Chapter 2

Dynamics of protoplanetary discs

The evolution of protoplanetary discs proceeds, as mentioned in chapter 1, through the redistribution of angular momentum allowing accretion of disc matter onto the central star. During this process, the matter in the outer parts of the disc takes up angular momentum from the matter in the inner parts, so the latter losing angular momentum, gradually falls onto the central star (e.g., Lynden-Bell and Pringle, 1974; Pringle, 1981). This transport of mass inwards and angular momentum outwards can be provided by a variety of physical processes, such as turbulence originating from various instabilities in the disc or, in the case of relatively massive discs, by disc self-gravity. These processes play a dual role here: firstly, they drive outward angular momentum transport, which is necessary for accretion to occur, by exerting torques on the disc and, secondly, they provide a channel of conversion of gravitational energy, liberated as mass falls towards the centre, into thermal energy. The dissipated energy, in turn, can influence the observed SED. The angular momentum transport by turbulent torques (stresses) can be conveniently characterised in terms of so-called *turbulent, or 'anomalous' viscosity*[1], the concept of which in the theory of astrophysical discs was first introduced by von Weizsäcker (1948) and further developed by Shakura and Sunyaev (1973) (see also comprehensive reviews by Balbus and Hawley 1998 and Balbus 2003). In this chapter, we outline mechanisms that can provide angular momentum transport in discs and a possible nature of this anomalous/enhanced viscosity.

[1] We will show below that ordinary molecular viscosity of gas is far too small to yield typical time-scales of secular evolution of accretion discs.

To first gain insight into how viscosity, irrespective of its origin, causes mass and angular momentum redistribution in discs, following Lynden-Bell and Pringle (1974); Pringle (1981); Frank et al. (2002), let us consider the simplest model of a protoplanetary disc – a razor-thin (2D) gaseous disc rotating around a central star with considerably larger mass $M_\star$, so that disc self-gravity can be neglected (we will return to self-gravitating discs later). We adopt cylindrical coordinates (r, ϕ, z) with the central star at the origin and the disc lying in the $z = 0$ plane. The razor-thin disc approximation implies that the typical length-scale in the vertical direction – the disc scale height $H \simeq c_s/\Omega_0$, where c_s is the gas sound speed and Ω_0 is the angular velocity of disc rotation, is much smaller than the radial distance from the central star, that is, the aspect ratio $H/r \ll 1$.[2] In this case, basic hydrodynamical quantities are vertically averaged, that is, integrated in the vertical direction all over the disc height, and thus are made 2D (see section 2.3). The disc is then characterised by its surface density Σ, which is a mass per unit surface area of the disc, obtained by vertically averaging the gas density ρ. The thin disc approximation is also equivalent to the requirement that the sound speed c_s be much less than the rotation velocity $r\Omega_0(r)$, that is, the disc flow is strongly supersonic. From the latter condition, in turn, it follows that the radial gradients of pressure can be ignored, so that the disc angular velocity is given by a balance between the gravity force of the central star and centrifugal force due to rotation, resulting in the Keplerian rotation profile (Pringle, 1981; Frank et al., 2002; Lodato, 2007)

$$\Omega_0(r) = \left(\frac{GM_\star}{r^3}\right)^{1/2}.$$

In this case, a disc is said to be rotationally supported. The disc accretes, meaning that in addition to azimuthal Keplerian velocity it also possesses radial 'drift' velocity, u_r, directed towards the central star. As will be clear below, this radial velocity is much smaller than the Keplerian velocity (the ordering $u_r \ll c_s \ll r\Omega_0$ holds) and is directly related to the viscosity ν. The disc is assumed to be axisymmetric, so that all quantities are functions of only radius r and time t (z-dependence is absorbed in the vertical averaging). The height-integrated continuity equation for surface density is

$$\frac{\partial \Sigma}{\partial t} + \frac{1}{r}\frac{\partial}{\partial r}(r\Sigma u_r) = 0. \tag{2.1}$$

We also need another height-integrated equation for angular momentum conservation, which

