Lecture 2: Protoplanetary discs

1 Disc formation and the angular momentum problem

Obervations of both the solar system and exo-planetary systems suggest that planets form in discs around young stars, so in order to understand how planets form it is first necessary to consider how stars form. Naively we might assume that stars are able to form by simple gravitational collapse of gas clouds, but we must also consider the effects of rotation. "Cores" in star-forming molecular clouds are observed to have angular velocities of order $\Omega_c \sim 10^{-14}$–10^{-13}s^{-1}, and we can thus compute the angular momentum of a core by appealing to the Jeans length

$$R_J \simeq \frac{c_s}{\sqrt{G\rho}} \tag{1}$$

and Jeans mass

$$M_J \simeq \frac{c_s^3}{G^{3/2}\rho^{1/2}} \,. \tag{2}$$

(Here we have neglected order-of-unity constants for clarity.) Molecular clouds have temperatures $T \simeq 10\text{K}$, yielding sound speeds $c_s \simeq 0.2\text{km/s}$. We therefore see that forming a star of solar mass by Jeans collapse requires densities $\rho \gtrsim 10^{-19}\text{g/cm}^3$, and requires that material fall inwards from distances of order $R_J \sim 0.1\text{pc}$. The specific angular momentum of the collapasing core is thus

$$j_c \simeq \Omega_c R_J^2 \simeq 10^{21} - 10^{22}\text{cm}^2/\text{s} \,. \tag{3}$$

By contrast, the break-up velocity of a star (the maximum velocity at which it can rotate) can be computed by equating centrifugal acceleration with gravity thus

$$\Omega_b^2 r_* = \frac{GM_*}{r_*^2} \tag{4}$$

$$\Omega_b = \sqrt{\frac{GM_*}{r_*^3}} \,. \tag{5}$$

A star like the Sun therefore has a break-up velocity $\Omega_b \sim 10^{-3}\text{s}^{-1}$ (corresponding to a few hundred km/s), and a break-up specific angular momentum (assuming solid-body rotation) of

$$j_b \simeq \Omega_b r_*^2 \simeq 10^{18} - 10^{19}\text{cm}^2/\text{s} \,. \tag{6}$$

Thus $j_b \ll j_c$, and most stars in fact rotate well below break-up. We see therefore that young stars have much lower angular momenta than the gas clouds from which they form, and how this angular momentum is lost is the so-called "angular momentum problem" of star formation.

In Lecture 1 we discussed the various observations which suggest that the Solar System formed from a single rotating disc, and we can similarly appeal to discs as a solution to the angular momentum problem of star formation. If we assume that discs around young stars are in Keplerian rotation, we can estimate their typical size from the angular momentum of the system. In a Keplerian orbit the specific angular momentum of a mass orbiting a star of mass M_* at radius R is

$$j_K = \sqrt{GM_* R} \tag{7}$$

and if we set $j_K = j_c$ we find that protostellar discs around solar-mass stars should have typical sizes of

$$R = \frac{j_c^2}{GM_*} \sim 10^3 - 10^4\text{AU} \,. \tag{8}$$

Star formation is in fact much more dynamic than the process we have described here, and interactions between protostars and their surroundings can redistribute some of the excess angular momentum. Nevertheless, when we observe young stars we see resolved discs with sizes of 100–1000AU, and numerical simulations of star formation typically produce discs of similar sizes. Disc accretion is thus crucial to the star formation process, and planet-forming discs are an inevitable consequence of star formation.

2 Observations of protoplanetary discs

The first protoplanetary discs were observed in the 1980s, and since then we have amassed a huge body of research into their structure and evolution. The subject of protoplanetary disc observations is large enough to form an entire lecture course by itself, so here we merely summarise the most relevant points. The bulk of the disc mass is gas (mostly molecular hydrogen), but the trace dust component dominates the opacity. The dust is therefore crucial for disc thermodynamics (and for planet formation), and despite representing only around 1% of the total mass the dust is also much easier to observe than the gas.

Young, low-mass stars are traditionally classified by the shape of their infrared (IR) spectral energy distribution (SED), which is typically measured through broad-band photometry (allowing large numbers of objects to be observed simultaneously). The SED classification scheme, originally proposed by Lada (1987) and subsequently updated several times, is based on the IR spectral index

$$\alpha_{\rm IR} = \frac{d\log(\lambda F_\lambda)}{d\log\lambda} \,. \tag{9}$$

In practice $\alpha_{\rm IR}$ is usually measured between two fixed wavelengths: early work used 2.2μm (K-band) and 14μm, but more recent studies generally use one of the *Spitzer* bands for the long-wavelength point (usually 24μm). Light from the central (proto-)star is absorbed by circumstellar dust and re-emitted at longer wavelengths, so a "redder" SED, resulting from more material in the circumstellar environment, is thought to be indicative of an earlier evolutionary phase. The modern classification scheme is as follows:

