1 INTRODUCTION

The mechanism by which angular momentum is lost by the material in accretion discs is a classic problem in astrophysics. While gravitational torques associated with a binary companion (e.g. Sawada et al. 1987) or gravitational instabilities in massive discs (Adams, Ruden & Shu 1989) may play an occasional role, magnetic tension appears likely to be the most common source of the torque required to remove angular momentum from the material in the disc and permit it to spiral towards the central object.

Magnetically mediated accretion disc models fall into two broad paradigms. In disc–jet models, a large-scale poloidal field threads the disc, and the magnetic tension of the field lines anchored to the disc accelerates material outwards and upwards along the field lines away from the disc surfaces to form a collimated bipolar jet (Blandford & Payne 1982). The net effect is to transfer angular momentum from the disc material to the accelerated outflow. In the absence of an outflow, the disc may still be braked through torsional magnetic coupling to an extended envelope or surrounding cloud (e.g. Krasnopolsky & Königl 2002).

In models invoking magnetically driven turbulence, on the other hand, tension in a weak magnetic field virulently destabilizes the orbital shear flow in the disc via the magnetorotational instability (MRI; Velikhov 1959; Chandrasekhar 1960; Balbus & Hawley 1991), driving magnetohydrodynamic (MHD) turbulence that transports angular momentum radially outwards (e.g. Balbus & Hawley 1992; Stone et al. 1996). Generally, one expects that the jet models apply when the magnetic and thermal pressures are comparable at the disc mid-plane, while substantially subthermal magnetic fields are subject to the MRI and are incapable of driving strong jets. It is conceivable that these two transport mechanisms may operate at the same radius for intermediate strength fields, with the MRI operating for a range of z adjacent to the mid-plane of the disc and vertical transport dominating at greater heights (e.g. Salmeron, Königl & Wardle 2007). Similarly, as the field strength is expected to decrease with radius, there could also be an inner region where jets dominate, with the MRI operating at larger radii (Combet & Ferreira 2008).

In this paper, we focus on the efficacy of the MRI-driven turbulence in protostellar/protoplanetary discs. In this context, apart from transporting angular momentum, MRI-driven turbulence selectively heats the disc, effectively mixes chemical species (e.g. Semenov, Wiebe & Henning 2006) and affects the transport, aggregation and sedimentation of dust grains (Johansen & Klahr 2005; Fromang & Papaloizou 2006; Turner et al. 2006), the assembly blocks of planetesimals in the 'core accretion' scenario of planet formation. Furthermore, the effective viscosity of the disc associated with MHD turbulence determines the ability of protoplanets to open gaps in the disc and also contributes to the delicate imbalance of gravitational torques that determines the rate and direction of planetary migration (e.g. Matsumura & Pudritz 2003; Nelson & Papaloizou 2004; Johnson, Goodman & Menou 2006).

It is of great importance, therefore, to determine the location, extent and behaviour of the magnetically active regions of protoplanetary discs, which are restricted by the very low fractional ionization beyond ∼0.1 au from the central star, where thermal ionization is ineffective (Gammie 1996). The dominant sources of ionization – X-rays or ultraviolet radiation from the forming star and possibly interstellar cosmic rays – are unable to penetrate to the disc mid-plane except perhaps at larger radii where the column density is low. At intermediate radii, a very low level of ionization is maintained by cosmic rays or, if these are absent, the decay of radioactive elements mixed with the gas (Semenov, Wiebe & Henning 2004; Glassgold et al. 2005). MRI-driven MHD turbulence is therefore thought to be restricted to the surface layers (Gammie 1996) over the range ∼0.3–20 au where shielding of external ionization sources results in very low levels of ionization.

Assessments of the location and extent of magnetically active zones (or equivalently, their complementary magnetic 'dead zones') typically invoke only damping of the linear MRI by Ohm diffusion (e.g. Hayashi 1981; Jin 1996; Fromang, Terquem & Balbus 2002; Matsumura & Pudritz 2003; Ilgner & Nelson 2006; Turner & Sano 2008). However, Ohm diffusion is the dominant magnetic diffusivity only at the very high densities achieved close to the disc mid-plane within 1 au of the central star. Instead, Hall diffusion[1] dominates at the intermediate mid-plane densities between 1 and 30 au (Wardle & Ng 1999; Sano & Stone 2002a), and ambipolar diffusion beyond about 30 au. Note, however, that as the density decreases towards the disc surfaces, different diffusivities may dominate in different layers. The neglect of ambipolar diffusion in assessing magnetic activity is less drastic, as it dominates near the disc surface where the fractional ionization is relatively high and magnetic diffusivity may be insufficient to stabilize the MRI (e.g. Perez-Becker & Chiang 2011).

Hall diffusion provides a dissipation-free pathway for the magnetic field that either enables the MRI to grow despite the damping effect of Ohm or ambipolar diffusion or suppresses it entirely (Wardle 1999; Balbus & Terquem 2001), depending on the orientation of the magnetic field. Based on the linear growth rates, Hall diffusion may drastically extend or restrict the reach of magnetic activity in protoplanetary discs (Wardle 2007) *and likely modifies the transport and dissipative properties of the resulting turbulence.* However, pioneering shearing-box calculations including both Ohm and Hall diffusion did not appear to confirm this expectation, suggesting instead that Hall diffusion was unable to counter the damping effect of Ohm diffusion (Sano & Stone 2002a,b). As a result, subsequent evaluations of the extent of dead zones in protoplanetary discs are based either on the damping effects of Ohm diffusion (e.g. Turner & Sano 2008; Turner & Drake 2009; Kretke & Lin 2010) or on Ohm and ambipolar diffusion (Bai 2011; Perez-Becker & Chiang 2011).

The apparent failure of fully developed turbulence to respect the underlying linear instability in the Hall-diffusion-dominated limit is surprising, as it does so in the Ohm (e.g. Turner & Sano 2008) and ambipolar diffusion regimes (Bai & Stone 2011). However, closer inspection of the parametrization of the diffusive effects adopted by Sano & Stone (2002a,b) reveals that their simulations did not probe the important regime in which (a) Hall diffusion dominates Ohm diffusion *and* (b) both dominate inductive effects. Indeed, none of the results of Sano & Stone (2002a,b) conflicts with expectations based on the linear analysis. The confusion that has arisen can largely be attributed to the adoption of a magnetic Reynolds number $Re_{\rm M}$ and a Hall parameter X to characterize the magnitudes of Ohm and Hall diffusion, respectively. Then the criteria for Ohm and Hall diffusion to dominate the inductive term are $Re_{\rm M} < 1$ and $X > 2$, respectively. Crucially, the ratio of Hall to Ohm diffusion is then $(1/2)XRe_{\rm M}$. This is never $\gg 1$ when $Re_{\rm M} < 1$ in the simulations, apart from the zero-net-flux simulations S14–S16 of Sano & Stone (2002b) in which Hall diffusion suppresses the MRI in the $B_z < 0$ regions, as one might expect based on the linear analysis.

[1] Despite the lack of dissipation associated with Hall drift, we adopt the term 'Hall diffusion' to maintain consistency with the notion of a tensor diffusivity appearing in the induction equation.

The purpose of this paper is to emphasize the importance of Hall diffusion for MRI-driven turbulence in protoplanetary discs. We do this by considering the simplest example of the MRI – perturbations with wave vectors parallel to an initially vertical magnetic field – and demonstrating that the column density of the MRI-unstable region at 1 au in the minimum-mass solar nebula (MMSN) increases or decreases (depending on the sign of B_z) by an order of magnitude when Hall diffusion is included (see Fig. 14). We begin by describing the field-line drifts induced by ambipolar, Hall and Ohm diffusion in Section 2 and give a qualitative discussion of their effect on the MRI. In Section 3, we present an overview of the results of a linear analysis of the MRI in the presence of diffusion and give a coherent survey of its dependence on Pedersen (i.e. Ohm+ambipolar) and Hall diffusion by considering contours of the growth rate and wavenumber of the fastest growing mode and the range of MRI-unstable wavenumbers in Section 4. We then compare previous notation with the diffusivity notation we now favour and discuss the numerical simulations of Sano & Stone (2002a,b). In Section 5, we estimate the column density of the MRI-unstable layer at 1 au in the MMSN using the criterion that there should exist a local unstable MRI mode with wavenumber satisfying $kh > 1$. The results are dramatic, indicating that Hall diffusion increases the column density of the MRI-unstable surface layers by an order of magnitude depending on whether B_z is positive or negative, respectively. Finally, we summarize our results and conclusions in Section 6.

2 MAGNETIC DIFFUSION AND THE MAGNETOROTATIONAL INSTABILITY

2.1 Magnetic diffusion

The magnetic field evolves according to the induction equation,

$$\frac{\partial \boldsymbol{B}}{\partial t} = \nabla\times(\boldsymbol{v}\times\boldsymbol{B}) - c\nabla\times\boldsymbol{E}', \tag{1}$$

where $\boldsymbol{E}'$ is the electric field in the local instantaneous rest frame of the fluid, which satisfies the generalized Ohm's law

$$c\boldsymbol{E}' = \eta_{\rm A}(\nabla\times\boldsymbol{B})_\perp + \eta_{\rm H}(\nabla\times\boldsymbol{B})\times\hat{\boldsymbol{B}} + \eta(\nabla\times\boldsymbol{B}), \tag{2}$$

where we use subscripts $\parallel$ and $\perp$ to refer to the orientation with respect to the local magnetic field and $\eta_{\rm A}$, $\eta_{\rm H}$ and η are the ambipolar, Hall and Ohm diffusivities, respectively. It will prove useful to recast the induction equation in a form that makes explicit the drift of the magnetic field through the fluid, i.e.

$$\frac{\partial \boldsymbol{B}}{\partial t} = \nabla\times\left[(\boldsymbol{v}+\boldsymbol{v}_{\rm B})\times\boldsymbol{B} - \eta(\nabla\times\boldsymbol{B})_\parallel\right], \tag{3}$$

where

$$\boldsymbol{v}_B = c\,\frac{\boldsymbol{E}'\times\boldsymbol{B}}{B^2} = \eta_{\rm P}\,\frac{(\nabla\times\boldsymbol{B})_\perp\times\hat{\boldsymbol{B}}}{B} - \eta_{\rm H}\,\frac{(\nabla\times\boldsymbol{B})_\perp}{B} \tag{4}$$

and the Pedersen diffusivity $\eta_{\rm P}$ is given by

$$\eta_{\rm P} = \eta_{\rm A} + \eta. \tag{5}$$

This form makes explicit the role of the diffusivities in transporting flux and emphasizes that only Ohm diffusion is capable of destroying magnetic flux [i.e. via the $\eta(\nabla\times\boldsymbol{B})_\parallel$ term].

The diffusivities are determined by the abundances of charged particles and their collision cross-sections with the neutrals (e.g. Cowling 1957; Wardle & Ng 1999), and – when the fractional ionization is not too small – with each other. In general, η, $\eta_{\rm A}$ and $\eta_{\rm P}$ are all positive, whereas $\eta_{\rm H}$ is positive or negative depending on whether positively or negatively charged species, respectively, are on average more decoupled from the magnetic field by collisions. For example, for a weakly ionized ion–electron–neutral plasma, we obtain

$$\eta = \frac{c^2 m_{\rm e}\gamma_{\rm e}\rho}{4\pi e^2 n_{\rm e}}, \tag{6}$$

$$\eta_{\rm A} = \frac{B^2}{4\pi\,\gamma_{\rm i}\,\rho\rho_{\rm i}} \tag{7}$$

and

$$\eta_{\rm H} = \frac{c\,B}{4\pi\,e\,n_{\rm e}} \tag{8}$$

(e.g. Königl 1989; Balbus & Terquem 2001). The diffusivities exhibit more complex dependencies on density, field strength and electron abundance when charged grains are present. In this case, the Hall diffusivity is negative at relatively low densities when neutral collisions decouple grains from the magnetic field, but not the electrons and ions, and is positive at higher densities when ions are also decoupled from the magnetic field (e.g. Wardle & Ng 1999). The diffusivities exhibit significant spatial gradients if the ionization rate and gas density do so, as in protoplanetary discs (e.g. Wardle 2007). Fortunately, for a linear analysis, the response of the diffusivities to perturbations is not needed as $\nabla\times\boldsymbol{B}$ vanishes in the initial equilibrium state. Thus in our local analysis, we can regard the diffusivities as specified constants, but we shall return to them when considering the application to protoplanetary discs in Section 5.

2.2 Effect on the MRI

Before deriving the dispersion relation for the MRI in diffusive media, we first give an intuitive, physical description of the effect of magnetic diffusion on its growth and properties.

We first examine the growth of the instability in a Keplerian accretion disc under ideal-MHD conditions and assuming an initially vertical magnetic field (see Fig. 1, top-left panel). We consider, in particular, the situation where alternate layers of fluid are displaced from their equilibrium position, such that point 1 moves radially inwards (and forwards in azimuth), while point 2 is shifted radially outwards (and backwards in azimuth). In this case, the magnetic field is buckled because the field lines are frozen into the fluid (top-right panel). The deformation of the lines creates magnetic tension forces which are directed outwards and backwards in azimuth (at point 1) or inwards and forwards in azimuth (at point 2), as depicted in the lower panel of Fig. 1. At point 1 then the magnetic tension provides some radial support against gravity and supplies a negative torque. This means that the fluid element at point 1 will lose angular momentum and spiral inwards. The opposite takes place at point 2, with the fluid element there spiralling outwards. As a result, the magnetic field buckling and associated tension increases, leading to runaway growth. The fastest growing mode has growth rate $(3/4)\Omega$ on scales with wavenumbers $k \sim \Omega/v_{\rm A}$, where $\Omega = v_{\rm K}/r$ is the Keplerian frequency and $v_{\rm A}$ is the Alfvén speed (Balbus & Hawley 1991).

Magnetic diffusion modifies the above picture by allowing slippage between the field lines and the fluid, so that there is no longer a direct connection between the relative displacement of the fluid layers and the buckling of the magnetic field lines. One might expect that this effect should always reduce the instability because the displacement of the field would lag that of the fluid. However,

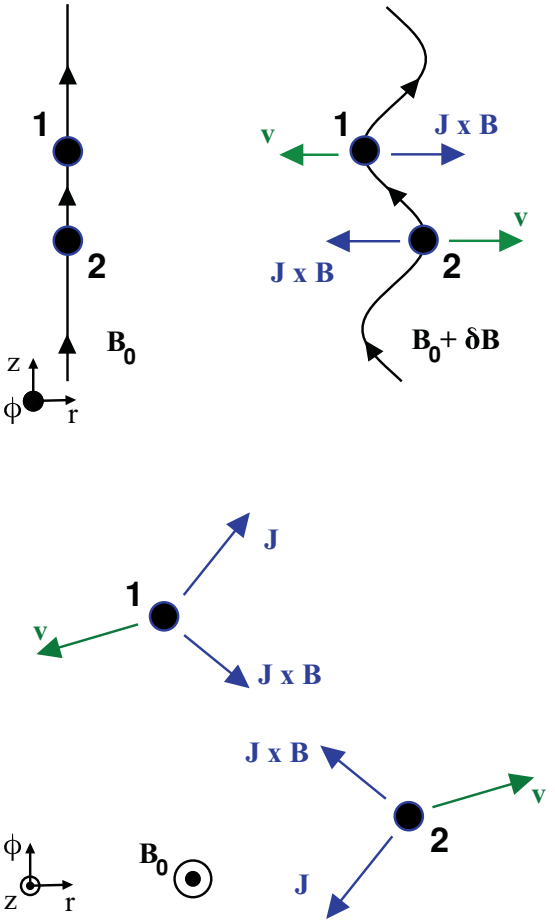

Figure 1. A sketch of the development of the MRI in an initially vertical magnetic field B_0 embedded in an unstratified, Keplerian disc for a vertical wave vector. Top panel shows a poloidal view of (left) the initial field configuration and (right) the perturbed field and associated current density ($\boldsymbol{J}$), Lorentz force ($\boldsymbol{J}\times\boldsymbol{B}$) and deviation from local Keplerian velocity. The lower panel shows in plan view the perturbed fluid elements at points 1 and 2. At point 1, the magnetic tension provides some radial support against gravity and supplies a negative torque. As a result, the material here is sub-Keplerian and moving inwards, as indicated by the green vector which indicates the departure from Keplerian rotation. The field transfers the angular momentum to the fluid at locations such as 2 where the field is buckled outwards, and the fluid there spirals out. The differential radial motion of the fluid that is losing and gaining angular momentum at 1 and 2, respectively, enhances the buckling, leading to runaway growth.

this intuition is based on the limit of Ohm or ambipolar diffusion, in which the drift is in the direction of the magnetic stresses and tends to straighten up field lines. Hall diffusion, by contrast, creates a drift orthogonal to the tension forces and may, therefore, enhance or suppress the radial buckling depending on the situation. It is this feature that gives Hall diffusion its unique properties and that we aim to explain here.

