1

CrossMark

LECTURE NOTES

Magnetohydrodynamics of protoplanetary discs

Geoffroy R. J. Lesur †

Univ. Grenoble Alpes, CNRS, IPAG, 38000 Grenoble, France

Protoplanetary discs are made of gas and dust orbiting a young star. They are also the birth place of planetary systems, which motivates a large amount of observational and theoretical research. In these lecture notes, I present a review of the magnetic mechanisms applied to the outer regions ($R \gtrsim 1$ AU) of these discs, which are the planet-formation regions. In contrast to usual astrophysical plasmas, the gas in these regions is noticeably cold ($T < 300$ K) and dense, which implies a very low ionisation fraction close to the disc midplane. In these notes, I deliberately ignore the innermost ($R \sim 0.1$ AU) region, which is influenced by the star–disc interaction and various radiative effects. I start by presenting a short overview of the observational evidence for the dynamics of these objects. I then introduce the methods and approximations used to model these plasmas, including non-ideal magnetohydrodynamics, and the uncertainties associated with this approach. In this framework, I explain how the global dynamics of these discs is modelled, and I present a stability analysis of this plasma in the local approximation, introducing the non-ideal magneto-rotational instability. Following this mostly analytical part, I discuss numerical models that have been used to describe the saturation mechanisms of this instability, and the formation of large-scale structures by various saturation mechanisms. Finally, I show that local numerical models are insufficient because magnetised winds are also emitted from the surface of these objects. After a short introduction on wind physics, I present global models of protoplanetary discs, including both a large-scale wind and the non-ideal dynamics of the disc.

Key words: astrophysical plasmas, plasma instabilities

Contents

† Email address for correspondence: geoffroy.lesur@univ-grenoble-alpes.fr

PART ONE: Observations and physical description

1. Observational context

Recent years have seen a dramatic change in our understanding of protoplanetary discs (PPDs), both from an observational and a theoretical point of view. Observations are now able to resolve the outer regions (radii larger than 1 astronomical unit [AU]) and show the existence of many unexpected features: spiral arms, rings and crescent-like structures. Although these observations mostly probe the distribution of dust grains, they indicate that the gaseous structure of PPDs is much more complex and rich than initially anticipated.
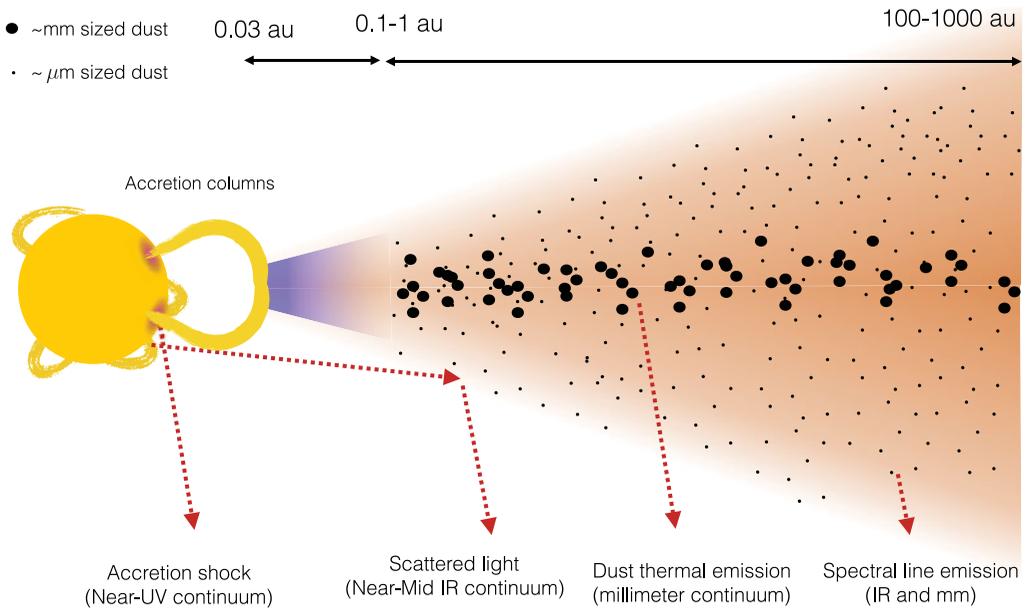
FIGURE 1. PPD diagram showing the various observational diagnostics. Disc winds have been omitted for clarity.

In this part, we review the most recent evidence for PPD structure and evolution, which can be used to constrain the most recent theoretical models.

1.1. *Observational diagnostics*

Today observations probe different regions of the disc. In order to interpret these observations and constrain theoretical models, it is essential to clearly understand the quantities and limitations of each kind of observation. A typical PPD can be separated into two parts: an inner dust-free disc (from a few stellar radii to the dust sublimation radius) made of hot gas (typically more than 1000 K) and an outer disc of gas and dust (figure 1). The disc outer edge can range from 100 AU to more than 1000 AU depending on the object under consideration.

Observations typically probe the following regions:

(i) The UV excess is a signature of the accretion shock at the foot of accretion columns. It is very often the only way to deduce the accretion rate in a specific disc.

(ii) The near and mid-infrared continuum (also known as infrared excess) is a result of stellar photons scattered by small dust grains (typically less than 1 μm in size). Scattered light probes the very surface of the dust layer as the dust disc is very optically thick at these wavelengths. For this reason, the intensity of scattered light is not related to the column density but to the amount of stellar light received by the layer. It therefore characterizes the disc geometry.

(iii) The (sub-)millimetre continuum probes the thermal emission of bigger dust grains (typically with a size of the order of 1 mm). If the dust layer is optically thin at these wavelengths (as usually assumed), the emissivity is related to the column density of dust, but also to its temperature.
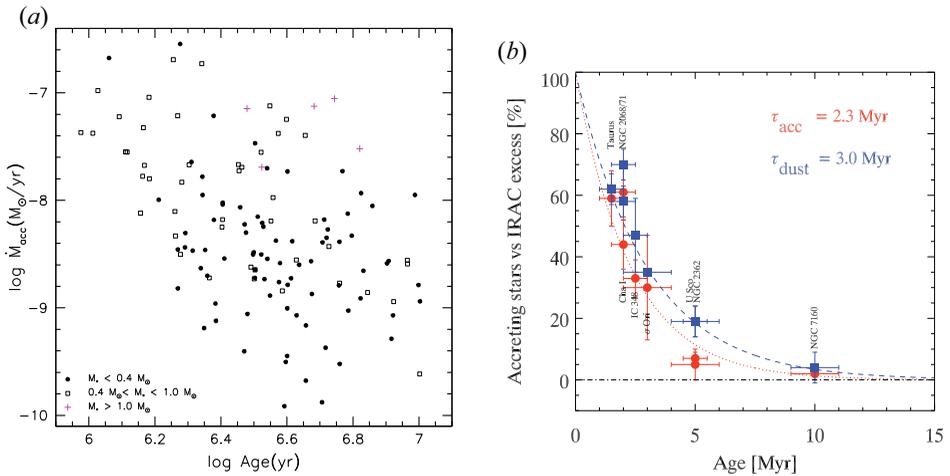
FIGURE 2. (*a*) Measurement of the accretion rate as a function of stellar age in NGC 2264 using the excess UV due to accretion columns. From Venuti *et al.* (2014). (*b*) Fraction of disc signature (accretion) and dust signature (infrared excess) as a function of the cluster age. Both show that discs have an average lifetime of a few million years. From Fedele *et al.* (2010).

(iv) Spectral lines, both in the infrared and at radio wavelengths, probe specific gas tracers such as gas or molecular transitions. These lines are usually optically thick, which implies that they only probe the surface of the gas layer. For this reason, direct estimates of the gas mass in the disc is very difficult, and one has to rely on proxies.

These observational properties are then used to derive several useful dynamical quantities.

1.2. *Accretion*

Because the thermal equilibrium of PPDs for $R \gtrsim 1$ AU is dominated by the illumination of the central star (D'Alessio *et al.* 1998), a direct measurement of the accretion rate through viscous heating is not possible. For this reason, observational evidence of accretion in these regions are scarce and plagued by uncertainties. There are mainly two classes of accretion signature, which are all indirect.

The first is the observational signature of accretion columns at the stellar surface. These accretion columns are formed when the disc material is lifted and accreted by the stellar magnetic field. The gas then ends up in a nearly free-fall speed and hits the stellar surface, forming an accretion shock. The luminosity of this accretion shock observed in UV bands is directly related to the accretion rate in the accretion columns and, therefore, in the innermost disc. It should be kept in mind that accretion rates deduced by this method are not necessarily accretion rates in the entire disc, which can in principle vary with radius if the disc is not in steady state, or if the disc is losing mass from a wind. Typical results show accretion rates of the order of $10^{-8}\,M_{\odot}$/year with uncertainties of the order of an order of magnitude depending on the object under consideration[1] (e.g. figure 2*a*). These accretion rates tend to decrease over timescales of a few million years.