- **Class 0:** SED peaks in the far-IR or sub-mm, with no measurable flux being emitted in the near- or mid-IR. These objects are typically interpreted as proto-stars which are still in the collapse phase.
- **Class I:** SED peaks in the mid- or far-IR, with a rising SED slope ($\alpha_{\rm IR} \gtrsim 0$) in the near-IR. These objects are inferred to be embedded protostellar discs with substantial circumstellar envelopes.
- **Class II:** SED is declining in the near- and mid-IR, but shows significant excess emission over the stellar photosphere ($-1.5 \lesssim \alpha_{\rm IR} \lesssim 0$). These are optically-visible pre-main-sequence stars with a surrounding disc, but little or no remaining envelope.
- **Class III:** SED is effectively that of a stellar photosphere, with $\alpha_{\rm IR} \sim -1.5$. These are pre-main-sequence stars which have lost their discs.

Additional sub-classes exist (notably "flat spectrum" sources between Classes I & II, and "transitional discs" between Classes II & III), but these objects are relatively rare. The relationship between the SED classification scheme and the physical state of the YSO is now supported by a large body of evidence, and although there is not a perfect one-to-one correspondence between SED class and evolution the term "Class X" is commonly used to refer to the physical state of the object. For the purposes of this course we are most interested in Class II & III objects, but we will also discuss Class I objects. Note also that Classes II & III correspond almost perfectly with an older classification scheme, based on the strength of optical emission lines: objects with Hα equivalent widths $\gtrsim 10$Å are referred to as Classical T Tauri stars (CTTs), while objects with Hα equivalent widths $\lesssim 10$Å are referred to as Weak-lined T Tauri stars (WTTs). We now understand that the observed emission lines are primarily due to accretion on to the stellar surface: almost all disc-bearing Class II sources are CTTs, while disc-less Class III are usually WTTs.

Observations of dust continuum emission are the most straightforward means of observing protoplanetary discs, and we now have large statistical samples across a broad range of wavelengths. Near- and mid-IR emission comes from the inner disc, at radii $\lesssim 1$AU, where the disc is optically thick, while longer wavelength emission primarily comes from the colder outer disc. At mm wavelengths the emission is optically thin, and (sub-)mm observations allow us to measure the total (dust) mass in the disc. Estimating disc masses in this manner is subject to significant uncertainties (particularly in the dust-to-gas ratio), but standard assumptions derive disc masses which range from $\gtrsim 0.1\mathrm{M}_{\odot}$to $\lesssim 0.001\mathrm{M}_{\odot}$. The median disc mass (for Class II sources / CTTs) derived in this manner is approximately 1% of the stellar mass (i.e., $\sim$10$\mathrm{M}_{\mathrm{Jup}}$ for solar-mass stars).

Observing the gaseous component of the disc is more difficult, due to the fact that H_2 has no bright emission lines (because hydrogen is a homo-nuclear molecule with no permanent dipole moment). In general we are limited to detecting H_2 emission from the hot disc surface (a thin layer, strongly irradiated by the star), and/or detecting emission from (trace) heavier elements and molecules (notably CO). Our primary means of detecting the bulk of the disc gas is observing signatures of accretion on to the stellar surface: typical accretion rates[1] for Class II sources lie in the range 10^{-7}–$10^{-9}\mathrm{M}_{\odot}\mathrm{yr}^{-1}$. The fact that discs are observed to accrete at such rates tells us that they must evolve on $\sim$Myr time-scales (as $M_d/\dot{M} \sim 10^6$yr), and we discuss accretion and disc evolution in more detail below.

These observations also tell us that essentially all young stars form with discs. The disc fraction in the youngest ($\lesssim 1$Myr) star clusters is close to 100%, but the number of discs drops rapidly with time and in $\sim$ 10Myr-old clusters very few discs remain. This result holds across the full range of disc signatures, and again implies that typical protoplanetary disc lifetimes are a few Myr. This in turn sets a strict limit on the process(es) of planet formation: after $\sim$10Myr T Tauri stars have insufficient gas to form even Neptune-mass planets, so giant planets must be able to form within the lifetimes of protoplanetary discs, on $\sim$Myr time-scales.

3 Protoplanetary disc structure

The equation of motion for an inviscid, non-magnetised fluid is

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v}.\nabla)\,\mathbf{v} = -\frac{1}{\rho}\nabla P - \nabla\Phi\,, \tag{10}$$

where $\mathbf{v}$ is the fluid velocity, ρ the density, P the pressure and Φ the gravitational potential. We consider a disc as a stationary, axisymemetric flow around a central gravitating mass, and therefore work in cylindrical co-ordinates (R, z, ϕ)[2]. In the limit of a low-mass disc the potential is simply $\Phi = -GM_*/r$, and the radial component of the equation of motion is an expression of centrifugal balance

$$\frac{v_\phi^2}{R} = \frac{1}{\rho}\frac{dP}{dR} + \frac{GM_*}{R^2}\,. \tag{11}$$

If we neglect the gas pressure, we find that the azimuthal velocity v_ϕ is simply the Keplerian velocity $v_{\mathrm{K}} = \sqrt{GM_*/R}$.