[2] In protoplanetary discs, typically $H/r \simeq 0.05 - 0.1$ (e.g., Durisen et al., 2007), so the condition for thin disc limit is marginally satisfied. Nevertheless, many studies are limited just to this approximation, because it helps to understand basic physical processes and instabilities, which can then be generalised to thick 3D discs.

can be similarly derived. The only difference is that now we should include the viscous stress tensor, which in this special case is dominated by $r\phi$-component proportional to ν,

$$W_{r\phi} = \nu \frac{d\Omega_0}{d \ln r}.$$

As a result, we obtain

$$\frac{\partial}{\partial t}(r^2 \Sigma \Omega_0) + \frac{1}{r}\frac{\partial}{\partial r}(r^3 \Sigma \Omega_0 u_r) = \frac{1}{r}\frac{\partial}{\partial r}(r^2 \Sigma W_{r\phi}). \tag{2.2}$$

Equations (2.1-2.2) can be combined into a single equation for surface density evolution

$$\frac{\partial \Sigma}{\partial t} = \frac{3}{r}\frac{\partial}{\partial r}\left[r^{1/2}\frac{\partial}{\partial r}(\nu r^{1/2}\Sigma)\right]. \tag{2.3}$$

In deriving this equation we have made use of the fact that the rotation is stationary and Keplerian. Equation (2.3), commonly used in accretion disc modelling, is the basic equation governing the secular, or viscous evolution of the disc. This equation is a non-linear diffusion-type equation for Σ, where ν can be a complex function of local variables (surface density, radius, temperature, ionisation fraction, etc.) so it should generally be solved numerically. However, if viscosity can be expressed as some power of radius, then analytic solutions are feasible (Lynden-Bell and Pringle, 1974; Armitage, 2010) that allow us to get the general feeling of a disc's viscous evolution.

A simple inspection of the structure of Eq. (2.3) shows that the radial velocity is proportional to viscosity,

$$u_r = -\frac{3}{r^{1/2}\Sigma}\frac{\partial}{\partial r}(\nu r^{1/2}\Sigma) \sim \frac{\nu}{r}$$

and the characteristic time-scale of the surface density evolution due to viscosity is

$$t_{visc} = \frac{r^2}{\nu} \sim \frac{r}{u_r}.$$

In real discs, this secular evolution time-scale is usually much larger than the dynamical/orbital time $t_{dyn} \equiv \Omega_0^{-1}$ (see below), implying, as noted above, that the radial accretion velocity u_r is much smaller than the disc's rotational velocity.

To illustrate viscous evolution more quantitatively, we consider a simple, analytically tractable case of constant viscosity coefficient, $\nu = const$. Suppose that initially at $t = 0$ matter is concentrated in a very thin ring with radius r_0 and mass m. The surface density of such a ring has the form

$$\Sigma(r, t = 0) = \frac{m}{2\pi r_0}\delta(r - r_0).$$

An analytic solution can be constructed in terms of Bessel functions describing the time-development of this initial ring-shaped profile of surface density (e.g., Pringle, 1981; Frank

