
Magnetohydrodynamics of Protostellar Disks

Steven A. Balbus

Laboratoire de Radio Astronomie, École Normale Supérieure, 24 rue Lhomond,
75231 Paris EDEX 05, France balbus@ens.fr

The magnetohydrodynamical behavior (MHD) of accretion disks is reviewed. A detailed presentation of the fundamental MHD equations appropriate for protostellar disks is given. The combination of a weak (subthermal) magnetic field and Keplerian rotation is unstable to the magnetorotational instability (MRI), if the degree of ionization in the disk is sufficiently high. The MRI produces enhanced angular momentum and leads to a breakdown of laminar flow into turbulence. If the turbulent energy is dissipated locally, standard “ α ” modeling should give a reasonable estimate for the disk structure. Because away from the central star the ionization fraction of protostellar disks is small, they are generally not in the regime of near perfect conductivity. Nonideal MHD effects are important. Of these, Ohmic dissipation and Hall electromotive forces are the most important. The presence of dust is also critical, as small interstellar scale grains absorb free charges that are needed for good magnetic coupling. On scales of AU’s there may be a region near the disk midplane that is magnetically decoupled, a so-called *dead zone*. But the growth and settling of the grains as time evolves reduces their efficiency to absorb charge. With ionization provided by coronal X-rays from the central star (and possibly also cosmic rays), protostellar disks may be sufficiently magnetized throughout most of their lives to be MRI active, especially away from the disk midplane.

1 Introduction

An understanding of the magnetohydrodynamics (MHD) of protostellar disks is crucial for the theoretical development of these objects. There is no getting around the fact that the subject is decidedly inelegant. In principle, we are only interested in solving the equation $\mathbf{F} = m\mathbf{a}$, but the range of topics that bears on this problem is truly daunting. Molecular chemistry, dust grain physico-chemistry, photo-ionization physics, aerosol theory, and non-ideal MHD are all key players in this game. If the results of the theorists’

efforts had been little or no progress, there would have been no shortage of excuses. (“More work is needed.”) But in fact there has been substantial progress in understanding important issues, more perhaps than some practitioners may even realize. A sort of consensus is beginning to take shape of the gross properties of protostellar disks that are dictated by the demands of MHD physics. In this review, I will try to put forward as strong a case as I can for what I somewhat boldly regard as the canonical protostellar disk model, while at the same time try not to gloss over what are genuine uncertainties or difficulties.

Protostellar disks are a class of accretion disks, and one area where there certainly *has* been a great deal of progress in recent years is in the development of accretion disk theory. The realization that a combination of differential rotation and a weak magnetic field is profoundly unstable and produces turbulence has given the subject a foundation on which to build (Balbus & Hawley 1991, 1998). Indeed, the primary reason for studying MHD processes in protostellar disks in detail is that once these systems are no longer self-gravitating, magnetic fields profoundly influence their dynamical behavior. It is not possible to understand angular momentum transport in low mass (“T-auri”) disks without focusing on magnetic fields.

Magnetic fields provide a conduit for free energy to flow from the differential rotation to the disk itself, producing fluctuations, turbulent heating, and quite possibly the directly observed outflows. Most importantly, we shall see that MHD turbulence produces well-correlated fluctuations of the radial and azimuthal components of the magnetic field and fluctuating gas velocity. It is precisely these correlations that result in a substantial outward transport of angular momentum, allowing the disk gas to spiral inward and accrete onto the central star.

It may also be possible in principle for winds to remove angular momentum from the disk without completely disrupting it. (See Koenigl, this volume.) The difficulty that needs to be overcome for this mechanism is that the outflow must occur over the whole face of the disk, not just from the innermost regions from where jets are launched: material near the inner disk edge has already lost almost all of its angular momentum. Instead, the outer disk material, torqued by the magnetic field, must continuously slip relative to the field, or else the field tends to become very centrally concentrated (Lubow, Papaloizou, & Pringle 1994). How this might work without generating internal disk turbulence that itself transports angular momentum remains to be fully understood.

This understanding that magnetic fields play a crucial role as the source of disk turbulence developed more than thirty years after the basic instability (now known as the magnetorotational instability or MRI) first appeared in the literature (Velikhov 1959). But even if the importance of the instability had been immediately grasped, without the computational power that became available only post 1990, it is doubtful that the impact would have been the

same. There is no substitute for being able to visualize on your desktop the development of a linear instability into full-fledged MHD turbulence.

The MRI depends upon the presence of electrical currents to do its job, and this means that the disk gas must be at least partially ionized. “Partially ionized” in practice could mean even a minute electron fraction. Tiny traces of electrons can magnetize the disk, an effect we will quantify in §8. It is because even a wisp of a magnetic field and the merest trace of electrons take on such dynamical significance that protostellar disk dynamics depends so heavily on protostellar disk chemistry.

Young protostellar disks need sodium, potassium and other trace “vitamins” to make them vigorous and active, because these alkalis are important in regulating the delicate ionization balance in the gas. It is ultimately the ionization level that determines the resistivity that determines whether the magnetic field is well-coupled to the gas or not. This is not an issue in accretion disks around compact objects, for which even a small fraction of the turbulent energy dissipated would be enough to ensure that the gas is thermally ionized. Protostellar disks, on the other hand, are big (so the free energy of differential rotation is small), dusty (so grains absorb free electrons), and cold (so thermal ionization is unimportant except near the star). The overarching question of protostellar MHD research is to understand under what conditions the abundance of free electrons level drops below the level needed to ensure good magnetic coupling. The magnetic coupling leads in turn to instability, turbulence, and enhanced transport, the essence of disk dynamics. Much of this review will be devoted directly to the question of disk ionization balance. For readers wishing additional astronomical background material, the text of Lee Hartmann (1998) is an excellent choice. Balbus & Hawley (2000) and Stone et al. (2000) are earlier reviews of MHD transport processes in protostellar disks. More general disk reviews include Pringle (1981; a classic review but pre-MRI), Papaloizou & Lin (1995), Lin & Papaloizou (1996), Balbus & Hawley (1998), and Balbus (2003). In addition, the *Protostars and Planets* series published by the University of Arizona Press series contains a useful historical record of the development of the subject of protostellar disks; the latest volume is *Protostars and Planets V* (Reipurth, Jewitt, & Keil 2007).

2 On the Need for MHD

The modern view of protostellar disks is heavily influenced by accretion disk formalism. Lin & Papaloizou (1980) is the first pioneering study to bring to bear accretion disk formalism to the study of protostellar disks, though it is now thought that MHD turbulence, rather than convective turbulence, is the key to protostellar disk dynamics. To understand the central role of a magnetic field in our understanding of protostellar disks, it is of interest to review where we stand with respect to our knowledge of the onset of turbulence in rotating flows.

The principle problem for classical disk theory (Shakura & Sunyaev 1973) was to discover the origin of the very large stresses that were needed to transport angular momentum, a process then (and still often now) referred to as “anomalous viscosity.” The approach of Shakura & Sunyaev was to make a virtue of necessity, arguing that the small microscopic viscosity implied a very large Reynolds number (Re)¹, and a large Reynolds number meant that shear turbulence was present. This conclusion was sustained by the belief that the laboratory experiments showing a nonlinear breakdown of Cartesian shear layers at values of Re in excess of $\sim 10^3$ would naturally carry over to Keplerian disks, where its value was considerably higher. The fact that the inertial force associated with the Coriolis effect (not present in planar flow) was well in excess of the destabilizing shear does not seem to have been viewed as troublesome to proponents of this view.

The classical laboratory method for investigating the stability of rotating flows is a Couette apparatus. In such a device, the space between two coaxial cylinders is filled with a liquid, almost always ordinary water. The two cylinders rotate at different angular velocities, let us say Ω_{in} and Ω_{out} . The gap between the cylinders is of order a centimeter, and a stable rotation profile will be attained in a matter of minutes by viscous diffusion, even if Re is very large. (An *unstable* rotation profile will of course never be found as such; the flow will remain permanently disturbed.) By choosing Ω_{in} and Ω_{out} appropriately, a section of a Keplerian disk can be accurately mimicked, and the effects of Coriolis forces can be studied.

According to the classic text of Landau & Lifschitz (1959), a Couette rotation profile that is found in the laboratory to be stable “does not actually mean, however, that the flow actually remains steady no matter how large $[\text{Re}]$ becomes.” The monograph by Zel’dovich, Ruzmaikin, & Sokoloff (1983) is even more explicit on the question, unambiguously stating that laboratory experiments had already shown that Keplerian rotation profiles were nonlinearly unstable at sufficiently large Re .

These are stunning claims. At the time these statements were written, there was certainly no credible laboratory evidence that Keplerian rotation profiles were nonlinearly unstable. The first explicit claim that Keplerian flow was unstable based on laboratory evidence seems to be the unpublished result of Richard (2001). However, a later experiment by Ji et al. (2006) suggests that the earlier finding of instability was probably due to spurious effects arising from the interaction between the flow and the endcaps of the experiments (Ekman layers²). When care is taken to minimize this interaction, Keplerian flow is found to be linearly and nonlinearly stable at values of Re up to

¹ The Reynolds number Re is the ratio of a characteristic flow velocity V times a characteristic length L divided by the kinematic viscosity, ν , $\text{Re} = VL/\nu$.

² Ekman layers are narrow fluid layers between a solid boundary and a rotating flow. When the bulk of the flow is not at rest in the frame of the boundary, the fluid rotation changes rapidly in the Ekman layer and can produce a secondary circulation pattern invading the entire flow.

2×10^6 . Serious students of disk theory would do well to develop a deep skepticism of “large Re means disks are turbulent” arguments, which sadly are still promulgated in contemporary textbooks.

Why is the Coriolis force so harmful to the maintenance of turbulence? Moffatt, Kida, & Ohkitani (1994) have referred to stretched vortices as the “sinews of turbulence.” Turbulence is maintained in shear flows by vortices that are ensnared along the axis of strain, coupling the free energy of the shear directly to the internal vortex motion. Two neighboring fluid elements in a vortex are rapidly pulled apart as its circulation rises. This is possible only if Coriolis forces are absent. Despite the presence of shear, Coriolis forces induce epicyclic, oscillatory motions that do not allow anything resembling continuous vortex stretching. Without this, turbulence cannot be maintained. This argument breaks down if the strain rate much exceeds the formal epicyclic oscillation period, even the latter is quite well defined. Such flows, in fact, are found to be nonlinearly unstable in numerical simulations. But for Keplerian flow, the epicyclic frequency is just the local Ω of the disk, and it exerts a strongly stabilizing influence.

Figure (1) illustrates this point. Locally, Cartesian shear flow and Keplerian differential rotation appear to be similar. But perturbed fluid elements behave very differently in the two systems. Viewed from a frame moving with the undisturbed flow, a displaced fluid element in shear flow approximately follows an unbounded parabolic trajectory A, while a displaced fluid element in a Keplerian disk approximately follows a bound epicycle, B. A small perturbing vortex would be stretched continuously in shear flow, tapping into the free energy source needed to maintain turbulence. By contrast, the embedded vortex would merely oscillate in a disk. It is this difference, due entirely to the presence of the rotational Coriolis force, that is ultimately responsible for the hydrodynamical stability of Keplerian disks, whereas shear flow can be nonlinearly unstable.

The title of the Ji et al. (2006) paper is “Hydrodynamic turbulence cannot transport angular momentum effectively in astrophysical disks,” and that seems as good a one-sentence summary on the topic as any; their experiment is probably the final word on nonlinear Keplerian hydrodynamical shear instability. It is gratifying, therefore, that all of the difficulties encountered by Coriolis stabilization vanish when magnetic fields are taken into account. This is due to the fact that the magnetic field introduces new modes of response (shear waves) that are much less prone to Coriolis stabilization. One of these waves, the so-called “slow mode,” turns into a local instability when differential rotation is present with a weak magnetic field. This is the magnetorotational instability. To understand this process in more detail, it is best to begin with the fundamentals of MHD.

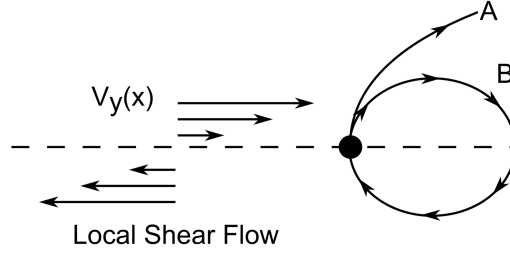


Fig. 1. The stabilization mechanism in a rotating disk versus Cartesian shear flow. In shear flow, displacements are unbounded, as indicated by path A. In a disk, epicyclic motions keeps the displacement tightly bound. See text for further details.

3 Fundamentals

In this section, a detailed derivation of the fundamental MHD equations is presented. The discussion will be more technical here than in most of the rest of the paper, but it is very important to see how the basic governing equations of the subject arise, and much of this material is not so easy to find outside of specialized treatments. I hope the reader will have the patience to read carefully through this section.

A protostellar disk is a gas of neutral particles (predominantly H_2 molecules), electrons, ions (the most important of which will generally be K^+ and Na^+), and dust grains. We defer to §8 a discussion of the complications due to the presence of dust grains and consider here the gas dynamical equations for mixture of neutrals, ions and atoms.

Each species (denoted by subscript s) is separately conserved, and obeys the mass conservation equation

$$\frac{\partial \rho_s}{\partial t} + \nabla \cdot (\rho_s \mathbf{v}_s) = 0 \quad (1)$$

where ρ_s is the mass density for species s and \mathbf{v}_s is the velocity. The flow quantities for the dominant neutral species will henceforth be presented without subscripts.

The dynamical equation for the neutral particles is

$$\rho \frac{\partial \mathbf{v}}{\partial t} + \rho(\mathbf{v} \cdot \nabla) \mathbf{v} = -\nabla P - \rho \nabla \Phi - \mathbf{p}_{nI} - \mathbf{p}_{ne} \quad (2)$$

where P is the pressure of the neutrals, Φ the gravitational potential and \mathbf{p}_{nI} (\mathbf{p}_{ne}) is the momentum exchange rate between the neutrals and the ions (electrons).

Let us examine these last two important terms in more detail. \mathbf{p}_{nI} takes the form

$$\mathbf{p}_{nI} = n\mu_{nI}(\mathbf{v} - \mathbf{v}_I)\nu_{nI} \quad (3)$$

where n is the number density of neutrals, and μ_{nI} is the reduced mass of an ion and neutral particle,

$$\mu_{nI} \equiv \frac{m_I m_n}{m_I + m_n}, \quad (4)$$

m_I and m_n being the ion and neutral mass respectively. ν_{nI} is the collision frequency of a neutral with a population of ions,

$$\nu_{nI} = n_I \langle \sigma_{nI} w_{nI} \rangle. \quad (5)$$

In equation (5), n_I is the number density of ions, σ_{nI} is the effective cross section for neutral-ion collisions, and w_{nI} is the relative velocity between a neutral particle and an ion. The angle brackets represent an average over a Maxwellian distribution function for the relative velocity. (The mass appearing in this Maxwellian will of course be the reduced mass μ_{nI} .) For neutral-ion scattering, we may take the cross section σ_{nI} to be approximately geometrical, which means that the quantity in angle brackets will be proportional to $\mu_{nI}^{-1/2}$. The order of the subscripts has no particular significance for the cross section σ_{nI} reduced mass μ_{nI} or relative velocity w_{nI} . But ν_{In} differs from ν_{nI} : the former is proportional to the neutral density n , not the ion density n_I .

