) by the background disc, whereas regions 5 and 6 (where the pressure gradients are very steep) are defocusing. In regions 2 and 3 the essentially linear phenomena of wave focusing competes against the nonlinear enhancement of pressure forces in the wave, resulting in a change in behaviour when the wave is sufficiently nonlinear.

The inner part of the disc covered by Figure 6.6 includes regions $1, 2, 5, 6$. As N^2 needs to be large in order that the separation of scales in the short wavelength limit remains valid, the dashed line corresponding to the breathing mode (and $\partial_a N^2 = 0$) is important in determining the validity of the theory. Considering an eccentric wave that starts in the short wavelength limit in the outer regions, then we require $\partial_a N^2 < 0$ for the theory to be applicable in the inner disc. Only region 5 violates this condition; however this region is unlikely to be important for real discs due to the steep pressure gradients required.

The outer part covered by Figure 6.6 includes regions $2-5$. Starting with an eccentric wave in the inner region in the short wavelength limit we require $\partial_a N^2 > 0$ for the theory to be applicable to the outer disc. Region 2 violates this; unlike region 5 in the inner disc, region 2 can be important to the outer part of discs with retrograde precession and shallow pressure gradients.

The reader might be concerned with the fact that a disc with inner quadrupole precession sits entirely in region 1, which predicts that $\frac{\partial}{\partial a}\langle e^2 \rangle > 0$. This contradicts the results of many hydrodynamic simulations of circumbinary discs (e.g. MacFadyen & Milosavljević (2008); Miranda et al. (2017); Thun et al. (2017); Thun & Kley (2018); Miranda & Rafikov (2018)). It should be noted however that this result is only true when the eccentricity is wavelike. The modes set up in simulations are better thought

of as confined modes trapped by the quadrupole potential and are non-oscillatory so the theory developed here is not applicable. It is possible in the outer disc far from the binary where the evanescent wave has declined to zero that the wavelike eccentricity should dominate. However this will require a mechanism for a binary to excite such a wave and maintain it against dissipation. As the confined and wavelike waves are nonlinear in different parts of the disc we expect that these two waves can be superposed with each dominating the flow geometry in a particular part of the disc.

6.8.2 Eccentric waves in an illustrative astrophysical system

An astrophysical system which highlights the role of the different precessional effects on the behaviour of the eccentric discs are X-ray binaries, which was studied by Ferreira & Ogilvie (2009) using linear theory. While it should be cautioned that the theory we have developed is not directly applicable to XRBs due to dissipative effects from MRI turbulence which is neglected, studying the ideal theory in such a system provides insight into the nonlinear behaviour of eccentric waves in real discs.

If we assume that a linear eccentric wave is generated in the outer region of the disc, most likely due to excitation at the 3:1 Lindblad resonance (Lubow, 1991a,b), we can follow this eccentric wave through the different regions of the disc where different forces dominate the precessional effects to see how it develops in the inner disc.

In the outermost part of the disc the dominant precessional force is from the external companion. Here the the disc will be in either region 3 or 4, for $r = -1.5$ (typical pressure exponent for an α-disc) the disc will be in region 4 and we expect the nonlinearity of the wave to decrease as it propagates inwards while its eccentricity increases.

If this precessional term dominates over a large radial extent in the disc then the eccentric wave will exit the short-wavelength regime and the theory developed is no longer valid. In this case we would expect that the eccentric wave will be global in character and vary on the disc lengthscale. However as the eccentric wave propagates away from the disc edge other precessional terms become important. The next precessional effect to occur is a global (retrograde) precession which results in the disc sitting in region 3 or 4. Here we expect the nonlinearity to start increasing again while the increase in the eccentricity tails off. The breathing mode can also be important here with a similar effect to a global eccentric mode; however its effects can extend into the inner disc extending region 2/3 further in than might be expected from 2D theory.

In the inner disc general relativistic precession will start to compete with the breathing mode, resulting in the disc transitioning into regions 1 or 2. This will cause the nonlinearity to increase while the amplification of the eccentricity will stall and eventually decline. In general regions 1,2 are conducive to the formation of a nonlinear short-wavelength wave. As such, if the eccentricity of the eccentric wave generated by the external companion has not increased to too high a value in the outer disc, we expect a short-wavelength wave as studied here to form in the inner disc (this can be regarded as a nonlinear extension of the results of Ferreira & Ogilvie (2009)). The last region that we have left off is the internal quadrupole; while present due to higher order GR corrections, this is less important to the eccentric wave as it never dominates over the lower order GR precession considered until the marginally stable orbit is reached.

Near the marginally stable orbit the pressure and surface density profiles turn over so that they decline inward. As GR precession will still be dominant in this region this will force the solution into region 2 resulting in the nonlinearity increasing inwards and possible increase of the eccentricity. In the linear theory (Ferreira & Ogilvie, 2009) the eccentricity and nonlinearity diverge at the disc edge. Here we predict that the eccentricity can grow, but must do so more slowly than the linear predictions of Ferreira & Ogilvie (2009); however this may still be enough for the short-wavelength theory to predict its own breakdown in this region. Unlike Ferreira & Ogilvie (2009) the absolute value of the nonlinearity has physical significance and cannot exceed $q = 1$.

It's not apparent whether the untwisted eccentric mode or the tightly wound eccentric wave is more relevant to the inner regions of black hole discs. This depends on the nature of the marginally stable orbit and whether it is permeable to eccentric waves. Early work such as Ferreira & Ogilvie (2009) assumed that it was not and imposed a zero torque boundary condition ($\mathcal{E}_a = 0$ in the linear limit). However recent simulation work (Dewberry et al., 2020a,b) featured a highly twisted eccentric wave showing evidence that either the eccentric wave is transmitted through the marginally stable orbit, or is catastrophically damped there.

6.8.3 Towards an extension of the results of Ferreira & Ogilvie (2009)

Following Ferreira & Ogilvie (2009) we consider the behaviour of an eccentric disc in a steady alpha model (Shakura & Sunyaev, 1973); however we neglect the effects of dissipation on the eccentric mode. We make use of the relativistic expressions for the characteristic frequencies in a Kerr metric (Okazaki et al., 1987; Kato, 1990)

$$\Omega = (r^{3/2} + s)^{-1} \quad , \tag{6.57}$$

$$\kappa = \Omega\sqrt{1 - \frac{6}{r} + \frac{8s}{r^{3/2}} - \frac{3s^2}{r^2}} \quad , \tag{6.58}$$

where in this section r is the cylindrical radius in units of GM/c^2, s is the dimensionless spin parameter of the black hole ($-1 < s < 1$) and the frequencies are in units of c^3/GM. In the Shakura & Sunyaev model the disc is split into an inner radiation pressure dominated region (region a) and an outer gas pressure dominated region (region b). In region a) (Ferreira & Ogilvie, 2009),

$$\Sigma^\circ \propto r^{3/2} f^{-1} \quad , \tag{6.59}$$

$$H \propto f \quad , \tag{6.60}$$

while in region b)

$$\Sigma^\circ \propto r^{-3/5} f^{3/5} \quad , \tag{6.61}$$

$$H \propto r^{21/20} f^{1/5} \quad , \tag{6.62}$$

where we have introduced $f = 1 - \sqrt{r_{in}/r}$ with r_{in} the dimensionless, spin dependent, radius of the marginally stable orbit.

This allows us to obtain the background profiles P°, Σ°, ω_f in the unperturbed circular disc. As noted in Ferreira & Ogilvie (2009) this background disc model is strictly Newtonian, apart from the truncation of the inner edge at the marginally stable orbit. The most important relativistic effect on the eccentric mode is the contribution to apsidal precession.

Making the approximation $r \approx a$ which is correct in the limit $N^2 \gg 1$ we obtain expressions for aH_a° and N which can be used to determine the eccentric disc dynamics:

$$aH_a^\circ \propto a^2 P^\circ \propto a^{1/2} f \quad , \tag{6.63}$$

$$N^2 \propto \frac{\Sigma^\circ}{aH_a^\circ} a^{-1}(6a - 8sa^{1/2} + 3s^2) \quad . \tag{6.64}$$

The behaviour of N depends on whether the disc is gas or radiation pressure dominated.

a) Radiation pressure dominated

$$N^2 \propto f^{-2}(6a - 8sa^{1/2} + 3s^2) \quad , \tag{6.65}$$

b) Gas pressure dominated

$$N^2 \propto f^{-2/5}a^{-21/10}(6a - 8sa^{1/2} + 3s^2) \quad . \tag{6.66}$$

In the outer part of the disc (region b of the Shakura & Sunyaev model) our theory predicts that the eccentric wave should be in region 1 where eccentricity decreases inwards but nonlinearity and N^2 increase. This is the same behaviour as found in Ferreira & Ogilvie (2009) as the nonlinearity doesn't modify the qualitative behaviour in this region. The main effect of nonlinearity will be to make the inward decrease of the eccentricity stronger so that for a given radius the eccentricity in the nonlinear theory will be lower than that predicted in Ferreira & Ogilvie (2009).

Consider the region close to the marginally stable orbit and assume that this corresponds to region a) of the Shakura & Sunyaev model. In the linear regime the envelopes for q and $\langle e^2 \rangle$ are given by

$$q \propto f^{-1}a^{-1/4}(6a - 8sa^{1/2} + 3s^2)^{1/4} \quad , \tag{6.67}$$

$$\langle e^2 \rangle \propto a^{-1/2}(6a - 8sa^{1/2} + 3s^2)^{-1/2} \quad , \tag{6.68}$$

which shows a divergence of q as the wave approaches the marginally stable orbit, while the eccentricity should reach a finite value. In nonlinear theory in this part of the disc we obtain the following for $J(q_+)$:

$$J(q_+) \propto f^{-2}a^{-1/2}(6a - 8sa^{1/2} + 3s^2)^{1/2}. \tag{6.69}$$

Near the marginally stable orbit $f \to 1$ so we require that $J(q_+)$ diverges. In linear theory as $J(q_+) \propto q_+^2$ this would cause q_+ to diverge. However in nonlinear theory $J(q_+)$ diverges when $q_+ \to 1$ so we can conclude that $q_+ \to 1$ as $f \to 0$. For the eccentricity we can write

$$\langle e^2 \rangle = \langle e_{\rm lin}^2 \rangle \frac{I(0)}{I(q_+)}. \tag{6.70}$$

As $\langle e_{\rm lin}^2 \rangle$ tends to a constant as $f \to 0$ and $I(q_+)$ diverges as $q_+ \to 1$ we find that $\langle e^2 \rangle \to 0$ as $f \to 0$.

Given these results, the expected behaviour for an eccentric wave approaching the marginally stable orbit is for the wave to circularise while approaching an orbital intersection. This may have important consequences for the boundary conditions applicable at this location, as the near orbital intersection at the marginally stable orbit suggests that the wave should be highly dissipated here.

Eccentric waves in the inner region of XRBs are of interest as a driving mechanism for high frequency quasi-periodic oscillations (QPOs), a type of variability observed in the X-ray spectra of some black hole systems, via the excitation of trapped inertial waves (Kato, 2008; Ferreira & Ogilvie, 2008, 2009; Dewberry et al., 2018, 2019). This chapter shows the short-wavelength limit of eccentric waves may be important for the inner regions of XRBs, which is a nonlinear extension of those considered by Ferreira & Ogilvie (2008), Ferreira & Ogilvie (2009) in this context. Ferreira & Ogilvie (2008) showed that the growth rate of trapped inertial waves from interaction with the eccentric background is proportional to $|r\frac{dE}{dr}|^2 \approx q^2$, as such nonlinear eccentric waves should couple more strongly to inertial waves and thus be more effective at generating QPOs.

6.8.4 Tightly wound waves in an isothermal pseudo-Newtonian disc, a comparison with Dewberry et al. (2020a) and Dewberry et al. (2020b)

Recent simulations of eccentric discs in a 3D unstratified pseudo-Newtonian disc were carried out by Dewberry et al. (2020a) (hydrodynamic) and Dewberry et al. (2020b) (magnetohydrodynamic). In both papers an eccentric wave was excited at the outer boundary and propagated inwards, forming a twisted eccentric disc. Dewberry et al. (2020a) and Dewberry et al. (2020b) considered a globally isothermal disc with a constant surface density and mimicked the the effects of GR through the Paczyński-Wiita potential (Paczyńsky & Wiita, 1980),

$$\Phi_{\rm PW} = -\frac{GM_1}{r - r_s} \quad , \tag{6.71}$$

where r_s is the Schwartzchild radius of the black hole.

One important finding of Dewberry et al. (2020b) was that while the eccentric wave became increasingly nonlinear as it approached the marginally stable orbit there was also a decrease in eccentricity. This effect was present in both the hydrodynamic and

magnetohydrodynamic runs and is thus unlikely to be related to turbulent dissipation of the disc eccentricity. This is in contrast to the expectations of Ferreira & Ogilvie (2009) where the eccentricity only declines as it approaches the marginally stable orbit if it is subject to sufficient damping.

Based on some of the published material that makes up this Chapter, Dewberry et al. (2020b) suggested that the decline in eccentricity was due to the nonlinear steepening of the wave as it approaches the marginally stable orbit, specifically the effect captured by equation 6.70. In this section we shall show that the results of Dewberry et al. (2020b) appear to be caused by a peculiarity of the Paczyński-Wiita potential which, when combined with the nonlinear wavesteepening explain the results of the simulations.

As discussed above, there is currently a lot of uncertainty in what boundary condition to apply at the marginally stable orbit. However, as the eccentric discs produced in Dewberry et al. (2020a) and Dewberry et al. (2020b) are clearly twisted they are best compared with the tightly wound theory. The disc is globally isothermal with a constant surface density so the mass per unit semimajor axis $M_a = 2\pi\Sigma^\circ a$ and circular Hamiltonian density $H_a^\circ = 2\pi c_0^2 \Sigma^\circ a$, where c_0 and Σ° are constant sound speed and surface density respectively. For a Paczyński-Wiita potential the circular angular velocity is (Paczyńsky & Wiita, 1980)

$$\Omega = \sqrt{\frac{GM_1}{r^3}} \frac{r}{r - r_s} \tag{6.72}$$

and the epicyclic frequency is (Okazaki et al., 1987)

$$\kappa = \Omega \sqrt{\frac{r - 3r_s}{r - r_s}} \quad . \tag{6.73}$$

Notably this is different from the exact epicyclic frequency measured at infinity (equation 6.58) which was adopted by Ferreira & Ogilvie (2009). This difference is demonstrated by Figure A1 of Okazaki et al. (1987) which shows the pseudo-Newtonian epicyclic frequency consistently overpredicts the relativistic epicyclic frequency, but declines more rapidly so that the marginally stable orbit is located in the correct place.

As in Section 6.8.3 we make the approximation $a \approx r$ and, using Equations 6.72 and 6.73, obtain the forced precession frequency from the Paczyński-Wiita potential,

$$\omega_f = \frac{n a^2 r_s}{(a - r_s)^3}. \tag{6.74}$$

Here we have used $n = \sqrt{\frac{GM_1}{a^3}}$ to denote the Keplerian mean motion. At large distances ω_f behaves like $\omega_f \propto a^{-5/2}$ which was used to model GR precession in the power law discs. This, along with surface density and pressure profiles, gives an expression for N^2,

$$N^2 = \frac{n^2 a^4 r_s}{c_0^2 (a - r_s)^3} \quad . \tag{6.75}$$

In the linear regime, q and the shape of the eccentric envelope are given by

$$q \propto a^{-3/4}(a - r_s)^{-3/4} \quad , \tag{6.76}$$

$$\langle e^2 \rangle \propto a^{-5/2}(a - r_s)^{3/2} \quad . \tag{6.77}$$

$\langle e^2 \rangle$ has a maxima interior to the marginally stable orbit and vanishes on the event horizon. These solutions are markedly different from the solutions in Section 6.8.3. The main cause of this difference is the absence of a sharp disc truncation at the marginally stable orbit which is responsible for the singular behaviour seen in Ferreira & Ogilvie (2009).

Equivalent expressions for using the characteristic frequencies of the Schwarzschild metric are,

$$q \propto a^{-3/2}, \tag{6.78}$$

$$\langle e^2 \rangle \propto a^{-1}, \tag{6.79}$$

which agree with those of the Paczyński-Wiita potential at large semimajor axis.

As in Section 6.8.3, in the nonlinear theory we obtain an expression for $J(q_+)$,

$$J(q_+) \propto a^{-3/2}(a - r_s)^{-3/2}. \tag{6.80}$$

This diverges as the wave approaches the event horizon, meaning we expect $q \to 1$ interior to the marginally stable orbit. Whether q_+ attains a value close to 1 by the time it reaches the marginally stable orbit therefore depends on the amplitude of the wave on the outer boundary. This situation changes if, as seen in simulations (Dewberry et al., 2020a,b), the surface density falls to zero near the marginally stable orbit. In which case $J(q_+)$ diverges, meaning $q \to 1$ as the wave approaches the edge of

the disc, similar to behaviour seen in Section 6.8.3. This will lead to the circularisation of the wave as it approaches the marginally stable orbit.

As in Section 6.8.3 the nonlinear eccentricity is related to the linear eccentricity through equation 6.70, however for tightly wound waves in an isothermal disc $\frac{I(0)}{I(q_+)}$ is of order unity until q_+ is very close to 1, so the nonlinear eccentricity envelope is reasonably well approximated by the linear result,

$$\langle e^2 \rangle \approx e_0^2 a^{-5/2} (a - r_s)^{3/2}, \tag{6.81}$$

throughout most of the disc. The (linear) eccentricity increases as it approaches the marginally stable orbit, and has a maxima at $a = (5/2)r_s$. If the wave becomes highly nonlinear before crossing the marginally stable orbit then $\frac{I(0)}{I(q_+)}$ will no longer be close to 1 causing the eccentricity to decrease inwards, similar to the behaviour seen in Dewberry et al. (2020b). This decline in eccentricity is greatly facilitated by the much shallower profile of $e_{\rm lin}$ in the Paczyński-Wiita potential, as it is approaching the maximum at $a = (5/2)r_s$, compared with the steeper profile found when using the frequencies from a Schwarzschild metric. This can be seen in the right hand plot in Figure 6.7 where the eccentricity in the Paczyński-Wiita disc declines as it approaches the marginally stable orbit, while that using the Schwarzschild frequencies approaches a constant value. This suggests caution when using the Paczyński-Wiita potential to study eccentric waves due to their sensitivity to the details of the epicyclic frequency in the disc.

The simulations of Dewberry et al. (2020a) and Dewberry et al. (2020b) found that a weakly excited eccentric wave in an isothermal Paczyński-Wiita disc increases inwards, as it approaches the marginally stable orbit, while a strongly excited eccentric wave decreases inwards. This is consistent with the results of this chapter. Unless q is very close to 1 the eccentricity of a tightly wound wave is essentially unchanged from linear theory (i.e. it follows Equation 6.81), and will increase inwards. However if the excitation at the outer boundary is strong enough q can approach 1 exterior to the marginally stable orbit, causing the eccentricity to decline. This is clearly seen in Figure 6.7 where the linear and nonlinear theory predict opposite behaviour for the eccentricity in the disc.

Finally Figure 6.8 shows plots of the surface density and radial velocity predicted by our theory. These appear qualitatively similar to the simulation outputs of Dewberry et al. (2020a) and Dewberry et al. (2020b).

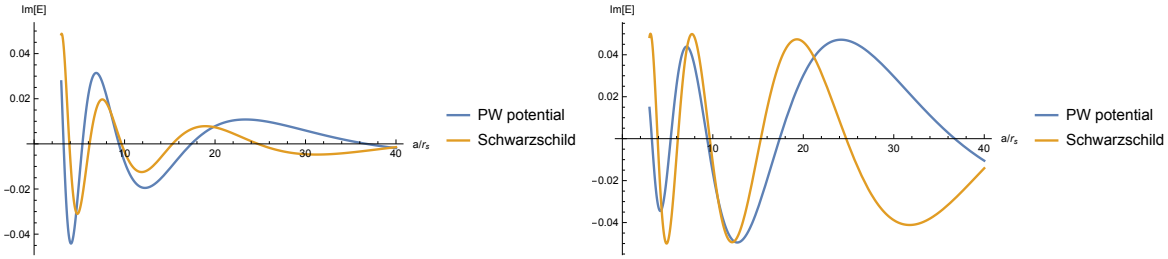


Fig. 6.7 Linear (left) and Nonlinear (right) profiles of $\mathrm{Im}[\mathcal{E}] = e\sin\varpi$ in a Paczyński-Wiita potential (Blue) and using the exact forced precession from a Schwarzschild metric. The linear theory of the two models of GR precession are broadly similar. The eccentricity of the nonlinear theory in the Paczyński-Wiita potential features a decline as it approach the marginally stable orbit that was observed in Dewberry et al. (2020b); this doesn't occur when using the exact epicyclic frequencies, where the eccentricity instead saturates. The circularisation of the eccentric wave seen in Section 6.8.3 does not occur here as the disc continues across the marginally stable orbit with a constant surface density. If the disc surface density falls to zero at the marginally stable orbit then the Nonlinear solutions will circularise there as in Section 6.8.3.

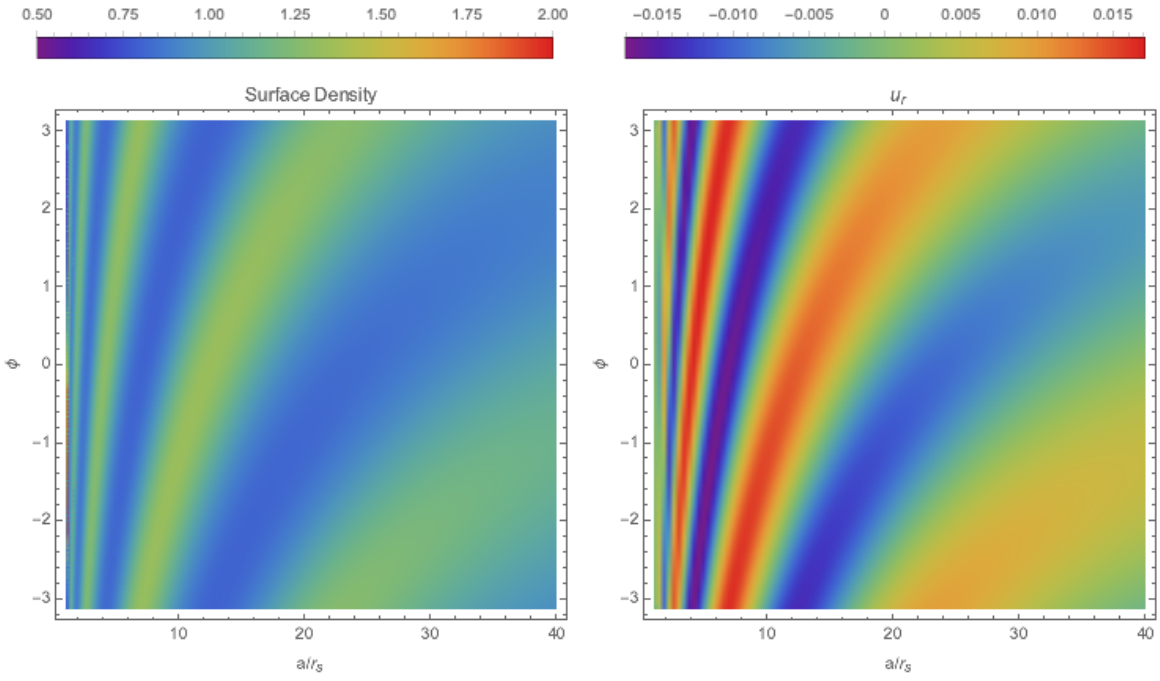


Fig. 6.8 Surface density (Left) and radial velocity (Right) for an eccentric wave propagating in a globally isothermal Paczyński-Wiita disc. $c_0 = 0.01$ and the eccentric wave has been normalised so that the maximum eccentricity in the disc matches that in Figure A3. from Dewberry et al. (2020b).

6.9 Conclusion

In this Chapter we have formulated a nonlinear theory for short wavelength eccentric waves in astrophysical discs based on the averaged Lagrangian method of Whitham (1965). We have found that the linear theory can be used to provide bounds on the nonlinear behaviour and can thus be useful in determining the large lengthscale behaviour. We have also derived conditions under which nonlinearity and/or eccentricity can grow through wave steepening in a disc and discussed what effect this has on eccentric waves in various astrophysical settings.

In general we have found that the inclusion of the nonlinearity has the following effects:

- There are conditions under which eccentric waves remain in the short-wavelength limit but become increasingly nonlinear as they propagate inwards. These depend on the precessional effects in the disc and the surface density/pressure profiles. This is a form of wave-focusing of the eccentric waves which can occur in astrophysical discs. While predicted in linear theory, we have shown that this wave focusing occurs even in the presence of strong nonlinearity.

- There exists an highly nonlinear limit where untwisted eccentric wave becomes a triangle wave and the maximum nonlinearity in a wave cycle q_+ approaches 1. In this limit the eccentric wave oscillates between nearly intersecting orbits, with alternating argument of pericentre, on a short radial lengthscale comparable to (but slightly longer than) the scale height.

- Nonlinearity prevents the generation of an orbital intersection by strengthening the pressure forces so that it is necessary for the precessional forces to diverge in order that an orbital intersection is generated.

The interior region of an X-ray binary should naturally support such short-wavelength eccentric waves, while the orbiting companion provides a mechanism by which they can be generated. These eccentric waves are of interest as a driving mechanism for QPOs and were simulated by Dewberry et al. (2020a) and Dewberry et al. (2020b), where the effects of eccentric wave focusing and steepening were demonstrated. The decline in the eccentricity, as it approaches the marginally stable orbit, appears to be due to a combination of the nonlinear steepening predicted in this chapter and a peculiarity of the Paczyński-Wiita potential and is less pronounced when using the exact GR precession frequencies.

Part III

Nonconservative Terms and Eccentric Discs

Chapter 7

Dissipation and Radiation in Eccentric Disc Theory

7.1 Introduction

In the previous chapters we considered the conservative dynamics of eccentric discs and neglected dissipative effects such as turbulent viscosity and radiation. As reviewed in Chapter 2 the action of turbulent viscosity in eccentric discs is complicated and can act to excite, instead of damp, the eccentricity. This is further complicated by the fact that it is not clear how MRI turbulence operates in an eccentric disc, and whether it can be successfully modelled as an effective viscosity. Recent simulations by Dewberry et al. (2020b) have found that a sufficiently nonlinear eccentric wave is capable of suppressing the MRI.

The emission from an eccentric disc is, unsurprisingly, more complicated than that from a circular disc. The temperature and disc thickness in an eccentric disc vary around the orbit, with emission disproportionately originating from pericentre where the disc is typically both thinnest and hottest. This asymmetric emission allows eccentric discs to be much more radiatively efficient than their circular counterparts and can reach greater temperatures and luminosities while remaining geometrically thin. In the extreme case of the highly eccentric discs we will consider in subsequent chapters, the entire disc thermodynamics can be controlled by a small region around pericentre where the vast majority of dissipative heating and radiative cooling occurs.

In the rest of this chapter we shall derive the equations of vertical structure for an eccentric disc with an $\alpha-$viscosity which is subject to radiative cooling. These equations take the form of a local model of the fluid flow in an eccentric disc, similar

to that developed in Ogilvie & Barker (2014), and describe the simplest laminar flow in an eccentric disc subject to nonideal effects.

7.2 Derivation of the equations of vertical structure, including thermal physics

7.2.1 Local model of a horizontally invariant "laminar flow" in an eccentric disc

A local model of a thin, Keplerian eccentric disc was developed in Ogilvie & Barker (2014). For a disc that includes viscosity and radiative cooling a local model can be obtained in a similar manner. We use the (a, E) orbital coordinates (introduced in Section 4.3.1 of Chapter 4) rather than those in Ogilvie & Barker (2014). The equations, formulated in a frame of reference that follows the elliptical orbital motion, are the vertical equation of motion,

$$\frac{Dv_z}{Dt} = -\frac{GM_1 z}{r^3} - \frac{1}{\rho}\frac{\partial}{\partial z}\left(p - T_{zz}\right), \tag{7.1}$$

the continuity equation,

$$\frac{D\rho}{Dt} = -\rho\left(\Delta + \frac{\partial v_z}{\partial z}\right), \tag{7.2}$$

and the thermal energy equation,

$$\frac{Dp}{Dt} = -\Gamma_1 p\left(\Delta + \frac{\partial v_z}{\partial z}\right) + \left(\Gamma_3 - 1\right)\left(\mathcal{H} - \frac{\partial F}{\partial z}\right), \tag{7.3}$$

where, for horizontally invariant "laminar flows",

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + v_z\frac{\partial}{\partial z} \tag{7.4}$$

is the Lagrangian time-derivative,

$$\Delta = \frac{1}{J}\frac{dJ}{dt} \tag{7.5}$$

is the divergence of the orbital velocity field, with J the Jacobian of the (Λ, λ) coordinate system, and is a known function of E that depends on e, q and E_0 (see Section 4.3.1 of Chapter 4 for details). $F = F_{\rm rad} + F_{\rm ext}$ is the total vertical heat flux with

$$F_{\rm rad} = -\frac{16\sigma T^3}{3\kappa\rho}\frac{\partial T}{\partial z} \tag{7.6}$$

being the vertical radiative heat flux and $F_{\rm ext}$ containing any additional contributions to the heat flux (such as from convection or turbulent heat transport).

$$T_{zz} = 2\mu_s\frac{\partial v_z}{\partial z} + \left(\mu_b - \frac{2}{3}\mu_s\right)\left(\Delta + \frac{\partial v_z}{\partial z}\right) \tag{7.7}$$

is the zz component of the viscous stress tensor and $\mathcal{H}$ is the viscous heating rate per unit volume. μ_s and μ_b are the dynamical shear and bulk viscosities. Γ_1 and Γ_3 are the first and third adiabatic exponents.

We assume the pressure in the disc, which includes contributions from radiation and a perfect gas with a ratio of specific heats γ, is given by

$$p = p_r + p_g = \frac{4\sigma}{3c}T^4 + \frac{\mathcal{R}\rho T}{\mu} \quad . \tag{7.8}$$

We define β_r to be the ratio of radiation to gas pressure:

$$\beta_r := \frac{p_r}{p_g} = \frac{4\sigma\mu}{3c\mathcal{R}}\frac{T^3}{\rho} \quad . \tag{7.9}$$

We assume a constant opacity κ, applicable to the electron-scattering opacity relevant for the application to TDEs in subsequent chapters. It is possible to generalise this theory to other opacity laws, in particular power law opacities of the form $\kappa \propto \rho^x T^y$ (where x,y are constants) by following Ogilvie (2001).

We further assume an alpha viscosity,

$$\mu_{s,b} = \alpha_{s,b}\frac{p_v}{\omega_{\rm orb}}, \tag{7.10}$$

where $\alpha_{s,b}$ are dimensionless coefficients and $\omega_{\rm orb}$ is some characteristic frequency of the orbital motion. In the $\alpha-$viscosity prescription the stress scales with some pressure p_v; for the radiation-gas mixture it is not clear whether this pressure should be the total pressure, the gas pressure or some average of the two. For the pure gas there is only one pressure so $p_v = p$.

It is very unclear what the correct frequency to use should be. For instance we could have $\omega_{\rm orb} = n = (GM_1/a^3)^{1/2}$, the mean motion, $\omega_{\rm orb} = \Omega = n\sqrt{1-e^2}(1-e\cos E)^{-2}$, the angular frequency or $\omega_{\rm orb} = n(1-e\cos E)^{-3/2}$, the circular frequency at that radius. In fact Ogilvie (2001) used a different scaling with $\omega_{\rm orb} = n(1-e^2)^{-3/2}$. In a circular disc this distinction is unimportant as $n = \Omega = \Omega_c$, but for a highly eccentric

disc $\Omega, \Omega_c \gg n$ at pericentre, so this makes a significant difference to the strength of the viscosity. Unless otherwise stated we choose $\omega_{\rm orb} = n$ simply because this is independent of the eccentricity and eccentric anomaly; this choice is equivalent to Ogilvie (2001) with an appropriately rescaled α.

The viscous heating rate is

$$\mathcal{H} = f_{\mathcal{H}}(t) n p_v, \tag{7.11}$$

where $f_{\mathcal{H}}$ is a complicated dimensionless expression given in Section 7.2.4.

We consider two situations, a pure gas disc with $F_{\rm ext} = 0$ and $\Gamma_1 = \Gamma_3 = \gamma$ and a radiation+gas mixture where there is an additional heat flux $F_{\rm ext}$ due to convective or turbulent mixing and the first and third adiabatic exponents are given by (Chandrasekhar, 1967)

$$\Gamma_1 = \frac{1 + 12(\gamma-1)\beta_r + (\gamma-1)(1+4\beta_r)^2}{(1+\beta_r)(1+12(\gamma-1)\beta_r)} \quad , \tag{7.12}$$

$$\Gamma_3 = 1 + \frac{(\gamma-1)(1+4\beta_r)}{1+12(\gamma-1)\beta_r} \quad . \tag{7.13}$$

7.2.2 Separable solutions

We propose a separable solution of the form

$$\begin{aligned} \rho &= \hat{\rho}(t)\tilde{\rho}(\tilde{z}), \\ p &= \hat{p}(t)\tilde{p}(\tilde{z}), \\ T &= \hat{T}(t)\tilde{T}(\tilde{z}), \\ F &= \hat{F}(t)\tilde{F}(\tilde{z}), \\ v_z &= \frac{dH}{dt}\tilde{z}, \end{aligned} \tag{7.14}$$

where

$$\tilde{z} = \frac{z}{H(t)} \tag{7.15}$$

is a Lagrangian variable that follows the vertical expansion of the disc, i.e. $D\tilde{z}/Dt = 0$, and the quantities with the tildes satisfy the dimensionless equations of vertical structure,

$$\frac{d\tilde{p}}{d\tilde{z}} = -\tilde{\rho}\tilde{z}, \tag{7.16}$$

$$\frac{d\tilde{F}}{d\tilde{z}} = \lambda\tilde{p}, \tag{7.17}$$

$$\frac{d\tilde{T}}{d\tilde{z}} = -\tilde{\rho}\tilde{T}^{-3}\tilde{F}_{\rm rad}, \tag{7.18}$$

$$\tilde{p} = \tilde{\rho}\tilde{T}. \tag{7.19}$$

Here λ is a dimensionless eigenvalue of the vertical structure problem when appropriate boundary conditions are imposed and corresponds to a dimensionless cooling rate.

The separated solution works provided that the dimensionful quantities ($\hat{\rho}(t)$,... etc) satisfy

$$\begin{aligned}\frac{d^2H}{dt^2} = &-\frac{GM_\bullet}{r^3}H + \frac{\hat{p}}{\hat{\rho}H} \\ &\times\left[1 - \frac{2\alpha_s}{nH}\frac{\hat{p}_v}{\hat{p}}\frac{dH}{dt} - \frac{(\alpha_b - \frac{2}{3}\alpha_s)}{n}\frac{p_v}{p}\left(\Delta + \frac{1}{H}\frac{dH}{dt}\right)\right],\end{aligned} \tag{7.20}$$

$$\frac{d\hat{\rho}}{dt} = -\hat{\rho}\left(\Delta + \frac{1}{H}\frac{dH}{dt}\right), \tag{7.21}$$

$$\frac{d\hat{p}}{dt} = -\Gamma_1\hat{p}\left(\Delta + \frac{1}{H}\frac{dH}{dt}\right) + (\Gamma_3 - 1)\left(f_{\mathcal{H}}n\hat{p}_v - \lambda\frac{\hat{F}}{H}\right), \tag{7.22}$$

$$\hat{F} = \frac{16\sigma\hat{T}^4}{3\kappa\hat{\rho}H} \quad , \tag{7.23}$$

$$\hat{p} = (1 + \beta_r)\frac{\mathcal{R}\hat{\rho}\hat{T}}{\mu} \quad , \tag{7.24}$$

where $\beta_r = 0$ for the pure gas and

$$\beta_r = \frac{4\sigma\mu}{3c\mathcal{R}}\frac{\hat{T}^3}{\hat{\rho}} \tag{7.25}$$

for the radiation-gas mixture.

Note that the surface density and vertically integrated pressure are

$$\Sigma = \hat{\rho}H, \quad P = \hat{p}H, \quad P_v = \hat{p}_v H. \tag{7.26}$$

The vertically integrated heating and cooling rates are

$$f_{\mathcal{H}} n P_v, \quad \lambda \hat{F} \quad . \tag{7.27}$$

The cooling rate can also be written as

$$\lambda \hat{F} = 2\sigma \hat{T}_s^4 \tag{7.28}$$

where $\hat{T}_s(t)$ is a representative surface temperature defined by

$$\hat{T}_s^4 = \frac{8\lambda}{3} \frac{\hat{T}^4}{\tau} \tag{7.29}$$

and

$$\tau = \kappa \Sigma \tag{7.30}$$

is the optical thickness.

We then have

$$\begin{aligned} \frac{1}{H}\frac{d^2H}{dt^2} = &-\frac{GM_\bullet}{r^3} + \frac{P}{\Sigma H^2} \\ &\times \left[1 - \frac{2\alpha_s}{nH}\frac{P_v}{P}\frac{dH}{dt} - \frac{(\alpha_b - \frac{2}{3}\alpha_s)}{n}\frac{P_v}{P}\left(\Delta + \frac{1}{H}\frac{dH}{dt}\right)\right], \end{aligned} \tag{7.31}$$

$$J\Sigma = \text{constant}, \tag{7.32}$$

$$\frac{dP}{dt} = -\Gamma_1 P\left(\Delta + \frac{1}{H}\frac{dH}{dt}\right) + \frac{P}{H}\frac{dH}{dt} + (\Gamma_3 - 1)nP\left(f_{\mathcal{H}}\frac{P_v}{P} - \frac{\lambda\hat{F}}{nP}\right), \tag{7.33}$$

with

$$\frac{\lambda\hat{F}}{nP} = \lambda \frac{16\sigma(\mu/\mathcal{R})^4}{3\kappa n} P^3 \Sigma^{-5} (1+\beta_r)^{-4} \quad . \tag{7.34}$$

For a standard circular disc we have $f_{\mathcal{H}} = \frac{9}{4}\alpha_s$ and hydrostatic balance, $P/\Sigma H^2 = GM_1/r^3 = n^2$. These determine the equilibrium values, H° and T°, etc., for a reference

circular orbit, in terms of Σ and n. When $p_v = p$ for the radiation-gas mixture, there are two physical solution branches; for what follows the reference state is the gas pressure dominated branch.

Scaling H by H°, $\hat{T}$ by T°, t by $1/n$, r by a and J by $2/n$ we obtain the dimensionless version

$$\begin{aligned}\frac{\ddot{H}}{H} &= -(1 - e\cos E)^{-3} + \frac{T}{H^2}\frac{1+\beta_r}{1+\beta_r^\circ} \\ &\quad \times \left[1 - 2\alpha_s \frac{P_v}{P}\frac{\dot{H}}{H} - (\alpha_b - \frac{2}{3}\alpha_s)\frac{P_v}{P}\left(\frac{\dot{J}}{J} + \frac{\dot{H}}{H}\right)\right],\end{aligned} \tag{7.35}$$

$$\begin{aligned}\dot{T} &= -(\Gamma_3 - 1)T\left(\frac{\dot{J}}{J} + \frac{\dot{H}}{H}\right) \\ &\quad + (\Gamma_3 - 1)\frac{1+\beta_r}{1+4\beta_r}T\left(f_{\mathcal{H}}\frac{P_v}{P} - \frac{9}{4}\alpha_s\frac{P_v^\circ}{P^\circ}\frac{1+\beta_r^\circ}{1+\beta_r}J^2T^3\right),\end{aligned} \tag{7.36}$$

where a dot over a letter indicates a derivative with respect to the rescaled time. We have written the thermal energy equation in terms of the temperature. The factor $\frac{\Gamma_3 - 1}{1+4\beta_r} \propto \frac{1}{c_V}$ where c_V is the specific heat capacity at constant volume. β_r can be obtained through

$$\beta_r = \beta_r^\circ JHT^3 \quad , \tag{7.37}$$

where we have introduced β_r°, which is the β_r of the reference circular disc. The equilibrium values of the reference circular disc H°, $\hat{T}^\circ$, etc., are determined by β_r° and n.

We introduce the (nondimensionalised) specific entropy

$$s := \ln(JHT^{1/(\gamma-1)}) + 4\beta_r \quad , \tag{7.38}$$

which includes contributions from the gas and radiation. This evolves according to

$$\dot{s} = (1+\beta_r)\left(f_{\mathcal{H}}\frac{P_v}{P} - \frac{9}{4}\alpha_s\frac{P_v^\circ}{P^\circ}\frac{1+\beta_r^\circ}{1+\beta_r}J^2T^3\right) \tag{7.39}$$

which we solve for the pure gas case instead of Equation 7.36. For the radiation-gas mixture we solve Equation 7.36 to avoid having to invert Equation 7.38 for the temperature at each timestep.

7.2.3 Generalised vertical structure equations

In this section we discuss the solution to the dimensionless equations of vertical structure (Equations 7.16-7.19). To begin with we must specify the appropriate boundary conditions. At the upper surface of the disc the density and pressure are assumed to vanish. Defining $\tilde{z}_s$ to be the location of the upper surface, with $\tilde{p}(\tilde{z}_s) = \tilde{\rho}(\tilde{z}_s) = 0$, we have

$$2\tilde{F}(\tilde{z}_s) = \lambda \quad . \tag{7.40}$$

Thus λ is a dimensionless cooling rate when the following normalisation conditions

$$\int_{-\infty}^{\infty} \tilde{\rho}\, d\tilde{z} = \int_{-\infty}^{\infty} \tilde{\rho}\tilde{z}^2\, d\tilde{z} = 1 \tag{7.41}$$

are imposed. Note that these together with the vertical structure equations imply

$$\int_{-\infty}^{\infty} \tilde{p}\, d\tilde{z} = 1 \quad . \tag{7.42}$$

The total dimensionless heat flux is $\tilde{F} = \tilde{F}_{\rm rad} + \tilde{F}_{\rm ext}$. We assume that heat is only transported radiatively in the disc photosphere so that $\tilde{F}_{\rm ext}(z_s) = 0$ and $\tilde{F}(\tilde{z}_s) = \tilde{F}_{\rm rad}(\tilde{z}_s)$. In the pure gas case we take $\tilde{F} = \tilde{F}_{\rm rad}$ throughout and Equations 7.16-7.19 can be solved numerically to determine that $\lambda = 3.213$.

For a radiation+gas mixture there is an additional contribution to the heat flux $\tilde{F}_{\rm ext}$ due to vertical mixing, either from convection (Bisnovatyi-Kogan & Blinnikov, 1976), or from the disc turbulence. This mixing is assumed to set up a vertically isentropic structure, which for a radiation dominated disc, corresponds to a polytrope with index 3 (this is a similar model to that of Ferreira & Ogilvie (2009) who looked at eccentricity propagation in a black hole disc). For a radiatively cooled, gas-dominated disc the vertical structure is close to the polytrope with index 3. As such a polytropic vertical structure with index 3 should be a good approximation to the radiation+gas mixture case for a constant opacity disc. This model has $\tilde{T}^3/\tilde{\rho} = \text{const}$ and therefore β_r is independent of z, which is required to maintain separability. The pressure and density for the polytrope are

$$\tilde{p} = \tilde{p}_0 \left[1 - \left(\frac{\tilde{z}}{\tilde{z}_s}\right)^2\right]^4, \quad \tilde{\rho} = \tilde{\rho}_0 \left[1 - \left(\frac{\tilde{z}}{\tilde{z}_s}\right)^2\right]^3 \quad . \tag{7.43}$$

The normalisation conditions (Equations 7.41-7.42) set $\tilde{z}_s = 3$, $\tilde{p}_0 = \frac{105}{256}$, $\tilde{\rho}_0 = \frac{35}{96}$. The flux within the disc can be obtained by directly integrating Equation 7.17,

$$\tilde{F} = \lambda \tilde{p}_0 \tilde{z} \left[1 - \frac{4}{3} \left(\frac{\tilde{z}}{\tilde{z}_s} \right)^2 + \frac{6}{5} \left(\frac{\tilde{z}}{\tilde{z}_s} \right)^4 - \frac{4}{7} \left(\frac{\tilde{z}}{\tilde{z}_s} \right)^6 + \frac{1}{9} \left(\frac{\tilde{z}}{\tilde{z}_s} \right)^8 \right] \quad , \tag{7.44}$$

where we have made use of the symmetry condition $\tilde{F}(0) = 0$.

The boundary condition at the upper disc surface $\tilde{F}(\tilde{z}_s) = \tilde{F}_{\rm rad}(\tilde{z}_s)$ can be satisfied by evaluating Equation 7.44 at $\tilde{z} = \tilde{z}_s$ and substituting into Equation 7.18. The boundary condition is satisfied provided that

$$\lambda = \frac{315}{64} \frac{\tilde{p}_0^3}{\tilde{\rho}_0^5 \tilde{z}_s^2} = \frac{6561}{1120} \quad . \tag{7.45}$$

This is higher than the pure gas as convective/turbulent mixing allows for more efficient heat transport near the disc midplane.

7.2.4 Heating function

In this section we derive the dimensionless heating rate $f_{\mathcal{H}}$ for an eccentric $\alpha-$disc. We will find that $f_{\mathcal{H}}$ can be written as a horizontal part depending only on the orbital geometry (an explicit function of E with coefficients involving $\alpha_{s,b}$) together with

$$2\alpha_s \left(\frac{\dot{H}}{H} \right)^2 + \left(\alpha_b - \frac{2}{3} \alpha_s \right) \left(\frac{\dot{J}}{J} + \frac{\dot{H}}{H} \right)^2 \quad . \tag{7.46}$$

The viscous stress tensor is given by

$$T^{ab} = 2\mu_s S^{ab} + (\mu_b - \frac{2}{3}\mu_s) \nabla_c u^c g^{ab} \quad , \tag{7.47}$$

where the rate-of-strain tensor is

$$S^{ab} = \frac{1}{2} \left(\nabla^a u^b + \nabla^b u^a \right) \tag{7.48}$$

and g^{ab} is the metric tensor.