In general, the pressure in a disc decreases outwards (as both the surface density and temperature are typically decreasing functions of radius). The dP/dR term in Equation 11 is therefore

[1]Note, however, that the accretion is highly variable, especially during the earlier evolutionary phases.

[2]Note that I use lower-case r for spherical radius and upper-case R for cylindrical radius.

negative, and the orbital velocity of the gas is sub-Keplerian. How sub-Keplerian the gas is depends on radial temperature and density structure of the disc, but for typical parameters we find that $v_\phi/v_{\rm K} \simeq 0.995$. This is negligible for gas dynamics, but has important consequences for solid bodies in the disc (as we will see in Lecture 3).

The vertical component of Equation 10 is a statement of hydrostatic balance

$$\frac{1}{\rho}\frac{dP}{dz} = -\frac{d\Phi}{dz} = \frac{d}{dz}\left(\frac{GM_*}{r}\right) = \frac{d}{dz}\left(\frac{GM_*}{(R^2+z^2)^{1/2}}\right) , \tag{12}$$

with pressure supporting the disc against gravity. In the limit of a thin disc $(z \ll R)^3$, this reduces to

$$\frac{1}{\rho}\frac{dP}{dz} = -\frac{GM_* z}{R^3} = -\Omega_{\rm K}^2 z , \tag{13}$$

where $\Omega_{\rm K} = v_{\rm K}/R$ is the Keplerian orbital frequency. If we then assume that the disc is vertically isothermal, the equation of state is $P = c_s^2\rho$ and we have

$$\frac{1}{\rho}\frac{d\rho}{dz} = \frac{d\log\rho}{dz} = -\frac{z}{H^2} , \tag{14}$$

where we have defined the disc scale-height $H = c_s/\Omega$. We can integrate to find the vertical structure

$$\rho(z) = \rho_0 \exp\left(-\frac{z^2}{2H^2}\right) . \tag{15}$$

A vertically isothermal disc therefore has a Gaussian density profile, with scale-height H (so 68% of the disc mass lies within $\pm H$ of the midplane). The midplane density ρ_0 is related to the local surface density Σ by the normalisation condition

$$\rho_0 = \frac{\Sigma}{\sqrt{2\pi}H} . \tag{16}$$

The disc structure $H(R)$ is primarily determined by the radial temperature profile of the disc $T(R)$. The low accretion rates in protoplanetary discs mean that accretion (viscous) heating is usually negligible, and the disc's heating is instead dominated by irradiation from the central star[4]. For a razor-thin disc, we can compute the radial temperature profile by considering the flux absorbed by a patch of the disc at radius r. The star has radius r_* and effective temperature T_*, so if we assume a constant surface brightness we have $I_* = (1/\pi)\sigma_{\rm SB}T_*^4$ (where $\sigma_{\rm SB}$ is the Stefan-Boltzmann constant). The flux F absorbed by the disc is simply the integral of the brightness over the fraction of the stellar surface "seen" by the disc, so

$$F = \int I_* \sin\theta\cos\phi d\Omega , \tag{17}$$

where $d\Omega = \sin\theta d\theta d\phi$ is the (infinitesimal) solid angle element. If we consider only one hemisphere of the star (and therefore only the flux absorbed by one surface of the disc), we see that the limits on the integral are $-\pi/2 \le \phi \le \pi/2$ and $0 \le \theta \le \sin^{-1}(r_*/R)$. Substituting, we find that

$$F = I_* \int_{-\pi/2}^{\pi/2} \cos\phi d\phi \int_0^{\sin^{-1}(r_*/R)} \sin^2\theta d\theta . \tag{18}$$

[3]In deriving Equations 11 & 12 we have formally assumed that the radial gas velocity $v_{\rm R} \ll c_s \ll v_\phi$. The first inequality requires that any radial gas flow (i.e., accretion) be very sub-sonic, while the second inequality is essentially a re-statement of the thin disc approximation (and also implies that the pressure term in Equation 11 is small.) Observations of disc aspect ratios and accretion rates confirm that both of these approximations are justified for protoplanetary discs.

[4]Accretion heating dominates at small radii and high accretion rates, $\gtrsim 10^{-7}{\rm M}_\odot{\rm yr}^{-1}$. These conditions are only satisfied in the inner disc ($\lesssim$ 1AU) and in the early stages of disc evolution.