Putting these definitions together gives

$$\mathbf{p}_{nI} = nn_I \mu_{nI} \langle \sigma_{nI} w_{nI} \rangle (\mathbf{v} - \mathbf{v}_I) \quad (6)$$

In accordance with Newton's third law, this is symmetric with respect to the interchange $n \leftrightarrow I$, except for a change in sign, $\mathbf{p}_{nI} = -\mathbf{p}_{In}$. All of these considerations hold, of course, for electron-neutral scattering as well. Explicitly, we have

$$\mathbf{p}_{ne} = nn_e \mu_{ne} \langle \sigma_{ne} w_{ne} \rangle (\mathbf{v} - \mathbf{v}_e) \simeq nn_e m_e \langle \sigma_{ne} w_{ne} \rangle (\mathbf{v} - \mathbf{v}_e). \quad (7)$$

The gas will be locally neutral, so that $n_e = Zn_i$ where Z is the number of ionizations per ion particle. In a weakly ionized gas, $Z = 1$. The reduced mass

μ_{ne} may safely be set equal to the electron mass m_e . We will use the following expressions for the collision rates (Draine, Roberge, & Dalgarno 1983)³:

$$\langle \sigma_{nI} w_{nI} \rangle = 1.9 \times 10^{-9} \text{ cm}^3 \text{ s}^{-1} \quad (8)$$

$$\langle \sigma_{ne} w_{ne} \rangle = 10^{-15} (128kT/9\pi m_e)^{1/2} = 8.3 \times 10^{-10} T^{1/2} \text{ cm}^3 \text{ s}^{-1} \quad (9)$$

The electron-neutral collision rate is essentially the ion geometric cross section times an electron thermal velocity. (The peculiar factor of $(128/9\pi)^{1/2}$ comes from the way the electron velocity is projected, when averaged over the electron Maxwellian distribution function, along the direction of the mean ion flow.) But the ion-neutral collision rate is temperature independent, much more beholden to long range induced dipole interactions, and significantly enhanced relative to a geometrical cross section assumption. Even if the ion-neutral rate were determined only by a geometrical cross section, $|\mathbf{p}_{nI}|$ would exceed $|\mathbf{p}_{ne}|$ by a factor of order $(m_e/\mu_{nI})^{1/2}$. In fact, the dipole enhancement of the ion-neutral cross section makes this factor yet larger.

In the astrophysical literature, it is common to write the ion-neutral momentum coupling in the form

$$\mathbf{p}_{In} = \rho \rho_I \gamma (\mathbf{v}_I - \mathbf{v}), \quad (10)$$

where γ is the so-called *drag coefficient*,

$$\gamma \equiv \frac{\langle \sigma_{nI} w_{nI} \rangle}{m_I + m_n} \quad (11)$$

and we will use this notation from here on. Numerically, $\gamma = 3 \times 10^{13} \text{ cm}^3 \text{ s}^{-1} \text{ g}^{-1}$ for astrophysical mixtures (Draine, Roberge, & Dalgarno 1983).

The dynamical equations for the ions and electrons are

$$\rho_I \frac{\partial \mathbf{v}_I}{\partial t} + \rho_I \mathbf{v}_I \cdot \nabla \mathbf{v}_I = -\nabla P_I - \rho_I \nabla \Phi + Z e n_I \left(\mathbf{E} + \frac{\mathbf{v}_I}{c} \times \mathbf{B} \right) - \mathbf{p}_{In} \quad (12)$$

and

$$\rho_e \frac{\partial \mathbf{v}_e}{\partial t} + \rho_e \mathbf{v}_e \cdot \nabla \mathbf{v}_e = -\nabla P_e - \rho_e \nabla \Phi - e n_e \left(\mathbf{E} + \frac{\mathbf{v}_e}{c} \times \mathbf{B} \right) - \mathbf{p}_{en}, \quad (13)$$

respectively. Throughout this paper, e will denote the *positive* charge of a proton, 4.803×10^{-10} esu. For a weakly ionized gas, the Lorentz force and collisional terms dominate in each of the latter two equations. Comparison of the magnetic and inertial forces, for example, shows that the latter are smaller than the former by the ratio of the proton or electron gyroperiod to a macroscopic flow crossing time. Thus, to an excellent degree of approximation,

$$Z e n_I \left(\mathbf{E} + \frac{\mathbf{v}_I}{c} \times \mathbf{B} \right) - \mathbf{p}_{In} = 0, \quad (14)$$

³ A recent calculation of H-H⁺ scattering by Glassgold, Krstić, & Schultz (2005) may imply a slightly higher value for the neutral-ion collision rate.

and

$$-en_e \left(\mathbf{E} + \frac{\mathbf{v}_e}{c} \times \mathbf{B} \right) - \mathbf{p}_{en} = 0. \quad (15)$$

The sum of these two equations gives

$$\frac{\mathbf{J}}{c} \times \mathbf{B} = \mathbf{p}_{In} + \mathbf{p}_{en} \quad (16)$$

where charge neutrality $n_e = Zn_I$ has been used, and we have introduced the current density

$$\mathbf{J} \equiv en_e(\mathbf{v}_I - \mathbf{v}_e). \quad (17)$$

The equation for the neutrals becomes

$$\rho \frac{\partial \mathbf{v}}{\partial t} + \rho \mathbf{v} \cdot \nabla \mathbf{v} = -\nabla P - \rho \nabla \Phi + \frac{\mathbf{J}}{c} \times \mathbf{B} \quad (18)$$

Due to collisional coupling, the neutrals are subject to the magnetic Lorentz force just as though they were a gas of charged particles. It is not the magnetic force *per se* that changes in a neutral gas. As we shall presently see, it is the inductive properties of the gas.

Let us return to the force balance equations for the electrons:

$$-en_e \left(\mathbf{E} + \frac{\mathbf{v}_e}{c} \times \mathbf{B} \right) - \mathbf{p}_{en} = 0. \quad (19)$$

After division by $-en_e$, this may be expanded to

$$\mathbf{E} + \frac{1}{c} [\mathbf{v} + (\mathbf{v}_e - \mathbf{v}_I) + (\mathbf{v}_I - \mathbf{v})] \times \mathbf{B} + \frac{m_e \nu_{en}}{e} [(\mathbf{v}_e - \mathbf{v}_I) + (\mathbf{v}_I - \mathbf{v})] = 0, \quad (20)$$

where we have introduced the collision frequency of an electron in a population of neutrals:

$$\nu_{en} = n \langle \sigma_{ne} w_{ne} \rangle. \quad (21)$$

We have written the electron velocity \mathbf{v}_e in terms of the dominant neutral velocity \mathbf{v} and the key physical velocity differences of our problem. It has already been noted that in equation (16), \mathbf{p}_{en} is small compared with \mathbf{p}_{In} , provided that the velocity difference $|\mathbf{v}_e - \mathbf{v}|$ is not much larger than $|\mathbf{v}_I - \mathbf{v}|$. In fact, when used in (20), the \mathbf{p}_{en} term in equation (16) may quite generally be dropped:

$$\frac{\mathbf{J}}{c} \times \mathbf{B} \simeq \mathbf{p}_{In}. \quad (22)$$

It may also be shown that the final term in equation (20)

$$\frac{m_e \nu_{en}}{e} (\mathbf{v}_I - \mathbf{v}),$$

which would then be proportional to $\mathbf{J} \times \mathbf{B}$, becomes small compared with the third term

$$\frac{1}{c}(\mathbf{v}_e - \mathbf{v}_I) \times \mathbf{B},$$

also proportional to $\mathbf{J} \times \mathbf{B}$. Both of these rather technical claims are justified in detail in Appendix A. (In both cases, the incurred error is of order $(m_e/\mu_{In})^{1/2}$.) These simplifications allow us to write the electron force balance equation as

$$\mathbf{E} + \frac{\mathbf{v}}{c} \times \mathbf{B} - \frac{\mathbf{J} \times \mathbf{B}}{en_e c} - \frac{\mathbf{J}}{\sigma_{cond}} + \frac{(\mathbf{J} \times \mathbf{B}) \times \mathbf{B}}{c^2 \gamma \rho \rho_I} = 0, \quad (23)$$

where the electrical conductivity has been defined as

$$\sigma_{cond} \equiv \frac{e^2 n_e}{m_e \nu_{en}} \quad (24)$$

The associated resistivity η is (e.g. Jackson 1975)

$$\eta = \frac{c^2}{4\pi\sigma_{cond}}, \quad (25)$$

which has units of $\text{cm}^2 \text{s}^{-1}$. Numerically (e.g. Blaes & Balbus 1994, Balbus & Terquem 2001):

$$\eta = 234 \left(\frac{n}{n_e} \right) T^{1/2} \text{ cm}^2 \text{ s}^{-1} \quad (26)$$

Equation (23) is a general form of Ohm's law for a moving, multiple fluid system.

Next, we make use of two of Maxwell's equations. The first is Faraday's induction law:

$$\nabla \times \mathbf{E} = -\frac{1}{c} \frac{\partial \mathbf{B}}{\partial t}. \quad (27)$$

We substitute \mathbf{E} from equation (23) to obtain an equation for the self-induction of the magnetized fluid,

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times \left[\mathbf{v} \times \mathbf{B} - \frac{\mathbf{J} \times \mathbf{B}}{en_e} + \frac{(\mathbf{J} \times \mathbf{B}) \times \mathbf{B}}{c\gamma\rho\rho_I} - \frac{c\mathbf{J}}{\sigma_{cond}} \right] \quad (28)$$

It remains to relate the current density \mathbf{J} to the magnetic field \mathbf{B} . This is accomplished by the second Maxwell equation,

$$\frac{4\pi}{c} \mathbf{J} = \nabla \times \mathbf{B} + \frac{1}{c} \frac{\partial \mathbf{E}}{\partial t} \quad (29)$$

The final term in the above is the displacement current, and it may be ignored. Indeed, since we have not, and will not, use the ‘‘Gauss's Law’’ equation

$$\nabla \cdot \mathbf{E} = 4\pi e(Zn_I - n_e), \quad (30)$$

we *must not* include the displacement current. In Appendix B, we show that departures from charge neutrality in $\nabla \cdot \mathbf{E}$ and the displacement current are

both small terms that contribute at the same order: v^2/c^2 . These must both be self-consistently neglected in nonrelativistic MHD. (The final Maxwell equation $\nabla \cdot \mathbf{B} = 0$ adds nothing new. It is automatically satisfied by equation (27), as long as the initial magnetic field satisfies this divergence free condition.) These considerations imply

$$\mathbf{J} = \frac{c}{4\pi} \nabla \times \mathbf{B} \quad (31)$$

for use in equation (28).

To summarize, the fundamental equations of a weakly ionized fluid are mass conservation of the dominant neutrals (eq.[1])

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0, \quad (32)$$

the equation of motion (eq. [18] with [31])

$$\rho \frac{\partial \mathbf{v}}{\partial t} + \rho \mathbf{v} \cdot \nabla \mathbf{v} = -\nabla P - \rho \nabla \Phi + \frac{1}{4\pi} (\nabla \times \mathbf{B}) \times \mathbf{B}, \quad (33)$$

and the induction equation (eq. [28] with [25] and [31])

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times \left[\mathbf{v} \times \mathbf{B} - \frac{c(\nabla \times \mathbf{B}) \times \mathbf{B}}{4\pi e n_e} + \frac{[(\nabla \times \mathbf{B}) \times \mathbf{B}] \times \mathbf{B}}{4\pi \gamma \rho \rho_I} - \eta \nabla \times \mathbf{B} \right] \quad (34)$$

It is only natural that the reader should be a little taken aback by the sight of equation (34). Be assured that it is rarely, if ever, needed in full generality: almost always one or more terms on the right side of the equation may be discarded. When only the induction term $\mathbf{v} \times \mathbf{B}$ is important, we refer to this regime as *ideal* MHD. The three remaining terms on the right are the nonideal MHD terms.

To get a better feel for the relative importance of the nonideal MHD terms in equation (28), we denote the terms on the right side of the equation, moving left to right, as *I* (induction), *H* (Hall), *A* (ambipolar diffusion), and *O* (Ohmic resistivity). Then, the scaling of each of these terms relative to the Hall term *H* is

$$\frac{I}{H} \sim \frac{v}{v_I - v_e}, \quad \frac{A}{H} \sim Z \frac{\omega_{cI}}{\gamma \rho}, \quad \frac{O}{H} \sim \frac{\nu_{en}}{\omega_{ce}} \quad (35)$$

where the ion cyclotron frequency $\omega_{cI} = eB/m_I c$, and similarly for the electron frequency ω_{ce} , with m_e replacing m_I .

We will always be in a regime in which the presence of the induction term is not in question, i.e. the relative ion-electron drift velocity $v_I - v_e$ is always comparable to or in fact much less than v . More interesting is the relative importance of the nonideal terms. The explicit dependence of A/H and O/H in terms of the fluid properties of a cosmic gas is given in Balbus & Terquem (2001):

$$\frac{A}{H} = Z \left(\frac{9 \times 10^{12} \text{ cm}^{-3}}{n} \right)^{1/2} \left(\frac{T}{10^3 \text{ K}} \right)^{1/2} \left(\frac{v_A}{c_S} \right) \quad (36)$$

and

$$\frac{O}{H} = \left(\frac{n}{8 \times 10^{17} \text{ cm}^{-3}} \right)^{1/2} \left(\frac{c_S}{v_A} \right) \quad (37)$$

Here n is the total number density of all particles, T is the kinetic temperature, v_A is the so-called *Alfvén velocity*,

$$\mathbf{v}_A = \frac{\mathbf{B}}{\sqrt{4\pi\rho}} \quad (38)$$

and c_S is the isothermal speed of sound,

$$c_S^2 = 0.429 \frac{kT}{m_p} \quad (39)$$

where k is the Boltzmann constant and m_p the mass of the proton. The coefficient 0.429 corresponds to a mean mass per particle of $2.33m_p$, appropriate to a molecular cosmic abundance gas.

As reassurance that the fully general nonideal MHD induction equation is not needed for our purposes, note that equations (36) and (37) imply that for all three nonideal MHD terms to be comparable, $T \sim 10^8 \text{ K}$! Obviously this is not an issue for protostellar disks. In figure (2), we plot the domains of relative dominance of the nonideal MHD terms in the nT plane. Protostellar disks are dominated by the resistive and Hall nonideal MHD terms, except in the innermost regions, where ohmic dissipation is largest, and the outermost regions, where ambipolar diffusion becomes important (Wardle 1999).