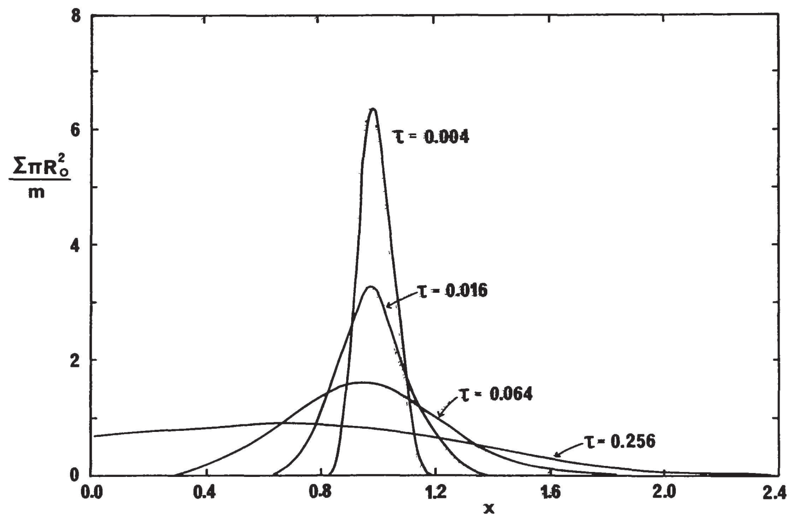

Figure 2.1: Evolution of the surface density of a Keplerian spreading ring under the action of constant viscosity ν, shown as a function of the normalised radius $x = r/r_0$, where r_0 is the initial radius of the ring, and of non-dimensional time $\tau = 12\nu t/r_0^2$. Adapted from Pringle (1981).

et al., 2002; Lodato, 2007), which we do not give here. Instead, we show it graphically in Fig. 2.1 at different times. It is clear from this figure that viscosity acts to spread the ring both inwards and outwards on the time-scale of t_{visc}. During this process, mass in the inner parts of the disc, losing angular momentum, moves inwards and eventually falls onto the central star. At the same time, matter in the outer parts moves to larger radii to take up outwardly flowing angular momentum. The net angular momentum is conserved; it is merely redistributed between different annuli. A detailed analysis of the corresponding analytical solution shows that the boundary between inwardly and outwardly propagating regions of the spreading ring expands outwards as $r_b \simeq t\nu/r_0$ (Pringle, 1981; Frank et al., 2002; Lodato, 2007). As a result, eventually (at $t \to \infty$) a singular equilibrium state is reached – almost all the matter stripped of its angular momentum is accreted at the centre, while a negligibly small amount of mass at infinitely large radii carries all the angular momentum. In another case, where viscosity has a power-law dependence on radius, an analytic self-similar solution can also be obtained that gives a qualitatively similar evolutionary picture (Lynden-Bell and Pringle, 1974; Lodato, 2007; Armitage, 2010).

The above analysis shows that viscosity is crucial for disc evolution – it is a primary means of outward transport of angular momentum. However, a major unresolved problem here concerns the physical nature of viscosity ν. It had soon been realised that standard molecular viscosity is too small to account for observed mass accretion rates and time-scales for disc secular evolution. Indeed, molecular viscosity is given by the product of random velocity of molecules, which is of the order of the sound speed c_s, and their collisional mean free path λ, $\nu = \lambda c_s$. In a typical protoplanetary disc with the number density $n \sim 10^{13}$ cm^{-3} of molecules having the cross-section $\sigma_{coll} \sim 10^{-16}$ cm^2, the mean free path is $\lambda = 1/(n\sigma_{coll}) \sim 10^3$ cm. So, the ratio of viscous to dynamical times is $t_{visc}/t_{dyn} = r^2/(H\lambda) = Re$, where $Re \equiv r^2\Omega_0/\nu$ is the Reynolds number of disc flow. If we now take $r \sim 100$ AU for the disc radial size and $H/r \sim 0.1$, we obtain $t_{visc}/t_{dyn} = Re \sim 10^{13}$. Since the dynamical time-scale for protoplanetary discs is of the order of a few years, the viscous time $t_{vis} \sim 10^{13}$ yr, during which the whole disc accretes onto the central star, turns out to be much larger than observationally inferred disc lifetimes of $10^6 - 10^7$ yr. This means that there must exist some other mechanisms capable of producing orders of magnitude larger, or anomalous, effective viscosity resulting in mass accretion rates that will be consistent with observed disc lifetimes.