In the ambipolar diffusion limit, the magnetic field is frozen into the ions and electrons, which drift together through the neutral component of the fluid. Collisions with the neutrals then transmit magnetic stresses to the bulk of the gas. As the fractional ionization is low, the ion and electron inertia and thermal pressure are negligible, and the drift velocity of the ions, electrons and field is determined by the balance between the Lorentz force on the ions

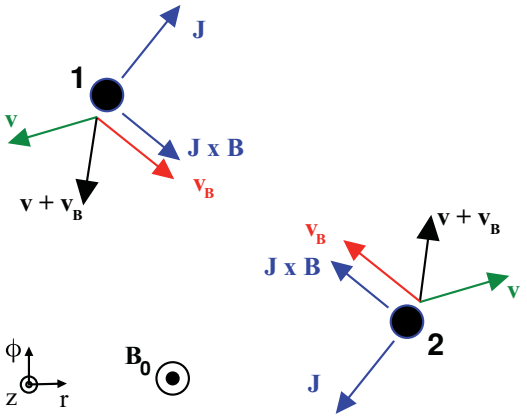

Figure 2. Plan view of the effect of ambipolar or Ohm diffusion on the development of the MRI. The current density and associated magnetic stress on the fluid at points 1 and 2 of Fig. 1 are indicated in blue. Because of finite diffusivity, the field lines drift through the fluid at velocity $\boldsymbol{v}_B$ (*red* vectors) in the direction parallel to $\boldsymbol{J}\times\boldsymbol{B}$ (see equation 9). The net drift of the field line with respect to Keplerian rotation, $\boldsymbol{v}+\boldsymbol{v}_B$ is indicated by the *black* vector. In this case, the net drift is marginally inwards at point 1. At point 2, the current density and magnetic stresses are reversed, so the field drifts in the opposite direction. As a result, the field lines drift inwards and outwards at reduced rates at points 1 and 2, respectively, and the instability proceeds more slowly than in ideal MHD.

and the collisional drag with the neutrals,

$$\boldsymbol{v}_{\rm P} = \boldsymbol{v}_{\rm i} - \boldsymbol{v} = \frac{\boldsymbol{J}\times\boldsymbol{B}}{c\gamma_{\rm i}\rho_{\rm i}\rho}\,, \tag{9}$$

and so is parallel to $\boldsymbol{J}\times\boldsymbol{B}$. In equation (9), ρ and $\rho_{\rm i}$ are the neutral and ion density, respectively, and

$$\gamma_{\rm i} = \frac{\langle\sigma v\rangle_{\rm i}}{m_{\rm i}+m}\,, \tag{10}$$

where $\langle\sigma v\rangle_{\rm i}$ measures the rate coefficient of momentum exchange via collisions with the neutrals, taken to have mean mass m and $m_{\rm i}$ is the ion mass. When $\boldsymbol{J}\cdot\boldsymbol{B}=0$, such as envisaged here, Ohm diffusion also produces a drift in the same direction, so we have adopted the subscript P to denote the Pedersen drift (see equation A10).

Fig. 2 shows the field line drift (red vectors) for the fluid configuration depicted in Fig. 1, but now incorporating the effect of ambipolar or Ohm diffusion. The black vectors indicate the total drift of the magnetic field in the local Keplerian frame, and correspond to the vector sum $\boldsymbol{v}+\boldsymbol{v}_B$ (see equation A10). In this case, the net effect of the drift is to reduce the radial and azimuthal stretching of the field, and therefore – as one might guess – reduce the degree of instability.

In the Hall limit, collisions with the neutral gas are frequent and strong enough to decouple the ions from the field, but the electrons – which have a smaller collision cross-section with the neutrals and a higher charge-to-mass ratio – remain well coupled. The magnetic field is then frozen into the electrons, and these drift together through the neutrals and ions, which remain tightly coupled by collisions. In this limit, the field drift speed through the neutrals is given by the ion–electron drift and hence is antiparallel to the current density,

$$\boldsymbol{v}_{\rm H} = \boldsymbol{v}_{\rm e} - \boldsymbol{v}_{\rm i} = -\frac{\boldsymbol{J}}{en_{\rm e}}\,. \tag{11}$$

In equation (11), $n_{\rm e}$ ($=n_{\rm i}$) is the electron number density and e is the elementary electric charge.

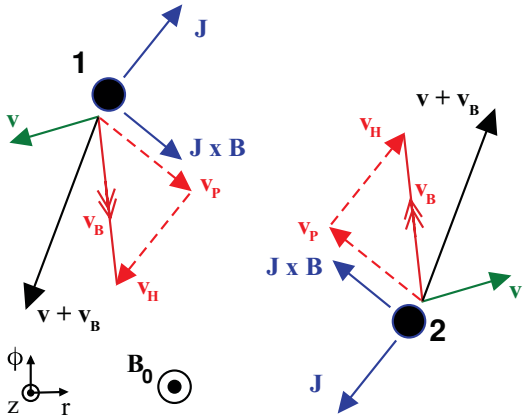

Figure 3. Plan view of the effect of magnetic diffusion [incorporating the ambipolar, Ohm and Hall (>0) terms] on the development of the MRI. The current density and associated magnetic stress on the fluid at points 1 and 2 of Fig. 1 are indicated in blue. At point 1, the magnetic tension provides some radial support against gravity and supplies a negative torque. As a result, the material here is sub-Keplerian and moving inwards, as indicated by the green vector (which denotes departure from Keplerian rotation). Because of the finite diffusivity, the field lines drift through the fluid (*red* vectors) with a velocity determined by the magnitudes of the ambipolar, Ohm and Hall diffusivities (see equation A10). Ambipolar and/or Ohm diffusion contributes a drift $\boldsymbol{v}_{\rm P}$ parallel to $\boldsymbol{J}\times\boldsymbol{B}$, while the Hall drift $\boldsymbol{v}_{\rm H}$ is antiparallel to the current density $\boldsymbol{J}$. The net drift of the field line with respect to the Keplerian rotation, $\boldsymbol{v}+\boldsymbol{v}_B$, is indicated by the *black* vector. In this case, the net drift is inwards. At point 2, the current density and magnetic stresses are reversed, so the field drifts in the opposite direction. As a result, the field lines drift inwards at point 1 and outwards at point 2 and the instability proceeds.

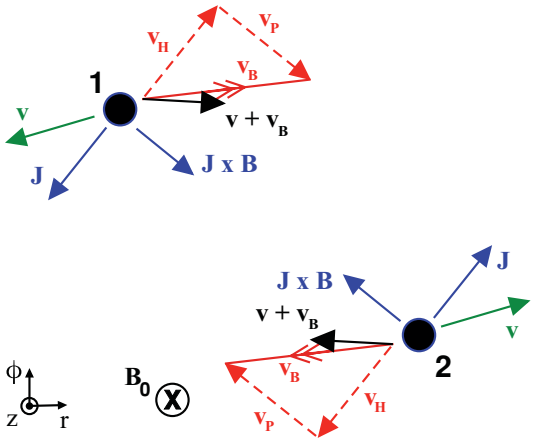

Figure 4. As for Fig. 3, but after a global reversal of the magnetic field, as would be expected if the initial magnetic field was antiparallel to the disc's angular velocity vector. In this case, the magnetic stresses are unaffected – and so, therefore, are the fluid velocity and Ohm/ambipolar drift current density. However, the current density is reversed, and therefore so is the Hall drift. The magnetic field now drifts outwards at 1 and inwards at 2, reducing the radial buckling. The MRI is partly or entirely suppressed in these circumstances (see text).

More generally, when all diffusion mechanisms are taken into account, the drift velocity $\boldsymbol{v}_B$ of the field lines through the fluid is the sum of these two orthogonal contributions (see equation A10 below). The implications for the MRI are sketched in Fig. 3, which now includes a Hall drift antiparallel to $\boldsymbol{J}$. The generic geometry of the MRI means that the Hall drift at point 1 is directed inwards and retrograde in azimuth, and it is outwards and prograde at point 2. The net effect is to exacerbate the radial buckling of the field, but to reduce the buckling in the azimuthal direction. We show in the next section that it is the radial effect that is critical, as azimuthal field is always created out of the underlying Keplerian shear of the disc. Therefore, Hall diffusion in this case (e.g. when the magnetic field is aligned with the angular velocity vector of the disc) tends to be destabilizing.

This is not always the case, however, because the direction of the current density, and therefore of the Hall diffusion, reverses upon global reversal of the magnetic field. This situation is illustrated in Fig. 4. The magnetic stresses, the fluid velocity and the field-line drift associated with Ohm and ambipolar diffusion are not affected by this reversal. Note, however, that Hall diffusion now acts to stabilize the disc by acting against the radial buckling of the field that would otherwise be driving the fluid motions.

To summarize, Ohm and ambipolar diffusion are stabilizing effects, whereas Hall diffusion may be destabilizing or stabilizing depending on whether the initial vertical magnetic field is parallel or antiparallel to the rotation axis, respectively. This asymmetry reflects the fundamental asymmetry in the microscopic properties of the positive and negative charged species in the fluid in this limit.

3 LINEAR ANALYSIS OF THE MRI

We consider a small region of an axisymmetric, geometrically thin and nearly Keplerian disc, threaded by a vertical magnetic field, with sound and Alfvén speeds ($c_{\rm s}$ and $v_{\rm A}$) at the mid-plane that are both small compared to the local Keplerian speed $v_{\rm K}$. We assume that radial gradients are on the scale of r and neglect vertical stratification of the initial equilibrium state, so that our analysis only holds near the mid-plane at heights $z \ll c_{\rm s}/\Omega$. The initial state is in Keplerian rotation with a uniform density, pressure and vertical magnetic field $\boldsymbol{B} = sB\hat{\boldsymbol{z}}$ (where $s = \pm 1$). We linearize the equations around this state and seek solutions for axisymmetric perturbations of the form $\exp(\nu t - {\rm i}kz)$. The equations for perturbations in density, pressure and v_z form a separate system that describes vertically propagating sound waves. The system of linear equations in the remaining perturbations involve fluctuations in the r and ϕ components of $\boldsymbol{B}$, $\boldsymbol{v}$ and $\boldsymbol{v}_{\rm B}$.

A derivation of the dispersion relation is outlined in Appendices A and B. This dispersion relation was first derived by Wardle (1999), and later extended to more general geometries in the Hall–Ohm limit (Balbus & Terquem 2001), and then including ambipolar diffusion (Desch 2004); the ambipolar diffusion limit was also considered by Kunz & Balbus (2004). Our emphasis here is on the dependence of the instability on the Pedersen and Hall diffusivities $\eta_{\rm P}$ and $\eta_{\rm H}$; a detailed analysis is presented in the appendices. An overview is provided by Fig. 5, which illustrates the qualitative changes in the growth rate versus wavenumber curves for different choices of the diffusivities.

First, note that the ideal-MHD limit holds at the origin (i.e. $\eta_{\rm H} = \eta_{\rm P} = 0$), the Hall MHD limit holds along the horizontal axis ($\eta_{\rm P} = 0$) and the Ohm or ambipolar diffusion limits hold along the vertical line ($\eta_{\rm H} = 0$). Recall also that only the half-plane $\eta_{\rm P} \geq 0$ is physically relevant. It turns out that there are three distinct forms of the resulting $\nu(k)$ curve, corresponding to regions labelled I, II and III in the plane, as illustrated in Fig. 5. In region I, wavenumbers less than a cut-off $k_{\rm c}$ are unstable with the maximum growth rate attained at an intermediate wavenumber k_0 (see equations B12, B13 and

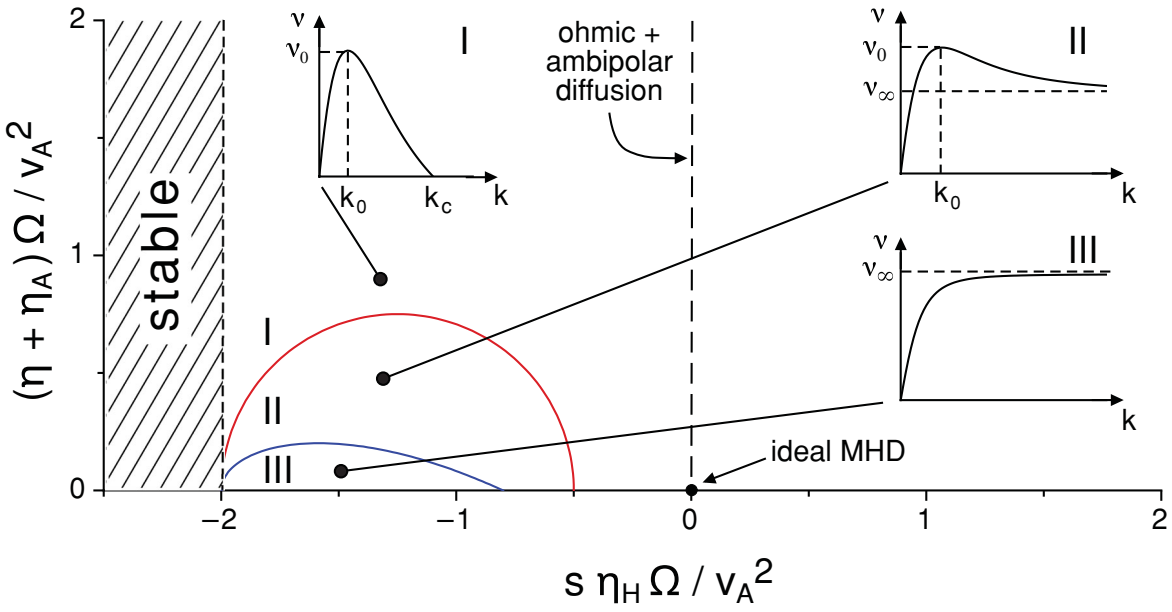

Figure 5. Schematic dependence of the behaviour of the MRI on the Hall and Pedersen (Ohm+ambipolar) diffusivities, η_H and η_P, for an initial magnetic field $sB\hat{z}$ in a Keplerian disc, where $s = \pm 1$ is the sign of B_z. The field is assumed to be weak so that stratification can be neglected. Physical values of the Pedersen conductivity η_P are non-negative, and there are no unstable modes for Hall diffusivities $s\eta_H\Omega/v_A^2 \leq -2$. The unstable region $s\eta_H\Omega/v_A^2 > -2$ is subdivided according to the dependence of the growth rate on wavenumber k, which is sketched in the insets labelled I–III. In region I (outside the red locus given by equation B11), wavenumbers less than a cut-off k_c are unstable, with the maximum growth rate ν_0 attained at k_0 (see equations B12, B13 and B14). Between the red and blue loci (region II), all wavenumbers are unstable, with the maximum growth rate still occurring at finite wavenumber. Within the blue locus given by equation (B15) (region III), all wavenumbers are unstable and the maximum growth rate is approached asymptotically as $k \to \infty$.

B14). In region II, all wavenumbers are unstable with the maximum growth rate still occurring at finite wavenumber. Finally, in region III all wavenumbers are unstable and the maximum growth rate is approached asymptotically as $k \to \infty$.

Having delineated these three regions, we now consider how the critical wavenumber k_c, fastest growth rate ν_0 and corresponding wavenumber k_0 vary across the entire η_P–η_H plane. Contours of the growth rate and wavenumber of the most unstable mode are plotted in this plane in Fig. 6. The growth rate increases clockwise, from 0Ω along the vertical line $s\eta_H = -2$ up to 0.75Ω for the horizontal line $\eta_P = 0$ for $s\eta_H > -4/5$. In the absence of Hall diffusion, the maximum growth rate ν_0 declines with increasing (Ohm and/or ambipolar) diffusivity (e.g. moving vertically upwards through the $s\eta_H = 0$ point in the horizontal axis), with $\nu_0 \approx (3/4)\eta_P^{-1}$ for $\eta_P \gg 1$. The most important effect of Hall diffusion, apparent from Fig. 6, is that *the growth rate of the MRI exceeds 0.3Ω for* $s\eta_H \gtrsim \eta_P$, *even for arbitrarily large* η_P. More generally, the addition of Hall diffusion at fixed η_P increases the growth rate if $s\eta_H > 0$ and decreases it when $s\eta_H < 0$. For large values of η_P, equation (B13) shows that $s\eta_H/\eta_P \approx 24\nu_0/(9 - 16\nu_0^2)$. It is this fact that has the potential to modify the extent of dead zones in protoplanetary discs, as we explore in Section 5.