This integral is ugly but basically straightforward [use Pythagoras to evaluate $\cos(\sin^{-1} x)$], and we find that

$$F = I_* \left[\sin^{-1}(r_*/R) - \left(\frac{r_*}{R}\right) \sqrt{1 - \left(\frac{r_*}{R}\right)^2} \right] . \tag{19}$$

In thermodynamic equilibrium this absorbed flux is equal to the flux radiated by the disc. If the disc has local temperature $T(r)$ then $F = \sigma_{\rm SB} T^4$ and

$$\left(\frac{T(R)}{T_*}\right)^4 = \frac{1}{\pi} \left[\sin^{-1}(r_*/R) - \left(\frac{r_*}{R}\right) \sqrt{1 - \left(\frac{r_*}{R}\right)^2} \right] . \tag{20}$$

This expression is not particularly instructive, but if we expand as a Taylor series[5] in the far-field limit ($r_*/R \ll 1$), we find that

$$T(R) \propto R^{-3/4} . \tag{21}$$

This is the radial temperature profile for a flat, optically-thick reprocessing disc[6], and by assuming that $c_s^2 \propto T$ we find that $H/R \propto R^{1/8}$. This solution is therefore not self-consistent, as the disc is not flat (the aspect ratio H/R is an increasing function of radius). Self-consistent solutions must take account of the relationship between temperature and disc thickness, and also the fact that the disc sub-tends a larger solid angle at larger radii. (This is most readily done by defining a flaring angle $\alpha = d/dR(H/R)$, and modifying Equation 17 accordingly.)

We can use the Planck function and integrate Equation 19 to find the disc SED (we find $\lambda F_\lambda \propto \lambda^{-4/3}$), but it was recognised more than 30 years ago that a flat reprocessing disc produces substantially less IR emission than is typically observed in Class II discs[7]. We now understand that protoplanetary discs are flared, with H/R increasing significantly with radius: the disc is relatively thicker at larger radii, and thus intercepts (and re-emits) a larger fraction of the stellar flux than a thin disc (leading to a larger IR excess). For a vertically isothermal disc the self-consistent solution is $T \propto R^{-1/2}$, which gives $H/R \propto R^{5/4}$: the temperature profile is shallower than that of a flat disc, and the disc flares substantially with radius. Modern disc models relax the assumption of vertical isothermality (see, e.g., the "two-layer" model of Chiang & Goldreich 1997), and include additional complications such as accretion heating and realistic opacities. Detailed study of disc structure remains an active area of research.

4 Protoplanetary disc evolution

4.1 The viscous disc

In order to determine how a disc evolves, we first consider mass and angular momentum conservation in a thin annulus. We assume that the disc is axisymmetric, with surface density $\Sigma(R)$. Accretion is a radial flow of gas, and we denote the radial velocity as $v_{\rm R}(R)$[8]; by convention, positive $v_{\rm R}$ is in the outward direction (accretion therefore has $v_{\rm R} < 0$). We then consider an annulus at radius R with thickness ΔR. The rate of mass flow through the inner edge of the annulus is

$$\dot{M}_{\rm inner} = 2\pi R \Sigma(R) v_{\rm R}(R) , \tag{22}$$

and the rate of mass flow through the outer edge of the annulus is

$$\dot{M}_{\rm outer} = 2\pi (R + \Delta R) \Sigma(R + \Delta R) v_{\rm R}(R + \Delta R) . \tag{23}$$

[5] The relevant Taylor series expansion is $\sin^{-1} x = x + x^3/6 + x^5/40 \ldots$

[6] Coincidentally, this power-law scaling is the same as that found in a self-luminous accretion disc.

[7] The flat disc solution for λF_λ has $\alpha_{\rm IR} = -4/3$, which is close to the Class II/III boundary and much steeper ("bluer") than is typical for Class II objects.

[8] Note that, as before, we assume that $v_{\rm R} \ll c_s \ll v_\phi$.

The difference between these two quantities is the rate of change of mass in the annulus, so

$$2\pi R \Delta R \frac{\partial \Sigma}{\partial t} = \dot{M}_{\rm inner} - \dot{M}_{\rm outer} \,, \tag{24}$$

and we can substitute to find

$$R\frac{\partial \Sigma}{\partial t} = -\frac{(R+\Delta R)\Sigma(R+\Delta R)v_{\rm R}(R+\Delta R) - R\Sigma(R)v_{\rm R}(R)}{\Delta R} \,. \tag{25}$$

We can then take the limit $\Delta R \rightarrow 0$ to find

$$R\frac{\partial \Sigma}{\partial t} + \frac{\partial}{\partial R}\left(R\Sigma v_{\rm R}\right) = 0 \,, \tag{26}$$

which is the equation of mass continuity for a thin disc.

If the orbital frequency is $\Omega(R)$ then the angular momentum per unit area is $R^2\Omega\Sigma$, and a similar analysis yields a corresponding equation for the conservation of angular momentum. However, in this case we must consider the effects of torques on the annulus, so the equation for angular momentum conservation becomes

$$R\frac{\partial}{\partial t}\left(R^2\Omega\Sigma\right) + \frac{\partial}{\partial R}\left(R^2\Omega R\Sigma v_{\rm R}\right) = \frac{1}{2\pi}\frac{\partial G}{\partial R} \,. \tag{27}$$

Here $G(R)$ is the torque: it appears as a radial derivative, as we are interested only in the differential torque across the annulus. (If the torque is constant across the annulus then $\partial G/\partial R = 0$ and there is no net change in the angular momentum.)