Our emphasis of the relative ordering of the nonideal terms in the induction equation should not obscure the fact that ideal MHD is often an excellent approximation, even when the ionization fraction is $\ll 1$. For example, the ratio of the ideal inductive term to the ohmic loss term is given by the Lundquist number

$$\mathbb{L} = \frac{v_A L}{\eta} \quad (40)$$

where L is a characteristic gradient length scale. To orient ourselves, let us set $L = 0.1R$, where R is the radial location in the disk. Then \mathbb{L} is given by

$$\mathbb{L} \simeq 2.5(n_e/n)(v_A/c_S)R_{cm},$$

R_{cm} being the radius in centimeters. In other words, the critical ionization fraction at which $\mathbb{L} = 1$ is about

$$(n_e/n)_{crit} = 0.4(c_S/v_A)R_{cm}^{-1} \sim 10^{-13}(c_S/10v_A)$$

at $R = 1 \text{ AU}$. The actual nebular ionization fraction at this location may be above or below this during the course of the solar systems evolution, but the

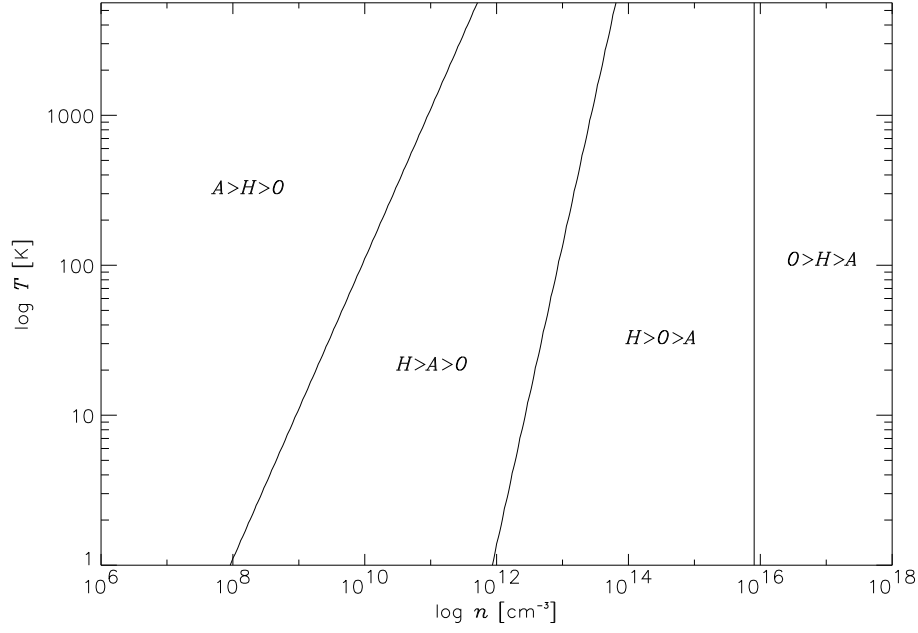


Fig. 2. Parameter space for nonideal MHD. The curves correspond to the case $v_A/c_S = 0.1$. (From Kunz & Balbus 2004.)

point worth noting here is that R_{cm} is a large number for a protostellar disk, whatever the ionization source! Ionization fractions far, far below unity can render an astrophysical gas a near perfect electrical conductor. It therefore makes a great deal of sense to begin by examining the behavior of an ideal MHD fluid.

4 Ideal MHD

The fundamental equations of ideal (single-fluid) MHD are mass conservation (1)

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0 \quad (41)$$

the dynamical equation of motion

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla) \mathbf{v} = -\frac{1}{\rho} \nabla \left(P + \frac{B^2}{8\pi} \right) - \nabla \Phi + \left(\frac{\mathbf{B}}{4\pi} \cdot \nabla \right) \mathbf{B} \quad (42)$$

where Φ is the (central) gravitational potential function, and our newly simplified induction equation

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{v} \times \mathbf{B}) \quad (43)$$

We shall work in a standard cylindrical coordinate system (R, ϕ, z) , where R is the radius, ϕ the azimuthal angle, and z the vertical coordinate. In these coordinates, the three components of the equation of motion are

$$\left[\frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla \right] v_R - \frac{v_\phi^2}{R} = -\frac{1}{\rho} \frac{\partial}{\partial R} \left(P + \frac{B^2}{8\pi} \right) - \frac{\partial \Phi}{\partial R} + \frac{\mathbf{B}}{4\pi\rho} \cdot \nabla B_R - \frac{B_\phi^2}{4\pi\rho R}, \quad (44)$$

$$\left[\frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla \right] (Rv_\phi) = -\frac{1}{\rho} \frac{\partial}{\partial \phi} \left(P + \frac{B^2}{8\pi} \right) + \frac{\mathbf{B}}{4\pi\rho} \cdot \nabla (RB_\phi), \quad (45)$$

$$\left[\frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla \right] v_z = -\frac{1}{\rho} \frac{\partial}{\partial z} \left(P + \frac{B^2}{8\pi} \right) - \frac{\partial \Phi}{\partial z} + \frac{\mathbf{B}}{4\pi\rho} \cdot \nabla B_z, \quad (46)$$

and the three components of the induction equation are

$$\left[\frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla \right] B_R = -B_R \nabla \cdot \mathbf{v} + \mathbf{B} \cdot \nabla v_R \quad (47)$$

$$\left[\frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla \right] \frac{B_\phi}{R} = -\frac{B_\phi}{R} \nabla \cdot \mathbf{v} + \mathbf{B} \cdot \nabla \left(\frac{v_\phi}{R} \right) \quad (48)$$

$$\left[\frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla \right] B_z = -B_z \nabla \cdot \mathbf{v} + \mathbf{B} \cdot \nabla v_z \quad (49)$$

For many problems of interest, an important simplification can be made to these equations. The essence of rotational dynamics is local. Imagine going to a location R_0 in the disk. The Keplerian angular velocity Ω at $R = R_0$ is Ω_0 . We hold Ω_0 fixed but allow R_0 to become arbitrarily large. Thus, $v_\phi(R_0)$ is unbounded, but $v_\phi/R_0 = \Omega_0$ is finite.

Next, we introduce local coordinates (x, α, z) defined by

$$R = R_0 + x, \quad x \ll R_0, \quad (50)$$

$$\alpha = \phi - \Omega_0 t \ll \pi \quad (51)$$

and z is unchanged. We will formally introduce a new coordinate $t' = t$. This is desirable because we wish to take a time derivative at constant x, y, z , not at constant R, ϕ, z . In fact,

$$\frac{\partial}{\partial t'} = \frac{\partial}{\partial t} + \Omega_0 \frac{\partial}{\partial \alpha} \quad (52)$$

Partial derivatives with respect to R and ϕ become partial derivatives with respect to x and α respectively. The Lagrangian derivative

$$\frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla \quad (53)$$

becomes

$$\frac{\partial}{\partial t'} + (\mathbf{v} - R\Omega_0 \mathbf{e}_\phi) \cdot \nabla \quad (54)$$

(The ∇ operator now formally involves x and α derivatives.) This transformation suggests that we introduce a new azimuthal velocity,

$$w_\alpha = v_\phi - R\Omega_0 \quad (55)$$

and for notational consistency we will use w_x and w_z for the radial and vertical velocities, though they are identical to v_R and v_z .

In the local approximation, we assume that the magnitude w is much smaller than the (formally infinite) rotation velocity, and that w , the Alfvén speed v_A , and the thermal velocity $(P/\rho)^{1/2}$ are comparable in magnitude. In the limit $R_0 \rightarrow \infty$, the radial equation becomes

$$\left[\frac{\partial}{\partial t'} + \mathbf{w} \cdot \nabla \right] w_R - 2\Omega_0 w_\phi = R(\Omega_0^2 - \Omega_K^2) - \frac{1}{\rho} \frac{\partial}{\partial x} \left(P + \frac{B^2}{8\pi} \right) + \frac{\mathbf{B}}{4\pi\rho} \cdot \nabla B_R, \quad (56)$$

where the Keplerian angular velocity at R satisfies

$$R\Omega_K^2 = \frac{\partial\Phi}{\partial R} \quad (57)$$

Expanding to first order in R ,

$$R(\Omega_0^2 - \Omega_K^2) = -x \frac{d\Omega_K^2}{dR} = 3\Omega_0^2 \quad (58)$$

The final form of the equation is

$$\left[\frac{\partial}{\partial t} + \mathbf{w} \cdot \nabla \right] w_x - 2\Omega w_y = 3\Omega^2 x - \frac{1}{\rho} \frac{\partial}{\partial x} \left(P + \frac{B^2}{8\pi} \right) + \frac{\mathbf{B}}{4\pi\rho} \cdot \nabla B_x, \quad (59)$$

where we have dropped the subscript 0 and the prime '. The vector components x, y, z refer to the quasi-Cartesian system $dx = dR$, $dy = R d\alpha$, and dz as before. (See fig. 2.) When non-Keplerian disks are considered, $3\Omega^2$ should be replaced by $d\Omega^2/d \ln R$.

The transformed azimuthal (y) equation is

$$\left[\frac{\partial}{\partial t} + \mathbf{w} \cdot \nabla \right] w_y - 2\Omega w_x = -\frac{1}{\rho} \frac{\partial}{\partial y} \left(P + \frac{B^2}{8\pi} \right) + \frac{\mathbf{B}}{4\pi\rho} \cdot \nabla B_y, \quad (60)$$

and the z equation is

$$\left[\frac{\partial}{\partial t} + \mathbf{w} \cdot \nabla \right] w_z = -\frac{\partial\Phi}{\partial z} - \frac{1}{\rho} \frac{\partial}{\partial z} \left(P + \frac{B^2}{8\pi} \right) + \frac{\mathbf{B}}{4\pi\rho} \cdot \nabla B_z, \quad (61)$$

The mass conservation and induction equations, like the z equation of motion, show no change of form in these corotating coordinates:

$$\frac{\partial\rho}{\partial t} + \nabla \cdot (\rho \mathbf{w}) = 0, \quad (62)$$

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{w} \times \mathbf{B}) \quad (63)$$

To summarize, we have transformed our general equations into a coordinate frame rotating at the Keplerian angular velocity $\Omega_K(R_0)$, restricting the spatial domain to a small neighborhood around a particular fluid element. The rotational dynamics appears in the form of a Coriolis force $-2\boldsymbol{\Omega} \times \mathbf{w}$, and a centrifugal force that cancels the main gravitational force, leaving the residual tidal term $3\Omega^2 x$. These simplified local equations, which are the magnetized version of what is known as the *Hill system*, retain a rich dynamical content. This includes the full development of MHD turbulence.

4.1 Magnetorotational Instability

Let us apply equations (56)- (63) to the study of the paths of fluid elements departing from the origin $x = y = z = 0$. These *Lagrangian displacements* are in the xy orbital plane, and depend only upon z . We will ignore vertical stratification here, so that the equilibrium is z independent. Thus, we may assume a spatial dependence of e^{ikz} for the perturbed fluid elements.

We assume a very simple equilibrium magnetic field: a constant vertical component B , no components in the orbital xy plane. If the displacement vector of an element of fluid in the xy plane is $\boldsymbol{\xi}$, then the perturbed magnetic field at given location is

$$\delta \mathbf{B} = \nabla \times (\boldsymbol{\xi} \times \mathbf{B}) \quad (64)$$

This result, which is true quite generally, is lengthy to prove by a direct assault, but can be intuited rather easily. The left side of the exact equation (63) may be written $\delta \mathbf{B} / \delta t$, the change in \mathbf{B} at fixed location divided by the change in time. In the equilibrium state this is zero. Imagine perturbing this pure state by giving each fluid an additional finite velocity \mathbf{U} , but letting it act only for an infinitesimal time δt . Then

$$\delta \mathbf{B} = \nabla \times (\mathbf{w} \delta t \times \mathbf{B}) + \nabla \times (\mathbf{U} \delta t \times \mathbf{B}) = 0 + \nabla \times (\boldsymbol{\xi} \times \mathbf{B}) \quad (65)$$

where $\boldsymbol{\xi} = \mathbf{U} \delta t$ is the displacement of the fluid relative to its equilibrium path. This result is perhaps more clear in integral form. If we integrate over a surface A and use Stokes Theorem, there obtains

$$\int \delta \mathbf{B} \cdot d\mathbf{A} = \int (\boldsymbol{\xi} \times \mathbf{B}) \cdot d\mathbf{s} \equiv \int (d\mathbf{s} \times \boldsymbol{\xi}) \cdot \mathbf{B}, \quad (66)$$

where $d\mathbf{s}$ is a vector line element on the curve bounding A . This states that whatever explicit change in magnetic flux there is through A as the surface is displaced by $\boldsymbol{\xi}$, it is precisely compensated by the gain (or loss) of magnetic flux through of the side of the cylinder (elemental area $d\mathbf{s} \times \boldsymbol{\xi}$) swept out by A . This is just a round about way of saying that the magnetic flux through A is conserved when moving *with* the fluid itself.

For the problem at hand,

$$\delta \mathbf{B} = \nabla \times (\boldsymbol{\xi} \times \mathbf{B}) = i \mathbf{k} \times (\boldsymbol{\xi} \times \mathbf{B}) = i(\mathbf{k} \cdot \mathbf{B}) \boldsymbol{\xi}, \quad (67)$$

since $\mathbf{k} \cdot \boldsymbol{\xi} = 0$. The magnetic term in equations (56) and (60) is

$$\frac{1}{4\pi\rho} \mathbf{B} \cdot \nabla \delta \mathbf{B} = -\frac{(\mathbf{k} \cdot \mathbf{B})^2}{4\pi\rho} \boldsymbol{\xi} \equiv -(\mathbf{k} \cdot \mathbf{v}_A)^2 \boldsymbol{\xi} \quad (68)$$

We have introduced the so-called *Alfvén velocity*

$$\mathbf{v}_A = \frac{\mathbf{B}}{\sqrt{4\pi\rho}} \quad (69)$$

whose physical interpretation we discuss below. For the moment, note that the magnetic force is like tension or a spring: it is always restoring, and proportional to the displacement.

In equations (56) and (60), the derivative

$$\frac{D}{Dt} \equiv \frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla$$

is just the time derivative following a fluid element, so that we will write $D\mathbf{w}/Dt = \ddot{\boldsymbol{\xi}}$, and of course $\mathbf{w} = \dot{\boldsymbol{\xi}}$. Finally, we will allow for the possibility of non-Keplerian rotation. The equations of motion for the fluid element displacements ξ_x and ξ_y are then

$$\ddot{\xi}_x - 2\Omega \dot{\xi}_y = -\left[\frac{d\Omega^2}{d \ln R} + (kv_A)^2 \right] \xi_x \quad (70)$$

$$\ddot{\xi}_y + 2\Omega \dot{\xi}_x = -(kv_A)^2 \xi_y \quad (71)$$

These are a simple set of coupled linear equations with time-dependent solutions of the form $e^{i\omega t}$. Without loss of generality, we take $\omega > 0$. In the absence of rotation, the equations decouple completely and one finds

$$\omega = |kv_A| \quad (72)$$

These are waves that travel along the magnetic field lines with group and phase velocity v_A , exactly like waves on a string. These so-called *Alfvén waves* are thus simple, nondispersive, incompressible disturbances.