The second observational evidence lies in the proportion of stars showing disc features (accretion on the stellar surface, or infrared excess signifying the presence of dust around

---

[1]Additional sources of uncertainties (not shown here) arise from the method used to reconstruct the mass accretion rate from the UV excess, and from the intrinsic accretion variability of the object (e.g. Venuti *et al.* 2014).
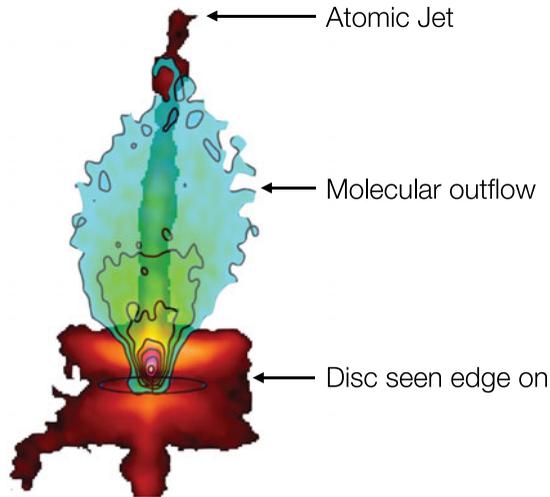
FIGURE 3. Observation of a disc and an atomic jet seen by the Hubble Space Telescope (Burrows *et al.* 1996) and a molecular wind observed in CO(2-1) by ALMA (Louvet *et al.* 2018) in HH30, a PPD seen edge-on. Courtesy of F.Louvet.

the star) as a function of the stellar age. The disappearance of these signatures in older stars allows one to evaluate the typical gas and dusty disc lifetimes. These two time scales do not necessarily match as the gas disc could, for instance, disappear before the dusty disc. However, they both show the same trend: disc tends to disappear on a timescale of a few million years (figure 2*b*).

By combining this information, and assuming that accretion is approximately constant during the lifetime of these objects, one deduces that typical PPD masses range from $10^{-3} M_{\odot}$ to $10^{-1} M_{\odot}$, which is consistent with mass inferred from the total dust content of the disc (Andrews *et al.* 2013).

1.3. *Ejection: winds and jets*

PPDs are often observed in association with large-scale winds and jets. Jets are often seen in forbidden emission lines and correspond to fast collimated flow ($v > 100 \, \mathrm{km \, s^{-1}}$). Their high velocity suggests they are launched from the inner few AU of the disc (Frank *et al.* 2014). The typical outflow rate is estimated to be of the order of 10 % of the accretion rate in classical T-tauri stars (Frank *et al.* 2014).

In addition to these jets, a slower component is also observed in molecular lines. This 'molecular outflow' is denser and reach velocities $v \sim 1$–$10 \, \mathrm{km \, s^{-1}}$ (figure 3). They could be a result of the interaction of the jet with its environment, or they could be a genuine outflow component, emitted from the disc at $R \gtrsim 1$ AU.

1.4. *Structures*

The progress in observational techniques (adaptive optics, interferometry) now allows astronomers to resolve the disc and look for signatures of planet formation, accretion or other unexpected processes. The first class of observations relies on polarimetric differential imaging (PDI) of scattered light emission in the near infrared. This technique allows one to obtain only the light scattered by dust grains (which is naturally polarised) and not the light of the central object. They have been used to probe the disc surface of
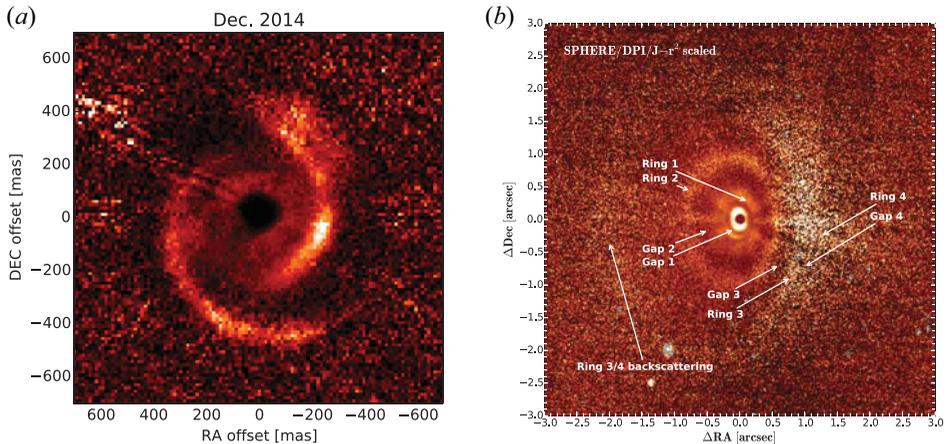
FIGURE 4. Scattered light images in the near infrared using PDI: (*a*) spiral structures observed in MWC758, from Benisty *et al.* (2015); (*b*) multiple ring structures observed in HD97048, from Ginski *et al.* (2016).

various disc (mainly transitional discs). Stunning structures such as spiral and rings were found[2] in several objects (figure 4).

The second class of observation is based on interferometry at millimetric and sub-millimetric wavelengths. The ALMA observatory has been very successful at probing the very structure of PPDs with incredible resolution and unexpected results (figure 5, Andrews *et al.* 2018).

Although these observations probe the dust distribution in the disc, they also tell us about the gas distribution and dynamics, because the grains that are observed are tightly coupled to the gas through a drag force. Such a direct connection has been recently confirmed observationally by simultaneously looking at the continuum (dust) and line emissions (probing the gas kinematics) (Teague, Bae & Bergin 2019).

All these observations indicate that discs are not smooth and symmetrical. They are instead structured on length scales comparable with our solar system. Structures are categorised in spirals, rings and horseshoes, which can be associated with specific physical processes in the disc. It should be noted some of these structures are found in transitional discs, i.e. truncated discs that are presumably in the final evolution stage of PPDs. All of these structures could be the signature of embedded planets perturbing the disc structure by gravitational interaction. However, other processes have been proposed that do not assume planets. One of the key questions is, therefore, whether or not these structures are necessarily a signature of embedded planets.

1.5. *Turbulence*

Turbulence is likely one of the key elements of any dynamical theory for the evolution of discs. Theoretical arguments (see § 4.4) show that turbulence should be subsonic in these systems, i.e., that chaotic motions of the gas are slower than the sound speed. This implies that turbulence is difficult to detect because the turbulent broadening of spectral lines is comparable with the thermal spreading of the molecules constituting the gas. For this reason, heavy molecules such as CN and CO tend to be preferred to detect turbulence, because their thermal velocity is lower compared with lighter molecules at a

---

[2]These structures trace the disc surface and not the column density.
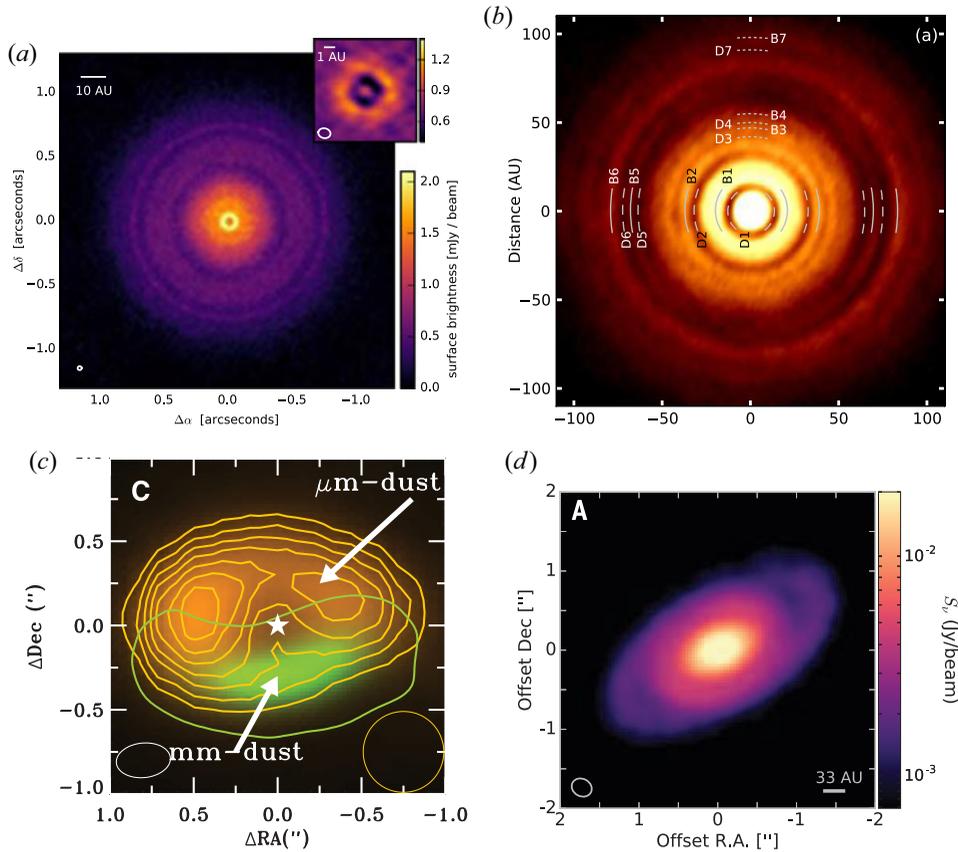
FIGURE 5. (*a*) Ring-like structures observed in TW Hydra. From Andrews *et al.* (2016). (*b*) Multiple ring structure in a deprojected image of HL-Tau from Partnership *et al.* (2015). (*c*) Horsehoe-like structure observed in Oph IRS 48 at sub-millimetre wavelengths (green, tracing millimetre-sized dust) and corresponding scattered light infrared emission (yellow, tracing μm size dust) from van der Marel *et al.* (2013). (*d*) Spiral structures seen at sub-millimetre wavelengths in the young and massive disc of Elias 2-27, from Pérez *et al.* (2016).