Very large Reynolds numbers, $Re \sim 10^{13}$, associated with the disc flow seem to offer a clue to resolve this problem. In laboratory experiments, hydrodynamic flows usually tend to become turbulent at high enough Reynolds numbers (e.g., Taylor, 1936; Bayly et al., 1988; Richard and Zahn, 1999; Drazin and Reid, 1981; Longaretti, 2002). Based on this, it was originally thought that gas flow in discs with such huge Reynolds numbers should naturally be strongly turbulent.[3] In this case, angular momentum exchange occurs not through collisions of molecules, but due to the mixing of fluid elements in turbulent motion. The typical length-scales of such turbulent motions can be several orders of magnitude larger than the collisional mean free path of molecules and therefore transport becomes much more efficient. The question is then how such an enhanced turbulent transport can be accommodated in the framework of a viscous disc model described above, or, in other words, what effective viscosity coefficient should be used to represent the turbulence.

To understand this, we note that the total gas velocity in a turbulent Keplerian disc can be written as a sum of the background Keplerian rotation and turbulent fluctuation $\mathbf{u}_t$,

$$\mathbf{u} = r\Omega_0(r)\mathbf{e}_\phi + \mathbf{u}_t, \tag{2.4}$$

where $\mathbf{e}_\phi$ is the unit vector in the azimuthal ϕ-direction. In this case, the mean accretion velocity is $u_r = \langle u_{tr} \rangle$, where the angle brackets denote vertical and azimuthal averages. We assume that this accretion velocity is much smaller than the fluctuation velocity amplitude and hence rotational velocity, so that the ordering $\langle u_{tr} \rangle \ll u_t \ll r\Omega_0$ holds (see e.g., Balbus,

[3] We will see below that care is needed in generalising this statement to accretion disc flows, especially to Keplerian rotation profiles (see section 2.2.1).

2003). Taking decomposition (2.4) into account, the angular momentum equation (without molecular viscosity) takes the form (Balbus and Hawley, 1998; Balbus and Papaloizou, 1999; Balbus, 2003)

$$\frac{\partial}{\partial t}(r^2\Sigma\Omega_0) + \frac{1}{r}\frac{\partial}{\partial r}(r^3\Sigma\Omega_0 u_r) = -\frac{1}{r}\frac{\partial}{\partial r}(r^2\Sigma\langle u_{tr}u_{t\phi}\rangle).$$

Comparing this equation to Eq. (2.2), we see that the averaged turbulent stress $\langle u_{tr}u_{t\phi}\rangle$ plays a similar role as the viscous stress $W_{r\phi}$. Consequently, we can define the effective turbulent viscosity as

$$\nu_t = \langle u_{tr}u_{t\phi}\rangle \left|\frac{d\Omega_0}{d\ \ln\ r}\right|^{-1} \sim \lambda_t u_t,$$

where $\lambda_t \sim u_t/\Omega_0$ is the characteristic length-scale of the turbulent motions. This expression resembles that for molecular viscosity. However, since turbulence is generally a very complex phenomenon, we do not know how to determine exactly the characteristic length-scales, λ_t, and velocities, u_t, associated with a turbulent flow without a proper knowledge of the underlying physical mechanisms causing the transition to turbulence in discs. Nevertheless, we can use physical arguments to put some constraints on these two quantities. Firstly, the typical size of the largest turbulent eddies cannot exceed the disc thickness, $\lambda_t \lesssim H$, otherwise turbulence would not be isotropic and transport would not be local, i.e., the characterisation of turbulent transport in terms of viscosity coefficient would not be possible. Secondly, it is unlikely that turbulence in discs is very supersonic, because in this case strong shocks would appear resulting in a high dissipation rate of turbulent motions. As a result, turbulent velocity would be damped to subsonic values, $u_t \lesssim c_s$. Based on these two basic arguments, in their seminal paper, Shakura and Sunyaev (1973) came up with the following parametrisation of turbulent viscosity:

$$\nu_t = \alpha c_s H,$$

where, because of the above limits on the size and velocity of turbulent motions, the non-dimensional parameter α is less than unity. This formulation, also mentioned above, is known as the 'α-prescription' for turbulent viscosity. In fact, α contains in itself all the uncertainties regarding the onset mechanism and properties of turbulence in discs, that is, the source of anomalous viscosity. Doing such a scaling, we just replace one unknown parameter ν_t with the other unknown one, α. But the usefulness of this approach is that now we measure the turbulent stress, which determines the angular momentum transport rate, in units of the local pressure and we know that in these units, if turbulence is local, it should not exceed unity. In other words, the characterisation of angular momentum transport in terms of the