The wavenumber of the fastest growing mode (blue contours in Fig. 6) decreases as the diffusivity is increased. Again, the contours are not arranged so that the highest wavenumbers occur in the ideal-MHD limit, but to the $s\eta_H < 0$ side, within the boundary between regions II and III (traced by the $k_0 = \infty$ contour).

Contours of constant k_c are semicircles, as plotted in Fig. 7. While the range of unstable wavenumbers is reduced for large values of $s\eta_H$ and η_P, as one might expect, the range is not maximized in the ideal limit (i.e. at the origin) but in regions II and III, bounded by the $k_c = \infty$ contour.

In general, we note that increasing η_P decreases the maximum growth rate and the characteristic wavenumbers, whereas increasing $s\eta_H$ above -2 increases the maximum growth rate and may either increase (when $s\eta_H + 1 \lesssim \eta_P$) or decrease (when $s\eta_H + 1 \gtrsim \eta_P$) the corresponding wavenumber.

Overall, these patterns place the ideal, Ohm (or ambipolar) and Hall regimes in context, and for the first time we see an overview of the effect of magnetic diffusivity on the linear MRI. In particular, there is nothing special about the Ohm/ambipolar limit, e.g. the behaviour of the instability in the presence of diffusion is not qualitatively different for $s\eta_H = 2$ versus $\eta_H = 0$. Even the ideal-MHD limit does not stand apart as remarkable, although it still holds a special place conceptually because flux freezing holds and it is easier to visualize.

What does stand out is the part of the plane in the lower left, regions II and III, characterized by high wavenumbers and the spraying out of the growth contours. In this part of the plane, the instability operates in the 'cyclotron limit' $\eta_\perp \sim 1$ and $k^2 \gg 1$ (see Appendix C), in which the instability arises in the competition between magnetic diffusion and advection of the field by the fluid; generation of B_ϕ from B_r by the Keplerian shear flow is negligible. The lack of any k dependence in this regime occurs because both the magnetic diffusion and the magnetic stresses on the fluid (which are responsible for the fluid displacement) scale as k^2. This short-wavelength, low-frequency limit corresponds to the cyclotron mode of the magnetized fluid, which has frequency $\omega_H = v_A^2/|\eta_H|$ (Wardle & Ng 1999; Pandey & Wardle 2008). This mode couples effectively to the Keplerian rotation as long as the sense of circular polarization matches that of epicyclic motion, i.e. as long as $B_z\eta_H < 0$. The other short-wavelength mode, the high-frequency whistler ($\omega \approx k^2 v_A^2/\omega_H$), is unable to couple effectively to the rotation (Wardle 1999).

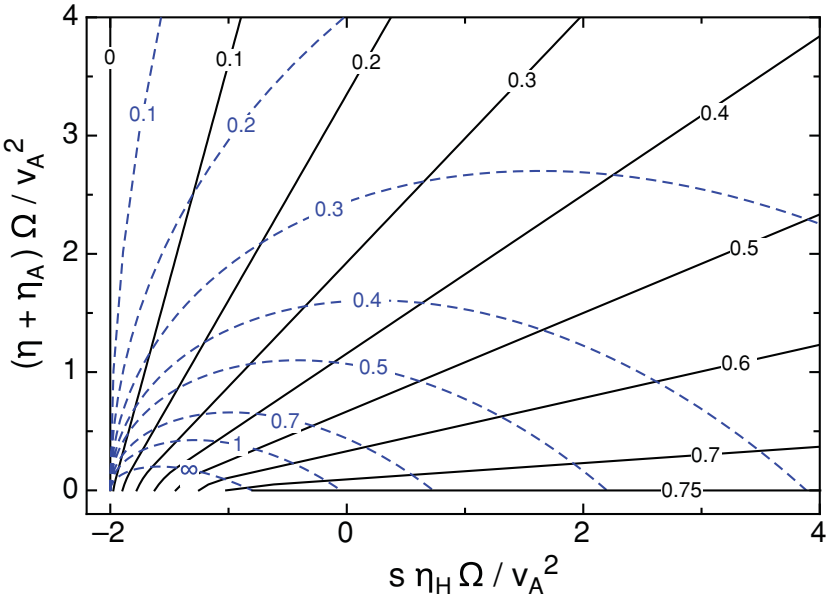

Figure 6. Contours of the maximum growth rate of the MRI (solid black lines, units Ω) and the corresponding wavenumber (dashed blue lines, units Ω/v_A) as a function of Hall and Pedersen (Ohm+ambipolar) diffusivities. The innermost blue contour shows where k_0 becomes infinite, corresponding to the boundary between regions II and III in Fig. 5. Within this region, the fastest growth rate is approached asymptotically as $k \to \infty$; this transition is responsible for the curvature of the black contours within this region (see equations B13 and B16).

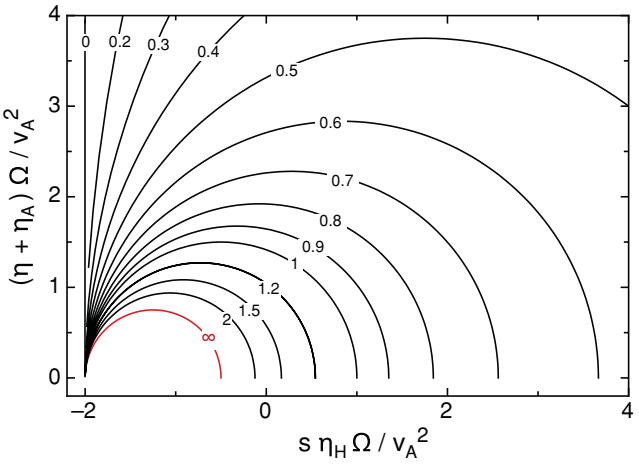

Figure 7. Contours of the critical wavenumber k_c below which instability sets in (units Ω/v_A). The red innermost contour shows where k_c becomes infinite and all wavenumbers are unstable, corresponding to the boundary between regions I and II in Fig. 5.

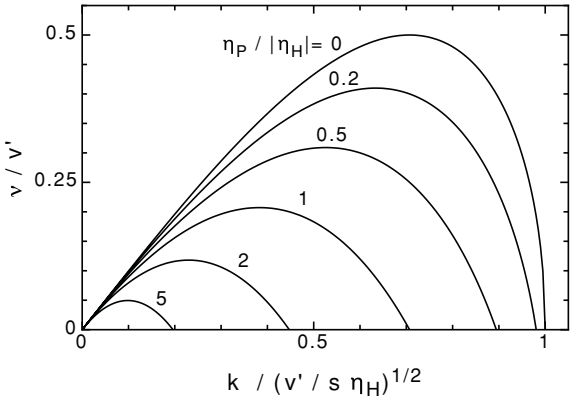

Figure 8. Growth rate versus wavenumber for the Hall-driven diffusive instability for a mutually orthogonal magnetic field, shear velocity and velocity gradient v' in the highly diffusive limit. For Keplerian rotation and cylindrical geometry, the effective velocity gradient is $v' = (3/2)\Omega$. The curves correspond to different values of the ratio of the Pedersen and Hall diffusivities.

The maximum growth rate contours emerging from regions II and III continue on to attain another limit when $\eta_P^2 + \eta_H^2 \gg 1$, in which the field evolves in response to shear and diffusion, without significant feedback from the perturbations that it induces in the fluid flow. Instability in this case relies on the Keplerian shear flow generating B_ϕ from B_r and the tendency of Hall diffusion to convert B_ϕ back into B_r. This brings the potential destabilizing effect of Hall diffusion in shear flows (Kunz 2008) to the fore and shows that it is quite independent of rotational effects – i.e. the Coriolis and centripetal acceleration – that drive the MRI. In Appendix C, we obtain simple analytic expressions for growth rate in plane-parallel shear flows as a function of k for high diffusivity; the results are plotted in Fig. 8. For more general field and wave vector configurations, ambipolar diffusion plays a similar role, albeit hindered by dissipation (Desch 2004; Kunz & Balbus 2004; Kunz 2008; Pandey & Wardle 2012).

4 COMPARISON WITH NUMERICAL SIMULATIONS

In this section, we explore the relationship of our linear calculations to the non-linear, unstratified shearing box simulations of MRI-driven turbulence by Sano & Stone (2002a,b), which explicitly included both Ohm and Hall diffusion. These authors' results could be interpreted as implying that the presence of Hall currents has little effect on the critical (minimum) degree of magnetic coupling (the coupling between the neutral gas and the magnetic field) for the

Table 1. Typical dimensionless numbers used to characterize the degree of coupling between the magnetic field and the neutral gas: magnetic Reynolds number, Lundquist number and Elsasser number. For comparison, we also show the Reynolds number (although this parameter combination does not involve the magnetic diffusivity and, therefore, does not measure the field–matter coupling). The fluid variables used to characterize the required scalings are the flow velocity (v), the Alfvén velocity (v_A), the kinematic viscosity (ν) and the magnetic diffusivity η. As usual, Ω is the Keplerian angular frequency and l is a typical, unspecified, length-scale of the flow.

Symbol	Dimensionless number	Ratio of physical quantities	Scaling V	Scaling L	Scaling D	Definition vl/D
Re	Reynolds	Inertial to viscous forces	v	l	ν	vl/ν
$\mathrm{Re_M}$	Magnetic Reynolds	Inertial to resistive time-scales	v	l	η	vl/η
S	Lundquist	Resistive diffusion to Alfvén time-scales	v_A	l	η	$v_A l/\eta$
Λ	Elsasser	Lorentz to Coriolis forces	v_A	v_A/Ω	$\eta_\perp$	$v_A^2/\eta_\perp\Omega$

instability to operate.[2] We show, however, that their calculations do not yet probe deeply enough into the Hall regime for Hall diffusion to be able to impact on the development and properties of the instability. In fact, their solutions are in agreement with expectations from local and stratified linear analyses for the same values of the parameters describing the initial fluid conditions. Finally, we use the ratios of the different terms in the induction equation in order to delineate the region of parameter space where Hall diffusion is expected to substantially modify the growth rate, and spatial scale, of the MRI-unstable modes. First, however, and for the sake of clarity, we discuss the notation used in the literature to characterize the fluid, and its magnetic activity, in diffusive environments.

4.1 Fluid parameters

The growth and structure of the MRI are strongly dependent on the magnetic field strength, geometry and the nature – and magnitude – of the magnetic diffusivity. Different parametrizations have been used in the literature to characterize these properties, depending on the formulation of the problem and the adopted form of the induction equation. Some authors use a tensor diffusivity, and the induction equation then takes the form

$$\frac{\partial \boldsymbol{B}}{\partial t} = \nabla\times(\boldsymbol{v}\times\boldsymbol{B}) - \nabla\times[\eta\nabla\times\boldsymbol{B} + \eta_H(\nabla\times\boldsymbol{B})\times\hat{\boldsymbol{B}} + \eta_A(\nabla\times\boldsymbol{B})_\perp] \quad (12)$$

(e.g. Wardle 2007). Others prefer to use separate equations for the charge carriers and the neutrals. In the latter 'multifluid' approach, when the ionized species are ions and electrons only (denoted by the subscripts 'i' and 'e', respectively), the induction equation becomes (e.g. Königl 1989; Balbus & Terquem 2001; Sano & Stone 2002a)

$$\frac{\partial \boldsymbol{B}}{\partial t} = \nabla\times(\boldsymbol{v}\times\boldsymbol{B}) - \nabla\times\left[\frac{cm_e\nu_{en}\boldsymbol{J}}{n_e e^2} + \frac{\boldsymbol{J}\times\boldsymbol{B}}{en_e} - \frac{(\boldsymbol{J}\times\boldsymbol{B})\times\boldsymbol{B}}{c\gamma_i\rho\rho_i}\right]. \quad (13)$$

Comparison of equations (12) and (13) yields the expressions given in (6)–(8) for the diffusivities. On the right-hand side (RHS) of equation (12), as well as in the second line of equation (13), the terms (from left to right) denote the inductive (I), Ohm (O), Hall (H) and ambipolar diffusion (A) contributions to the evolution of the magnetic field, respectively.

The following parameters have typically been used to characterize the magnetic properties of the fluid.

(i) *Field strength*. This property is commonly measured either by the ratio of the Alfvén speed to the isothermal sound speed (v_A/c_s) or by the plasma β parameter $\beta = (2/\gamma)c_s^2/v_A^2$, where γ is the adiabatic index of the fluid.

(ii) *Field–matter coupling*. The degree of coupling between the neutral matter and the magnetic field is, most generally, measured by a ratio of the type vl/D, where V, L and D represent characteristic speed, length and diffusion scales of the flow, respectively. In MRI studies, it is appropriate to take $V = v_A$, $L = v_A/\Omega$ (the characteristic wavelength of MRI-unstable modes in ideal-MHD conditions) and $D = \eta_\perp$, the total perpendicular diffusivity [e.g. equation (C8)]. The resulting ratio is the Elsasser number

$$\Lambda \equiv \frac{v_A^2}{\Omega\eta_\perp}. \quad (14)$$

This dimensionless number is formally defined as the ratio of the Lorentz force $(\boldsymbol{J}\times\boldsymbol{B})/c$ to the Coriolis force $\propto \rho(\boldsymbol{v}\times\boldsymbol{\Omega})$, which results in an expression of the form $v_A^2/(LV)\Omega$. Adopting the magnetic diffusivity $\eta_\perp$ as the typical magnitude of the LV factor in the denominator yields expression (14).

The Elsasser number is widely used in geophysics, as in a geodynamo process the magnetic field is thought to amplify until Λ becomes of the order of unity (see e.g. Christensen 2010, and references therein). This ratio was denoted χ in the work of Wardle (1999), who expressed it as

$$\chi \equiv \frac{B^2\sigma_\perp}{\rho c^2\Omega} \equiv \frac{\omega_c}{\Omega}, \quad (15)$$

where $\sigma_\perp = c^2/(4\pi\eta_\perp)$ and ω_c is the critical frequency above which flux-freezing conditions break down, so that the generic non-ideal-MHD term in the induction equation dominates over the inductive term. When the field–matter coupling is characterized by the Elsasser number, values $\gg 1$ ($\ll 1$) correspond to strong (weak) coupling. Similarly, from equation (15) it is clear that when $\Lambda \gg 1$, the magnetic field is strongly coupled to the fluid at frequencies of order Ω, the frequencies of interest for the study of the MRI.

Note that, in general, the Elsasser number is different from the combination $v_A l/\eta$, which compares the resistive diffusion time $\tau_R \propto l^2/\eta$ to the Alfvén time $\tau_A \propto l/v_A$ (e.g. the Lundquist number) and from the ratio vl/η (the magnetic Reynolds number $\mathrm{Re_M}$, where v is the fluid velocity). However, both terms have been used in the literature to refer to Λ. Table 1 summarizes the definitions and typical scalings associated with the dimensionless numbers

[2] This field–matter coupling parameter is often referred to as the magnetic Reynolds number $\mathrm{Re_M}$. As we expound below, however, it really corresponds to the Elsasser number Λ (see discussion below and Table 1).

discussed above, namely the magnetic Reynolds, Lundquist and Elsasser numbers. For clarity, we also list the Reynolds number, although this does not measure magnetic coupling, as it deals with the viscosity of the fluid instead of its magnetic diffusivity.

Finally, sometimes c_s is adopted as the characteristic velocity scale of the flow instead of v_A, resulting in the ratio $c_s^2/\eta\Omega$ (e.g. Fleming, Stone & Hawley 2000).