given equilibrium temperature. High-resolution spectra obtained from ALMA for CO lines indicates that turbulence is very weak, or non-existent (Flaherty *et al.* 2015, 2017). Spectral broadening smaller than 3 % of the local sound speed are found as best fits to observational data at large distances (typically more than 30 AU). This turbulent broadening is way smaller than the typical values expected from ideal magnetohydrodynamics (MHD) turbulence which typically predicts $\delta v \gtrsim 0.1 c_s$.

Another signature of turbulence (or, more precisely, the lack of turbulence) lies in the dust vertical distribution. Indeed, dust grains naturally tend to settle towards the midplane, unless turbulence stirs them up into the disc atmosphere. Direct measurements of the thickness of the dust layer allow one to deduce the level of hydrodynamical turbulence in the disc. Such a measurement has been done in the case of HL-tau, where the thickness of the rings is used as a tracer for the disc thickness (Pinte *et al.* 2016). The result is that 100 μm grains have settled towards the midplane, with a vertical dust scale height about 10 times smaller than the gas scale height. This implies a very low level of turbulence in the disc, with typically $\delta v \sim 10^{-2} c_s$ ($\alpha \sim 10^{-4}$, see § 4.4).

1.6. *Magnetic fields*

Evidence for magnetic fields in PPDs is scarce. Typical values are expected to be of the order of a Gauss at 1 AU down to a few milli-Gauss at a few tens of astronomical units (Wardle 2007), although these theoretical values could vary by several orders of magnitude. For this reason, measurement through Zeeman effect is unfeasible except in the very inner disc. In this region, toroidal magnetic fields of a few kilo-Gauss have been measured, although it is not clear whether this field belongs to the host star or to the disc itself (Donati *et al.* 2005). At larger distances (tens of astronomical units), attempts at measuring the field strength through Zeeman splitting in molecular lines have only led to upper limits, with $B_z < 0.8$ mG and $B < 30$ mG (Vlemmings *et al.* 2019).

Topological information on the field is also accessible through polarisation in the continuum (i.e. dust thermal emission). It is assumed that dust grains tend to align perpendicularly to magnetic field lines, thereby emitting thermal radiation with a preferred polarisation, perpendicular to the local field orientation (Cho & Lazarian 2007; Stephens *et al.* 2014). However, polarisation in sub-millimetric radiation can also be due to self-scattering by dust grains (Kataoka *et al.* 2015; Yang *et al.* 2016). Campaigns using multiple wavelengths observations have attempted to disentangle these two effects (Stephens *et al.* 2017), but the interpretation of the results in terms of magnetic topology remains very uncertain.

Finally, magnetic field intensities can be deduced from meteoritic and cometary evidence in our own solar system, assuming that the field gets frozen in the body during its formation in the parent disc. Field strength of the order of 0.1 G around 1 AU are inferred from remnant magnetisation in meteorites following this idea (Fu *et al.* 2014), whereas upper limits with $B < 30$ mG in the region around 15–45 AU is deduced from the magnetisation of Comet 67P/Churyumov-Gerasimenko (Biersteker *et al.* 2019).

2. Disc prototype

2.1. *Fluid properties*

PPDs are rather cold objects, with temperatures ranging from 1000 K in the inner (0.1 AU) disc down to 10 K in the outer (100 AU) disc. In order to characterise these discs, It is important to quantify the typical length scales and time scales relevant to the problem. Let us start with a typical disc model which matches disc observations (Andrews *et al.* 2009):

$$\left.\begin{aligned}
\Sigma &= 300 \, R_{AU}^{-1} \, \mathrm{g \, cm^{-2}}, \\
T &= 280 \, R_{AU}^{-1/2} \, \mathrm{K}, \\
\Omega &= 2 \times 10^{-7} \, R_{AU}^{-3/2} \, \mathrm{s^{-1}}.
\end{aligned}\right\} \tag{2.1}$$

Here, we have defined the main physical properties of a disc: its surface density Σ, which correspond to the usual mass density integrated in the direction perpendicular to the disc plane, its temperature T, and its angular velocity Ω around the central object. We also define for convenience a dimensionless distance from the central object, in astronomical units: $R_{AU} \equiv R/1$ AU.

This simple model leads to a $0.04 \, M_\odot$ mass disc, extending from 0.07 to 200 AU, rotating around a solar mass star, typical of discs which have been observed. We can deduce some useful dynamical parameters associated from this simplified models. Defining the isothermal sound speed as $c_s \equiv \sqrt{P/\rho}$ and using the vertical

hydrostatic equilibrium to define the disc vertical scale height (§ 4.2) $H = c_s/\Omega$, one obtains

$$
\left.\begin{aligned}
c_s &= 10^5 \, R_{\mathrm{AU}}^{-1/4} \, \mathrm{cm \, s^{-1}}, \\
H &= 5 \times 10^{11} \, R_{\mathrm{AU}}^{5/4} \, \mathrm{cm}, \\
\frac{H}{R} &= 0.03 \, R_{\mathrm{AU}}^{1/4}, \\
\rho_{\mathrm{mid}} &= 6 \times 10^{-10} \, R_{\mathrm{AU}}^{-9/4} \, \mathrm{g \, cm^{-3}}, \\
n_{\mathrm{mid}} &= 1.5 \times 10^{14} \, R_{\mathrm{AU}}^{-9/4} \, \mathrm{cm^{-3}}, \\
P_{\mathrm{mid}} &= 6 \, R_{\mathrm{AU}}^{-11/4} \, \mathrm{dyn \, cm^{-2}}.
\end{aligned}\right\} \tag{2.2}
$$

2.2. *Magnetic fields*

Magnetic fields in PPDs are poorly constrained (§ 1.6). It is widely believed that fields are largely sub-thermal: the thermal pressure of the fluid dominates over the magnetic pressure (this requirement follows from the fact that the discs are approximately in Keplerian rotation). This translates into a plasma β parameter

$$
\begin{aligned}
\beta &\equiv \frac{P_{\mathrm{th}}}{P_{\mathrm{mag}}} \\
&= \frac{8\pi P}{B^2} \gg 1. \tag{2.3}
\end{aligned}
$$

In practice, $\beta \simeq 1$ constitutes a lower limit for the MRI to operate in geometrically thin discs (see § 6.4.6). Note also that if dynamo action is generating a field (both ordered or disordered), then β cannot reach a value lower than $\beta \sim 1$, hence this value is actually a lower limit for the typical plasma β expected in these discs. It is possible to connect the field strength to the plasma β using the properties defined previously and obtain

$$
B = 12 \, R_{\mathrm{AU}}^{-11/8} \beta^{-1/2} \, \mathrm{G}. \tag{2.4}
$$

The upper bound $B \lesssim 10$ mG for $R \sim 10$ AU mentioned in § 1.6 tend to suggest $\beta \gtrsim 10^4$ in these regions, which confirms that the field strength is expected to be strongly sub-thermal.

2.3. *Fluid approximation*

PPDs are mostly constituted of neutral gas. In order to describe this gas, it is tempting to use the fluid approximation. For this approximation to be valid, the gas under consideration needs to be collisional, i.e. gas particles need to be subject to many collisions during one dynamical timescale. This ensures that at the microphysical level, the velocity distribution of the gas phase can be approximated by a Maxwellian distribution, allowing us to use a scalar pressure field.