The viscous heating rate per unit volume is given by

$$\mathcal{H} = T^{ab} S_{ab} = 2\mu_s S^{ab} S_{ab} + \left(\mu_b - \frac{2}{3} \mu_s \right) S^2 \tag{7.49}$$

where $S = S^a_a = \nabla_a u^a = \Delta + \frac{\dot{H}}{H}$ is the velocity divergence. Assuming an α viscosity (with $\omega_{\rm orb} = n$) and using the relationship $\mathcal{H} = f_{\mathcal{H}} n p_v$ we obtain for the dimensionless heating rate

$$f_{\mathcal{H}} = 2\alpha s^{ab} s_{ab} + \left(\alpha_b - \frac{2}{3}\alpha\right)\left(\frac{\dot{J}}{J} + \frac{\dot{H}}{H}\right)^2 \tag{7.50}$$

where we have introduced a dimensionless rate-of-strain tensor such that $S^{ab} = n s^{ab}$. It now remains to determine $s^{ab}s_{ab}$. This transforms as a scalar under change of variables so it is straightforward to take the expressions for the components of s^{ab} and s_{ab} in the (λ, ϕ) coordinate system from the appendix of Ogilvie (2001) and obtain

$$\begin{aligned} s^{ab}s_{ab} = \frac{(1+ec)^2}{8\left(1-e^2\right)^3\left[1+(e-\lambda e_\lambda)c - e\lambda\varpi_\lambda s\right]^2}\Bigg\{ & c^4e^2\left[e^2\left(8\lambda\varpi_\lambda^2+1\right) - 4e\lambda e_\lambda + 4\lambda e_\lambda^2\right] \\ & + 8c^3e\left[e^2\left(2\lambda\varpi_\lambda^2+1\right) - 3e\lambda e_\lambda + 2\lambda e_\lambda^2\right] + 2c^2[e^4\left(8\lambda\varpi_\lambda^2+1\right)s^2 \\ & - 4e^3 s(\lambda e_\lambda s + 3\lambda\varpi_\lambda) + e^2\left(4\lambda e_\lambda^2 s^2 + 8\lambda e_\lambda\lambda\varpi_\lambda s + 4\lambda\varpi_\lambda^2 + 11\right) - 22e\lambda e_\lambda + 8\lambda e_\lambda^2] \\ & + 8c\left[4e^3\lambda\varpi_\lambda^2 s^2 - 6e^2\lambda\varpi_\lambda s + e(2\lambda e_\lambda\lambda\varpi_\lambda s + 3) - 3\lambda e_\lambda\right] \\ & + e^4\left(8\lambda\varpi_\lambda^2+1\right)s^4 - 4e^3\lambda e_\lambda s^4 + 2e^2s^2\left(2\lambda e_\lambda^2 s^2 - 8\lambda e_\lambda\lambda\varpi_\lambda s + 8\lambda\varpi_\lambda^2 - 1\right) \\ & + 4es(\lambda e_\lambda s - 6\lambda\varpi_\lambda) + 8\lambda e_\lambda^2 s^2 + 9\Bigg\} + \left(\frac{\dot{H}}{H}\right)^2 \end{aligned} \tag{7.51}$$

where we have defined $c := \cos f$, $s := \sin f$ and the subscript λ denotes a derivative with respect to λ.

$s^{ab}s_{ab}$ in the (a, E) coordinate system can be obtained by substituting the appropriate expressions for λ, ϕ, λe_λ, $\lambda\varpi_\lambda$ using a symbolic algebra package such as Mathematica. The resulting expression is quite a bit longer than Equation 7.51 so we don't include it here.

7.3 Periodic attractor solution to the vertical structure equations

In our numerical setup we solve Equations 7.35 and 7.36 for the radiation-gas mixture and Equations 7.35 and 7.39 for a pure gas where $\Gamma_1 = \Gamma_3 = \gamma$.

We integrate these equations forward in eccentric anomaly using a 4th order Runge-Kutta method with a control on the relative error on H. We expect a 2π-periodic solution to act as an attractor or limit cycle for trajectories in phase space so that from our initial guess ($T = H = 1$, $\dot{H} = 0$ corresponding to the hydrostatic solution in the

reference circular disc) the solution should relax to this after sufficient time. Figure 7.1 shows the relaxation towards a periodic solution of one such disc.

The 2π-periodic solution generally involves a nonlinear forced breathing mode which is maintained against dissipation by periodic forcing by the variation of J (for a disc with $q \neq 0$) and the vertical gravity around the orbit. It corresponds to a solution which is steady but non-axisymmetric in an inertial frame of reference. Additional free oscillations of the breathing mode, which have been studied in circular discs (Lubow, 1981; Ogilvie, 2002b), can also exist in an eccentric disc; however, without an excitation mechanism we expect these to damp given sufficient time. In fact we find that if numerical errors are not controlled carefully these can excite the free mode so that instead of a periodic attractor there is a quasiperiodic attractor solution, being a superposition of the free and forced modes. If the numerical errors are controlled sufficiently the free mode will damp until it reaches the amplitude of the numerical errors.

We find that generically, if the relative errors are controlled, then the solution will relax to a periodic solution on the thermal timescale (see Figure 7.2), $1/\alpha_s$ in the dimensionless units. The resulting solution is insensitive to the initial guess used; although when e and q are very close to 1 (larger than the $e = q = 0.9$ typically considered in this thesis) starting far from the periodic solution leads to the solver taking prohibitively short step sizes to control the numerical error.

7.4 Thermal instability in eccentric discs

The thermal stability of a radiation pressure dominated disc is currently an open question. For a circular disc with an alpha model (Shakura & Sunyaev, 1973) of the viscous stress a radiation pressure dominated disc is thermally unstable when the stress scales with the total (radiation + gas) pressure (Shakura & Sunyaev, 1976; Pringle, 1976; Piran, 1978), but thermally stable when stress scales with gas pressure (Meier, 1979; Sakimoto & Coroniti, 1981). The $\alpha-$viscosity prescription is a model of the turbulent stress which arises from the action of the magnetorotational instability (MRI, Balbus & Hawley (1991); Hawley & Balbus (1991)). While early MRI simulations including thermal physics tended to support thermal stability (Hirose et al., 2009), more recent calculations (Jiang et al., 2013; Mishra et al., 2016; Ross et al., 2017), where the box size is large enough so that turbulent eddies are limit by compressibility rather than the box length, support stress scaling with total pressure and imply thermal instability in radiation pressure dominated discs. The thermal instability in a MRI

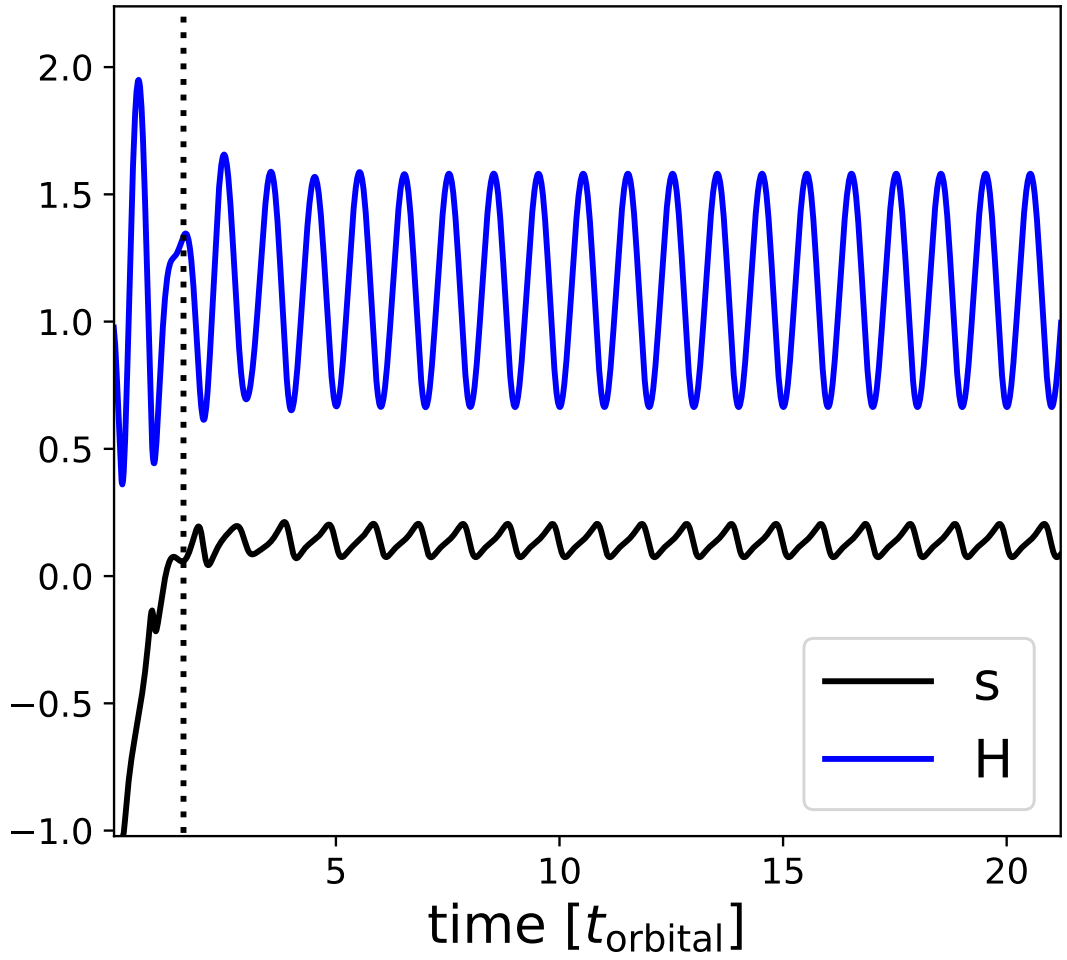


Fig. 7.1 Convergence of the scaleheight and entropy towards a periodic solution. The vertical dashed line is at one thermal timescale after the start. The parameters of this pure gas model are $\gamma = 5/3$, $\alpha_s = 0.1$, $\alpha_b = 0.0$, $e = 0.3$, $q = 0.3$, $E_0 = 0$.

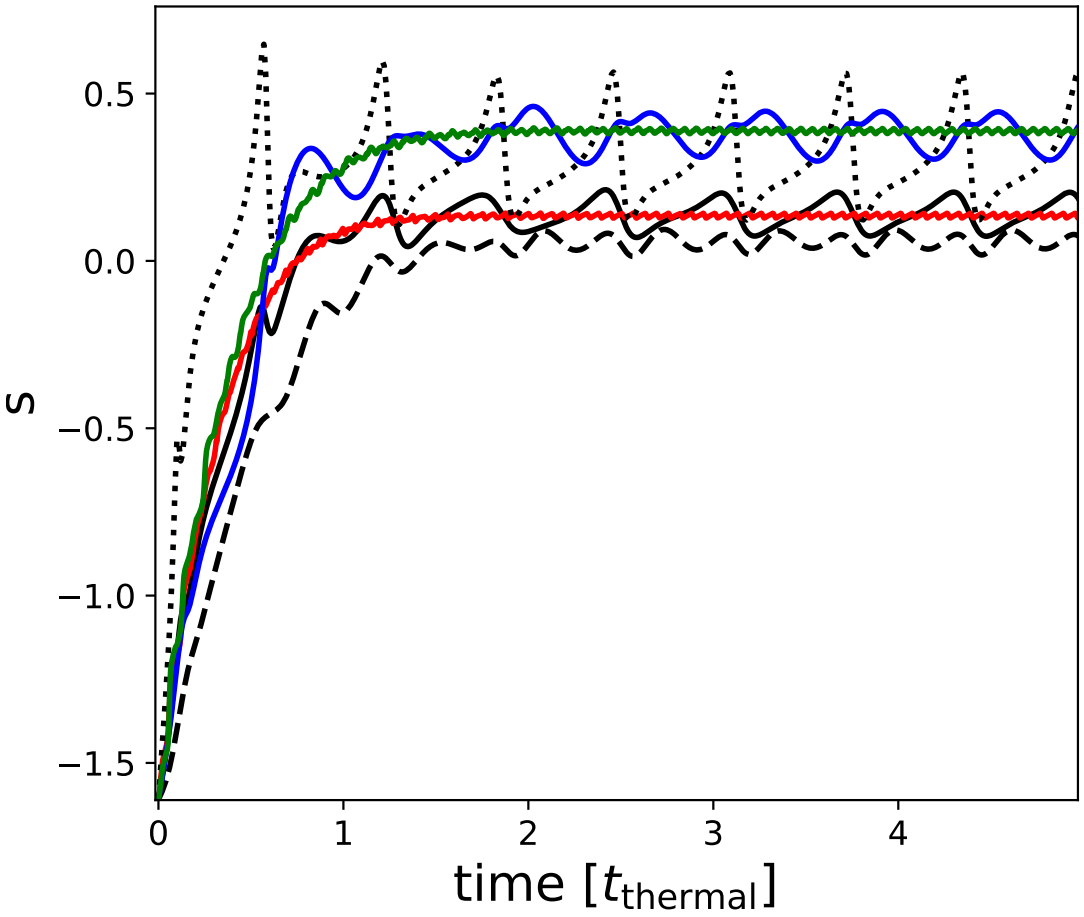


Fig. 7.2 Convergence of the specific entropy towards the periodic solution for different disc models. Time is in units of the thermal timescale for the corresponding disc model. Except where noted the disc parameters correspond to those in Figure 7.1. Black solid: same as Figure 7.1. Red solid: $\alpha_s = 0.01$. Black dashed: $e = 0.1$, $q = 0.5$. Black dotted: $e = 0.5$, $q = 0.1$. Blue solid: $\gamma = 1.4$. Green solid: $\gamma = 1.4$, $\alpha_s = 0.01$.

turbulent circular disc results in either thermal runaway (Jiang et al., 2013; Ross et al., 2017) or collapse to a thin gas pressure dominated disc (Jiang et al., 2013; Mishra et al., 2016).

The solution method described in Section 7.3 finds periodic solutions which are local attractors. This method guarantees (in principle nonlinear) stability of the solution to perturbations of H, $\dot{H}$ and T, which should damp on approximately the thermal time. A thermally unstable solution will not normally be found by our solution method, as the time for the free mode to damp out (so that the solution becomes $\sim 2\pi$-periodic) will be comparable to the time for the onset of thermal instability. Instead the solution will either transition to a thermally stable branch (if it exists) or experience unbound thermal runaway.

Figure 7.3 confirms that a circular radiation pressure dominated disc with $p_v = p$ experiences a thermal instability causing an unbound thermal runaway. This results carries over into the eccentric disc (here with $e = q = 0.3$), but requires a much lower β_r° as the eccentric disc is always hotter/more radiation pressure dominated than its circular counterpart ($\beta_r > \beta_r^\circ$). Surprisingly however the perfect gas with $\gamma = 4/3$, which is often used as a model of an (adiabatic) radiation pressure dominated gas, is thermally stable. We shall explore this result further in section 7.5.

As noted above a circular radiation pressure dominated gas with $p_v = p_g$ is thermally stable (Meier, 1979; Sakimoto & Coroniti, 1981). For circular and moderately eccentric radiation-pressure dominated discs, having stress scale with gas pressure does prevent the thermal instability. However, Figure 7.4 shows that for a sufficiently eccentric radiation-pressured dominated disc the disc experiences a thermal instability despite $p_v = p_g$. The thermal instability is much weaker than the $p_v = p$ case. Instead of an unbound thermal runaway, the solution appears to transition from a thermally unstable "gas pressure dominated" branch to a thermally stable radiation pressure dominated branch. This thermal instability in the highly eccentric discs will be discussed further in Chapter 11 in the context of TDE discs.

7.5 When is a $\gamma = 4/3$ perfect gas a good model of a radiation-dominated disc?

When modelling a radiation-dominated disc a common approximation is to model the radiation-gas mixture as a $\gamma = 4/3$ perfect gas (e.g. Loeb & Ulmer (1997), Lu & Bonnerot (2019), Bonnerot & Lu (2019) and Zanazzi & Ogilvie (2020) for radiation dominated TDE discs). But, in the preceding section, we found that the $\gamma = 4/3$

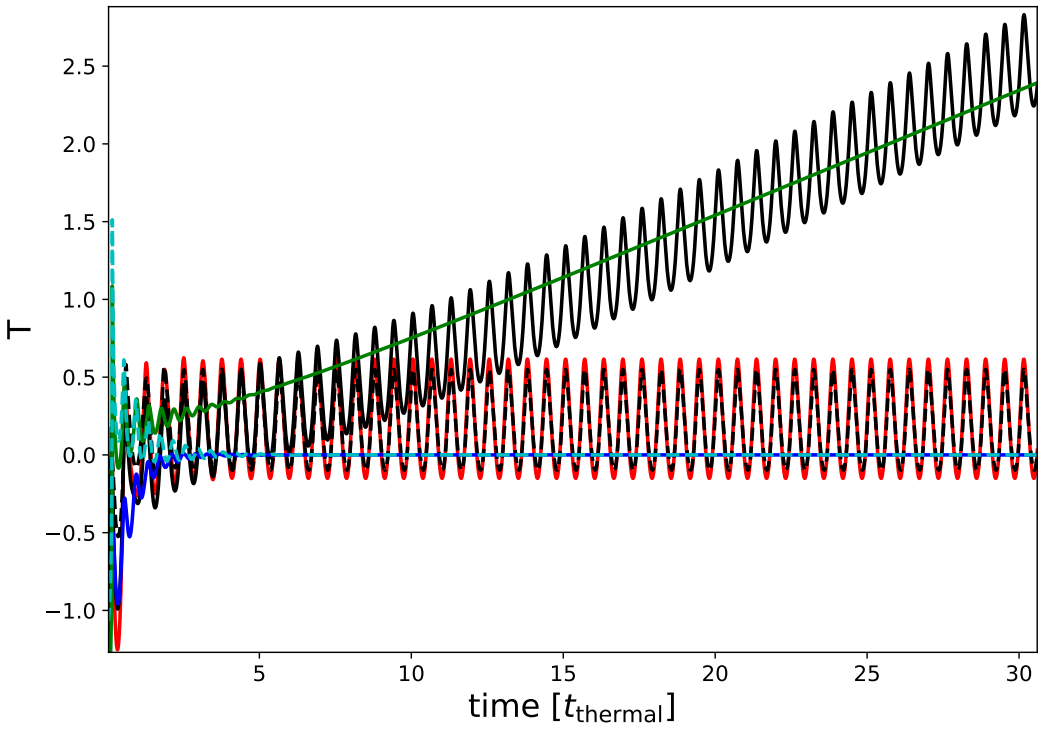


Fig. 7.3 Convergence of the temperature, or emergence of the thermal instability, for different disc models. Time is in units of the thermal timescale for the corresponding disc model. All models have parameters $\alpha_s = 0.1$, $\alpha_b = 0.0$, $E_0 = 0$ and the radiation-gas mixtures have $\gamma = 5/3$ and $p_v = p$. Black dashed: perfect gas with $\gamma = 4/3$ and $e = q = 0.3$. Red solid: radiation-gas mixture with $\beta_r^\circ = 10^{-3}$ and $e = q = 0.3$. Black solid: radiation-gas mixture with $\beta_r^\circ = 10^{-1}$ and $e = q = 0.3$. Cyan dashed: perfect gas with $\gamma = 4/3$ and $e = q = 0$. Blue solid: radiation-gas mixture with $\beta_r^\circ = 10^{-1}$ and $e = q = 0$. Green solid: radiation-gas mixture with $\beta_r^\circ = 10$ and $e = q = 0$.

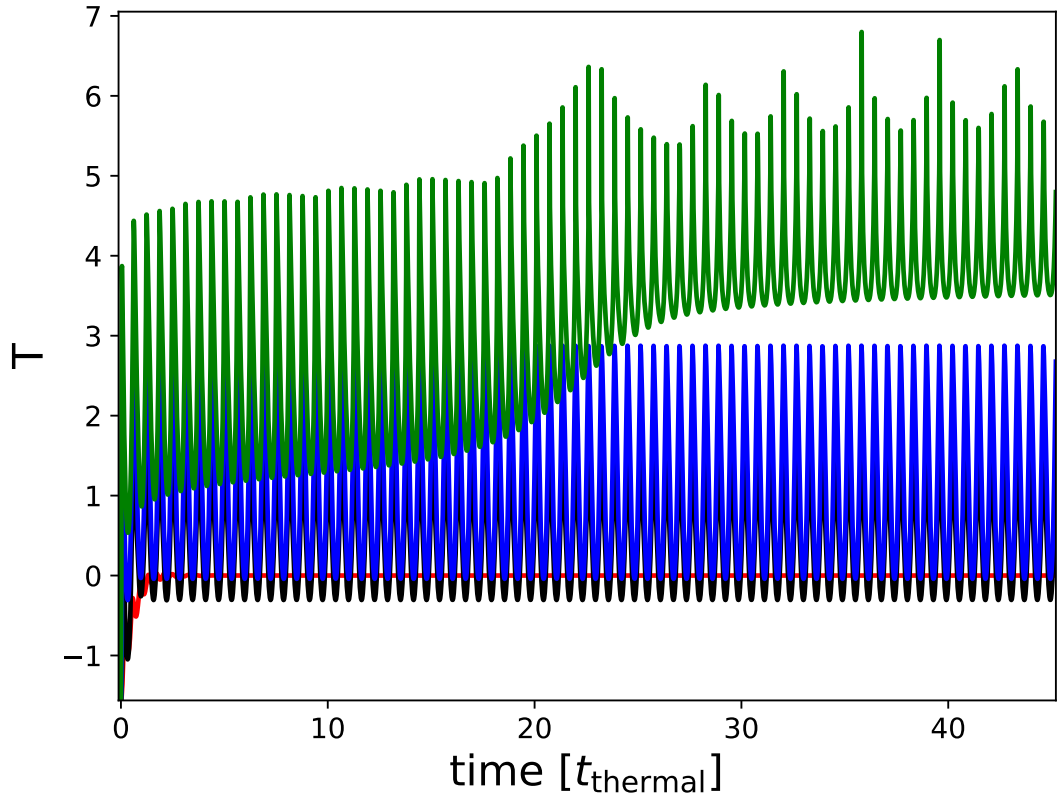


Fig. 7.4 Convergence of the temperature, or emergence of the thermal instability, for a radiation-gas mixture with $p_v = p_g$. The disc parameters are $\alpha_s = 0.1$, $\alpha_b = 0.0$, $\beta_r^\circ = 10^{-3}$ and $E_0 = 0$. Red solid line: $e = q = 0$. Black dashed: $e = q = 0.5$. Blue solid: $e = q = 0.75$. Green solid: $e = q = 0.9$.

perfect gas model of an eccentric disc behaves differently from the radiation-dominated models; notably the $\gamma = 4/3$ perfect gas model does not experience an unbound thermal runaway present in the radiation-pressure dominated disc. So we now consider under what circumstances a $\gamma = 4/3$ gas behaves like a (radiation pressure dominated) radiation-gas mixture. Generally the $\gamma = 4/3$ perfect gas is used to model an adiabatic radiation pressure dominated limit so in what follows we consider a nearly adiabatic disc (i.e. $\alpha_s, \alpha_b \ll 1$).

First consider a perfect gas with $\gamma = 4/3$ and mean molecular weight μ_1. The equation of state is

$$p = (\mathcal{R}/\mu_1)\rho T \quad . \tag{7.52}$$

When this undergoes a nearly adiabatic change we have

$$p = K_1 \rho^{4/3} \quad , \tag{7.53}$$

where K_1 is a slowly varying function. We can rearrange equations 7.52 and 7.53 to obtain

$$p = K_1^{-3} (\mathcal{R}/\mu_1)^4 T^4 \tag{7.54}$$

and

$$T^3 = K_1^3 (\mu_1/\mathcal{R})^3 \rho \tag{7.55}$$

Now consider instead a mixture of a perfect gas (with mean molecular weight μ_2) and black body radiation. In the radiation-dominated limit we have the equation of state

$$p = \frac{4\sigma}{3c} T^4 \quad . \tag{7.56}$$

When this undergoes a nearly adiabatic change we have

$$T^3 = K_2 \rho \quad , \tag{7.57}$$

where K_2 is a slowly varying function. We can rearrange equations 7.56 and 7.57 to obtain

$$p = \frac{4\sigma}{3c} K_2 \rho T \quad . \tag{7.58}$$

These two systems are equivalent provided that

$$K_1^3 = K_2(\mathcal{R}/\mu_1)^3 \tag{7.59}$$

and

$$\frac{4\sigma}{3c}K_2 = (\mathcal{R}/\mu_1) \quad . \tag{7.60}$$

The first of these conditions provides a relationship between the entropies of the two systems. The second condition is more interesting. Noting the correspondence between Equations 7.57 and 7.9, we can rewrite Equation 7.60 in terms of β_r,

$$\frac{\mu_1}{\mu_2} = \beta_r^{-1} \quad . \tag{7.61}$$

In the radiation-dominated limit the molecular weight of the $\gamma = 4/3$ perfect gas is much lower than that of the gas in the radiation dominated disc and will correspond to some suitably weighted average of the massless photons and massive ions and electrons. As the disc becomes increasingly radiation dominated the ratio of photons to massive particles increases resulting in a decrease in the effective molecular weight of the $\gamma = 4/3$ gas.

The molecular weight doesn't appear in the equations describing the dynamical vertical structure, so has no direct influence on the dynamics. Instead changing the molecular weight changes the reference state of the disc. In other words, the circular reference discs for the $\gamma = 4/3$ perfect gas disc and the radiation dominated disc are different in general. For a nearly adiabatic disc where β_r is approximately constant we can rescale the reference state so that the $\gamma = 4/3$ perfect gas solution agrees with the nearly adiabatic radiation-dominated solution (see Figure 7.5).

A non-adiabatic change in the radiation-dominated gas causes a change in β_r (as the entropy $s \approx 4\beta_r$). This corresponds to a change in the molecular weight (i.e. a compositional change) in the $\gamma = 4/3$ perfect gas. In other words non-adiabatic heating and cooling in the radiation-dominated model are equivalent to a chemical reaction in the $\gamma = 4/3$ perfect gas model. This can be understood as the emission/absorption of photons by the gas resulting in a change in the ratio of photons to massive particles and, consequently, a change in the effective molecular weight μ_1.

This then explains the absence of the thermal instability in the $\gamma = 4/3$ perfect gas model. When the radiation-dominated disc undergoes thermal runaway the entropy (and thus β_r) increases with each orbit. In the $\gamma = 4/3$ disc this would correspond to a reduction in μ_1 with each orbit. However, as our equations do not allow for the

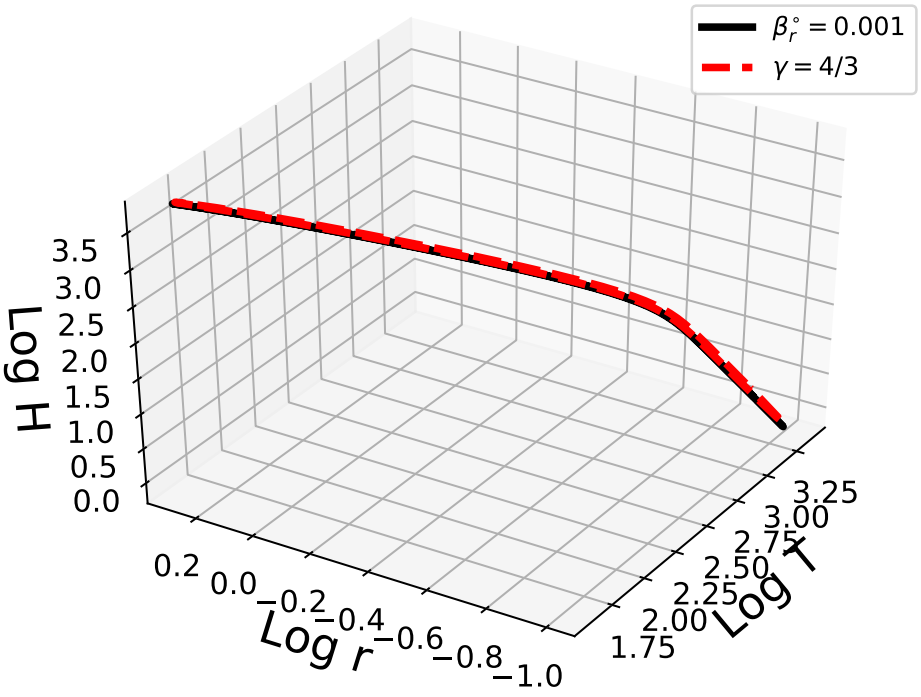


Fig. 7.5 Comparison of a $\gamma = 4/3$ perfect gas model with $\alpha_s = 10^{-6}$, $\alpha_b = 0$ $e = q = 0.9$ and $E_0 = 0$ against a radiation-gas mixture with $\beta_r^\circ = 0.001$, $\alpha_s = 0.1$, $\alpha_b = 0$, $e = q = 0.9$, $E_0 = 0$ and stress scaling with gas pressure. Note that this solution is radiation dominated as the resulting $\beta_r \approx 10^5$. The temperature and scale height of the $\gamma = 4/3$ perfect gas model have been rescaled by a constant factor - this rescaling is equivalent to a change in the reference disc by selecting K_1 and μ_1 so that the two solutions agree.

evolution of μ_1, the effective β_r of the $\gamma = 4/3$ disc is fixed preventing a thermal runaway.

Most simulations use a caloric equation of state expressing internal energy $e = e(p, \rho)$ in terms of pressure and density. The thermal energy equation is then solved using the internal energy so the simulation makes no reference to either the temperature or the molecular weight, making the above a technicality. Temperatures calculated as a post-processing step using the correct equation of state (e.g. as done in Bonnerot & Lu (2019) and Zanazzi & Ogilvie (2020)) will also be correct. However, explicit cooling terms typically make reference to the temperature so the results of this section will be relevant.

7.6 Conclusion

In this section we looked at laminar flows in an eccentric disc in the presence of viscous dissipation and radiative cooling. In a Lagrangian frame, these laminar flows take the form of a periodic, homogeneously expanding/contracting fluid column. We found that, like circular discs, a radiation pressure dominated eccentric $\alpha-$disc is thermally unstable when stress scales with total pressure. However, unlike circular discs, sufficiently eccentric, radiation pressure dominated, $\alpha-$discs can be thermally unstable when stress scales with gas pressure. Lastly, care must be taken when modelling radiation pressure dominated discs as a $\gamma = 4/3$ perfect gas does not feature the thermal instability.

Chapter 8

Generation and Damping of Tightly Wound Waves

8.1 Introduction

The short wavelength theory developed in Chapter 6 was an ideal theory of a steady eccentric wave, which didn't allow for the excitation or damping of the wave. In this Chapter we generalise the tightly wound theory to allow for the time evolution of the wave and non-conservative terms in the form of shear and bulk viscosity. In principle we could generalise the untwisted theory to include these effects as well; however it's unlikely that this would be useful as one of the effects of viscosity is to introduce a twist in the disc. Additionally an analysis of the Hamiltonian structure of the ideal disc shows that disc twist is a requirement for evolution of the disc eccentricity. As discussed in Chapter 6 a major motivation for developing the short wavelength theory was to provide a nonlinear generalisation of the work of Ferreira & Ogilvie (2008). As the inner regions of black hole discs are likely MRI turbulent, Ferreira & Ogilvie (2008) considered a model where the eccentric wave was damped (by a process modelled using a bulk viscosity) as it travels through the disc. In this Chapter we provide a nonlinear generalisation of Ferreira & Ogilvie (2008) for purely ingoing tightly wound waves, which will be useful for understanding the simulations of Dewberry et al. (2020a) and Dewberry et al. (2020b).

In this Chapter we also apply the theory of Whitham (1965) to determine the behaviour of tightly wound eccentric waves, allowing for time evolution and the presence of non-conservative terms. It is possible to derive a more general short-wavelength ($q^2 = (ae_a)^2 + (ae\varpi_a)^2 \sim 1$, $e \ll 1$) limit which allows for twisted (but not necessarily tightly wound) discs, where the untwisted eccentric waves ($ae\varpi_a = 0$) and the tightly

wound waves($ae_a = 0$) are obtainable as separate limits. However the general twisted case is substantially more complicated and requires further work to understand fully, so I have not included it in this Thesis.

8.2 The time dependent short-wavelength Lagrangian

As discussed in Chapter 6 the Lagrangian for an eccentric wave can be written as $L = L_T + L_M$, where L_M describes an eccentric wave which is steady in a rotating frame and L_T contains the term responsible for time evolution (in that precessing frame). In Chapter 6 we neglected L_T and derived a short-wavelength theory of an eccentric wave steady in a frame rotating with angular frequency ω. In this Chapter we restore L_T and allow the eccentric wave to evolve in the rotating frame. Repeated from Chapter 6, L_T and L_M are defined as follows:

$$L_T = \int M_a n a^2 \tilde{\varpi}_t \sqrt{1-e^2}\, da \quad , \tag{8.1}$$

$$L_M = \int \left[M_a n a^2 \left(\omega\sqrt{1-e^2} + \int \frac{e\omega_f(a,e)}{\sqrt{1-e^2}}\, de \right) - H_a^\circ F \right] da \quad , \tag{8.2}$$

with $L = L_T + L_M$. We have introduced $\tilde{\varpi} := \varpi - \omega t$. L_M is the Lagrangian for an eccentric wave which is steady in a frame precessing at rate ω which was considered in Chapter 6. As in Chapter 6 the precessional forces coming from H_a^f can be expressed in terms of the precession rate $\omega_f(a,e)$ of a test particle.

We now consider the situation where the precessional (or test particle) forces are strong relative to the background pressure gradient. As per Chapter 6 we define

$$N^2 = \frac{M_a n a^2 (\omega_f(a,0) - \omega)}{H_a^\circ} \quad , \tag{8.3}$$

with the limit of interest being $N^2 \gg 1$.

We introduce the rapidly varying phase variable $\varphi = \varphi(a,t)$ which satisfies $a\varphi_a = N(a)k(a)$ with $k(a) = O(1)$ a rescaled dimensionless wavenumber. We also make use of the rescaled variable $\tilde{e} = eNk$. For a twisted disc the nonlinearity becomes

$$\begin{aligned} q^2 &= \left(a\frac{de}{da} \right)^2 + a^2 e^2 \left(\frac{d\varpi}{da} \right)^2 \\ &\approx \tilde{e}_\varphi^2(\varphi,a) + \tilde{e}^2(\varphi,a)\varpi_\varphi^2(\varphi,a) \quad . \end{aligned} \tag{8.4}$$

which is of order unity.

Substituting this into the Lagrangian and keeping only terms of the lowest order in N^{-1}, the Lagrangian becomes

$$\begin{aligned} L &\approx \int \left(-\frac{M_a n a^2}{N^2} \tilde{\varpi}_t \frac{\tilde{e}^2}{2k^2} + H_a^\circ \hat{\mathcal{L}} \right) da \\ &= \int -\frac{M_a n a^2}{N^2} \tilde{\varpi}_t \frac{\tilde{e}^2}{2k^2} + H_a^\circ \left(\frac{1}{2} \frac{\tilde{e}^2}{k^2} - F(q) \right) da \quad , \end{aligned} \tag{8.5}$$

where we have neglected two leading order terms; $M_a n a^2 \omega$ which is a function only of a that has no effect on the dynamics and $M_a n a^2 \tilde{\varpi}_t$. It is possible to show that the latter term only contributes when $M_a n a^2$ is time dependent, a situation we shall not consider in this chapter. The neglected $M_a n a^2 \tilde{\varpi}_t$ could be important for a theory that allowed for the evolution of the background surface density profile. Finally, as in Chapter 6, we have separated L_M into a dimensionless part $\hat{\mathcal{L}}$ which controls the small scale dynamics and a dimensional part H_a° which is important in determining how the solution varies on the length scale of the disc.

8.3 Conservative dynamics of tightly wound waves

Initially we neglect non-conservative terms and derive the equations for a ideal, time dependent, tightly wound wave. As the equations are fully conservative this can be done entirely within the framework of Lagrangian mechanics and the averaged Lagrangian theory of Whitham (1965). As noted in Chapter 6 the short lengthscale dynamics of a tightly wound wave are trivial, with ϖ linearly dependent on wave phase and all other quantities constant. It is thus straightforward to average Equation 8.5. Thus, the Whitham averaged Lagrangian density for a tightly wound eccentric wave with self-gravity and pressure is

$$\mathcal{L}_a = -\frac{M_a n a^2}{N^2} \varphi_t \frac{\tilde{e} q_+}{2k^2} + H_a^\circ \left(\frac{1}{2} \frac{\tilde{e}^2}{k^2} - F(q_+) \right) + \frac{1}{4} G M_a^2 q_+ \mathcal{W}(q_+) \frac{\tilde{e}}{kN}, \tag{8.6}$$

where we have made use of $\tilde{\varpi}_t \tilde{e} = \varphi_t \tilde{\varpi}_\varphi \tilde{e} = \varphi_t q_+$. The last term in the Lagrangian, incorporating self gravity, was taken from Lee & Goodman (1999) and rewritten in our variables. It is derived assuming that the gravitational self energy of a tightly wound wave is well approximated by that of neighbouring thin wires and that on the

lengthscale of the wave the effects of the orbital curvature are negligible. Under these assumptions this term shouldn't be modified by the inclusion of the vertical variation of the disc. It is convenient to write this self gravity term in terms of the aspect ratio, h°, and Toomre-Q, Q°, of the reference circular disc,

$$\mathcal{L}_a = -\frac{M_a n a^2}{N^2}\varphi_t \frac{\tilde{e} q_+}{2k^2} + H_a^\circ \left(\frac{1}{2}\frac{\tilde{e}^2}{k^2} - F(q_+) + \frac{\gamma}{2Q^\circ h^\circ} q_+ \mathcal{W}(q_+) \frac{\tilde{e}}{kN} \right), \tag{8.7}$$

where $Q^\circ = \frac{c_s^\circ n}{G\Sigma^\circ \pi}$, $h^\circ = \frac{H^\circ}{a}$, with c_s° and H° the sound speed and scale height of the reference circular disc respectively.

The function $\mathcal{W}(q_+)$ is given by (Lee & Goodman, 1999),

$$\mathcal{W}(q_+) = -\frac{4}{\pi q_+}\int_{-\infty}^{\infty} dy \ln \left| \frac{1}{2} + \frac{1}{2}\sqrt{1 - \left(\frac{\sin(q_+ y)}{y}\right)^2} \right| \quad , \tag{8.8}$$

is a monotonically increasing function with $\mathcal{W}(q_+) = 1 + \frac{1}{4}q_+^2 + O(q_+^4)$. $\mathcal{W}(q_+)$ is finite but nondifferentiable when $q_+ = 1$, with $\lim_{q_+ \to 1} \mathcal{W}'(q_+) \propto \ln(1 - q_0)$.

The dispersion relation is obtained by varying the Lagrangian with respect to the wave amplitudes $\tilde{e}$, q_+

$$\frac{\delta \mathcal{L}}{\delta \tilde{e}} = 0, \quad \frac{\delta \mathcal{L}}{\delta q_+} = 0 \quad , \tag{8.9}$$

combined with the compatibility relation

$$N\partial_t k = a\partial_a \varphi_t \quad . \tag{8.10}$$

At next order the evolution of the wave fluxes are given by,

$$\frac{\delta \mathcal{L}}{\delta \varphi} = 0 \quad . \tag{8.11}$$

The averaged Lagrangian has the following gauge symmetry

$$k \to \lambda(a,t)k \quad , \qquad \varphi_t \to \lambda(a,t)\varphi \quad , \qquad \tilde{e} \to \lambda(a,t)\tilde{e} \tag{8.12}$$

where $\lambda(a,t)$ is an arbitrary function. As discussed in Chapter 6, this relates to our freedom to redefine the wave phase owing to the fact $\tilde{e}$ and q do not vary on the short length/timescale. It is useful to gauge fix out this symmetry by setting $q_+ = \tilde{e}$. Upon adopting this gauge the Lagrangian becomes

$$\mathcal{L}_a = -\frac{M_a n a^2}{N^2}\varphi_t \frac{q_+^2}{2k^2} + H_a^\circ \left(\frac{1}{2}\frac{q_+^2}{k^2} - F(q_+) + \frac{\gamma q_+^2 \mathcal{W}(q_+)}{2Q^\circ h^\circ k N} \right) \quad , \tag{8.13}$$

with the dispersion relation now just given by variation with q_+,

$$\frac{\delta \mathcal{L}}{\delta q_+} = 0 \quad , \tag{8.14}$$

while the compatibility relation and variational equation for the wave flux evolution are the same. The dispersion relation is given by

$$0 = -\frac{M_a n a^2}{N^2}\varphi_t \frac{q_+^2}{k^2} + H_a^\circ \left(\frac{q_+^2}{k^2} - q_+ F'(q_+) + \frac{\gamma q_+^2}{2Q^\circ h^\circ k N}(2\mathcal{W}(q_+) + q_+ \mathcal{W}'(q_+)) \right) \quad . \tag{8.15}$$

Substituting this into the compatibility relation we obtain an equation in conservative form

$$\frac{\partial k}{\partial t} - \frac{\partial}{\partial y}\left\{ (\omega_f - \omega)\left[1 - k^2 q_+^{-1} F'(q_+) + \frac{\gamma k}{2Q^\circ h^\circ N}(2\mathcal{W}(q_+) + q_+ \mathcal{W}'(q_+)) \right] \right\} = 0 \quad , \tag{8.16}$$

where $y = \int \frac{N}{a} da$. The equation for the wave flux evolution is

$$0 = \partial_t \left(\frac{M_a n a^2}{N^2}\frac{q_+^2}{2k^2} \right) + \partial_a \frac{a H_a^\circ}{N} \left(\frac{q_+ F'(q_+)}{k} - \frac{\gamma}{2Q^\circ h^\circ N}\frac{q_+^2}{k^2}\frac{\partial}{\partial q_+}\left[q_+ \mathcal{W}(q_+) \right] \right) \quad , \tag{8.17}$$

where we have made use of the dispersion relation to simplify the spacial component of the flux. Consider steady solutions to these equations. Without loss of generality we can set $\varphi_t = 0$ (φ_t must be constant so that $\partial_t k = 0$ and can be set to zero by a suitable choice of rotating frame). Then the equations become

$$1 + \frac{\gamma k}{2Q^\circ h^\circ N}(2\mathcal{W}(q_+) + q_+ \mathcal{W}'(q_+)) = k^2 q_+^{-1} F'(q_+) \tag{8.18}$$

$$\frac{a H_a^\circ}{N}\left(\frac{q_+ F'(q_+)}{k} - \frac{\gamma}{2Q^\circ h^\circ N}\frac{q_+^2}{k^2}\frac{\partial}{\partial q_+}\left[q_+ \mathcal{W}(q_+) \right] \right) = \text{const} \quad . \tag{8.19}$$

Here Equation 8.18 is an amplitude dependent dispersion relation, while Equation 8.19 describes the conservation of AMD flux in the disc. When combined these equations determine how k and q_+ vary with a for a given disc profile.

Solving the dispersion relation for k,

$$k = \frac{\gamma}{4Q^{\circ}h^{\circ}N}\chi(q_+) + \sqrt{\frac{q_+}{F'(q_+)} + \frac{\gamma^2}{16(Q^{\circ}h^{\circ}N)^2}\chi^2(q_+)} \quad , \tag{8.20}$$

where we have introduced a self-gravity nonlinear control parameter

$$\chi(q_+) = \frac{q_+}{F'(q_+)}(2\mathcal{W}(q_+) + q_+\mathcal{W}'(q_+)) \quad , \tag{8.21}$$

which determines whether pressure or self gravity becomes more important as the wave becomes nonlinear. The relative importance of self gravity to pressure forces is then set by a combination of the Toomre-Q and aspect ratio of the reference disc, the relative strength of the forced precession to the background pressure gradient (N) and the nonlinear control parameter $\chi(q_+)$. Figure 8.1 shows how this control parameter changes with q_+ and the equation of state/disc model. Generally as the waves become increasingly nonlinear, pressure increasingly dominates over self gravity. This can be understood by realising that the $q \to 1$ limit should match onto a spiral shock wave. Compared to the 2D discs considered in Lee et al. (2018) the self gravity in 3D discs, with the same ratio of specific heats γ, is stronger due to the increased compressibility.

8.4 Non-conservative terms

In order to include non-conservative terms in our theory we need to move beyond the Lagrangian framework. In Chapter 4 we reviewed the pseudo-Lagrangian framework and the extension of the Whitham averaging method to pseudo-Lagrangians by Jimenez & Whitham (1976) which we shall need in this Chapter.

There is an unfortunate notation clash between the usual notation for complementary phase space variables (i.e. $'$) and the notation for a derivative for a function of one variable, which we have made extensive use of to denote q derivatives of the geometric part of the Hamiltonian F. To avoid this we shall adopt the convention in this chapter that complementary phase space variables are denoted by a breve ($\breve{\cdot}$).

8.4.1 Viscous stresses

In this section we consider the effects of viscous stresses on tightly wound waves. We assume the vertically integrated shear viscosity $\bar{\mu}$ and bulk viscosity $\bar{\mu}_b$ can be written in the form $\bar{\mu} = \frac{\alpha\mu_a^{\circ}}{2\pi}m$ and $\bar{\mu}_b = \frac{\alpha_b\mu_a^{\circ}}{2\pi}m$, where α, α_b are dimensionless constants, μ_a° depends only on a and m is a function of j and h only.

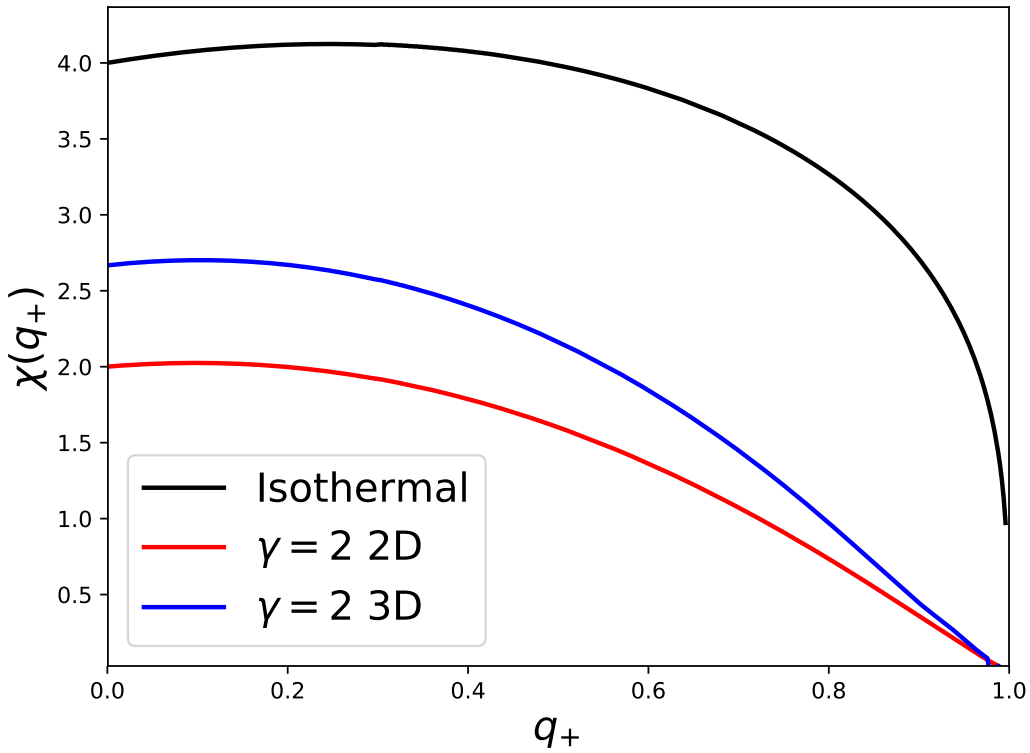


Fig. 8.1 Dependence of the Nonlinear Self-Gravity Control parameter on q_+ for different equations of state. Generally the as the wave becomes nonlinear self gravity becomes less important. Additionally more compressible equations of state/discs have stronger self gravity.