α parameter is an inherently local description valid for thin ($H \ll r$) discs.[4] Although, as numerical simulations show, α may display both temporal and spatial variations in the disc (e.g., Papaloizou and Nelson, 2003), in analytical modelling it is often assumed to be constant. This actually gives reasonably good results as far as disc secular evolution is concerned. Observations indicate that $\alpha \sim 0.01$ for protoplanetary discs (e.g., Hartmann et al., 1998; Hueso and Guillot, 2005; Andrews et al., 2009), which corresponds to disc evolutionary time-scales of the order of observed disc lifetimes of a few Myr. Thus, the α parameter is a key quantity as it can be accessed observationally and can provide a check on the theory of turbulence and anomalous transport in discs.

It is crucial to have a proper understanding of disc turbulence and its onset mechanisms, as apart from driving the accretion process, it can also have a profound effect on other processes occurring in discs. This, in turn, allows us to determine α in different parts of the disc, critically examine the validity of the constant-α assumption and also make comparisons with observations. Great progress has been made in recent years in identifying the mechanisms of transition to turbulence in astrophysical discs (see e.g., Balbus, 2003). As in the case of hydrodynamic turbulence in shear flows, turbulence in discs (that in fact represent a special case of shear flow as discs almost always rotate differentially) should continuously extract energy from the disc flow at the largest scales ($\sim H$) and transfer it to motions at smaller and smaller scales until eventual dissipation on the smallest viscous scales. The extraction of energy at the largest scales, in turn, is possible owing to various types of instabilities in the disc. Turbulence should also provide 'right' outward transport of angular momentum for accretion. Thus, the starting point in tackling the disc turbulence problem is to explore what types of instabilities can develop in accretion discs. The Keplerian differential rotation plays a special role here because, as observations indicate (e.g., Dutrey, 2000; Simon et al., 2000), in most cases protoplanetary discs do rotate with Keplerian or nearly Keplerian velocities unless they are strongly self-gravitating. Therefore, research in this direction is mainly focused on investigating the stability of the Keplerian rotation profile, its subsequent (possible) transition to turbulence and the properties (e.g., angular momentum transport rate) of the resulting turbulent state itself when other physical factors specific to discs (magnetic fields, self-gravity, radial and vertical stratification, etc.) are involved. However, we should note in this respect that not all types of disc instabilities in the non-linear regime necessarily lead to turbulence in the classical sense, but can result in some kind of more or-

[4] Note that there are also other non-local, long-range angular momentum transport mechanisms, for example, gravitational instability described in section 2.2.3 below.

ganised, quasi-steady, saturated non-linear states (e.g., coherent vortices, quasi-steady spiral structure, convective rolls) that can also act to maintain angular momentum transport (in such cases, λ_t and u_t simply correspond to the characteristic sizes and velocities of these organised non-linear motions). Mechanisms/instabilities causing enhanced angular momentum transport in discs, broadly speaking, can be divided into two basic categories: magnetohydrodynamic (MHD) mechanisms, primarily giving rise to magnetorotational instability (MRI), and non-magnetic mechanisms, such as a purely hydrodynamical (i.e., driven by shear of disc flow velocity) turbulence, baroclinic/Rossby wave instability and coherent vortices, spiral density waves, convection and gravitational instability if the disc is self-gravitating. Below we describe/review each of these mechanisms separately and then put them into the context of the main thesis work.[5] However, we will not expand much on the MRI and put more emphasis on non-magnetic means of transport, since the latter is the central subject of study in the present thesis. Besides, MHD instabilities (mostly the MRI) in accretion discs have been the focus of numerous studies since the 1990s (see reviews by Balbus and Hawley, 1998; Balbus, 2003) and today are accepted as the most likely cause of turbulence and enhanced angular momentum transport in many accretion discs, but not necessarily in protoplanetary discs. There can be certain conditions in the protoplanetary disc under which the MRI cannot operate and that is where one has to appeal to non-magnetic means of transport.