Assuming the gas is mainly made of H_2 molecules of radius 10^{-8} cm, we can estimate the cross section of neutral molecules as $\sigma_{nn} = 3 \times 10^{-16}$ cm^2. This gives us an approximate mean free path ℓ_{mfp} and collision frequency ω_{coll}

$$
\left.\begin{aligned}
\ell_{\mathrm{mfp}} &= \frac{1}{n\sigma_{nn}} = 22 \, R_{\mathrm{AU}}^{9/4} \, \mathrm{cm}, \\
\omega_{\mathrm{coll}} &= \frac{v_{\mathrm{th}}}{\ell_{\mathrm{mfp}}} = 5 \times 10^3 \, R_{\mathrm{AU}}^{-5/2} \, \mathrm{s^{-1}}.
\end{aligned}\right\} \tag{2.5}
$$

We therefore have $\ell_{\mathrm{mfp}} \ll R$ and $\omega_{\mathrm{coll}} \gg \Omega$, which validate the fluid approximation to describe the dynamics of PPDs to a very good approximation. It should be noted that these quantities are evaluated at the disc midplane. If one looks at regions well above the disc, as in the case of outflows, ℓ_{mfp} increases significantly. One finds that $\ell_{\mathrm{mfp}} \gtrsim H$ when $n \lesssim 10^4$ cm^{-3}, i.e. when the atmosphere is 10^{10} times less dense than the midplane at 1 AU. Such a strong density contrast is almost never reached in outflow models, where one finds density contrasts between 10^4 and 10^7 (e.g. figure 46). Nevertheless, it should be kept in mind that very weak outflows in the outermost parts of the disc can be close to the collisionless regime.

2.4. *Grain population*

The question of grains is of importance in PPDs. As is usually assumed, we consider a constant dust to gas mass fraction, equal to that of the interstellar medium (1/100). We further assume that grains are spherical with a radius a and made of olivine with a density $\rho_o = 3$ g cm^{-3}. The density of grains is therefore

$$\left. \begin{array}{l} \rho_{\mathrm{grain}} = 6 \times 10^{-12}\, R_{\mathrm{AU}}^{-9/4}\ \mathrm{g\,cm}^{-3}, \\[2mm] n_{\mathrm{grain}} = 1.4\, R_{\mathrm{AU}}^{-9/4} a_{\mu\mathrm{m}}^{-3}\ \mathrm{cm}^{-3}. \end{array} \right\} \tag{2.6}$$

In this last estimate, we have assumed that all the grains had the same size. This is an over-estimation because the sizes are actually distributed over a wide range of scales. In addition, the grain size distribution is expected to evolve with time as grains are known to be growing in PPDs. However, this order of magnitude estimate points to an important fact: the abundance of grains $n_{\mathrm{grain}}/n \sim 10^{-14} a_{\mu\mathrm{m}}^{-3}$. Hence, if grains are smaller than 1 μm, the typical ionisation fraction of PPDs (10^{-14}) suggest that grains are more abundant than free charge carriers. As we show in § 3.4.3, this has a huge effect on the plasma conductivity tensor as grains can become the main charge carriers.

2.5. *Ionisation fraction*

The ionisation fraction $\xi \equiv n_-/n_n$, where n_- is the number of free negative charge carriers, is a highly uncertain quantity, with very little constraints coming from observations. The ionisation fraction typically range from 10^{-16}–10^{-13} at 1 AU to 10^{-13}–10^{-10} at 100 AU. However, the resulting plasma is not necessarily a plasma made of electrons and molecular ions. Indeed, if dust grains are present and sufficiently abundant, they tend to suck electrons and ions in the gas phase, leading to a plasma made of positively and negatively charged grains (Sano *et al.* 2000).

Here, we illustrate how each physical process affects the ionisation fraction by considering a simple chemical network which includes singly charged grains. We combine this network with ionisation rate prescriptions for the various ionisation sources (X-rays, UV, cosmic rays (CRs) and radioactive decay).

2.5.1. *Sources of ionisation*

As we focus on the outer part of PPDs ($R > 1$ AU), the gas is mostly cold with $T < 300$ K. This implies that thermal ionisation (owing to collision between molecules) is inefficient, and one has to rely on non-thermal ionisation processes. Here, we consider the following effects with their associated ionisation rate ζ:

(i) X-ray ionisation owing to bremsstrahlung emission from an isothermal $T = 5$ keV corona localised around the central protostar (Igea & Glassgold 1999; Bai & Goodman 2009, see their equation (21));
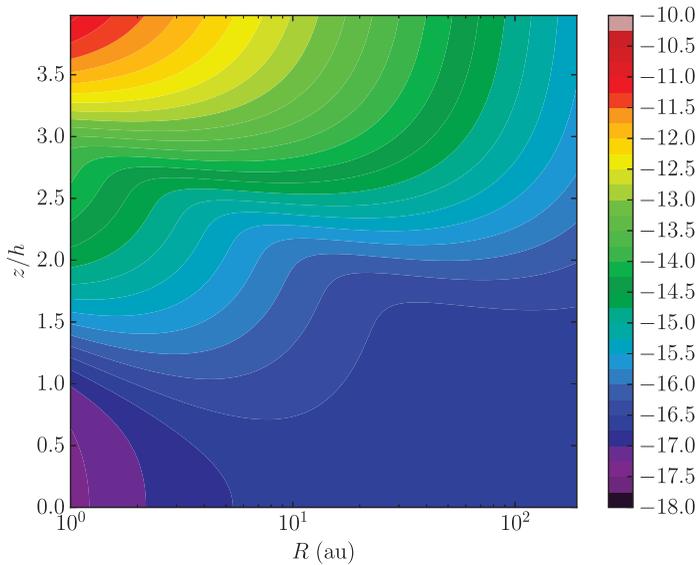
FIGURE 6. Ionisation rate $\log(\zeta)$ (s^{-1}) as a function of radius and altitude (in disc scale height) resulting from X-rays, CRs and radioactive decay.

(ii) CR ionisation with $\zeta_{CR} = \zeta_{CR,0} \exp(-\Sigma/96 \, \mathrm{g \, cm^{-2}}) \, s^{-1}$ (e.g. Umebayashi & Nakano 1981) and $\zeta_{CR,0} = 10^{-17} \, s^{-1}$, corresponding to the interstellar value;

(iii) radioactive decay with $\zeta_{rad} = 10^{-19} \, s^{-1}$ (Umebayashi & Nakano 2009).

The amount of ionisation due to CRs is highly disputed. Some authors have proposed that owing to the wind coming from the young star, CRs are magnetically mirrored from the PPD, resulting in a significantly reduced ionisation rate due to CRs ($\zeta_{CR,0} \sim 10^{-20} \, s^{-1}$, Cleeves, Adams & Bergin 2013). In contrast, it has been proposed that CRs could be accelerated in shocks produced in the protostellar jet by a Fermi process. This could result in ionisation rates as high as $\zeta_{CR,0} \sim 10^{-13} \, s^{-1}$ (Padovani *et al.* 2018). Observations of TW Hya tend to suggest a low ionisation rate owing to CRs ($\zeta_{CR,0} \lesssim 10^{-19} \, s^{-1}$, Cleeves *et al.* 2015), though this is still highly model dependent. Owing to these uncertainties, some authors (e.g.Ilgner & Nelson 2006) have simply omitted CR ionisation and consider only X-rays as the main source of ionisation. These difference and uncertainties in the treatment of the ionisation rate have to be kept in mind when comparing the results of different research groups.

We show in figure 6 the resulting ionisation rate following the disc structure presented in § 2.1. We find that CRs are shielded only in the innermost parts of the disc, where the column density goes above $100 \, \mathrm{g \, cm^{-2}}$. Most of the disc midplane up to $z \sim h$ has $\zeta \simeq \zeta_{CR,0}$, indicating that CRs are indeed the dominant source of ionisation in this region. Above $z \sim h$, X-rays start to penetrate the disc and the ionisation rate rises.