In the $\tilde{e} = q_+$ gauge, dissipative effects from viscous stresses in the tight-winding limit can be included using a pseudo-Lagrangian of the following form

$$\begin{aligned} F = & \frac{2}{3}\alpha \int a T_a^\circ (\varphi - \breve{\varphi}) F_{\rm shear}(\breve{q})\, da \\ & + \left(\alpha_b - \frac{2}{3}\alpha\right) \int a T_a^\circ (\varphi - \breve{\varphi}) F_{\rm bulk}(\breve{q})\, da \quad . \end{aligned} \tag{8.22}$$

which is derived in Appendix B. Here we have introduced $T_a^\circ = n\mu_a^\circ$ and

$$F_{\rm bulk}(q) := \frac{q}{2\pi} \int \tau(q, \cos E) \sin^2 E \, dE \quad , \tag{8.23}$$

$$F_{\rm shear}(q) := \frac{3q}{2\pi} \int m(q, \cos E) j^{-1} \left[2\cos E \, \left(\frac{1}{4} - j\right) + q \sin^2 E \right] dE \quad , \tag{8.24}$$

with

$$\tau(q, \cos E) := m(q, \cos E) \left(\frac{q}{j} + \frac{1}{h} \frac{\partial h}{\partial \cos E} \right) \quad . \tag{8.25}$$

Assuming that the viscosity can be written as a powerlaw in density, pressure and scale height then we have a functional form form for m,

$$m = j^x h^y \quad , \tag{8.26}$$

where for a standard $\alpha-$viscosity $x = -\gamma$, $y = -(\gamma - 1)$. In the linear limit $q_+ \to 0$ then $F_{\rm bulk}$ and $F_{\rm shear}$ have the following limits for a 2D disc (with $y = 0$)

$$\lim_{q\to 0} F_{\rm bulk}^{(2D)}(q) = \frac{q^2}{2} \quad , \tag{8.27}$$

$$\lim_{q\to 0} F_{\rm shear}^{(2D)}(q) = \frac{9}{4} q^2 (1 + x) \quad , \tag{8.28}$$

while in a 3D disc,

$$\lim_{q\to 0} F_{\rm bulk}^{(3D)}(q) = \frac{2\gamma - 1}{2\gamma} q^2 \quad , \tag{8.29}$$

$$\lim_{q\to 0} F_{\rm shear}^{(3D)}(q) = \frac{9}{4} q^2 \left(1 + x - y \frac{\gamma - 1}{\gamma} \right) \quad . \tag{8.30}$$

In order for shear viscosity to be dissipative then $F_{\rm shear} > F_{\rm bulk}$. For an $\alpha-$disc with $x = -\gamma$ and $y = -(\gamma-1)$ where $1 \leq \gamma \leq 2$ then both $F_{\rm shear}^{(2D)} < F_{\rm bulk}^{(2D)}$ and $F_{\rm shear}^{(3D)} < F_{\rm bulk}^{(3D)}$ in the limit $q \to 0$ so shear viscosity is anti-dissipative in the linear limit, recovering the known viscous overstability (Syer & Clarke, 1992, 1993; Lyubarskij et al., 1994; Ogilvie, 2001; Latter & Ogilvie, 2006). More generally in order that viscosity behaves dissipatively we require

$$\frac{2}{3}\alpha(F_{\rm shear}(q) - F_{\rm bulk}(q)) + \alpha_b F_{\rm bulk}(q) > 0 \tag{8.31}$$

In the linear limit, for a 2D disc this condition becomes

$$\begin{aligned}\lim_{q\to 0}\left[\frac{2}{3}\alpha(F_{\rm shear}^{(2D)}(q) - F_{\rm bulk}^{(2D)}(q)) + \alpha_b F_{\rm bulk}^{(2D)}(q)\right] \\ = \frac{q^2}{6}\alpha\left[7 + 3\frac{\alpha_b}{\alpha} + 9x\right] \quad ,\end{aligned} \tag{8.32}$$

while for a 3D disc,

$$\begin{aligned}\lim_{q\to 0}\left[\frac{2}{3}\alpha(F_{\rm shear}^{(3D)}(q) - F_{\rm bulk}^{(3D)}(q)) + \alpha_b F_{\rm bulk}^{(3D)}(q)\right] \\ = \frac{\alpha}{6}q^2\left[5 + 9x - 9y\frac{\gamma-1}{\gamma} + \frac{2}{\gamma} + 6\frac{\alpha_b}{\alpha}\left(1 - \frac{1}{2\gamma}\right)\right] \quad .\end{aligned} \tag{8.33}$$

For a 2D $\alpha-$disc the viscous overstability is suppressed when

$$\frac{\alpha_b}{\alpha} > \frac{9\gamma - 7}{3} \quad , \tag{8.34}$$

which agrees with the stability criterion given in Latter & Ogilvie (2006). For a 3D $\alpha-$disc the viscous overstability is suppressed when

$$\frac{\alpha_b}{\alpha} > \frac{13\gamma - 11}{6\gamma - 3} \quad . \tag{8.35}$$

This expression is different from the condition derived from Appendix B of Latter & Ogilvie (2006) in the limit $\alpha_b, \alpha \ll 1$, even when radiative cooling and viscous heating, in their model, are neglected. This appears to be due to the differing treatments of the vertical structure. Specifically, the vertical structure of the wave in Latter & Ogilvie (2006) is affected by the viscosity. In our model the vertical structure is not modified by the viscosity as it is assumed that this only modifies the solution over many wave cycles.

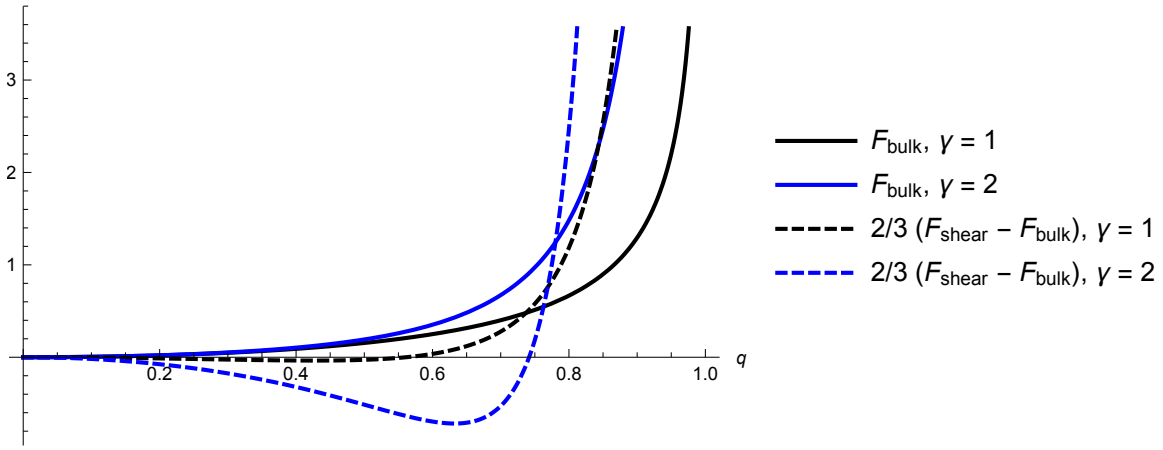


Fig. 8.2 Viscous integrals for a 2D $\gamma = 1$ and $\gamma = 2$ discs, which are expected to bracket the behaviour of 2D and 3D discs with $1 \leq \gamma \leq 2$. Negative values indicates anti-dissipative/overstable behaviour. The shear viscosity is initially antidissipative/overstable, but becomes dissipative at a critical q which depends on the disc model. The same qualitative behaviour has long been known in planetary rings (Borderies et al., 1985, 1986).

In the linear limit the overstability in the 3D disc stabilises at lower α_b than it does in the 2D disc. This differs from the findings of Latter & Ogilvie (2006) who included radiative cooling and viscous heating in their 3D discs (but not their 2D discs). This is, however, consistent with the expectation from the planetary ring literature that more compressible fluids are less susceptible to the viscous overstability (Borderies et al., 1985); as was discussed in Section 8.3, a 3D disc is more compressible than the 2D disc with the same γ.

8.5 Tight winding wave dynamics

The Whitham averaged pseudo-Lagrangian density for a tightly wound eccentric wave with self-gravity, pressure and viscosity, in the $\tilde{e} = q_+$ gauge, is:

$$\begin{aligned}\mathcal{F}_a = &-\frac{M_a n a^2}{N^2}\varphi_t \frac{q_+^2}{2k^2} + H_a^\circ\left(\frac{1}{2}\frac{q_+^2}{k^2} - F(q_+) + \frac{\gamma q_+^2 \mathcal{W}(q_+)}{2Q^\circ h^\circ k N}\right)\\ &+ \frac{2}{3}\alpha a T_a^\circ(\varphi - \breve{\varphi})(F_{\text{shear}}(\breve{q}_+) - F_{\text{bulk}}(\breve{q}_+))\\ &+ \alpha_b a T_a^\circ(\varphi - \breve{\varphi})F_{\text{bulk}}(\breve{q}_+) \quad .\end{aligned} \tag{8.36}$$

The dispersion relation is obtained by varying the pseudo Lagrangian with respect to the wave amplitude q_+

$$\left.\frac{\delta\mathcal{F}}{\delta q_+}\right|_{\breve{:=}.} = 0 \quad , \tag{8.37}$$

combined with the compatibility relation

$$N\partial_t k = a\partial_a\varphi_t \quad , \tag{8.38}$$

At next order the evolution of the wave fluxes are given by

$$\left.\frac{\delta\mathcal{F}}{\delta\varphi}\right|_{\breve{:=}.} = 0 \quad . \tag{8.39}$$

The dispersion relation is given by

$$0 = -\frac{M_a n a^2}{N^2}\varphi_t\frac{q_+^2}{k^2} + H_a^\circ\left[\frac{q_+^2}{k^2} - q_+F'(q_+)\left(1 - \frac{\gamma\chi(q_+)}{2Q^\circ h^\circ kN}\right)\right] \quad , \tag{8.40}$$

where we have made use of the self-gravity nonlinear control parameter $\chi(q_+)$. Substituting this into the compatibility relation we obtain an equation in conservative form,

$$\partial_t k - \partial_y\left\{(\omega_f - \omega)\left[1 - \frac{k^2}{q_+}F'(q_+)\left(1 - \frac{\gamma\chi(q_+)}{2Q^\circ h^\circ kN}\right)\right]\right\} = 0 \quad , \tag{8.41}$$

where $y = \int \frac{N}{a}da$.

At the next order, the equation for the wave flux evolution is,

$$\begin{aligned}\partial_t\left(\frac{M_a n a^2}{N^2}\frac{q_+^2}{2k^2}\right) + \partial_a\frac{aH_a^\circ}{N}\left(\frac{q_+F'(q_+)}{k} - \frac{\gamma}{2Q^\circ h^\circ N}\frac{q_+^2}{k^2}\frac{\partial}{\partial q_+}\left[q_+\mathcal{W}(q_+)\right]\right) \\ = -\frac{2}{3}\alpha a T_a^\circ(F_{\text{shear}}(q_+) - F_{\text{bulk}}(q_+)) - \alpha_b a T_a^\circ F_{\text{bulk}}(q_+) \quad ,\end{aligned} \tag{8.42}$$

where we have made use of the dispersion relation to simplify the spacial component of the flux.

Equations 8.41 and 8.42 are evolutionary equations for the wavenumber k and nonlinearity q_+. In the absence of viscosity these equations are conservative and

describe how the wavenumber and AMD are conserved in the disc. In general this equation has the form

$$\nabla_\mu j^\mu = S, \tag{8.43}$$

where $\mu = \{0, 1\}$ are space time indices, $j^t = \frac{M_a n a^2}{N^2}\frac{q_+^2}{2k^2}$ is the AMD density j^a is the AMD flux and S are sources/sinks of AMD.

Now we confirm that the linear limit of our theory matches the WKB solutions of the linear eccentric disc theory. In chapter 6 we showed this was the case for the ideal short-wavelength theory. For a 3D disc, the linearised dispersion relation is

$$0 = -\frac{M_a n a^2}{N^2 k^2}\varphi_t + H_a^\circ\left(\frac{1}{k^2} - \frac{2\gamma-1}{2\gamma} + \frac{\gamma}{Q^\circ h^\circ k N}\right) \quad , \tag{8.44}$$

while the equation for the wave flux evolution is

$$\begin{aligned} 0 = \partial_t\left(\frac{M_a n a^2}{N^2}\frac{q_+^2}{2k^2}\right) + \partial_a\frac{aH_a^\circ}{kN}\frac{q_+^2}{2}\left(\frac{2\gamma-1}{\gamma} - \frac{\gamma}{Q^\circ h^\circ k N}\right) \\ + aT_a^\circ\left[\frac{3}{2}\alpha\left(1 + x - y\frac{\gamma-1}{\gamma}\right) + \left(\alpha_b - \frac{2}{3}\alpha\right)\frac{2\gamma-1}{2\gamma}\right]q_+^2 \quad . \end{aligned} \tag{8.45}$$

Making use of $q_+ = Nke$ we can rearrange this to obtain

$$\begin{aligned} 0 = M_a n a^2 e_t + aH_a^\circ\left(Nk\frac{2\gamma-1}{\gamma} - \frac{\gamma}{Q^\circ h^\circ}\right)e_a + e\partial_a\left[\frac{aH_a^\circ}{2}\left(Nk\frac{2\gamma-1}{\gamma} - \frac{\gamma}{Q^\circ h^\circ}\right)\right] \\ + aT_a^\circ\left[\frac{3}{2}\alpha\left(1 + x - y\frac{\gamma-1}{\gamma}\right) + \left(\alpha_b - \frac{2}{3}\alpha\right)\frac{2\gamma-1}{2\gamma}\right]N^2k^2e \quad . \end{aligned} \tag{8.46}$$

From Teyssandier & Ogilvie (2016) the dynamics of a linear eccentric disc subject to pressure and bulk viscosity is

$$\begin{aligned} \Sigma r^2\Omega\dot{E} - i\Sigma r^2\Omega\omega_f E = \frac{i}{r}\frac{\partial}{\partial r}\left[\frac{1}{2}\left(2 - \frac{1}{\gamma}\right)Pr^3\frac{\partial E}{\partial r}\right] \\ + \frac{i}{2}\left(4 - \frac{3}{\gamma}\right)r\frac{dP}{dr}E + \frac{3}{2}i\left(1 + \frac{1}{\gamma}\right)PE + \frac{1}{2r}\frac{\partial}{\partial r}\left(\alpha_b Pr^3\frac{\partial E}{\partial r}\right) \quad . \end{aligned} \tag{8.47}$$

In the WKB limit terms we only keep terms involving derivatives of E, as well as the term involving the forced precession ω_f so the relevant equation is

$$\Sigma r^2\Omega\dot{E} - i\Sigma r^2\Omega\omega_f E = \frac{i}{r}\frac{\partial}{\partial r}\left[\frac{1}{2}\left(2-\frac{1}{\gamma}\right)Pr^3\frac{\partial E}{\partial r}\right] + \frac{1}{2r}\frac{\partial}{\partial r}\left(\alpha_b Pr^3\frac{\partial E}{\partial r}\right) \quad . \tag{8.48}$$

This scaling comes from assuming the precessional terms are large, $\Sigma r^2\Omega\omega_{\rm f} \gg P$. We also require the viscosity to be sufficiently small. The appropriate scaling to match the tight-winding theory is $\omega_f - \omega \sim \varpi_a^2 = O(\delta^{-2})$, $e_t \sim \varpi_a e$ and $\alpha_b = O(\delta)$. Splitting into real and imaginary parts and writing in terms of our notation Equation 8.48 becomes,

$$\begin{aligned} M_a n a^2 e_t = &-\varpi_a e\frac{\partial}{\partial a}\left[\frac{2\gamma-1}{2\gamma}H_a^\circ a^2\right] - \left[\frac{2\gamma-1}{2\gamma}H_a^\circ a^2\right](e_a\varpi_a + e\varpi_{aa}) \\ &- \frac{1}{2}\alpha_b T_a^\circ a^3 e\varpi_a^2 + \text{Higher order terms}, \end{aligned} \tag{8.49}$$

$$N^2 H_a^\circ e = \left[\frac{2\gamma-1}{2\gamma}H_a^\circ a^2\right]e\varpi_a^2 + \text{Higher order terms}. \tag{8.50}$$

Noting that $a\varpi_a = a\varphi_a = kN$, Equations 8.49 and 8.50 match Equations 8.46 and 8.44 respectively (in the absence of self gravity and shear viscosity), with the exception of a constant factor in the viscous term. In fact Equation 8.49 will match Equation 8.46 if the 2D form of the bulk viscosity ($F_{\rm bulk}^{(2D)}$) is used. In practice this difference can just be absorbed into the definition of α_b.

8.6 Example solutions for isothermal black hole discs

In Section 6.8.4 of Chapter 6 we obtained solutions for steady tightly-wound waves in an isothermal pseudo-Newtonian disc, which were compared against the findings of Dewberry et al. (2020a) and Dewberry et al. (2020b). To illustrate an application for the theory developed in this chapter we now extend this calculation to include viscous damping, and compare the very different behaviours of shear and bulk viscosities on tightly wound waves.

The pseudo-Lagrangian for a steady-tightly wound wave in an isothermal disc in a Paczyński-Wiita potential, subject to a bulk α-viscosity is

$$\mathcal{F}_a = H_a^\circ\left(\frac{1}{2}\frac{q_+^2}{k^2} - F(q_+)\right) + \alpha_b a T_a^\circ(\varphi - \breve{\varphi})F_{\rm bulk}(\breve{q}_+) \quad , \tag{8.51}$$

while the same setup subject to a shear $\alpha-$viscosity is,

$$\mathcal{F}_a = H_a^\circ \left(\frac{1}{2}\frac{q_+^2}{k^2} - F(q_+) \right) + \frac{2}{3}\alpha a T_a^\circ (\varphi - \breve{\varphi})(F_{\rm shear}(\breve{q}_+) - F_{\rm bulk}(\breve{q}_+)) \quad . \tag{8.52}$$

For an α-disc $aT_a^\circ = H_a^\circ$. We shall restrict our attention to discs with a constant surface density profile, so that $H_a^\circ = 2\pi c_0^2 \Sigma^\circ a$, with c_0, Σ° constant sound speed and reference surface density respectively. As obtained in Section 6.8.4, N^2 in this disc is

$$N^2 = \frac{n^2 a^4 r_g}{c_0^2 (a - r_g)^3} \quad . \tag{8.53}$$

The dispersion relation of the wave is independent of the viscosity and can be obtained by varying F with respect to q_+. For an inward travelling wave (i.e. one propagating in from the outer boundary as in Dewberry et al. (2020a) and Dewberry et al. (2020b)) this is given by

$$k = -\sqrt{\frac{q_+}{F'(q_+)}} \quad . \tag{8.54}$$

Varying Equations 8.51 and 8.52 with respect to φ and making use of the dispersion relation, we arrive at an equation describing how the nonlinearity varies in the disc. For bulk viscosity this is

$$a\partial_a \left(\frac{aH_a^\circ}{N} q_+^{1/2} F'(q_+) \right) = \alpha_b a H_a^\circ F_{\rm bulk}(q_+), \tag{8.55}$$

while for shear viscosity it is

$$a\partial_a \left(\frac{aH_a^\circ}{N} q_+^{1/2} F'(q_+) \right) = \frac{2}{3}\alpha a H_a^\circ (F_{\rm shear}(q_+) - F_{\rm bulk}(q_+)) \quad . \tag{8.56}$$

For an isothermal α-disc we can evaluate the "viscous integrals",

$$F_{\rm bulk}(q_+) = 1 + \frac{1}{\sqrt{1-q^2}}, \tag{8.57}$$

$$F_{\rm shear}(q_+) = \frac{3}{2}\frac{2(1-q^2)^{3/2} - 2 + 3q^3}{(1-q^2)^{3/2}}. \tag{8.58}$$

Equations 8.55 and 8.56 can be solved using a numerical ODE solver. We solve these using Mathematica for $c_0 = 0.01$ and different values of α/α_b, with solutions shown in Figures 8.3 and 8.4. As expected the inclusion of bulk viscosity acts to damp the waves, with the nonideal solutions having consistently lower amplitudes than

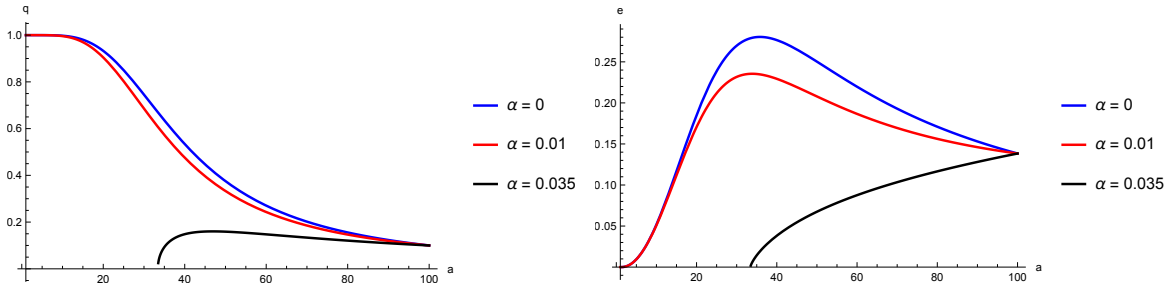


Fig. 8.3 Nonlinearity (Left) and eccentricity (Right) profiles for an isothermal, pseudo-Newtonian disc subject to bulk viscosity. The eccentricity scale is arbitrary, and the semimajor axis is in units of r_g. The bulk viscosity acts to damp the wave and all the solutions which include bulk viscosity are of lower amplitude than the ideal solution. As found in Ferreira & Ogilvie (2009) a large enough bulk viscosity can completely damp the wave before it reaches the marginally stable orbit.

their ideal counterpart. For sufficiently large bulk viscosity the wave is completely damped before it reaches the marginally stable orbit, a result also found by Ferreira & Ogilvie (2009) in linear theory. The solutions including shear viscosity confirm that it is anti-dissipative, for low enough amplitude waves, with the solutions containing a shear viscosity attaining large amplitudes at larger semimajor axis than their ideal counterpart. As the eccentric wave approaches the marginally stable orbit it becomes nonlinear enough for shear viscosity to act dissipatively resulting in a decline in the wave amplitude in the inner region of the disc.

The two types of viscosity have fundamentally different effects on the eccentric wave. This highlights the uncertainty that currently exists surrounding the inclusion of dissipative effects in eccentric discs. Which (if either) of the two is an appropriate proxy for the action of MRI turbulence on eccentric waves is currently unknown, and something that needs to be addressed in future research.

8.7 Conclusions

In this chapter we have formulated a nonlinear theory for highly twisted "tightly wound" eccentric waves in astrophysical discs based on the averaged Lagrangian method of Whitham (1965) and its generalisation to include non-conservative terms in Jimenez & Whitham (1976). We obtain a dispersion relation which matches that derived in Chapter 6 as the non-conservative terms are weak and do not appear at lowest order. Instead of the conserved flux of Chapter 6 we instead obtain an equation for the evolution of the AMD where viscosity acts as a source/sink. We confirm that

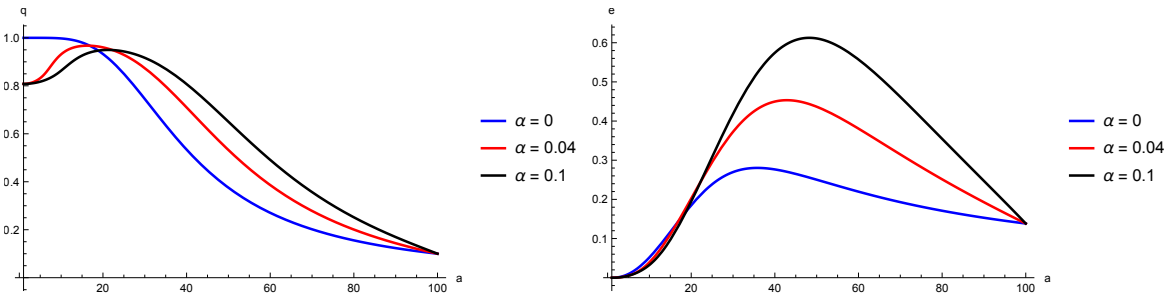


Fig. 8.4 Nonlinearity (Left) and eccentricity (Right) profiles for an isothermal, psuedo-Newtonian disc subject to shear viscosity. Over most of the disc the shear viscosity is anti-dissipative and the eccentric waves are excited to higher amplitudes than the ideal theory. As they approach the marginally stable orbit the waves become sufficiently nonlinear for viscosity to become dissipative again, resulting in a reduction in the wave amplitude.

tightly-wound waves can be amplified by the viscous overstability and are excited by shear viscosity when they are of low enough amplitude. Shear viscosity becomes dissipative again for sufficiently nonlinear waves.

Chapter 9

Magnetic Fields in Local Models of Eccentric Discs

9.1 Introduction

In the preceding Chapters we considered purely hydrodynamical models of eccentric discs. However, as reviewed in Chapter 2, the turbulence modelled by the $\alpha-$viscosity is most likely MHD in origin (although, unlike circular discs, eccentric discs can experience hydrodynamic turbulence due to the parametric instability). In addition many astrophysical discs are well ionised and can thus support large scale magnetic fields, which may have dynamical consequences.

In this chapter we consider the behaviour of magnetic fields in an eccentric disc. We will focus on local models of the magnetic field by adapting the ideal MHD eccentric shearing box given the Appendix of Ogilvie & Barker (2014). We shall start by deriving solutions for a largescale/coherent magnetic field in an ideal MHD eccentric shearing box. We shall then derive a nonlinear closure model for the MRI where the magnetic field is allowed to dissipate. The more difficult task of developing global MHD solutions in eccentric discs we shall not attempt here.

9.2 Ideal MHD

9.2.1 Solution to the induction equation

In this section we derive the the structure of a steady magnetic field in an eccentric disc. The equations for a horizontally invariant laminar flow in a magnetised disc in

an eccentric shearing box were derived in Ogilvie & Barker (2014), in their coordinate system (Reviewed in Section 4.3.2 of Chapter 4) the induction equation is

$$DB^{\xi} = -B^{\xi}\left[\Delta + \partial_{\varsigma} v_{\varsigma}\right] \quad , \tag{9.1}$$

$$DB^{\eta} = \Omega_{\lambda} B^{\xi} + \Omega_{\phi} B^{\eta} - B^{\eta}[\Delta + \partial_{\varsigma} v_{\varsigma}] \quad , \tag{9.2}$$

$$DB^{\zeta} = -B^{\zeta}\Delta \quad , \tag{9.3}$$

where $\Omega_{\lambda} = \frac{\partial \Omega}{\partial \lambda}$ and $\Omega_{\phi} = \frac{\partial \Omega}{\partial \phi}$. D is the Lagrangian time derivative, which for a horizontally invariant laminar flow is,

$$D = \partial_t + v^{\zeta}\partial_{\zeta} \quad . \tag{9.4}$$

and Δ the horizontal velocity divergence associated with the disc orbits. For these equations, and in what follows, we have dropped the primes on the shearing coordinate system that distinguishes them from the unsheared coordinates. Introducing the Lagrangian vertical coordinate $\tilde{z} = \frac{\varsigma}{H}$ with $D\tilde{z} = 0$ and we have for the vertical velocity

$$v_{\zeta} = \frac{\dot{H}}{H}\zeta \quad , \tag{9.5}$$

where an overdot indicates a Lagrangian time derivative, i.e. $\frac{d}{dt}$ at constant $\tilde{z}$.

In order to rewrite the terms involving derivatives of Ω we introduce functions $\aleph$ and $\beth$ defined by,

$$\frac{\dot{\aleph}}{\aleph} = -\Omega_{\phi} \quad , \frac{\dot{\beth}}{\aleph} = -\Omega_{\lambda}, \tag{9.6}$$

which in Ogilvie & Barker (2014) were denoted α and β. Noting that $\Delta = \frac{\dot{J}}{J}$ and $\partial_{\varsigma} v_{\varsigma} = \frac{\dot{H}}{H}$, where J is the Jacobian dimensionless determinant of the (Λ, λ) orbital coordinate system[1] and H is the dimensionless disc scale height (i.e. J and H have been rescaled by their values in the reference circular disc, as was done in Chapter 7), the ξ and ζ components of the induction equation become

[1]There is an unfortunate notational clash between our use of J to denote the dimensionless Jacobian of the (Λ, λ) coordinate system and it's meaning in Ogilvie & Barker (2014) as the Jacobian of the (λ, ϕ) coordinate system. In fact our J is closer to $\mathcal{J}$ of Ogilvie & Barker (2014), used to denote the Jacobian of the eccentric shearing box coordinate system.

$$\frac{\dot{B}^{\xi}}{B^{\xi}} + \frac{\dot{J}}{J} + \frac{\dot{H}}{H} = 0 \quad , \tag{9.7}$$

$$\frac{\dot{B}^{\zeta}}{B^{\zeta}} + \frac{\dot{J}}{J} = 0 \quad , \tag{9.8}$$

which have solutions,

$$B^{\xi} = \frac{B_0^{\xi}(\tilde{z})}{JH}, \quad B^{\zeta} = \frac{B_0^{\zeta}}{J} \quad , \tag{9.9}$$

where B_0^{ξ} and B_0^{ζ} are constant functions that are set by the initial conditions and we have made use of the solenoidal condition to show B_0^{ζ} is independent of $\tilde{z}$. Substituting the solution for B^{ξ} into the η component of the induction equation and rearranging we get,

$$\aleph JH\dot{B}^{\eta} + \aleph J\dot{H}B^{\eta} + \aleph\dot{J}HB^{\eta} + \dot{\aleph}JHB^{\eta} + \dot{\beth}B_0^{\xi} = 0 \quad , \tag{9.10}$$

which has solution,

$$B^{\eta} = \frac{\Omega B_0^{\eta}(\tilde{z})}{nJH} - \frac{\Omega\beth B_0^{\xi}(\tilde{z})}{nJH} \quad , \tag{9.11}$$

where we have used $\aleph \propto \Omega^{-1}$, which follows from $\Omega_{\phi} = \frac{\dot{\Omega}}{\Omega}$ and Equation 9.6.a. The factor of n has been included so that B^{η} and B_0^{η} have the same dimensions.Thus the large scale magnetic field solution in an eccentric shearing box is given by,

$$B^{\xi} = \frac{B_0^{\xi}(\tilde{z})}{JH}, \quad B^{\eta} = \frac{\Omega B_0^{\eta}(\tilde{z})}{nJH} - \frac{\Omega\beth B_0^{\xi}(\tilde{z})}{nJH}, \quad B^{\zeta} = \frac{B_0^{\zeta}(\tilde{z})}{J} \quad . \tag{9.12}$$

The equation for $\beth$ is given in Ogilvie & Barker (2014) as

$$\beth = \frac{3}{2}\left(1 + \frac{2e\lambda e_{\lambda}}{1-e^2}\right)\left(\frac{GM}{\lambda^3}\right)^{1/2} t - \frac{\lambda e_{\lambda}(2 + e\cos\theta)\sin\theta}{(1-e^2)(1+e\cos\theta)^2} - \frac{\lambda\omega_{\lambda}}{(1+e\cos\theta)^2} + \text{constant} \quad , \tag{9.13}$$

which in the (a, E) coordinate system is given by

$$\begin{aligned}\beth = {} & \frac{3nt}{2(1-e(e+2ae_a))\sqrt{1-e^2}} - \frac{ae_a}{(1-e^2)^{3/2}}\frac{2 - e\cos E - e^2}{1-e(e+2ae_a)}\sin E \\ & - \frac{a\varpi_a(1-e\cos E)^2}{(1-e^2)(1-e(e+2ae_a))} + \text{constant} \quad ,\end{aligned} \tag{9.14}$$

this contains a term that grows linearly in time. This means that B^η can be expected to grow linearly in the presence of a "quasiradial" field.

The ideal induction equation can be written in tensoral form using the operator $\mathcal{D}$ (introduced in Chapter 4) as $\mathcal{D}M^{ij} = \frac{1}{\mu_0}\mathcal{D}B^iB^j = 0$ (Ogilvie, 2001). The solution to this equation for a horizontally invariant laminar flow (along with the solenoidal condition) in an eccentric shearing box is given by Equation 9.12.

Finally, when $B_0^\xi = 0$ it is convenient to write the magnetic field in a form which is independent of the horizontal coordinate system used,

$$\mathbf{B} = \frac{B_{h0}(\tilde{z})}{nJH}\mathbf{v}_{\text{orbital}} + \frac{B_{v0}}{J}\hat{e}_z, \tag{9.15}$$

where B_{v0} is a constant, $B_{h0}(\tilde{z})$ is a function of $\tilde{z}$ only and $\mathbf{v}_{\text{orbital}}$ is the orbital velocity vector.

9.2.2 Derivation of the ideal induction equation model

We now derive the full set of equations for a horizontally invariant laminar flow in an eccentric disc with a magnetic field. Assuming the magnetic field can be split into mean and fluctuating parts:

$$B = \bar{B} + b \quad . \tag{9.16}$$

In order that we have a steady field we require $B^\xi = 0$, otherwise there is a source term in the η component of the induction equation from the winding up of the "quasiradial" (B^ξ) field. This trivially satisfies the ξ component of the induction equation. We assume the fluctuating field b is responsible for the MRI and it's affect on the dynamics is captured by the turbulent stress prescription. Thus keeping the mean field only and dropping the overbar the equations for a horizontally invariant laminar flow in a magnetised disc in the eccentric shearing coordinates of Ogilvie & Barker (2014) are the η and ζ components of the induction equation

$$DB^\eta = \Omega_\lambda B^\xi + \Omega_\phi B^\eta - B^\eta[\Delta + \partial_\varsigma v_\varsigma] \quad , \tag{9.17}$$

$$DB^\zeta = -B^\zeta\Delta \quad , \tag{9.18}$$

the momentum equation

$$Dv_\zeta = -\phi_2\zeta - \frac{1}{\rho}\partial_\zeta\left[p + \frac{B^2}{2\mu_0} - T_{zz}\right] + \text{Tension} \quad , \tag{9.19}$$

where Tension are the magnetic tension terms. The solenoidal condition gives $\partial_\zeta B^\zeta = 0$, thus B^ζ is independent of ζ and the magnetic tension terms in the vertical momentum equation are zero. Finally the thermal energy equation is

$$Dp = -\Gamma_1 p\left(\Delta + \partial_\zeta v_\zeta\right) + (\Gamma_3 - 1)(\mathcal{H} - \partial_\zeta F_\zeta) \quad , \tag{9.20}$$

and we must specify the equation of state. Making use of the solutions to the induction equation (Equation 9.15), we obtain an expression for the magnetic pressure,

$$p_M = \frac{B_{h0}^2(\tilde{z})}{2\mu_0(nJH)^2}v^2 + \frac{B_{v0}^2}{2\mu_0 J^2} \quad , \tag{9.21}$$

with $v = |\mathbf{v}_{\text{orbital}}|^2$ is the square of the magnitude of the orbital velocity.

The contribution of the vertical component of the magnetic field to the magnetic pressure is independent of the height in the disc (in order to satisfy the solenoidal condition) and makes no contribution to the dynamics of the vertical structure. As such we neglect the vertical component of the magnetic field from this point on. The magnetic pressure simplifies to

$$p_M = \frac{B_{h0}^2(\tilde{z})}{2\mu_0(nJH)^2}v^2 \quad . \tag{9.22}$$

On a circular orbit $v^2 = (na)^2$ so that $p_m \propto (JH)^{-2} \propto \rho^2$ and the magnetic pressure behaves like a perfect gas with $\gamma = 2$. Thus, for a magnetised radiation-gas mixture, the magnetic field is the least compressible constituent of the plasma and will be the dominant source of pressure when the plasma is sufficiently compressed. On an eccentric orbit there is an additional source of variability owing to the stretching and compressing of the field by the periodic variation of the velocity tangent to the field lines.

The vertical component of the momentum equation becomes

$$\frac{\ddot{H}}{H} = -\phi_2 - \frac{1}{\rho\hat{H}^2\tilde{z}}\partial_{\tilde{z}}\left[p + \frac{B_{h0}^2(\tilde{z})}{2\mu_0(nJH)^2}v^2 - T_{zz}\right] \quad , \tag{9.23}$$

where we have used $\hat{H}$ to denote the dimensionful scale height, to distinguish it from the dimensionless scale height H.

Proposing separable solutions with

$$p = \hat{p}(\tau)\tilde{p}(\tilde{z}), \quad T_{zz} = \tilde{p}(\tilde{z})\hat{T}_{zz}(\tau), \quad \rho = \hat{\rho}(\tau)\tilde{\rho}(\tilde{z}) \quad . \tag{9.24}$$

The solution obey the generalised hydrostatic equilibrium which means the pressure obeys

$$\frac{d\tilde{p}}{d\tilde{z}} = -\tilde{\rho}\tilde{z} \quad . \tag{9.25}$$

To maintain separability we require the reference plasma beta to be independent of height,

$$\beta_m^\circ = \frac{2\mu_0\tilde{p}(\tilde{z})p^\circ}{a^2 B_{\hbar 0}^2(\tilde{z})} \quad . \tag{9.26}$$

From this we obtain the equation for variation of the scale height around the orbit

$$\frac{\ddot{H}}{H} = -\phi_2 + \frac{\hat{p}}{\hat{\rho}\hat{H}^2}\left[1 + \frac{1}{\beta_m^\circ J^2 H}\frac{P^\circ}{\hat{P}}\frac{v^2}{(an)^2} - \frac{\hat{T}_{zz}}{\hat{p}}\right] \quad , \tag{9.27}$$

where the square of the velocity is

$$v^2 = (an)^2\frac{1 + e\cos E}{1 - e\cos E} \quad . \tag{9.28}$$

The reference circular disc has $f_{\mathcal{H}} = \frac{9}{4}\alpha_s$ as in the hydrodynamic models considered in previous Chapters. In the reference circular disc, hydrostatic balance is given by

$$\frac{P^\circ}{\Sigma^\circ H^\circ H^\circ}\left[1 + \frac{1}{\beta_m^\circ}\right] = n^2 \quad . \tag{9.29}$$

Rescaling Equation 9.27 by this reference circular disc we obtain (for the radiation-gas mixture considered in Chapter 7),

$$\frac{\ddot{H}}{H} = -(1 - e\cos E)^{-3} + \frac{T}{H^2}\frac{1 + \beta_r}{1 + \beta_r^\circ}\frac{\left[1 + \frac{1+\beta_r^\circ}{1+\beta_r}\frac{1}{\beta_m^\circ JHT}\frac{1+e\cos E}{1-e\cos E} - \frac{\hat{T}_{zz}}{p}\right]}{\left[1 + \frac{1}{\beta_m^\circ}\right]} \quad , \tag{9.30}$$

with the rest of the equations proceeding as in the hydrodynamic laminar flow model considered in Chapter 7.

9.3 Nonlinear constitutive models of the MRI

The $\alpha-$prescription is the simplest model of MRI turbulence that can be adopted. Even so, adapting it for an eccentric disc is not straightforward and raises a number of questions. As discussed in Chapter 7 the $\alpha-$prescription references the orbital frequency, however it is not clear in an eccentric disc which orbital frequency it should reference. Additionally, in the $\alpha-$prescription, the stress responds instantaneously to changes in pressure. In the classic accretion disc model the surface density (and thus pressure) evolve on the viscous time, much longer than the timescale for the MRI to adjust to the change in pressure (of the order a couple of orbits Held (2020)), so this approximation is reasonable. However for an eccentric disc the pressure changes on the orbital timescale dues to a combination of horizontal compression and variation in vertical gravity around an eccentric orbit. Thus, in an eccentric disc, there is no separation in timescale between changes in pressure and the MRI response time, so the instantaneous response of the $\alpha-$prescription is a highly questionable approximation.

Various attempts have been made to rectify some of the deficiencies of the $\alpha-$prescription using alternative closure models for the turbulent stress. Ogilvie (2000) and Ogilvie (2001) proposed a viscoelastic model for the dyadic part of the Maxwell stress to account for the finite response time of the MRI. It was shown in Ogilvie & Proctor (2003) that for incompressible fluids there is an exact asymptotic correspondence between MHD in the limit of large magnetic Reynolds number and viscoelastic fluids (specifically an Oldroyd-B fluid) in the limit of large relaxation time. Ogilvie (2002a) improved upon the viscoelastic model of Ogilvie (2000) and Ogilvie (2001) by including an isotropic part to the stress to model the effects of magnetic pressure and correcting the heating rate so that total energy is conserved. Ogilvie (2003) proposed solving for both the Maxwell and Reynolds stresses and proposed a nonlinear closure model based on requiring the turbulent stresses to exhibit certain properties (such as positive definiteness, and relaxation towards equipartition and isotropy) known from simulations and experiments.

In this section we shall derive a new nonlinear constitutive model of the MRI in the vein of Ogilvie (2002a) and Ogilvie (2003). The model will be constructed in such a way as to exhibit many known features of MRI from simulations, along with some plausible properties motivated by theoretic considerations. The model we shall derive represents an improvement on Ogilvie (2000), Ogilvie (2001) and Ogilvie (2002a) in terms of better reflecting the properties of the MRI, but is less complex than that presented in Ogilvie (2003).

9.3.1 Preliminaries - what we require from our model

In this section we review the macroscopic properties we require from a nonlinear constitutive model of the MRI. To develop this stress model we introduced a modified Maxwell stress,

$$M^{ij} = \frac{B^i B^j}{\mu_0}. \tag{9.31}$$

To begin with we note the asymptotic correspondence between an incompressible MHD fluid (in the limit of large magnetic Reynolds number) and an incompressible Oldroyd-B fluid (in the limit of large Deborah number) found by Ogilvie & Proctor (2003). In Chapter 4 we showed how this correspondence breaks down for compressible fluids owing to the fact that the (modified) Maxwell stress and the polymeric stress transform differently under changes in volume. Instead in the compressible model both the modified Maxwell stress and Oldroyd-B model evolve according to equations of the form,

$$\mathcal{D}T^{ij} = \ldots, \quad \mathcal{D}M^{ij} = \ldots \tag{9.32}$$

where the right hand side vanishes in the asymptotic limit. The operator $\mathcal{D}$, defined in Chapter 4 and repeated here, is

$$\mathcal{D} = J^W \mathcal{L}_u J^{-W} \quad , \tag{9.33}$$

where $\mathcal{L}_u$ is the Lie derivative with respect to the fluid flow u, J is the Jacobian of the Lagrangian coordinate system and W is the weight of the tensor density (-1 for the polymeric stress and -2 for the modified Maxwell stress) of the tensor density $\mathcal{D}$ is operating on. When acting on the modified Maxwell stress the operator $\mathcal{D}$ is equivalent to that considered in the compressible MRI models of Ogilvie (2000), Ogilvie (2001) and Ogilvie (2002a), and the equation $\mathcal{D}M^{ij} = 0$ yields the modified Maxwell stress from a magnetic field obeying the ideal induction equation. Taken together this suggests that our constitutive model of the MRI should take the form,

$$\mathcal{D}M^{ij} = \ldots \quad , \tag{9.34}$$

where the right hand side is to be determined. The right hand side need not vanish in the limit of large magnetic Reynolds number, if (as will be assumed) there is a turbulent cascade. To determine the right hand side we should bear in mind the

following properties, based on theoretical considerations and numerical simulations of MHD turbulence,

1. **Dissipation of the Magnetic field:** While nonideal MHD effects in a well ionised laminar accretion disc are weak, the turbulent cascade should always move magnetic energy to scales where nonideal MHD terms become important. Thus the right hand side should contain some form of dissipative term.

2. **Energy Conservation:** Energy lost from the magnetic field due to the aforementioned dissipative term should be added as heat to the fluid. Likewise growth in magnetic energy should be balanced by reducing the thermal energy.

3. **Realisability:** As M^{ij} comes from the product of $\mathbf{B}$ with it's transpose, M^{ij} must be positive (semi-)definite.

4. **Viscous Limit:** As the $\alpha-$prescription (mostly) successfully models MRI turbulence in classical disc theory, where pressure and orbital shear does not change on the orbital timescale, we expect our model should recover viscous behaviour in some appropriate limit.

In the following sections we shall introduce one such constitutive MRI model, and show that it exhibits the desired properties.

9.3.2 Statement of the model

We propose a constitutive model of the MRI where the disc turbulence acts to isotropise the magnetic field. This produces a quasi-radial field component that the background shear can convert into quasi-toroidal field. We assume that the (modified) Maxwell stress relaxes to an isotropic distribution proportional to some pressure p_v on a timescale τ. One such equation for M^{ij} which exhibits this behaviour is

$$\mathcal{D}M^{ij} = -\frac{1}{\tau}\left(M^{ij} - \mathcal{B}p_{\mathrm{v}}g^{ij}\right) \quad , \tag{9.35}$$

where we must specify how the relaxation time τ and fluctuation pressure p_v depend on M, p_g, p_r and the orbital frequencies in order to close the model. Without loss of generality $\mathcal{B}$ can be assumed to be constant, as any variation can be absorbed into the definition of p_v.

This equation is similar to the equation for the (modified) Maxwell stress given in Ogilvie (2003),

$$\mathcal{D}M^{ij} = \Omega_z p^{-1/2} \left[C_1 R^{-1/2} M R^{ij} - C_2 M^{1/2} M^{ij} \right], \tag{9.36}$$

where Ω_z is the vertical epicyclic frequency and R^{ij} is the Reynolds stress and C_1, C_2 are constants. In the MRI model proposed by Ogilvie (2003) there is an additional equation for the evolution of R^{ij}. Our equation is equivalent to taking Equation 9.36 and imposing an isotropic Reynolds stress $R^{ij} \propto p_v \left(\frac{p_v}{M}\right) g^{ij}$. The relaxation time for this model then takes the form $\tau \propto \Omega_z^{-1} (\frac{p_v}{M})^{1/2}$.

We shall normally express this relaxation time in terms of the Deborah number $\mathrm{De} = n\tau$. Unless otherwise stated, we adopt the closure model for the Deborah number that leads to consistency with Ogilvie (2003),

$$\mathrm{De} = \mathrm{De}_0 \left(\frac{n}{\Omega_z} \right) \sqrt{\frac{p_v}{M}}, \tag{9.37}$$

where De_0 is a constant.