(iii) *Diffusivity regime*. This property characterizes the importance of the different non-ideal-MHD terms in the induction equation. The ratios of each of the diffusive terms to the inductive term are typically used, which in a multifluid formulation are expressed as (e.g. Balbus & Terquem 2001; Sano & Stone 2002a)

$$\frac{O}{I} = \frac{\eta\Omega}{v_A^2} \equiv \tilde{\eta}\,, \tag{16}$$

$$\frac{H}{I} = \frac{X}{2} \tag{17}$$

and

$$\frac{A}{I} = \frac{\Omega}{\gamma_i \rho_i}\,. \tag{18}$$

Note that the inverse of the normalized Ohm resistivity $\tilde{\eta}^{-1}$ in equation (16) has been referred to as the magnetic Reynolds number. However, as discussed above, this parameter combination is akin to the Elsasser number, with the difference that in the definition of equation (14) we used the total perpendicular diffusivity $\eta_\perp$, instead of η. In equation (17),

$$X \equiv \frac{cB\Omega}{2\pi e n_e v_A^2} \tag{19}$$

is the so-called 'Hall parameter', not to be confused with its namesake, the ratio of the gyrofrequency to the collision frequency of ionized species j (here either ions or electrons) with the neutrals, given by

$$\beta_j \equiv \frac{eB}{m_j c}\frac{1}{\gamma_j \rho}\,. \tag{20}$$

Similarly, the relative importance of the non-ideal-MHD terms can be measured by the following ratios (e.g. Balbus & Terquem 2001; Sano & Stone 2002a):

$$\frac{H}{O} = \frac{eB}{m_e c}\frac{1}{\gamma_e \rho} \equiv \beta_e = \frac{X}{2\tilde{\eta}} \tag{21}$$

and

$$\frac{A}{H} = \frac{eB}{m_i c}\frac{1}{\gamma_i \rho} \equiv \beta_i\,. \tag{22}$$

Using the tensor diffusivity notation, we can write

$$\frac{O+A}{I} = \frac{\eta_P \Omega}{v_A^2}, \tag{23}$$

$$\frac{H}{I} = \frac{\eta_H \Omega}{v_A^2}, \tag{24}$$

and the relative magnitudes of the non-ideal-MHD terms are, simply,

$$\frac{H}{O} = \frac{\eta_H}{\eta} \tag{25}$$

and

$$\frac{A}{H} = \frac{\eta_A}{\eta_H}. \tag{26}$$

From expressions (23) and (24), it is clear that the Elsasser number is a measure of the ratio of the inductive term to the total non-ideal-MHD terms in the induction equation. Furthermore, for the vertical field geometry adopted here, the ambipolar diffusivity acts as a field-dependent resistivity[3] (see Balbus & Terquem 2001). This property will be used in the next section to treat the Ohm and ambipolar diffusivities as an 'Ohm-like' term when comparing the impact of the 'Ohm-like' and Hall resistivities on the MRI.

4.2 Criterion for Hall diffusion to affect the MRI

We now use the ratios given in the previous subsection to constrain the region of parameter space where the MRI is expected to grow, as well as to determine a criterion for Hall diffusion to substantially modify its properties. First, however, we discuss the results of the work of Sano & Stone (2002b) on this topic. These authors characterized the Ohm and Hall terms by their magnitudes relative to the inductive term in the induction equation (e.g. via equations 16 and 17). Furthermore, they considered that Hall diffusion was dominant when the ratio of the Hall to inductive terms was larger than unity ($H/I > 1$). Fig. 9 of Sano & Stone (2002b) shows the saturation level of the Maxwell stress (normalized by the initial gas pressure) as a function of the initial (subscript '0') of the inverse Ohm resistivity ($\tilde{\eta}_0^{-1}$) – called the magnetic Reynolds number – for different values of the initial plasma β and Hall parameters (β_0 and X_0, respectively).

Their results indicate that when $\tilde{\eta}_0^{-1} \geq 1$, the normalized saturated value of the Maxwell stress is of the order of 0.1 (with some scatter depending on the adopted β_0 and X_0), and it is fairly independent of the actual value of the resistivity. The actual saturated magnitude of the stress is larger in the models with $X_0 = 4$ with respect to the ones with $X_0 = 0$ (e.g. no Hall diffusivity) or $X_0 = -2$ (negative Hall diffusivity, when the magnetic field is counteraligned with the disc angular velocity vector). On the contrary, when the initial inverse resistivity $\tilde{\eta}_0^{-1}$ is less than unity, the saturation level of the Maxwell stress decreases by 1–2 orders of magnitude with respect to the $\tilde{\eta}_0^{-1} > 1$ case. This trend of the Maxwell stress at saturation with $\tilde{\eta}_0^{-1}$ seems to be unaffected by the presence (or magnitude) of the Hall diffusivity. These results have been interpreted to show that the Hall diffusivity does not change the critical (maximum) $\tilde{\eta}$ required for the instability to grow ($\tilde{\eta}_{crit} \sim 1$), but it enhances the saturated level of the Maxwell stress by a factor of a few.

Note that the calculations of Sano & Stone (2002b) probe the region of parameter space where Ohm and Hall diffusivity terms dominate over the inductive term ($I/O = \tilde{\eta}_0^{-1} < 1$ and $H/I = X_0/2 > 1$). However, $|X_0| \leq 4$ in all the presented calculations. As a result, for $\tilde{\eta}_0 > 2$, the ratio $H/O = X_0/(2\tilde{\eta}_0)$ is <1 and the dominant diffusion mechanism is Ohm type. For the calculations with $X_0 = 4$ and $\tilde{\eta}_0^{-1} = 0.1$, in particular, the ratio $H/O = 0.2$, which implies that the Hall term is too weak to overcome the damping effect introduced by Ohm diffusion. It is not surprising, therefore, that the large drop in the saturated value of the Maxwell stress, with respect to that associated with the solutions satisfying $\tilde{\eta}_0^{-1} > 1$, is not significantly modified by the presence of Hall diffusion. Naively, and as confirmed by the linear analysis, the Hall effect should substantially modify the growth rate of the MRI provided that Hall diffusion dominates over the inductive term (i.e. $|X_0| \gtrsim 2$)

[3] In other words, for this field geometry the current density $\boldsymbol{J}$ is perpendicular to $\boldsymbol{B}$ and the term $(\boldsymbol{J}\times\boldsymbol{B})\times\boldsymbol{B} \equiv (\boldsymbol{J}\cdot\boldsymbol{B})\boldsymbol{B} - B^2\boldsymbol{J} = -B^2\boldsymbol{J}$ (see equation 13).

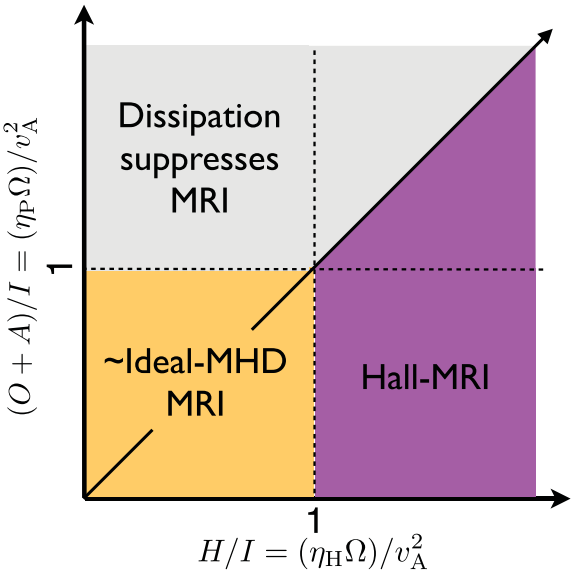

Figure 9. Regions of parameter space where the MRI is expected to grow, and dominant diffusion mechanism, in a $(O + A) / I$ versus H / I (or, equivalently, the normalized Pedersen versus Hall diffusivities) plane. Note that for the vertical field geometry considered here, the ambipolar and Ohm diffusivities can be combined into a generalized 'Ohm-type' diffusivity (see text). In the lower-left (ochre) panel, magnetic diffusion is weak and the instability grows at a rate comparable to the ideal-MHD rate ($\sim\Omega$). Conversely, in the upper-left (grey) and lower-right (purple) panels, only one diffusivity term (Ohm-type and Hall, respectively) dominates over the inductive term in the induction equation. Ohm-type diffusivity suppresses the instability in the top-left quadrant (see footnote 4, however), but since Hall diffusion can also be destabilizing (Section 2), the instability can still grow in the lower-right panel. Finally, both diffusivity components are important in the upper-right quadrant of the figure. In the top (grey) portion of this panel, Ohm-type diffusion overcomes Hall diffusion [$H / (O + A) < 1$] and the MRI is expected to be suppressed. Hall diffusivity, however, is dominant [$H / (O + A) > 1$] in the bottom (purple) portion of the panel. In this region of parameter space, $H / (O + A)$ *and* H / I are both >1 and the instability may proceed despite Λ being less than unity.

and also dominates Ohm diffusion (i.e. $X_0 \gtrsim 2\tilde{\eta}_0$). According to these criteria, the Hall term would be strong enough to modify the presented results if $X_0 \geq 20$ (for the $\tilde{\eta}_0^{-1} = 0.1$ case), a value likely to be well beyond what has been computationally feasible so far.

The above considerations are summarized in Fig. 9, which shows the regions of parameter space where the instability is expected to grow – and, if so, the dominant diffusion mechanism – in a [$(O + A) / I$–H / I] or, equivalently, the normalized Pedersen and Hall diffusivities, plane. As discussed at the end of Section 4.1, we use a generalized 'Ohm' term, which should be understood to mean 'Ohm + ambipolar'.

Diffusivity effects are weak in the lower-left (ochre) quadrant of the figure, as the inductive term dominates over both the Ohm and Hall diffusion terms in this region of parameter space. The MRI is then expected to grow at a rate not significantly reduced from the ideal-MHD rate.[4] Conversely, in the upper-left (grey) and lower-right (purple) panels, only one diffusivity term is dominant over the inductive term (the Ohm and Hall term, respectively). When Ohm diffusion dominates, the instability's growth is damped, as this diffusion mechanism is always stabilizing (see the discussion in Section 2 and Fig. 2). However, the Hall term can be either stabilizing or destabilizing (see Figs 3 and 4), so when it dominates, the instability can still potentially grow. Finally, the upper-right quadrant of the figure contains the region of parameter space where both the Ohm and Hall diffusivity terms dominate over the inductive term. Although both diffusivity terms are important with respect to the ideal-MHD term in this region, only in the lower (purple) portion of the panel, Hall diffusion is strong enough to potentially overcome the damping effect introduced by Ohm diffusivity, as *only* here the ratios $H / (O + A)$ and H / I are both >1. Note that this region of parameter space was not probed by Sano & Stone (2002b), as their solutions incorporating Hall diffusion lie in the upper (grey) portion of the panel, where Hall diffusion is weak in comparison to Ohm diffusion. The potentially destabilizing effect of a sufficiently strong Hall diffusivity (the purple region of the upper-right quadrant in Fig. 9), therefore, remains to be explored by numerical simulations.

In order to test these assertions, we computed the growth rate of the most unstable MRI mode in a stratified, geometrically thin and axisymmetric accretion disc, using the method described in Salmeron & Wardle (2003), to which we refer the reader for additional details. Our results are depicted in Fig. 10 for $\beta = 3200$ (or $v_A/c_s = 0.02$), as a function of the ratio H/O ($= X/2\tilde{\eta}$). Each curve corresponds to a different value of the ratio I/O ($= \tilde{\eta}^{-1}$) as follows: $\tilde{\eta}^{-1} = 0.1$ (blue), 1.0 (red) and 10 (green). The range of solutions

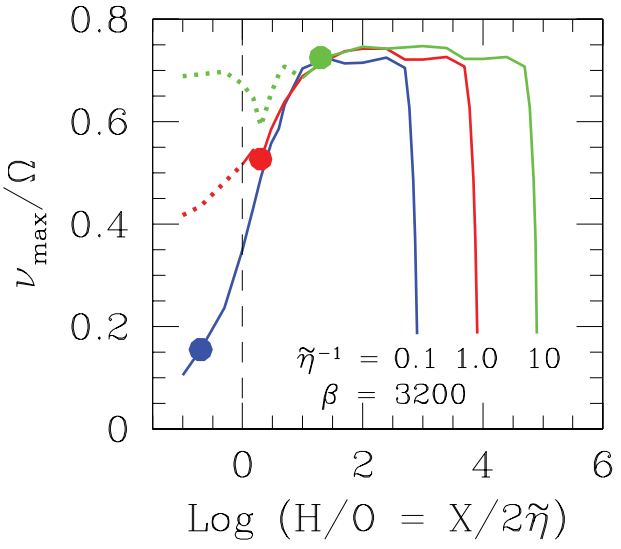

Figure 10. Growth rate of the fastest growing MRI-unstable mode as a function of the ratio of the Hall to Ohm diffusion terms in the induction equation [$H/O = X/(2\tilde{\eta})$]. Calculations correspond to a stratified disc with $\beta \equiv (2/\gamma)c_s^2/v_A^2 = 3200$ (or $v_A / c_s = 0.02$). Each curve corresponds to a different value of $\tilde{\eta}^{-1}$ ($= I / O$): 0.1 (blue), 1.0 (red) and 10 (green). A solid (dotted) line is used to plot solutions for which the ratio $H / I > 1$ (<1). Note that solutions for which Hall diffusion is dominant lie in the solid-line portion of the curves ($H / I > 1$) *and* to the right of the vertical dashed line ($H / O > 1$). Superimposed (filled circles) are the linear growth rates corresponding to the solutions presented in fig. 9 of Sano & Stone (2002b) for $X = 4$. Note that the calculation for $\tilde{\eta}^{-1} = 0.1$ and $X = 4$ satisfies $H / O = 0.2$ and is *not* in the Hall-dominated region of parameter space. As a result, the MRI is significantly damped by Ohm diffusion. On the contrary, the instability can grow at about the ideal-MHD rate when $H / O > 1$ even when $\tilde{\eta}^{-1} = 0.1$.

[4] Note, of course, that the boundaries between these regions are not sharp as this simplified diagram might suggest, and the MRI is expected to grow at reduced rate when $\eta_P v_A^2 / \Omega \sim 1$ (e.g. near the border between the lower and upper-left panels of the figure). This also applies along the diagonal line separating the Hall- and Ohm-dominated regions in the upper-right panel.

shown with solid lines satisfy $X > 2$, so that the Hall term dominates over the inductive term ($H/I > 1$). The solutions depicted with a dotted line correspond to $X < 2$ ($H/I < 1$). Note also that to the left of the vertical dashed line, the ratio $H/O = X/2\tilde{\eta} < 1$ and Ohm diffusion is the dominant non-ideal-MHD term.

From the considerations in the previous paragraphs, it is clear that the solutions that probe the Hall-dominated region of parameter space should lie in the solid segment of each curve *and* to the right of the vertical dashed line. Note, in particular, that even for $\tilde{\eta}^{-1} = 0.1$, the instability grows at essentially the ideal-MHD rate ($\nu_{\max}/\Omega \approx 0.75$) if the Hall term is strong enough ($H/O > 1$, or to the right of the vertical line). For comparison, we superimpose filled circles on the curves showing the linear growth rate of the most unstable mode corresponding to $X = 4$ and $\tilde{\eta}^{-1} = 0.1$, 1 and 10 (the set of parameters tested by Sano & Stone 2002b). Note that the solution with $X = 4$ and $\tilde{\eta}^{-1} = 0.1$ lies to the *left* of the $H/O = 1$ line, and Ohm diffusion suppresses the growth of the instability. We conclude, therefore, that the relative saturation levels of the MRI simulations presented in Sano & Stone (2002b) are entirely consistent with expectations based on a linear analysis, and did not probe the regime $H > \{I, O, A\}$ in which the hall term strongly modifies the linear MRI.

5 APPLICATION TO PROTOPLANETARY DISCS

We now illustrate the implications of Hall diffusion for the extent of MRI-driven activity in protoplanetary discs by applying the linear analysis from Appendix B to the MMSN at 1 au.

While a linear analysis cannot hope to give a good sense of the properties of the non-linear turbulence the MRI drives, numerical simulations in the Ohm limit have shown that it does predict when such turbulence exists, and appears to be an excellent predictor of the vertical extent of dead zones in stratified simulations of protoplanetary discs (Turner & Sano 2008). As noted in the previous section, the Hall–Ohm simulations of Sano & Stone (2002a,b) are also consistent with the expectations based on the linear dispersion relation. Thus, we use the local dispersion relation to examine the role of the magnetic diffusivity in determining the extent of MRI-active regions in protoplanetary discs.