2.5.2. *A simple chemical model*

To illustrate the typical ionisation fractions expected in PPDs, we follow Oppenheimer & Dalgarno (1974), Fromang, Terquem & Balbus (2002) and Ilgner & Nelson (2006) defining the following reaction network and rates with free electrons, neutral molecules m,

molecular ions m^+ and metal atoms M:

$$\left.\begin{array}{rcl}
m + \text{ionising radiation} \rightarrow m^+ + e^- & & \zeta, \\
m^+ + e^- \rightarrow m & & \delta, \\
M^+ + e^- \rightarrow M & & \delta_r, \\
m^+ + M \rightarrow m + M^+ & & \delta_t,
\end{array}\right\} \quad (2.7)$$
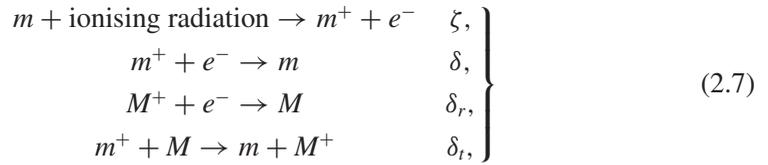
where ζ is the ionisation rate, δ is the dissociative recombination rate for molecular ions, δ_r the radiative recombination rate for metal atoms, and δ_t the rate of charge transfer from molecular ions to metal atoms. Following Fromang *et al.* (2002), we take

$$\left.\begin{array}{l}
\delta_r = 3 \times 10^{-11} T^{-1/2} \text{ cm}^3 \text{ s}^{-1}, \\[4pt]
\delta = 3 \times 10^{-6} T^{-1/2} \text{ cm}^3 \text{ s}^{-1}, \\[4pt]
\delta_t = 3 \times 10^{-9} \text{ cm}^3 \text{ s}^{-1}.
\end{array}\right\} \quad (2.8)$$

In the absence of metals and grains, the rate equations admit a simple solution in steady state:

$$\xi = \sqrt{\frac{\zeta}{\delta n_n}}. \quad (2.9)$$

In the opposite metal-dominated limit, still without grains, one obtains

$$\xi = \sqrt{\frac{\zeta}{\delta_r n_n}}. \quad (2.10)$$

As $\delta_r \ll \delta$, one clearly sees that the absence of metals leads to a dramatic decrease in the ionisation fraction (Fromang *et al.* 2002).

2.5.3. *Typical ionisation fraction profile*

Grain-free case, Metal-free case: Combining (2.9) with the ionisation rate in § 2.5.1, one can obtain the ionisation fraction in the disc. However, this ionisation fraction depends not only on the disc chemistry one assumes but also on the disc structure. A lot of theoretical work has focused on the minimum mass solar nebula (MMSN) model, which assumes $\Sigma = 1700 R_{\text{AU}}^{-3/2}$ g cm^{-2} (Wardle 2007; Bai & Stone 2013*b*; Lesur, Kunz & Fromang 2014). This makes the disc much denser in the inner part, resulting in a stronger shielding of CRs and a lower ionisation fraction than less-dense discs. As an illustration, we show in figure 7 the resulting ionisation fraction with the disc structure presented in § 2.1 and with a MMSN disc model.

We observe that the lowest ionisation fraction reaches 10^{-14} in the MMSN case or 10^{-13} in our disc model. The lowest ionisation fractions are reached in the innermost parts of the disc, where the recombination is the fastest and CRs + X-rays are efficiently shielded. The ionisation fraction progressively increases when X-rays start to penetrate, until one reach ionisation fractions as high as 10^{-6} at a few scale heights. Note that the differences between these models are only significant for $R < 10$ AU because the column densities between the MMSN and our disc model are similar above this radius.

Inclusion of grains and metals: As demonstrated by Elmegreen (1979) and Umebayashi & Nakano (1980) in the context of molecular clouds, and later applied to PPDs (Sano *et al.* 2000; Ilgner & Nelson 2006; Wardle 2007), grains tend to accelerate the recombination
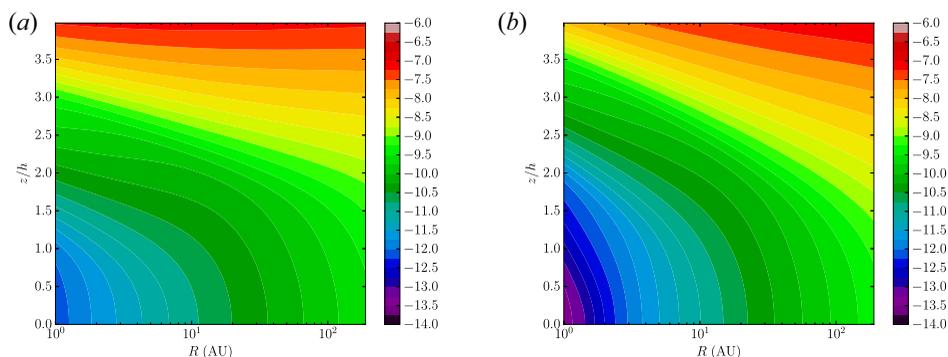
FIGURE 7. Ionisation fraction $\log(\xi)$ as a function of position in (*a*) our disc model (§ 2.1) and (*b*) in a MMSN. Note the difference in ionisation fraction close to the disc midplane for $R < 10$ AU.

of electrons by removing them from the gas phase, resulting in a lower global ionisation fraction, which we have ignored here. To illustrate the effect of grains, let us add the following reactions to our simplified reaction network:

$$
\left.
\begin{aligned}
\text{grain} + m^+ &\to \text{grain}^+ + m, \\
\text{grain}^- + m^+ &\to \text{grain} + m, \\
\text{grain} + e^- &\to \text{grain}^-, \\
\text{grain}^+ + e^- &\to \text{grain}, \\
\text{grain} + M^+ &\to \text{grain}^+ + M, \\
\text{grain}^- + M^+ &\to \text{grain} + M, \\
\text{grain}^+ + \text{grain}^- &\to \text{grain} + \text{grain}.
\end{aligned}
\right\}
\tag{2.11}
$$
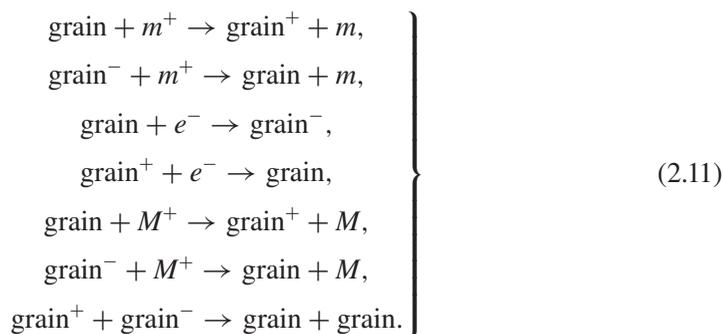
This reaction network only considers singly charged grains, whereas it is well known that grains can have many charges (Ilgner 2012). We chose this approach to illustrate in the simplest model the effect of grains on the ionisation fraction, and later on the diffusivities because the abundance of multiply charged grains is usually lower than that of singly charged grains for $z < h$ (Wardle 2007).

The rates for these reactions are computed by assuming each species collides at its thermal velocity with a spherical grain of radius a (see § 2.4 for more details). We assume a fixed sticking probability of electrons on grains, which corresponds to the probability of bouncing back from a grain.[3]

The resulting ionisation fraction owing to electrons and charged grains are presented in figure 8. We observe essentially two trends. First, the smallest ionisation fraction is found when grains are present, whereas the highest ionisation fractions correspond to grain-free metal-rich cases, with variations owing to this composition effect of the order of three orders of magnitude. Second, the ionisation fraction increases with increasing radius. This effect is not only because the ionisation rate increases, but also because the