At this point we can make little further progress without making some assumptions about the origin of the torque G. The simplest assumption is to assume that the torques are due to an ordinary fluid viscosity: in that case, the shearing nature of a Keplerian disc will result in viscous torques between adjacent annuli (as they have different azimuthal velocities). The viscous torque G is therefore

$$G = 2\pi R . R\nu\Sigma\frac{d\Omega}{dR}.R \tag{28}$$

Here ν is the kinematic viscosity. The second term is the viscous force per unit length, the first term comes from integrating around the annulus, and the final factor of R is the lever arm of the torque. The fact that $G \propto \frac{d\Omega}{dR}$ reflects the fact that the viscous torque is only non-zero if the disc has differential rotation. (If Ω is constant, the disc rotates as a solid body and there are no viscous torques between adjacent annuli.) We substitute this expression for G into Equation 27 to find

$$\frac{\partial}{\partial t}\left(R^2\Sigma\Omega\right) + \frac{1}{R}\frac{\partial}{\partial R}\left(R^3\Sigma v_{\rm R}\Omega\right) = \frac{1}{R}\frac{\partial}{\partial R}\left(R^3\nu\Sigma\frac{d\Omega}{dR}\right) \,, \tag{29}$$

which is the equation for angular momentum conservation in a viscous accretion disc. We can then combine Equations 26 & 29 to eliminate the radial velocity $v_{\rm R}$, and after some algebra we find

$$\frac{\partial \Sigma}{\partial t} = -\frac{1}{R}\frac{\partial}{\partial R}\left[\frac{1}{\frac{\partial}{\partial R}(R^2\Omega)}\frac{\partial}{\partial R}\left(R^3\nu\Sigma\frac{d\Omega}{dR}\right)\right] \,. \tag{30}$$

This is the general equation governing the evolution of a viscous accretion disc with an arbitrary rotation profile $\Omega(R)$. If we further assume that the disc is in Keplerian rotation we can substitute $\Omega = \sqrt{GM_*/R^3}$ and find

$$\frac{\partial \Sigma}{\partial t} = \frac{3}{R}\frac{\partial}{\partial R}\left[R^{1/2}\frac{\partial}{\partial R}\left(\nu\Sigma R^{1/2}\right)\right] \,. \tag{31}$$

4.2 Viscous disc solutions

The simplest solution to Equation 31 is a steady-state solution (which implicitly requires a source term as the outer boundary condition). We can find the steady-state solution by considering Equation 29. We set the $\partial/\partial t$ terms to zero and note that in a steady state the accretion rate $\dot{M} = -2\pi R \Sigma v_{\rm R}$ is constant. We then integrate with respect to R, and if we further assume that there is no torque exerted at the inner boundary (i.e., $d\Omega/dR = 0$ at $R = R_{\rm in}$), we find

$$-R^2 \Omega \dot{M} + R_{\rm in}^2 \Omega_{\rm in} \dot{M} = 2\pi R^3 \nu \Sigma \frac{d\Omega}{dR} \, . \tag{32}$$

Substituting for $\Omega = \sqrt{GM_*/R^3}$ gives

$$\Sigma(R) = \frac{\dot{M}}{3\pi\nu(R)} \left(1 - \sqrt{\frac{R_{\rm in}}{R}} \right) \, . \tag{33}$$

Away from the boundary we have $3\pi\nu\Sigma = \dot{M}$, so we see that the viscosity $\nu(R)$ is critical in determining the mass distribution in the disc. Further insight, however, requires at least some understanding of the viscosity.

The form of Equation 31 equates a single time derivative of Σ to a double radial derivative, which we recognise as a diffusion equation. Equation 31 can be re-cast using different variables to make the diffusive form more explicit[9]: the diffusion constant is proportional to the viscosity ν. The characteristic diffusion time-scale (neglecting factors of order unity) is

$$t_\nu \simeq \frac{R^2}{\nu} \, . \tag{34}$$

This is more commonly referred to as the viscous time-scale, and is the time required for viscous accretion to alter the surface density (locally) by a factor of order unity.

Equation 31 is in general non-linear, though it becomes linear is the viscosity ν is independent of Σ. Even in this case, however, analytic solutions are rare; most problem require numerical solution. The simplest analytic solution is that of the "spreading ring". This case assumes a constant viscosity ν, and an initial surface density that is a δ-function at $R = 1$. The analytic solution is a modified Bessel function (see, e.g., Fig.1 in Pringle 1981), and the diffusive nature of the disc equations is readily apparent. Angular momentum is transported outwards by viscosity, allowing mass to accrete (though angular momentum conservation prevents all of the mass from ever being accreted). In the limit $t \to \infty$ an infinitesimal fraction ϵ of the mass "spreads" to large radii (carrying all of the initial angular momentum), allowing $(1 - \epsilon)$ of the mass to accreted on to the central point mass.