9.3.3 Viscoelastic behaviour

This model is part of a large class of possible viscoelastic models for M^{ij}. The elastic limit of this equation $\tau \to \infty$ is fairly obvious, corresponding to a magnetic field which obeys the ideal induction equation through the "freezing in" of M^{ij} ($\mathcal{D}M^{ij} = 0$). However, to obtain the viscous behaviour responsible for an $\alpha-$viscosity requires more work. The viscous limit is obtained when $\tau \ll 1$. For simplicity, in what follows, we shall assume τ and p_v are independent of the magnetic field (M^{ij}).

Proposing a series expansion in τ,

$$M^{ij} = \sum_{k=0}^{\infty} \tau^k M_k^{ij} \quad . \tag{9.38}$$

This expansion is only likely to be valid in the short τ limit and may break down for material variations on timescales shorter than τ. With that caveat we can find a series solution for equation 9.35,

$$M^{ij} = \mathcal{B} \sum_{k=0}^{\infty} (-\tau \mathcal{D})^k p_v g^{ij} \quad . \tag{9.39}$$

Keeping the lowest order terms in the expansion we have,

$$M^{ij} = \mathcal{B} p_v g^{ij} - \mathcal{B}\tau \mathcal{D} p_v g^{ij} + O(\tau^2) \quad . \tag{9.40}$$

The lowest order term is an isotropic stress and evidently a form of magnetic pressure. The operator $\mathcal{D} = D$ when acting on a scalar and when acting on the metric tensor $\mathcal{D}g^{ij} = -2S^{ij} + 2\nabla_k u^k g^{ij}$, where S^{ij} is the rate of strain tensor, thus,

$$\begin{aligned} M^{ij} &\approx \mathcal{B}p_v g^{ij} + 2\mathcal{B}\tau p_v S^{ij} - 2\mathcal{B}\tau(p_v \nabla_k u^k + Dp_v)g^{ij} \\ &= \mathcal{B}p_v g^{ij} + 2\mathcal{B}\tau p_v S^{ij} + 2\mathcal{B}\tau \left(\left(\frac{\partial p_v}{\partial \ln \rho} \right)_s - p_v \right) \nabla_k u^k g^{ij} - 2\mathcal{B}\tau \left(\frac{\partial p_v}{\partial s} \right)_\rho Ds\, g^{ij} \quad . \end{aligned} \tag{9.41}$$

So the first $O(1)$ term is a magnetic pressure the first $O(\tau)$ term is a shear viscosity, the second is a bulk viscosity and the final term is an additional nonadiabatic correction which has no obvious analogue in the standard viscous or magnetic models. Higher order terms contain time derivatives of the pressure p_v and the velocity gradients $\nabla_i u_j$. This produces a memory effect in the fluid causing the dynamics of the fluid to depend on it's history. The fluid has a finite memory and becomes insensitive to ("forgets about") it's state at times $> \tau$ in the past.

When τ and p_v depend on M^{ij} then the terms in equation 9.41 are modified. However the equation still decomposes into an isotropic magnetic pressure term, a shear stress term $\propto S^{ij}$, a bulk stress term $\propto \nabla_k u^k g^{ij}$ and a non-adiabatic term $\propto Ds\, g^{ij}$ (as in equation 9.41).

9.3.4 Realisability

In addition to it behaviour in the viscous and elastic limits another necessary property of a model of a Maxwell stress is it's realisability from real magnetic fields. As $M^{ij} = \frac{B^i B^j}{\mu_0}$, M^{ij} must be positive semi-definite. Thus for all positive semi-definite initial conditions $M^{ij}(0)$ our constitutive model equation 9.35 must conserve the positive semi-definite character of M^{ij}.

To show this consider a vector field Y_i then defining a scalar $Q := M^{ij}Y_iY_j$ we must have $Q \geq 0$ for all vector fields Y_i. In what follows we shall follow a similar argument to that presented by Ogilvie (2003) to demonstrate the realisability of the stress model presented in that paper. The evolutionary equation for Q is,

$$\begin{aligned}\mathcal{D}Q &= Y_iY_j\mathcal{D}M^{ij} + M^{ij}\mathcal{D}(Y_iY_j) \\ &= \frac{\Omega_z}{\mathrm{De}_0}\sqrt{\frac{M}{p_v}}\left(\mathcal{B}p_vY^2 - Q\right) + M^{ij}\mathcal{D}(Y_iY_j) \quad .\end{aligned} \tag{9.42}$$

In order for a positive definite initial condition $M^{ij}(0)$ to evolve to violate the positive (semi-)definite property Q must pass through zero for some Y_i. As such, without loss of generality we can consider an initial M^{ij} such that $Q = 0$ for some initial Y_i (although it may be positive for other vector fields). If the vector field Y^i evolves according to $\mathcal{D}Y_i = 0$ (in other words it is passively advected by the flow) then the equation for Q becomes,

$$DQ = \frac{\Omega_z}{\mathrm{De}_0}\sqrt{\frac{M}{p_v}}\left(\mathcal{B}p_vY^2 - Q\right) \quad , \tag{9.43}$$

where we have made use of the fact that $\mathcal{D} = D$ when acting on a scalar. Within the disc we expect $p_v > 0$, additionally as M^{ij} is positive semi-definite $M \geq 0$. Thus at $t = 0$, $Q|_{t=0} = 0$ and the time derivative of Q is given by,

$$DQ|_{t=0} = \frac{\Omega_z}{\mathrm{De}_0}\sqrt{\frac{M}{p_v}}\mathcal{B}p_vY^2 \geq 0 \quad . \tag{9.44}$$

Thus Q cannot pass through 0 and M^{ij} remains positive semi-definite.

9.3.5 Energy conservation

In order that the interior of our disc conserve total energy we need to derive the appropriate magnetic heating/cooling rate so that energy lost/gained by the magnetic field is transferred to/from the thermal energy. The MHD total energy equation with radiative flux is (Ogilvie, 2016),

$$\partial_t\left[\rho\left(\frac{\mathbf{u}^2}{2} + \Phi + e\right) + \frac{\mathbf{B}^2}{2\mu_0}\right] + \nabla\cdot\left[\rho\mathbf{u}\left(\frac{\mathbf{u}^2}{2} + \Phi + h\right) + \mathbf{u}\frac{\mathbf{B}^2}{\mu_0} - \frac{1}{\mu}(\mathbf{u}\cdot\mathbf{B})\mathbf{B} + \mathbf{F}\right] = 0 \quad . \tag{9.45}$$

In terms of the modified Maxwell stress,

$$\partial_t \left[\rho \left(\frac{u^2}{2} + \Phi + e \right) + \frac{M}{2} \right] + \nabla_i \left[\rho u^i \left(\frac{u^2}{2} + \Phi + h \right) + u^i M - u_j M^{ij} + F^i \right] = 0 \quad , \tag{9.46}$$

from which we obtain the thermal energy equation,

$$\rho T D s = M^{ij} S_{ij} - \frac{1}{2}(DM + 2M \nabla_i u^i) - \nabla_i F^i \quad . \tag{9.47}$$

Using the constitutive relation, Equation 9.35, we obtain

$$\rho T D s = \frac{1}{2\tau} \left(M - 3\mathcal{B} p_v \right) - \nabla_i F^i \quad , \tag{9.48}$$

so we have nondimensional heating rate,

$$f_{\mathcal{H}} = \frac{1}{2\mathrm{De}} \left(\frac{M}{p_v} - 3\mathcal{B} \right) \quad . \tag{9.49}$$

Substituting Equation 9.41 into Equation 9.47 we recover terms proportional to $S^{ij} S_{ij}$ and $(\nabla_i u^i)^2$ which act like a viscous heating rate.

9.3.6 Stress model behaviour in a circular disc

In this section we consider the behaviour of our nonlinear constitutive model in a circular disc. We derive the reference circular disc, with respect to which our models are scaled. For a circular disc, the fixed point of equation 9.35 is,

$$M^{RR} = M^{zz} = \mathcal{B} p_v \tag{9.50}$$

$$R M^{R\phi} = -\frac{3}{2} \mathrm{De} \mathcal{B} p_v \tag{9.51}$$

$$R^2 M^{\phi\phi} = \mathcal{B} p_v \left(1 + \frac{9}{2} \mathrm{De}^2 \right), \tag{9.52}$$

which results in a magnetic pressure of

$$p_m = \frac{1}{2} M = \mathcal{B} p_v \left(\frac{3}{2} + \frac{9}{4} \mathrm{De}^2 \right) \quad . \tag{9.53}$$

As De and p_v can depend on M this equation needs to be solved to determine the equilibrium M. De is given by Equation 9.37. When p_v is independent of M we have,

$$M = \frac{3}{2}\mathcal{B}p_v\left(1+\sqrt{1+2\frac{\mathrm{De}_0^2}{\mathcal{B}}}\right) \quad , \tag{9.54}$$

while for $p_v = p + p_m$,

$$M = 3\mathcal{B}p\frac{1+\frac{3}{2}\mathrm{De}_0^2 \pm \sqrt{1+2\frac{\mathrm{De}_0^2}{\mathcal{B}}}}{2-\mathcal{B}\left(3+\frac{9}{4}\mathrm{De}_0^2\right)} \quad . \tag{9.55}$$

This solution can be negative, which is not allowed if M^{ij} is realisable from a real magnetic field, meaning that regions of the (De_0,$\mathcal{B}$) parameter space are unphysical for this stress model. This arises when $\mathcal{B}(\frac{3}{2}+\frac{9}{4}\mathrm{De}^2) > 1$ which, by Equation 9.53, implies $p_m > p_v = p + p_m$. Equation 9.53, with $p_v = p + p_m$, has a singular point when $2 = \mathcal{B}\left(3+\frac{9}{4}\mathrm{De}_0^2\right)$. However M is always negative at this singular point so we do not consider it further.

In addition to the equilibrium values of M^{ij}, the circular reference disc obeys hydrostatic and thermal balance. The equation for hydrostatic equilibrium in a circular disc is,

$$\frac{P}{\Sigma H^2}\left[1+\frac{M}{2p}-\frac{M_{zz}}{p}\right] = n^2 \quad , \tag{9.56}$$

while the equation for thermal balance is,

$$\mathcal{C}^\circ = f_{\mathcal{H}}\frac{P_v^\circ}{P^\circ} = \frac{9}{4}\mathcal{B}\mathrm{De}^\circ\frac{P_v^\circ}{P^\circ} \quad , \tag{9.57}$$

where,

$$\mathrm{De}^\circ = \mathrm{De}_0\sqrt{\frac{p_v^\circ}{M^\circ}} \quad . \tag{9.58}$$

Taking the solution to equations 9.56-9.57 which has $\beta_r = \beta_r^\circ$ and substitute this into equations 7.35 and 7.36 as the circular reference state we obtain,

$$\frac{\ddot{H}}{H} = -(1-e\cos E)^{-3} + \frac{T}{H^2}\frac{1+\beta_r}{1+\beta_r^\circ}\frac{\left[1+\frac{1}{2}\frac{M}{p}-\frac{M^{zz}}{p}\right]}{\left[1+\mathcal{B}\frac{P_v^\circ}{P^\circ}\left(\frac{1}{2}+\frac{9}{4}(\mathrm{De}^\circ)^2\right)\right]} \quad , \tag{9.59}$$

$$\dot{T} = -(\Gamma_3-1)T\left(\frac{\dot{J}}{J}+\frac{\dot{H}}{H}\right)+(\Gamma_3-1)\frac{1+\beta_r}{1+4\beta_r}T\left(\frac{1}{2\mathrm{De}}\left(\frac{M}{p}-3\mathrm{B}\frac{p_v}{p}\right)-\mathcal{C}^\circ\frac{1+\beta_r^\circ}{1+\beta_r}J^2T^3\right) \quad , \tag{9.60}$$

where the reference cooling rate is given by equation 9.57.

9.4 Conclusion

In this section we have considered the dynamics of magnetic fields in eccentric discs. We considered both a large scale/coherent magnetic field and a small scale turbulent field responsible for the MRI. For the large scale field we derived the simplest "locally axisymmetric" solutions to the Ideal induction equation in an eccentric shearing box. For the turbulent field responsible we derive a nonlinear constitutive model for the MRI satisfying the following properties,

- **Energy conservation:** energy lost/gained by the magnetic field are added/removed from the thermal energy equation.
- **Realisability:** the (modified) Maxwell stress is positive semi-definite.
- **Elastic response:** when subject to rapid deformation the magnetic field responds elastically and obeys the ideal induction equation.
- **Viscous limiting behaviour:** in the limit that deformations are much slower than the MRI response time the magnetic stress can be written as a series solution in relaxation time τ. The lowest order $O(1)$ term is an isotropic magnetic pressure, while the $O(\tau)$ terms act like shear and bulk viscosities.
- **Stress scales with magnetic pressure:** The viscous stress is naturally proportional to the magnetic pressure in the series solution.
- **Limited fluid memory:** Higher order terms act like a fluid memory, however the fluid memory is limited and "forgets" about it's history after $\sim$ one relaxation time.

Our model can be thought of as a compromise between the models presented in Ogilvie (2002a) and Ogilvie (2003). Our model is equivalent to that of Ogilvie (2003), but using a prescription for the Reynolds stress, rather than solving for it directly. The evolutionary equations for the (modified) Maxwell stress in our model and the model presented in Ogilvie (2003) are equivalent for an isotropic Reynolds stress proportional to some pressure.

Part IV

Eccentric Discs in Tidal Disruption Events

Chapter 10

Introduction to Tidal Disruption Events

10.1 Introduction

Tidal Disruption Events (TDEs) occur when an object moving in a gravitational field encounters tidal forces sufficient to overcome the forces holding the object together (be these self-gravity or material strength). As the basic requirement for a TDE is an object moving through a sufficiently strongly varying gravitational field they can occur on wide range of scales. However, to disrupt anything more than a loosely gravitationally bound object requires a strong tidal field. This means that with the exception of gas clouds, galaxies and rubble pile asteroids, TDEs typically require a compact object.

Typically a Tidal Disruption Event is taken to refer to the disruption of a star by a supermassive black hole (Hills, 1975; Rees, 1988; Phinney, 1989). More recently TDEs have been talked about in the context of asteroids and planets around white dwarfs, as a potential source of white dwarf pollution (Manser et al., 2016a; Cauley et al., 2018; Miranda & Rafikov, 2018; Malamud & Perets, 2019a,b; Nixon et al., 2020). Here, as well as self gravity, material strength can be important, particularly for smaller asteroids. Similarly it has been proposed that the rings of the gas giants were formed from the disruption of a moon or comet, with tidal disruption one mechanism by which this could happen (Harris, 1984; Colwell, 1994; Canup, 2010; Leinhardt et al., 2012). The case of a tidal disruption of a comet by a giant planet has a close resemblance to stellar TDEs as in both cases the pre-disruption orbit will be nearly parabolic.

TDEs are distinct from tidal stripping, such as occurs during stable Roche Lobe overflow, as instead of the steady removal of the least bound material, a sizeable fraction of the object's mass is removed on a short timeframe. Partial disruptions are

possible (Guillochon & Ramirez-Ruiz, 2013; Chornock et al., 2014; Coughlin & Nixon, 2019; Gomez et al., 2020), in which only the outer layers of an object are removed but the centre remains bound. The difference between this and tidal stripping is the inherently transient nature of TDEs, which occur on the orbital timescale as opposed to some longer timescale (e.g. the thermal time in the case of stable Roche Lobe overflow (Morton, 1960; Paczyński, 1971)). This difference in timescale is primarily a result of the very different orbital motions of a tidally stripped object, which due to the action of tides typically sits on a circular orbit and experience a relatively constant tidal force, compared to a total or partially tidally disrupted object on a nearly parabolic orbit, which experiences a rapidly changing tidal potential during its closest approach.

In order to tidally disrupt, an object must pass within its Roche radius, the distance from the disruptor where the body's self gravity is overcome by the tidal force of the disruptor. The tidal[1] or Roche radius (following Hills (1975)) of an object of radius R_1 and mass M_1 to disrupt around an object of mass M_0 is,

$$R_t = R_1(M_0/M_1)^{1/3} \quad . \tag{10.1}$$

Alternatively this can be written in terms of the ratio of the average density of the two objects,

$$R_t = R_0 \left(\rho_0/\rho_1\right)^{1/3} \tag{10.2}$$

It is common to parameterise how deep the disrupted object penetrates the Roche radius, and thus the strength of the tidal interaction, using the impact parameter,

$$\beta = R_t/r_p \quad , \tag{10.3}$$

where r_p is the pericentric radius. In order for a tidal disruption to occur we require $\beta \geq 1$.

As Equation 10.2 makes apparent, in order for the Roche radius to be outside of the disruptor's surface, the object must be less dense than the disruptor. This makes it difficult to tidally disrupt chondritic or iron asteroids with planets or stars, although large stony asteroids and comets can be disrupted by terrestrial planets with earth-like composition. Similarly ice rich, iron poor moons are susceptible to tidal disruption if they approach too close to a dense enough object. As catastrophically disrupting

[1]This differs by a factor of order unity from the classical Roche radius due to a difference in conventions between the stellar TDE literature, which uses Equation 10.1, and the planetary TDE literature which uses the Roche radius.

an object, rather than slowly stripping the outer layers, requires the object to pass entirely (or almost entirely) within the Roche radius it is difficult to disrupt planet or stars with other stars, and this typically requires a compact object. Recently, however, there has been discussion of tidal disruption of hot Jupiters, by their parent star, as an explanation of the lack of ultra-short period hot Jupiters (Matsakos & Königl, 2016).

For small asteroids or comets, material strength is more important than self gravity. For instance a typical yield stress (under tension) for 20m stony asteroid is ~ 0.2 M Pa (Collins et al., 2017), which will tidally disrupt $\sim 10^7$m away from a 1 solar mass point mass. This distance is comparable in size to Earth or, more importantly, a white dwarf and substantially smaller than any main sequence star. Asteroid yields strengths are highly variable and further complicated by being size dependent; as the asteroid size increases so does the number and size of the flaws, along with an increase in the proportion of weaker matrix material, weakening the asteroid (Petrovic, 2001; Collins et al., 2017). There is a minimum in asteroid strength at the strength-gravity transition for asteroids with diameters 100-1000 m, as these asteroids are large enough to have limited material strength but too small to be strongly gravitationally bound (Asphaug et al., 2002). Even for these much weaker asteroids, using equation 10.1, tidal disruption only occurs close to a solar radius.

In the subsequent sections we shall focus on the disruption of a star by a supermassive (or intermediate mass) black hole.

10.2 Classic model of Rees (1988)

The classical model of the disruption of a star by a supermassive black hole is by Rees (1988) in which a star, on a parabolic orbit, is instantaneously disrupted at pericentre and the stellar material is emplaced onto a set of orbits, with the outer half of the star becoming unbound and the inner half emplaced onto a range of bound orbits with large eccentricities approaching 1 (see Figure 10.1). The star's mass is assumed to be distributed with equal mass going into equal orbital energy bins. This material then follows test particle (assumed to be Keplerian) orbits with the most bound material returning to pericentre first.

As the material returns to pericentre it is assumed to circularise and radiate its orbital energy to infinity. The material now accretes onto the central black hole in a compact (with the outer disc radius $R_{\rm out} \sim 2R_p$, where R_p is the pericentre radius of the stars original orbit) thick accretion torus. The mass return rate to pericentre is controlled by the orbital period of the orbits the material is emplaced upon, with the

most bound material returning to pericentre first with increasingly unbound material with larger orbital periods returning later. This results in a mass return rate that follows a power law in time with index $-5/3$ (the original paper by Rees (1988) contains a mistake with the correct powerlaw index first given in Phinney (1989)). As the energy emitted from circularisation is proportional to the mass being circularised this leads to a prediction that the bolometric lightcurve of a TDE should follow a $t^{-5/3}$ decay law at late times.

Apart from the $t^{-5/3}$ decay law the other main prediction of Rees (1988) model is the bolometric luminosity of the event. For instance this predicts a peak bolometric Luminosity of $l \sim 10^{44}$ ergs s^{-1} for the disruption of a 1 solar mass star.

10.3 Theoretical Developments

Subsequent work has developed various aspects of TDE theory and addressed various issues with the original theory. Arguably the biggest question left open by the Rees (1988) model is how, and indeed if, the disc material circularises as it returns to pericentre. Simulations and analytical work have found that it is extremely difficult to dissipate the energy required to circularise the disc during pericentre passage, even with the extreme radial and horizontal compression and dissipation in the so called "nozzle shock" (Shiokawa et al., 2015; Piran et al., 2015; Krolik et al., 2016; Hayasaki et al., 2016; Bonnerot et al., 2016). The Rees (1988) model assumes that the main effect of general relativity is to rapidly circularise the disc through shocks induced by apsidal precession. Subsequent work has shown that strong apsidal precession causes the most bound material to intersect the less bound material resulting in dissipation of the orbital energy in a collisional shock (Bonnerot et al., 2016).

For TDEs disrupted close to the event horizon apsidal precession is strong and orbital intersection is prompt, occurring close to pericentre resulting in a very strong shock and rapid circularisation into a thick accretion torus a la Rees (1988). However TDEs occurring far from the event horizon have weak apsidal precession, with intersection close to apocentre. The shock, occurring at apocentre, will be weak and may not be capable of rapidly circularising the orbits. As a consequence it has been proposed that, instead of circularising, TDEs might accrete more directly onto the black hole as a highly eccentric disc (Shiokawa et al., 2015; Piran et al., 2015; Krolik et al., 2016). See, for example, Figures 10.2 and 10.3 which show moderate to highly eccentric discs in the simulations of Bonnerot et al. (2016) and Guillochon et al. (2014) respectively. Additionally what isn't always appreciated, particularly when explored

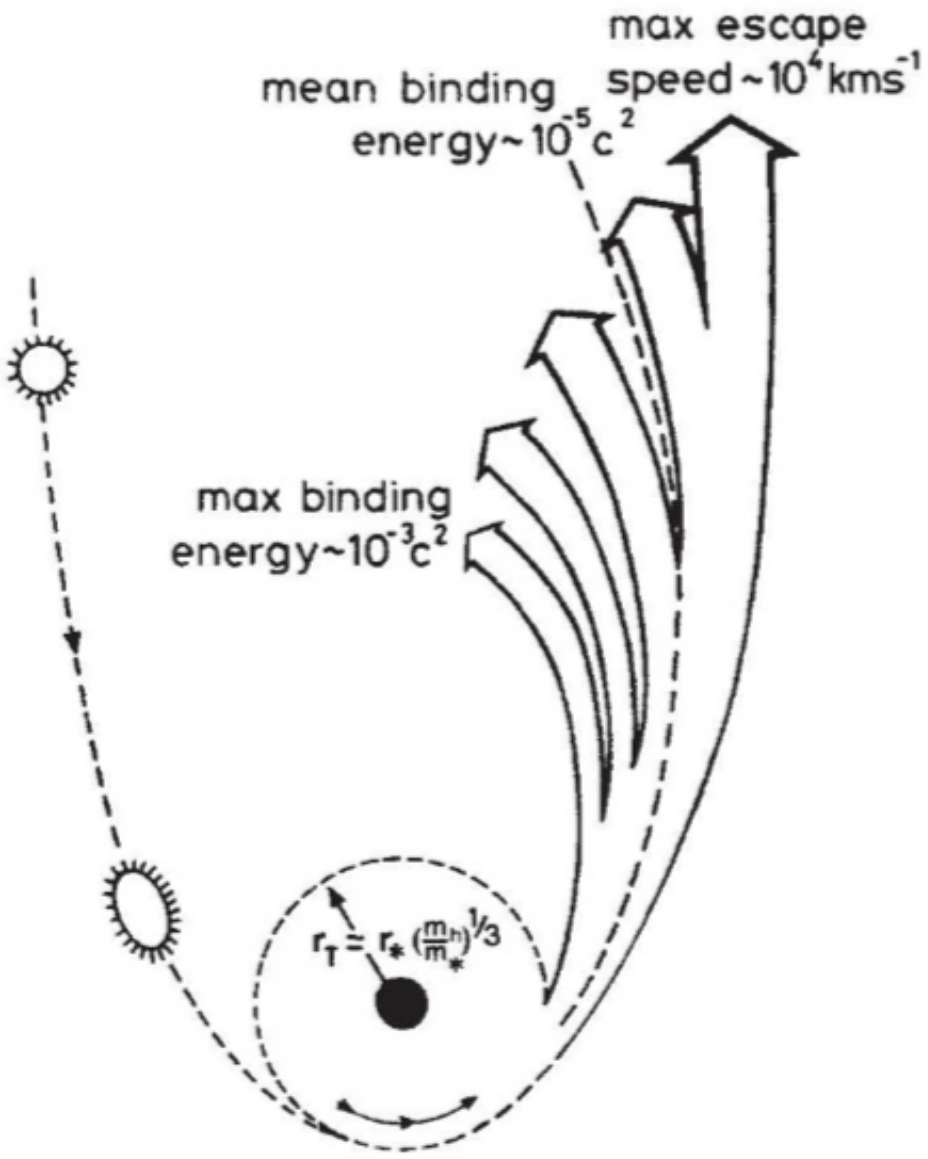


Fig. 10.1 Cartoon from Rees (1988) depicting the classical TDE model

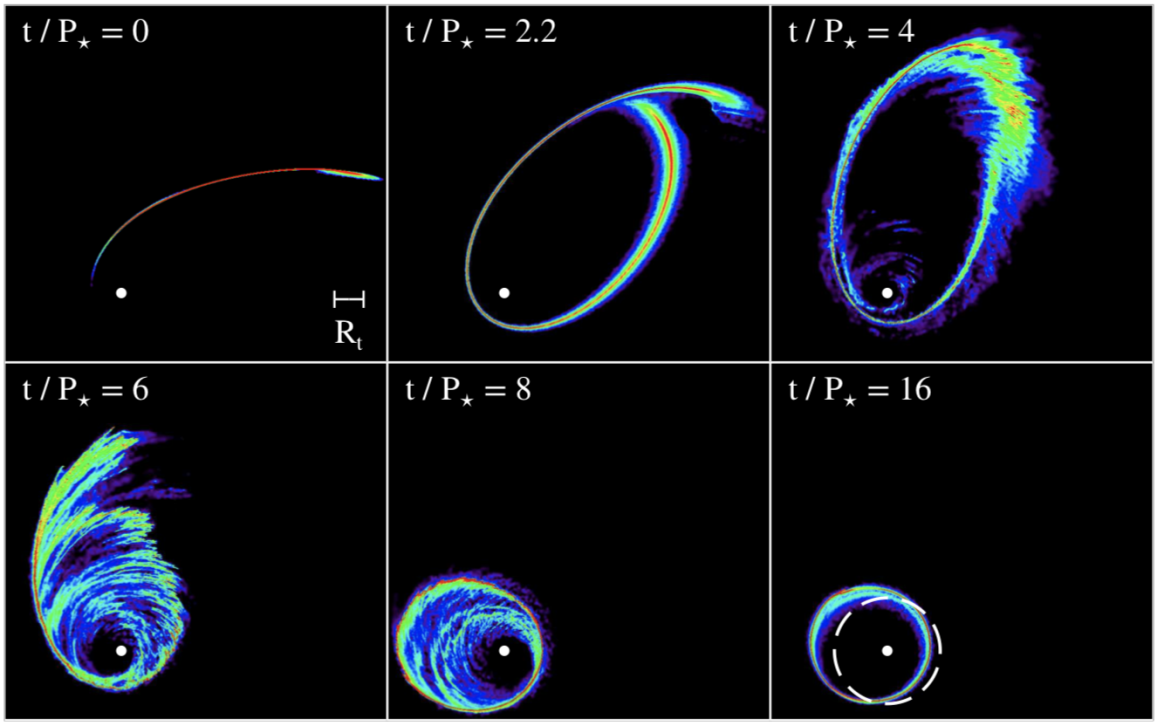


Fig. 10.2 Simulations of TDE discs from Bonnerot et al. (2016)

through potentially overly dissipative simulations, is that for weak apsidal precession the fact that eccentric orbits aren't closed in GR isn't sufficient to cause orbital self intersection. Rather when the apsidal precession angle during one orbit is small the fluid will follow slowly evolving Keplerian ellipses and this can result in a highly eccentric disc undergoing coherent global precession due to a combination of pressure forces and GR apsidal precession (Zanazzi & Ogilvie, 2020).

So what is more important for most TDEs? Strong apsidal precession resulting in prompt orbital intersection and circularisation similar to the Rees (1988) model or weak apsidal precession where circularisation is less clear? The answer will depend on the expected impact parameter β and the ratio R_t/R_s of the tidal radius to the Schwarzschild radius. Models of scattering/loss cone diffusion of stars onto orbits around the central black hole where they will disrupt typically find that the most common TDEs are those with the minimum penetration factor that will still disrupt ($\beta \approx 1$, Stone & Metzger (2016); Lu & Bonnerot (2019)). For TDEs around lower mass ($\sim 10^6 M_\odot$) black holes, or for higher mass stars, $R_t \gg R_s$, and are firmly placed in the weak apsidal precession regime. Strong apsidal precession is more common in the disruption of lower mass stars, or around higher mass black holes, as the stars need to get closer to the black hole to disrupt.

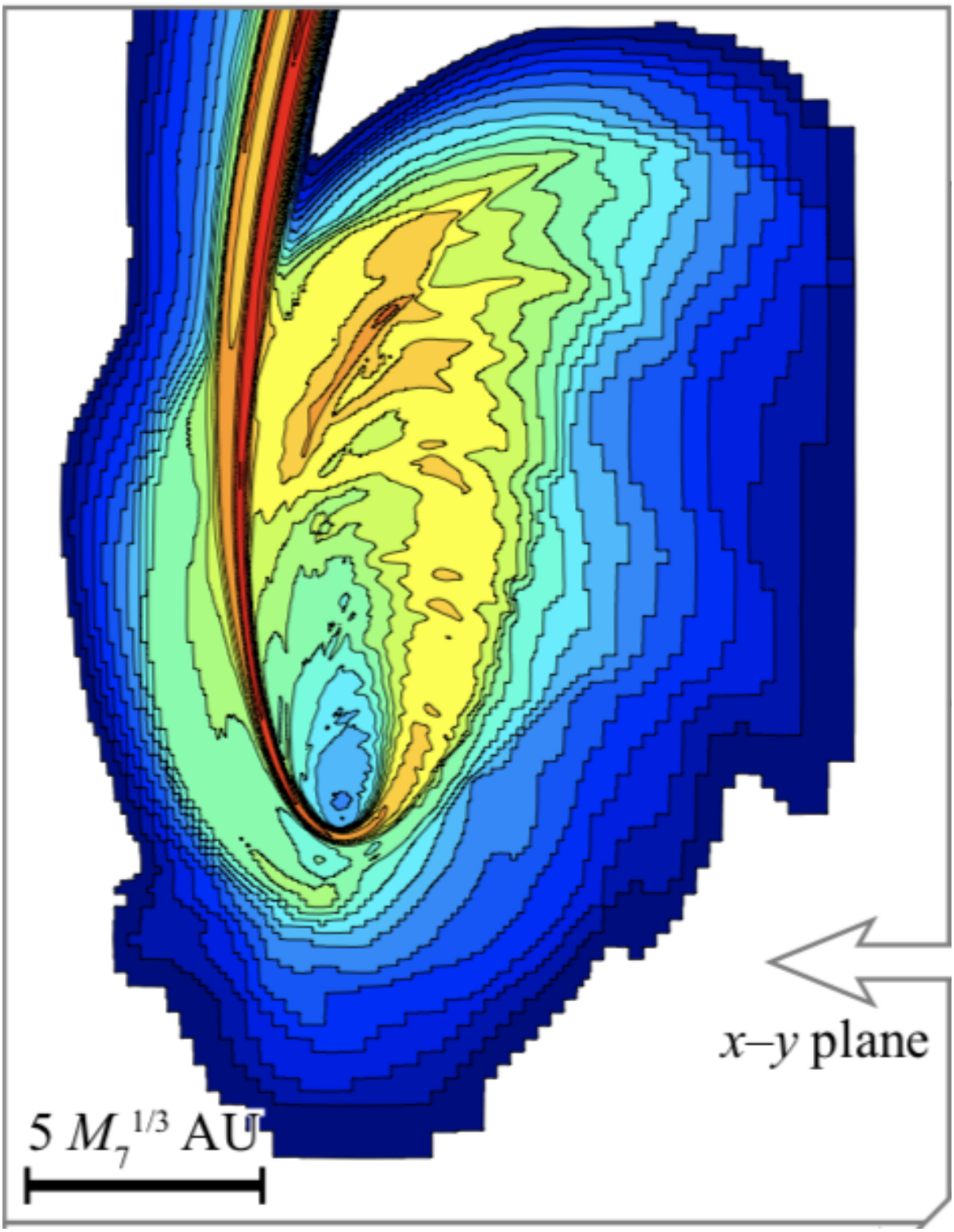


Fig. 10.3 Simulations of TDE discs from Guillochon et al. (2014)

10.4 Observations of TDEs

Observationally TDEs can be split into optically/UV bright TDEs and X-ray bright TDEs (Gezari et al., 2006, 2008; van Velzen et al., 2011; Wang et al., 2011, 2012; Arcavi et al., 2014). The X-ray bright TDEs are the closest to the classic Rees (1988) model and appear to broadly fit theoretical expectations. The low effective temperature of the optical/UV bright TDE is difficult to explain with the compact discs of classical TDE models, and was one of the motivating factors behind considering the more extended eccentric disc models (Shiokawa et al., 2015; Piran et al., 2015; Krolik et al., 2016; Zanazzi & Ogilvie, 2020). While optical/UV bright and X-ray bright TDEs are observationally distinct categories it is probable that they are not physically distinct and that this is simply a selection effect. In particular any initially X-ray bright TDE which is first detected in the optical will have decayed sufficiently to be undetectable in any X-ray followup. Recently a new sub-class of X-ray bright TDEs have been detected which possess jets and associated radio emission (Bloom et al., 2011; Burrows et al., 2011; Berger et al., 2012; Yang et al., 2016; Mattila et al., 2018; Chakraborty et al., 2020).

One issue apparent in the TDE literature is the occasional over-reliance on the simple model presented in Rees (1988), in particular the $t^{-5/3}$ decay law is occasionally taken as the definition of a TDE in the observational literature. This leads to obvious issues with selection biases. An example of this is detection of TDEs by the MAXI instrument aboard the International Space Station where the selection pipeline explicitly uses a $t^{-5/3}$ luminosity decay law to classify an event as a TDE (Kawamuro et al., 2017). This obviously leads to a confirmation bias where TDEs are observed to predominantly have indices of $-5/3$ simply because events are more likely to be classified as a TDE if they have this decay law. Despite this there is mounting observational evidence that there is a lot of variability in the TDE decay curve slopes, and that $-5/3$ isn't even typical. For instance Auchettl et al. (2017) refitted 70 TDE X-ray lightcurves taken from the literature and found that typically TDEs have shallower decay laws at late times (see Figure 10.4).

As with many other (potentially) eccentric discs one of the best probes of the geometry of the disc comes from observations of asymmetries in double peaked line profiles. For TDEs the most important line for probing the disc geometry is the Hα line, which occurs in optical/UV bright TDEs. To model the Hα line of Liu et al. (2017) (for TDE PTF09djl) and Cao et al. (2018) (for TDE ASASSN-14li) proposed that the line originated from a reprocessing the X-ray emission in a flat highly eccentric disc (e.g. Figure 10.5, from Cao et al. (2018)). As the Hα line in both TDEs was very broad

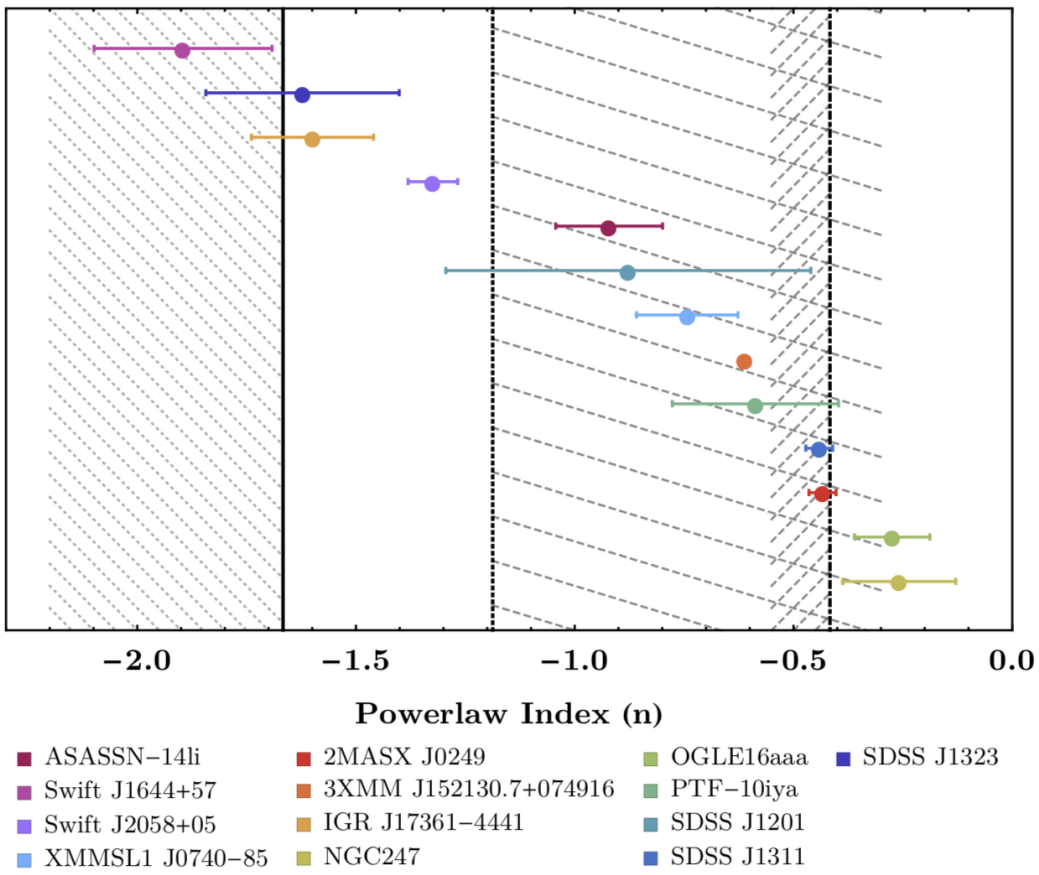


Fig. 10.4 Powerlaw indices fitted to X-ray TDE lightcurves from Auchettl et al. (2017). In general the powerlaw decay of TDEs is not well fitted by the classic $t^{-5/3}$ result and there is considerable scatter in the powerlaw indices of real TDEs which typically have shallower decay laws.

it was fitted with a disc with a large $e \approx 0.97$ constant eccentricity. Neither Liu et al. (2017) nor Cao et al. (2018) explicitly fitted the eccentricity directly, instead fitting the locations of the inner and outer edge of the disc, along with the disc orientation, and obtaining the eccentricity from the inferred pericentre distance and location of the stream-stream intersection. Holoien et al. (2019) showed that the Hα emission of TDE PS18kh indicated non-axisymmetric structures in the TDE disc, consistent with either an eccentric disc with eccentricity $e \lesssim 0.25$ or a spiral arm and was inconsistent with an axisymmetric disc. The most recent use of this technique was by Hung et al. (2020) with the claim of a much lower eccentricity of $e \approx 0.1$ for the TDE AT 2018hyz, with this lower eccentricity being measured shortly after the disruption - consistent with the return time of the debris. Interestingly the results of Hung et al. (2020) indicate the disc is much bigger than the tidal radius, which is inconsistent with the classical rapidly circularising model. One possible explanation, given in Hung et al. (2020), is that the Hα emitting region consists of a small percentage of the mass of the disc, with the majority of the disc located at smaller radii. Such a model might have much in common with the collisionally induced outflow + disc model of Lu & Bonnerot (2019) and Bonnerot & Lu (2019). An alternative explanation might be that given in Shiokawa et al. (2015) who pointed out that even if circularisation is efficient, if the radiation cannot escape the disc, then the disc must remain large, with similar semimajor axis to the eccentric debris, due to the additional pressure support. Unlike Liu et al. (2017) and Cao et al. (2018), Holoien et al. (2019) and Hung et al. (2020) fitted the eccentricity directly.

10.5 Typical orbital elements and reference disc for a TDE

In this section we calculate the typical orbital elements and appropriate reference circular disc for a TDE disc shortly after disruption. To do this fully self-consistently we should solve for e and ϖ as a function of a using the eccentric disc theory of Ogilvie (2001) (adapted for our coordinate system). This was done recently by Zanazzi & Ogilvie (2020). The orbits calculated here are the initial orbits after disruption, but it is not clear how long the debris will remain on these initial orbits.

Consider a star on a parabolic orbit which is instantaneously disrupted at pericentre. The fluid is assumed to exit pericentre with a uniform velocity (equal to the star's velocity at the moment of disruption) and the effects of self gravity and pressure are neglected. As the star is radially extended, the fluid is emplaced onto a range of bound

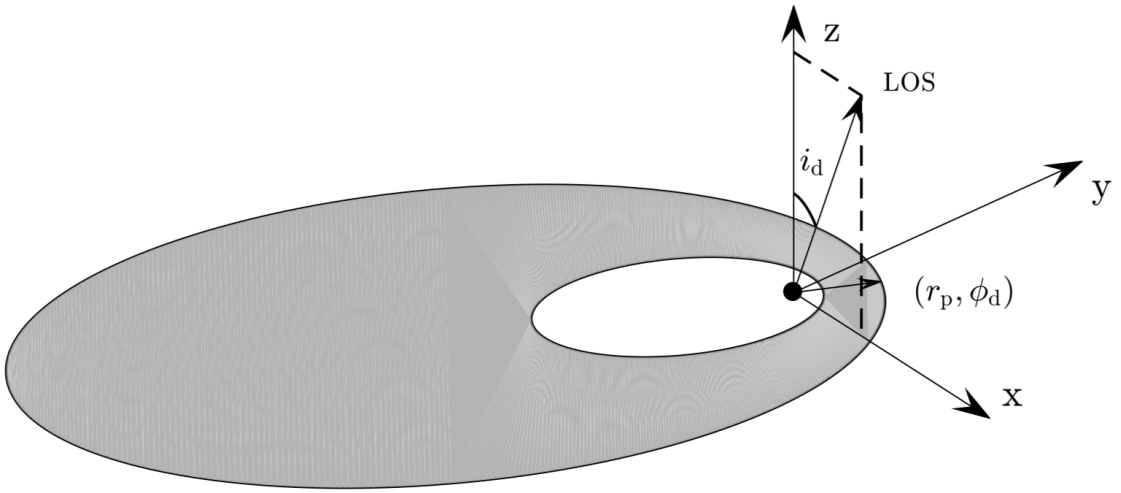


Fig. 10.5 Highly eccentric disc model from Cao et al. (2018) used to fit the Hα emission from the TDE ASASSN-14li. A similar model was used by Liu et al. (2017) for TDE PTF09djl.

and unbound orbits. The centre of the star (which continues along the original stellar orbit) follows a parabolic reference orbit with pericentre distance R_p and velocity v_p related by

$$\frac{1}{2}v_p^2 - \frac{GM_\bullet}{R_p} = 0 \quad . \tag{10.4}$$

Fluid elements all have the same pericentre velocity v_p but have a range of pericentre distances R given by

$$R = R_p(1+x) \quad , \tag{10.5}$$

where x is a dimensionless variable.

Their specific energy is

$$\mathcal{E} = \frac{v_p^2}{2} - \frac{GM_\bullet}{R} = -\frac{GM_\bullet}{2a}, \tag{10.6}$$

which we can use to obtain an equation for the semimajor axis and eccentricity parameterised in terms of x,

$$a = -\frac{R_p}{2}x^{-1}(1+x) \quad , \tag{10.7}$$

$$e = 1 - \frac{R}{a} = 1 + 2x \quad . \tag{10.8}$$

From these we can calculate q,

$$q = \frac{1}{1-2x} \approx 1 + 2x + O(x^2) \tag{10.9}$$

and so we conclude that for the early TDE disc we have

$$e \approx q \tag{10.10}$$

and

$$e = 1 - \frac{2R_p}{2a + R_p} \approx 1 - R_p/a \quad . \tag{10.11}$$

For the unbound debris with $x > 0$ we have $q > 1$. Normally this would indicate an orbital intersection, however the (a, E) coordinate system is not well defined for unbound orbits. Using the (λ, ϕ) coordinate system, or simply by plotting the orbits, it's apparent that the unbound debris has no orbital intersection for the eccentricity profile considered.

This eccentricity profile is for the orbits immediately after disruption, before the debris has spread around the orbit to form a coherent disc. The fluid must spread over the range in eccentric anomaly before we can make use of eccentric disc theory. This spread in E is initially prevented by the stream self gravity (Kochanek, 1994; Guillochon et al., 2014) so some degree of circularisation must have occurred. This will modify the eccentricity profile away from Equation 10.11. In the limit where apsidal precession is weak this circularisation will be weak and its main effect will be in breaking the hold self gravity has on the stream thickness and allowing the stream to spread around the orbit.

For a given black hole mass and R_p, Equation 10.7 allows us to calculate the mean motion n of the disc orbits. This is required to determine the scale height of the reference circular disc (important for determining if the disc is thin). We can rearrange the equations for hydrostatic and thermal balance, along with the equation of state, to obtain an equation for H° in terms of β_r°,

$$H^\circ = \left[\frac{4\lambda}{9\alpha_s} \frac{16\sigma}{3\kappa n^7} \left(\frac{\mathcal{R}}{\mu} \right)^4 \left(\frac{3c}{4\sigma} \right)^2 (\beta_r^\circ)^2 (1 + \beta_r^\circ)^3 \right]^{1/8} \quad . \tag{10.12}$$

Figure 10.6 shows how H°/a varies with β_r° for a typical TDE streamline.