---

[3]This probability varies greatly in the literature, from fixed values in Wardle (2007) to various temperature and charge-dependent fits in Ilgner & Nelson (2006) and Bai (2011*a*). Choosing a fixed sticking probability, as we do, tends to increase the effect of grains at high temperature, so the results presented here are a limit case of extreme grain sticking efficiency.
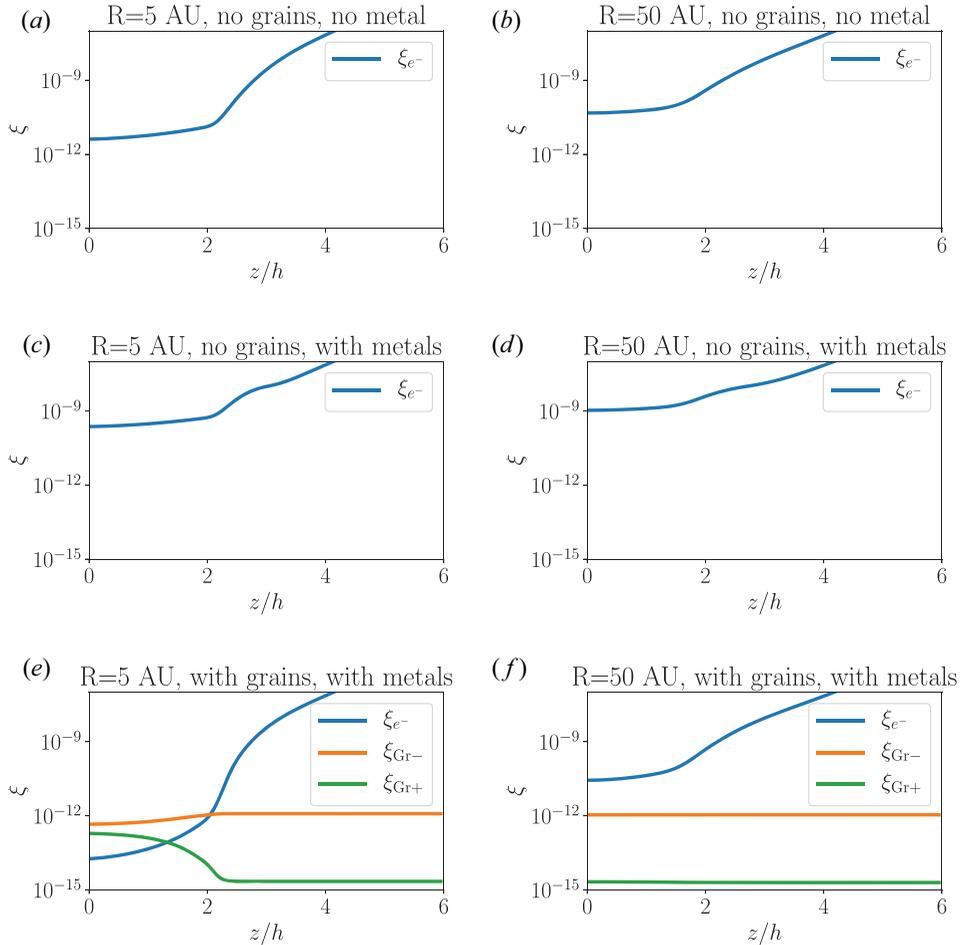
FIGURE 8. Ionisation fraction ξ for three different compositions: row 1 (*a,b*), no grains, no metals; row 2 (*c,d*), no grains with $[M] = 10^{-8}$; row 3 (*e,f*), with $a = 0.1\ \mu\text{m}$ grains and metal atoms. The first column corresponds to $R = 5$ AU and the second column to $R = 50$ AU.

recombination rate decreases owing to lower densities. Let us finally point out that when grains are present, they can become the main charge carrier, as is the case at $R = 5$ AU.

In the following, we use the value $\xi = 10^{-13}$ to evaluate several plasma parameters, keeping in mind this corresponds to a lower bound in our disc model.

3. Plasma description in PPDs

In this section, we explore the properties of the plasma constituting PPDs and ask whether they can be described using non-ideal MHD. For this limit to be valid, we have to satisfy the following three criteria.

  (i) Binary Coulomb interactions should be negligible. This implies that the plasma parameter (defined in the following) is much larger than one.
 (ii) Electro-neutrality is satisfied on timescales of interest, i.e. any charge separation is quickly eliminated by electrostatic interactions.
(iii) The behaviour of each fluid component (electrons, ions, neutrals, charged grains) can be described using a single-fluid approximation.

3.1. *Plasma parameter*

Several quantities allow one to characterise a plasma, the first being the nature of the electromagnetic interaction. The most fundamental quantity characterising a plasma is the Debye length that may be written in an electron–ion plasma

$$
\lambda_{\mathrm{D}} \equiv \sqrt{\frac{k_B T_e}{4\pi(1+Z)n_e e^2}}
$$

$$
= 30 \left(\frac{\xi}{10^{-13}}\right)^{-1/2} R_{\mathrm{AU}}^{7/8}(1+Z)^{-1/2} \text{ cm}, \tag{3.1}
$$

where Z is the averaged number of charges on the ions and we have assumed electro-neutrality so that $n_i = n_e/Z$. The Debye length is clearly below the scales of interest in PPDs. Even if one considers charged grains, the same Debye length can be derived because it does not depend on the particle mass. In addition to this characteristic length, a 'good' plasma should have many particles in a Debye sphere, ensuring the screening of short-range Coulomb interaction. This is quantified by the plasma parameter Υ, equal to the number of charge carriers in a Debye sphere

$$
\Upsilon = 4\pi n_e \lambda_{\mathrm{D}}^3
$$

$$
= 4 \times 10^5 (1+Z)^{-3/2} R_{\mathrm{AU}}^{3/8} \left(\frac{\xi}{10^{-13}}\right)^{-1/2}. \tag{3.2}
$$

Hence, despite the low ionisation fraction and low temperature of these objects, they are still very much in the plasma regime where short-range Coulomb interactions can be neglected. Note, however, that reducing the ionisation fraction and, at the same time, increasing the number of charges could change this picture, breaking the plasma approximation altogether. However, this would require $Z \gtrsim 10^3$ in PPDs, a value that is never encountered, even in chemical models including grains (e.g. Wardle 2007).

3.2. *Electro-neutrality and drag*

PPDs are weakly ionised objects. This implies that the dynamical equations describing the flow and the approximations underlying their derivation should be clearly stated. In this section, we derive these equations, starting from the multi-fluid plasma description. We assume the gas is made of neutral and charged 'particles' (particle could mean electron, ion, or charged grain, indifferently). The multi-fluid approximation is valid because the collision timescales are short, as demonstrated previously. We therefore start from the following dynamical equations:

$$
\frac{\partial n_j}{\partial t} + \nabla \cdot n_j \boldsymbol{v}_j = 0, \tag{3.3}
$$

$$
\frac{\partial n_j m_j \boldsymbol{v}_j}{\partial t} + \nabla \cdot (n_j m_j \boldsymbol{v}_j \otimes \boldsymbol{v}_j) = -\nabla P_j + \boldsymbol{f}_j + n_j q_j \left(\frac{\boldsymbol{v}_j \times \boldsymbol{B}}{c} + \boldsymbol{E}\right) + \boldsymbol{R}_j, \tag{3.4}
$$

where n_j, m_j, $\boldsymbol{v}_j$, P_j, q_j, and $\boldsymbol{f}_j$ are the number density, mass, velocity, pressure, charge, and additional forces (gravity, etc.) on species j. We have also included a drag force $\boldsymbol{R}_j$ between this species and all of the other species. This force is a result of to inter-species collisions

and can be written as

$$R_j = \sum_k \gamma_{jk} \rho_j \rho_k (v_k - v_j), \tag{3.5}$$

because each fluid component is collisional and, therefore, has a Maxwellian velocity distribution. Here, $\gamma = \langle \sigma v \rangle_{jk}/(m_j + m_k)$ and $\langle \sigma v \rangle_{jk}$ is the momentum exchange rate between species j and k. As expected from momentum conservation, we have $\sum_j R_j = 0$.

It is usually assumed that electro-neutrality follows from the fact that the plasma frequency ω_p is much larger than any frequency of interest. Although this is indeed a good criterion for a fully ionised plasma, it is not necessarily true for a weakly ionised plasma. Let us therefore revisit this criterion, starting from the linearised multi-fluid equations. We perturb only one species along the x-axis, leaving the others unperturbed. We moreover assume that the fluid pressure and other external forces are negligible compared with electromagnetic forces. The linearised equation of motion reads

$$\left.\begin{aligned} \frac{\partial \delta n}{\partial t} + n_0 \partial_x v_x &= 0, \\ n_0 m \frac{\partial v_x}{\partial t} &= n_0 q E_x - \gamma m n_0 \rho v_x. \end{aligned}\right\} \tag{3.6}$$

Solving these equations requires an equation for the electric field, which is obtained from one of Maxwell's equation

$$\partial_x E = 4\pi q \delta n. \tag{3.7}$$

We can combine these equations to obtain a second-order relation on the density fluctuation

$$-\frac{\partial^2 \delta n}{\partial t^2} = \omega_p^2 \delta n + \frac{1}{\tau_s} \frac{\partial \delta n}{\partial t}, \tag{3.8}$$

where we have introduced the plasma frequency ω_p and the stopping time τ_s as

$$\left.\begin{aligned} \omega_p &\equiv \left(\frac{4\pi n q^2}{m}\right)^{1/2}, \\ \tau_s &\equiv \frac{1}{\gamma \rho}. \end{aligned}\right\} \tag{3.9}$$

Dynamically, this equation describes damped plasma oscillations with frequencies

$$\omega_{\pm} = \frac{i\tau_s^{-1} \pm \sqrt{4\omega_p^2 - \tau_s^{-2}}}{2}, \tag{3.10}$$

for which we can distinguish two physical limits.

(i) $\omega_p \gg \tau_s^{-1}$ in which case the plasma is subject to plasma oscillations at frequency ω_p with a damping timescale equal to τ_s. If we consider phenomena on frequencies much lower than ω_p, we can average out the highest-order time derivative and obtain a simple closure relation between v_x and E_x: $v_x = q E_x / \gamma m \rho$, which constitutes the base of Ohm's law. Once these oscillations are time-averaged, the plasma can be assumed to be electrically neutral.