On the other hand, in the absence of rotation, we find

$$\omega^2 = 4\Omega^2 + \frac{d\Omega^2}{d \ln R} \equiv \kappa^2 \quad (73)$$

κ^2 is known as the epicyclic frequency. In most astrophysical applications, Ω decreases radially outwards and $\kappa < 4\Omega^2$. The fluid displacements for this mode execute elliptical with a major to minor axis ratio of $\kappa/2\Omega$. The major axis lies along the azimuthal direction. The sense of the rotation around

the ellipse is retrograde, opposite to the sense of Ω . These paths are known as epicycles. Epicyclic motion corresponds to the first order departure from circular orbits for simple planetary motion.

The full dispersion relation resulting from equations (70) and (71) is more complicated than one might have expected:

$$\omega^4 - [\kappa^2 + 2(kv_A)^2]\omega^2 + (kv_A)^2 \left[(kv_A)^2 + \frac{d\Omega^2}{d \ln R} \right] = 0 \quad (74)$$

This is clearly not just a matter of adding Alfvénic and epicyclic motion in quadrature, something new is going on. The dispersion relation is a simple quadratic equation in ω^2 , and it is straightforward to show that this quantity must be real. Stability may then be investigated by passing through the point $\omega^2 \rightarrow 0$. Therefore, marginal instability corresponds to

$$(kv_A)^2 + \frac{d\Omega^2}{d \ln R} = 0 \quad (75)$$

and *instability* is present when

$$\frac{d\Omega^2}{d \ln R} < 0. \quad (76)$$

Of course, this condition is nearly universally satisfied for any type of astrophysical disk. This is the simplest manifestation of the *magnetorotational instability*, or MRI.

4.2 Physical Interpretation

The equations of motion (70) and (71) have a very simple mechanical analogue: they are exactly the equations of two point masses in orbit around a central body, bound together by a spring of frequency kv_A (Balbus & Hawley 1992).

This mechanical analogy immediately offers a physical explanation of the MRI (see figure [3]). Two masses are connected by a spring, an inner mass m_i and an outer mass m_o . The mass m_i orbits faster than the outer mass m_o , causing the spring to stretch. The spring pulls backwards on m_i and forwards on m_o . The negative torque on m_i causes it to lose angular momentum and sink, while the positive torque on m_o causes it to gain angular momentum and rise. Thus, even though the spring nominally supplies an attractive force, in a rotating frame this drives the masses apart. The continued stretching of the spring just makes matters worse, and the process runs away. The mass with the higher angular momentum obtains yet more, the mass with less angular momentum loses what little it has. It is the “same old story” in a dynamical context: the rich get richer and the poor get poorer.

In a protostellar disk, it is the magnetic tension force that plays the role of the spring, and the two masses are any two fluid elements tethered by the field line. The linear phase of the instability is the exponentially growing

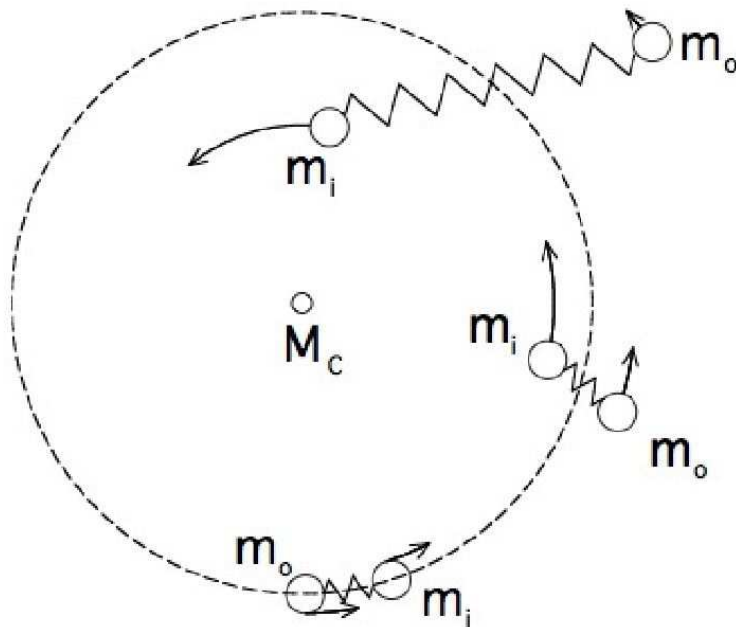


Fig. 3. The magnetorotational instability. Magnetic fields in a disk bind fluid elements precisely as though they were masses in orbit connected by a spring. The inner element m_i orbits faster than the outer element m_o , and the spring causes a net transfer of angular momentum from m_i to m_o . This transfer is unstable, as described in the text. The inner mass continues to sink, whereas the outer mass rises farther outward. (*Figure courtesy of H. Ji.*)

separation of the displaced fluid elements, which is followed by the nonlinear mixing of gas parcels from different regions of the disk. The mixing seems to lead to something resembling a classical turbulent cascade, though the details of this process, with different viscous and resistive dissipation scales, remain to be fully understood.

Notice that angular momentum transport is not something that happens as a consequence of the nonlinear development of the MRI, it is the essence of the MRI even in its linear phase. The very act of transporting angular momentum from the inner to outer fluid elements via a magnetic couple is a spontaneously unstable process.

4.3 General Adiabatic Disturbances

If Ω is a function only of cylindrical radius R , then for general magnetic field geometries, local incompressible WKB disturbances with space-time dependence

$$\delta X \sim \exp(ik_R R + ik_z z - i\omega t) \quad (77)$$

where δX is the Eulerian perturbation of any flow quantity, satisfy the following set of linearized dynamical equations:

$$k_R \delta v_R + k_z \delta v_z = 0, \quad (78)$$

$$-i\omega \delta v_R - 2\Omega \delta v_\phi + i\frac{k_R}{\rho} \left[\delta P + \frac{\mathbf{B} \cdot \delta \mathbf{B}}{4\pi} \right] - i\frac{\mathbf{k} \cdot \mathbf{B}}{4\pi\rho} \delta B_R = 0, \quad (79)$$

$$-i\omega \delta v_\phi + \frac{\kappa^2}{2\Omega} \delta v_R - i\frac{\mathbf{k} \cdot \mathbf{B}}{4\pi\rho} \delta B_R = 0, \quad (80)$$

$$-i\omega \delta v_z + i\frac{k_z}{\rho} \left[\delta P + \frac{\mathbf{B} \cdot \delta \mathbf{B}}{4\pi} \right] - i\frac{\mathbf{k} \cdot \mathbf{B}}{4\pi\rho} \delta B_z = 0. \quad (81)$$

The linearized induction equations are

$$-i\omega \delta B_R - i(\mathbf{k} \cdot \mathbf{B}) \delta v_R = 0, \quad (82)$$

$$-i\omega \delta B_\phi - \delta B_R \frac{d\Omega}{d \ln R} - i(\mathbf{k} \cdot \mathbf{B}) \delta v_\phi = 0, \quad (83)$$

$$-i\omega \delta B_z - i(\mathbf{k} \cdot \mathbf{B}) \delta v_z = 0. \quad (84)$$

Finally, the entropy equation is

$$i\omega \gamma \frac{\delta \rho}{\rho} + \delta v_z \frac{\partial \ln P \rho^{-\gamma}}{\partial z} = 0, \quad (85)$$

where γ is the adiabatic index (not to be confused with the collision rate). We have ignored background radial gradients in the entropy and pressure, but retained their vertical gradients, in accordance with the assumption of a thin disk.

This set of linearized equations leads to the dispersion relation (Balbus & Hawley 1991)

$$\left(\frac{k^2}{k_z^2} \right) \tilde{\omega}^4 - \left[\kappa^2 + \left(\frac{k_R}{k_z} \right)^2 N^2 \right] \tilde{\omega}^2 - 4\Omega^2 (\mathbf{k} \cdot \mathbf{v}_A)^2 = 0. \quad (86)$$

Here

$$\tilde{\omega}^2 = \omega^2 - (\mathbf{k} \cdot \mathbf{v}_A)^2, \quad k^2 = k_z^2 + k_R^2, \quad (87)$$

and

$$N^2 = -\frac{1}{\gamma\rho} \frac{\partial P}{\partial z} \frac{\partial \ln P \rho^{-\gamma}}{\partial z}. \quad (88)$$

N is known as the Brunt-Väisälä frequency. It is the natural frequency at which a vertically displaced fluid element would oscillate in the disk due to buoyancy forces. In general there is also a contribution due to radial gradients as well (Balbus & Hawley 1991), but this usually may be ignored in a

rotationally supported (supersonic orbital speed) disk. Without any magnetic field, the dispersion relation becomes

$$\omega^2 = (k_z/k)^2 \kappa^2 + (k_R/k)^2 N^2 \quad (89)$$

Since the displacement of the fluid element is incompressible, the wave number ratio k_z/k is a measure of the *radial* displacement, whereas k_R/k is a measure of the *vertical* displacement. There is no instability in this case, only a wavelike response due to the restoring Coriolis and buoyant forces.

Notice that N vanishes in the disk plane by symmetry, it affects only the behavior of disturbances at least one scale height or so above $z = 0$. Moreover, the actual value of N is likely to be determined by radiative diffusion processes in the vertical direction. The radiation requires a source, which in disks is the turbulence we are trying to explain in the first place! (In protostellar disks, external heating of the upper layers is also a source of departure from adiabatic behavior.)

The other limit of interest is $k_R = 0$, which returns us to the dispersion relation of the previous section. (The expression kv_A should of course be replaced by $\mathbf{k} \cdot \mathbf{v}_A$.) This is fortunate because it can be shown (Balbus & Hawley 1992) that the most rapidly growing modes are displacements in the plane of the disk with $k_R = 0$. Since it is the differential rotation that ultimately destabilizes, it is very sensible that these displacements, which most effectively sample the differential rotation, are the most unstable. For a Keplerian rotation profile, it may be directly computed from the dispersion relation that the wave number of maximum growth is given by

$$\mathbf{k} \cdot \mathbf{v}_A = (\sqrt{15}/4)\Omega \quad (\text{wavenumber of maximum growth}), \quad (90)$$

or 0.97Ω . By comparison, the largest unstable wavelength (i.e., smallest unstable wave number) has a value of 1.73Ω for $\mathbf{k} \cdot \mathbf{v}_A$. The maximum growth rate is

$$|\omega| = (3/4)\Omega \quad (\text{maximum growth rate}) \quad (91)$$

This is a *very* large growth rate, with amplitudes growing at a rate of $\exp(3\pi/2) \sim 111$ per orbit.

4.4 Angular Momentum Transport

The azimuthal equation of motion (45) may be written

$$\frac{\partial(\rho R v_\phi)}{\partial t} + \nabla \cdot \left[R \left(\rho \mathbf{v} v_\phi - \rho \mathbf{v}_A v_{A\phi} + \left(P + \frac{B^2}{8\pi} \right) \mathbf{e}_\phi \right) \right] = 0, \quad (92)$$

where $v_{A\phi}$ is the azimuthal component of the Alfvén velocity. This is an equation for angular momentum conservation, with angular momentum density $\rho R v_\phi$, and an azimuthally averaged flux of

$$\mathbf{F}_J = \langle R(\rho \mathbf{v} v_\phi - \rho \mathbf{v}_A v_{A\phi}) \rangle \quad (93)$$

In protostellar disks, we are interested in the radial transport of angular momentum,

$$F_{JR} = R\rho(v_R v_\phi - v_{AR} v_{A\phi}) \quad (94)$$

It is most instructive to begin with the simple case of linear instability in a uniform, vertical magnetic field. What is the lowest order flux that results from these exponentially growing modes? In equilibrium, v_R and the Alfvén velocities vanish. The linear perturbation of the radial velocity at a given spatial location (a so-called Eulerian perturbation) will be denoted δv_R . Let the radial displacement ξ_R of a fluid element from its equilibrium location be given by

$$\xi_R = ae^{\gamma t} \cos(kz) \quad (95)$$

where γ and k correspond to the maximum growth rate and its associated wave number, and a is a slowly varying amplitude. Then

$$\delta v_R = \dot{\xi}_R = \gamma \xi_R \quad (96)$$

The azimuthal velocity v_ϕ consists of an unperturbed Keplerian velocity $R\Omega$ plus a linear perturbation δv_ϕ . The product of δv_R and Ω contributes zero when a height integration is performed because of the cosine factor, so we must consider the direct product $\delta v_R \times \delta v_\phi$ to obtain the lowest order contribution. In calculating δv_ϕ , note that

$$\delta v_\phi = \dot{\xi}_\phi - \xi_R \frac{d\Omega}{d \ln R} \quad (97)$$

The subtraction is needed to eliminate the change in velocity a displaced fluid element would make even if there were no physical change in the rotation velocity at the new radial location: the actual change in velocity δv_ϕ is due to the change in $\dot{\xi}_\phi$ the displaced fluid element makes *in excess* of $\xi_R d\Omega/d \ln R$. Equation (70) then gives

$$\delta v_\phi = \dot{\xi}_\phi - \xi_R \frac{d\Omega}{d \ln R} = \left(\frac{\gamma^2 + k^2 v_A^2}{2\Omega} \right) \xi_R \quad (98)$$

and therefore the velocity contribution to the angular momentum flux is

$$\langle \delta v_R \delta v_\phi \rangle = \frac{\gamma}{2\Omega} (\gamma^2 + k^2 v_A^2) \langle \xi_R^2 \rangle \quad (99)$$

Not surprisingly, it is positive (outward).