More relevant for the study of protoplanetary discs is the similarity solution derived by Lynden-Bell & Pringle (1974). They showed that is the viscosity is a time-independent power-law $\nu(R) \propto R^\gamma$, then solutions of the form

$$\Sigma(R, t) = \frac{M_d (2 - \gamma)}{2\pi R_0^2 r^\gamma} \tau^{\frac{-(5/2-\gamma)}{2-\gamma}} \exp\left(-\frac{r^{2-\gamma}}{\tau} \right) \, , \tag{35}$$

satisfy Equation 31. Here M_d is the initial disc mass, and the scale radius R_0 sets the initial disc size. The dimensionless radius $r = R/R_0$, and the dimensionless time

$$\tau = \frac{t}{t_\nu} + 1 \, . \tag{36}$$

[9]This is left as an exercise – the relevant substitutions are $X = 2R^{1/2}$ and $S = R^{1/2}\Sigma$.

The viscous scaling time t_ν is given by

$$t_\nu = \frac{R_0^2}{3(2-\gamma)^2\nu_0} \,. \tag{37}$$

where $\nu_0 = \nu(R_0)$. $\Sigma(R,t)$ therefore has a self-similar form: a power-law $\Sigma \propto R^{-\gamma}$, exponentially truncated at large radii. The surface density declines as a power-law in time (as mass is accreted), and the disc expands to conserve angular momentum. As before, the asymptotic solution is that almost all of the initial disc mass is accreted, while all of the angular momentum is carried to large radii by an infinitesimal fraction of the mass.

4.3 The α-prescription

Thus far we have studiously avoided discussion of the viscosity, but we must now stop and consider the physical origin of the angular momentum transport (and the observed accretion). It is immediately obvious that ν cannot be a simple fluid viscosity: consideration of inter-molecular forces shows that the typical viscous time-scale at $\sim$AU radii in protoplanetary discs is $\gtrsim 10^{13}$yr! We therefore require some other source of viscosity in order to explain the observed accretion rates in young stars (and other accreting systems).

The orbital velocity in a thin disc is highly supersonic [as $v_\phi = (H/R)^{-1}c_s$], and if molecular forces are the only source of fluid viscosity then the orbital flow has extremely high Reynolds number ($\gtrsim 10^{10}$). Consequently, it is commonly (though not necessarily correctly) argued that the disc will be turbulent, and we can therefore appeal to turbulence as the source of angular momentum transport in accretion discs. This is the essence of the α-disc model, proposed by Shakura & Sunyaev (1973). They argued that the characteristic length-scale of the turbulent eddies is $\sim H$, and that the characteristic speed of the turbulence is $\sim c_s$. On dimensional grounds they therefore proposed that

$$\nu = \alpha c_s H \,, \tag{38}$$

where $\alpha \leq 1$ is a dimensionless parameter that quantifies the strength of the turbulence.

This formalism offers a number of advantages, primarily that the accretion disc equations become closed if α is known. Without understanding the source of the turbulence, we can estimate values of α from observations of disc accretion: discs around compact objects (such as dwarf novae and cataclysmic variables) generally require $\alpha \gtrsim 0.1$, while observations of protoplanetary discs point towards lower values, $\alpha \sim 0.01$. The α-prescription also provides some justification for the power-law viscosity used in the Lynden-Bell & Pringle similarity solution: for constant α, typical solutions for flaring discs result in power-law indices $\gamma \simeq 1.0$–1.5.

5 Turbulence as a mechanism for angular momentum transport

Equation 31 can also be derived directly from the Navier-Stokes equation (see, e.g., Lodato 2008). This approach is rather more involved and perhaps a little less intuitive, but has the advantage of identifying the "viscosity" ν with the viscous stress tensor. In the case of a classical shear viscosity the only non-zero component of the stress tensor is $\sigma_{R,\phi}$, and the vertically integrated stress tensor can be related to our previous analysis thus

$$T_{R,\phi} = \nu R \Sigma \frac{d\Omega}{dR} \,. \tag{39}$$

We have already seen that an ordinary fluid viscosity cannot provide sufficient stress to drive the observed accretion, but this equation points towards an alternative. If we consider a non-viscous fluid we can separate the equations into mean and fluctuating parts (e.g., Balbus & Hawley 1998; Lodato 2008), and in this case we identify $T_{R,\phi}$ with correlated fluctuations in the flow. If we then compare Equation 39 with the α-prescription, we can relate the α parameter to properties of the turbulent flow. In the general case that the fluid is both magnetized and self-gravitating, we can relate α to the various different stresses that arise in the turbulent flow:

$$\alpha = \left\langle \frac{\delta v_R \delta v_\phi}{c_s^2} - \frac{B_r B_\phi}{4\pi c_s^2} - \frac{g_r g_\phi}{4\pi G \rho c_s^2} \right\rangle . \tag{40}$$

Here the angled brackets represent a (density-weighted) average over time for the various fluctuating fields in the flow. The first term is the hydrodynamic (Reynolds) stress, the second term the magnetic (Maxwell) stress, and the third term the gravitational stress, and we therefore see that turbulence in the disc has the potential to drive angular momentum transport.