(ii) $\omega_p \ll \tau_s^{-1}$ in which case the plasma is subject to over-damped oscillations with two imaginary frequencies $\omega_+ = i\tau_s^{-1}$ and $\omega_- = i\omega_p^2 \tau_s \ll \omega_+$ associated with two

damping timescales $\tau_\pm = (\omega_\pm)^{-1}$. To interpret physically these timescales, let us consider a plasma at rest in which we introduce a localised charge deficit. First, the plasma is going to start moving to 'fill' the charge deficit. Owing to the drag, however, it very rapidly reaches an asymptotic velocity, given by $v_x = qE_x/\gamma m\rho$. Here τ_+ is the time needed by the system to be put in motion and reach this quasi-stationary velocity. This velocity fluctuation, however, is smaller than that which would be obtained in a pure plasma oscillation, because the drag prevents the plasma from reaching high velocities. Hence, it takes a time τ_- to actually fill the charge deficit. This implies that Ohm's law, given by the asymptotic velocity, is valid on timescales longer than τ_+, and that charge inertia can be neglected in that limit. However, charge neutrality is restored on the much longer timescale τ_-.

To summarise, it is possible to neglect inertia for the charged species in the momentum equation provided that the timescales under consideration are larger than $\max(\tau_s, \omega_p^{-1})$, and recover Ohm's law without time derivative. Note that this condition is different from electroneutrality, which requires timescales longer than $\max(\omega_p^{-1}, (\omega_p^2\tau_s)^{-1})$, which are significantly longer than τ_s when $\omega_p\tau_s \ll 1$. It should be pointed out that this analysis was done for a single species, whereas plasmas in PPDs can be made of many different species. Hence, the condition for electroneutrality needs to be satisfied only by the most mobile species of the plasma, which can then compensate for charge fluctuations, and not necessarily by all of the species present.

In PPDs, we obtain the following values for the plasma frequency, depending on the type of charge carrier

$$\left. \begin{aligned} \omega_{p,e} &= 2.2 \times 10^5 \left(\frac{\xi}{10^{-13}}\right)^{1/2} R_{\mathrm{AU}}^{-9/8}\ \mathrm{s}^{-1}, \\[2mm] \omega_{p,i} &= 9.3 \times 10^2 \left(\frac{\xi}{10^{-13}}\right)^{1/2} R_{\mathrm{AU}}^{-9/8}\ \mathrm{s}^{-1}, \\[2mm] \omega_{p,g} &= 1.8 \times 10^{-3} \left(\frac{\xi}{10^{-13}}\right)^{1/2} R_{\mathrm{AU}}^{-9/8} a_{\mu\mathrm{m}}^{-3/2}\ \mathrm{s}^{-1}, \end{aligned} \right\} \tag{3.11}$$

where e, i and g denote electrons, ions and grains, respectively. As can be seen, this frequency is always short compared with the timescales of interest, but grains tend to have significantly lower frequencies owing to their higher inertia.

The stopping times can be estimated starting from the momentum exchange rates $\langle \sigma v \rangle_{ij}$. As we are interested only in weakly ionised plasmas, collisions between charged species will be extremely rare. We therefore only consider neutral-charge collisions.

The 'collision' between electron/ions and neutrals are mainly a result of the electrostatic interaction between the approaching charge and the dipole induced on the neutral by the charge. This is estimated by

$$\left. \begin{aligned} \langle \sigma v \rangle_e &= 8.3 \times 10^{-9} \times \max\left[1, \left(\frac{T}{100\ \mathrm{K}}\right)^{1/2}\right]\ \mathrm{cm}^3\ \mathrm{s}^{-1}, \\[2mm] \langle \sigma v \rangle_i &= 2.4 \times 10^{-9}\ (m_H/m_n)^{1/2}\ \mathrm{cm}^3\ \mathrm{s}^{-1}, \end{aligned} \right\} \tag{3.12}$$

where $\langle \sigma v \rangle_e$ is deduced from Draine, Roberge & Dalgarno (1983) and $\langle \sigma v \rangle_i$ is obtained from Draine (2011), following Bai (2011a).[4] For grains above a size of a few 10^{-2} μm, collisions mainly behave as billiard balls. In other words, σv is roughly equal to the velocity of the incident neutral times the cross-section of the grain. For spherical grains, this leads to[5]

$$\langle \sigma v \rangle_g = \pi a^2 \sqrt{\frac{2k_B T}{m_n}}$$

$$= 2.6 \times 10^{-3} a_{\mu m} \left(\frac{T}{100 \text{ K}} \right)^{1/2} \text{ cm}^3 \text{ s}^{-1}. \tag{3.13}$$

These rates allow us to compute stopping times for each species following the previous definition:

$$\left. \begin{aligned} \tau_{s,e} &= 6.7 \times 10^{-7} R_{AU}^{9/4} \text{ s}, \\ \tau_{s,i} &= 4.9 \times 10^{-5} R_{AU}^{9/4} \text{ s}, \\ \tau_{s,g} &= 8.1 \times 10^{4} R_{AU}^{9/4} a_{\mu m}^2 \left(\frac{100 \text{ K}}{T} \right) \text{ s}, \end{aligned} \right\} \tag{3.14}$$

which shows that because of the low ionisation fraction and the neutral drag, $\omega_{p,j} \tau_{s,j} < 1$ for ions and electrons, whereas it is greater than one for grains. This means that plasma oscillations are over-damped for ions and electrons (case (ii)) and are not directly relevant for quasi-neutrality. Nevertheless, $\omega_p \tau_s > 10^{-2}$, so even in this case, electroneutrality is recovered on timescales shorter than a second. Grains, on the other hand, are usually in regime (i), with a relatively low plasma frequency (period of a few days for 1 μm size grains), decreasing rapidly with increasing grain size. Grains are usually not the only charge carrier in discs, so electro-neutrality is guaranteed by ions and electrons, but it should be kept in mind that, in a hypothetical situation where grains would be the only charge carrier, electro-neutrality could be violated, leading to phenomena similar to lightning. This, however, is not explored here, and we only consider situations where ions and electrons are still present in the system.

3.3. *Single-fluid approximation*

3.3.1. *Dynamical equation for the centre of mass*

The set of equations (3.3) and (3.4) can, in principle, be solved simultaneously (O'Keeffe & Downes 2014). However, it is numerically expensive because the numerical time steps are usually limited by τ_s, which is much smaller than the timescales of interest (as described previously). Note, however, that there are situations where the multifluids approach cannot be avoided, such as when the timescale to reach the ionisation/recombination equilibrium becomes of the order of the timescales of interest (e.g. Ilgner & Nelson 2008), or when the neutral density is so low that the collision timescale τ_s becomes of the order of the timescales of interest, which can occur well above the disc in the early phases of star formation, when X-rays and UV are not yet produced by the central body.

---

[4]The momentum exchange rate $\langle \sigma v \rangle_i$ estimated by Bai (2011a) is actually the collision rate. The momentum exchange rate quoted here is larger by a factor of approximately 1.21 than the collision rate quoted by Bai (2011a) (see Draine (2011, equation (2.39))).

[5]Note that the expression provided by Bai (2011a) for this rate is incorrect by more than three orders of magnitude

However, if one focuses on disc dynamics and its immediate environment once the central star is formed, the single-fluid approximation is a perfectly reasonable approximation, as multi-fluid approaches tend to confirm (Rodgers-Lee, Ray & Downes 2016). For this reason, I will focus here on the single-fluid approximation. To derive this single-fluid approximation, let us consider the dynamical equations for the centre of mass of the fluid, defining the total mass density $\rho = \sum_j n_j m_j$, the flow velocity $\boldsymbol{v} = \sum_j n_j m_j \boldsymbol{v}_j / \rho$ and the drift speed for each species $\boldsymbol{w}_j = \boldsymbol{v}_j - \boldsymbol{v}$ we sum equations (3.3) and (3.4) to obtain

$$\left.\begin{aligned}\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot \rho \boldsymbol{v} &= 0, \\ \frac{\partial \rho \boldsymbol{v}}{\partial t} + \boldsymbol{\nabla} \cdot (\rho \boldsymbol{v} \otimes \boldsymbol{v}) = \boldsymbol{\nabla} \cdot \left(\sum_j n_j m_j \boldsymbol{w}_j \otimes \boldsymbol{w}_j\right) - \boldsymbol{\nabla} P + \boldsymbol{f} + \frac{\boldsymbol{J} \times \boldsymbol{B}}{c} + \sum_j n_j q_j \boldsymbol{E}, \end{aligned}\right\}$$

(3.15)

where we have introduced the total pressure and force P and $\boldsymbol{f}$ as well as the total current $\boldsymbol{J} = \sum_j n_j q_j \boldsymbol{v}_j$. These equations are exact. However, they do not correspond to the usual dynamical equations one is used to, and it is important to understand why each extra term can be neglected.