To calculate the magnetic field correlation, we use equation (67) to go between magnetic field fluctuations and displacements. Then, using (71), we find

$$\xi_\phi = -\frac{2\omega\gamma}{\gamma^2 + k^2 v_A^2} \xi_R \quad (100)$$

from which it follows, among other things, that $\dot{\xi}_\phi$ and δv_ϕ have opposite signs. Hence

$$\frac{\langle \delta B_R \delta B_\phi \rangle}{4\pi\rho} = -\frac{2\Omega k^2 v_A^2 \gamma}{\gamma^2 + k^2 v_A^2} \langle \xi_R^2 \rangle \quad (101)$$

Therefore

$$\langle \delta v_R \delta v_\phi - \frac{\delta B_R \delta B_\phi}{4\pi\rho} \rangle = \frac{\gamma}{2\Omega} \left[\frac{(\gamma^2 + k^2 v_A^2)^2 + 4\Omega^2 k^2 v_A^2}{\gamma^2 + k^2 v_A^2} \right] \langle \xi_R^2 \rangle \quad (102)$$

Using the dispersion relation (74) to simplify this a bit, our expression for the radial angular momentum flux is

$$F_{JR} = \rho R \frac{\gamma}{2\Omega} \left[\frac{8\Omega^2 k^2 v_A^2}{\gamma^2 + k^2 v_A^2} - \kappa^2 \right] \langle \xi_R^2 \rangle \quad (103)$$

Finally, integrating over height and defining the effective surface density Σ by

$$\int \rho \langle \xi_R^2 \rangle dz = \Sigma a^2 e^{2\gamma t} / 2, \quad (104)$$

the angular momentum flux for the most rapidly growing mode is found to be

$$F_{JR} = \frac{3R\Sigma}{4} \Omega^2 a^2 e^{2\gamma t} \quad (105)$$

This result bears some commentary. First, the radial flux is always positive for any unstable mode, not just the most rapidly growing. This, as we shall see, is a direct consequence of energy conservation: energy is extracted from the differential rotation to the fluctuations only if the radial angular momentum flux is positive. Next, note that there is an outward angular momentum flux only to the extent that there is a correlation in the velocity fluctuations. This is also true in turbulent flows, and everything that we have done in this section goes through in much the same way if the fluctuations are turbulent as opposed to wavelike. The important physical point is that an instability that leads to turbulence need not lead to enhanced angular momentum transport. Only turbulence with strongly correlated velocity fields does this. Strong correlations, in turn, are necessary to extract energy from differential rotation. This is indeed the free energy source of shear-driven turbulence, but not any form of turbulence.

This point was made in Stone & Balbus (1996), in which the angular momentum transport resulting from *convective* instability was studied. The angular momentum was tiny in magnitude and *inwardly* directed: it had the “wrong” sign. Disk instabilities based on adverse thermal gradients do not, in general, lead to systematically large outward transfers of angular momentum. When it comes to angular momentum transport, not all turbulence is equivalent.

4.5 Diffusion of a Scalar

A classical problem in protostellar disk theory is to understand how dust particles are mixed with the gas. The fact that the MRI leads to vigorous radial angular momentum transport suggests that other quantities may be transported as well. In this section, we will give an argument that shows that the radial diffusion of angular momentum is indeed closely related to the radial transport of a conserved scalar quantity, say Q . Assume that Q is a fluid element label and satisfies an equation of the form

$$\left(\frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla\right) Q = 0 \quad (106)$$

Then,

$$\frac{\partial Q}{\partial t} + \nabla \cdot (\mathbf{v} Q) - Q \nabla \cdot \mathbf{v} = 0. \quad (107)$$

Using mass conservation, this implies

$$\frac{\partial Q}{\partial t} + \nabla \cdot (\mathbf{v} Q) + Q \frac{D \ln \rho}{Dt} = 0, \quad (108)$$

where

$$\frac{D}{Dt} = \left(\frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla\right) \quad (109)$$

Simplifying the equation for Q , we obtain

$$\frac{\partial(\rho Q)}{\partial t} + \nabla \cdot (\rho \mathbf{v} Q) = 0, \quad (110)$$

which looks very much like the angular momentum conservation equation (92). In a similar way, there will be a turbulent Q -flux given by

$$F_Q = \rho \langle \delta \mathbf{v} \delta Q \rangle \quad (111)$$

Since Q is conserved as we follow a fluid element, the so-called *Lagrangian perturbation* ΔQ , defined by

$$\Delta Q = \delta Q + \boldsymbol{\xi} \cdot \nabla Q \quad (112)$$

must vanish. This is because ΔQ is constructed to be the change in Q following a fluid element as it is displaced by a small distance $\boldsymbol{\xi}$, and such changes *must* vanish since by definition Q does not change along fluid element paths. Hence, our expression for the i th component of F_Q is

$$F_{Qi} = -\rho \langle \delta v_i \xi_j \rangle \partial_j Q \quad (113)$$

This defines the diffusion tensor \mathcal{D}_{ij} ,

$$\mathcal{D}_{ij} = \langle \delta v_i \xi_j \rangle \quad (114)$$

In a height-integrated calculation, it is the RR component of \mathcal{D} that is of importance. In the linear regime $\delta v_R = \gamma \xi_R$, so that

$$\mathcal{D}_{RR} = \gamma \langle \xi_R^2 \rangle \quad (115)$$

In this case, equation (103) gives a relationship between the flux of angular momentum and the diffusion coefficient of a passive scalar,

$$F_{JR} = \frac{\rho R}{2\Omega} \left[\frac{8\Omega^2 k^2 v_A^2}{\gamma^2 + k^2 v_A^2} - \kappa^2 \right] \mathcal{D}_{RR} \quad (116)$$

For the fastest growing linear mode, this gives $F_{JR} = 2\Omega \mathcal{D}_{RR}$.

There is a sense in which the linear theory might find its way into nonlinear turbulent diffusion. In simulations the MRI appears to locally stretch field lines and exponentiate velocity growth over a limited duration of time, before one fluid element becomes mixed with another and the process starts anew. If such a picture is reasonably accurate, then $F_{JR} \sim \Omega \mathcal{D}_{RR}$ may well be valid in turbulent flow. (It is also, of course, what we might expect on the basis of simple dimensional analysis alone!) The important astrophysical point is that if protostellar disks are MHD turbulent, they ought to be well-mixed.

5 Energetics of MHD Turbulence

5.1 Hydrodynamic Considerations

The quantity

$$(\rho \mathbf{v} v_\phi - \rho \mathbf{v}_A v_{A\phi}) \quad (117)$$

plays two conceptually different roles in the theory of accretion disks. We have seen that it is intimately linked to the direct transport of angular momentum, and in this section, we shall study in detail how it extracts free energy from the large scale differential rotation.

Let us begin with the relatively simple case of an adiabatic nonmagnetized gas. The azimuthal velocity is decomposed into a time-steady large rotational component $R\Omega$ plus a fluctuating component u_ϕ :

$$v_\phi = R\Omega + u_\phi \quad (118)$$

we will assume neither that u_ϕ is small compared with $R\Omega$ nor that the mean value of u_ϕ vanishes, though in practice both might well be the case. We will write u_R and u_z for the radial and vertical velocity components respectively, and \mathbf{v} for the full velocity vector. Once again, we think of these as fluctuations, but our treatment will in fact be exact.

The adiabatic radial equation of motion is

$$\rho \left[\frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla \right] u_R - \frac{\rho}{R} (R\Omega + u_\phi)^2 = -\frac{\partial P}{\partial R} - \rho \frac{\partial \Phi}{\partial R} \quad (119)$$

Multiplying by u_R and regrouping:

$$\rho \left[\frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla \right] \left(\frac{u_R^2}{2} \right) - 2\rho\Omega u_\phi u_R - \rho \frac{u_R u_\phi^2}{R} = -u_R \left(\frac{\partial P}{\partial R} + \rho \frac{\partial \Phi_{eff}}{\partial R} \right) \quad (120)$$

where

$$\Phi_{eff} = \Phi - \int^R s \Omega^2(s) ds \quad (121)$$

Exactly the same manipulations with the ϕ and z equations produce

$$\rho \left[\frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla \right] \left(\frac{u_\phi^2}{2} \right) + \rho u_\phi u_R \frac{\kappa^2}{2\Omega} + \rho \frac{u_R u_\phi^2}{R} = -u_\phi \frac{\partial P}{R \partial \phi} \quad (122)$$

$$\rho \left[\frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla \right] \left(\frac{u_z^2}{2} \right) = -u_z \left(\frac{\partial P}{\partial z} + \rho \frac{\partial \Phi_{eff}}{\partial z} \right) \quad (123)$$

where we have assumed that Φ is independent of ϕ . Adding the three dynamical equations and using mass conservation leads to

$$\frac{\partial}{\partial t} \left(\frac{1}{2} \rho u^2 \right) + \nabla \cdot (\rho u^2 \mathbf{v} / 2) + \rho u_R u_\phi \frac{d\Omega}{d \ln R} = -\mathbf{u} \cdot \nabla P - \rho \mathbf{u} \cdot \nabla \Phi_{eff} \quad (124)$$

where $u^2 = u_R^2 + u_\phi^2 + u_z^2$. Yet another use of mass conservation and a regrouping of the pressure term gives us

$$\frac{\partial}{\partial t} \left(\frac{1}{2} \rho u^2 + \rho \Phi_{eff} \right) + \nabla \cdot \left(\mathbf{v} \left(\frac{1}{2} \rho u^2 + \rho \Phi_{eff} \right) + P \mathbf{u} \right) = P \nabla \cdot \mathbf{u} - \rho u_R u_\phi \frac{d\Omega}{d \ln R} \quad (125)$$

We have already at hand the main structure of our energy equation. The left side is in conservation form with a well defined energy density and energy flux, with the fluctuations isolated. (Notice that the azimuthal average of the energy equation, which we ultimately will be working with, has no rotational terms in the energy flux divergence.) Sources of energy fluctuations are work done by pressure (which may cause heating or cooling), and the all important $R\phi$ stress coupling to the differential rotation.

5.2 The Effects of Magnetic Fields

When magnetic fields are included, everything goes through as before, except that the right side of equation (124) contains the terms

$$-\mathbf{u} \cdot \nabla P_{tot} - \rho \mathbf{u} \cdot \nabla \Phi_{eff} + \frac{\mathbf{u} \cdot [\mathbf{B} \cdot \nabla] \mathbf{B}}{4\pi} \quad (126)$$

The final magnetic term in the above may be written in an index notation as

$$\frac{u_j B_i \partial_i B_j}{4\pi} \quad (127)$$

where i, j, k take the values x, y, z , and repeated indices are summed over. ∂_i denotes the the partial derivative with respect to the i th spatial variable.

To make further progress, we need the induction equation:

$$\left(\frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla\right) \mathbf{B} = -\mathbf{B} \nabla \cdot \mathbf{v} + (\mathbf{B} \cdot \nabla) \mathbf{v}. \quad (128)$$

Take the dot product of this with \mathbf{B} and write the last term in component form:

$$\left(\frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla\right) \frac{B^2}{2} = -B^2 \nabla \cdot \mathbf{v} + B_j B_i \partial_i v_j \quad (129)$$

Now,

$$B_j B_i \partial_i v_j = \partial_i (B_i B_j v_j) - v_j \partial_i (B_j B_i) = \nabla \cdot (\mathbf{v} \cdot \mathbf{B} \mathbf{B}) - R\Omega [\mathbf{B} \cdot \nabla \mathbf{B}]_\phi - u_j B_i \partial_i B_j \quad (130)$$

where we have used $\partial_i B_i = 0$. Note that with the last term, we make contact with the energy equation terms (126). The subscript ϕ on the penultimate term denotes a vector component:

$$R\Omega [\mathbf{B} \cdot \nabla \mathbf{B}]_\phi = \Omega [\mathbf{B} \cdot \nabla (R B_\phi)] = \nabla \cdot (R\Omega \mathbf{B} B_\phi) - B_\phi B_R \frac{d\Omega}{d \ln R} \quad (131)$$

Therefore, the right side of equation (129) becomes

$$-B^2 \nabla \cdot \mathbf{v} + \nabla \cdot [(\mathbf{u} \cdot \mathbf{B}) \mathbf{B}] + B_\phi B_R \frac{d\Omega}{d \ln R} - u_j B_i \partial_i B_j \quad (132)$$

Combining this result with the left side of the equation (129) then, after some cancellation, given us an expression for $u_j B_i \partial_i B_j$: of terms,

$$\begin{aligned} -u_j B_i \partial_i B_j &= \left(\frac{\partial}{\partial t} + \Omega \frac{\partial}{\partial \phi}\right) \frac{B^2}{2} + \nabla \cdot \left(\mathbf{u} \frac{B^2}{2} - (\mathbf{u} \cdot \mathbf{B}) \mathbf{B}\right) \\ &\quad + \frac{B^2}{2} \nabla \cdot \mathbf{u} - B_\phi B_R \frac{d\Omega}{d \ln R} \end{aligned} \quad (133)$$

Armed with this result, we return to (126) and make a substitution for the final term. The resulting energy equation can be simplified to

$$\begin{aligned} \frac{\partial}{\partial t} \left(\frac{1}{2} \rho u^2 + \rho \Phi_{eff} + \frac{B^2}{8\pi} \right) &+ \nabla \cdot \left[\mathbf{v} \left(\frac{1}{2} \rho u^2 + \rho \Phi_{eff} \right) + P \mathbf{u} + \frac{\mathbf{B} \times (\mathbf{u} \times \mathbf{B})}{4\pi} \right] + \\ &+ \nabla \cdot \left(\mathbf{e}_\phi R\Omega \frac{B^2}{8\pi} \right) = P \nabla \cdot \mathbf{u} - \rho (u_R u_\phi - v_{AR} v_{A\phi}) \frac{d\Omega}{d \ln R} \end{aligned} \quad (134)$$

The pressure term on the right side of the above equation can be eliminated using the thermal energy equation

$$\frac{3\rho}{2} \left(\frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla \right) \frac{P}{\rho} = -P \nabla \cdot \mathbf{u} - \rho \mathcal{L} \quad (135)$$

(since $\nabla \cdot \mathbf{v} = \nabla \cdot \mathbf{u}$). We have introduced the radiative energy loss term per unit volume $\rho\mathcal{L}$. Carrying through the elimination of the pressure term and averaging over ϕ lead to the total energy equation:

$$\frac{\partial \mathcal{E}}{\partial t} + \nabla \cdot \mathcal{F} = -\rho(u_R u_\phi - v_{AR} v_{A\phi}) \frac{d\Omega}{d \ln R} - \rho\mathcal{L} \quad (\phi \text{ averaged}) \quad (136)$$

where \mathcal{E} is the energy density in fluctuations,

$$\mathcal{E} = \frac{1}{2}\rho u^2 + \rho\Phi_{eff} + \frac{B^2}{8\pi} + \frac{3P}{2}, \quad (137)$$

and \mathcal{F} is the corresponding flux,

$$\mathcal{F} = \mathbf{u} \left(\frac{1}{2}\rho u^2 + \rho\Phi_{eff} \right) + \frac{5P}{2}\mathbf{u} + \frac{\mathbf{B} \times (\mathbf{u} \times \mathbf{B})}{4\pi} \quad (138)$$

We have completely ignored dissipation effects (viscosity and resistivity). What effect would this have on our final equation (136)? The answer is essentially none. While it is true that new dissipation terms would appear on the right side of equation (134), they would also appear with the opposite sign in equation (135). They would completely cancel on the right side of the final equation (136): dissipation is not a loss of energy, but a conversion of mechanical or magnetic field energy into heat. Additional small energy flux terms *would* appear (within the divergence operator) in connection viscosity and resistivity, but these are generally negligible compared with the dynamical terms. The essential physical point is that total energy is conserved in the presence of dissipation even if mechanical energy is lost. Only radiative processes can remove energy from the disk.