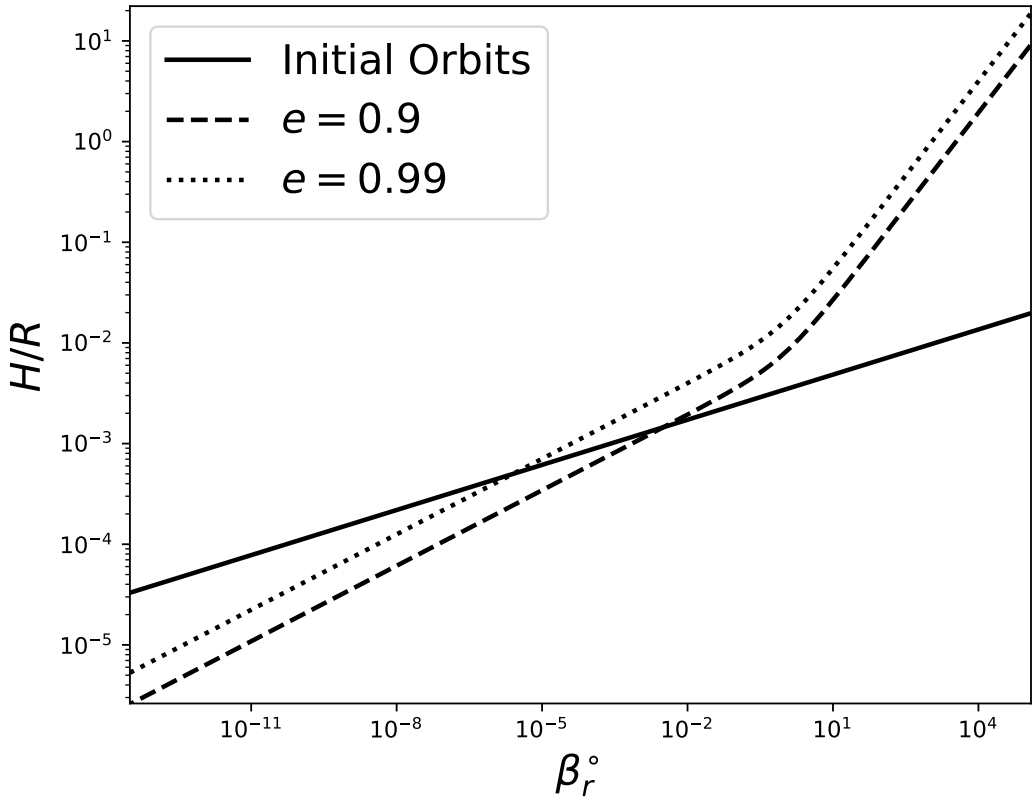


Fig. 10.6 Reference H/R against reference β_r from the disruption of a $1M_\odot$ star around a $10^6 M_\odot$ black hole at the tidal radius. The solid line is assuming the debris remains on the post disruption orbits given by Equation 10.11. The dashed and dotted lines keep e and q fixed as β_r° is varied. Stress is assumed to scale with gas pressure ($p_v = p_g$).

10.6 Summary of the evidence for eccentric TDE discs

In the next two chapters we shall consider a model of a highly eccentric TDE disc using the theoretical framework developed in this thesis. Here we shall summarise the evidence for eccentric TDE discs.

The first piece of evidence for eccentric TDE discs is that the initial stellar orbit is nearly parabolic. This means the debris from the tidal disruption event starts off on highly eccentric orbits. Thus a large degree of circularisation must occur quickly in order for the debris to not end up forming a highly eccentric disc. When apsidal precession is strong this appears to occur through the formation of a very strong shock due to the orbital intersection. However in many TDEs (particularly around lower mass black holes) apsidal precession is quite weak and it is not clear that it is capable of circularising the disc. Simulations of TDE discs that feature a realistic (i.e. $e \to 1$) initial stellar orbit all form a highly eccentric disc, although the length of time the TDE spends as a highly eccentric disc in these simulations is highly variable and can be as short as a couple of orbits. Even in simulations where the disc rapidly "circularises" the resulting disc is not circular and often continues to posses a moderate eccentricity. So the question is less about whether TDEs form eccentric discs (they do), but whether they spend any appreciable time with very large eccentricities.

On the observational front there are two lines of evidence for eccentric TDE discs. The first is the unexpectedly large and cool optically/UV bright TDEs which require a much larger accretion structure than the compact discs of the classical theory. Such an extended accretion structure has a natural explanation in the form of a highly eccentric disc, where the apocentre is much cooler and at much larger radii than the classical TDE discs. The second line of evidence is from the recent fits to the Hα emission which indicate an underlying non-axisymmetric structure (Liu et al., 2017; Cao et al., 2018; Holoien et al., 2019; Hung et al., 2020), although currently there is disagreement as to quite how eccentric the measurements imply the disc is.

Chapter 11

Dynamical Vertical Structure of TDE Discs

11.1 Introduction

As discussed in Chapter 10 there is a great deal of uncertainty in how, and indeed if, TDE discs circularise. The hydrodynamic simulations of Hayasaki et al. (2016) and Bonnerot et al. (2016) have shown that, when apsidal precession is weak, circularisation of TDE discs depends on the assumed thermodynamics, in particular how efficient the disc is at radiating energy generated by shocks or turbulence. In their radiatively efficient case the disc loses any energy generated in shocks or from viscous heating (the behaviour of the fluid being polytropic in Hayasaki et al. (2016) and isothermal in Bonnerot et al. (2016)). In this limit circularisation is not efficient and the disc forms a thin eccentric ring. In the opposite radiatively inefficient limit the disc satisfies the appropriate thermal energy equation so that kinetic energy lost in shocks and to viscosity is converted to thermal energy in the disc. In this limit the disc rapidly circularises and heats up to form a thick accretion torus (Hayasaki et al., 2016; Bonnerot et al., 2016). Simulations by Shiokawa et al. (2015), also radiatively inefficient, additionally allowed for a temperature-dependent first adiabatic exponent to account for contributions to the pressure from both the radiation and the gas. The disc formed in Shiokawa et al. (2015) is thick as in the other radiatively inefficient models; however, the disc does not rapidly circularise and maintains moderate eccentricities throughout the disc. To date simulations of circularisation with explicit cooling physics have not been carried out.

In this chapter we apply the equations for the vertical structure of a thin viscous eccentric disc of radiation and gas, developed in Chapter 7, to the highly eccentric TDE discs. We solve these dynamical vertical structure equations for different assump-

tions about the equation of state of the gas and for different α-viscosity laws. The vertical motion, radiative physics and viscous heating are important for understanding dissipative processes in TDEs, which is key to determining whether TDEs circularise.

11.2 Scale height variation in TDE discs

In this section we study the dynamics of a pure gas ($\Gamma_1 = \Gamma_3 = \gamma$) in a disc with geometry appropriate for an eccentric TDE. We solve Equations 7.35 and 7.39 of Chapter 7; as there is only gas pressure in the disc there is only one choice for the pressure in the α-viscosity law: $p_v = p$.

Discs formed by TDEs represent an extreme environment for eccentric disc theory, which forces a very strong variation of the vertical structure around the orbit. In the initial TDE disc the details of the thermodynamics are inseparable from the dynamics of the vertical structure and can be a challenge to numerical methods. Ogilvie & Barker (2014) showed that for ideal discs the variation of the scale height becomes extreme in the limits $e \to 1$ or $q \to 1$ depending on γ. Both these limits are relevant to TDE discs, with the initial post-disruption orbits having $e \approx q \approx 1$ (see Section 10.5 in the preceding Chapter). This results in extreme compression of the disc near pericentre in the ideal theory (Ogilvie & Barker, 2014). It might be expected that viscous heating raises the pressure at pericentre and prevents the extreme compression seen in the ideal models; the reality is rather different.

We find that there are two limiting behaviours in these extreme discs, which are compared in Table 11.1. A pericentre-dominated limit, where the entropy in the disc is strongly peaked at pericentre, and a pericentre-suppressed limit, where the entropy peak at pericentre is much smaller and comparable to the entropy in the rest of the disc. These two limits have some similarities with the two limiting behaviours of the ideal disc. The pericentre-dominated disc typically occurs for lower γ, while the pericentre suppressed disc typically occurs when γ is higher.

Between these two extremes there is an intermediate behaviour where the peak at pericentre is comparable to the pericentre-dominated disc but is much broader and extends over a much larger fraction of the orbit. This typically occurs for discs with low γ but relatively large α_s.

The controlling factor that dictates which behaviour the disc has appears to be the strength of the viscous stresses associated with the vertical compression (i.e. the terms proportional to $\dot{H}$ in Equation 7.35). In the pericentre-dominated limit this stress

dominates over the pressure forces and horizontal viscous stresses, particularly around pericentre. It is comparable to the other forces in the pericentre-suppressed limit.

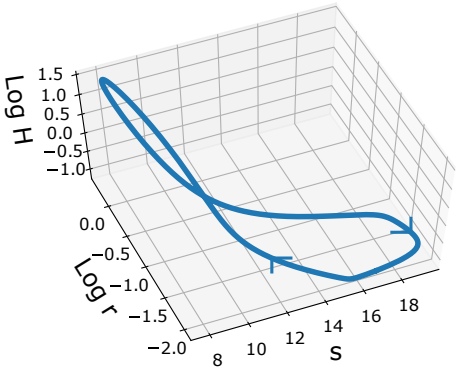
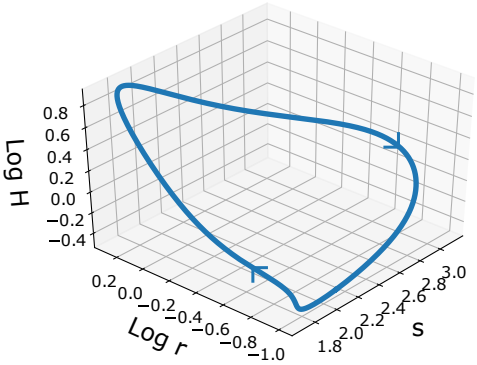
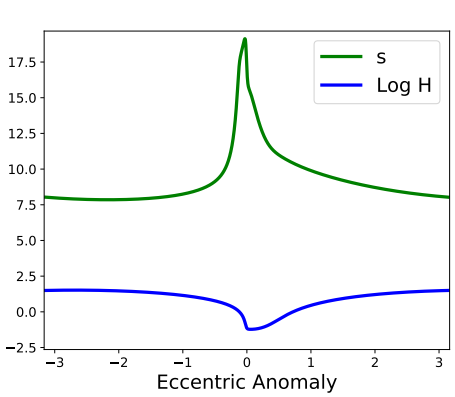
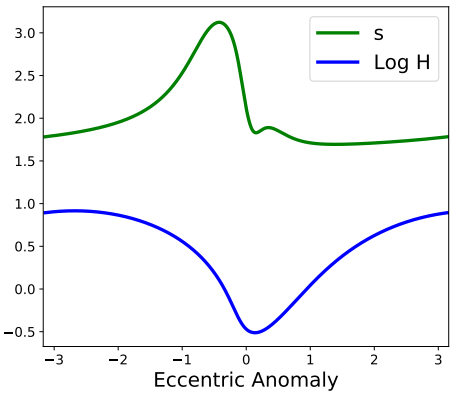
One feature absent in our model is the presence of an orbital intersection at apocentre owing to the strong GR precession. However TDEs are expected to occur close to the maximum distance capable of disrupting the star (Stone & Metzger, 2016; Lu & Bonnerot, 2019); for lower mass SMBHs, or higher mass stars, the TDE occurs sufficiently far from the black hole that apsidal precession is weak. This means that the GR precession causes a slow evolution of the disc streamlines rather an intersection of the stream at apocentre, and its main effect on our model is through setting the dependence of e and ϖ on a. As we are dealing with a local model centred on one orbit this weak apsidal precession has no direct effect on our solutions. Our model is incapable of dealing with the strong apsidal precession seen in TDEs with pericentric distances comparable to the gravitational radius (e.g. as simulated in Hayasaki et al. (2013); Sądowski et al. (2016))

In our disc we have no orbital intersection (which would correspond to $q = \pm 1$) and there is no way for the model to be close to orbital intersection simultaneously at pericentre and apocentre without the eccentricity oscillating on a short lengthscale or the disc being highly twisted. In fact with the positive eccentricity gradient expected in a tidal disruption event the disc should only be near to an orbital intersection in the nozzle region at pericentre. For TDEs with smaller impact parameters GR precession will be stronger and we expect this will cause the disc to become increasingly twisted (corresponding to $E_0 \to \pi$) eventually resulting in an orbital intersection close to apocentre. We leave consideration of disc twist to future work.

11.3 Radiation/gas mixture

We now consider the dynamics of a TDE including both radiation and gas pressure. The disc's vertical structure is assumed to undergo vertical mixing (either from convection or the disc turbulence), which provides an additional source of heat transport that is comparable to the radiative flux near the midplane. It is assumed that the vertical mixing sets up a polytropic vertical structure with polytropic index 3, which implies that β_r is independent of height. In the radiation pressure dominated limit (expected to be relevant for TDE discs, e.g. see calculations in Loeb & Ulmer (1997); Shen & Matzner (2014)) this leads to a vertically isentropic structure. This assumption is required for the equations to remain separable, although it has no other effect on the

Table 11.1 Comparison of the pericentre-dominated and pericentre-suppressed cases.

Pericentre-dominated	Pericentre-suppressed
Characterised by: Sharp peak in entropy at pericentre which is significantly higher than the rest of the disc	Characterised by: Modest peak in entropy at pericentre that is comparable to the maximum entropy in the rest of the disc
Caused by: • lower γ • higher e	Caused by: • higher γ • higher q
Example disc parameters: $e = 0.99$, $q = 0.9$, $\gamma = 4/3$, $E_0 = 0$, $\alpha_s = 0.01$	Example disc parameters: $e = 0.9$, $q = 0.9$, $\gamma = 5/3$, $E_0 = 0$, $\alpha_s = 0.1$
	
	
Disc showing the pericentre-dominated behaviour. As it returns to pericentre the fluid converges in a vertical ballistic motion, after which it undergoes very strong heating and cooling. In this limit the fluid column can be thought of as collapsing as it approaches pericentre and undergoing abrupt heating and momentum loss causing the fluid to bounce. On exiting pericentre the fluid is accelerated by gas pressure along the receding arc.	Disc showing the pericentre-suppressed behaviour. As with the high-e case strong heating occurs at pericentre. However the dominant source of heating comes from horizontal compression along the approaching arc. The fluid forces in the pericentre-suppressed limit are strong enough that the fluid motion on the return trajectory departs strongly from the quasiballistic motion seen in the pericentre-dominated limit.

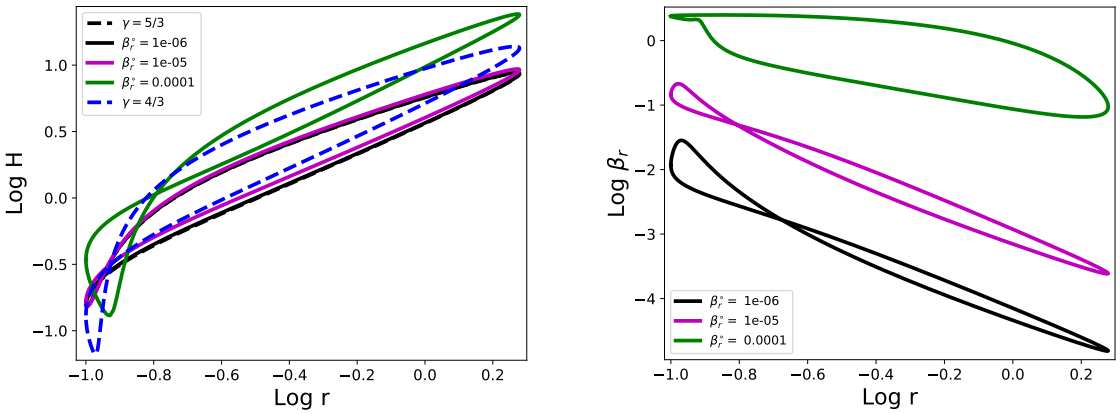


Fig. 11.1 Variation of the scale height (left) and β_r (right) of the disc for a radiation-gas mixture with $p_v = p$. The dashed lines are for pure gas discs with γ of 5/3 and 4/3. Disc parameters are $\alpha_s = 0.01$, $\alpha_b = 0$, $e = q = 0.9$ and $E_0 = 0$. For low β_r° the disc behaves like the pure gas disc with $\gamma = 5/3$. Discs with larger β_r° are unstable to the thermal instability.

variation of the scale height around the orbit. Γ_3 is now given by Equation 7.13 and we solve Equations 7.35 and 7.36.

11.3.1 Stress scaling with total pressure ($p_v = p$)

Figure 11.1 shows how the scale height and β_r vary around the orbit for different β_r° for a disc with $e = q = 0.9$, $\alpha_s = 0.01$ and $E_0 = \alpha_b = 0$ and stress scaling with total pressure ($p_v = p$). In general the β_r of the disc can exceed β_r° by several orders of magnitude as the physical conditions in the eccentric disc are very different from those of the circular reference state, primarily due to the enhanced heating rates occurring at pericentre.

The importance of radiation pressure is greatest in the strong compression experienced during pericentre passage. However even at pericentre β_r cannot significantly exceed 1 without triggering the thermal instability. This is distinct from the behaviour of the pure gas with $\gamma = 4/3$, normally used to mimic a radiation-dominated disc, as we find this to be thermally stable. This highlights the difference between a $\gamma = 4/3$ gas and a radiation-dominated disc when heating and cooling are present. The opposite limit $\beta_r^\circ \to 0$ matches the behaviour of the pure gas solutions with $\gamma = 5/3$.

The initial TDE disc tends to be highly radiation pressure dominated (Loeb & Ulmer, 1997; Shen & Matzner, 2014), for instance the estimate in Loeb & Ulmer (1997) gives $\beta_r \approx 10^4$. Similarly Zanazzi & Ogilvie (2020) considered TDEs with $\beta_r \approx 10^4$

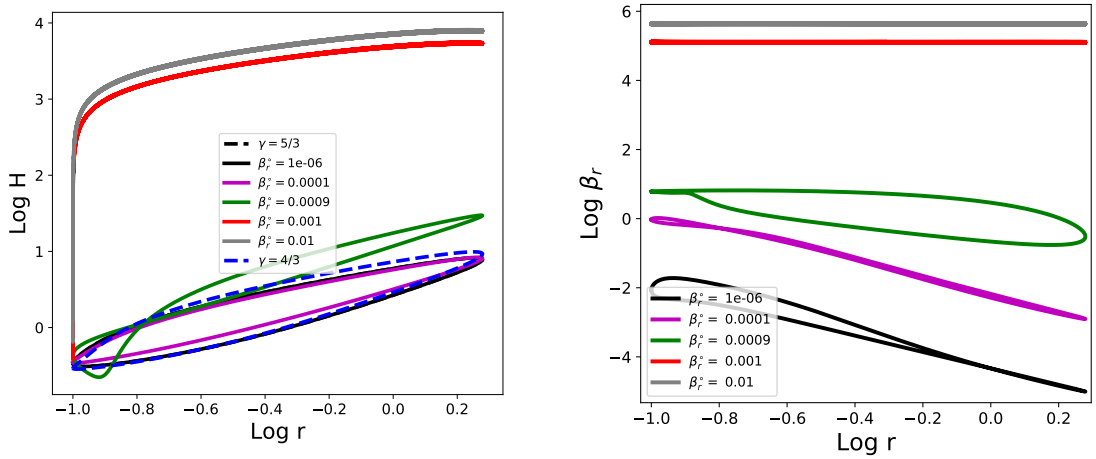


Fig. 11.2 Same as Figure 11.1 but with $p_v = p_g$ and $\alpha_s = 0.1$. Much larger β_r can be achieved owing to the stabilisation of the thermal instability. There appears to be an abrupt change in behaviour between $\beta_r^\circ = 9 \times 10^{-4}$ and $\beta_r^\circ = 1 \times 10^{-3}$ where the disc becomes approximately adiabatic. In the radiation dominated limit a nearly constant β_r implies a nearly constant entropy (see Equation 7.38).

while Bonnerot & Lu (2019) found β_r of 10^5. While we can slightly increase β_r by reducing α_s, in general our models are incapable of reaching as high β_r as the values quoted, owing to the presence of the thermal instability. In our model, thermally unstable circular discs can collapse to a thin gas pressure dominated state or experience thermal runaway with the temperature and disc thickness increasing with each orbit. Our thermally unstable highly eccentric discs almost universally experience thermal runaway.

11.3.2 Stress scaling with gas pressure ($p_v = p_g$)

In order to reach realistic β_r we test solutions where the stress scales with gas pressure ($p_v = p_g$). Figure 11.2 show how the scale height and β_r vary around the orbit for different β_r° for this case. The disc parameters are $e = q = 0.9$ and $E_0 = \alpha_b = 0$ like those considered in Figure 11.1; however we set $\alpha_s = 0.1$. These can now attain realistic values of β_r without encountering an unbounded thermal runaway.

When β_r° is sufficiently high the disc becomes strongly radiation pressure dominated and behaves nearly adiabatically. The transition to this behaviour as we vary β_r° is quite abrupt and may represent a change in solution branch. Our interpretation of this result is that, for highly eccentric discs with $p_v = p_g$, there are two solution branches: a gas pressure dominated solution branch with significant entropy variation and a nearly

adiabatic radiation pressure dominated solution branch. When β_r° is sufficiently high the gas pressure dominated branch is thermally unstable and the solution switches to the other branch. This nearly adiabatic radiation pressure dominated branch seems to be peculiar to eccentric discs as the circular disc has only one solution branch, which is stable regardless of β_r°. When stress scales with total pressure the nearly adiabatic radiation pressure dominated branch doesn't appear to exist (or if it does it is always unstable) as the solution undergoes an unbounded thermal runaway.

We find that for our radiation pressure dominated discs there is a long lived transient state in which the solution is well approximated by an adiabatic solution on a orbital timescale, but undergoes a slow drift in (orbit-averaged) entropy over many orbits. This is related to the freedom to specify the disc entropy in a strictly ideal disc. Eventually the solution settles down to a fixed orbit-averaged entropy; there is still entropy variation around the orbit but this is small compared with the orbit-averaged entropy and has little effect on the dynamics.

The radiation dominated solution and the $\gamma = 4/3$ perfect gas with $\alpha_s \ll 1$ go through the same nearly adiabatic transient entropy drift phase before converging on a specific value for the orbit-averaged entropy. However, owing to the difference in the entropy/temperature for a $\gamma = 4/3$ gas and a radiation dominated disc, these converge to different nearly adiabatic solutions! This leads to a disagreement between eccentric radiation pressure dominated discs and eccentric $\gamma = 4/3$ discs, which was discussed in Section 7.5 of Chapter 7.

Both the radiation dominated solution and $\gamma = 4/3$ adiabatic (or $\alpha \ll 1$) solution exhibit extreme compression near pericentre (see also Ogilvie & Barker (2014)). The reader might be concerned that the H can reach $10^3 - 10^4$ at apocentre and appears to strongly violate the thin-disc assumptions. However, the disc scale height is normalised with respect to the scale height of a reference circular disc, which can be very thin. For a typical $1M_\odot$ TDE around a $10^6 M_\odot$ black hole, however, $H^\circ/a \sim 10^{-3}$ (see Section 10.5 in the preceding Chapter), so H/r is ~ 1 at apocentre and the thin disc assumptions are violated. These discs start out geometrically thin around the entire orbit, prior to the onset of thermal instability. In general, the onset of thermal instability can be treated by our model, while the outcome, even for $p_v = p_g$, cannot. It is likely that additional physics, such as a coherent magnetic field, will reduce or prevent the extreme behaviour seen in the radiation dominated models. We shall explore this possibility further in the next chapter.

If we increase α_s it is possible to obtain radiation pressure dominated solutions which exhibit significant entropy variation around the orbit. In these solutions viscosity

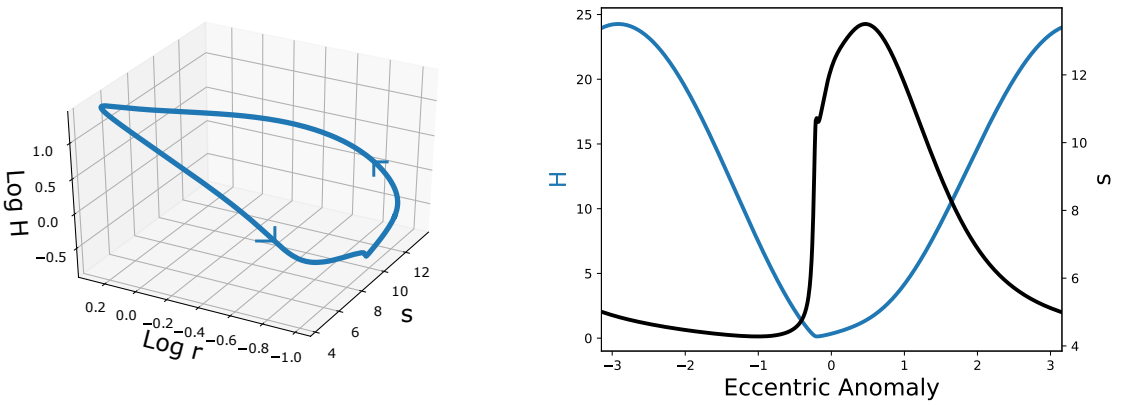


Fig. 11.3 Disc with radiation-gas mixture with $p_v = p$, $\beta_r^\circ = 10^{-4}$, $\alpha_s = 0.01$, $\alpha_b = 0$, $e = q = 0.9$ and $E_0 = 0$. The entropy varies strongly around the orbit as in the $\gamma = 4/3$ disc; however, after an abrupt increase in entropy during pericentre passage, the entropy declines gradually over half an orbit.

and radiative cooling are again important and they have a similar behaviour to the marginally gas pressure dominated solutions in the models where stress scales with total pressure (but attain higher β_r at pericentre). However, these solutions typically require $\alpha_s > 1$ and it's not clear that this is physically reasonable, although this α_s is defined with respect to the gas pressure.

Figure 11.3 shows the solution for one of the most radiation dominated discs we can achieve when stress scales with total pressure (with disc parameters $\beta_r^\circ = 10^{-4}$, $\alpha_s = 0.01$, $\alpha_b = 0$, $e = q = 0.9$ and $E_0 = 0$). Similarly Figure 11.4 shows a radiation dominated disc when stress scales with gas pressure (with disc parameters $\beta_r^\circ = 0.001$, $\alpha_s = 0.1$, $\alpha_b = 0$, $e = q = 0.9$, and $E_0 = 0$). Both exhibit abrupt changes in the entropy near pericentre, coincident with a reversal in the direction of vertical motion; we examine this behaviour in more detail in Section 11.4 below. The variation in entropy in the latter solution is small compared to the total entropy and this solution is close to adiabatic.

Finally Figure 11.5 shows examples of solutions with a different choice for the orbital frequency $\omega_{\rm orb}$ in the $\alpha-$prescription (here $\omega_{\rm orb} = \Omega$). The solutions have broadly the same physical behaviour as those with $\omega_{\rm orb} = n$, with the thermal instability requiring larger β_r° (although it can still occur for gas pressure dominated reference discs). These solutions have a more pronounced compression at pericentre than their $\omega_{\rm orb} = n$ counterparts, which appears to be characteristic of discs with weaker viscosities. Adopting a slightly larger value for α_s removes this difference.

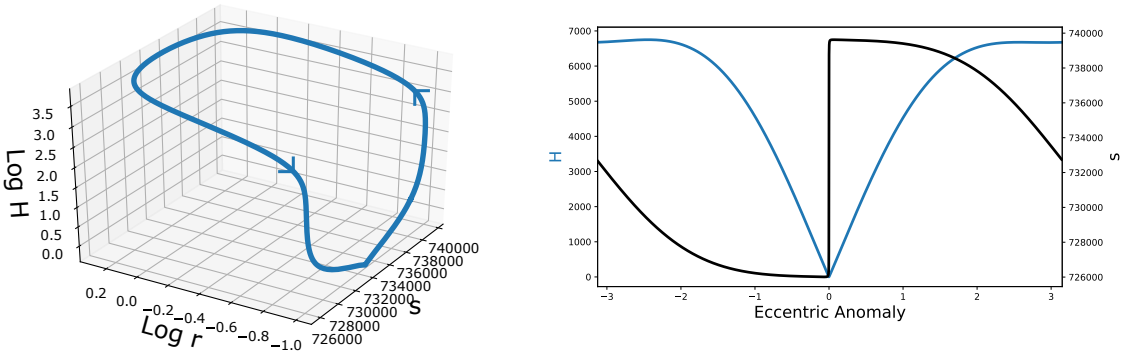


Fig. 11.4 Disc with radiation-gas mixture with $p_v = p_g$, $\beta_r^\circ = 0.001$, $\alpha_s = 0.1$, $\alpha_b = 0$, $e = q = 0.9$, and $E_0 = 0$. The entropy shows very similar features to that seen in Figure 11.3, in particular the abrupt jump at pericentre, however the relative change in entropy around the orbit is small. The scale height is almost symmetric about pericentre indicating a lack of dissipation.

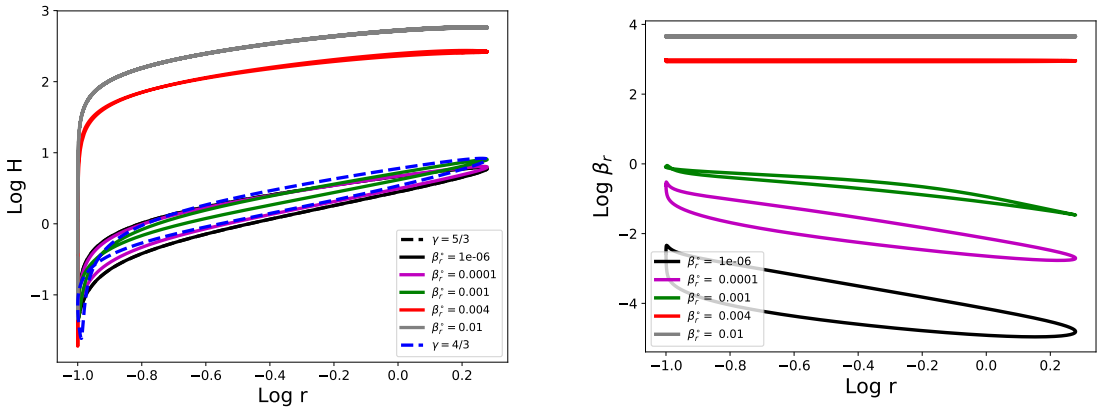


Fig. 11.5 Same as Figure 11.2 but with $\omega_{\rm orb} = \Omega$, showing similar physical behaviour with the alternative choice of frequency. The onset of the thermal instability for this model occurs at higher β_r°.

11.3.3 Implications for TDE thermodynamics

So what can we conclude about the disc thermodynamics? When stress is proportional to total pressure (e.g. Figure 11.3) we have strong variation of the entropy around the orbit so our model is not well described by an adiabatic or isothermal disc. The thermal physics of our disc is quite different from the simplified thermodynamic models considered in Hayasaki et al. (2016) and Bonnerot et al. (2016). The $\gamma = 4/3$ pure gas model is not a good model of a radiation pressure dominated disc when stress scales with total pressure. In particular we find that the radiation dominated discs are thermally unstable, while the $\gamma = 4/3$ disc is stable. It is likely that the simulations of TDEs do not encounter the thermal instability owing to their simplified treatment of thermodynamics.

For our radiation dominated solutions (with stress scaling with gas pressure) the relative entropy variation around the orbit is small and, on an orbital timescale, the solution can be well approximated by an adiabatic solution. However this apparent adiabaticity is only maintained because the oscillator is extracting energy from the orbital motion to make up for radiative losses. On longer timescales the orbits must evolve due to radiative losses, which will cause the entropy of the solution to adjust. That the entropy of our radiation dominated solutions can adjust is attested by the fact that the entropy grows by many orders of magnitude (from the starting s close to 1) as the solution relaxes to a periodic state.

Figure 11.6 shows the timescale (in orbital periods) for an orbit to radiate its entire orbital energy, given by the dimensionless quantity,

$$\frac{|\varepsilon_{\rm orbital}|}{\langle \mathcal{C}_a \rangle P_{\rm orbital}} = \frac{\tau^\circ}{16\pi\lambda}\left(\frac{H^\circ}{a}\right)^{-1}\left(\frac{na}{c}\right)\langle T^4 J^2\rangle^{-1}, \tag{11.1}$$

where $\varepsilon_{\rm orbital} = -\frac{1}{2}M_a n^2 a^2$ is the orbital energy (per unit semimajor axis), $P_{\rm orbital}$ is the orbital period, $\langle \mathcal{C}_a \rangle$ is the orbit averaged dimensionful cooling rate (per unit semimajor axis), and τ° and H°/a are the optical depth and aspect ratio of the reference circular disc. The reference disc has been scaled to a disruption of $1M_\odot$ star around a $10^6 M_\odot$ black hole. The thermally stable solutions take a large number ($\sim 10^4$) orbits to radiate their orbital energy, as the discs are very thin. For strongly radiation pressure dominated discs the disc is thermally unstable and, for $p_v = p_g$, becomes much thicker, allowing the orbital energy to be radiated on ~ 10 orbits. It is likely that the thermally unstable $p_v = p$ discs (not shown here) also radiate their orbital energy on much shorter timescales, potentially before they reach the unphysically large H seen when we attempt to run these models to convergence.

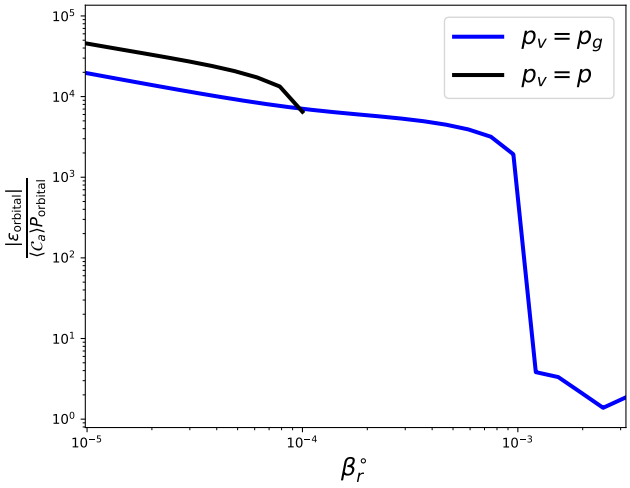


Fig. 11.6 Timescale (in orbital periods) for the orbit to radiate its entire orbital energy for the disruption of a $1M_\odot$ star around a $10^6 M_\odot$ black hole. For gas pressure dominated to moderately radiation pressure dominated discs the disc takes $\sim 10^4$ orbits to radiate away the orbital energy, as the discs are extremely thin. The cutoff for the $p_v = p$ line at $\beta_r^\circ = 10^{-4}$ is due to the onset of the unbound thermal instability. For the $p_v = p_g$ discs the thermal instability causes an abrupt drop in the time to radiate the orbital energy as the disc becomes thick.

We caution that the timescale given by $\frac{|\varepsilon_{\rm orbital}|}{\langle \mathcal{C}_a \rangle}$ should be viewed as the timescale for the orbit to change and not as a circularisation timescale. Whether the disc circularises will also depend on angular momentum transport, which is difficult to address within a fully local model. Even at fixed angular momentum a highly eccentric disc must radiate significantly more that its orbital energy to completely circularise. For instance an orbit with $e = 0.9$ must radiate $\sim 4|\varepsilon_{\rm orbital}|$ to fully circularise (at fixed angular momentum), while for an orbit with $e = 0.99$ this increases to $\sim 50|\varepsilon_{\rm orbital}|$.

In all our models the disc radiates any energy it gained from viscous heating so cooling is fairly efficient; however the strong temperature variation around the orbit shows that this model is far from the isothermal model used in Bonnerot et al. (2016) to represent an efficiently cooling disc. In some respects our highly radiation pressure dominated models are similar to the radiatively efficient polytropic model of Hayasaki et al. (2016) (albeit with $\gamma = 4/3$) on an orbital timescale. Although, as noted above, the disc can lose or gain entropy through radiative cooling or viscous heating on longer timescales.

It is important to note that orbit-averaged heating equalling orbit-averaged cooling isn't an assumption of our model, it is in fact a requirement for $2\pi-$periodicity. Any

imbalance would cause the disc to heat or cool until, if the disc is thermally stable, a balance between averaged heating and cooling is restored. In a real TDE, however, the orbit may change too rapidly to settle down into the periodic solutions attained here.

11.4 Pericentre passage

The nozzle-like structure at pericentre, involving strong heating and cooling, seen in both our pericentre-dominated models and the radiation-gas mixture, may be analogous to the 'nozzle shocks' discussed in the TDE literature (Evans & Kochanek, 1989; Kochanek, 1994; Guillochon et al., 2014; Shiokawa et al., 2015; Piran et al., 2015). These nozzle shocks are expected to arise because of the convergence of the streamlines, both horizontally and vertically, towards their common pericentre. In our model this corresponds to q approaching 1 with $E_0 = 0$ so that both horizontal and vertical compression occur at pericentre. In the pericentre-dominated limit the horizontal compression (while present) is less important and the nozzle-like structure is predominantly due to the vertical motion.

Figure 11.7 shows the pericentre passage for a radiation-gas mixture with, $p_v = p$, $\beta_r^\circ = 10^{-4}$, $\alpha_s = 0.01$, $\alpha_b = 0$, $e = q = 0.9$, and $E_0 = 0$. This shows a slice which follows the orbital geometry (i.e. $a = \text{const}$) and is therefore different from what is typically shown in simulation papers which is typically a planar slice through pericentre.

The significant asymmetry between the disc before and after pericentre passage is caused by the dissipation of the vertical velocity in the nozzle. This feature is apparent in the abrupt change in entropy just before pericentre. Generally the more dissipative the model the more asymmetric the nozzle structure is and the earlier (relative to pericentre passage) it occurs.

Our results differ from the expectations of a 'nozzle shock'. As we include viscosity, a shock in our model would not be discontinuous and would have a finite thickness. However, the size of the (continuous) entropy jump near pericentre is dependent on α_s and vanishes in the inviscid limit. In fact the ideal discs possess no shock structure, strongly suggesting that the fluid can pass through the nozzle without encountering a large-scale "nozzle shock" [1]. Our models terminate at a photosphere; it is possible there could be different dynamics above the photosphere which might include shocks, but this probably accounts for a small fraction of the mass in the disc.

[1]This does not rule out small-scale shocks associated with turbulent motion, particularly for solutions where the viscous stress exceeds the pressure.

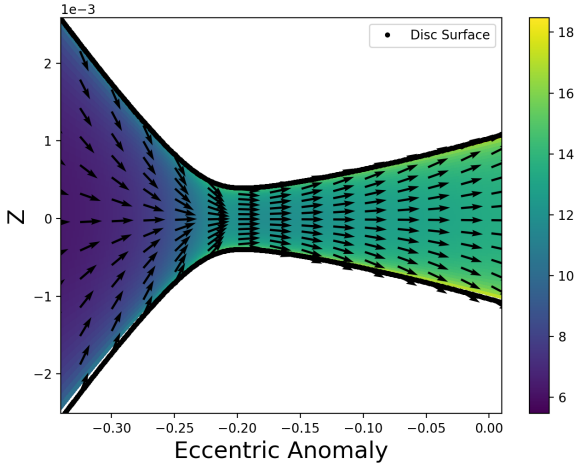


Fig. 11.7 Pericentre passage for a radiation-gas mixture with $p_v = p$, $\beta_r^\circ = 10^{-4}$, $\alpha_s = 0.01$, $\alpha_b = 0$, $e = q = 0.9$, and $E_0 = 0$. Arrows show the flow direction and colour indicates entropy. The vertical coordinate is expressed in units of a, using the scale height of the reference circular disc H° calculated in Appendix 10.5.

Figure 11 of Shiokawa et al. (2015) shows shocks just after pericentre in their hydrodynamical simulation. These shocks do not extend to the midplane, and may be present because of the finite (and large) thickness of the disc. Their Figure 11 is a snapshot taken fairly early in their simulation and may not be directly comparable to our calculations which assume the TDE has had time to settle into an eccentric disc.

11.5 Discussion

11.5.1 Stability of the solutions

An important consideration for any steady or periodic solution of a dynamical system is its stability against perturbations. One advantage of our method is that the (in principle nonlinear) stability of our solutions to perturbations of $H, \dot{H}$ and T is guaranteed as the relaxation method should only be able to find locally attracting solutions. As discussed in Chapter 7 our solutions converge to the periodic solution, from a range of initial guesses, on approximately the thermal timescale. This would imply that nonlinear perturbations to our solution should damp out on the thermal timescale or faster.

For some parts of our parameter space our method fails to converge to a 2π-periodic solution. This does not necessarily mean that such a solution does not exist, but

may instead indicate the 2π-periodic solution is unstable. For instance our method fails to find a 2π-periodic solution when β_r° and α_s are sufficiently large that the disc becomes thermally unstable. In general we find the only way to obtain highly radiation dominated discs, as expected in TDEs, is by assuming that stress scales with gas pressure. We note that a pure gas disc with $\gamma = 4/3$ is thermally stable; however this is a poor model of a radiation dominated disc when dissipation and cooling are involved.

It may be that TDE discs are genuinely thermally unstable; however such an instability is unlikely to manifest itself in simulations which use simplified thermodynamics. If the thermal instability is present then the resulting thermal runaway could cause the TDE disc to become substantially thicker, potentially making it much easier to circularise the disc through enhanced dissipation. Our model is incapable of exploring this possibility, and current simulations lack the necessary thermodynamics, so a new approach would be needed to study thermally unstable TDE discs.

Our solution method doesn't tell us about the stability of our solutions to perturbations of short wavelength (comparable or less than the scale height). So our disc structure could be unstable to such perturbations. In particular it is likely our discs are unstable to the growth of inertial waves via the parametric instability (Papaloizou, 2005a,b; Wienkers & Ogilvie, 2018; Barker & Ogilvie, 2014; Pierens et al., 2020).

11.5.2 Extreme behaviour during pericentre passage

Many of our models exhibit extreme behaviour near pericentre. This is particularly true for the more radiation-dominated discs (where these are stable). In general we find that it is difficult to reverse the collapse of the fluid column during pericentre passage in the more radiation-dominated discs without encountering either extreme compression (for the nearly adiabatic discs) or unphysically strong viscous stresses (for the $\gamma = 4/3$ pure gas disc and the marginally radiation-dominated discs).

With the exception of the strongly radiation-dominated nearly adiabatic solutions, most of our solutions have a very strong vertical viscous stress at pericentre – often comparable to or exceeding the pressure. This is particularly relevant for the pure gas models with $\gamma = 4/3$, where we find that the dominant balance at pericentre is between the vertical motion and the viscous stress, which is responsible for reversing the collapse of the fluid. This issue occurs to a lesser extent in radiation-gas mixtures (with either stress model) with $\beta_r \approx 1$, where the vertical viscous stress at pericentre is typically comparable to or slightly larger than the pressure.

Viscous stresses exceeding pressure creates a problem peculiar to turbulent/α stresses. This is because, if the viscous stresses are sourced from a turbulent velocity

field, viscous stresses exceeding pressure imply the turbulent velocity is supersonic. Supersonic turbulence will cause shocks which should either damp the turbulent eddies or shock heat the gas causing the turbulence to become subsonic. A such the appearance of vertical viscous stresses exceeding pressure in our models should call into question the validity of the $\alpha-$prescription during pericentre passage.

For the nearly adiabatic radiation pressure dominated discs the viscous stresses are well behaved, however we encounter extreme compression during pericentre passage. This is the same as the extreme behaviour seen in the truly adiabatic discs (e.g. see Ogilvie & Barker (2014) and Zanazzi & Ogilvie (2020). Such extreme behaviour merits caution when interpreting or accepting the results as it is likely that some neglected physical effect may regulate the behaviour near pericentre. In particular the discs we consider implicitly contain a magnetic field (responsible for the $\alpha-$viscosity) and a real magnetic field might be able to resist such an extreme compression.

In all our (stable) radiation pressure dominated discs there is thus a question as to how the magnetic field responds to the compression experienced during pericentre passage. Is the magnetic pressure capable of reversing the collapsing fluid column? Can the MRI still operate during pericentre passage, and if so how quickly can it respond to the change in pressure?

Very little work has been done on magnetic fields in highly eccentric discs. While linear growth of MRI in highly eccentric discs has been studied in Chan et al. (2018) they made the assumption that the fluid motion was divergence free and thus MRI can be treated as incompressible. The presence of the breathing mode in eccentric discs causes compressional motion in the disc, so the results of Chan et al. (2018) are probably only applicable to weakly eccentric discs. As is clear from our results the dominant stress at pericentre in a highly eccentric disc is that resulting from the vertical fluid motion, so the results of Chan et al. (2018) are not valid in this limit. It's not clear if the turbulent stress should continue to scale with pressure, be shut off completely (e.g. by the concentration of magnetic fields to sufficient strengths to shut off the MRI) or be in some sense "frozen in" with the turbulence not having time to respond to the extreme conditions at pericentre.

One way of capturing the finite response timescale of MRI turbulence is through the viscoelastic model of Ogilvie (2000), Ogilvie (2001) and Ogilvie (2002a). Using this viscoelastic prescription with a relaxation time comparable to the orbit period, the fluid should behave nearly elastically near pericentre. This will weaken the dissipative contribution to the stress that is associated with the disc turbulence. The weaker

dissipation during pericentre passage may be a better model of MHD turbulence, that remains subsonic, than the simpler viscous treatment considered here.

In the next Chapter we will look at the behaviour of the magnetic field in a highly eccentric disc and its effects on the vertical structure. We will also look at what effect the finite response time of the MRI has on the dynamics.

11.5.3 Implication of disc surface height variation for observational claims of highly eccentric TDE discs

One motivation for considering highly eccentric disc in TDEs is the fits to the H α emission made by Liu et al. (2017) Cao et al. (2018) and Holoien et al. (2019). In these papers the authors proposed that the H α emission originates from the reprocessing of X-rays in a highly eccentric disc ($e \approx 0.97$ for Liu et al. (2017); Cao et al. (2018); the disc of Holoien et al. (2019) has more moderate eccentricity with $e \leq 0.25$). The discs considered in their models are flat, being confined to a plane. As we have demonstrated in this chapter, this cannot be the case for a real disc as the height of the disc surface in such an eccentric disc must strongly vary around the orbit. The model of reprocessing in a flat eccentric disc would only be valid if the disc was extremely thin, such that the variation of the height of the disc surface around the orbit was negligible relative to the radial extent of the disc. This is unlikely to be the case in a real TDE disc. While present this concern is much less important for the model proposed by Hung et al. (2020) owing to its much lower eccentricity.

Another overlooked factor is that from certain viewing angles the emission from pericentre will be blocked by the much thicker material at apocentre. This is probably particularly relevant to the disc model of Liu et al. (2017) as their disc is both highly eccentric and close to edge on. Given the proposed orientation of the disc the much thicker material coming into pericentre should block the view to material leaving pericentre which could significantly reduce the amplitude of the redshifted peak. More work needs to be done to understand the look-angle effects on line emission from highly eccentric discs.

11.6 Conclusion

We have applied a non-ideal theory of eccentric discs to tidal disruption events (TDEs) discs in order to better understand their dynamical vertical structure and its variation around the elliptical orbits of a TDE disc. We have included the thermal physics

appropriate to a mixture of gas and radiation undergoing viscous heating and radiative cooling. This has a significant role in determining the structure and evolution of TDE discs and will be important in understanding to what extent they circularise prior to accretion. In summary our results are:

1. The eccentricity in the TDE discs forces a strong variation of the scale height and entropy around the orbit.

2. The disc is out of equilibrium. It is thick where it is cold and thin where it is hot, the opposite of a disc in hydrostatic equilibrium. Additionally a highly eccentric disc can attain very high temperatures at pericentre while remaining relatively thin around the entire orbit.