It is important to stress at this point that this "turbulent viscosity" is not really a viscosity at all. In the macroscopic limit the turbulence may behave like a fluid viscosity, but the viscous approximation does not hold when we consider length scales comparable to that of the turbulence (i.e., $\lesssim H$). This subtlety can often be glossed over when considering the global evolution of an accretion disc[10], but it is of critical importance for the formation of planets. Microscopic dust particles only "see" the small-scale structure in the disc, and so we must consider the turbulent nature of the disc when trying to understand how planets form.

Note also that we have now assigned a rather different meaning to the α parameter. Classical accretion disc theory uses the α-viscosity to make predictions from the disc equations, and α is usually regarded as a (free) input parameter to models. By contrast, in Equation 40 α is a dimensionless measure of how efficiently turbulence transports angular momentum. In a sense, this approach reverses the direction of causality implied by the α-prescription: α is a property of the local disc conditions, rather than a parameter which determines these conditions. In practice this means that α can be measured directly in numerical simulations of turbulent accretion discs.

5.1 The magnetorotational instability

Equation 40 shows us that turbulence in accretion discs can in principle by hydrodynamic, magnetohydrodynamic (MHD), or gravitational in origin. Self-gravity is generally only significant in very massive discs (roughly if $M_{\rm d}/M_* \gtrsim H/R$), and can lead either to angular momentum transport or fragmentation and gravitational collapse. However, in the Class II phase protoplanetary discs are not massive enough to be gravitationally unstable, so we defer discussion of gravitational instability to Lecture 4. In the case of purely hydrodynamic turbulence the stability criterion is simply the Rayleigh criterion, and instability requires that

$$\frac{d}{dR}\left(R^2\Omega\right) < 0 . \tag{41}$$

The term inside the brackets is the specific angular momentum, which *increases* with radius in Keplerian discs. Keplerian accretion discs are therefore linearly stable to hydrodynamic perturbations, and purely hydrodynamic turbulence is not likely to be a significant source of angular momentum transport[11].

By contrast, in the case of a magnetised fluid we recover a different stability criterion (see, e.g., Chapter 12 of Pringle & King). In this case the disc is unstable to axisymmetric perturbations if

$$\frac{d\Omega^2}{d\ln R} < 0 . \tag{42}$$

[10] Note also that a fluid viscosity provides both angular momentum transport and local heating. It is not obvious that a turbulent "viscosity" behaves in the same manner, and although MHD turbulence often can be approximated as an α-viscosity the right-hand side of Equation 40 does not necessarily imply local dissipation of energy.

[11] Some hydrodynamic instabilities, such as the baroclinic instability, can transport angular momentum in Keplerian discs, but these generally require specific (and often unusual) thermodynamic conditions to be satisfied.

This condition is satisfied in Keplerian discs, so we expect protoplanetary discs to be linearly unstable to MHD instabilities. The most promising candidate for driving angular momentum transport is the magnetorotational instability (MRI). The full derivation of the MRI is beyond the scope of this course, and in the lecture we will instead discuss a qualitative physical interpretation of the instability. The interested reader is pointed towards Balbus & Hawley (1991, 1998) & Balbus (2011) for more detailed discussion.

In a disc threaded by a vertical magnetic field the linear growth of the MRI manifests itself as so-called "channel flow" solutions, in which vertical layers of the disc move alternately inwards or outwards. In the vertically averaged sense this does not lead to net transport of angular momentum; instead we must follow the non-linear development of the instability. This requires numerical calculations, and numerical simulations of MHD turbulence in discs remains an active area of research. Current models suggest that the transport is primarily driven by magnetic stresses (i.e., the second term on the RHS of Equation 40 dominates), and that in the ideal-MHD limit MRI-driven turbulence transports angular momentum with efficiencies $\alpha \sim 10^{-3}$–10^{-2}.

However, it is unlikely that ideal MHD applies in most protoplanetary discs: all three non-ideal terms[12] are likely to become important in different regions of the disc, and break the ideal coupling between the magnetic field and the fluid. The level of disc ionization is particularly important: a fluid must be at least partially ionized in order to couple to a magnetic field, and in an insufficiently ionized disc Ohmic dissipation acts to suppress the MRI. Again, detailed analysis is beyond the scope of this course, but it can be shown (e.g., Gammie 1996; Armitage 2010) that in order for the MRI to operate in protoplanetary discs the ionization fraction in the disc must be $\gtrsim 10^{-12}$. In most astrophysical discs thermal ionization is sufficient to allow the MRI to operate, but despite the critical ionization fraction being very small indeed[13] this condition is frequently not satisfied in protoplanetary discs. Where the disc is insufficiently ionized we expect a so-called "dead zone", where the MRI does not operate. Close to the star thermal ionization is more than sufficient, and at large radii (where the surface density is low) ionization by cosmic rays provides an ample source of free electrons (though ambipolar diffusion may suppress the MRI here). At intermediate ($\sim$AU) radii, however, the disc midplane is likely to be MRI-dead, and in this region we therefore expect accretion to proceed through a partially-ionized surface layer (e.g., Gammie 1996; Armitage et al. 2001; Zhu et al. 2009).