We interpret equation (136) as saying that energy is exchanged with the differential rotation at a volumetric rate $-T_{R\phi}d\Omega/d \ln R$, where

$$T_{R\phi} = \rho(u_R u_\phi - v_{AR} v_{A\phi}) \quad (139)$$

is the dominant component of the stress tensor. The energy made available from the differential rotation may in principle remain in velocity fluctuations in the form of waves, but in a thin, cool disk it is more likely that this energy will be locally dissipated, and subsequently radiated. This is because if the energy flux varies radially over a scale of order R itself, then the $T_{R\phi}$ source term on the right will be larger than any term on the left (by an amount of order $u/(R\Omega)$). The stress can only be counterbalanced by the radiative loss term:

$$-T_{R\phi}d\Omega/d \ln R = \rho\mathcal{L} \quad (140)$$

Therefore, although $-T_{R\phi}d\Omega/d \ln R$ is itself a nondissipative source term, in these so-called local models, it works out to be the rate at which energy is dissipated (and ultimately lost by radiation) as well. Bear in mind that the identification of the stress term with dissipation follows from a reasonable

assumption about the behavior of thin disks: local dissipation of the fluctuations. Dissipation is in no way a fundamental property of the left side of equation (140), which under different conditions could just as well be an energy source or reserve for purely adiabatic waves. There are many papers in which this point is misunderstood: *caveat lector*.

6 Resistive and Hall terms

6.1 Local Dispersion Relation

Extensive regions in real protostellar disks are in a regime far from ideal MHD, in which Ohmic decay and Hall electromotive forces are important (e.g. figure 1). It is important to understand how the MRI behaves under these conditions.

The fastest growing modes, as usual, correspond to axial wavenumbers $\mathbf{k} = k\mathbf{e}_z$. Fortunately, these are also the simplest to treat analytically, and we will limit our discussion to this class of disturbance.

The linearly perturbed Hall term in the induction equation is to leading WKB order

$$-\nabla \times \left[\frac{\delta \mathbf{J} \times \mathbf{B}}{en_e} \right] = \nabla \times \left[\frac{c\mathbf{B}}{4\pi en_e} \times (i\mathbf{k} \times \delta \mathbf{B}) \right], \quad (141)$$

since $\delta \mathbf{J} = (c/4\pi)(i\mathbf{k} \times \delta \mathbf{B})$. Taking the curl and simplifying allows us to write the right side as

$$\frac{c}{4\pi en_e} (\mathbf{k} \cdot \mathbf{B})(\mathbf{k} \times \delta \mathbf{B}). \quad (142)$$

(In a fully ionized disk, the presence of n_e in the denominator renders these terms negligibly small.) For our particular problem, \mathbf{k} is axial and $\delta \mathbf{B}$ lies in the disk plane, and the above reduces to

$$\frac{ck^2 B}{4\pi en_e} (\delta B_R \mathbf{e}_\phi - \delta B_\phi \mathbf{e}_R). \quad (143)$$

Including the effects of resistivity, the linearized induction equations (82-83) become:

$$(-i\omega + \eta k^2) \delta B_R + \frac{ck^2 B_z}{4\pi en_e} \delta B_\phi - i(\mathbf{k} \cdot \mathbf{B}) \delta v_R = 0, \quad (144)$$

$$(-i\omega + \eta k^2) \delta B_\phi - \delta B_R \left(\frac{d\Omega}{d \ln R} + \frac{ck^2 B_z}{4\pi en_e} \right) - i(\mathbf{k} \cdot \mathbf{B}) \delta v_R = 0, \quad (145)$$

and the z equation is not needed (trivially satisfied). The linearized dynamical and energy equations from section (4.3) remain unchanged.

The Hall electromotive force terms introduce a phase shift in the induction that leads to a circularly polarized component of what would otherwise be

an Alfvén or slow mode. When an instability is present due to differential rotation, this phase shift can either destabilize or stabilize depending upon the *sign* of B_z .

To see this more quantitatively, we must study the dispersion relation that follows from our linearized set of equations. To this end, we define a Hall parameter v_H which has the dimensions of a velocity (Balbus & Terquem 2001),

$$v_H^2 = \frac{\Omega B_z c}{2\pi e n_e}. \quad (146)$$

Notice that if we take the sense of rotation to be positive, that v_H^2 can have either sign depending upon whether B_z is positive or negative. With $\sigma = -i\omega$, our dispersion relation is

$$\sigma^4 + 2\eta k^2 + \mathcal{C}_2 \sigma^2 + 2\eta k^2(\kappa^2 + k^2 v_A^2)\sigma + \mathcal{C}_0 = 0, \quad (147)$$

where the constants \mathcal{C}_2 and \mathcal{C}_0 are given by

$$\mathcal{C}_2 = \kappa^2 + 2k^2 v_A^2 + \eta^2 k^4 + \frac{k^2 v_H^2}{4\Omega^2} \left(\frac{d\Omega^2}{d \ln R} + k^2 v_H^2 \right) \quad (148)$$

$$\mathcal{C}_0 = k^2 \left(v_A^2 + \frac{\kappa^2 v_H^2}{4\Omega^2} \right) \left(\frac{d\Omega^2}{d \ln R} + k^2 v_A^2 + k^2 v_H^2 \right) + \kappa^2 \eta^2 k^4 \quad (149)$$

6.2 Stability

A sufficient condition for instability is that $\mathcal{C}_0 < 0$. When compared with the purely Ohmic requirement

$$k^2 v_A^2 \left(k^2 v_A^2 + \frac{d\Omega^2}{d \ln R} \right) + \kappa^2 \eta^2 k^4 < 0 \quad (150)$$

The Hall condition $\mathcal{C}_0 < 0$ is much more easily satisfied, since v_H^2 can change sign. Perhaps most striking is the possibility that the intuitive result that large wavenumbers are stabilized can be lost when $v_H^2 < 0$, and that *all* wavenumbers can be unstable, even in the presence of finite resistivity. (See figure [4].) This may be shown straightforwardly from the definition of \mathcal{C}_0 ; the reader may wish to consult Balbus & Terquem (2001) for explicit details. In addition to requiring counteraligned magnetic and rotation axes, this state also demands a magnitude ratio of v_A^2 to v_H^2 lies between about 2 and 4 (for a Keplerian disk). When the field is aligned with the rotation axis, the range of unstable wavenumbers is more restricted than that found for the ideal-MHD MRI.

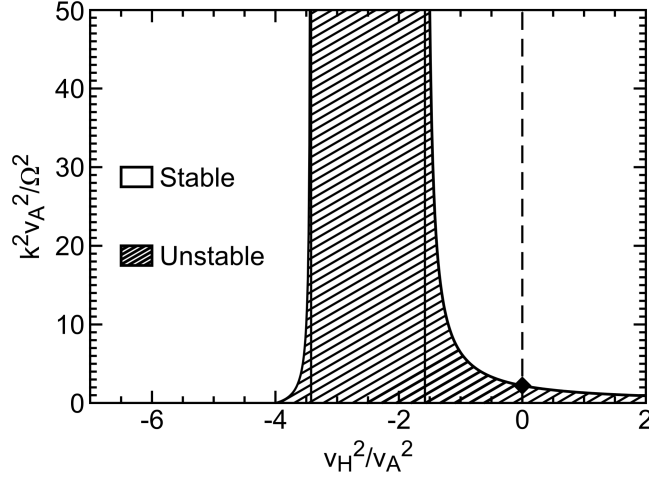


Fig. 4. Range of unstable wavenumbers for a representative disk in the Hall-Ohm regime. The black diamond corresponds to the MRI without the Hall effect. The resistivity is chosen to be $\kappa^2 \eta^2 / v_A^4 = 0.35$. From Balbus & Terquem (2001).

6.3 Numerical Simulations of the Hall-Ohm-MRI.

What does the preceding section imply for the MHD transport properties of protostellar disks? The first local simulations of the nonlinear development of the MRI in the presence of Ohmic resistivity are those of Fleming, Stone, & Hawley (2000). These authors found the important result that the nonlinear turbulent state is easier to maintain when a mean vertical magnetic field is present. This is sometimes misunderstood as a requirement that a disk must have a global vertical magnetic field to maintain MRI turbulence, but in fact all that the very local calculations of Fleming et al. (2000) really tell us is that on this scale, working with a precisely zero mean field is probably too restrictive an approximation for a real disk. This study found a difference of two orders of magnitude in the critical magnetic Reynolds number (defined here as the product of the isothermal times the box size divided by the resistivity) needed to sustain turbulence, depending upon whether a vertical field was present or not, the vertical field runs being the easier to sustain. It is now understood that very high resolution grids are needed for the local study of

the MRI in a shearing box with zero mean field, so these early results should not be used quantitatively. With this precaution understood, it is interesting to note that for the vertical field runs, the Fleming et al. finding seems to be not very different from the $\mathbb{L} = 1$ criterion discussed earlier.

This nature of the nonlinear behavior of a Hall fluid was examined in a numerical investigation by Sano & Stone (2002a,b). These authors carried out an extensive study of the local properties of Hall-MRI turbulence in the shearing box formalism. The results are somewhat surprising.

In linear theory, counter-aligned magnetic and rotation axes produce a very broad wavenumber spectrum of instability, and should be more unstable than the aligned case. What Sano and Stone found, however, was just the opposite: the aligned case showed greater levels of field coherence and higher rms fluctuations than the counter-aligned case. The reason for this is interesting, and illustrative of the hazards of extrapolating directly from linear theory.

In the case of counter-aligned axes, the extended wavenumber of spectrum of linear MRI instability results, in its nonlinear resolution, in a highly efficient turbulent cascade to smaller and smaller scales. In a finite difference numerical simulation, this cascade terminates at the grid scale where energy is ultimately lost. By contrast, the aligned case shows instability only at longer wavelengths (small wavenumbers). At larger wavenumbers, the response of the fluid is wavelike, and this is likely to make a local small scale turbulent cascade considerably less efficient. The result: a less lossy system and greater large scale field coherence.

Finally, when simulations were done with half the field aligned with the rotation axis and half counter-aligned, the Hall effect tended to wash out, and the results were similar to resistive MHD turbulence without the additional electromotive forces. The criterion for the onset of turbulence depended only the relative strength of the resistivity, as measured by the dimensionless number $v_A^2/\eta\Omega$. This so-called *Lundquist number* must be greater than unity or MRI-induced turbulence is not maintained, whether Hall electromotive forces are present or not, according to Sano & Stone. The presence of the Hall terms in the induction equation do, however, promote a slightly elevated rms level when turbulence can be maintained.

The Hall numerical studies are fascinating and suggestive, but caution is necessary before taking the final leap from simulation to real disks. Note, for example, the role played by the numerical grid scale losses in our discussion of the nonlinear resolution of the counter-aligned case. Nature works without a grid, however, and it is not clear what scale ultimately intervenes at high wavenumbers in Hall protostellar disks. Moreover, it is now understood that there are important resolution questions for local shearing box simulations with no mean field, that have only recently been brought to light (Fromang et al. 2007; Lesur & Longaretti 2007). The important pioneering studies of Sano & Stone (2002a,b) merit a thoroughgoing follow-up on the much larger numerical grids that are currently available to numericists.

7 Alpha models of protostellar disks

7.1 Basic principles

We are now in a position to assemble a basic “ α model” of a protostellar disk. An important assumption in such models is that angular momentum and mechanical energy are transported radially through the disk. We therefore work with height and azimuthally integrated quantities. Under these conditions, Balbus & Papaloizou (1999) have shown how the fundamental equations of MHD lead directly to the α formalism.

The velocities consist of a mean part plus a fluctuation of zero mean. For the azimuthal velocity v_ϕ , the mean will be taken to be the Keplerian velocity

$$(R\Omega)^2 = \frac{GM}{R} \quad (151)$$

where M is the central mass and G is the Newtonian constant. For the radial velocity v_R , the mean motion corresponds to the very slow inward accretion drift, v_2 . The azimuthal fluctuation velocity u_ϕ is much less than $R\Omega$, whereas the radial fluctuation velocity u_R is much larger than the inward drift. These claims will shortly be quantified. The magnetic field Alfvén velocities are assumed to be of the same order as the u velocities, namely less than or of order the isothermal thermal sound speed c_S . To summarize,

$$v_2 \ll u, v_A, c_S \ll R\Omega. \quad (152)$$

Under steady-state conditions, the local disk accretion rate has a time averaged value of

$$\mu \equiv 2\pi R \langle \rho v_2 + \delta \rho \delta u_R \rangle \quad (153)$$

The height integrated form of this is $-\dot{M}$, the mass accretion rate:

$$\int \mu dz = -\dot{M}, \quad (154)$$

defined so that \dot{M} is a positive constant. A useful average of v_2 is the drift velocity defined by

$$v_d \equiv -\frac{\dot{M}}{2\pi R \Sigma} \quad (155)$$

where Σ is the disk surface density.

In steady-state, the height-integrated average radial angular momentum flux is proportional to $1/R$. In other words,

$$-\dot{M} R^2 \Omega + 2\pi \Sigma R W_{R\phi} = C/R \quad (156)$$

where $W_{R\phi}$ is defined by

$$\int \langle \rho u_R u_\phi - \rho v_{AR} v_{A\phi} \rangle dz = \Sigma W_{R\phi}, \quad (157)$$

and C is a constant to be determined. Traditionally, this has been done by asserting that $W_{R\phi}$ is proportional to the radial gradient of Ω , and at some point before the surface of the star is reached this must vanish. C is then determined at this point (Pringle 1981).

This approach now seems dated, and especially inappropriate for a proto-stellar disk in which the magnetic interactions between the disk and star may become quite complex at small radii. Instead, let us note that at small R , C/R must be very small if $\Sigma W_{R\phi}$ does not blow up, which seems quite reasonable. If C/R is very small at the inner edge of the disk, it is clearly negligible in the body of the disk, and we shall set $C = 0$ with the understanding that our solution should not be taken to the stellar surface. Then, (156) reduces to

$$v_d = -W_{R\phi}/R\Omega \quad (158)$$

showing that v_d (and v_2 which is of the same order) is very small compared with the u velocities. Alternatively,

$$\Sigma W_{R\phi} = \dot{M}R\Omega/2\pi. \quad (159)$$

This is a particularly useful result because $W_{R\phi}$ appears in the energy balance equation (140):

$$-\Sigma W_{R\phi} \frac{d\Omega}{d \ln R} = -\frac{d\Omega}{d \ln R} \int T_{R\phi} dz = \int \rho \mathcal{L} dz = 2\sigma T_{eff}^4. \quad (160)$$

In the last equality, we have equated the radiated energy per unit surface of the disk to twice the blackbody emissivity, the factor of two represents two radiating surfaces. Here σ is the Stefan-Boltzmann constant and T_{eff} is the effective blackbody surface temperature. Combining (159) and (160) then gives us

$$T_{eff}^4 = \frac{3GM\dot{M}}{8\sigma\pi R^3}, \quad (161)$$

relating the potentially observable disk surface temperature to the central mass, accretion rate, and radial location. The unknown turbulent stress parameter $W_{R\phi}$ has conveniently vanished!