3. We find a peaked structure in the entropy with a rapid change of the vertical velocity at pericentre which has much in common with the nozzle shaped flows seen in simulations. However, our results suggest the fluid can pass through the nozzle without the formation of a large-scale "nozzle shock".

4. In many of our solutions the vertical viscous stress can exceed gas pressure. This would suggest that turbulence was supersonic – implying a breakdown of the $\alpha-$prescription in these extreme environments.

Finally and perhaps most importantly:

6. The thermal instability operates in radiation-dominated discs for both stress scaling with total pressure and stress scaling with gas pressure. In the latter case there exists an additional nearly adiabatic radiation-dominated branch which is stable and acts as an attractor for the thermal instability. It does not operate in a $\gamma = 4/3$ perfect gas used as a model of a radiation-dominated disc.

In the next chapter we will explore alternative stress models which may better capture the effects of a magnetic field in an eccentric disc and lead to less extreme behaviour in the radiation dominated regime. This will also allow us to address the breakdown of the $\alpha-$prescription in highly eccentric discs. Additionally a more realistic stress model may change the thermal stability of the disc.

Chapter 12

Magnetic Fields in TDEs

12.1 Introduction

Two issues were highlighted in the last chapter. One was the thermal instability of radiation pressure dominate, highly eccentric, discs when stress scales with total pressure, a result that has long been known for circular discs (Shakura & Sunyaev, 1976; Pringle, 1976; Piran, 1978). Circular radiation pressure dominated discs can be stabilised by assuming stress scales with gas pressure (Meier, 1979; Sakimoto & Coroniti, 1981). For highly eccentric discs it appears that the thermal instability is still present when stress scales with gas pressure; however there exists a stable radiation pressure dominated branch which is the outcome of the thermal instability. For realistic TDE parameters this branch is typically too thick for the thin disc approximation to be valid.

The second issue was the extreme behaviour that can occur during pericentre passage. For the radiation pressure dominated disc where stress scales with gas pressure the solution is nearly adiabatic and undergoes extreme compression near pericentre. In models where the viscous stresses contribute to the dynamics we typically found that the vertical viscous stress is comparable to or exceeds the (total) pressure, which is possibly problematic for the α model as it would indicate transonic turbulence. In some of the solutions the vertical viscous stress can exceed the total pressure by an order of magnitude, strongly violating the assumptions of the $\alpha-$model.

In this chapter we focus on the second of the two issues by considering alternative turbulent stress models which are better motivated by the physics of the underlying magnetic field to see if this resolves some of the extreme behaviour seen in the $\alpha-$disc models. We will also see if alternative stress models are more thermally stable than

the $\alpha-$disc model, although it's possible the solution to this issue is outside the scope of a thin disc model.

Two additional physical effects, not present in an $\alpha-$disc model, may be important for regulating the extreme behaviour at pericentre. One is the finite response time of the MRI (see for instance the viscoelastic model of Ogilvie (2001) and discussion therein) which means the viscous stress cannot respond instantly to the rapid increase in pressure and velocity gradients during pericentre passage, potentially weakening the viscous stresses so they no longer exceed the pressure. Another is the relative incompressibility of the magnetic field, compared with the radiation or the gas, with the magnetic pressure providing additional support during pericentre passage which could prevent the extreme compression seen in some models.

Various attempts have been made to rectify some of the deficiencies of the $\alpha-$prescription using alternative closure models for the turbulent stress. Ogilvie (2000) and Ogilvie (2001) proposed a viscoelastic model for the dyadic part of the Maxwell stress (i.e. the contribution from magnetic tension $\frac{B^i B^j}{\mu_0}$) to account for the finite response time of the MRI. It was shown in Ogilvie & Proctor (2003) that for incompressible fluids there is an exact asymptotic correspondence between MHD in the limit of large magnetic Reynolds number and viscoelastic fluids (specifically an Oldroyd-B fluid) in the limit of large relaxation time. Ogilvie (2002a) improved upon the viscoelastic model of Ogilvie (2000) and Ogilvie (2001) by including an isotropic part of the stress to model the effects of magnetic pressure and correcting the heating rate so that total energy is conserved. Ogilvie (2003) proposed solving for both the Maxwell and Reynolds stresses and proposed a nonlinear closure model based on requiring the turbulent stresses to exhibit certain properties (such as positive definiteness, and relaxation towards equipartition and isotropy) known from simulations and experiments.

Simulations of MRI in circular discs typically find that the magnetic pressure tends to saturate at 10% of the gas pressure. However in the local linear stability analysis of Pessah & Psaltis (2005) the toroidal magnetic field only stabilises axisymmetric MRI modes when it is highly suprathermal (specifically when the Alfvén speed exceeds the geometric mean of the sound speed and Keplerian speed). Das et al. (2018) confirmed this result persists in a global linear eigenmode calculation. Similarly Ogilvie & Pringle (1996) found that the most unstable non-axisymmetric MRI mode is only stabilised by the magnetic field when the magnetic and rotational energies are approximately in equipartition. In light of this, Begelman & Pringle (2007) have suggested that, for a strongly magnetised disc, the viscous stress may scale with the magnetic pressure and showed that such a disc would be thermally stable even when radiation pressure

dominates over gas pressure. Such a disc was simulated by Sądowski et al. (2016), who indeed found thermal stability.

12.2 Derivation of the equations of vertical structure, including thermal effects

In the previous chapter we considered a purely hydrodynamic model of a TDE disc, that made use of an α-viscosity prescription, based on the local model developed in Chapter 7. In a similar vein we here develop a local model that allows for a more general treatment of the turbulent/magnetic stress.

The equations, formulated in a frame of reference that follows the elliptical orbital motion, are the vertical equation of motion,

$$\frac{Dv_z}{Dt} = -\frac{GM_\bullet z}{r^3} - \frac{1}{\rho}\frac{\partial}{\partial z}\left(p + \frac{1}{2}M - M_{zz}\right), \tag{12.1}$$

the continuity equation,

$$\frac{D\rho}{Dt} = -\rho\left(\Delta + \frac{\partial v_z}{\partial z}\right), \tag{12.2}$$

and the thermal energy equation,

$$\frac{Dp}{Dt} = -\Gamma_1 p\left(\Delta + \frac{\partial v_z}{\partial z}\right) + (\Gamma_3 - 1)\left(\mathcal{H} - \frac{\partial F}{\partial z}\right), \tag{12.3}$$

where $\frac{D}{Dt}$, Δ, $F_{\rm rad}$ and $F_{\rm ext}$ have the same meanings as given in Chapter 7.

$$M^{ij} := \frac{B^i B^j}{\mu_0} \tag{12.4}$$

is a modified Maxwell stress tensor, which can include contributions from a large scale mean field and from the disc turbulence. Its trace is denoted $M = M^i{}_i$, which corresponds to twice the magnetic pressure. In this paper we shall explore two different closure models for time-evolution of the modified Maxwell stress tensor.

$\mathcal{H}$ is the heating rate per unit volume due to the dissipation of magnetic/turbulent energy. Following Chapter 7, we write the heating rate per unit volume as

$$\mathcal{H} = f_{\mathcal{H}} n p_v, \tag{12.5}$$

where $f_{\mathcal{H}}$ is a dimensionless expression that depends on the closure model.

In addition to the magnetic pressure, which is included through the $\frac{1}{2}M$ term, the pressure includes contributions from radiation and a perfect gas with a ratio of specific heats γ. We define the hydrodynamic pressure to be the sum of the gas and radiation pressure,

$$p = p_r + p_g = \frac{4\sigma}{3c}T^4 + \frac{\mathcal{R}\rho T}{\mu}, \tag{12.6}$$

and β_r to be the ratio of radiation to gas pressure:

$$\beta_r := \frac{p_r}{p_g} = \frac{4\sigma\mu}{3c\mathcal{R}}\frac{T^3}{\rho} \quad . \tag{12.7}$$

We assume a constant opacity law, applicable to the electron-scattering opacity expected in a TDE, with the opacity denoted by κ. We consider a radiation+gas mixture where the first and third adiabatic exponents, Γ_1, Γ_3, are given by Equations 7.12 and 7.13 of Chapter 7.

Similar to the hydrodynamic model, we propose a separable solution of the form

$$\begin{aligned} \rho &= \hat{\rho}(t)\tilde{\rho}(\tilde{z}), \\ p &= \hat{p}(t)\tilde{p}(\tilde{z}), \\ M_{ij} &= \hat{M}_{ij}(t)\tilde{M}(\tilde{z}), \\ F &= \hat{F}(t)\tilde{F}(\tilde{z}), \\ v_z &= \frac{dH}{dt}\tilde{z}, \end{aligned} \tag{12.8}$$

where

$$\tilde{z} = \frac{z}{H(t)} \tag{12.9}$$

is a Lagrangian variable that follows the vertical expansion of the disc.

In order to preserve separability the (modified) Maxwell stress M^{ij} must have the same vertical structure as the pressure ($\tilde{M} = \tilde{p}$), although we can have an additional contribution to the (modified) Maxwell stress which is independent of height. This has a couple of important consequences. In the absence of any contributions which are independent of height, this corresponds to a plasma-β, defined as the ratio of hydrodynamic to magnetic pressure $\beta_m := p/p_m$, independent of height. Additionally this has implications for the realisability of M^{ij}: for a large scale field any non-zero M^{zz} term must be independent of height in order that the underlying magnetic field

obeys the solenoidal condition. For small scale/turbulent fields this condition is instead that the mean of B^z must be independent of height, however M^{zz} has the same vertical structure as pressure.

The separated solution works provided that

$$\frac{d^2H}{dt^2} = -\frac{GM_\bullet}{r^3}H + \frac{\hat{p}}{\hat{\rho}H}\left[1 + \frac{\hat{M}}{2\hat{p}} - \frac{\hat{M}_{zz}}{\hat{p}}\right], \tag{12.10}$$

$$\frac{d\hat{\rho}}{dt} = -\hat{\rho}\left(\Delta + \frac{1}{H}\frac{dH}{dt}\right), \tag{12.11}$$

$$\frac{d\hat{p}}{dt} = -\Gamma_1\hat{p}\left(\Delta + \frac{1}{H}\frac{dH}{dt}\right) + (\Gamma_3 - 1)\left(f_{\mathcal{H}}n\hat{p}_v - \lambda\frac{\hat{F}}{H}\right), \tag{12.12}$$

$$\hat{F} = \frac{16\sigma\hat{T}^4}{3\kappa\hat{\rho}H} \quad , \tag{12.13}$$

$$\hat{p} = (1 + \beta_r)\frac{\mathcal{R}\hat{\rho}\hat{T}}{\mu} \quad , \tag{12.14}$$

where

$$\beta_r = \frac{4\sigma\mu}{3c\mathcal{R}}\frac{\hat{T}^3}{\hat{\rho}} \quad . \tag{12.15}$$

We must supplement these equations with a closure model for M^{ij} and $f_{\mathcal{H}}$.

We assume for a given β_m°, β_r° and n there exists an equilibrium solution for a circular disc and uses this solution to nondimensionalise the equations. As in the hydrodynamical model we use $^\circ$ to denote the equilibrium values in the reference circular disc (e.g. H°, T° etc). Depending on the closure model there can be multiple equilibrium solutions, some of which can be unstable (particularly in the radiation dominated limit). Our choice of solution branch, for a given closure model, is specified in Chapter 9.

Scaling H by H°, $\hat{T}$ by T°, M^{ij} by p°, t by $1/n$ and J by $2/n$ we obtain the dimensionless version

$$\frac{\ddot{H}}{H} = -(1 - e\cos E)^{-3} + \frac{T}{H^2}\frac{1 + \beta_r}{1 + \beta_r^\circ}\frac{\left[1 + \frac{1}{2}\frac{M}{p} - \frac{M^{zz}}{p}\right]}{\left[1 + \frac{1}{2}\frac{M^\circ}{p^\circ} - \frac{(M^{zz})^\circ}{p^\circ}\right]}, \tag{12.16}$$

$$\begin{aligned}\dot{T} = &-(\Gamma_3 - 1)T\left(\frac{\dot{J}}{J} + \frac{\dot{H}}{H}\right) \\ &+ (\Gamma_3 - 1)\frac{1+\beta_r}{1+4\beta_r}T\left(f_{\mathcal{H}}\frac{P_v}{P} - \mathcal{C}^\circ\frac{1+\beta_r^\circ}{1+\beta_r}J^2T^3\right),\end{aligned} \tag{12.17}$$

where a dot over a letter indicates a derivative with respect to rescaled time. As in the hydrodynamical model $\beta_r = \beta_r^\circ JHT^3$. The equilibrium values of the reference circular disc H°, $\hat{T}^\circ$, etc., are determined by β_r°, β_m° and n. The reference cooling rate can be obtained by setting it equal to the reference heating rate $\mathcal{C}^\circ = f_{\mathcal{H}}^\circ \frac{P_v^\circ}{P^\circ}$, which will depend on the closure model for the magnetic field/turbulence. Finally the (nondimensional) entropy is unchanged from the hydrodynamical model.

12.3 Effect of magnetic fields

12.3.1 Magnetic fields in eccentric discs

In Chapter 11 we found that (when $p_v = p_g$) our radiation dominated solutions exhibit extreme compression at pericentre, the same as the extreme behaviour of the adiabatic solutions of Ogilvie & Barker (2014). Many of our solutions with more moderate behaviour have strong viscous stresses at pericentre which call into question the validity of the $\alpha-$prescription.

What additional physical processes could reverse the collapse of the fluid column and prevent the extreme compression seen in the radiation dominated model? Can the collapse be reversed without encountering unphysically strong viscous stresses. An obvious possibility is the presence of a large scale horizontal magnetic field within the disc which will resist vertical compression. Such a field could be weak for the majority of the orbit but, owing to the relative incompressibility of magnetic fields (see discussion in Chapter 9), become dynamically important during the maximum compression at pericentre. In Chapter 9 we showed that in an eccentric disc, a solution to the steady ideal induction equation in an inertial frame is,

$$B^a = 0, \quad B^E = \frac{\Omega B_0^E(a,\tilde{z})}{nJH}, \quad B^z = \frac{B_0^z(a)}{J} \quad . \tag{12.18}$$

B^E is parallel to the orbital motion (quasi-toroidal) and B^z is the vertical magnetic field. We use the term quasi-poloidal to indicate the fields B^a and B^z.

The magnetic field of a star undergoing tidal disruption has been studied by Guillochon & McCourt (2017) and Bonnerot et al. (2017). In these papers it was found that the stretching of the fields during the disruption causes an increase in the magnetic pressure from the field aligned with the orbital direction. Meanwhile the gas pressure and magnetic pressure from the field perpendicular to the orbit drops. Guillochon & McCourt (2017) found that this tends to result in the magnetic pressure from the parallel field becoming comparable to the gas pressure. Similar results were found in Bonnerot et al. (2017), although with a dependence on the initial field direction. This supports our adopted field configuration, with the vertical field set to zero. As the vertical field does not contribute to the dynamics of the vertical oscillator we can do so without loss of generality.

As discussed in Section 12.2, in order to preserve separability of the equations we require $B_0^E(a, \tilde{z})$ to depend on $\tilde{z}$ in such a way as to make β_m independent of height. The dimensionless equations for the variation of the dimensionless scale height H and temperature T around the orbit (derived in Section 9.2.2 of Chapter 9) are then:

$$\begin{aligned}\frac{\ddot{H}}{H} = &-(1-e\cos E)^{-3} + \frac{T}{H^2}\frac{1+\beta_r}{1+\beta_r^\circ}\left[1+\frac{1}{\beta_m^\circ}\right]^{-1} \\ &\times\left[1+\frac{1}{\beta_m} - 2\alpha_s\frac{P_v}{P}\frac{\dot{H}}{H} - \left(\alpha_b - \frac{2}{3}\alpha_s\right)\frac{P_v}{P}\left(\frac{\dot{J}}{J}+\frac{\dot{H}}{H}\right)\right],\end{aligned} \tag{12.19}$$

$$\begin{aligned}\dot{T} = &-(\Gamma_3-1)T\left(\frac{\dot{J}}{J}+\frac{\dot{H}}{H}\right) \\ &+(\Gamma_3-1)\frac{1+\beta_r}{1+4\beta_r}T\left(f_{\mathcal{H}}\frac{P_v}{P} - \frac{9}{4}\alpha_s\frac{P_v^\circ}{P^\circ}\frac{1+\beta_r^\circ}{1+\beta_r}J^2T^3\right),\end{aligned} \tag{12.20}$$

with the plasma-β given by,

$$\beta_m = \beta_m^\circ JHT\frac{1+\beta_r}{1+\beta_r^\circ}\frac{1-e\cos E}{1+e\cos E} \tag{12.21}$$

where β_m° is the plasma beta in the reference circular disc.

These can be solved using the same relaxation method used to solve the purely hydrodynamic equations. However caution must be taken when solving the equations with low β_m° (i.e. strong magnetic fields throughout the disc) as the solutions do not always converge to a periodic solution (or at least takes an excessively long time to do so). This is most likely due to the absence of dissipative effects acting on the magnetic

fields, so any free oscillations in the magnetic field are not damped out. We believe that the quasiperiodic solutions we find for low β_m° are the superposition of the forced solution and a free fast magnetosonic mode. For now we only consider values of β_m° which successfully converge to a periodic solution.

12.3.2 Viscous stress independent of the magnetic field ($p_v = p_g$)

Figure 12.1 shows the variations of the scale height, β_r and plasma beta around the orbit for a disc with $\alpha_s = 0.1$, $\alpha_b = 0$, $e = q = 0.9$ and $E_0 = 0$. The magnetic field has a weak effect on the gas pressure dominated ($\beta_r^\circ = 10^{-4}$) solutions. For the radiation pressure dominated disc ($\beta_r^\circ = 10^{-3}$) a strong enough magnetic field stabilises the solution against the thermal instability and, instead of the nearly adiabatic radiation dominated solutions seen in the hydrodynamic case the solution is moderately radiation pressure dominated and maintains significant entropy variation around the orbit. This solution is similar to the moderately radiation pressure dominated hydrodynamic solutions. If the field is too weak (e.g. $\beta_m^\circ = 100$ considered here) the magnetic field isn't capable of stabilising against the thermal instability and the solution tends to the nearly adiabatic radiation dominated solution.

We see it is possible to attain significantly radiation pressure dominated solutions which do not possess the extreme variation of the scale height around the orbit present in the hydrodynamic case. In particular the green curve with $\beta_r^\circ = 1$, $\beta_m^\circ = 0.005$. Like the radiation dominated hydrodynamic solutions the solution with $\beta_r^\circ = 1$, $\beta_m^\circ = 0.005$ is nearly adiabatic, however magnetic pressure dominates over radiation pressure during pericentre passage. This additional support at pericentre prevents the extreme compression, and resultant heating, seen in the hydrodynamic model - resulting in more moderate variation of the scale height around the orbit.

It should be cautioned that the grey solution (with $\beta_r^\circ = 1$, $\beta_m^\circ = 1$) in the right hand plot of Figure 12.1 has not converged. The magnetic field is unimportant for this solution. Based on the radiation dominated hydrodynamic models of Chapter 11, the disc with $\beta_r^\circ = 1$ will converge on a solution with β_r much larger than the $\beta_r \sim 10^5 - 10^6$, which were the most radiation dominated, converged, solutions obtained in Chapter 11. As the entropy gained per orbit is small compared to the entropy in the disc, this will take a large number of orbits (> 10000 orbits) to converge, so the converged solution isn't of much interest when considering transient phenomena like a TDE.

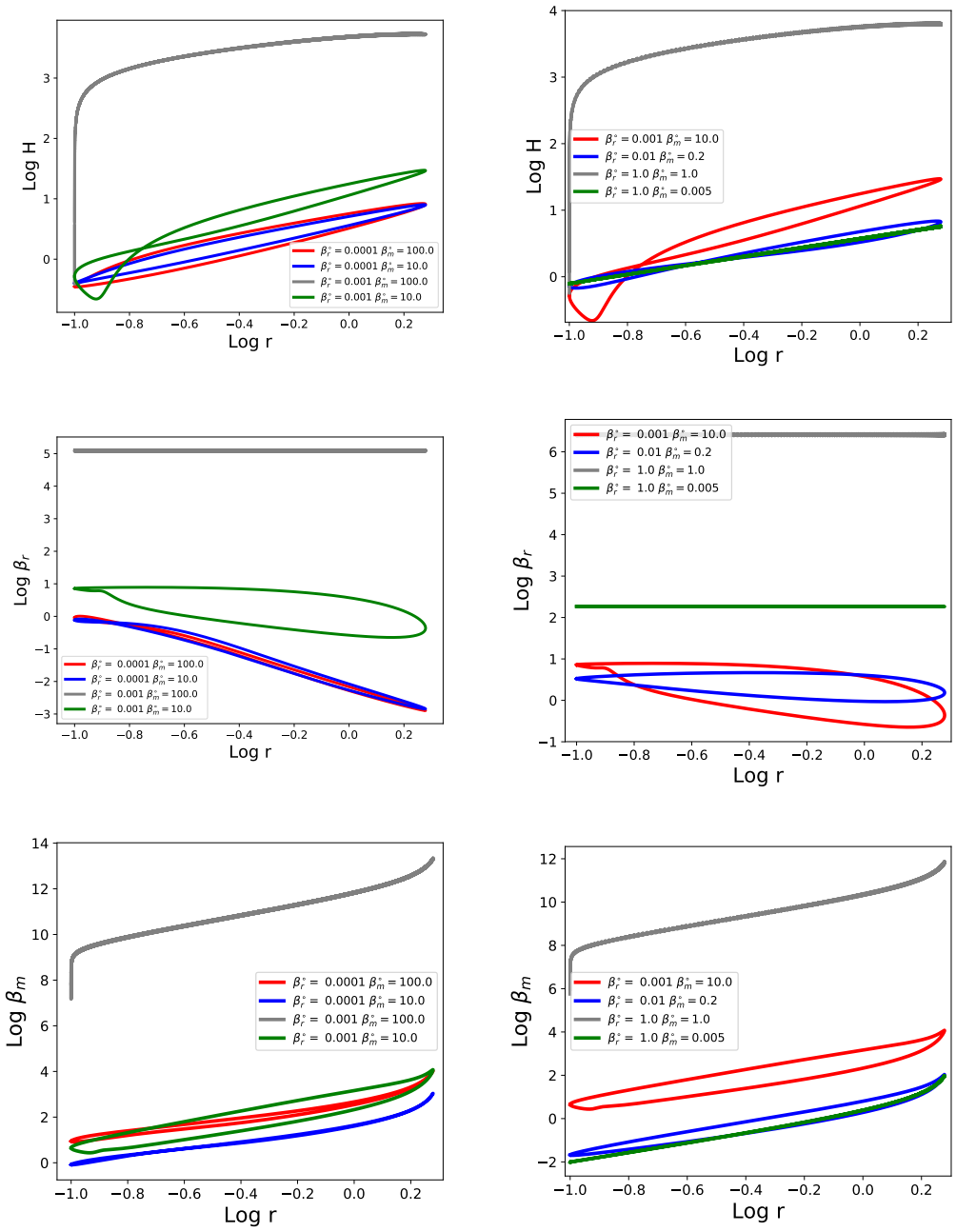


Fig. 12.1 Variation of the scale height, β_r and β_m of the disc with radiation and gas pressure with different β_r° and magnetic fields. Disc parameters are $p_v = p_g$, $\alpha_s = 0.1$, $\alpha_b = 0$ $e = q = 0.9$ and $E_0 = 0$. The right hand column considers discs with stronger magnetic fields. The discs with $\beta_m^\circ = 100$ are nearly indistinguishable from an unmagnetised disc.

Figure 12.2 shows the magnitude of different terms in the vertical momentum equation for a disc with $p_v = p_g$, $\beta_r^\circ = 1$, $\beta_m^\circ = 0.005$, $\alpha_s = 0.1$, $\alpha_b = 0$, $e = q = 0.9$ and $E_0 = 0$ (i.e. the green solution from the right hand plots of Figure 12.1). This shows that the dominant balance in this solution is between the vertical acceleration, gravity and the magnetic field. This suggests that the dynamics of radiation pressure dominated TDEs may be entirely dominated by the magnetic field. Being the least compressible pressure term (see Chapter 9), the magnetic field tends to dominate at pericentre, even if it is fairly weak throughout the rest of the disc. However for the nearly adiabatic, radiation dominated, TDEs pressure is only important near pericentre so even a weak magnetic field will have a disproportionate contribution to the dynamics. This suggests ignoring even subdominant magnetic fields in TDE discs can lead to fundamental changes to the TDE dynamics.

While it is possible to find a combination of β_r° and β_m° which yields a solution with the desired β_r exhibiting "reasonable" behaviour, it is not clear that the magnetic field in the disc will always be strong enough to produce the desired behaviour. It is possible that this represents a tuning problem for β_m°.

To explore this we look at what happens if the magnetic field is initially too weak to stabilise against the thermal instability but gradually raise it, over several thermal times. Figure 12.3 shows what happens when the magnetic field is increased gradually from $\beta_m^\circ = 100$ to $\beta_m^\circ = 10$ for a disc with $p_v = p_g$, $\beta_r^\circ = 10^{-3}$, $\alpha_s = 0.1$, $\alpha_b = 0$ $e = q = 0.9$ and $E_0 = 0$. This corresponds to moving from the grey to the green solution in the right hand plots of Figure 12.1. This is done by periodically stopping the calculation and restarting with a larger β_m°. The resulting β_r in fact increases with time and remains close to that of the grey solution in Figure 12.1 even as we increase the magnetic field strength, and does not transition to a value consistent with the green solution. This suggests that the solution is sensitive to the path taken; a magnetic field which grows (from an initially weak seed field), in a nearly adiabatic radiation pressure dominated disc, may not cause the disc to collapse to the gas pressure dominated branch. This is likely because the disc is very expanded meaning the magnetic field is still quite weak and incapable of influencing the dynamics.

We carried out a similar calculation for a disc with $p_v = p_g$, $\beta_r^\circ = 1$, $\alpha_s = 0.1$, $\alpha_b = 0$, $e = q = 0.9$ and $E_0 = 0$ moving from $\beta_m^\circ = 1$ to $\beta_m^\circ = 5 \times 10^{-3}$ (corresponding to the grey and green solutions of the right hand plots of 12.1). In this case β_r steadily increases with time (apart from a small variation over the orbital period) with the magnetic field having no appreciable influence on the solution. Owing to the relatively large β_r° this solution never reached steady state, as discussed previously.

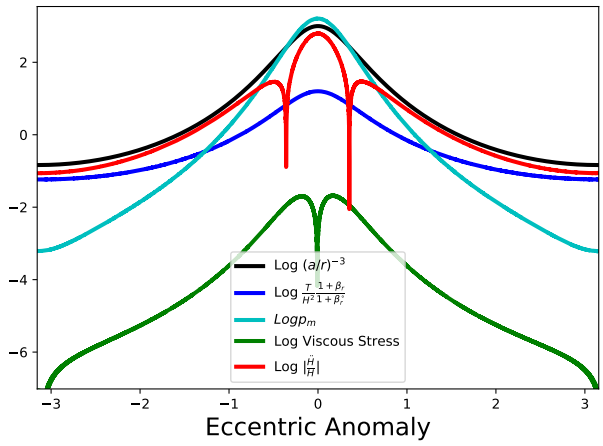


Fig. 12.2 Magnitude of the different terms in the vertical momentum equation for a magnetised radiation-gas mixture; the black line is the gravity, the blue is the hydrodynamic pressure, the cyan is the magnetic pressure, green is the viscous stress and red is the vertical acceleration. Disc parameters are $p_v = p_g$, $\beta_r^\circ = 1$, $\beta_m^\circ = 0.005$, $\alpha_s = 0.1$, $\alpha_b = 0$, $e = q = 0.9$ and $E_0 = 0$. The balance at pericentre is now between the gravity and the magnetic pressure.

Figure 12.4 shows the pericentre passage for a magnetised disc with $p_v = p_g$, $\beta_r^\circ = 10^{-3}$, $\beta_m^\circ = 10$, $\alpha_s = 0.1$, $\alpha_b = 0$, $e = q = 0.9$ and $E_0 = 0$. The magnetic pressure is extremely concentrated within the nozzle and near to the midplane. Like the hydrodynamic nozzle like structure found in Chapter 11, the nozzle is asymmetric, located prior to pericentre and is associated with an increase in entropy which is appears to be characteristic of dissipative highly eccentric discs.

12.3.3 Viscous stress dependent on the magnetic field

Begelman & Pringle (2007) have suggested that discs with strong toroidal fields may be stable to the thermal instability if the stress depends on the magnetic pressure. In this subsection we explore this possibility for a highly eccentric disc.

Figure 12.5 shows the variation of the scale height, β_r and plasma beta for a disc with $p_v = p_m$, $\alpha_s = 0.1$, $\alpha_b = 0$, $e = q = 0.9$ and $E_0 = 0$. When the magnetic field is weak, these have essentially the same behaviour as the nearly adiabatic radiation pressure dominated discs for the hydrodynamic case. This is not surprising as in this limit the gas and magnetic pressures are essentially negligible, which also results in negligible viscous stress/heating when it scales with either of these pressures. Increasing the magnetic field strength stabilises the "gas pressure dominated" branch (see the right

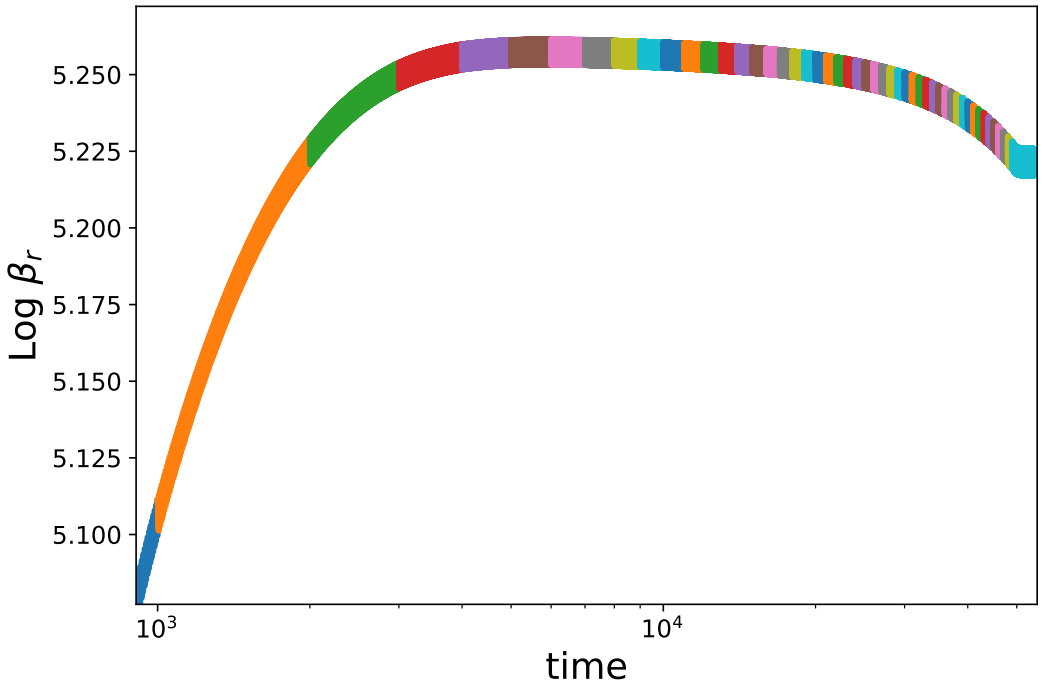


Fig. 12.3 β_r for a disc where the magnetic field strength is gradually increased from $\beta_m^\circ = 100$ to $\beta_m^\circ = 10$. Disc parameters are $p_v = p_g$, $\beta_r^\circ = 10^{-3}$, $\alpha_s = 0.1$, $\alpha_b = 0$, $e = q = 0.9$ and $E_0 = 0$. Colours indicate where we have stopped the calculation and restarted with a different magnetic field strength. The resulting solution remains in the nearly adiabatic radiation pressure dominated state and doesn't converge on the green solution in the right hand plots in Figure 12.1. The solution with the final magnetic field strength ($\beta_m^\circ = 10$) was run for longer to allow it to relax to a steady state.

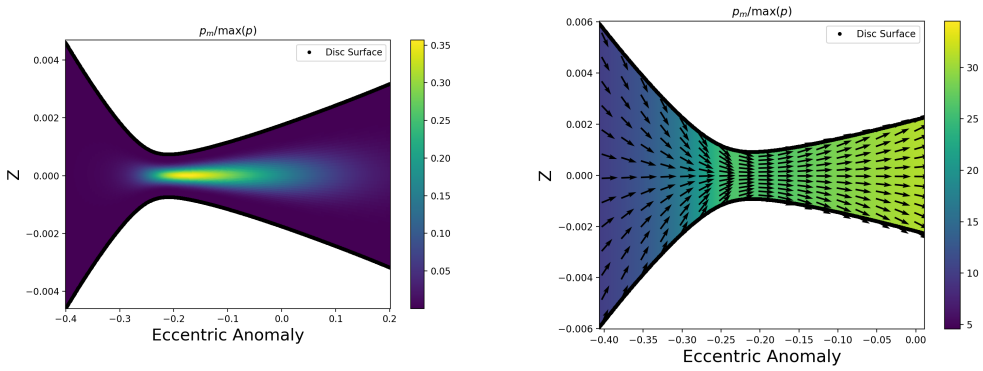


Fig. 12.4 Pericentre passage for a magnetised disc with $p_v = p_g$, $\beta_r^\circ = 10^{-3}$, $\beta_m^\circ = 10$, $\alpha_s = 0.1$, $\alpha_b = 0$, $e = q = 0.9$ and $E_0 = 0$. Left shows the magnetic pressure, right the velocity field with colour the entropy. The magnetic field is highly concentrated in the nozzle. Like the hydrodynamic discs the there is an increase in entropy combined with a reversal in the direction of the flow in the nozzle. The nozzle occurs prior to pericentre and is asymmetric which is indicative of dissipation.

hand column of Figure 12.5), where the magnetic field and viscous dissipation become important. This branch can have $p_r \gg p_g$ around the entire orbit, this is similar to the behaviour of the radiation pressure dominated hydrodynamic discs considered in Chapter 11 with large α_s.

Figure 12.6 shows the variation of the scale height, β_r and plasma beta for a disc with $p_v = p + p_m$, $\alpha_s = 0.1$, $\alpha_b = 0$, $e = q = 0.9$ and $E_0 = 0$. Here we find that, with a strong enough magnetic field, we can obtain thermally stable solutions despite the dependence of the stress on the radiation pressure. Generally for thermal stability the magnetic field needs to dominate (over radiation pressure) over part of the orbit. If the magnetic field is too weak, however, we encounter the thermal instability similar to the hydrodynamic radiation pressure dominated discs when $p_v = p$.

Part of the motivation for introducing the magnetic field was to regularise some of the extreme behaviour encountered at pericentre. Unfortunately, while $p_v = p_m$ and $p_v = p + p_m$ are promising as a way of taming the thermal instability they exhibit the same extreme behaviour that the hydrodynamic models possess. In particular when $p_v = p_m$ the solutions exhibit extreme compression at pericentre, while for the more magnetised discs (with either $p_v = p_m$ or $p_v = p + p_m$) we again encounter the issue of the viscous stresses being comparable to or exceeding the pressure (including the magnetic pressure). See, for example Figure 12.7 which shows the viscous stresses exceed the magnetic, gas and radiation pressures during pericentre passage.

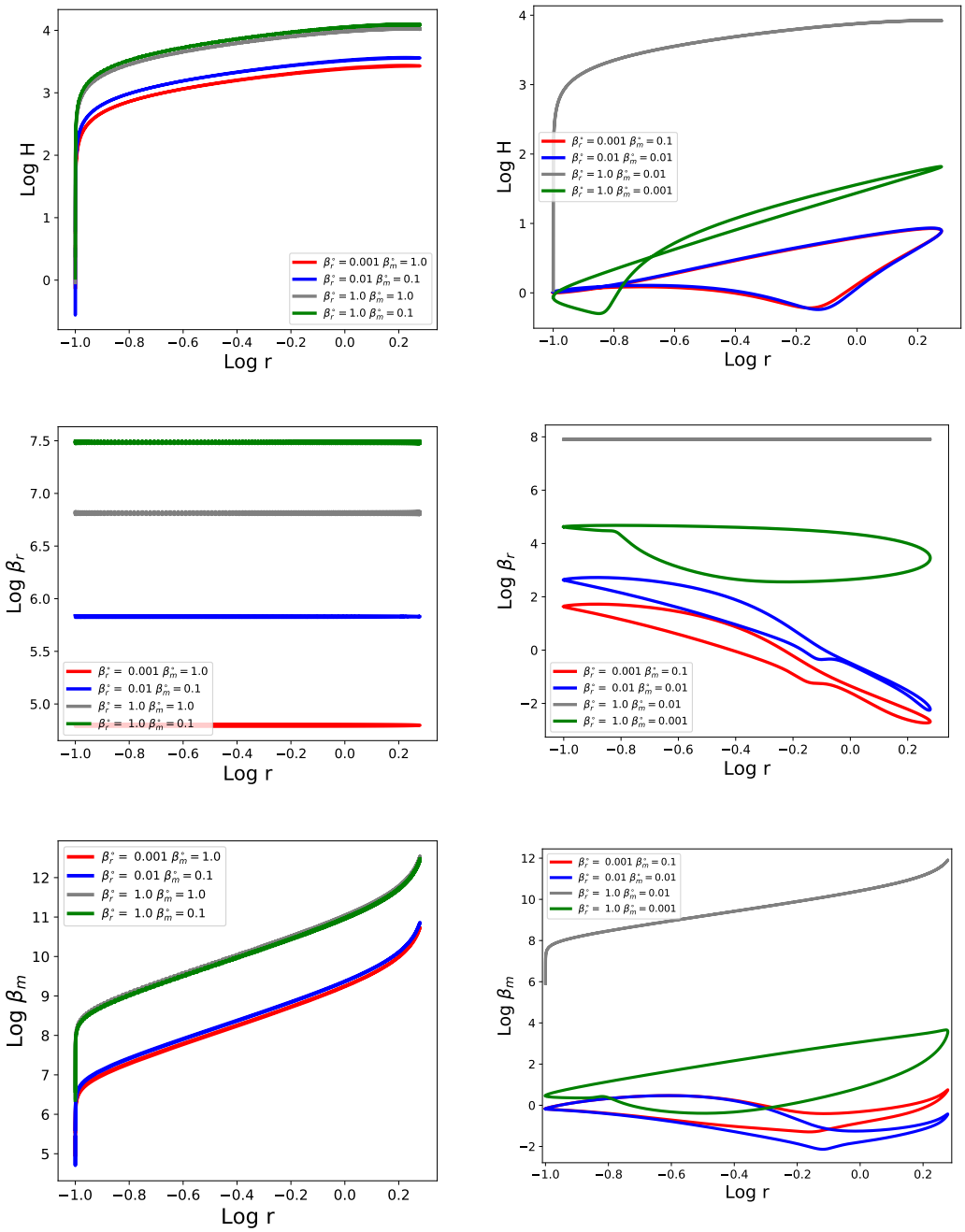


Fig. 12.5 Variation of the scale height, β_r and β_m of the disc when the viscous stress is proportional to the magnetic pressure. Disc parameters are $p_v = p_m$, $\alpha_s = 0.1$, $\alpha_b = 0$, $e = q = 0.9$ and $E_0 = 0$. The right hand column considers discs with stronger magnetic fields. This confirms that a strong magnetic field can stabilise the thermal instability in an eccentric disc if p_v depends on the magnetic pressure.

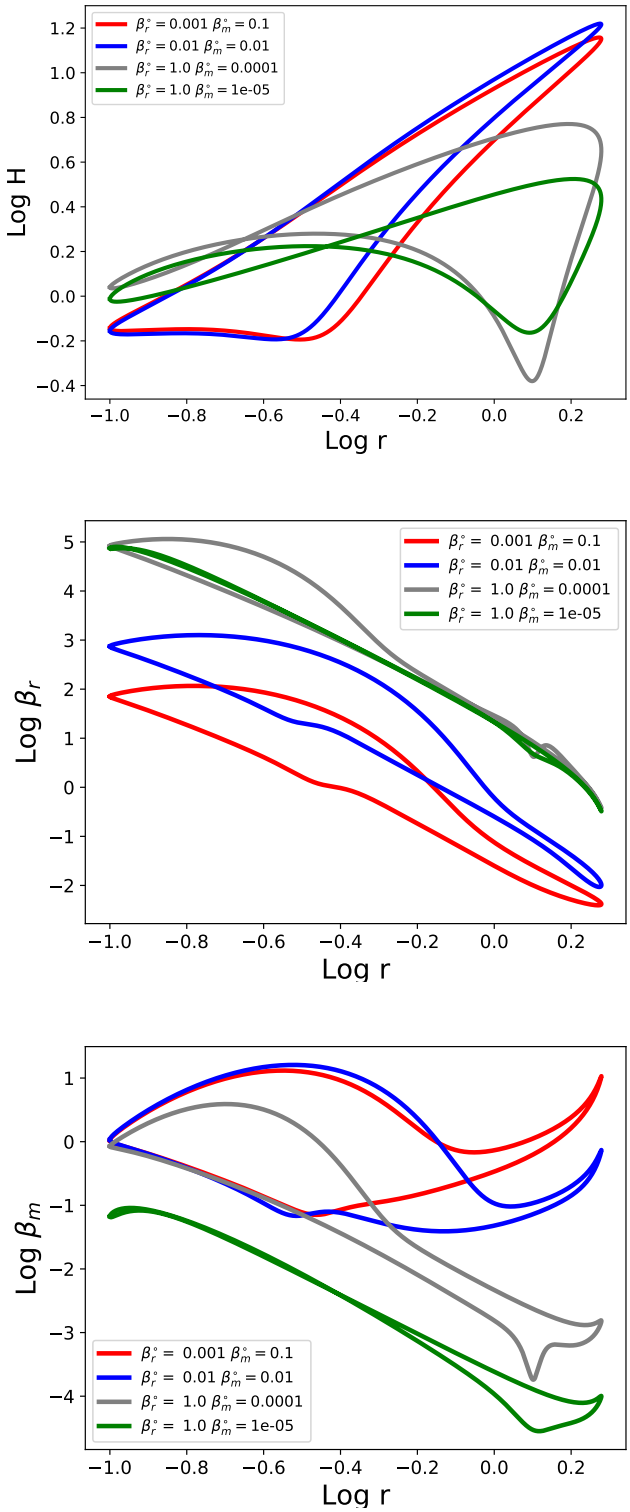


Fig. 12.6 Variation of the scale height, β_r and β_m of the disc when the viscous stress is proportional to the total gas+radiation+magnetic pressure. Disc parameters are $p_v = p + p_m$, $\alpha_s = 0.1$, $\alpha_b = 0$, $e = q = 0.9$ and $E_0 = 0$. This confirms that a strong magnetic field can stabilise the thermal instability in an eccentric disc if p_v depends on the magnetic pressure.

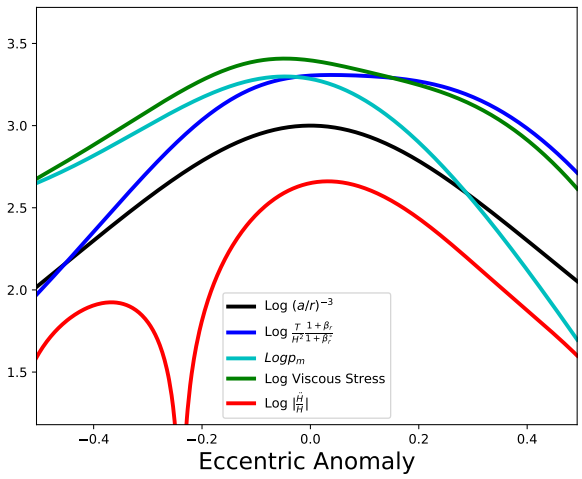


Fig. 12.7 Magnitude of terms in the vertical momentum equation for a disc with $p_v = p + p_m$, $\beta_r^\circ = 10^{-3}$, $\beta_m^\circ = 0.1$ $\alpha_s = 0.1$, $\alpha_b = 0$, $e = q = 0.9$ and $E_0 = 0$. The viscous stress exceeds the magnetic, gas and radiation pressures in the nozzle, during pericentre passage.

12.4 Nonlinear constitutive model for the magnetic stress

The model considered in Section 12.3 has a number of drawbacks. The first is that the $\alpha-$viscosity and the coherent magnetic field are treated as separate physical effects when they are in fact intrinsically linked (although subsection 12.3.3 partially addresses this issue). Secondly, the turbulent magnetic field, responsible for the α-viscosity, cannot store energy. Lastly the model neglects resistive effects and, while nonideal MHD effects would be weak if the flow were strictly laminar, the turbulent cascade should always move magnetic energy to scales on which nonideal effects become important. Thus the coherent magnetic field should be affected by some dissipative process.

To address these issues we consider the model of the (modified) Maxwell stress where the magnetic field is forced by a turbulent emf and relaxes to an isotropic field proportional to some pressure p_v on a timescale τ, derived in Chapter 9.

The Maxwell stress in this model evolves according to

$$\mathcal{D}M^{ij} = -(M^{ij} - \mathcal{B}p_v g^{ij})/\tau \quad , \tag{12.22}$$

where $\mathcal{B}$ is a nondimensional parameter controlling the strength of forcing relative to p_v, which can be taken to be constant by absorbing any variation into the definition of p_v. $\mathcal{D}$ is the operator from Chapter 4 which acts on a rank (2,0) tensor by

$$\mathcal{D}M^{ij} = DM^{ij} - 2M^{k(i}\nabla_k u^{j)} + 2M^{ij}\nabla_k u^k \quad . \tag{12.23}$$

As noted in Ogilvie (2001), $\mathcal{D}M^{ij} = 0$ is the equation for the evolution of the (modified) Maxwell stress for a magnetic field which satisfies the ideal induction equation.

The Deborah number De is given by

$$\mathrm{De} = \tau n = \mathrm{De}_0 \frac{n}{\Omega_z}\sqrt{\frac{p_v}{M}} \quad , \tag{12.24}$$

where Ω_z is the vertical oscillation frequency; this matches the functional form for the relaxation time τ given in the compressible version of Ogilvie (2003). When the Deborah number is given by Equation 12.24 then Equation 12.22 corresponds to the equation for the (modified)-Maxwell stress given in Ogilvie (2003) if the Reynolds stress is isotropic and proportional to some pressure p_v. One emergent property of such a stress model is the stress will naturally scale with magnetic pressure, as the latter is the trace of the former (see discussion in Chapter 9).