The tendency of the MRI to manifest on successively larger scales in the presence of increasing diffusivity is limited by the finite thickness of the disc: the wavenumbers of interest are bounded from below by requiring $kh > 1$, where $h = c_s/\Omega$ is the disc scaleheight. Thus we adopt as a local criterion for growth of the MRI that the dispersion relation (B6) yields modes with $\nu > 0$ and $kh > 1$. From this, we infer that the diffusivities must lie within the semicircular locus obtained from equation (B10) with $\nu = 0$ and $k = h^{-1}$, i.e.

$$\left(\frac{s\eta_H\Omega}{v_A^2} + \frac{5}{4} - \frac{3}{4}\frac{c_s^2}{v_A^2}\right)^2 + \left(\frac{\eta_P\Omega}{v_A^2}\right)^2 < \frac{9}{16}\left(1 + \frac{c_s^2}{v_A^2}\right)^2, \qquad (27)$$

where $s = \mathrm{sign}(B_z)$. This criterion applies to any near-Keplerian disc threaded by a vertical magnetic field. Its apparent simplicity belies the fact that the diffusivities are complicated (but calculable) functions of magnetic field strength, gas density, temperature and the abundances of charged species. Note also that the ratio c_s^2/v_A^2 plausibly ranges between 0 and the equipartition value 2 at the disc mid-plane but may be significantly higher away from the mid-plane if the field is anchored at lower heights within the disc.

We specialize to protoplanetary discs by adopting the ionization models of Wardle (2007) for a standard MMSN disc ionized by cosmic rays and stellar X-rays at 1 au from the central star. These models assume a column density of 1700 g cm^{-2} and temperature $T = 280$ K independent of height, with X-ray ionization rate computed by Igea & Glassgold (1999) and a standard interstellar cosmic ray ionization rate of 10^{-17} s^{-1} H^{-1} that is exponentially attenuated with depth as $\exp(-\Sigma/96\,\mathrm{g\,cm^{-2}})$. We characterize the grain population by assuming a single 1-μm radius and varying the dust-to-gas mass ratio, crudely mimicking the effect of the settling of grains to the disc mid-plane.

The abundances of charged species for dust-to-gas mass ratios of 10^{-2}, 10^{-4} and 0 are presented in Fig. 11. The ionization rate as a function of depth is plotted in the lower panel – interstellar cosmic rays are the dominant source below two scaleheights; above this, ionization by stellar X-rays dominates. In the no-grain case, electrons and metal ions are the dominant charged species because the metal ions have the smallest recombination rate coefficient. The top panel shows the effect of adding a population of 1-μm-radius grains with total mass 1 per cent of the gas mass: grains acquire a charge via sticking of electrons and ions from the gas phase. Above $z/h \approx 2.5$, the grain charge is determined by the competitive rates of sticking of ions and electrons, with the Coulomb repulsion of electrons by negatively charged grains offsetting their greater thermal velocity compared to ions (Spitzer 1941; Draine & Sutin 1987). This leads to a Gaussian grain charge distribution with mean charge (in units of e) $\langle Z_g\rangle \approx -4akT/e^2 \approx -67$ and standard deviation $\approx\langle Z_g\rangle^{1/2} \approx 8$. Most recombinations still occur in the gas phase. The abundance of metal ions and electrons is reduced over the grain-free case for $z/h \lesssim 5$, where the charge stored on grains becomes comparable to the electron abundance. For $z/h \approx 2$–2.5, the abundances of ions and electrons have declined to the point that the majority of electrons stick to grain surfaces before they can recombine in the gas phase, and most neutralizations occur when ions stick to negatively charged grains (Nakano & Umebayashi 1980). Closer to the mid-plane, the ionization fraction is so low that most grains are lightly charged, so that Coulomb attraction or repulsion of ions and electrons by grains is negligible. Then ions and electrons stick to any grain that they encounter, with recombinations occurring on grain surfaces. The middle panel of Fig. 11 shows what happens if 99 per cent of the grains are removed (e.g. by settling to the mid-plane). The capacity of the grain population to soak up electrons from the gas phase is reduced a hundredfold, and the height below which grains substantially reduce the free electron density below the ion density moves downwards to the lowest scaleheight.

Unlike Ohm resistivity, the Hall and ambipolar diffusivities depend on the magnetic field strength as well as the charged particle abundances, so we consider field strengths ranging between 10^{-3} and 10^2 G, encompassing the plausible range of values in the solar nebula (see Wardle 2007). For each choice of magnetic field, we use the ionization calculations to compute the diffusivities as a function of height. Typically, ambipolar diffusion dominates at the surface and, for weak magnetic fields, Ohm diffusion dominates near the mid-plane. Between these regimes, Hall diffusion dominates (Wardle 2007). With the diffusivity profiles in hand, we apply equation (27) to identify the fastest growing MRI mode that has $kh > 1$ at a given height; the results for dust-to-gas mass ratios 0 and 10^{-2} are displayed in Figs 12 and 13, respectively.

First consider the case in which grains are absent (Fig. 12). The bottom panel shows the fastest growth rate (subject to $kh > 1$) as a function of magnetic field strength and height if Hall and ambipolar diffusion are neglected and only Ohm diffusion is included. In this case, the departure from ideal MHD is measured by $\eta\Omega/v_A^2$ (the inverse of the Elsasser number; see Section 4.1), which strongly decreases with height as the fractional ionization and Alfvén speed

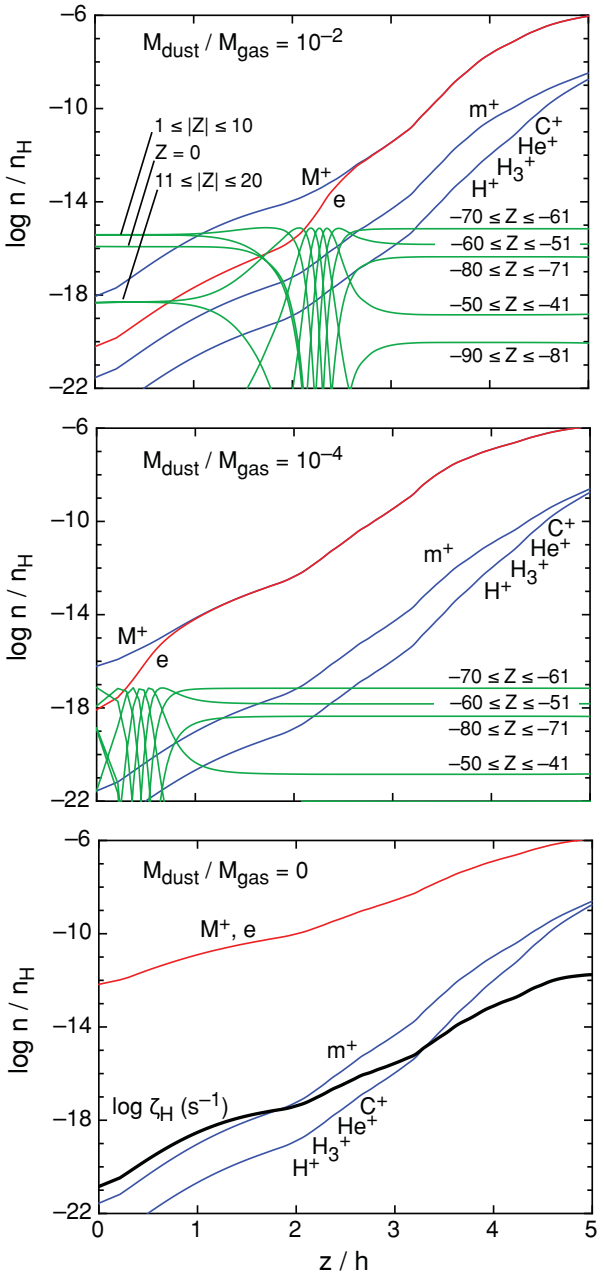

Figure 11. Fractional abundances of charged particles in the MMSN at 1 au from the Sun as a function of height above the mid-plane, assuming that grains have radius 1 μm. The upper, centre and lower panels correspond to dust-to-gas ratios 10^{-2}, 10^{-4} and 0 by mass, respectively. The black curve in the lower panel shows the height dependence of the ionization rate per hydrogen nucleus due to stellar X-rays and interstellar cosmic rays (see text) assumed for all three models. The other curves give the fractional abundances of electrons (*red*), light ions (H^+, H_3^+, He^+ and C^+) representative molecular (m^+) and metal (M^+) ions (*blue*), and grains (*green*, labelled by charge state).

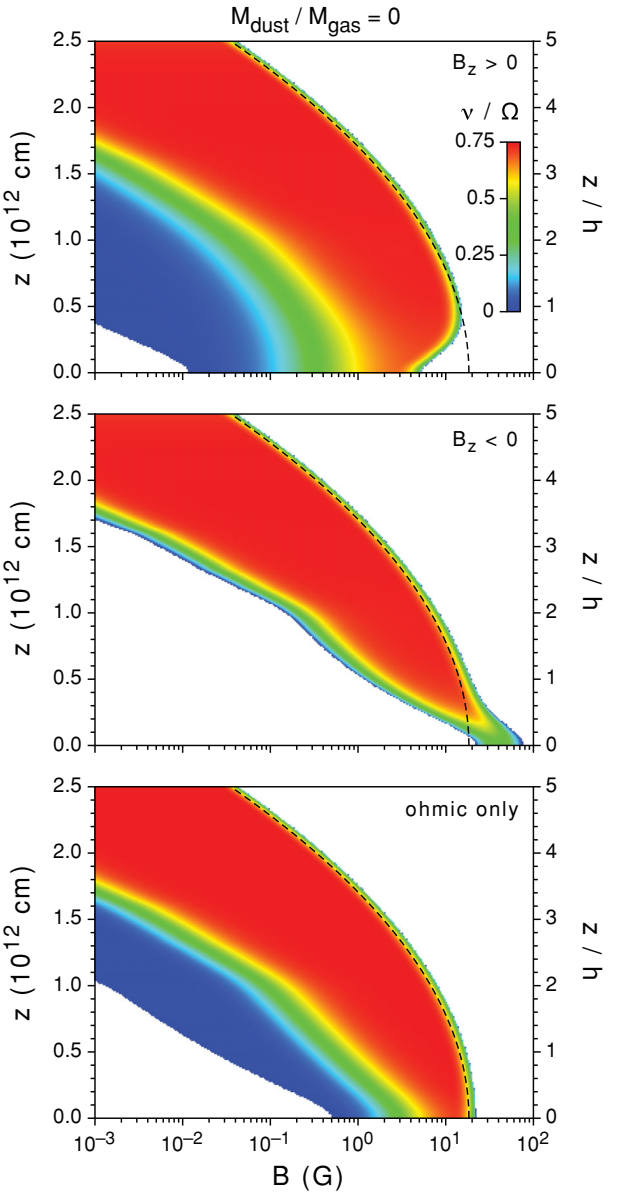

Figure 12. Colour shading shows the growth rate of the fastest growing local MRI mode at height z above the mid-plane in the MMSN at 1 au as a function of the strength of an initially vertical magnetic field B. Dust grains are assumed to have settled to the mid-plane (see lower panel of Fig. 11). The vertical wavenumber k of the modes are required to satisfy $kh > 1$, where h is the disc scaleheight. The unshaded regions show stable combinations of height and field strength. The top and middle panels correspond to the cases where the initial field and rotation axis are parallel ($B_z > 0$) or antiparallel ($B_z < 0$), respectively. The lower panel shows the effect of artificially suppressing ambipolar and Hall diffusion, so including only Ohm diffusion. In this case, there is no dependence on the sign of B_z. The dashed line in each panel indicates the value of the local equipartition magnetic field as a function of height above the mid-plane.

both increase sharply away from the mid-plane. For the ionization profile plotted in Fig. 11, it turns out that $\eta\Omega/v_A^2$ is small near the surface (so that ideal MHD holds) and large near the mid-plane (so that Ohm damping is severe). As a result, near the surface the largest achievable growth rate is close to the ideal value 0.75Ω, but declines rapidly below the height where $\eta\Omega/v_A^2 \sim 1$. The height of the transition between these regimes declines with increasing field strength, simply because $\eta\Omega/v_A^2 \propto B^{-2}$. A second consideration is that the range of MRI-unstable wavenumbers is bounded above by k_c (see equation B12), and this must be larger than $1/h$ if *any* unstable modes are to exist with $kh > 1$. It is this criterion that provides the upper and lower envelopes to the unstable region in Figs 12 and 13. Near the surface where $\eta\Omega/v_A^2 \ll 1$, k_c is approximately $\sqrt{3}\Omega/v_A$, and as v_A rapidly increases with height, the unstable range

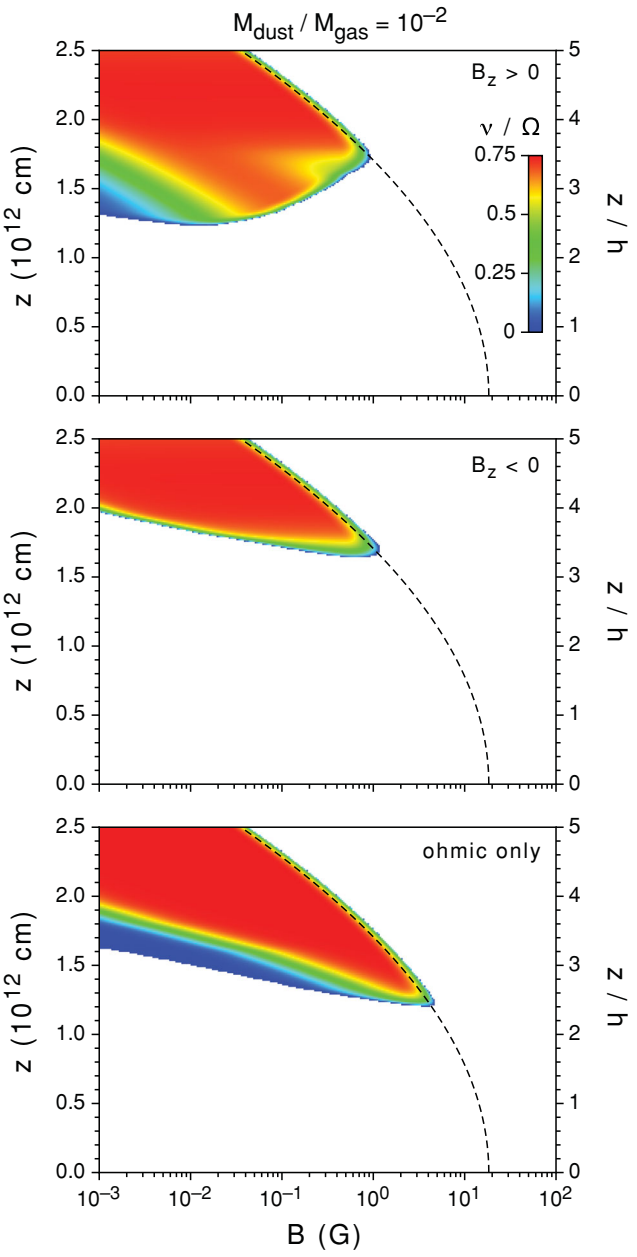

Figure 13. As for Fig. 12, but now including a population of 1-μm-radius grains with total mass 1 per cent of the gas mass (see upper panel of Fig. 11).

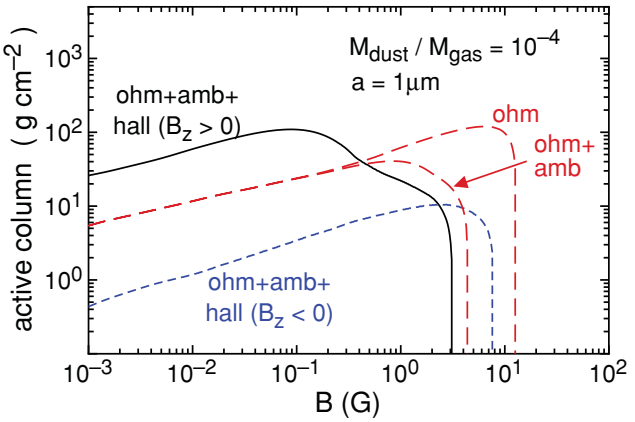

Figure 14. Column density of the MRI-active region at 1 au in the MMSN as a function of magnetic field strength for different assumptions about magnetic diffusion (see text). Grains are assumed to have radius 1 μm with 99 per cent having settled to the disc mid-plane (i.e. total mass only 10^{-4} relative to gas). *Solid black* and *dashed blue* curves are for the magnetic field oriented parallel or antiparallel to the disc's rotation axis, respectively. *Red long-dashed* curves indicate the effect of accounting only for Ohm or Ohm+ambipolar (i.e. Pedersen) diffusion.

shifts to wavenumbers with $kh < 1$, and it is no longer possible to find unstable modes that fit within a scaleheight. This occurs when the magnetic pressure roughly exceeds the gas pressure. By contrast, close to the mid-plane where $\eta\Omega/v_A^2 \gg 1$, $k_c \approx \sqrt{3}v_A/\eta$ and the wavenumbers are pushed out of the relevant range by the increasing diffusivity and the declining Alfvén speed as the mid-plane is approached.