However, modern MRI simulations suggest that even this "layered accretion" picture is overly simplistic, and there is now broad agreement the MRI is likely to be suppressed by either Ohmic dissipation or ambipolar diffusion over large regions of protoplanetary discs. The details remain uncertain (they depend crucially on non-ideal MHD and complex thermal/ionization physics, and are very difficult to model accurately), but the general trend in recent years has been towards an increasingly pessimistic view of the MRI as a mechanism for driving protoplanetary disc accretion (except very close to the star). In non-ideal MHD calculations in a local geometry, the combination of ambipolar diffusion and a net vertical B-field invariably results in a "magneto-centrifugal wind" being launched from the disc surface layers. Unlike thermal winds (e.g., from photoevaporation), magnetised winds can exert torques on the material that remains in the disc, and recent simulations suggest that such winds may play a major, and possibly even dominant, role in the angular momentum evolution of protoplanetary discs (e.g., Bai et al. 2016; see also the discussion in Turner et al. 2014). Global simulations remain challenging, however, and our understanding of how protoplanetary discs accrete remains an evolving area of research.

[12] The three non-ideal MHD effects are Ohmic resistivity (electron-neutral collisions), ambipolar diffusion (electron-neutral drift) and the Hall effect (electron-ion drift), and all three act to suppress the MRI. Current models (see, e.g., Turner et al. 2014) suggest that ambipolar diffusion dominates in the outer disc ($\gg$10AU) while Ohmic dissipation dominates close to the star ($\lesssim$1AU), with both the Hall and ambipolar terms playing a role at intermediate radii.)

[13] For typical parameters, this corresponds to an election density of $n_e \sim 0.1$–10cm^{-3} at the disc midplane!

Further Reading

In addition to the main list of references given on the course home-page, the following papers are particularly relevant to this lecture:

Pringle, *Accretion discs in astrophysics*, 1981, ARA&A, 19, 137.
Lodato, *Classical disc physics*, 2008, NewAR, 52, 21.
Armitage, *Dynamics of protoplanetary disks*, 2011, ARA&A, 49. 195.
Armitage, *Physics Processes in Protoplanetary Discs*, Saas-Fee lectures, 2015 (`arXiv:1509.06382`).
Williams & Cieza, *Protoplanetary Disks and Their Evolution*, 2011, ARA&A, 49, 67.
Balbus & Hawley, *A powerful local shear instability in weakly magnetized disks*, 1991, ApJ, 376, p214 (Part I); p223 (Part II).
Balbus & Hawley, *Instability, turbulence, and enhanced transport in accretion disks*, 1998, Rev.Mod. Phys, 70, 1.
Dullemond et al., *Models of the Structure and Evolution of Protoplanetary Disks*, 2007, Protostars & Planets V, p555.
Natta et al., *Dust in Protoplanetary Disks: Properties and Evolution*, 2007, Protostars & Planets V, p767.
Gammie, *Layered Accretion in T Tauri Disks*, 1996, ApJ, 457, 335.
Balbus, *Magnetohydrodynamics of Protostellar Disks*, 2011, in "Physical Processes in Circumstellar Disks around Young Stars". (`arXiv:0906.0854`)
Armitage et al., *Episodic accretion in magnetically layered protoplanetary discs*, 2001, MNRAS, 324, 705.
Bai et al., *Magneto-thermal Disk Winds from Protoplanetary Disks*, 2016, ApJ, 818, 152
Zhu et al., *Two-dimensional Simulations of FU Orionis Disk Outbursts*, 2009, ApJ, 701, 620.
Chiang & Goldreich, *Spectral Energy Distributions of T Tauri Stars with Passive Circumstellar Disks*, 1997, ApJ, 490, 368.
Lada, *Star formation - From OB associations to protostars*, 1987, IAU Symposium 115, p1.

Finally, the following chapters in *Protostars & Planets VI* provide up-to-date (and more detailed) reviews of many of the issues discussed here:

Turner et al., *Transport and accretion in planet-forming disks*, `arXiv:1401.7306`.
Alexander et al., *The dispersal of protoplanetary disks*, `arXiv:1311.1819`.
Testi et al., *Dust evolution in protoplanetary disks*, `arXiv:1402.1354`.
Dutrey et al., *Physical and chemical structure of planet-forming disks probed by millimeter observations and modelling*, `arXiv:1402.3503`.