From this stress model, we have a nondimensional heating rate,

$$f_{\mathcal{H}} = \frac{1}{2\mathrm{De}}\left(\frac{M}{p_v} - 3\mathcal{B}\right) \quad ; \tag{12.25}$$

this ensures that magnetic energy loss/gained via the relaxation terms in Equation 12.22 is converted to/from the thermal energy (this is shown in Section 9.3.6 of Chapter 9). Thus energy is conserved within the disc, although it can be lost radiatively from the disc surface.

In Chapter 9 we obtain the hydrostatic solutions for a circular disc. If p_v is independent of M the vertical equation of motion, rescaled by this reference circular disc, is

$$\frac{\ddot{H}}{H} = -(1 - e\cos E)^{-3} + \frac{T}{H^2}\frac{1+\beta_r}{1+\beta_r^\circ}\frac{\left[1 + \frac{1}{2}\frac{M}{p} - \frac{M^{zz}}{p}\right]}{\left[1 + \mathcal{B}\frac{P_v^\circ}{P^\circ}\left(\frac{1}{2} + \frac{9}{4}\mathrm{De}_0^2\frac{P_v^\circ}{P^\circ}\right)\right]}, \tag{12.26}$$

while the thermal energy equation is

$$\begin{aligned}\dot{T} = &-(\Gamma_3 - 1)T\left(\frac{\dot{J}}{J} + \frac{\dot{H}}{H}\right) \\ &+ (\Gamma_3 - 1)\frac{1+\beta_r}{1+4\beta_r}T\left(\frac{1}{2\mathrm{De}}\left(\frac{M}{p} - 3\mathcal{B}\frac{p_v}{p}\right) - \mathcal{C}^\circ\frac{1+\beta_r^\circ}{1+\beta_r}J^2T^3\right),\end{aligned} \tag{12.27}$$

where we have introduced a reference cooling rate,

$$\begin{aligned}\mathcal{C}^\circ &= \frac{9}{4}\mathcal{B}\mathrm{De}^\circ\frac{P_v^\circ}{P^\circ} \\ &= \left(\frac{3}{2}\right)^{3/2}\mathcal{B}^{1/2}\mathrm{De}_0\left(1+\sqrt{1+2\frac{\mathrm{De}_0^2}{\mathcal{B}}}\right)^{-1/2}\frac{P_v^\circ}{P^\circ} \quad .\end{aligned} \tag{12.28}$$

We solve these equations along with the equations for the evolution of the stress components,

$$\dot{M}^{\lambda\lambda} + 2\left(\frac{\dot{J}}{J} + \frac{\dot{H}}{H}\right)M^{\lambda\lambda} = -(M^{\lambda\lambda} - \mathcal{B}p_v g^{\lambda\lambda})/\tau, \tag{12.29}$$

$$\begin{aligned}&\dot{M}^{\lambda\phi} - M^{\lambda\lambda}\Omega_\lambda - M^{\lambda\phi}\Omega_\phi + 2\left(\frac{\dot{J}}{J} + \frac{\dot{H}}{H}\right)M^{\lambda\phi} \\ &\qquad = -(M^{\lambda\phi} - \mathcal{B}p_v g^{\lambda\phi})/\tau,\end{aligned} \tag{12.30}$$

$$\begin{aligned}&\dot{M}^{\phi\phi} - 2M^{\lambda\phi}\Omega_\lambda - 2M^{\phi\phi}\Omega_\phi + 2\left(\frac{\dot{J}}{J} + \frac{\dot{H}}{H}\right)M^{\phi\phi} \\ &\qquad = -(M^{\phi\phi} - \mathcal{B}p_v g^{\phi\phi})/\tau,\end{aligned} \tag{12.31}$$

$$\dot{M}^{zz} + 2\frac{\dot{J}}{J}M^{zz} = -(M^{zz} - \mathcal{B}p_v g^{zz})/\tau \quad . \tag{12.32}$$

We solve for these stress components in the (λ, ϕ) coordinate system of Ogilvie (2001) as this simplifies the metric tensor and covariant derivatives. We can do this as, apart from M^{zz} (which is the same in both coordinate systems), our equations only depend on M^{ij} through scalar quantities.

Figure 12.8 show the variations of the scale height, β_r and plasma beta (defined as $\beta_m^\circ = \frac{2p}{M}$) around the orbit for a disc with $p_v = p_g$, $\mathrm{De} = 0.5$, $\mathcal{B} = 0.1$, $e = q = 0.9$ and $E_0 = 0$. Like the ideal induction equation model of Section 12.2 the coherent magnetic field has a stabilising effect on the dynamics. The effect is not as strong as that seen in the ideal induction equation model as, in that model, we could choose β_m° so as to

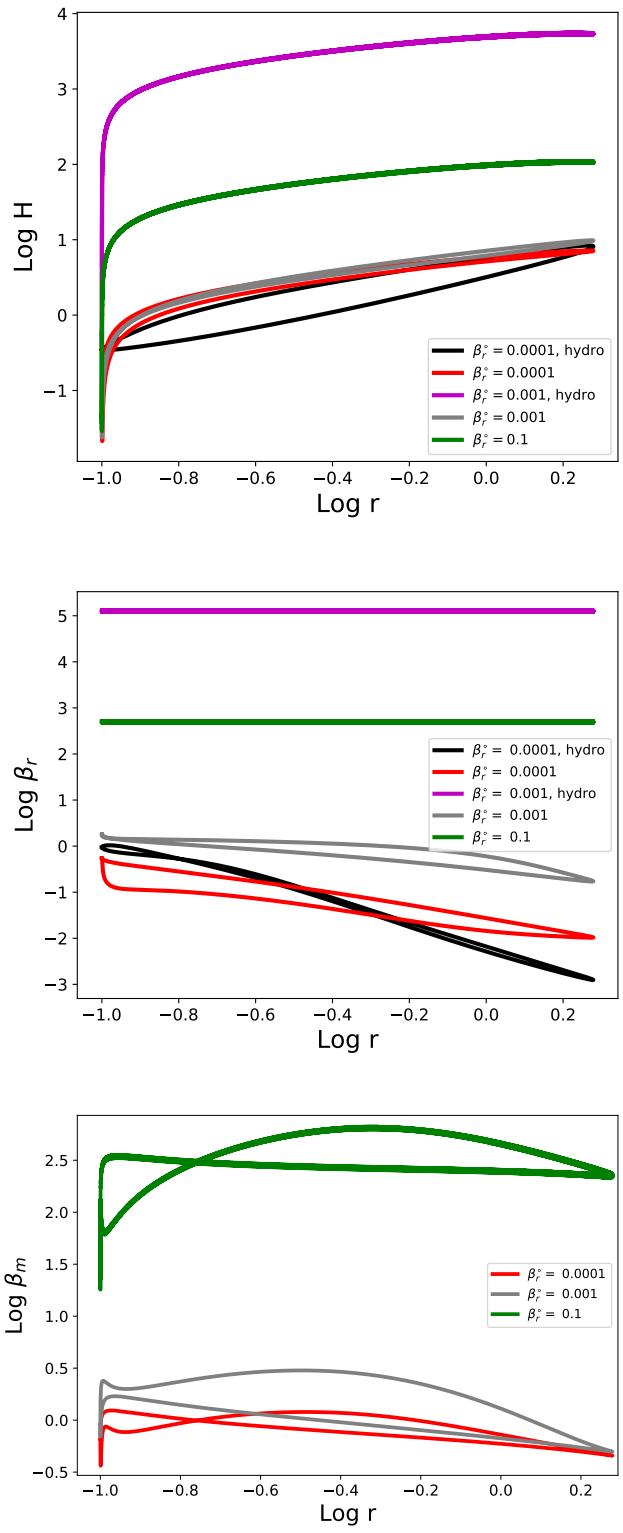


Fig. 12.8 Variation of the scale height, β_r and β_m of the disc with radiation + gas pressure with different β_r° using our modified Maxwell stress prescription. Disc parameters are $e = q = 0.9$ and $E_0 = 0$ and $p_v = p_g$, α-discs have $\alpha_s = 0.1$, $\alpha_b = 0$ while the Maxwell stress prescription has De $= 0.5$, $\mathcal{B} = 0.1$. Black line is an α-disc with $\beta_r^\circ = 10^{-4}$; red line has $\beta_r = 10^{-4}$ with the Maxwell stress prescription, magenta line is an α-disc with $\beta_t^\circ = 10^{-3}$, grey line has $\beta_r = 10^{-3}$ with the Maxwell stress prescription, green line has $\beta_r = 10^{-3}$ with the Maxwell stress prescription.

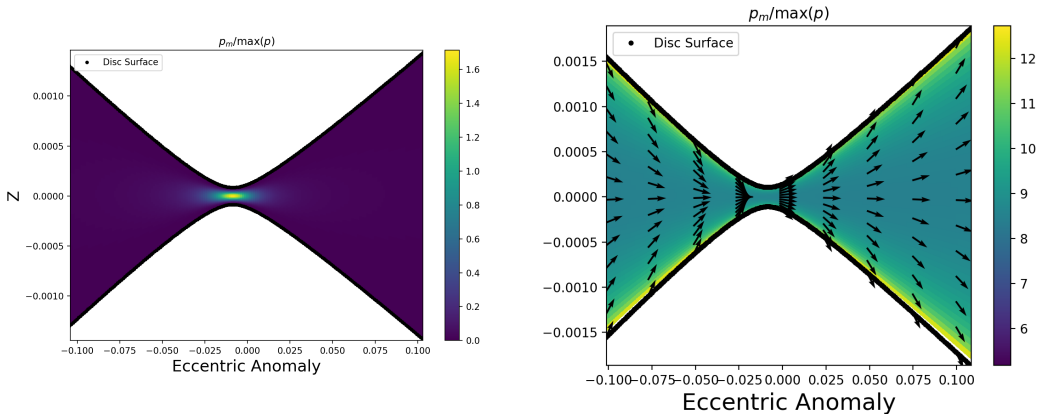


Fig. 12.9 Pericentre passage for a disc with $p_v = p_g$, $\beta_r^\circ = 10^{-3}$, $\mathrm{De}_0 = 0.5$, $\mathcal{B} = 0.1$, $e = q = 0.9$ and $E_0 = 0$. Left shows the magnetic pressure ($M/2$), right the velocity field with colour the entropy. As with the ideal induction equation model the magnetic field is highly concentrated in the nozzle. The nozzle is very symmetric, although it is slightly offset from pericentre, indicating weak dissipation. Like the hydrodynamic discs the there is an increase in entropy combined with a reversal in the direction of the flow in the nozzle.

achieve a much stronger field than achieved by the constitutive model here. Compared with the ideal induction equation model the plasma-β is more uniform around the orbit; there is still an abrupt decrease in the plasma-β near pericentre, which highlights the importance of the magnetic field during pericentre passage.

Figure 12.9 shows the pericentre passage for a disc with $p_v = p_g$, $\beta_r^\circ = 10^{-3}$, $\mathrm{De}_0 = 0.5$, $\mathcal{B} = 0.1$, $e = q = 0.9$ and $E_0 = 0$. As with the ideal induction equation model, the magnetic pressure is extremely concentrated within the nozzle and near to the midplane. The nozzle flow is far more symmetric compared to the ideal induction equation model as the weaker field means that the disc is in a modified form of the nearly adiabatic radiation pressure dominated state. Compared with the hydrodynamic models considered in Chapter 11 the entropy generation in the nozzle is much less abrupt, extending smoothly either side of the nozzle.

In addition to considering the situation where the fluctuation pressure scales with the gas pressure $p_v = p_g$, we also considered $p_v = p + p_m$. As in the ideal induction equation model we found it is possible to stabilise the thermal instability with a strong enough magnetic field, however we found that this requires fine tuning of De_0 and $\mathcal{B}$, and there is no obvious justification for the fine tuning of De_0 and $\mathcal{B}$ to give thermally stable solutions. However, instead of stabilising the thermal instability, it is possible to delay it's onset by choosing a small enough De_0, so that the thermal runaway occurs on

a timescale much longer than the orbital time (occurring after ~ 1000 orbits). However, these solutions never settle down into a periodic (or nearly periodic) solution and instead have a long phase of quasi-periodic evolution, where the mean scale height remains close to its initial value, before eventually experiencing thermal runaway. A quasi-periodic solution of our model is not self consistent, so the possibility that the thermal instability is delayed in the nonlinear-constitutive MRI model needs to be explored using an alternative method.

The possibility that the thermal instability stalls or is delayed has some support from simulations looking at the thermal stability of MRI active discs (Jiang et al., 2013; Ross et al., 2017). In both these papers it was found that the disc was quasi-stable, with thermal instability occurring when a particularly large turbulent fluctuation caused a strong enough perturbation away from the equilibrium. It is possible that our stress model captures the thermal quasi-stability seen in Jiang et al. (2013) and Ross et al. (2017) in some averaged sense. The possibility that the thermal instability is delayed or slowed is particularly relevant for TDEs which are inherently transient phenomena - if the timescale for thermal runaway is made long enough then eccentric TDE discs maybe thermally stable over the lifetime of the disc.

12.5 Discussion

12.5.1 Stability of the solutions

As discussed in Chapter 11, one advantage of our solution method is that the solutions it finds are typically nonlinear attractors (or at least long lived transients) so are stable against (nonlinear) perturbations to the solution variables (H,$\dot{H}$, T and M^{ij} when present). Generally we expect such perturbations to damp on the thermal time or faster. Instabilities such as the thermal instability manifest as a failure to converge to a $2\pi-$periodic solution.

For the ideal induction equation model, our method cannot tell us about the stability of the solution to perturbations to the horizontal magnetic field. Showing this would require a separate linear stability analysis. Perturbations to the vertical field typically have no influence on the dynamics of the disc vertical structure.

However, for the constitutive model, perturbations to the magnetic field are encapsulated in perturbations to M^{ij} so these solutions are stable against (large scale) perturbations to the magnetic field. This is most likely because, unlike the ideal induction equation, dissipation acts on the magnetic field.

Our solution method doesn't tell us about the stability of our solutions to short wavelength (comparable or less than the scale height) perturbations to our system. So our disc structure could be unstable to such perturbations. Like the hydrodynamic solutions in Chapter 11, it is likely our discs are unstable to the parametric instability (Papaloizou, 2005a,b; Wienkers & Ogilvie, 2018; Barker & Ogilvie, 2014; Pierens et al., 2020). Additionally if, as assumed, highly eccentric discs are unstable to the MRI then there must be perturbations to the magnetic field in the ideal induction equation model which are unstable.

12.5.2 Resistivity and dynamo action

Even when the magnetic field in our models does a good job of resisting the collapse of the disc, the stream will still be highly compressed at pericentre. The highly compressed flow combined with a very strong field (with $\beta_m \ll 1$) makes the nozzle a prime site for magnetic reconnection. This will require that magnetic field lines on neighbouring orbits can have opposite polarities. Our solutions are agnostic to the magnetic field polarity and, in principle, support this possibility.

The simulations of Guillochon & McCourt (2017) suggest that the initial magnetic field in the disc will be (qausi-)toroidal with periodic reversals in direction. When such a field is compressed both horizontally and vertically during pericentre passage, neighbouring toroidal magnetic field lines of opposite polarity can undergo reconnection, generating a quasi-poloidal field. We thus have a basis for an eccentric $\alpha - \Omega$ dynamo, where strong reconnection in the nozzle creates quasi-poloidal field, which in turn creates a source term in the quasi-toroidal induction equation from the shearing of this quasi-radial field (see Chapter 9). In highly eccentric TDE discs, with their large vertical and horizontal compression, this dynamo could potentially be quite strong.

The constitutive model we considered in section 12.4 implicitly possesses a dynamo through the source term proportional to p_v. This is a small scale dynamo which arises from the action of the turbulent velocity field and will be limited by the kinetic energy in the turbulence. A dynamo closed by reconnection in the nozzle would be a large scale dynamo and the coherent field produced could potentially greatly exceed equipartition with the turbulent velocity field.

Lastly reconnection sites can accelerate charged particles. As such TDEs in which there is line of sight to the pericentre, should emit high energy particles. From most look angles the line of sight to the pericentre "bright point" will be blocked (see also Zanazzi & Ogilvie (2020)). This will both block the X-ray flux from the disc and any

particles accelerated by reconnection in the nozzle. Hence these high energy particles should only be observable in X-ray bright TDEs.

12.6 Conclusion

In this Chapter we considered alternative models of the Maxwell stress from the standard $\alpha-$prescription applied to highly eccentric TDE discs. In particular we focus on the effects of the coherent magnetic field on the dynamics. We consider two separate stress models: an $\alpha-$disc with an additional coherent magnetic field obeying the ideal induction equation and a nonlinear constitutive (viscoelastic) model of the magnetic field. In summary our results are:

1. The coherent magnetic field in both models has a stabilising effect on the dynamics, making the gas pressure dominated branch stable at larger radiation pressures and reducing or removing the extreme variation in scale height around the orbit for radiation pressure dominated solutions.

2. The coherent magnetic field is capable of reversing the collapse at pericentre without the presumably unphysically strong viscous stress seen in some of the hydrodynamic models.

3. For the radiation pressure dominated ideal induction equation model with a moderate magnetic field the dynamics of the scale height is set by the magnetic field (along with gravity and vertical motion) and doesn't feel the effects of gas or radiation pressure. This is because magnetic pressure dominates during pericentre passage which is the only part of the orbit where pressure is important.

Part V

Epilogue

Chapter 13

Summary and Suggestions for Future Work

"They gave themselves a papyrus-roll as their lector-priest, a writing-board as their dutiful son. Teachings are their mausoleam the reed-pen their child, the burnishing-stone their wife. Both great and small are given them as their children, for the writer is chief" **Hieratic text on Papyrus Chester Beatty IV, Trans: Toby Wilkinson.**

13.1 Summary

This thesis has presented an investigation of various aspects of non-circular accretion disc theory. Particular focus has been given to nonlinearly eccentric discs and their dynamical non-axisymmetric vertical structure. The overarching theoretical framework of this thesis has been the eccentric thin-disc model, where the dominant fluid motion consists of non-intersecting, nested, confocal Keplerian ellipses. The latter part of this thesis considered an application of eccentric disc theory to the highly eccentric discs expected from the tidal disruption of a star around a black hole.

In Chapter 5 I formulated a Hamiltonian theory for an ideal eccentric disc, where the eccentric orbits evolve due to pressure gradients and external perturbations. It was shown that this Hamiltonian theory can be derived by Whitham averaging (Whitham, 1965) the coplanar affine disc model of Ogilvie (2018). The resulting theory is substantively simpler than the more complete theory of Ogilvie (2001) and Ogilvie & Barker (2014) which was based on orbit averaged stress gradients. Nonlinear eccentric modes, untwisted eccentric discs which are stationary in a rotating frame, were presented which generalise the results of Barker & Ogilvie (2016).

Chapter 6 examined a nonlinear limit of the theory developed in Chapter 5, where the eccentricity is small, but with potentially nonlinear eccentricity gradients or twists. A similar limit was considered by Lee & Goodman (1999) for 2D discs. I derived the conditions under which an eccentric wave can be focused by the pressure profile and precessional forces in the disc, causing it to become nonlinear. This theory was applied to the case of an eccentric wave propagating in the inner part of a black hole disc as a nonlinear generalisation of the WKB theory of Ferreira & Ogilvie (2009). Recently there has been a possible confirmation of this effect in the numerical simulations of Dewberry et al. (2020b).

In Chapter 7 I considered a local model of an eccentric disc which includes nonideal effects, such as viscous heating and radiative cooling, and include a more general equation of state consisting of a mixture of a perfect gas and radiation. I derived the simplest horizontally invariant "laminar flow" solution to the flow in the eccentric disc. I found that the thermal instability in radiation pressure dominated eccentric discs can be different from that in circular discs as a form of the thermal instability still operates when the turbulent stress is assumed to scale with gas pressure, a situation found to be stable in circular discs.

In Chapter 8 I derived non-conservative terms in the limit of the highly twisted ("tightly-wound") waves that were considered in Chapter 6. This allows for the derivation of a more general, time dependent, theory of tightly wound waves that includes the effects of pressure gradients, self gravity, external perturbation and excitation or damping by viscosity. The non-conservative terms act as sources or sinks for the conserved fluxes derived in Chapter 6. I found that whether shear viscosity excites or damps tightly-wound waves depends on how nonlinear the wave is.

Chapter 9 looks at magnetic fields in eccentric discs within the framework of a local model. I derived the simplest "locally axisymmetric" solutions to the ideal induction equation, which include a solution where the magnetic field is aligned to the fluid flow in the disc. I derived a nonlinear constitutive model of the MRI which has a number of useful properties including viscoelastic behaviour, energy conservation, and realisability from a mean field.

Chapters 10-12 considered an application of eccentric disc theory to the highly eccentric discs expected when a star, on a nearly parabolic orbit, is disrupted by a supermassive black hole. The focus of these chapters was on the dynamical vertical structure set up by the eccentric orbits.

Chapter 11 applied the solutions derived in Chapter 7 to TDE discs. An important finding of this Chapter is that highly eccentric, radiation pressure dominated, discs are

thermally unstable for stress scaling with either total pressure or gas pressure. This differs from a radiation pressure dominated circular disc where this only occurs when stress scales with total pressure. I also found that care must be taken when modelling radiation dominated gasses as a common approximation using a perfect gas did not exhibit thermal instability.

Finally Chapter 12 generalised the results of Chapter 11 to MHD using the theory developed in Chapter 9. Magnetic fields can have a strong influence on the dynamics of the disc and are capable of stabilising the thermal instabilities found in Chapter 11. The relative incompressibilitiy of magnetic fields means they can dominate the dynamics even when they are quite weak, by resisting the compression during pericentre passage. Magnetic fields in eccentric discs are an understudied area and the work of this thesis suggests they may be extremely important for TDE discs.

13.2 Future work

There are many promising avenues for future research that follow on from this thesis. There are fundamental questions which remain unanswered (particularly regarding magnetic fields in eccentric discs), important physical effects which have been neglected, or not yet fully developed (such as self-gravity and non ideal terms), and there are many avenues for extending the theoretical framework presented here to more general accretion discs.

The combination of the affine disc model and Whitham's averaging technique has the potential to be a very powerful way of deriving thin disc equations for various systems. This was shown in Chapter 5 where it was used to obtain the equations for a thin eccentric disc. Extending the affine disc model to include dissipative effects by use of a pseudo-Lagrangian prescription will provide a convenient way of deriving the non-ideal terms in an eccentric disc theory. It should be possible to generalise this method to develop a thin disc description in potentials other than a Newtonian point mass (e.g. Paczyński–Wiita potential or circumbinary discs). To do so will require dealing with the fact that, in general, the test particle orbits will not be closed in the inertial frame.

The test particle motion may, however, be closed in some rotating frame. Assume that the Lagrangian takes the form $L = L_T + L_F$, where L_T is the Lagrangian of the test particle motion and L_F contains the contribution from fluid forces. The test particle orbits described by L_T are not closed in general, but may be closed in a suitable rotating frame. Defining $\omega(\mathbf{x}, \mathbf{u})$ to equal the constant precession rate of the frame in

which the orbit passing through $(\mathbf{x}, \mathbf{u})$ is closed we can define a modified test particle Lagrangian,

$$L_1 = L_T + \int \Sigma_0 \frac{R^2}{2} \left(\omega(\mathbf{x}, \mathbf{u})^2 - 2\omega(\mathbf{x}, \mathbf{u})\dot{\phi} \right) d^2\bar{x}_0. \tag{13.1}$$

The extra terms in the Lagrangian mimic the effects of rotation so that the orbits of L_1 are closed in the inertial frame. The orbit of L_1, in the inertial frame, has the same shape as a (closed) orbit of L_T in a frame rotating at an angular frequency $\omega(\mathbf{x}, \mathbf{u})$. In order that the Lagrangian remain unchanged by this procedure we must define a modified fluid Lagrangian, L_2 by subtracting off the additional terms from L_F,

$$L_2 = L_F - \int \Sigma_0 \frac{R^2}{2} \left(\omega(\mathbf{x}, \mathbf{u})^2 - 2\omega(\mathbf{x}, \mathbf{u})\dot{\phi} \right) d^2\bar{x}_0. \tag{13.2}$$

Thus our Lagrangian, $L = L_1 + L_2 = L_T + L_F$, remains unchanged by this process. L_1 can be treated as describing the test particle motion and, after Whitham averaging, L_2 will cause a slow evolution of the orbits described by L_1. For this procedure to be valid we likely require $\Omega \gg \omega(\mathbf{x}, \mathbf{u})$, where Ω is the angular frequency of the test particle motion of L_1. Unlike the nearly Keplerian thin disc theories described in this thesis, however, Ω can depart strongly from Keplerian value and the shape of the orbits need not be ellipses.

In the absence of fluid forces, L_2 will cause differential precession of the orbits described by L_1, ultimately resulting in an orbital intersection and a breakdown of the disc model. When fluid forces are included they should resist the differential precession and (potentially) prevent the orbital intersection. While the above is merely a sketch of how a more general thin disc model might be obtained, it is sufficient to demonstrate the potential power of combining the affine disc model with Whitham's averaging method.

Another area of future study is the problem of the excitation and damping of nonlinear eccentric waves. While a lot of progress has been made in linear theory, most of the work on excitation and damping of nonlinear eccentric waves has been confined to numerical simulations. This is despite the fact that the equations for including these in eccentric discs have been known about since Ogilvie (2001) (or even longer with the work of Syer & Clarke (1992), Syer & Clarke (1993) and Lyubarskij et al. (1994)). Theoretical studies of nonlinear eccentric discs, including most of the current work, have mostly been limited to studies of eccentric modes in ideal theory.

The treatment of self-gravity is now a major omission from ideal eccentric disc theory, despite its importance. Currently, even in linear theory, there is much disagreement

about how to formulate the self-gravity terms, particularly in the softening prescription used to regularise the singularity present in Laplace-Lagrange theory (Sefilian & Rafikov, 2019). Very little is currently known about how to formulate a nonlinear self-gravitating eccentric disc.

Finally there are many unanswered questions when it comes to magnetic fields in eccentric discs, both for large scale coherent fields and the MRI. In the latter case recent work (Dewberry et al., 2020b) has shown that nonlinear eccentric waves can suppress MRI. It is also unclear whether MRI turbulence damps eccentricity, or can excite it through viscous overstability. Thus the complex interplay between eccentric waves and MRI turbulence is a promising area for future research.

References

Abbot B. P., et al., 2017, The Astrophysical Journal, 848, L12
Abramowicz M. A., Czerny B., Lasota J. P., Szuszkiewicz E., 1988, ApJ, 332, 646
Alei E., Claudi R., Bignamini A., Molinaro M., 2020, Astronomy and Computing, 31, 100370
Arcavi I., et al., 2014, ApJ, 793, 38
Arnold V. I., 1978, Mathematical methods of classical mechanics
Artymowicz P., Lubow S. H., 1994, ApJ, 421, 651
Artymowicz P., Clarke C. J., Lubow S. H., Pringle J. E., 1991, ApJ, 370, L35
Asphaug E., Ryan E. V., Zuber M. T., 2002, Asteroids III, 1, 463
Ataiee S., Pinilla P., Zsom A., Dullemond C. P., Dominik C., Ghanbari J., 2013, A&A, 553, L3
Auchettl K., Guillochon J., Ramirez-Ruiz E., 2017, ApJ, 838, 149
Bai Y., Crumeyrolle O., Mutabazi I., 2015, Phys. Rev. E, 92, 031001
Balbus S. A., Hawley J. F., 1991, ApJ, 376, 214
Bardeen J. M., Petterson J. A., 1975, ApJ, 195, L65
Barker A. J., Ogilvie G. I., 2014, MNRAS, 445, 2637
Barker A. J., Ogilvie G. I., 2016, MNRAS, 458, 3739
Begelman M. C., Pringle J. E., 2007, MNRAS, 375, 1070
Benisty M., et al., 2015, A&A, 578, L6
Benisty M., et al., 2018, A&A, 619, A171
Berger E., Zauderer A., Pooley G. G., Soderberg A. M., Sari R., Brunthaler A., Bietenholz M. F., 2012, ApJ, 748, 36
Béthune W., Lesur G., Ferreira J., 2016a, in Reylé C., Richard J., Cambrésy L., Deleuil M., Pécontal E., Tresse L., Vauglin I., eds, SF2A-2016: Proceedings of the Annual meeting of the French Society of Astronomy and Astrophysics. pp 129–132
Béthune W., Lesur G., Ferreira J., 2016b, A&A, 589, A87
Béthune W., Lesur G., Ferreira J., 2017, A&A, 600, A75
Bisnovatyi-Kogan G. S., Blinnikov S. I., 1976, Soviet Astronomy Letters, 2, 191
Blandford R. D., Payne D. G., 1982, MNRAS, 199, 883
Blandford R. D., Znajek R. L., 1977, MNRAS, 179, 433
Bloom J. S., et al., 2011, Science, 333, 203
Boldyrev S., Huynh D., Pariev V., 2009, Phys. Rev. E, 80, 066310
Bollada P. C., Phillips T. N., 2012, Archive for Rational Mechanics and Analysis, 205, 1
Bonnerot C., Lu W., 2019, arXiv e-prints, p. arXiv:1906.05865
Bonnerot C., Rossi E. M., Lodato G., Price D. J., 2016, MNRAS, 455, 2253
Bonnerot C., Price D. J., Lodato G., Rossi E. M., 2017, MNRAS, 469, 4879
Borderies N., Goldreich P., Tremaine S., 1983, AJ, 88, 1560

Borderies N., Goldreich P., Tremaine S., 1985, Icarus, 63, 406
Borderies N., Goldreich P., Tremaine S., 1986, Icarus, 68, 522
Borderies N., Goldreich P., Tremaine S., 1989, Icarus, 80, 344
Braga-Ribas F., et al., 2014, Nature, 508, 72
Burrows D. N., et al., 2011, Nature, 476, 421
Butler R. P., et al., 2006, ApJ, 646, 505
Cannizzo J. K., Ghosh P., Wheeler J. C., 1982, ApJ, 260, L83
Canup R. M., 2010, Nature, 468, 943
Cao R., Liu F. K., Zhou Z. Q., Komossa S., Ho L. C., 2018, MNRAS, 480, 2929
Cauley P. W., Farihi J., Redfield S., Bachman S., Parsons S. G., Gänsicke B. T., 2018, ApJ, 852, L22
Chakraborty S., Bhattacharyya S., Chakraborty C. r., Rao A. R., 2020, MNRAS, 492, 1634
Chan C.-H., Krolik J. H., Piran T., 2018, ApJ, 856, 12
Chandrasekhar S., 1960, Proceedings of the National Academy of Science, 46, 253
Chandrasekhar S., 1967
Chang P., Oishi J. S., 2010, ApJ, 721, 1593
Chen J.-H., Shen R.-F., 2018, ApJ, 867, 20
Chornock R., et al., 2014, ApJ, 780, 44
Clausen D., Eracleous M., 2011, ApJ, 726, 34
Coleman M. S. B., Kotko I., Blaes O., Lasota J. P., Hirose S., 2016, MNRAS, 462, 3710
Collins G. S., Lynch E., McAdam R., Davison T. M., 2017, Meteoritics and Planetary Science, 52, 1542
Colwell J. E., 1994, Planet. Space Sci., 42, 1139
Copeland J. A., Heard J. F., 1963, Publications of the David Dunlap Observatory, 2, 317
Coughlin E. R., Nixon C. J., 2019, ApJ, 883, L17
Cuzzi J. N., Scargle J. D., 1985, ApJ, 292, 276
Cuzzi J. N., Lissauer J. J., Shu F. H., 1981, Nature, 292, 703
D'Angelo G., Henning T., Kley W., 2002, A&A, 385, 647
Das U., Begelman M. C., Lesur G., 2018, MNRAS, 473, 2791
De Marco O., Izzard R. G., 2017, Publ. Astron. Soc. Australia, 34, e001
Debes J. H., Sigurdsson S., 2002, ApJ, 572, 556
Deibert E., Petrovich C., Wu Y., 2018, in Diversis Mundi: The Solar System in an Exoplanetary Context. p. 12, doi:10.5281/zenodo.1317487
Deng H., Mayer L., Latter H., 2020, ApJ, 891, 154
Dermine T., Izzard R. G., Jorissen A., Van Winckel H., 2013, A&A, 551, A50
Dermott S. F., Grogan K., Holmes E. K., Wyatt M. C., 1998, in Exozodiacal Dust Workshop:. p. 59
Dewberry J. W., Latter H. N., Ogilvie G. I., 2018, MNRAS, 476, 4085
Dewberry J. W., Latter H. N., Ogilvie G. I., 2019, MNRAS, 483, 1609
Dewberry J. W., Latter H. N., Ogilvie G. I., Fromang S., 2020a, MNRAS, 497, 435
Dewberry J. W., Latter H. N., Ogilvie G. I., Fromang S., 2020b, MNRAS, 497, 451
Doğan S., Nixon C., King A., Price D. J., 2015, MNRAS, 449, 1251
Doyle L. R., et al., 2011, Science, 333, 1602
Durisen R. H., Boss A. P., Mayer L., Nelson A. F., Quinn T., Rice W. K. M., 2007, in Reipurth B., Jewitt D., Keil K., eds, Protostars and Planets V. p. 607 (arXiv:astro-ph/0603179)

Elliot J. L., 1984, in Brahic A., ed., Planetary Rings. p. 155
Esposito L. W., et al., 1983, Science, 222, 57
Evans C. R., Kochanek C. S., 1989, ApJ, 346, L13
Event Horizon Telescope Collaboration et al., 2019, ApJ, 875, L5
Fabian A. C., Iwasawa K., Reynolds C. S., Young A. J., 2000, PASP, 112, 1145
Faulkner J., Lin D. N. C., Papaloizou J., 1983, MNRAS, 205, 359
Ferreira B. T., Ogilvie G. I., 2008, in Axelsson M., ed., American Institute of Physics Conference Series Vol. 1054, American Institute of Physics Conference Series. pp 111–116 (arXiv:0810.0116), doi:10.1063/1.3002492
Ferreira B. T., Ogilvie G. I., 2009, MNRAS, 392, 428
Field G. B., 1965, ApJ, 142, 531
Fleming T. P., Stone J. M., Hawley J. F., 2000, ApJ, 530, 464
Ford E. B., Rasio F. A., 2008, ApJ, 686, 621
Franchini A., Lubow S. H., Martin R. G., 2019, ApJ, 880, L18
Frankowski A., Jorissen A., 2007, Baltic Astronomy, 16, 104
Fu W., Lai D., 2009, ApJ, 690, 1386
Fu W., Lai D., 2013, Monthly Notices of the Royal Astronomical Society, 431, 3697
Fu W., Lubow S. H., Martin R. G., 2015a, ApJ, 807, 75
Fu W., Lubow S. H., Martin R. G., 2015b, ApJ, 813, 105
Fu W., Lubow S. H., Martin R. G., 2017, ApJ, 835, L29
Fukumura K., Tsuruta S., 2004, ApJ, 613, 700
Gammie C. F., 1996, ApJ, 457, 355
Gammie C. F., 2001, ApJ, 553, 174
Gammie C. F., Menou K., 1998, ApJ, 492, L75
Gelfand I. M., Silverman R. A., et al., 2000, Calculus of variations. Courier Corporation
Gezari S., et al., 2006, ApJ, 653, L25
Gezari S., et al., 2008, ApJ, 676, 944
Goldreich P., Lynden-Bell D., 1965, MNRAS, 130, 125
Goldreich P., Tremaine S. D., 1978a, Icarus, 34, 240
Goldreich P., Tremaine S. D., 1978b, Icarus, 34, 240
Goldreich P., Tremaine S., 1980, ApJ, 241, 425
Gomez S., et al., 2020, arXiv e-prints, p. arXiv:2003.05469
González Martínez-País I., Shahbaz T., Casares Velázquez J., 2014, Accretion Processes in Astrophysics
Goodchild S., Ogilvie G., 2006, MNRAS, 368, 1123
Goodman J., 2003, MNRAS, 339, 937
Gratton R., et al., 2019, A&A, 623, A140
Greaves J. S., et al., 2016, MNRAS, 461, 3910
Groot P. J., 2001, ApJ, 551, L89
Guillochon J., McCourt M., 2017, ApJ, 834, L19
Guillochon J., Ramirez-Ruiz E., 2013, ApJ, 767, 25
Guillochon J., Manukian H., Ramirez-Ruiz E., 2014, ApJ, 783, 23
Hamers A. S., Lai D., 2017, MNRAS, 470, 1657
Hameury J. M., 2020, Advances in Space Research, 66, 1004
Hanuschik R. W., 1987, A&A, 173, 299
Hanuschik R. W., 1988, A&A, 190, 187
Harlaftis E. T., Steeghs D., Horne K., Martín E., Magazzú A., 1999, MNRAS, 306, 348

Harris A. W., 1984, in Greenberg R., Brahic A., eds, IAU Colloq. 75: Planetary Rings. pp 641–659
Hartnoll S. A., Blackman E. G., 2002, MNRAS, 332, L1
Hawley J. F., Balbus S. A., 1991, The Astrophysical Journal, 376, 223
Hawley J. F., Stone J. M., 1998, ApJ, 501, 758
Hawley J. F., Gammie C. F., Balbus S. A., 1995, ApJ, 440, 742
Hayasaki K., Stone N., Loeb A., 2013, MNRAS, 434, 909
Hayasaki K., Stone N., Loeb A., 2016, MNRAS, 461, 3760
Held L. E., 2020, PhD thesis, University of Cambridge, Cambridge, UK
Hessman F. V., Mantel K. H., Barwig H., Schoembs R., 1992, A&A, 263, 147
Hills J. G., 1975, Nature, 254, 295
Hines D. C., et al., 2007, ApJ, 671, L165
Hirose M., Osaki Y., 1990, PASJ, 42, 135
Hirose S., Krolik J. H., Blaes O., 2009, ApJ, 691, 16
Holoien T. W. S., et al., 2019, ApJ, 880, 120
Hughes A. M., et al., 2017, ApJ, 839, 86
Hung T., et al., 2020, arXiv e-prints, p. arXiv:2003.09427
Hunter C., 1973, ApJ, 181, 685
Ichikawa S., Hirose M., Osaki Y., 1993, PASJ, 45, 243
Ingram A., Motta S., 2014, MNRAS, 444, 2065
Ireland M. J., Kraus A. L., 2008, ApJ, 678, L59
Ito T., Ohtsuka K., 2019, Monographs on Environment, Earth and Planets, 7, 1
Ivanov P. B., Papaloizou J. C. B., Polnarev A. G., 1999, MNRAS, 307, 79
Izzard R., Jermyn A., 2018, Galaxies, 6, 97
Izzard R. G., Dermine T., Church R. P., 2010, A&A, 523, A10
Jiang Y.-F., Stone J. M., Davis S. W., 2013, ApJ, 778, 65
Jimenez J., Whitham G. B., 1976, Proceedings of the Royal Society of London Series A, 349, 277
Jin L., 1996, ApJ, 457, 798
Joergens V., Spruit H. C., Rutten R. G. M., 2000, A&A, 356, L33
Johns-Krull C. M., et al., 2016, ApJ, 826, 206
Jorissen A., Van Eck S., Mayor M., Udry S., 1998, A&A, 332, 877
Julian W. H., Toomre A., 1966, ApJ, 146, 810
Jura M., 2003, ApJ, 584, L91
Jura M., 2008, AJ, 135, 1785
Jura M., Young E. D., 2014, Annual Review of Earth and Planetary Sciences, 42, 45
Jurić M., Tremaine S., 2008, ApJ, 686, 603
Karas V., Martocchia A., Subr L., 2001, PASJ, 53, 189
Kato S., 1978, MNRAS, 185, 629
Kato S., 1990, PASJ, 42, 99
Kato S., 2004, PASJ, 56, 905
Kato S., 2008, PASJ, 60, 111
Kato S., Fukue J., 1980, PASJ, 32, 377
Kawamuro T., Ueda Y., Shidatsu M., Hori T., Kawai N., Negoro H., Mihara T., 2017, in Serino M., Shidatsu M., Iwakiri W., Mihara T., eds, 7 years of MAXI: monitoring X-ray Transients. p. 199
Kennedy G. M., 2020, Royal Society Open Science, 7, 200063
Kennedy G. M., et al., 2019, Nature Astronomy, 3, 230

Keppler M., et al., 2020, arXiv e-prints, p. arXiv:2005.09037
Kley W., Dirksen G., 2006, A&A, 447, 369
Kley W., Nelson R. P., 2012, ARA&A, 50, 211
Kley W., Papaloizou J. C. B., Ogilvie G. I., 2008, A&A, 487, 671
Kley W., Thun D., Penzlin A. B. T., 2019, A&A, 627, A91
Kochanek C. S., 1994, ApJ, 422, 508
Koester D., Provencal J., Shipman H. L., 1997, A&A, 320, L57
Kolykhalov P. I., Syunyaev R. A., 1980, Soviet Astronomy Letters, 6, 357
Komberg B. V., 2007, in Karas V., Matt G., eds, IAU Symposium Vol. 238, Black Holes from Stars to Galaxies – Across the Range of Masses. pp 393–394, doi:10.1017/S1743921307005613
Kóspál Á., et al., 2013, ApJ, 776, 77
Kostov V. B., et al., 2014, ApJ, 784, 14
Kostov V. B., et al., 2016, ApJ, 827, 86
Kozai Y., 1962, AJ, 67, 591
Kozlowski M., Wiita P. J., Paczynski B., 1979, Acta Astron., 29, 157
Kral Q., Wyatt M., Carswell R. F., Pringle J. E., Matrà L., Juhász A., 2016, MNRAS, 461, 845
Kral Q., Marino S., Wyatt M. C., Kama M., Matrà L., 2019, MNRAS, 489, 3670
Kratter K., Lodato G., 2016, ARA&A, 54, 271
Krolik J., Piran T., Svirski G., Cheng R. M., 2016, ApJ, 827, 127
Kumar S., 1987, MNRAS, 225, 823
Kumar S., 1990, MNRAS, 245, 670
Kumar S., Pringle J. E., 1985, MNRAS, 213, 435
Kunz M. W., Balbus S. A., 2004, MNRAS, 348, 355
Kunz M. W., Lesur G., 2013, MNRAS, 434, 2295
Kunze S., Speith R., Riffert H., 1997, MNRAS, 289, 889
Lai D., Pu B., 2017, AJ, 153, 42
Lasota J.-P., 2001, New Astron. Rev., 45, 449
Latter H. N., Ogilvie G. I., 2006, MNRAS, 372, 1829
Latter H. N., Papaloizou J. C. B., 2012, MNRAS, 426, 1107
Latter H. N., Papaloizou J., 2017, MNRAS, 472, 1432
Lee E. J., Chiang E., 2016, ApJ, 827, 125
Lee E., Goodman J., 1999, MNRAS, 308, 984
Lee W.-K., Dempsey A. M., Lithwick Y., 2018, arXiv e-prints,
Lee W.-K., Dempsey A. M., Lithwick Y., 2019a, arXiv e-prints, p. arXiv:1906.05290
Lee Y.-M., Lee H.-W., Lee H.-G., Angeloni R., 2019b, MNRAS, 487, 2166
Leinhardt Z. M., Ogilvie G. I., Latter H. N., Kokubo E., 2012, MNRAS, 424, 1419
Lesur G., Papaloizou J. C. B., 2009, A&A, 498, 1
Lidov M. L., 1962, Planet. Space Sci., 9, 719
Lieman-Sifry J., Hughes A. M., Carpenter J. M., Gorti U., Hales A., Flaherty K. M., 2016, ApJ, 828, 25
Lin J. W., Chiang E., 2019, ApJ, 883, 68
Lin D. N. C., Papaloizou J., 1979, MNRAS, 186, 799
Lin D., Papaloizou J., 1986, The Astrophysical Journal, 307, 395
Lin D. N. C., Papaloizou J. C. B., 1993, in Levy E. H., Lunine J. I., eds, Protostars and Planets III. p. 749
Lin C. C., Shu F. H., 1964, ApJ, 140, 646

Lin C. C., Shu F. H., 1966, Proceedings of the National Academy of Science, 55, 229
Liska M., Hesp C., Tchekhovskoy A., Ingram A., van der Klis M., Markoff S. B., Van Moer M., 2020, MNRAS,
Lissauer J. J., Goldreich P., Tremaine S., 1985, Icarus, 64, 425
Lithwick Y., 2009, ApJ, 693, 85
Lithwick Y., Wu Y., 2011, ApJ, 739, 31
Liu F. K., Zhou Z. Q., Cao R., Ho L. C., Komossa S., 2017, MNRAS, 472, L99
Loeb A., Ulmer A., 1997, ApJ, 489, 573
Lu W., Bonnerot C., 2019, MNRAS, p. 3033
Lubow S. H., 1981, ApJ, 245, 274
Lubow S. H., 1991a, ApJ, 381, 259
Lubow S. H., 1991b, ApJ, 381, 268
Lubow S. H., Martin R. G., 2018, MNRAS, 473, 3733
Lubow S. H., Ogilvie G. I., 2017, MNRAS, 469, 4292
Lubow S. H., Seibert M., Artymowicz P., 1999, ApJ, 526, 1001
Luna G. J. M., 2019, Boletin de la Asociacion Argentina de Astronomia La Plata Argentina, 61, 93
Luna G. J. M., Sokoloski J. L., Mukai K., Nelson T., Nuñez N. E., 2014, in Revista Mexicana de Astronomia y Astrofisica Conference Series. pp 158–159
Lynch E. M., Ogilvie G. I., 2019, MNRAS, 488, 1127
Lynden-Bell D., 1969, Nature, 223, 690
Lynden-Bell D., Pringle J. E., 1974, MNRAS, 168, 603
Lyubarskij Y. E., Postnov K. A., Prokhorov M. E., 1994, MNRAS, 266, 583
MacFadyen A. I., Milosavljević M., 2008, ApJ, 672, 83
Maguire K., Eracleous M., Jonker P. G., MacLeod M., Rosswog S., 2020, Space Sci. Rev., 216, 39
Malamud U., Perets H., 2019a, arXiv e-prints
Malamud U., Perets H., 2019b, arXiv e-prints, p. arXiv:1911.12068
Maness H. L., et al., 2009, ApJ, 707, 1098
Manser C. J., et al., 2016a, MNRAS, 455, 4467
Manser C. J., Gänsicke B. T., Koester D., Marsh T. R., Southworth J., 2016b, MNRAS, 462, 1461
Manser C. J., Gänsicke B. T., Gentile Fusillo N. P., Ashley R., Breedt E., Hollands M., Izquierdo P., Pelisoli I., 2020, MNRAS, 493, 2127
Marino S., et al., 2016, MNRAS, 460, 2933
Martin R. G., Lubow S. H., 2017, ApJ, 835, L28
Martin R. G., Lubow S. H., 2018, MNRAS, 479, 1297
Martin R. G., Lubow S. H., 2019, MNRAS, 490, 1332
Martin R. G., Nixon C., Lubow S. H., Armitage P. J., Price D. J., Doğan S., King A., 2014, ApJ, 792, L33
Mathieu R. D., 1994, ARA&A, 32, 465
Mathieu R. D., Walter F. M., Myers P. C., 1989, AJ, 98, 987
Matrà L., et al., 2017a, MNRAS, 464, 1415
Matrà L., et al., 2017b, ApJ, 842, 9
Matsakos T., Königl A., 2016, ApJ, 820, L8
Mattila S., et al., 2018, Science, 361, 482
McKinney J. C., Tchekhovskoy A., Bland ford R. D., 2012, MNRAS, 423, 3083
Meier D. L., 1979, ApJ, 233, 664