Next, we add the remaining magnetic diffusion terms, i.e. ambipolar and Hall diffusion. The latter depends on the sign of B_z, which we take to be positive or negative in the upper and middle panels of Fig. 12, respectively. Ambipolar diffusion dominates in the surface layers (see fig. 5 of Wardle 2007), and it acts just like Ohm diffusion when, as assumed here, the initial magnetic field is vertical,[5] tending to damp the growth of the MRI and pushing the unstable modes to longer wavelengths. Hall diffusion tends to dominate the lower layers, except for weak fields when Ohm diffusion dominates closer to the mid-plane.

For $B_z > 0$, Hall diffusion is destabilizing and pushes the region where the fastest growth rate is close to 0.75Ω to greater depths. The rapid decline to low growth rates (i.e. from green to blue shading in Fig. 12) traces the transition from Hall-dominated to Ohm-dominated diffusion. Below this, Hall diffusion acts to significantly extend the region of slow growth by modifying k_c, and this extends all the way to the mid-plane for fields in excess of 10 mG.

On the other hand, when $B_z < 0$ (i.e. $\boldsymbol{B}$ is antiparallel to the rotation axis), Hall diffusion tends to stabilize the MRI while extending the unstable wavelengths to shorter wavelengths (see Fig. 6). As a result, superequipartition fields are unstable near the mid-plane. We emphasize that *irrespective of the magnitudes of the other diffusivities, Hall diffusion is completely stabilizing when* $\eta_H > 2v_A^2/\Omega$. This is responsible for the sharp cut-off at the lower boundary in the middle panel of Fig. 12.

When a full complement of 1-μm dust grains are present (i.e. dust-to-gas mass ratio 10^{-2}; charged species as in lower panel of Fig. 11), the diffusivities are greatly increased near the mid-plane because electrons are locked up by grains and rendered immobile. Fig. 13 displays the same trends as in the zero-grain case, but the MRI-unstable region is now restricted to the upper layers of the disc, and in all cases the bulk of the disc is stable to the MRI. While the differences between the three panels might appear less severe in this case, the strong density stratification means that there are orders of magnitude differences in the column density of the MRI-unstable region.

Fig. 14 shows the magnetically active column density as a function of field strength for different assumptions regarding the diffusivity when 99 per cent of grains are assumed to have settled

[5] This is not the case in more general geometries, e.g. ambipolar diffusion may be destabilizing when the field has vertical and toroidal components and the wavenumber has radial and vertical components (see Desch 2004; Kunz & Balbus 2004).

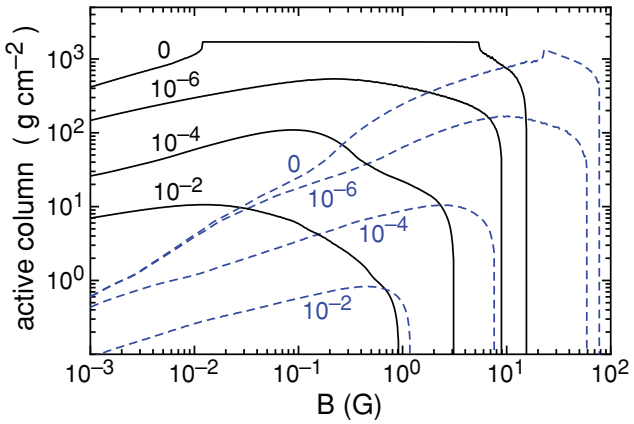

Figure 15. As for Fig. 14 but for dust-to-gas mass ratio varying from 0 to 10^{-2}. Note that active column density is capped at the total column of the MMSN at 1 au, i.e. 1700 g cm^{-2}.

(i.e. dust-to-gas ratio 10^{-4}). The active column density varies by 1–2 orders of magnitude depending on the accuracy of the treatment of magnetic diffusion. First, neglect Hall diffusion and consider either Ohm diffusion alone or Ohm and ambipolar diffusion (i.e. Pedersen diffusion) operating in concert (red long-dashed curves). Ohm diffusion dominates ambipolar diffusion except for strong magnetic fields and low densities, so there is little difference between these two cases except for magnetic fields in excess of 1 G where the additional damping due to ambipolar diffusivity suppresses the MRI. Hall diffusion either increases or decreases the active column density by an order of magnitude depending on whether the initial magnetic field is pointing up or down; the maximum unstable field strength also varies considerably.

Fig. 15 shows how the active column depends on grain abundance and magnetic field orientation. In the absence of grains, the entire disc cross-section is magnetically active for an upwardly directed field in the range 0.01–5 G. On the other hand, when the field is pointing downwards, a minuscule fraction of the disc is active for weak fields, but the active column increases rapidly and encompasses the entire cross-section of the disc for field strengths in the range 20–80 G. As one would expect, grains sharply reduce the active column density because of the reduction in mobile charge carriers. The continued extreme sensitivity to field alignment demonstrates that Hall diffusion still plays a critical role in determining the extent of magnetic activity regardless of the grain abundance.

The reduction in the active column when $B_z < 0$ occurs because Hall diffusion stabilizes the disc against the MRI when $s\eta_H < -2v_A^2/\Omega$. When the grains and ions are strongly coupled to the neutrals and the electrons are coupled to the magnetic field, the Hall diffusivity is given by equation (8), and we obtain a field-dependent criterion for stability on the fractional ionization,

$$\frac{n_e}{n_H} \lesssim \frac{1}{2}\frac{1.4 m_H c\,\Omega}{eB} \approx \frac{1.5\times10^{-11}}{B(\mathrm{G})}\frac{(M/\mathrm{M}_\odot)^{1/2}}{(r/\mathrm{au})^{3/2}}. \quad (28)$$

This critical fractional ionization is typically larger than the corresponding criterion for Ohm damping, so sets the lower boundary of the magnetically active region when $B_z < 0$. Note that once the grain abundance is low enough not to affect the ionization fraction, the lower boundary reaches the mid-plane and the active column density saturates.

These results suggest that Ohm estimates of the column density of the magnetically active layers in protoplanetary discs are in error by about an order of magnitude, systematically underestimating or overestimating the active column if the magnetic field is directed upwards or downwards, respectively.

This conclusion rests on several assumptions. Our neglect of stratification in the linear analysis is unlikely to be serious given that in the Ohm limit this approach yields excellent predictions of the extent of the depth of the turbulent layers (Turner & Sano 2008). The restriction to a simple and somewhat degenerate geometry – i.e. to vertical fields and wave vectors – is perhaps more suspect. However, when ambipolar diffusion is unimportant, a toroidal component of the magnetic field and a radial component of the wave vector can easily be accommodated: a rescaling of the full dispersion relation yields the identical dispersion relation (Pandey & Wardle, in preparation). The situation is less clear when ambipolar diffusion is involved, as in this case it stabilizes or destabilizes the MRI (Desch 2004; Kunz & Balbus 2004) in an analogous manner to the Hall contribution for the vertical field considered here (Pandey & Wardle, in preparation). However, at $\sim$1 au, ambipolar diffusion only dominates near the disc surface where the ionization fraction is so high that diffusion is small in any case. Finally, although the grain model is crude, the key property of the grain population is its capacity to soak up electrons, which is proportional to $\int a\,n(a)\,\mathrm{d}a$ (Wardle 2007). This can be used to scale our results to any grain size distribution. The real uncertainty is the small-radius tail of the grain size distribution which may still be sufficient to reduce the electron and ion fractions at a few scaleheights despite the tendency of grains to aggregate and settle towards the mid-plane.

The robust point is that the relative magnitudes of the diffusivities are independent of the ionization fraction, being controlled by the ratio B/n_H, and that Hall diffusion dominates the other mechanisms over much of protoplanetary discs (Wardle & Ng 1999; Balbus & Terquem 2001; Sano & Stone 2002a). As X-rays and cosmic rays both penetrate to the Hall-dominated region just below the disc surface, it is absolutely essential that Hall diffusion be included in modelling the extent of MRI-driven turbulence in protoplanetary discs. The real uncertainty is that, notwithstanding the pioneering effort of Sano & Stone (2002a,b), the saturation of MRI-driven turbulence in the Hall-dominated regime as yet remains unexplored.

6 SUMMARY AND CONCLUSIONS

In this paper, we re-examined the role of Hall diffusion in suppressing or enhancing MRI-driven magnetic turbulence in Keplerian discs.

We first undertook a local, linear analysis of the MRI, for simplicity restricting our attention to a vertical, weak magnetic field subject to axisymmetric perturbations with a purely vertical wave vector. While this is not new, our approach and presentation differed from previous analyses (Wardle 1999; Balbus & Terquem 2001) in two critical ways. First, we characterized the departure from ideal MHD using magnetic diffusivities rather than conductivities or characteristic frequencies, allowing us to make a clearer connection with the $\boldsymbol{E}\times\boldsymbol{B}$ field-line drift implicit in the diffusive MHD induction equation. Secondly, we presented a clear perspective of the dependence on the Hall and Pedersen (Ohm+ambipolar) diffusivities by examining how the properties of the instability vary over a Hall–Pedersen diffusivity plane (see Fig. 5). This clearly delineates when

the range of unstable wavenumbers is finite or infinite, and whether the fastest growth occurs for finite wavenumber or asymptotically as $k \to \infty$. We also discussed the limiting forms of the dispersion relation, making a connection with the diffusive plane-parallel shear instabilities of Kunz (2008).

Next, we reviewed the alternative parametrizations of non-ideal MHD that have appeared in the literature, and emphasized that existing simulations of the non-linear development and saturation of the instability in the Hall–Ohm case (Sano & Stone 2002a,b) are consistent with expectations based on the simple linear analysis and have not yet probed the Hall-dominated regime characteristic of protoplanetary discs.

Finally, we illustrated the critical effect of Hall diffusion on the size of dead zones in protoplanetary discs by applying a local criterion for growth of the MRI to a simple model of MMSN at 1 au, including X-ray and cosmic ray ionization and a population of 1-μm grains.

Our key results can be summarized as follows.

(i) The radial diffusion of perturbed field lines through the fluid is directly related to the criterion for marginal stability. The MRI is suppressed if the radial drift of field lines against the infall of the fluid is sufficient to restore the field to its equilibrium radial position.

(ii) The behaviour of the MRI for our adopted geometry is determined by the dimensionless Pedersen and Hall diffusivities $\eta_{\rm P}\Omega/v_{\rm A}^2$ and $s\eta_{\rm H}\Omega/v_{\rm A}^2$, where $s = {\rm sign}(B_z)$. For $s\eta_{\rm H} < -2v_{\rm A}^2/\Omega$, Hall diffusion suppresses the MRI irrespective of the value of $\eta_{\rm P}$. For $s\eta_{\rm H} > -2v_{\rm A}^2/\Omega$, the $\eta_{\rm P}$–$s\eta_{\rm H}$ half-plane is divided into three regions in which there is either (I) a finite range of unstable wavenumbers with a fastest growing mode, (II) instability for all wavenumbers, with a unique fastest growing mode and a slower asymptotic growth rate as $k \to \infty$ or (III) instability at all wavenumbers with fastest growth asymptotically achieved as $k \to \infty$ (see Fig. 5).

(iii) For fixed $\eta_{\rm P}$, the maximum growth rate increases with increasing $s\eta_{\rm H}$, from 0 at $s\eta_{\rm H}\Omega/v_{\rm A}^2 = -2$ to the maximum $0.75\,\Omega$ as $s\eta_{\rm H} \to +\infty$. For fixed $\eta_{\rm H}$ with $s\eta_{\rm H} > -2v_{\rm A}^2\Omega$, the growth rate is a maximum for $\eta_{\rm P} = 0$ and declines as $\eta_{\rm P} \to \infty$ (see Fig. 6).

(iv) In the highly diffusive limit, the instability reduces to the Hall-diffusion-driven instability in plane-parallel shear flow discussed by Kunz (2008). Diffusion is so severe in this limit that the perturbations in the fluid velocity do not affect the field evolution, which is driven purely by diffusion and Keplerian shear. Our restriction to vertical initial fields and perturbation wavenumbers enabled us to extend the results of Kunz (2008) to include the damping by Pedersen diffusion and to derive pleasant analytic expressions for the growth rate and wavenumber of the most unstable mode (see equations C7, C9 and C10, respectively, and also Fig. 8).

(v) We argued that simulations of MRI-driven MHD turbulence in the presence of Hall and Ohm diffusion (Sano & Stone 2002a,b) have not yet probed the 'deep' Hall regime $s\eta_{\rm H} \gtrsim \eta_{\rm P} \gtrsim v_{\rm A}^2/\Omega$, where the linear analysis suggests that Hall diffusion allows the instability to proceed when it otherwise would not.

(vi) We found that at 1 au in the MMSN, Hall diffusion changes the magnetically active column density by an order of magnitude. This change is either an increase or decrease depending on whether B is parallel or antiparallel to the rotation axis, respectively. Hall diffusion likely plays a critical role in determining the radial extent of dead zones and the thickness of magnetically active layers in protoplanetary discs, and estimates based on damping by Ohm diffusion are probably wildly inaccurate.

The simplifications adopted in our analysis engender three significant uncertainties in our conclusions that are worth some final discussion.

First, the restriction to the stability of vertical magnetic fields to perturbations with vertical wave vectors does not capture the destabilization by ambipolar diffusion that arises when toroidal field and radial wave vector components are also present (Desch 2004; Kunz & Balbus 2004). However, we have captured the analogous Hall-diffusion-driven destabilization of the MRI, and as Hall diffusion typically dominates this restriction is unlikely to have a great impact.

Secondly, while the use of the linear analysis to predict the boundary of the manifestly non-linear active region appears to be justified for the Ohm case (Turner & Sano 2008), it is not known whether this applies in the Hall-dominated regime that we tout here.

Finally, we caution that the MRI may simply be irrelevant in protoplanetary discs. A minor population of small dust grains would remove so many electrons from the gas phase that the MRI-active column density becomes so small as to be irrelevant. Instead, magnetic activity – if any – may be due to fields lying within the disc and varying on length-scales of order r (Turner & Sano 2008) or a strong poloidal magnetic field brought in during the formation of the disc.

ACKNOWLEDGMENTS

We are indebted to Catherine Braiding, Matthew Kunz, B. P. Pandey, Takayoshi Sano and Jim Stone for stimulating discussion and suggestions, and are grateful for the hospitality provided by the Isaac Newton Institute for Mathematical Sciences at Cambridge University and the School of Physics at the University of Sydney, where some of this work was conducted. This research was supported by the Australian Research Council through Discovery Project grant DP0881066, a Visiting Fellowship from the Isaac Newton Institute at the University of Cambridge and an Outside Studies Programme grant from Macquarie University.