Equation (161) is in a restricted sense “exact.” It is based on the assumption of time-steady conditions and thermal radiation, but it is independent of the explicit nature of the turbulence, as long as it is *local*. If these conditions are all met, equation (161) is a simple matter of energy conservation. To go beyond this result, something has to be said directly about $W_{R\phi}$. In α disk theory, that condition is

$$W_{R\phi} = \alpha c_S^2 \quad (162)$$

where α is assumed to be constant, but otherwise unknown. In contrast to the simple and plausible assumption that turbulence is locally dissipated, this “ α assumption” is far from obvious, and in a strict mathematical sense almost certainly wrong.

The original justification for this form of the stress tensor was based on the notion of hydrodynamical turbulence and the idea that the velocity fluctuations would be restricted to some fraction of the sound speed (fluctuations in excess would cause shocks). Shakura & Sunyaev (1973) consider magnetic stresses, however, and argue that they fit within the α formalism as well since Alfvén velocities in excess of c_S are also dynamically unlikely.

The real problem with the prescription (162) is that turbulence is just too complicated. Not only are long term averages very hard to define (a problem even for a result like equation [161]), it is also entirely possible that α could vary in a complex nonsystematic way by an order of magnitude or more from one part of the disk to another. The primary justification for (161) is that many scaling results are very insensitive to α . On dimensional grounds c_S is certainly an important local characteristic velocity, but it is not yet clear under what conditions a “background” magnetic field might also be providing a mean Alfvén velocity that limits or guides the turbulence.

Continuing with our α model, to link the surface temperature T_{eff} with the midplane temperature T (used in quantities like c_S) we need to introduce explicit vertical structure into the problem. The hydrostatic equilibrium equation is

$$-\frac{\partial P}{\partial z} = \frac{GM\rho z}{R^3} = \rho z \Omega^2 \quad (163)$$

This may be solved in conjunction with a detailed energy equation, but it seems best for illustrative purposes to note that this equation serves to provide a decent and very simple estimate of the disk half thickness, $H = c_S/\Omega$. The energy equation for simple vertical radiative diffusion defines the local radiative energy flux F_γ as:

$$F_\gamma = \frac{4\sigma}{3} \frac{dT^4}{d\tau}. \quad (164)$$

The “optical depth” τ is given by

$$d\tau = -\rho \kappa dz \quad (165)$$

where κ is the opacity of the disk (units: cross sectional area per unit mass). In the simplest possible model, F_γ is a constant given by σT_{eff}^4 , and the temperature at the midplane T is then

$$T^4 = \frac{3\tau T_{eff}^4}{4} \quad (166)$$

where τ is the optical depth from the outer disk surface to the midplane. (In this simplest of all possible models, we set $\tau = \kappa H$, where κ is evaluated at the midplane temperature.)

The surface temperature of a disk with a mass accretion rate of $10^{-8} M_\odot$ per year around a $1 M_\odot$ star is $85 R_{AU}^{-0.75} \text{K}$, where R_{AU} is the radius in astronomical unit. The midplane temperature T is a factor $\sim \tau^{1/4}$, larger, typically

a factor of 5 or so larger. At what value of T would we expect the ionization fraction to reach the critical value of 10^{-13} we found earlier, corresponding to a Lundquist number of unity? Put in somewhat different terms, is our model of self-sustaining MHD turbulence self-consistent?

To answer this, we need to address the physics of thermal ionization. In the low ionization regime we are working, thermal electrons are supplied by trace alkali elements, notably potassium, with an ionization potential of only 4.341 eV. Even this modest value corresponds to an effective temperature of 50,375 K, well above the range of 10^2 to 10^3 K we expects, and potassium will barely be ionized.

The equation governing the ionization fraction x is the Saha equation⁴. In this barely ionized regime, it may be written (Stone et al. 2000)

$$x^2 = as \left(\frac{2.4 \times 10^{15}}{n} \right) T^{3/2} \exp(-50375/T) \quad (167)$$

where a is the abundance of potassium (10^{-7}), n is the ratio of the dominant H_2 molecules, and s is a ratio of statistical weights, expected to be near unity. We may rewrite this equation as

$$T = -\frac{50375}{\ln X} \quad (168)$$

where

$$X = x^2 T^{-3/2} (n/2.4 \times 10^{15}) \quad (169)$$

With $x = 10^{-13}$ and the final density factor anything between 0.01 and 1 ($0.1 M_\odot$ of gas spread out over a region of 10 AU yields a density of about 5×10^{13} per cc) gives values for T close to 10^3 K, which we will adopt as a working number.

Therefore, a surface temperature of about 200 K is required to attain a midplane temperature of 10^3 , and thereby ensure a critical level of ionization. This is the key issue for MHD theories of protostellar disks: *beyond 0.2–0.3 AU, or $\sim 3 \times 10^{12}$ cm, the heat generated by the dissipation associated with local turbulence is insufficient to keep the the disk magnetically well-coupled.* Where and how do protostellar disks maintain good magnetic coupling?

8 Ionization Models of Protostellar Disks

8.1 Layered accretion

The above considerations indicate that dissipative heating from MHD turbulence self-consistently generates enough heat to maintain the minimum ther-

⁴ The discussion presented here presumes that the ionization reaction rates are sufficiently rapid to keep up with the dynamical time scales of any turbulence present. At low ionization fractions this breaks down, and a single temperature may not be enough to characterize the ionization state (Pneuman & Mitchell 1965).

mal levels of requisite ionization only with a few 0.1 AU. Nonthermal sources of ionization are therefore of great interest to protostellar disk theorists. The principle ionizing agents that have been studied are cosmic rays, X-rays, and radioactivity.

Gammie (1996) argued that the low density extended vertical layers of a protostellar disk would be exposed to an ionizing flux of interstellar cosmic rays. Just as in models of molecular cloud ionization, cosmic rays would maintain a minimal level of ionization in the upper disk layers. The range of the low energy galactic cosmic rays is about 100 g/cm^2 . This is much less than the disk column density at $\sim 1 \text{ AU}$ in generic solar nebula models, but can easily exceed the disk column at larger distances. If the level of ionization is high enough—and the Alfvén velocity v_A is *small* enough—the gas within the range of the cosmic rays will be MRI active. (The Alfvén velocity cannot be too large if disturbances are not to be stabilized by magnetic tension.) If these criteria are met, Gammie argued that turbulent accretion would proceed in the outer layers of protostellar disks, but that the midplane regions would remain laminar—in effect, a “dead zone.” This is the basis of the concept of layered accretion, which has become an important idea in protostellar disk modeling (Sano et al. 2000).

Gammie’s original construction was based on the assumptions that the accretion would occur in a layer of fixed column, and that α is constant. Taken together, these assumptions preclude a steady solution; the mass flux rate is not independent of position. Instead, matter is deposited from the outer regions into the disk’s inner regions where it builds up. At some point in this scenario the disk become gravitationally unstable, and it was speculated that an accretion outburst might occur, which was tentatively identified with FU Orionis behavior. On the other hand, there is no compelling argument (beyond mathematical convenience) to prevent the α parameter from adjusting with position, if this allows a relaxation to a time steady solution. The overarching layered accretion picture would, however, remain intact with this modification.

YSO’s are almost universally X-ray sources. Glassgold et al. (2000) noted that X-rays are potentially a far more powerful ionization source, even if attenuated, than galactic cosmic rays. These authors were thus the first to draw the link between X-ray observations and MRI activity of accretion disks. In the Glassgold et al. formulation, X-rays from a locally extended corona centered on the YSO irradiate the disk, and depending upon the chosen parameters, could provide the requisite ionization levels to lessen the extent of the dead zone or eliminate it altogether.

Whether the dead zone persists in the presence of X-ray irradiation or not depends, among other things, upon the model adopted for the disk structure. Generally, the so-called minimum mass solar nebula model (Weidenschilling 1977) is used. This is a reconstruction of the surface density distribution of the sun’s protostellar disk based on current planetary masses and compositions, and results in an $R^{-3/2}$ radial distribution. Fromang, Terquem, & Balbus (2002) suggested that ionization fractions should be calculated self-

consistently within the framework of accretion disk theory, which in general leads to a more shallow dependence of surface density with radius. By introducing accretion parameters whose values are free (α and the mass accretion rate \dot{M}), the range of parameter space increases, and the presence and extent of a dead zone becomes yet more model dependent.

The ionization fraction of a protostellar disk is a chemical problem, and in principle can involve a very complex, uncertain, molecular reaction network. Ilgner & Nelson (2004a) investigated the consequences for the dead zone of a considerably richer chemical network. While the quantitative details were sensitive to the chemistry, the qualitative structure was not. A dead zone is likely to be present in any plausible chemical network that has been studied up to the present, but there are conditions in which its extent can be very small or possibly even zero. Ilgner & Nelson (2004b), for example, made the interesting comment that flaring activity by the central star can have a significant effect on the extent of the dead zone.

8.2 Activity in the Dead Zone

Just because the dead zone is unable to host the MRI does not mean that it is well and truly dead. Fleming and Stone (2003) carried out numerical simulations in which the magnetically active upper layers of the disk coupled dynamically to the magnetically inert dead zone, resulting in a small but significant *Reynolds* stress, even though the Maxwell stress was zero. More recently, Turner & Sano (2008) suggested that the high resistivity characteristic of the dead zone is also a means for magnetic field to diffuse into this region from the active layers. Moreover, these authors point out, an active MRI region is not an absolute prerequisite for accretion: direct magnetic torques on much larger scales would also serve, and since they do not dissipate energy in a turbulent cascade, are much less costly to maintain.

Terquem (2008) constructed global models of protostellar disks based on the α prescription. The presence of a dead zone was surprisingly nondisruptive, provided that α was not too small. Steady solutions were found for values of α in the dead zone as small as 10^{-3} times the active zone value. In these models, the dead zone was thicker and more massive than its surroundings, but because of the relatively weak α scalings, by less than an order of magnitude. These results, taken as a whole, suggest that an embedded dead zone in a protostellar disk is not necessarily disruptive to the accretion process.

8.3 Dust

In all of our discussions, we have been most negligent by not mentioning the effects of small dust grains (e.g. Sano et al. 2000). Let us see crudely why this is so.

Consider a mass M of protostellar disk gas, of which $10^{-2}M$ is in the form of spherical dust grains of radius r_d . If the dust grains have a density of 3 gm

cm^{-3} , there are a total of

$$N_d = (10^{-2}/4\pi)(M/r_d^3)$$

grains in a volume V . The grains present a geometric cross section of $\sigma_d = \pi r_d^2$ (we ignore here the enhancement due to the induced charge [Drain & Sutin 1987]), and the electrons have an average radial velocity of some $v_e = 1.6(kT/m_e)^{1/2}$. The total dust recombination rate per unit volume is

$$(N_d/V)n_e\sigma_d v_e \quad (170)$$

where n_e is the number density of electrons. This should be compared with a typical dielectronic recombination rate of $\beta \equiv 8.7 \times 10^{-6} \text{ cm}^3 \text{ s}^{-1}$ (Glassgold, Lucas, & Omont 1986; Gammie 1996). If x is ionization fraction, then the ratio of dielectronic recombination to dust recombination is

$$\frac{n_e \beta V}{N_d \sigma_d v_e} \sim \left(\frac{x}{10^{-13}} \right) \frac{r_d}{T} \quad (171)$$

where r_d is in cm and T in K. For small dust grains ($\sim 10^{-5} \text{ cm}$) this ratio is $\ll 1$ and dust recombination is overwhelmingly important. But the relative importance of the grains diminishes as they grow in size, especially in the cooler portions of the disk. Sano et al. (2000) concluded that in their model (based on cosmic ray ionization) a typical protostellar disk would be MRI stable inside of about 20 AU if small grain are present, except for the innermost regions which would be thermally ionized.

The detailed effects of dust grains have been considered by many authors; a very good review and list of references is given in the recent paper of Salmeron & Wardle (2008). These authors have considered the vertical structure of the magnetic coupling in a minimum mass solar nebula model, including the effects of Hall electromotive forces and ambipolar diffusion. They employed a sophisticated chemical network (Nishi et al. 1991), and small dust grains ranging in radius from 0.1 to 3 microns. At the fiducial locations of 5 and 10 AU, Salmeron & Wardle found good magnetic coupling over an impressive range of magnetic field strengths provided for the 3 micron grains, a considerable improvement from our naive estimate above. For the 0.1 micron grains, the magnetic coupling dropped sharply (as did MRI growth rates), and was restricted to higher elevations above the disk plane.

9 Summary

Protostellar disks are gaseous systems dynamically dominated by Keplerian rotation. These disks are accreting onto the central protostar, so there must be a source of enhanced angular momentum transport present. In the early stages of the disk's life, this enhanced transport may well be due to self-gravity, with density waves largely responsible for moving angular momentum

outwards. Once the disk becomes observable, its mass is below the minimum needed to sustain self-gravitational spiral wave transport. The only established mechanism able to sustain high levels of angular momentum transport is MHD turbulence produced by the magnetorotational instability, or MRI.

The dynamical effects of magnetic fields on protostellar disks depends crucially on the degree of ionization, more specifically on the electron fraction, that is present. This means that there is a direct link between the gross dynamical behavior of a protostellar disk and its detailed chemical profile. In principle, the full multifluid nature of the disk gas — neutrals, ions, electrons, and dust grains — must be grappled with at some level to elicit and understand the disk structure. Fortunately, only a very small ionization level is needed to couple the charged and neutral components, leaving (in essence), a single magnetized fluid. An electron fraction of 10^{-13} is a typical fiducial number for the threshold of magnetic coupling at 1 AU in the T Tauri phase of the solar nebula: roughly one electron per cubic millimeter of disk gas! Unfortunately, however, on scales of 10's of AU, near the dense midplane the ionization of protostellar disks may not even rise to this minimal level of ionization. At the boundary between good and poor magnetic coupling, the MHD processes are complex, not well-understood, and the domain of ongoing inquiry.

In regions of the disk where the magnetic coupling is sufficient, the combination of a weak (subthermal) magnetic field and Keplerian rotation leads to the MRI. In ideal MHD, the magnetic field behaves as though it were frozen into the conducting gas, and the presence of differential rotation produces azimuthal magnetic field from a radial magnetic field. But this is not all that happens. If one tries to simulate this simple process on a computer, the laminar flow breaks down into a turbulent mess, even if the field is very weak.