In order for the MRI to develop, the disc should necessarily have regions where gas is sufficiently ionised to couple effectively with magnetic field, so that the latter can affect the gas dynamics (Blaes and Balbus, 1994; Gammie, 1996; Sano et al., 2000; Fromang et al., 2002; Sano and Stone, 2002; Salmeron and Wardle, 2003; Desch, 2004). The minimum ionisation fraction (i.e., the ratio of electron concentration, which are the main carriers of current, to that of neutrals) required for the onset of the MRI in protoplanetary discs is very low, typically 10^{-13} (e.g., Gammie, 1996). Above this minimum ionisation, magnetic diffusion is small and cannot suppress the MRI, while at lower ionisation fractions the MRI is quenched due to dissipative effects. Disc ionisation can be provided by thermal collisions of molecules, X-rays and cosmic rays. Thermal ionisation of alkali metals (potassium) can occur at $T \gtrsim 10^3$ K in the inner parts ($\lesssim 0.1$ AU) of the disc (Umebayashi, 1983) and keep the ionisation level there above the threshold value, so that these inner regions are maintained magnetically active. Ionisation due to cosmic rays or due to illumination from the central star's X-ray

[5] We do not consider here external factors, such as the effects of a companion star, which can also produce anomalous angular momentum transport by tidally inducing shocks of spiral density waves in the disc (e.g., Rozyczka and Spruit, 1993; Lin and Papaloizou, 1993).

and UV radiation dominate at larger radii, where the disc is cooler and thermal ionisation is ineffective. These non-thermal sources, entering the disc from the top and bottom, can ionise gas only in surface layers with column densities up to a maximum value of 100 g cm^{-2} above which they cannot penetrate deeper due to attenuation (Umebayashi and Nakano, 1981). As a result, a large magnetically inactive region – so-called 'dead zone' – forms near the disc midplane outside $\sim$ 0.1 AU, where the ionisation level is much smaller than that required for magnetic coupling. This dead zone with almost no MHD turbulent activity thus appears sandwiched between two magnetically active, MRI-turbulent surface layers and can extend to radii as large as 30 AU (layered accretion disc model, see e.g., Gammie, 1996; Fromang et al., 2002; Fleming and Stone, 2003). Beyond this radius, the disc surface density drops below the maximum value, so that cosmic rays can propagate all along the height and fully ionise the gas, thereby making it turbulent again and the layered structure disappears. Analogous almost neutral regions can also exist in cataclysmic variable (CV) discs (Gammie and Menou, 1998) and also in the outer parts of cool active galactic nuclei (AGN) discs (Menou and Quataert, 2001). The absence of sufficient magnetic coupling, and therefore of MHD turbulence, within such large portions of the disc necessitates a search for other alternative non-magnetic mechanisms of angular momentum transport. However, it should be noted in this regard that it is also possible that the turbulence in the magnetically active surface layers induces small angular momentum transport in the dead zone by driving velocity fluctuations (hydrodynamical waves) there and/or through diffusive penetration of magnetic fields to the midplane (Fleming and Stone, 2003; Turner et al., 2007; Oishi et al., 2007; Turner and Sano, 2008; Oishi and Mac Low, 2009).

2.1 Magnetorotational instability

The MRI is a linear instability occurring in differentially rotating, ionised gaseous discs threaded by magnetic field. The remarkable feature of this instability is that it grows on a characteristic time-scale of the order of the orbital period and even a very weak imposed magnetic field is sufficient for its activation. In other words, the MRI persists and retains its key properties in the limit of vanishing magnetic field, $B \to 0$, or equivalently, for weak ionisation, but, as mentioned above, the ionisation fraction of the disc must still be larger than a certain minimum value, which can be very small. We note in this respect that the stability properties of weakly ionised discs are qualitatively different from those of non-magnetised discs.