REFERENCES

Adams F. C., Ruden S. P., Shu F. H., 1989, ApJ, 347, 959
Bai X.-N., 2011, ApJ, 739, 50
Bai X.-N., Stone J. M., 2011, ApJ, 736, 144
Balbus S. A., Hawley J. F., 1991, ApJ, 376, 214
Balbus S. A., Hawley J. F., 1992, ApJ, 400, 610
Balbus S. A., Terquem C., 2001, ApJ, 552, 235
Blandford R. D., Payne D. G., 1982, MNRAS, 199, 883
Chandrasekhar S., 1960, Proc. Natl. Acad. Sci., 46, 253
Christensen U. R., 2010, Space Sci. Rev., 152, 565
Combet C., Ferreira J., 2008, A&A, 479, 481
Cowling T., 1957, Magnetohydrodynamics, 1st edn. Interscience, New York
Desch S. J., 2004, ApJ, 608, 509
Draine B. T., Sutin B., 1987, ApJ, 320, 803
Fleming T. P., Stone J. M., Hawley J. F., 2000, ApJ, 530, 464
Fromang S., Papaloizou J., 2006, A&A, 452, 751
Fromang S., Terquem C., Balbus S. A., 2002, MNRAS, 329, 18
Gammie C. F., 1996, ApJ, 457, 355
Glassgold A. E., Feigelson E. D., Montmerle T., Wolk S., 2005, in Krot A. N., Scott E. R. D., Reipurth B., eds, Chondrites and the Protoplanetary Disk, ASP, San Fransisco
Hayashi C., 1981, Progress Theor. Phys. Suppl., 70, 35
Igea J., Glassgold A. E., 1999, ApJ, 518, 848
Ilgner M., Nelson R. P., 2006, A&A, 455, 731
Jin L., 1996, ApJ, 457, 798
Johansen A., Klahr H., 2005, ApJ, 634, 1353

Johnson E. T., Goodman J., Menou K., 2006, ApJ, 647, 1413
Königl A., 1989, ApJ, 342, 208
Krasnopolsky R., Königl A., 2002, ApJ, 580, 987
Kretke K. A., Lin D. N. C., 2010, ApJ, 721, 1585
Kunz M. W., 2008, MNRAS, 385, 1494
Kunz M. W., Balbus S. A., 2004, MNRAS, 348, 355
Matsumura S., Pudritz R. E., 2003, ApJ, 598, 645
Nakano T., Umebayashi T., 1980, PASJ, 32, 613
Nelson R. P., Papaloizou J. C. B., 2004, MNRAS, 350, 849
Pandey B. P., Wardle M., 2008, MNRAS, 385, 2269
Pandey B. P., Wardle M., 2012, MNRAS, doi:10.1111/j.1365-2966.2012.20799.x
Perez-Becker D., Chiang E., 2011, ApJ, 735, 8
Salmeron R., Wardle M., 2003, MNRAS, 345, 992
Salmeron R., Königl A., Wardle M., 2007, MNRAS, 375, 177
Sano T., Stone J. M., 2002a, ApJ, 570, 314
Sano T., Stone J. M., 2002b, ApJ, 577, 534
Sawada K., Matsuda T., Inoue M., Hachisu I., 1987, MNRAS, 224, 307
Semenov D., Wiebe D., Henning T., 2004, A&A, 417, 93
Semenov D., Wiebe D., Henning T., 2006, ApJ, 647, L57
Spitzer L. J., 1941, ApJ, 93, 369
Stone J. M., Hawley J. F., Gammie C. F., Balbus S. A., 1996, ApJ, 463, 656
Turner N. J., Drake J. F., 2009, ApJ, 703, 2152
Turner N. J., Sano T., 2008, ApJ, 679, L131
Turner N. J., Willacy K., Bryden G., Yorke H. W., 2006, ApJ, 639, 1218
Velikhov E., 1959, J. Exp. Theor. Phys. (USSR), 36, 1398
Wardle M., 1999, MNRAS, 307, 849
Wardle M., 2007, Ap&SS, 311, 35
Wardle M., Ng C., 1999, MNRAS, 303, 239

APPENDIX A: FORMULATION

Here we give the MHD equations describing a non-self-gravitating disc orbiting a point mass M. We adopt cylindrical coordinates (r, ϕ, z) centred on M with the disc mid-plane corresponding to $z = 0$ and the disc angular velocity vector parallel to the z-axis. We write the fluid equations in the inertial 'laboratory frame' with the fluid velocity written as $\boldsymbol{v} + \boldsymbol{v}_{\rm K}$, so that $\boldsymbol{v}$ denotes the *departure* of the flow from the Keplerian velocity field

$$\boldsymbol{v}_{\rm K} = \sqrt{\frac{GM}{r}}\,\hat{\boldsymbol{\phi}}\,, \tag{A1}$$

which we eliminate in favour of the Keplerian angular velocity

$$\boldsymbol{\Omega} = \frac{v_{\rm K}}{r}\,\hat{\boldsymbol{z}} \tag{A2}$$

using the identities $\nabla\cdot\boldsymbol{v}_{\rm K} = 0$ and $\nabla\times\boldsymbol{v}_{\rm K} = (1/2)\boldsymbol{\Omega}$. This approach avoids the complexity that would be introduced by transforming to a local non-inertial frame or adopting a shearing-sheet approximation, neither of which are needed in our simple linear analysis in the next section.

The continuity equation becomes

$$\frac{\partial\rho}{\partial t} + \Omega\frac{\partial\rho}{\partial\phi} + \nabla\cdot(\rho\boldsymbol{v}) = 0\,. \tag{A3}$$

The advective term containing Ω corresponds to $(\boldsymbol{v}_{\rm K}\cdot\nabla)\rho$ and represents the azimuthal advection associated with Keplerian rotation – analogous terms will arise in the momentum and induction equations. The momentum equation,

$$\frac{\partial\boldsymbol{v}}{\partial t} + \Omega\frac{\partial\boldsymbol{v}}{\partial\phi} + (\boldsymbol{v}\cdot\nabla)\boldsymbol{v} - 2\Omega v_\phi\hat{\boldsymbol{r}} + \frac{1}{2}\Omega v_r\hat{\boldsymbol{\phi}} = r\Omega^2\hat{\boldsymbol{r}} - \nabla\Phi - \frac{1}{\rho}\nabla P + \frac{\boldsymbol{J}\times\boldsymbol{B}}{c\rho}\,, \tag{A4}$$

also picks up Ω -bearing terms that account for centripetal acceleration and the angular momentum loss implicit in radial motion if the azimuthal speed is Keplerian. The gravitational potential is

$$\Phi = -\frac{GM}{\sqrt{r^2+z^2}}\,, \tag{A5}$$

and we adopt an isothermal equation of state:

$$P = \rho c_{\rm s}^2\,, \tag{A6}$$

where $c_{\rm s}$ is the isothermal sound speed.[6] The current density satisfies Ampères law,

$$\boldsymbol{J} = \frac{c}{4\pi}\nabla\times\boldsymbol{B}\,, \tag{A7}$$

and the magnetic field is, of course, solenoidal,

$$\nabla\cdot\boldsymbol{B} = 0\,, \tag{A8}$$

and evolves according to the induction equation (A9). The qualitative discussion of the effect of field diffusion and the MRI in Section 2 suggests that it will be useful to recast the induction equation in a form that makes explicit the drift of the magnetic field through the fluid, i.e.

$$\frac{\partial\boldsymbol{B}}{\partial t} + \Omega\frac{\partial\boldsymbol{B}}{\partial\phi} + \frac{3}{2}\Omega B_r\hat{\boldsymbol{\phi}} = \nabla\times\left[(\boldsymbol{v}+\boldsymbol{v}_{\rm B})\times\boldsymbol{B} - \eta(\nabla\times\boldsymbol{B})_\parallel\right]\,, \tag{A9}$$

where

$$\boldsymbol{v}_B = \eta_{\rm P}\frac{(\nabla\times\boldsymbol{B})_\perp\times\hat{\boldsymbol{B}}}{B} - \eta_{\rm H}\frac{(\nabla\times\boldsymbol{B})_\perp}{B}\,. \tag{A10}$$

Note that when $\boldsymbol{J}_\parallel$ vanishes, ambipolar and Ohm diffusion behave identically, appearing only together in sum as $\eta_{\rm P}$. This is the case in the linear analysis of the MRI presented here.

Note that the local dissipation rate associated with magnetic diffusion is

$$\boldsymbol{J}\cdot\boldsymbol{E}' = \frac{4\pi}{c^2}\left[\eta\boldsymbol{J}_\parallel^2 + \eta_{\rm P}\boldsymbol{J}_\perp^2\right]\,, \tag{A11}$$

and so, as is well known, Hall diffusion has no associated dissipation. When $\boldsymbol{J}_\parallel$ vanishes, ambipolar and Ohm diffusion behave identically, appearing only together in sum as $\eta_{\rm P}$; in particular, this is the case in the linear analysis of the MRI presented here.

Despite the appearance of Ω-bearing terms, the equations we have derived apply to an arbitrary fluid flow – we have not yet assumed that the flow is close to Keplerian, or even disc-like, but have simply written the fluid velocity in the MHD equations as $\boldsymbol{v} + \boldsymbol{v}_{\rm K}$. Of course, this form of the equations is only really useful for nearly Keplerian flows with $|\boldsymbol{v}| \ll v_{\rm K}$. In the next section, we shall specialize to a near-Keplerian disc equilibrium state with a vertical magnetic field, and consider perturbations with a vertical wave vector.

APPENDIX B: LINEAR ANALYSIS

We consider a small region of an axisymmetric, geometrically thin and nearly Keplerian disc, threaded by a vertical magnetic field, with sound and Alfvén speeds ($c_{\rm s}$ and $v_{\rm A}$) at the mid-plane that are both small compared to the local Keplerian speed $v_{\rm K}$. We assume that radial gradients are on the scale of r and neglect vertical stratification of the initial equilibrium state, so that our analysis only holds near the mid-plane, at heights $z \ll c_{\rm s}/\Omega$. Then, we may neglect the term $r\Omega^2\hat{\boldsymbol{r}} - \nabla\Phi$ in equation (A4), and the remaining radial and azimuthal

[6] This will not really be needed as the mode of interest does not involve pressure fluctuations.

derivatives appearing in equations (A3)–(A9), leaving only partial derivatives in t and z.

The initial state has $\boldsymbol{v} = 0$ (i.e. in Keplerian rotation) with a uniform density, pressure and vertical magnetic field $\boldsymbol{B} = sB\hat{\boldsymbol{z}}$ (where $s = \pm 1$). We linearize the equations around this state and seek solutions for axisymmetric perturbations of the form $\exp(\nu t - \mathrm{i}kz)$. The equations for perturbations in density, pressure and v_z form a separate system that describes vertically propagating sound waves. The system of linear equations in the remaining perturbations involve fluctuations in the r and ϕ components of $\boldsymbol{B}$, $\boldsymbol{v}$ and $\boldsymbol{v}_\mathrm{B}$. We treat these as two-component vectors in the equations below.

In the following equations, we recast physical quantities in dimensionless form by adopting Ω and v_A as the units of frequency and velocity, respectively. For the remainder of this section, η_P, $s\eta_\mathrm{H}$, ν and k denote $\eta_\mathrm{P}\Omega/v_\mathrm{A}^2$, $\eta_\mathrm{H}\Omega/v_\mathrm{A}^2$, ν/Ω and kv_A/Ω, respectively. Then the linearized momentum, field-line drift and induction equations are

$$\frac{\boldsymbol{\delta v}}{v_\mathrm{A}} = \frac{-ik}{(1+\nu^2)} \begin{pmatrix} \nu & 2 \\ -\frac{1}{2} & \nu \end{pmatrix} \frac{\boldsymbol{\delta B}}{B}, \tag{B1}$$

$$\frac{\boldsymbol{\delta v}_B}{v_\mathrm{A}} = -ik \begin{pmatrix} \eta_\mathrm{P} & s\eta_\mathrm{H} \\ -s\eta_\mathrm{H} & \eta_\mathrm{P} \end{pmatrix} \frac{\boldsymbol{\delta B}}{B} \tag{B2}$$

and

$$\begin{pmatrix} \nu & 0 \\ \frac{3}{2} & \nu \end{pmatrix} \frac{\boldsymbol{\delta B}}{B} + \mathrm{i}k \left(\frac{\boldsymbol{\delta v}}{v_\mathrm{A}} + \frac{\boldsymbol{\delta v}_B}{v_\mathrm{A}} \right) = 0. \tag{B3}$$

Using equations (B1) and (B2) to substitute for $\boldsymbol{\delta v}$ and $\boldsymbol{\delta v}_\mathrm{B}$ in equation (B3) yields the relationship between the components of $\boldsymbol{\delta B}$,

$$\delta B_r = -\frac{s\eta_\mathrm{H} + 2/(1+\nu^2)}{\eta_\mathrm{P} + \nu A}\, \delta B_\phi, \tag{B4}$$

where

$$A = \frac{1}{1+\nu^2} + \frac{1}{k^2} \tag{B5}$$

and the dispersion relation

$$ak^4 + bk^2 + c = 0, \tag{B6}$$

where

$$a = (1+\nu^2)\left(\eta_\mathrm{P}^2 + \eta_\mathrm{H}^2\right) + \tfrac{5}{2}s\eta_\mathrm{H} + 2\nu\eta_\mathrm{P} + 1\,, \tag{B7}$$

$$b = (1+\nu^2)\left(2\nu\eta_\mathrm{P} - \tfrac{3}{2}s\eta_\mathrm{H}\right) + 2\nu^2 - 3\,, \tag{B8}$$

$$c = \nu^2(1+\nu^2) \tag{B9}$$

(Wardle 1999).[7] While this form of the dispersion relation is convenient for determining the run of ν with k for a given choice of the diffusivity components, it does not provide an overview of the dependence of the characteristic properties of the instability – the range of unstable wavenumbers, maximum growth rate and corresponding wavenumber – on magnetic diffusion. To this end, we recast the dispersion relation into a form that emphasizes its dependence on the diffusivities η_H and η_P:

$$\left(s\eta_\mathrm{H} + \frac{2}{1+\nu^2} - \frac{3A}{4}\right)^2 + (\eta_\mathrm{P} + \nu A)^2 = \left(\frac{3A}{4}\right)^2. \tag{B10}$$

Conveniently, a given growth rate and wavenumber corresponds to a circular locus in the $s\eta_\mathrm{H}$–η_P plane, centred at $(3A/4 - 2/(1+\nu^2), -\nu A)$ with radius $3A/4$, making this form amenable to graphical analysis.

Using this expression, it is straightforward to show that the maximum possible growth rate is $\nu = 3/4$ (in units of Ω) and that this occurs only when $\eta_\mathrm{P} = 0$. To see this, first note that $\eta_\mathrm{P} \geq 0$ and $A > 0$. Then, for $\nu \geq 0$, the left-hand side of equation (B10) is $\geq \nu^2 A^2$, with equality attained only if $\eta_\mathrm{P} = 0$ and $s\eta_\mathrm{H} = 3A/4 - 2/(1+\nu^2)$. This statement must also hold for the RHS of (B10), so we conclude that $\nu^2 \leq 9/16$ with equality only holding when $\eta_\mathrm{P} = 0$ and $s\eta_\mathrm{H} = 3/4k^2 - 5/4$.

Consider now the dependence of the instability on the Pedersen and Hall diffusivities η_P and η_H (see Fig. 5). First, note that the ideal-MHD limit holds at the origin (i.e. $\eta_\mathrm{H} = \eta_\mathrm{P} = 0$), the Hall MHD limit holds along the horizontal axis ($\eta_\mathrm{P} = 0$) and the Ohm or ambipolar diffusion limits hold along the vertical line ($\eta_\mathrm{H} = 0$). Recall also that only the half-plane $\eta_\mathrm{P} \geq 0$ is physically relevant. Inspection of the dispersion relation, equation (B10), shows that all modes are stable for $s\eta_\mathrm{H} \leq -2$, and in the unstable region ($s\eta_\mathrm{H} > -2$) there is, at most, a single unstable mode for a given choice of wavenumber k. The run of growth rate with wavenumber, i.e. $\nu(k)$, for a particular choice of η_H and η_P, can be found by directly solving (B6) for ν. In practice, it is easier to choose ν and solve the quadratic equation (B6) for k^2. It turns out that there are three distinct forms of the resulting $\nu(k)$ curve, corresponding to regions I, II and III in the $s\eta_\mathrm{H}$–η_P plane, as illustrated in Fig. 5.