The reason for this is due to another classical property of magnetic fields, that the force exerted by the lines of force on the background gaseous fluid is the sum of a pressure-like term and a tension-like term. It is the latter tension-like term that is important for an understanding of the MRI. When two nearby fluid elements are moved apart, even if only because of a random perturbation, the magnetic tension force acts precisely like a spring coupling the two masses. The fluid element on the inside rotates faster than the element on the outside, and tries to speed it up. The outer element, in acquiring angular momentum from the inner element, finds itself too well-endowed, and spirals outward toward a higher orbit where its excess angular momentum can be accommodated. On the other hand, the inner element, having lost angular momentum, finds itself at a deficit and must drop to a lower angular momentum orbit. This separation stretches the field lines that couple the elements, the magnetic tension goes up, and the process runs away. This is the MRI. Notice that angular momentum transport is at the very core of the linear instability, rather than the result of some sort of nonlinear mixing process.

But the MRI does, in fact, lead to rapid turbulent mixing of the disk gas, as outwardly moving and inwardly moving fluid elements encounter one another and dissipate their energy. If this happens locally, then the ingredients for a classical α model of accretion are present. To the extent that many disk features are insensitive to, or independent of, the precise rms level of the disk fluctuations, these models can be of some practical utility. The classical formula relating the disk emissivity to radius R is the most important instance of this.

In real protostellar disks, the level of ionization present is such that there are significant departures from the behavior of an ideal (perfectly conducting) MHD gas. In principle, the gas can become completely decoupled from the magnetic field, and go over to a hydrodynamic system. Less dramatically, the current-bearing electrons can acquire a velocity significantly different from the dominant neutrals, since the former need to maintain a current density and magnetic field even as charged species become rarer and rarer. When this happens, the field lines are no longer frozen into the motion of the bulk of the (neutral) disk gas, they are frozen into the electrons and the distinction becomes important. In the so-called Hall regime, the ions and neutrals move together distinctly from the electrons, and the ion motion relative to the electron-following field lines induces an additional electromotive force into the gas, beyond the self-induction responsible for simple field line freezing. Hall MHD is likely to be important in protostellar disks on scales of AU to 10's of AU scales.

Dust grains are an ever present complication for MHD disk modeling. Typically, solids make up about 1 percent of the mass of interstellar gas. An interstellar population of small grains (radius $\sim 10^{-5}$ cm) with a total mass fraction of a percent would present an enormous collecting area of would-be gas phase electrons. Putting charges on the grains effectively removes them because of the low mobility of the grains. Determining the disposition of the dust grains is thus a necessity for constructing MHD disk models. As the disk evolves, so too do the dust grains. They grow in size as they agglomerate, and they tend to settle towards the midplane, if they are not stirred by turbulence or some other dynamical process. Larger dust grains are much less efficient in removing gas phase electrons than are smaller grains, where “larger” means growth in radius of an order of magnitude or more.

Where does all of this leave us? Clearly, we are a long way from framing a picture of protostellar disk evolution at the level of, say, classical stellar evolution. But for all of the gaping uncertainties, a useful zeroth order MHD-based picture of a protostellar disk in its T Tauri phase can be cobbled together:

Inside of about 0.3 AU, a protostellar disk will be thermally ionized by direct exposure to the central source and self-consistently by the dissipation of MHD turbulence. We may expect vigorous MRI induced MHD turbulence in this zone.

On scales of AU's, thermal ionization is no longer adequate to maintain the requisite levels of ionization to ensure MHD coupling, at least not near the

high density midplane. Depending upon the X-ray luminosity of the central star, the spectrum of dust grains and the abundance of gas phase potassium and sodium (electron donors), there could be good coupling at low density, higher elevation disk altitudes. The “dead zone” at lower altitudes need not be devoid of all transport; density waves from an adjacent active layer or large scale magnetic torques would each contribute their own stresses. It is even possible that some degree of coupling could be maintained down to the disk midplane, though this depends strongly on modeling assumptions. It seems likely, however, that on a scale an AU to tens of AU’s, the level of MHD turbulence will be far less than in the disk’s innermost regions.

Finally, on scales larger than tens of AU’s, the falling density lengthens recombination times and suggests a return to good MHD coupling. The typical size of a disk is many hundred AU’s so that the dead zone is small when viewed globally, and may not have much of an impact in the overall accretion process (Terquem 2008). Perhaps, however, it is not coincidence that the MHD “quiet zone” coincides with the region of planet formation in the solar nebula and (more speculatively) in other protostellar disks.

Acknowledgements

I would like to thank my collaborators over past several years with whom and from whom I have learned much of what I know of the subject of protostellar disks: O. Blaes, S. Fromang, C. Gammie, J. Hawley, J. Stone, and C. Terquem. I would also like to thank specifically P. Garcia and the referee for their important respective comments on the H-H^+ reaction rates and ionization equilibrium. This work was supported by a Chaire d’Excellence from the French Ministry of Higher Education, and a grant from the Conseil Régional de l’Île de France.

References

1. Balbus, S. A. 2001, *ApJ*, 562, 909
2. Balbus, S. A. 2003, *ARAA*, 41, 555
3. Balbus, S. A., & Hawley, J. F. 1991, *ApJ*, 376, 214
4. Balbus, S. A., & Hawley, J. F. 1992, *ApJ*, 392, 662
5. Balbus, S. A., & Hawley, J. F. 1998, *Rev. Mod. Phys.*, 70, 1
6. Balbus, S. A., & Hawley, J. F. 2000, in *From Dust to Terrestrial Planets*, eds. W. Benz, R. Kallenbach, G. Lugmaier, and F. Podosek, *ISSI Space Sciences Series* No. 9, *Space Science Reviews*, 92, 39
7. Balbus, S. A., & Papaloizou, J. C. B. 1999, *ApJ*, 521, 650
8. Balbus, S. A., & Terquem, C. 2001, *ApJ*, 552, 235
9. Blaes, O. M., & Balbus, S. A. 1994, *ApJ*, 421, 163
10. Draine, B. T., Roberge, W. G., & Dalgarno A. 1983, *ApJ*, 264, 485
11. Draine, B. T., & Sutin, B. 1987, *ApJ*, 320, 803

12. Fleming, T., & Stone J. M. 2003, *ApJ*, 585, 908
13. Fleming, T., & Stone J. M., & Hawley, J. F. 2000, *ApJ*, 530, 464
14. Fromang, S., Terquem C., & Balbus, S. A. 2002, *MNRAS*, 329, 18
15. Fromang, S., Papaloizou, J., Lesur, G., & Heinemann, T. 2007, *A&A*, 476, 1123
16. Gammie, C. F. 1996, *ApJ*, 457, 355
17. Glassgold, A. E., Feigelson, E. D., Montmerle, T. 2000, in *Protostars and Planets IV*, V. Mannings, A. P. Boss, S. Russel eds. (U. Arizona Press: Tucson), 429
18. Glassgold, A. E., Krstić, P. S., & Schultz, D. R. 2005, *ApJ*, 621, 808
19. Glassgold, A. E., Lucas, R., & Omont, A. 1986, *AA*, 157, 35
20. Hartmann, L. 1998, *Accretion Processes in Star Formation* (Cambridge: Cambridge University Press)
21. Ilgner, M., & Nelson, R. P. 2004a, *AA*, 445, 205
22. Ilgner, M., & Nelson, R. P. 2004b, *AA*, 455, 731
23. Jackson, J. D. 1975, *Classical Electrodynamics* (2nd. ed.) (Wiley: New York)
24. Ji, H., Burin, M., Schartman, E., & Goodman, J. 2006, *Nature* 444, 343
25. Kunz, M. W., & Balbus S. A. 2004, *MNRAS*, 348, 355
26. Landau, L. D., & Lifshitz, E. M. 1959, *Fluid Mechanics* (London: Pergamon Press)
27. Lesur, G., & Longaretti, P.-Y. 2007, *MNRAS*, 378, 1471
28. Lin, D. N. C., & Papaloizou, J. 1980, *MNRAS*, 191, 37
29. Lin, D. N. C., & Papaloizou, J. C. B. 1996, *ARAA*, 34, 703
30. Lubow, S. H., Papaloizou, J. C. B., & Pringle, J. E. 1994, *MNRAS*, 267, 235
31. Moffatt, H. K., Kida, S., & Ohkitani K. 1994, *J. Fluid Mech.*, 259, 241
32. Nishi, R., Nakano, T., Umebayashi, T. 1991, *ApJ*, 368, 181
33. Papaloizou, J. C. B., & Lin, D. N. C. 1995, *ARAA*, 33, 505
34. Pneman, G. W., & Mitchell, T. P. 1965, *Icarus*, 4, 494
35. Pringle, J. E, 1981, *ARAA*, 19, 137
36. Reipurth, B., Jewitt, D., & Keil, K. (eds.) 2007, *Protostars and Planets V*, (U. Arizona Press: Tucson)
37. Richard, D. 2001, Ph. D. Thesis, Univ. Paris 7
38. Sano, T., Miyama, S. M., Umebayashi, T., & Nakano, T. 2000, *ApJ*, 543, 486
39. Sano, T., & Stone, J. M. 2002a, *ApJ*, 570, 314
40. Sano, T., & Stone, J. M. 2002b, *ApJ*, 577, 534
41. Salmeron, R., & Wardle M. 2008, *MNRAS* 388, 1223
42. Shakura, N. I., & Sunyaev R. A. 1973, *AA*, 24, 337
43. Stone, J. M., & Balbus, S. A. 1996, *ApJ*, 464, 364
44. Stone, J. M., Gammie, C. F., Balbus, S. A., & Hawley, J. F. 2000, in *Protostars and Planets IV*, eds. V. Mannings, A. Boss, and S. Russell, *Space Science Reviews*, (U. Arizona Press: Tucson), 589
45. Terquem, C. 2008, *ApJ*, in press.
46. Turner, N. J., & Sano, T. 2008, *ApJ*, 679L, 131
47. Velikhov, E. P. 1959, *Sov. Phys. JETP*, 36, 995
48. Wardle, M. 1999, *MNRAS*, 307, 849
49. Weidenschilling, S. J. 1977, *Astrophys. & Space Science*, 51, 153
50. Zel'dovich, Ya. B., Ruzmaikin, A. A., & Sokoloff, D. D. 1983, *Magnetic Fields in Astrophysics* (Gordon and Breach: New York)

Appendix A

Begin with equation (16):

$$\frac{\mathbf{J}}{c} \times \mathbf{B} = \mathbf{p}_{In} + \mathbf{p}_{en} \quad (172)$$

where the right side of this equation is

$$n_I n \mu_{In} \langle \sigma_{nI} w_{nI} \rangle (\mathbf{v}_I - \mathbf{v}) + n_e n m_e \langle \sigma_{ne} w_{ne} \rangle [(\mathbf{v}_e - \mathbf{v}_I) + (\mathbf{v}_I - \mathbf{v})] \quad (173)$$

In what follows, we will often need an estimate of the ratio of the ion and electron collision rates. Following our discussion in section 3, we will assume that

$$\frac{\langle \sigma_{nI} w_{nI} \rangle}{\langle \sigma_{ne} w_{ne} \rangle} = \left(\frac{m_e}{\epsilon \mu_{In}} \right)^{1/2} \quad (174)$$

where $\epsilon < 1$ is inserted because the electron-neutral cross section is geometrical, while the ion-neutral collision cross section is larger than geometrical.

In (173), the first $(\mathbf{v}_I - \mathbf{v})$ term dominates the last by a factor of order $(\mu_{In}/\epsilon m_e)^{1/2}$. Hence,

$$\mathbf{p}_{In} = \frac{\mathbf{J}}{c} \times \mathbf{B} - \mathbf{p}_{en} = \frac{\mathbf{J}}{c} \times \mathbf{B} + \frac{n m_e}{e} \langle \sigma_{ne} w_{ne} \rangle \mathbf{J} \quad (175)$$

With the help of equation (10), this implies

$$\mathbf{v}_I - \mathbf{v} = \frac{\mathbf{J} \times \mathbf{B}}{c \gamma \rho \rho_I} + \underbrace{\sqrt{\frac{\epsilon m_e}{\mu_{nI}}} Z (\mathbf{v}_I - \mathbf{v}_e)}_A. \quad (176)$$

We have marked the term with an “A” for future reference.

Recall equation (20):

$$\mathbf{E} + \frac{1}{c} [\mathbf{v} + (\mathbf{v}_e - \mathbf{v}_I) + (\mathbf{v}_I - \mathbf{v})] \times \mathbf{B} + \frac{m_e \nu_{en}}{e} [(\mathbf{v}_e - \mathbf{v}_I) + (\mathbf{v}_I - \mathbf{v})] = 0. \quad (177)$$

In substituting equation (176) for $\mathbf{v}_I - \mathbf{v}$ in the above, we may always drop the “A” term, since it is small compared with $\mathbf{v}_e - \mathbf{v}_I$.

Proceeding with the above substitution leads to

$$\mathbf{E} + \frac{\mathbf{v}}{c} \times \mathbf{B} - \frac{\mathbf{J} \times \mathbf{B}}{e n_e c} \left[1 - \underbrace{\frac{m_e \nu_{en} n_e}{\gamma \rho \rho_I}}_B \right] + \frac{(\mathbf{J} \times \mathbf{B}) \times \mathbf{B}}{c^2 \gamma \rho \rho_I} - \frac{\mathbf{J}}{\sigma_{cond}} = 0, \quad (178)$$

where σ_{cond} is defined in equation (24). The “B” term in the above equation may now clearly be dropped: it is of order $Z(\epsilon m_e \mu_{nI})^{1/2}/m_n$. This leads to

$$\mathbf{E} + \frac{\mathbf{v}}{c} \times \mathbf{B} - \frac{\mathbf{J} \times \mathbf{B}}{e n_e c} + \frac{(\mathbf{J} \times \mathbf{B}) \times \mathbf{B}}{c^2 \gamma \rho \rho_I} - \frac{\mathbf{J}}{\sigma_{cond}} = 0, \quad (179)$$

which is precisely equation (23) in the text.

Appendix B:

To estimate the order of departures from charge neutrality or the displacement currents, we will assume that the ∇ operator is $\sim 1/L$, where L is a characteristic length scale of the flow, and $\partial/\partial t$ is $\sim v/L$, where v is a characteristic velocity (say, the largest of the neutral, ion, or electron velocities).

For the electric field, we take $E \sim vB/c$, since we are only interested in problems where the inductive terms are comparable to, or in larger than, the resistive damping. Then

$$\nabla \cdot \mathbf{E} \sim E/L \sim vB/Lc \sim 4\pi vJ/c^2 \sim 4\pi en_e v^2/c^2 \quad (180)$$

Hence, the divergence of \mathbf{E} is of order v^2/c^2 times the electron charge density (at most, since we assumed $v_I - v_e \sim v$ in the above). It may thus be ignored.

For the displacement current, the demonstration is almost a matter of direct inspection. If we take the curl of equation (29) and use equation (27), then the on the right side of (29) the first term is $\sim B/L^2$, while the second, displacement, term is $\sim v^2 B/c^2$. It may thus be ignored.