Metzger B. D., Rafikov R. R., Bochkarev K. V., 2012, MNRAS, 423, 505
Meyer F., Meyer-Hofmeister E., 1981, A&A, 104, L10
Meyer F., Liu B. F., Meyer-Hofmeister E., 2000, A&A, 361, 175
Mezcua M., 2017, International Journal of Modern Physics D, 26, 1730021
Miles J., Salmon R., 1985, Journal of Fluid Mechanics, 157, 519–531
Miner E. D., Wessen R. R., Cuzzi J. N., 2007, Planetary Ring Systems
Mineshige S., Osaki Y., 1983, PASJ, 35, 377
Miranda R., Lai D., 2015, MNRAS, 452, 2396
Miranda R., Rafikov R. R., 2018, preprint, (`arXiv:1802.00464`)
Miranda R., Muñoz D. J., Lai D., 2017, MNRAS, 466, 1170
Mishra B., Fragile P. C., Johnson L. C., Kluźniak W., 2016, MNRAS, 463, 3437
Moór A., et al., 2017, ApJ, 849, 123
Morales-Rueda L., Marsh T. R., Billington I., 2000, MNRAS, 313, 454
Morton D. C., 1960, ApJ, 132, 146
Muñoz D. J., Lai D., 2016, ApJ, 827, 43
Muñoz D. J., Lithwick Y., 2020, arXiv e-prints, p. arXiv:2008.08085
Muñoz D. J., Miranda R., Lai D., 2019, ApJ, 871, 84
Muñoz D. J., Lai D., Kratter K., Mirand a R., 2020, ApJ, 889, 114
Murray J. R., 1998, MNRAS, 297, 323
Mustill A. J., Davies M. B., Johansen A., 2017, MNRAS, 468, 3000
Mutter M. M., Pierens A., Nelson R. P., 2017, MNRAS, 465, 4735
Naoz S., 2016, ARA&A, 54, 441
Narayan R., Yi I., 1995, ApJ, 452, 710
Nealon R., Price D. J., Nixon C. J., 2015, MNRAS, 448, 1526
Nealon R., Nixon C., Price D. J., King A., 2016, MNRAS, 455, L62
Nicholson P. D., Matthews K., Goldreich P., 1982, AJ, 87, 433
Nityananda R., Narayan R., 1984, Advances in Space Research, 3, 319
Nixon C., King A., Price D., Frank J., 2012, ApJ, 757, L24
Nixon C., King A., Price D., 2013, MNRAS, 434, 1946
Nixon C. J., Pringle J. E., Coughlin E. R., Swan A., Farihi J., 2020, arXiv e-prints, p. arXiv:2006.07639
Ogilvie G. I., 2000, MNRAS, 317, 607
Ogilvie G. I., 2001, MNRAS, 325, 231
Ogilvie G. I., 2002a, MNRAS, 330, 937
Ogilvie G. I., 2002b, MNRAS, 331, 1053
Ogilvie G. I., 2003, MNRAS, 340, 969
Ogilvie G. I., 2007, MNRAS, 374, 131
Ogilvie G. I., 2008, MNRAS, 388, 1372
Ogilvie G. I., 2016, Journal of Plasma Physics, 82, 205820301
Ogilvie G. I., 2018, MNRAS, 477, 1744
Ogilvie G. I., Barker A. J., 2014, MNRAS, 445, 2621
Ogilvie G. I., Lynch E. M., 2019, MNRAS, 483, 4453
Ogilvie G. I., Potter A. T., 2008, Phys. Rev. Lett., 100, 074503
Ogilvie G. I., Pringle J. E., 1996, MNRAS, 279, 152
Ogilvie G. I., Proctor M. R. E., 2003, Journal of Fluid Mechanics, 476, 389
Okazaki A. T., 1991, PASJ, 43, 75
Okazaki A. T., 1997, A&A, 318, 548
Okazaki A. T., Kato S., Fukue J., 1987, PASJ, 39, 457

Orosz J. A., et al., 2012a, Science, 337, 1511
Orosz J. A., et al., 2012b, ApJ, 758, 87
Ortiz J. L., et al., 2015, A&A, 576, A18
Ortiz J. L., et al., 2017, Nature, 550, 219
Osaki Y., 1985, A&A, 144, 369
Osaki Y., 1989, PASJ, 41, 1005
Osaki Y., 1996, PASP, 108, 39
Paczyński B., 1971, ARA&A, 9, 183
Paczynski B., 1977, ApJ, 216, 822
Paczynski B., 1978a, Acta Astron., 28, 91
Paczynski B., 1978b, Acta Astron., 28, 241
Paczyńsky B., Wiita P. J., 1980, A&A, 88, 23
Pan M., Nesvold E. R., Kuchner M. J., 2016, ApJ, 832, 81
Papaloizou J. C. B., 2002, A&A, 388, 615
Papaloizou J. C. B., 2005a, A&A, 432, 743
Papaloizou J. C. B., 2005b, A&A, 432, 757
Papaloizou J., Lin D. N. C., 1984, ApJ, 285, 818
Papaloizou J. C. B., Lin D. N. C., 1989, ApJ, 344, 645
Papaloizou J. C. B., Pringle J. E., 1983, MNRAS, 202, 1181
Papaloizou J. C. B., Pringle J. E., 1985, MNRAS, 213, 799
Papaloizou J. C. B., Pringle J. E., 1987, MNRAS, 225, 267
Papaloizou J. C., Savonije G. J., 1991, MNRAS, 248, 353
Papaloizou J. C. B., Savonije G. J., 2006, A&A, 456, 1097
Papaloizou J. C. B., Terquem C., 2001, MNRAS, 325, 221
Patterson J., 1979, AJ, 84, 804
Peiris H. V., Tremaine S., 2003, ApJ, 599, 237
Pérez L. M., Isella A., Carpenter J. M., Chandler C. J., 2014, ApJ, 783, L13
Pessah M. E., Psaltis D., 2005, ApJ, 628, 879
Petrovic J. J., 2001, Journal of Materials Science, 36, 1579
Petrovich C., Deibert E., Wu Y., 2019, AJ, 157, 180
Phinney E. S., 1989, in Morris M., ed., IAU Symposium Vol. 136, The Center of the Galaxy. p. 543
Pichardo B., Sparke L. S., Aguilar L. A., 2005, MNRAS, 359, 521
Pichardo B., Sparke L. S., Aguilar L. A., 2008, MNRAS, 391, 815
Pierens A., Nelson R. P., 2008, A&A, 482, 333
Pierens A., McNally C. P., Nelson R. P., 2020, arXiv e-prints, p. arXiv:2005.14693
Piétu V., Gueth F., Hily-Blant P., Schuster K. F., Pety J., 2011, A&A, 528, A81
Piran T., 1978, ApJ, 221, 652
Piran T., Svirski G., Krolik J., Cheng R. M., Shiokawa H., 2015, ApJ, 806, 164
Porco C. C., 1983, PhD thesis, California Institute of Technology, Pasadena.
Porco C. C., 1990, Advances in Space Research, 10, 221
Porco C., Borderies N., Danielson G. E., Goldreich P., Holberg J. B., Lane A. L., Nicholson P. D., 1984a, in Brahic A., ed., Planetary Rings. p. 259
Porco C., Nicholson P. D., Borderies N., Danielson G. E., Goldreich P., Holberg J. B., Lane A. L., 1984b, Icarus, 60, 1
Potter W. J., Balbus S. A., 2014, MNRAS, 441, 681
Potter W. J., Balbus S. A., 2017, MNRAS, 472, 3021
Pringle J. E., 1976, MNRAS, 177, 65

Pringle J. E., 1981, ARA&A, 19, 137
Pringle J. E., 1992, MNRAS, 258, 811
Pu B., Lai D., 2018, MNRAS, 478, 197
Rafikov R. R., 2011a, MNRAS, 416, L55
Rafikov R. R., 2011b, ApJ, 732, L3
Ragusa E., Dipierro G., Lodato G., Laibe G., Price D. J., 2017, MNRAS, 464, 1449
Ragusa E., Alexander R., Calcino J., Hirsh K., Price D. J., 2020, MNRAS,
Railton A. D., Papaloizou J. C. B., 2014, MNRAS, 445, 4409
Rasio F. A., Ford E. B., 1996, Science, 274, 954
Rees M. J., 1988, Nature, 333, 523
Remillard R. A., McClintock J. E., 2006, ARA&A, 44, 49
Ricarte A., Moldvai N., Hughes A. M., Duchêne G., Williams J. P., Andrews S. M., Wilner D. J., 2013, ApJ, 774, 80
Richards M. T., 2007, in Hartkopf W. I., Harmanec P., Guinan E. F., eds, IAU Symposium Vol. 240, Binary Stars as Critical Tools & Tests in Contemporary Astrophysics. pp 160–169, doi:10.1017/S1743921307003973
Riols A., Latter H., 2018, MNRAS, 474, 2212
Riols A., Latter H., 2019, MNRAS, 482, 3989
Riols A., Rincon F., Cossu C., Lesur G., Ogilvie G. I., Longaretti P. Y., 2015, A&A, 575, A14
Rolfe D. J., Haswell C. A., Patterson J., 2000, MNRAS, 317, 759
Rosotti G. P., Booth R. A., Clarke C. J., Teyssandier J., Facchini S., Mustill A. J., 2017, MNRAS, 464, L114
Ross J., Latter H. N., Tehranchi M., 2017, MNRAS, 468, 2401
Sakimoto P. J., Coroniti F. V., 1981, ApJ, 247, 19
Salow R. M., Statler T. S., 2001, ApJ, 551, L49
Savonije G. J., Heemskerk M. H. M., 1993, A&A, 276, 409
Scepi N., Lesur G., Dubus G., Flock M., 2018a, A&A, 609, A77
Scepi N., Lesur G., Dubus G., Flock M., 2018b, A&A, 620, A49
Schneider J., Dedieu C., Le Sidaner P., Savalle R., Zolotukhin I., 2011, A&A, 532, A79
Schwamb M. E., et al., 2013, ApJ, 768, 127
Sefilian A. A., Rafikov R. R., 2019, MNRAS, 489, 4176
Shakura N. I., Sunyaev R. A., 1973, A&A, 500, 33
Shakura N. I., Sunyaev R. A., 1976, MNRAS, 175, 613
Shen R.-F., Matzner C. D., 2014, ApJ, 784, 87
Shiokawa H., Krolik J. H., Cheng R. M., Piran T., Noble S. C., 2015, ApJ, 804, 85
Showalter M. R., 1991, Nature, 351, 709
Shu F. H., 1984, in Greenberg R., Brahic A., eds, IAU Colloq. 75: Planetary Rings. pp 513–561
Sicardy B., et al., 2014, in AAS/Division for Planetary Sciences Meeting Abstracts #46. p. 408.01
Simkin S. M., 1970, ApJ, 159, 463
Sądowski A., Tejeda E., Gafton E., Rosswog S., Abarca D., 2016, MNRAS, 458, 4250
Smak J., 1982, Acta Astron., 32, 199
Smak J., 1996, Dwarf nova outbursts and superoutbursts. p. 45, doi:10.1007/978-94-009-0325-8_14
Smak J., 2000, New Astron. Rev., 44, 171
Smak J. I., 2001, Acta Astron., 51, 295

Smith A. J., Haswell C. A., Murray J. R., Truss M. R., Foulkes S. B., 2007, MNRAS, 378, 785
Statler T. S., 1999, ApJ, 524, L87
Statler T. S., 2001, AJ, 122, 2257
Statler T. S., King I. R., Crane P., Jedrzejewski R. I., 1999, AJ, 117, 894
Steeghs D., 2001, Spiral Waves in Accretion Discs - Observations. p. 45
Steeghs D., Horne K., Marsh T. R., Donati J. F., 1996, MNRAS, 281, 626
Steeghs D., Harlaftis E. T., Horne K., 1997, MNRAS, 290, L28
Stella L., Vietri M., 1998, ApJ, 492, L59
Stella L., Vietri M., 1999, Phys. Rev. Lett., 82, 17
Stone N. C., Metzger B. D., 2016, MNRAS, 455, 859
Syer D., Clarke C. J., 1992, MNRAS, 255, 92
Syer D., Clarke C. J., 1993, MNRAS, 260, 463
Tchekhovskoy A., Narayan R., McKinney J. C., 2011, MNRAS, 418, L79
Teyssandier J., Lai D., 2019, MNRAS, 490, 4353
Teyssandier J., Lai D., 2020, MNRAS, 495, 3920
Teyssandier J., Ogilvie G. I., 2016, MNRAS, 458, 3221
Teyssandier J., Ogilvie G. I., 2017, MNRAS, 467, 4577
Teyssandier J., Lai D., Vick M., 2019, MNRAS, 486, 2265
Thun D., Kley W., 2018, A&A, 616, A47
Thun D., Kley W., Picogna G., 2017, A&A, 604, A102
Tremaine S., 1995, AJ, 110, 628
Tremaine S., 2001, AJ, 121, 1776
Van Eylen V., et al., 2019, AJ, 157, 61
Van Winckel H., Waelkens C., Jorissen A., Van Eck S., North P., Waters R., 1997, in Habing H. J., Lamers H. J. G. L. M., eds, IAU Symposium Vol. 180, Planetary Nebulae. p. 374
Velikhov E., 1959, Sov. Phys. JETP, 36, 995
Vogt N., 1980, A&A, 88, 66
Vogt N., 1982, ApJ, 252, 653
Wang T.-G., Zhou H.-Y., Wang L.-F., Lu H.-L., Xu D., 2011, ApJ, 740, 85
Wang T.-G., Zhou H.-Y., Komossa S., Wang H.-Y., Yuan W., Yang C., 2012, ApJ, 749, 115
Wardle M., Koenigl A., 1993, ApJ, 410, 218
Warner B., 1986, Ap&SS, 118, 271
Warner B., 2003, Cataclysmic Variable Stars, doi:10.1017/CBO9780511586491.
Warner B., Woudt P. A., 2005, in Hameury J. M., Lasota J. P., eds, Astronomical Society of the Pacific Conference Series Vol. 330, The Astrophysics of Cataclysmic Variables and Related Objects. p. 227 (arXiv:astro-ph/0409287)
Warner B., Woudt P. A., 2008, in Axelsson M., ed., American Institute of Physics Conference Series Vol. 1054, American Institute of Physics Conference Series. pp 101–110 (arXiv:0806.1317), doi:10.1063/1.3002491
Welsh W. F., et al., 2012, Nature, 481, 475
Welsh W. F., et al., 2015, ApJ, 809, 26
White C. J., Stone J. M., Quataert E., 2019, ApJ, 874, 168
Whitehurst R., 1988, MNRAS, 232, 35
Whitham G. B., 1965, Journal of Fluid Mechanics, 22, 273
Wienkers A. F., Ogilvie G. I., 2018, MNRAS, 477, 4838

Wisdom J., 1980, AJ, 85, 1122
Wu Y., Lithwick Y., 2011, ApJ, 735, 109
Wyatt M. C., Dermott S. F., Telesco C. M., Fisher R. S., Grogan K., Holmes E. K., Piña R. K., 1999, ApJ, 527, 918
Yang J., Paragi Z., van der Horst A. J., Gurvits L. I., Campbell R. M., Giannios D., An T., Komossa S., 2016, MNRAS, 462, L66
Yang Y., et al., 2017, AJ, 153, 7
Zamanov R. K., et al., 2020, MNRAS,
Zanazzi J. J., Lai D., 2017, MNRAS, 467, 1957
Zanazzi J. J., Lai D., 2018, MNRAS, 473, 603
Zanazzi J. J., Ogilvie G. I., 2020, arXiv e-prints, p. arXiv:2009.06636
Zhang C. M., Dolgov A., 2002, International Journal of Modern Physics D, 11, 503
Zhu Z., 2015, ApJ, 799, 16
Zuckerman B., Koester D., Melis C., Hansen B. M., Jura M., 2007, ApJ, 671, 872
de Boer J., et al., 2016, A&A, 595, A114
de Lagrange J. L., 1853, Mécanique analytique. Vol. 1, Mallet-Bachelier
van Velzen S., et al., 2011, ApJ, 741, 73
von Zeipel H., 1910, Astronomische Nachrichten, 183, 345

Appendix A

Properties of Various Integrals Used in this Thesis

A.1 Behaviour of the wave action integrals

In this Appendix we introduce various functionals of F which are important in describing the nonlinear behaviour of the eccentric waves. We shall also prove a few results which are important for deriving conditions on the behaviour of the short wavelength eccentric waves, given reasonable assumptions about the properties of F.

A.1.1 Untwisted case

First we consider the untwisted short wavelength limit. Throughout this section we shall assume that F

1. Is a monotonically increasing function of $|q|$;
2. Has a monotonically increasing derivative (and thus $F'' > 0$);
3. Diverges (with diverging derivative) as $|q| \to 1$.

These are exactly the same assumptions on F that are used to show that q oscillates between two values q_+ and $-q_+$. Note $F'' = 0$ is a singular point of the equation describing the short lengthscale behaviour (equation 6.19); this singular point separates oscillatory and evanescent behaviour of the solution on the short lengthscale.

All given forms of F considered obey these properties, and any physical forms of F should diverge as $|q| \to 1$ due to the extreme pressure gradients associated with orbital

intersection. If F does not satisfy these properties (but does diverge at $|q| \to 1$) then it is possible that wave cannot exceed some critical q.

We make two additional assumptions about the properties of F which are needed for some of the results in this section. These are that

4. F'' is a nondecreasing function of $|q|$ (or $qF'''(q) \geq 0$);

5. $F''''(q) \geq 0$.

These are additional conditions strengthening the convexity of F. Although reasonable, these are harder to justify than conditions $1-3$ which ought to be satisfied by any physically relevant forms of F which produce oscillatory eccentric waves. Both $F^{(2D)}$ and $F^{(iso)}$ satisfy conditions 5 and 6, and by a continuity argument 3D discs with γ sufficiently close to 1 must remain close to the 2D (or isothermal) cases so that the conditions remain satisfied. It is possible that a sufficiently incompressible disc (or some other disc model not discussed here) might break one of these conditions, in which case the conclusions on the behaviour of eccentric waves in these discs as derived in this paper would have to be revisited for these disc models. The conditions we derive here are fairly conservative sufficiency conditions so even if a disc model breaks condition 5 or 6 the results we derive here may still hold.

We make use of various integrals of F in order to describe the nonlinear behaviour of the eccentric wave. To describe the short lengthscale nonlinear oscillator we make use of

$$I(q, q_+) := \int_0^q F''(x)[F(q_+) - F(x)]^{-1/2} dx \quad , \tag{A.1}$$

$$J(q, q_+) := \int_0^q F''(x)[F(q_+) - F(x)]^{1/2} dx \quad . \tag{A.2}$$

and for the long lengthscale behaviour we define complete forms of these integrals

$$I(q_+) := I(q_+, q_+) = \int_0^{q_+} F''(q)[F(q_+) - F(q)]^{-1/2} dq \quad , \tag{A.3}$$

$$J(q_+) := J(q_+, q_+) = \int_0^{q_+} F''(q)[F(q_+) - F(q)]^{1/2} dq \quad , \tag{A.4}$$

Given the assumptions about the properties of F, it can be seen that these integrals are strictly positive. Whether $I(q_+)$ and $J(q_+)$ are increasing or decreasing functions of q_+ is important in determining the behaviour of the untwisted eccentric modes on the disc lengthscale.

We can show that, for F satisfying the assumptions above, $\partial_{q_+} J(q_+) > 0$:

$$\begin{aligned}\partial_{q_+} J(q_+) &= \int_0^{q_+} \frac{\partial}{\partial q_+}\left(F''(q)[F(q_+) - F(q)]^{1/2}\right) dq \\ &= \frac{1}{2}\int_0^{q_+} F'(q_+)F''(q)[F(q_+) - F(q)]^{-1/2}\, dq \\ &+ \frac{1}{2}F'(q_+)F''(q)[F(q_+) - F(q)]^{-1/2}\Big|_{q=q_+} \\ &> 0 \quad ,\end{aligned} \tag{A.5}$$

as the integrand is positive and the boundary term is zero. Showing a similar result for $I(q_+)$ is less straightforward as differentiating $I(q_+)$ using in the normal manner using the Leibniz rule results in a divergent integral along with a divergent contribution from the boundary. The way in which these cancel is not clear. As such we need to proceed more carefully. Consider an expansion of $I(q_+ + x)$ with $x \ll 1$ and introducing a new renormalised variable $\tilde{q} = \frac{q_+}{q_+ + x} q$ then we obtain the expansion

$$\begin{aligned}I(q_+ + x) &= \int_0^{q_+ + x} F''(q)[F(q_+ + x) - F(q)]^{-1/2}\, dq \\ &= \int_0^{q_+} F''(\frac{q_+ + x}{q_+}\tilde{q})[F(q_+ + x) - F(\frac{q_+ + x}{q_+}\tilde{q})]^{-1/2}\left(\frac{q_+ + x}{q_+}\right) d\tilde{q} \\ &\approx I(q_+) \\ &+ x\int_0^{q_+} \frac{2[F(q_+) - F(\tilde{q})][F''(\tilde{q}) + \tilde{q}F'''(\tilde{q})] - [q_+F'(q_+) - \tilde{q}F'(\tilde{q})]F''(\tilde{q})}{2q_+[F(q_+) - F(\tilde{q})]^{3/2}} d\tilde{q} \\ &+ x^2\int_0^{q_+} \frac{1}{8q_+^2[F(q_+) - F(\tilde{q})]^{5/2}}\Bigg(3(\tilde{q}F'(\tilde{q}))^2F''(\tilde{q}) \\ &+ 2[F(q_+) - F(\tilde{q})][\tilde{q}^2F''(\tilde{q}) + 2\tilde{q}F'(\tilde{q})(F''(\tilde{q}) + \tilde{q}F'''(\tilde{q}))] \\ &+ 4[F(q_+) - F(\tilde{q})]^2[2\tilde{q}F'''(\tilde{q}) + \tilde{q}^2F''''(\tilde{q})]\Bigg) d\tilde{q} + O(x^3) \quad ,\end{aligned} \tag{A.6}$$

from which we obtain the derivative by $\partial_{q_+} I(q_+) = \partial_x I(q_+ + x)|_{x=0}$

$$\partial_{q_+} I(q_+) = \int_0^{q_+} \frac{2[F(q_+) - F(\tilde{q})][F''(\tilde{q}) + \tilde{q}F'''(\tilde{q})] - [q_+ F'(q_+) - \tilde{q}F'(\tilde{q})]F''(\tilde{q})}{2q_+[F(q_+) - F(\tilde{q})]^{3/2}} d\tilde{q} \quad , \tag{A.7}$$

and second derivative $\partial^2_{q_+} I(q_+) = \partial^2_x I(q_+ + x)|_{x=0}$

$$\begin{aligned} \partial^2_{q_+} I(q_+) = \int_0^{q_+} \frac{1}{4q_+^2[F(q_+) - F(\tilde{q})]^{5/2}} \Bigg(& 3(\tilde{q}F'(\tilde{q}))^2 F''(\tilde{q}) \\ & + 2[F(q_+) - F(\tilde{q})][\tilde{q}^2 F''(\tilde{q}) + 2\tilde{q}F'(\tilde{q})(F''(\tilde{q}) + \tilde{q}F'''(\tilde{q}))] \\ & + 4[F(q_+) - F(\tilde{q})]^2[2\tilde{q}F'''(\tilde{q}) + \tilde{q}^2 F''''(\tilde{q})] \Bigg) d\tilde{q} \quad . \end{aligned} \tag{A.8}$$

To show that $I(q_+)$ is a nondecreasing function of q_+ when conditions $1-5$ on F hold we start in the linear limit where it is straight forward to show that $\partial_{q_+} I(q_+) = 0$. When conditions $1-5$ hold the integrand in equation (A.8) is non-negative for all q, q_+. Thus $\partial^2_{q_+} I(q_+) \geq 0$ which implies that $\partial_{q_+} I(q_+) \geq 0$.

A.1.2 Tightly wound case

The equivalent expressions for $I(q_+)$ and $J(q_+)$ for a tightly wound wave are

$$I(q_+) := \frac{\pi}{\sqrt{2}} \sqrt{\frac{F'(q_+)}{q_+}} \quad , \tag{A.9}$$

$$J(q_+) := \frac{\pi}{2\sqrt{2}} q_+^{1/2} F'(q_+)^{3/2} \quad . \tag{A.10}$$

Showing that $J(q_+)$ is an increasing function of q_+, for $q_+ > 0$, is straightforward as it's derivative is strictly positive when $F'(q_+), F''(q_+) > 0$.

Showing $I(q_+)$ is an increasing function of q_+, for $q_+ > 0$, is more involved and requires assuming that $F'(q_+)$ is strictly convex on the interval $0 < q_+ < 1$ (an assumption similar, but not identical, to the assumptions made for the untwisted case). Consider $0 < q_+ < 1$ and $\tilde{q}_+ = \epsilon q_+ < q_+$, with $0 < \epsilon < 1$, then

$$I(\tilde{q}_+) = I(\epsilon q_+) = \frac{\pi}{\sqrt{2}} \sqrt{\frac{F'(\epsilon q_+)}{\epsilon q_+}} < \frac{\pi}{\sqrt{2}} \sqrt{\frac{\epsilon F'(q_+)}{\epsilon q_+}} = I(q_+), \tag{A.11}$$

with the inequality following from the convexity of $F'(q_+)$. Thus, $I(q_+)$ is an increasing function of q_+ for $0 < q_+ < 1$.

A.2 Trigonometric parametrisation for numerical calculations of untwisted modes

The Hamiltonian derived in Chapter 5 possess an apparent singularity when $q = 0$, which presents a problem for calculating untwisted modes numerically. These apparent singularities can be removed by using trigonometric parametrisation.

Starting with the Hamiltonian for the nonlinear oscillator (equation 6.22) and making a canonical transform to canonical variables (β, π_β) given by

$$q =: \sin 2\beta \quad , \pi_\beta := 2\cos 2\beta \pi_q \quad , \tag{A.12}$$

with Hamiltonian

$$H = \frac{\pi_\beta^2}{4k^2 \cos^2 2\beta (F''(\sin 2\beta))^2} + F(\sin 2\beta) \quad . \tag{A.13}$$

For $\gamma = 1$

$$F(\sin 2\beta) = -2 \ln \cos \beta, \tag{A.14}$$

$$F''(\sin 2\beta) = \frac{1}{2}(\sec^2 \beta \sec 2\beta + 2 \sec 2\beta \tan \beta \tan 2\beta) \sec 2\beta \tag{A.15}$$

and for a $\gamma = 2$ 2D disc

$$F(\sin 2\beta) = \sec 2\beta \quad , \qquad F''(\sin 2\beta) = (2 - \cos 4\beta) \sec^5 2\beta. \tag{A.16}$$

For each q_+ (and associated β_+) the nonlinear oscillator describing the φ dependence can be solved via a shooting method to find the value of k which makes the solution periodic. This gives us k as a function of q_+ from which we can obtain the integral $I(q_+)$ from

$$I(q_+) = \frac{\pi}{\sqrt{2}k} \quad . \tag{A.17}$$

$J(q_+)$ can be obtained from its relationship with $I(q_+)$:

$$J(q_+) = \int_0^{q_+} \frac{F'(q)}{2} I(q)\, dq \quad . \tag{A.18}$$

Alternatively we can directly compute the integrals I and J using the regularised forms for F and F''. This still possesses an apparent singularity at $q = q_+$ which must be dealt with separately. Typically it is sufficient to simply excise this apparent singularity by truncating the integral at a q sufficiently close to q_+.

A.3 Expressions for the Hamiltonian of a 2D disc

Let

$$I_p(q) := \frac{1}{2\pi} \int_0^{2\pi} \frac{d\theta}{(1 - q\cos\theta)^p}, \tag{A.19}$$

which can be evaluated using Legendre functions of the first kind $P_k(z)$:

$$I_p(q) = z^p P_{p-1}(z), \quad z = \frac{1}{\sqrt{1-q^2}}. \tag{A.20}$$

For integer p,

$$I_{-1}(q) = 1, \quad I_0(q) = 1, \quad I_1(q) = (1-q^2)^{-1/2}, \quad I_2(1-q^2)^{-3/2}. \tag{A.21}$$

We also have the integrals,

$$\frac{1}{2\pi} \int_0^{2\pi} \frac{\cos\theta}{(1 - q\cos\theta)^p} d\theta = \frac{I_p(q) - I_{p-1}(q)}{q}, \quad \frac{1}{2\pi} \int_0^{2\pi} \frac{\sin\theta}{(1 - q\cos\theta)^p} d\theta = 0, \tag{A.22}$$

$$\frac{1}{2\pi} \int_0^{2\pi} \frac{\cos^2\theta}{(1 - q\cos\theta)^p} d\theta = \frac{1}{q^2} \left[I_p(q) - 2I_{p-1}(q) + I_{p-2}(q) \right], \tag{A.23}$$

$$\frac{1}{2\pi} \int_0^{2\pi} \frac{\sin^2\theta}{(1 - q\cos\theta)^p} d\theta = I_p(q) - \frac{1}{q^2} \left[I_p(q) - 2I_{p-1}(q) + I_{p-2}(q) \right]. \tag{A.24}$$

Making use of these integrals the geometric part of the Hamiltonian in a 2D disc, with $\gamma > 1$, is

$$F^{(2\mathrm{D})} = \frac{1}{\gamma - 1}\left[\frac{\sqrt{1-e^2}}{1-e(e+ae_a)}\right]^{\gamma-1}[(1-x)I_{\gamma-1}(q) + xI_{\gamma-2}(q)], \tag{A.25}$$

with

$$x = \frac{e\cos\alpha}{q} = \frac{eae_a}{q^2[1-e(e+ae_a)]}. \tag{A.26}$$

For the isothermal case $\gamma = 1$ we make use of

$$\frac{1}{2\pi}\int_0^{2\pi}\ln(1-q\cos\theta)d\theta = \ln\left(\frac{\frac{1}{2}q^2}{1-\sqrt{1-q^2}}\right), \tag{A.27}$$

$$\frac{1}{2\pi}\int_0^{2\pi}\cos\theta\ln(1-q\cos\theta)d\theta = -\left(\frac{1-\sqrt{1-q^2}}{q}\right), \tag{A.28}$$

leading to

$$F^{(2\mathrm{D})} = \ln\left[\frac{2\sqrt{1-e^2}(1-\sqrt{1-q^2})}{q^2[1-e(e+ae_a)]}\right] - \frac{eae_a(1-\sqrt{1-q^2})}{q^2[1-e(e+ae_a)]}. \tag{A.29}$$

A.4 Evaluation of the viscous integrals in a 2D disc

For a $2D$ disc (or $3D$ isothermal disc), with $m = j^x$, F_{bulk} can be written as,

$$F_{\mathrm{bulk}}(q) = \frac{q^2}{2\pi}\int j^{x-1}\sin^2 E\, dE \quad . \tag{A.30}$$

Similarly for a viscosity law with $m = j^x$ (i.e. independent of the scale height $y = 0$) F_{shear} can be written as,

$$F_{\mathrm{shear}}(q) = \frac{3q}{2\pi}\int j^{x-1}\left[2\cos E\,\left(\frac{1}{4} - j\right) + q\sin^2 E\right]dE \quad , \tag{A.31}$$

Writing these in terms of the integrals I_p:

$$F_{\mathrm{bulk}}(q) = -(1-q^2)I_{1-x}(q) + 2I_{-x}(q) - I_{-x-1}(q), \tag{A.32}$$

$$F_{\mathrm{shear}}(q) = -\frac{3}{2}(1-2q^2)I_{1-x}(q) - \frac{3}{2}I_{-x}(q) + 3I_{-x-1}(q). \tag{A.33}$$

An important special case is $x = -1$, which gives the viscosity law for a (3 D) isothermal $\alpha-$disc. Then the integrals can be evaluated as

$$F_{\text{bulk}}(q) = \frac{1}{\sqrt{1-q^2}} - 1 \quad , \tag{A.34}$$

$$F_{\text{shear}}(q) = 3 - \frac{9}{2}(1-q^2)^{-1/2} + \frac{3}{2}(1-q^2)^{-3/2}. \tag{A.35}$$

Appendix B

Derivation of the pseudo-Lagrangian for viscous stress model

In this appendix we give a detailed derivation of the non-conservative terms which appear in the general pseudo-Lagrangian for a tightly wound wave discussed in Chapter 8.

From Ogilvie & Barker (2014) the eccentricity evolution caused by gradients in the stress tensor is given by

$$\begin{aligned} l\mathcal{M}\dot{\mathcal{E}}|_{\text{stress}} = \int\int [2(e^{i\phi}+\mathcal{E})\partial_\lambda \left(JR^2T^{\lambda\phi}\right) - i\lambda e^{i\phi}\partial_\lambda \left(JR_\lambda T^{\lambda\lambda}\right) - ie^{i\phi}JR^2T^{\phi\phi} \\ - \frac{JR^2}{\lambda}(e^{i\phi}+\mathcal{E}-\lambda\mathcal{E}_\lambda)T^{\lambda\phi}]d\phi dz, \end{aligned} \tag{B.1}$$

in the (λ, ϕ) orbital coordinate system. In this coordinate system the short wavelength limit approximation is given by $e \ll 1$, $q \simeq 1$ where $q^2 \approx \lambda^2\varphi_\lambda^2\left(e_\varphi^2 + e^2\varpi_\varphi^2\right)$. In this approximation we have the following approximations:

$$J = \lambda j \approx \lambda(1 - q\cos(\phi - \phi_0)) + O(e), \tag{B.2}$$

$$R \approx \lambda + O(e), \tag{B.3}$$

$$R_\lambda \approx j + O(e). \tag{B.4}$$

The viscous stress tensor is

$$T^{ij} = 2\mu S^{ij} + \left(\mu_b - \frac{2}{3}\mu\right) \nabla \cdot u \, g^{ij}, \tag{B.5}$$

where S^{ij} is the rate of strain tensor and μ the dynamic viscosity. Define a vertically average dynamic shear viscosity $\bar{\mu} := \int \mu dz$ (similarly for bulk viscosity $\bar{\mu}_b := \int \mu_b dz$). Using the forms of S^{ij} given in Ogilvie (2001) and keeping only the lowest order terms in the expansion we have vertically averaged stress tensors

$$\bar{T}^{\lambda\lambda} = 2\bar{\mu}\frac{n}{j^2}\frac{j_\phi}{j} + \left(\bar{\mu}_b - \frac{2}{3}\bar{\mu}\right) n \left(\frac{j_\phi}{j} + w\right) j^{-2} + O(e), \tag{B.6}$$

$$\bar{T}^{\lambda\phi} = 2\bar{\mu}\frac{n}{\lambda} j^{-2} \left(\frac{1}{4} - j\right) + O(e), \tag{B.7}$$

$$\bar{T}^{\phi\phi} = \left(\bar{\mu}_b - \frac{2}{3}\bar{\mu}\right) n \left(\frac{j_\phi}{j} + w\right) \lambda^{-2} + O(e), \tag{B.8}$$

$$\bar{T}^{zz} = 2\bar{\mu} n w + \left(\bar{\mu}_b - \frac{2}{3}\bar{\mu}\right) n \left(\frac{j_\phi}{j} + w\right) + O(e), \tag{B.9}$$

where we have introduced $w := \frac{\dot{h}}{h}$.

Expanding out equation B.1 in terms of the short wavelength expansion, keeping only terms of lowest order, for shear viscosity and preforming the vertical average we get

$$l\mathcal{M}\dot{\mathcal{E}}|_{\text{stress}} \approx \int \left[2e^{i\phi}\partial_\lambda\left(\lambda^3 j \bar{T}^{\lambda\phi}\right) - i\lambda e^{i\phi}\partial_\lambda\left(\lambda j^2 \bar{T}^{\lambda\lambda}\right) - ie^{i\phi} j\lambda^3 \bar{T}^{\phi\phi} - \lambda^2 j(e^{i\phi} - \tilde{\mathcal{E}}_\varphi)\bar{T}^{\lambda\phi}\right] d\phi + O(e). \tag{B.10}$$

Rearranging to separate out the terms involving total derivatives we get

$$\begin{aligned} l\mathcal{M}\dot{\mathcal{E}}|_{\text{stress}} \approx \partial_\lambda \int \lambda^2 j e^{i\phi} \left[2\lambda \bar{T}^{\lambda\phi} - ij\bar{T}^{\lambda\lambda}\right] d\phi \\ + \int \lambda j \left[ie^{i\phi}\left(j\bar{T}^{\lambda\lambda} - \lambda^2 \bar{T}^{\phi\phi}\right) - \lambda(e^{i\phi} - \tilde{\mathcal{E}}_\varphi)\bar{T}^{\lambda\phi}\right] d\phi \end{aligned} \tag{B.11}$$

Assuming that the vertically average dynamic viscosity is a function of pressure, surface density, scale height and semimajor axis only, then, in the short wavelength

limit, it has the functional form $\bar{\mu} = \bar{\mu}(\lambda, q, \cos(\phi - \phi_0))$ (and the same for μ_b) which we can use to simplify some of the orbit averages,

$$\begin{aligned} l\mathcal{M}\dot{\mathcal{E}}|_{\text{stress}} \approx \partial_\lambda \int \lambda^2 j e^{i\phi_0} \left[2\cos(\tilde{\phi})\lambda \bar{T}^{\lambda\phi} + \sin(\tilde{\phi}) j \bar{T}^{\lambda\lambda} \right] d\tilde{\phi} \\ - e^{i\phi_0} \int \lambda j \left[\sin(\tilde{\phi}) \left(j\bar{T}^{\lambda\lambda} - \lambda^2 \bar{T}^{\phi\phi} \right) + \lambda(\cos(\tilde{\phi}) - q)\bar{T}^{\lambda\phi} \right] d\tilde{\phi}, \end{aligned} \tag{B.12}$$

where we have shifted the phase of the orbit average by ϕ_0 and introduced $\tilde{\phi} := \phi - \phi_0$. We now switch to the a, E variables used in the rest of the work. At this order, $\lambda \approx a$, $\tilde{\phi} \approx E$ (strictly shifted E), $\phi_0 \approx E_0 + \varpi$ and $l\mathcal{M} \approx M_a n a^2$ and (after multiplying both sides by $i/2$) equation B.10 becomes

$$\begin{aligned} \frac{i}{2} M_a n a^2 \dot{\mathcal{E}}|_{\text{stress}} \approx \frac{\partial}{\partial a} \frac{i}{2} a^2 e^{i(E_0+\varpi)} \int j \left[2\cos E\, a\bar{T}^{\lambda\phi} + \sin E\, j\bar{T}^{\lambda\lambda} \right] dE \\ - \frac{i}{2} e^{i(E_0+\varpi)} \int aj \left[\sin E \left(j\bar{T}^{\lambda\lambda} - a^2 \bar{T}^{\phi\phi} \right) + a(\cos E - q)\bar{T}^{\lambda\phi} \right] dE. \end{aligned} \tag{B.13}$$

The left hand side comes from a variation of the Lagrangian, L_T, with

$$\frac{\delta L_T}{\delta \mathcal{E}^*} = \frac{i}{2} M_a n a^2 \dot{\mathcal{E}} \quad , \tag{B.14}$$

so we are searching for a pseudo-Lagrangian F so that varying $-F$ with respect to $\mathcal{E}^*$ yields the right hand side of Equation B.13. Noting that

$$\frac{\partial}{\partial \mathcal{E}_a^*} \frac{\text{Re}[ia\mathcal{E}_a^* \breve{\mathcal{E}}_a]}{q'} = \frac{i}{2} e^{i(\breve{E}_0 + \breve{\varpi})}, \quad \frac{\partial}{\partial \mathcal{E}^*} \frac{\text{Re}[ia\mathcal{E}^* \breve{\mathcal{E}}_a]}{q'} = \frac{i}{2} e^{i(\breve{E}_0 + \breve{\varpi})}, \tag{B.15}$$

we obtain a pseudo Lagrangian density,

$$\begin{aligned} F_a \approx a^2 \frac{\text{Re}[i\mathcal{E}_a^* \breve{\mathcal{E}}_a]}{\breve{q}} \int \breve{j} \left[2\cos E\, a\breve{\bar{T}}^{\lambda\phi} + \sin E\, \breve{j}\breve{\bar{T}}^{\lambda\lambda} \right] dE \\ + \frac{\text{Re}[i\mathcal{E}^* \breve{\mathcal{E}}_a]}{\breve{q}} \int a\breve{j} \left[\sin E \left(\breve{j}\breve{\bar{T}}^{\lambda\lambda} - a^2 \breve{\bar{T}}^{\phi\phi} \right) + a(\cos E - \breve{q})\breve{\bar{T}}^{\lambda\phi} \right] dE, \end{aligned} \tag{B.16}$$

where the breves ($\breve{\cdot}$) denote dependence on the complimentary phase space variables. We now preform the expansion of the complimentary phase variable $\breve{\varphi}$ about φ. We make use of the following expansions:

$$\frac{\mathrm{Re}[i\mathcal{E}_a^*\breve{\mathcal{E}}_a]}{\breve{q}} = -\frac{q}{a}\sin(\breve{E}_0 + \breve{\varpi} - E_0 - \varpi) \approx \frac{\breve{q}}{a}(\varphi - \breve{\varphi})(\breve{\varpi}_{\breve{\varphi}} + \breve{E}_{0,\breve{\varphi}}), \tag{B.17}$$

$$\frac{\mathrm{Re}[i\mathcal{E}^*\breve{\mathcal{E}}_a]}{\breve{q}} = -e\sin(\breve{E}_0+\breve{\varpi}-\varpi) \approx -e\sin\breve{E}_0-(\varphi-\breve{\varphi})\breve{e}_{\breve{\varphi}}\sin\breve{E}_0+(\varphi-\breve{\varphi})\breve{e}(\breve{\varpi}_{\breve{\varphi}}+\breve{E}_{0,\breve{\varphi}})\cos\breve{E}_0, \tag{B.18}$$

from which we obtain an approximation for the pseudo-Lagrangian density in the short wavelength limit,

$$\begin{aligned} F_a \approx{}& a\breve{q}(\varphi-\breve{\varphi})(\breve{\varpi}_{\breve{\varphi}}+\breve{E}_{0,\breve{\varphi}})\int\breve{j}\left[2\cos E\, a\breve{\bar{T}}^{\lambda\phi}+\sin E\,\breve{j}\breve{\bar{T}}^{\lambda\lambda}\right]dE \\ &-(\varphi-\breve{\varphi})\left[\breve{e}_{\breve{\varphi}}\sin\breve{E}_0-\breve{e}(\breve{\varpi}_{\breve{\varphi}}+\breve{E}_{0,\breve{\varphi}})\cos\breve{E}_0\right] \\ &\times\int a\breve{j}\left[\sin E\,\left(\breve{j}\breve{\bar{T}}^{\lambda\lambda}-a^2\breve{\bar{T}}^{\phi\phi}\right)+a(\cos E-\breve{q})\breve{\bar{T}}^{\lambda\phi}\right]dE, \end{aligned} \tag{B.19}$$

where we have dropped the $O(1)$ term as it doesn't contribute to the dynamics. In the tight winding limit $\breve{e}_{\breve{\varphi}} = \breve{E}_{0,\breve{\varphi}} = 0$ and $E_0 = \pi/2$, so the terms which did not originate in the total derivative vanish. Substituting the expressions for the stress tensor (Equations B.7-B.9) the pseudo-Lagrangian density becomes

$$\begin{aligned} F_a \approx{}& 2na\breve{q}(\varphi-\breve{\varphi})\breve{\varpi}_{0,\breve{\varphi}}\int\breve{\bar{\mu}}\breve{j}^{-1}\left[2\cos E\,\left(\frac{1}{4}-\breve{j}\right)+\sin E\,\breve{j}_E\right]dE \\ &+na\breve{q}(\varphi-\breve{\varphi})\breve{\varpi}_{0,\breve{\varphi}}\int\left(\breve{\bar{\mu}}_b-\frac{2}{3}\breve{\bar{\mu}}\right)\left(\frac{\breve{j}_E}{\breve{j}}+\breve{w}\right)\sin E\,dE. \end{aligned} \tag{B.20}$$

Proposing that the viscosity has the functional form $\bar{\mu}_b = \frac{\alpha_b}{2\pi}\mu_a^\circ(a)m(q,\cos E)$, $\bar{\mu} = \frac{\alpha}{2\pi}\mu_a^\circ(a)m(q,\cos E)$, appropriate for a viscosity which is a power law of pressure, surface density and scale height. Introducing $T_a^\circ = n\mu_a^\circ$ then, integrating over the disc, we can write the pseudo-Lagrangian as

$$\begin{aligned} F = &\frac{2}{3}\alpha \int a T_a^{\circ}(\varphi - \breve{\varphi}) \breve{\varpi}_{0,\breve{\varphi}} F_{\text{shear}}(\breve{q})\, da \\ &+ \left(\alpha_b - \frac{2}{3}\alpha\right) \int a T_a^{\circ}(\varphi - \breve{\varphi}) \breve{\varpi}_{0,\breve{\varphi}} F_{\text{bulk}},(\breve{q})\, da \end{aligned} \tag{B.21}$$

where we have introduced

$$F_{\text{bulk}}(q) := \frac{q}{2\pi} \int_0^{2\pi} \tau(q, \cos E) \sin^2 E\, dE, \tag{B.22}$$

$$F_{\text{shear}}(q) := \frac{3q}{2\pi} \int_0^{2\pi} m(q, \cos E) j^{-1} \left[2 \cos E \left(\frac{1}{4} - j \right) + q \sin^2 E \right] dE, \tag{B.23}$$

with

$$\tau(q, \cos E) := m(q, \cos E) \left(\frac{q}{j} + \frac{1}{h} \frac{\partial h}{\partial \cos E} \right). \tag{B.24}$$

In the $\tilde{e} = q_+$ gauge the pseudo-Lagrangian is

$$F = \frac{2}{3}\alpha \int a T_a^{\circ}(\varphi - \breve{\varphi}) F_{\text{shear}}(\breve{q})\, da + \left(\alpha_b - \frac{2}{3}\alpha\right) \int a T_a^{\circ}(\varphi - \breve{\varphi}) F_{\text{bulk}}(\breve{q})\, da \quad . \tag{B.25}$$

Interestingly when $F_{\text{bulk}}(q) > F_{\text{shear}}(q)$ the shear viscosity is anti-dissipative and we have a form of viscous overstability.