In region I, which lies outside the semicircular locus

$$\eta_\mathrm{P}^2 + (s\eta_\mathrm{H} + 5/4)^2 = 9/16\,, \tag{B11}$$

the range of unstable wavenumbers extends from $k = 0$ up to a maximum value

$$k_\mathrm{c} = \left[\frac{3}{2}\,\frac{s\eta_\mathrm{H} + 2}{\eta_\mathrm{P}^2 + (s\eta_\mathrm{H} + 5/4)^2 - 9/16}\right]^{1/2} \tag{B12}$$

found by setting $\nu = 0$ and $k = k_\mathrm{c}$ in the dispersion relation (B10). The form of $\nu(k)$ in this region is illustrated in inset I of Fig. 5. The maximum growth rate ν_0 corresponds to a repeated root for k^2 in the quadratic (B6), so it can be found by setting the discriminant to zero. This yields

$$s\eta_\mathrm{H} = \frac{24\,\nu_0}{9 - 16\nu_0^2}\,\eta_\mathrm{P} - \frac{2}{1+\nu_0^2}\,, \tag{B13}$$

so contours of constant ν_0 in region I are straight lines. The corresponding wavenumber, k_0, is then given by the ratio $-2c/b$, or alternatively $-b/2a$, obtained from equations (B7)–(B9) with ν set to ν_0, i.e.

$$k_0 = \left[\frac{-2\nu_0^2\left(1+\nu_0^2\right)}{\left(1+\nu_0^2\right)\left(2\nu_0\eta_\mathrm{P} - \frac{3}{2}s\eta_\mathrm{H}\right) + 2\nu_0^2 - 3}\right]^{1/2}. \tag{B14}$$

Region I encompasses the limits of ideal MHD, the Ohm and/or ambipolar diffusion limit, as well as the Hall limit (i.e. $\eta_\mathrm{P} = 0$) for $s\eta_\mathrm{H} > -0.5$. Elsewhere, i.e. within the semicircle bounded by (B11) and the $s\eta_\mathrm{H}$-axis ($\eta_\mathrm{P} = 0$), all wavenumbers are unstable. The interior of the semicircle is further subdivided into regions II and III, depending on whether the maximum growth rate is attained at a finite wavenumber or asymptotically as $k \to \infty$, as illustrated in insets II and III in Fig. 5, respectively.

In region II, ν_0 and k_0 still satisfy equations (B13) and (B14) just as in region I. The inner boundary of region II occurs where k_0 just becomes infinite, i.e. ν_0 satisfies equation (B13) and simultaneously the denominator of equation (B14) is zero. These two conditions yield a parametric solution for the locus separating regions II and

[7] There is a typographical error in equation (14) of Wardle (1999) – the RHS should be preceded by a minus sign.

III,

$$s\eta_{\rm H} = -\frac{2\left(9+4\nu_0^2\right)}{\left(1+\nu_0^2\right)\left(9+16\nu_0^2\right)}$$
$$\eta_{\rm P} = \frac{\nu_0\left(9-16\nu_0^2\right)}{\left(1+\nu_0^2\right)\left(9+16\nu_0^2\right)}\,, \qquad \text{(B15)}$$

which trace out an arc from $(\eta_{\rm P}, s\eta_{\rm H}) = (0, -2)$ to $(0, -4/5)$ as ν_0 runs from 0 to 3/4 (the blue locus in Fig. 5).

In region III, all wavenumbers are unstable, with the fastest growth occurring in the limit $k \to \infty$, yielding from (B10)

$$\left(s\eta_{\rm H} + \frac{5/4}{1+\nu_0^2}\right)^2 + \left(\eta_{\rm P} + \frac{\nu_0}{1+\nu_0^2}\right)^2 = \left(\frac{3/4}{1+\nu_0^2}\right)^2 . \qquad \text{(B16)}$$

Thus in region III, contours of constant ν_0 trace out segments of non-concentric circles running between the $s\eta_{\rm H}$-axis and the boundary with region II.

Having delineated these three regions, we now consider how the critical wavenumber $k_{\rm c}$, fastest growth rate ν_0 and corresponding wavenumber k_0 vary across the entire $\eta_{\rm P}$–$s\eta_{\rm H}$ plane. Contours of constant $k_{\rm c}$ are semicircles, as plotted in Fig. 7. While the range of unstable wavenumbers is reduced for large values of $s\eta_{\rm H}$ and $\eta_{\rm P}$, as one might expect, the range is not maximized in the ideal limit (i.e. at the origin) but in regions II and III, bounded by the $k_{\rm c} = \infty$ contour.

Turning now to the fastest growing modes, the maximum growth rate is given either by (B13) in regions I and II or by (B16) in region III, and the corresponding contours are plotted in Fig. 6. The growth rate increases clockwise, from 0Ω along the vertical line $s\eta_{\rm H} = -2$ up to 0.75Ω for the horizontal line $\eta_{\rm P} = 0$ for $s\eta_{\rm H} > -4/5$. In the absence of Hall diffusion, the maximum growth rate ν_0 declines with increasing (Ohm and/or ambipolar) diffusivity (e.g. moving vertically upwards through the $s\eta_{\rm H} = 0$ point in the horizontal axis), with $\nu_0 \approx (3/4)\eta_{\rm P}^{-1}$ for $\eta_{\rm P} \gg 1$. The most important effect of Hall diffusion, apparent from Fig. 6, is that *the growth rate of the MRI exceeds* 0.3Ω *for* $s\eta_{\rm H} \gtrsim \eta_{\rm P}$*, even for arbitrarily large* $\eta_{\rm P}$. More generally, the addition of Hall diffusion at fixed $\eta_{\rm P}$ increases the growth rate if $s\eta_{\rm H} > 0$ and decreases it when $s\eta_{\rm H} < 0$. For large values of $\eta_{\rm P}$, equation (B13) shows that $s\eta_{\rm H}/\eta_{\rm P} \approx 24\nu_0/(9-16\nu_0^2)$. It is this fact that has the potential to modify the extent of dead zones in protoplanetary discs, as we explore later in Section 5.

The wavenumber of the fastest growing mode (blue contours in Fig. 6) decreases as the diffusivity is increased. Again, the contours are not arranged so that the highest wavenumbers occur in the ideal-MHD limit, but to the $s\eta_{\rm H} < 0$ side, within the boundary between regions II and III (traced by the $k_0 = \infty$ contour).

These patterns place the ideal, Ohm (or ambipolar) and Hall regimes in context, and for the first time we see an overview of the effect of magnetic diffusivity on the linear MRI. In particular, there is nothing special about the Ohm/ambipolar limit, e.g. the behaviour of the instability in the presence of diffusion is not qualitatively different for $s\eta_{\rm H} = 2$ versus $\eta_{\rm H} = 0$. Even the ideal-MHD limit does not stand apart as remarkable, although it still holds a special place conceptually because flux freezing holds and it is easier to think about. What does stand out is the part of the plane in the lower left, regions II and III, characterized by high wavenumbers and the spraying out of the growth contours.

Overall, we note that increasing $\eta_{\rm P}$ decreases the maximum growth rate and the characteristic wavenumbers, whereas increasing $s\eta_{\rm H}$ above -2 increases the maximum growth rate and may either increase (when $s\eta_{\rm H} + 1 \lesssim \eta_{\rm P}$) or decrease (when $s\eta_{\rm H} + 1 \gtrsim \eta_{\rm P}$) the corresponding wavenumber.

APPENDIX C: LIMITING CASES AND DIFFUSIVE INSTABILITIES

In this section, we consider the interesting limiting cases of our analysis. As a preliminary, we substitute equations (B1) and (B2) into (B3) to express the linearized induction equation in the form

$$\left[\begin{pmatrix} \nu & 0 \\ \frac{3}{2} & \nu \end{pmatrix} + \frac{k^2}{1+\nu^2}\begin{pmatrix} \nu & 2 \\ -\frac{1}{2} & \nu \end{pmatrix} + k^2\begin{pmatrix} \eta_{\rm P} & s\eta_{\rm H} \\ -s\eta_{\rm H} & \eta_{\rm P} \end{pmatrix}\right]\boldsymbol{\delta B} = 0\,. \qquad \text{(C1)}$$

The three terms in this expression represent the effects of Keplerian shear, fluid displacement and magnetic diffusion on the magnetic field perturbations. Each of the three limiting cases can be obtained by neglecting one of these terms. To obtain the criteria for each limit, we note that $\nu \le 3/4$, and so the three terms have orders of magnitude of $1, k^2$ and $k^2\eta_\perp$, respectively. The three cases of interest are as follows.

Ideal MHD. When $k^2\eta_\perp \ll k^2 \sim 1$, i.e. $\eta_\perp \ll 1, k^2 \sim 1$, the third term is negligible and field evolution is determined by shearing of the field and the response of the fluid to magnetic stresses. This limit applies in the neighbourhood of the origin in Fig. 6, and (of course) recovers the results of Balbus & Hawley (1991) limited to a vertical field and wave vector and neglecting buoyancy.

Cyclotron limit. The first term is negligible when $k^2\eta_\perp \sim k^2 \gg 1$ (i.e. $\eta_\perp \sim 1$ and $k^2 \gg 1$). In this limit, the important effects are advection by the fluid displacement and magnetic diffusion. Generation of B_ϕ from B_r by the Keplerian shear flow is negligible in this limit, but Coriolis and centripetal acceleration still play a crucial role through the dynamics of the fluid which enters via the appearance of the off-diagonal matrix elements in the second term in equation (C1).

This short-wavelength, low-frequency limit corresponds to the cyclotron mode of the magnetized fluid, which has frequency

$$\omega_{\rm H} = \frac{v_{\rm A}^2}{|\eta_{\rm H}|} = \frac{eB}{m_{\rm i}c}\,\frac{\rho_{\rm i}}{\rho}\,, \qquad \text{(C2)}$$

where the second form applies for a simple ion–electron–neutral plasma (Wardle & Ng 1999; Pandey & Wardle 2008). This mode is able to couple effectively to the Keplerian rotation as long as the sense of circular polarization of the mode matches the epicyclic motion, i.e. as long as $B_z\eta_{\rm H} < 0$. The other short-wavelength mode, the high-frequency whistler ($\omega = k^2 v_{\rm A}^2/\omega_{\rm H}$),[8] is unable to couple effectively to the rotation (Wardle 1999).

This limit applies for $k \to \infty$ in regions II and III of Fig. 5, where arbitrarily large wavenumbers are unstable. The dispersion relation in this case is obtained by letting $k \to \infty$ in (B6), and reduces to

$$(1+\nu^2)\left(\eta_{\rm P}^2 + \eta_{\rm H}^2\right) + \tfrac{5}{2}s\eta_{\rm H} + 2\nu\eta_{\rm P} + 1 = 0\,, \qquad \text{(C3)}$$

i.e. $a(\nu) = 0$, with a given by equation (B7). The lack of any k dependence in this regime occurs because both the magnetic diffusion and the magnetic stresses on the fluid (which are responsible for the fluid displacement) scale as k^2.

Diffusive limit. Finally, the second term in (B2) may be neglected when $k^2\eta_{\rm H} \sim 1 \gg k^2$ (i.e. $\eta_\perp \gg 1, k^2 \sim 1/\eta_\perp$). In this case, instability relies on the Keplerian shear flow generating B_ϕ from B_r, and the tendency of Hall diffusion to convert B_ϕ back into B_r. This brings

[8] In dimensionless form, $\omega_{\rm H} = 1/|\eta_{\rm H}|$ and the whistler mode has frequency $\omega = k^2/|\eta_{\rm H}|$.

the potential destabilizing effect of Hall diffusion in shear flows to the fore and shows that it is quite independent of rotational effects – i.e. the Coriolis and centripetal acceleration – that drive the MRI (Kunz 2008). Here our simplified geometry and large diffusion limit allow us to find simple analytic expressions for the growth rate as a function of k for arbitrary diffusivity.

Equations (B1) and (B2) show that $|\boldsymbol{\delta v}| \ll |\boldsymbol{\delta v}_B|$, so that the $\boldsymbol{\delta v}$ term can be neglected in the linearized induction equation (B3), yielding

$$\begin{pmatrix} \nu & 0 \\ \frac{3}{2} & \nu \end{pmatrix} \frac{\boldsymbol{\delta B}}{B} + k^2 \begin{pmatrix} \eta_\mathrm{P} & s\eta_\mathrm{H} \\ -s\eta_\mathrm{H} & \eta_\mathrm{P} \end{pmatrix} \frac{\boldsymbol{\delta B}}{B} = 0\,, \tag{C4}$$

with $\boldsymbol{\delta v}$ still given by equation (B1). In this limit, the field evolves in response to shear and diffusion, without significant feedback from the perturbations that it induces in the fluid flow. The (3/2)-bearing term ($\equiv -\mathrm{d}\ln\Omega/\mathrm{d}\ln r$) is the only manifestation of the Keplerian rotation law and cylindrical geometry in this equation.[9] To generalize this expression, we replace (3/2)Ω by a characteristic shear frequency v' that can be thought of either as accommodating different rotation laws in cylindrical geometry or as representing plane-parallel shear in a Cartesian geometry. In the latter case, r and ϕ are mapped to $-x$ and $-y$, with a local velocity field $\boldsymbol{v} = v'x\hat{\boldsymbol{y}}$; the z-axis is unchanged. Without loss of generality, we may assume that $v' > 0$, and as before, we write $\boldsymbol{B} = sB\hat{\boldsymbol{z}}$, with $s = \pm 1$.

Upon replacing the 3/2 term in (C4) by v', we obtain the dispersion relation in dimensional form:

$$\left(k^2\eta_\mathrm{P} + \nu\right)^2 + \left(k^2 s\eta_\mathrm{H} - v'/2\right)^2 = (v'/2)^2\,. \tag{C5}$$

This immediately shows that instability requires both v' and $s\eta_\mathrm{H}$ to be non-zero and have the same sign, otherwise the second term exceeds the RHS and there are no acceptable solutions. The dispersion relation is easily solved for the growth rate

$$\nu = \sqrt{k^2 s\eta_\mathrm{H}(v' - k^2 s\eta_\mathrm{H})} - k^2\eta_\mathrm{P}, \tag{C6}$$

and in the rotating case thet equilibrium is unstable to wavenumbers running from 0 up to a cut-off

$$k_\mathrm{c} = \sqrt{\frac{s\eta_\mathrm{H}\, v'}{\eta_\perp^2}}\,, \tag{C7}$$

where, as usual, the perpendicular diffusivity is

$$\eta_\perp = \sqrt{\eta_\mathrm{H}^2 + \eta_\mathrm{P}^2}\,. \tag{C8}$$

[9] The velocity perturbations, however, are still affected by rotation (see equation B1) but do not themselves feed back on the field evolution.

In the plane-parallel case, equation (C4) breaks down for $k \gtrsim k_\mathrm{c}$ because the feedback of the velocity perturbations can no longer be neglected. In this case, instead of cutting off, the growth rate reaches a floor value $(\sqrt{\eta}_\mathrm{H}\,/\,\eta_\perp)v'$. To find the fastest growing mode, we play the usual trick: write equation (C5) as a quadratic in k^2 and find ν_0, the value of ν that gives zero discriminant. This means that there is only one corresponding wavenumber k_0^2 and therefore the point (k_0, ν_0) lies on the peak of the curve $\nu(k)$. The wavenumber is found from the quadratic, which is easily solved when the discriminant vanishes. Fortuitously, this procedure yields simple expressions for the maximal growth rate

$$\nu_0 = \frac{s\eta_\mathrm{H} v'}{2(\eta_\perp + \eta_\mathrm{P})}\,, \tag{C9}$$

and the corresponding wavenumber

$$k_0 = \sqrt{\frac{s\eta_\mathrm{H} v'}{2\eta_\perp(\eta_\perp + \eta_\mathrm{P})}}\,, \tag{C10}$$

where we have made use of the identity

$$\frac{\eta_\perp - \eta_\mathrm{P}}{\eta_\mathrm{H}} = \frac{\eta_\mathrm{H}}{\eta_\perp + \eta_\mathrm{P}} \tag{C11}$$

to neatly avoid delicate subtractions when $|\eta_\mathrm{H}| \ll \eta_\mathrm{P}$. Note that

$$\nu_0 = k_0^2\eta_\perp\,, \tag{C12}$$

which we can then use in equation (B4) along with $|\eta_\mathrm{H}|, \eta_\mathrm{P} \gg 1$ to show that the relationship between the perturbed field components in the fastest growing mode is

$$\delta B_x = -\frac{s\eta_\mathrm{H}}{\eta_\mathrm{P} + \eta_\perp}\,\delta B_y\,. \tag{C13}$$

The MRI proceeds in the ideal-MHD limit because shear creates B_ϕ from B_r, $B_r B_\phi$ stresses on the fluid cause it to move radially inwards and outwards, creating more B_r available to be converted to B_ϕ by the shear. In the diffusive limit, the radial component of $\boldsymbol{B}$ is not generated by the response of the fluid to magnetic stresses, but because Hall diffusion converts B_ϕ to B_r. For more general field and wave vector configurations, ambipolar diffusion may play a similar role in assisting or suppressing the MRI, albeit hindered by dissipation (Desch 2004; Kunz & Balbus 2004; Kunz 2008, Pandey & Wardle 2012).

This paper has been typeset from a TEX/LATEX file prepared by the author.