The first term on the right-hand side corresponds to the transport of momentum by the drift velocity. Physically, it can be interpreted as a diffusion of momentum owing to drifting particles. It can be neglected, provided that drift velocities are small, i.e. that $w_j < L\Omega\sqrt{\rho/\rho_j}$ where L is the typical length scale of interest and Ω is the typical frequency.[6] The presence of the density ratio ensures that even for drift velocities comparable with $L\Omega$, this term is negligible.

We also have a term involving the total charge of the flow $\sum_j n_j q_j$. As shown previously, this term is negligible provided that the timescale of interest is sufficiently long to recover charge neutrality, which is usually the case. We can therefore drop this term altogether to obtain the usual single-fluid equations

$$\left.\begin{aligned}\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot \rho \boldsymbol{v} &= 0, \\ \frac{\partial \rho \boldsymbol{v}}{\partial t} + \boldsymbol{\nabla} \cdot (\rho \boldsymbol{v} \otimes \boldsymbol{v}) &= -\boldsymbol{\nabla} P + \boldsymbol{f} + \frac{\boldsymbol{J} \times \boldsymbol{B}}{c}. \end{aligned}\right\}$$

(3.16)

3.3.2. *Ohm's law*

In the equation of motion for the centre of mass, we have left aside the fact that additional equations were required to obtain $\boldsymbol{B}$ and $\boldsymbol{J}$. Indeed, Maxwell's equations give us

$$\left.\begin{aligned}\frac{\partial \boldsymbol{B}}{\partial t} &= -c\boldsymbol{\nabla} \times \boldsymbol{E}, \\ \boldsymbol{J} &= \frac{c}{4\pi} \boldsymbol{\nabla} \times \boldsymbol{B}. \end{aligned}\right\}$$

(3.17)

The remaining unknown is, therefore, the electric field. Owing to our assumption of electro-neutrality, we cannot use Gauss's law to compute the electric field (since under our scheme of approximation, the total charge density is zero). However, we can use the

---

[6]Comparing the drift velocity with the mean velocity $\boldsymbol{v}$, as is sometimes done, is meaningless because by a Galilean boost, any drift velocity can be made negligible compared with the mean.

dynamical equation for charged species to deduce the electric field that is consistent with quasi-neutrality.

Let us start with (3.3), and let us separate the velocity into a velocity for the centre of mass, and the drift velocity for species j:

$$\rho_j \frac{\mathrm{d}w_j}{\mathrm{d}t} = -\nabla P_j + f_j + n_j q_j \left(\frac{w_j \times B}{c} + E_{\mathrm{cm}} \right) + R_j$$
$$- \rho_j \left[w_j \cdot \nabla v + v \cdot \nabla w_j + \frac{F_{\mathrm{cm}}}{\rho} \right] \tag{3.18}$$

where we have defined the electric field in the centre of mass frame $E_{\mathrm{cm}} \equiv E + v \times B/c$ and the forces on the centre of mass $F_{\mathrm{cm}} = -\nabla P + f + J \times B/c$. Several terms can be neglected here assuming that the stopping time for the species is short compared to the other timescales of the problem.

  (i) $\mathrm{d}_t w_j$ can be neglected provided that $\Omega \ll \tau_s^{-1}$ (i.e. this assumption is identical to the quasi-neutrality assumption discussed previously). In other words, the inertia of charged particles is negligible and they instantaneously reach their asymptotic velocity.

 (ii) Similarly, the inertial term (second line) and external forces f_j can be neglected because they modify the impulsion on timescales long compared with τ_s.

(iii) $\nabla P_j \sim \rho_j c_{s,j}^2 / \Lambda$ is negligible provided that $c_{s,j} \lesssim \Omega \Lambda$.

The equations of motion for charged particles in the frame of the centre of mass therefore read

$$q_j \left(\frac{w_j \times B}{c} + E_{\mathrm{cm}} \right) - \gamma_{jn} m_j \rho w_j = 0, \tag{3.19}$$

where we have assumed that dominant collisions were due to neutrals. This is usually recast as

$$w_j - \mu_j w_j \times \hat{b} = \frac{c \mu_j}{B} E_{\mathrm{cm}}, \tag{3.20}$$

where $\hat{b}$ is a unit vector parallel to B and

$$\mu_j \equiv \frac{q_j B}{\gamma_{jn} \rho m_j c}, \tag{3.21}$$

is the Hall parameter (Wardle & Ng 1999). Equation (3.20) can be solved for w_j, which gives the asymptotic velocity

$$\left. \begin{aligned} w_{j,\parallel} &= \frac{c \mu_j}{B} E_{\mathrm{cm},\parallel}, \\ w_{j,\perp} &= \frac{c \mu_j}{B(1 + \mu_j^2)} \left[E_{\mathrm{cm},\perp} + \mu_j E_{\mathrm{cm},\perp} \times \hat{b} \right]. \end{aligned} \right\} \tag{3.22}$$

We eventually obtain an expression closing our set of equations by relating the drift velocities to the current in the flow $J = \sum_j n_j q_j w_j$ and assuming quasi-neutrality

$\sum_j n_j q_j = 0$:

$$J_\parallel = \frac{c}{B} \left(\sum_j q_j n_j \mu_j \right) E_\parallel,$$

$$J_\perp = \frac{c}{B} \left(\sum_j \frac{q_j n_j \mu_j}{1 + \mu_j^2} \right) E_{\mathrm{cm},\perp} + \frac{c}{B} \left(\sum_j \frac{q_j n_j}{1 + \mu_j^2} \right) \hat{b} \times E_{\mathrm{cm},\perp}. \qquad (3.23)$$

These expressions constitute the base of Ohm's law. We can identify three conductivity tensors, the Ohmic, Hall and Petersen conductivity tensors,

$$\sigma_O = \frac{c}{B} \sum_j q_j n_j \mu_j,$$

$$\sigma_H = \frac{c}{B} \sum_j \frac{q_j n_j}{1 + \mu_j^2}, \qquad (3.24)$$

$$\sigma_P = \frac{c}{B} \sum_j \frac{q_j n_j \mu_j}{1 + \mu_j^2},$$

defined so that Ohm's law can be written in the more familiar form

$$J = \sigma_\parallel E_{\mathrm{cm},\parallel} + \sigma_H \hat{b} \times E_{\mathrm{cm},\perp} + \sigma_P E_{\mathrm{cm},\perp}. \qquad (3.25)$$

This relation can be inverted one final time to obtain the electric field in the observer frame and write the induction equation as

$$\frac{\partial B}{\partial t} = \nabla \times (v \times B) - \nabla \times \left(\eta_O \nabla \times B + \eta_H (\nabla \times B) \times \hat{b} + \eta_A (\nabla \times B)_\perp \right), \qquad (3.26)$$

where the magnetic diffusivities are defined as

$$\eta_O = \frac{c^2}{4\pi} \frac{1}{\sigma_O}, \qquad (3.27)$$

$$\eta_H = \frac{c^2}{4\pi} \frac{\sigma_H}{\sigma_H^2 + \sigma_P^2}, \qquad (3.28)$$

$$\eta_A = \frac{c^2}{4\pi} \left(\frac{\sigma_P}{\sigma_H^2 + \sigma_P^2} - \frac{1}{\sigma_O} \right), \qquad (3.29)$$

where the subscripts O, H and A denotes Ohmic, Hall and ambipolar.

3.4. *Non-ideal diffusivities*

3.4.1. *Simplified case of two charged species*

In the simplest case of a plasma made of two singly charged species $(+)$ and $(-)$, we obtain the following simplified expressions from (3.27), (3.28)

and (3.29):

$$
\eta_O = \frac{cB}{4\pi e n_+} \left(\frac{1}{\mu_+ - \mu_-} \right), \\
\eta_H = \frac{cB}{4\pi e n_+} \left(\frac{\mu_+ + \mu_-}{\mu_- - \mu_+} \right), \\
\eta_A = \frac{cB}{4\pi e n_+} \left(\frac{\mu_+ \mu_-}{\mu_- - \mu_+} \right).
$$

(3.30)

First, all of these coefficients are proportional to n_+^{-1}, i.e. are inversely proportional to the ionisation fraction. Second, because $\mu_j \propto B$, we find that η_O does not depend on B, whereas $\eta_H \propto B$ and $\eta_A \propto B^2$. Finally, we find that η_H may have either sign. If $|\mu_-| > |\mu_+|$, we find $\eta_H > 0$, and $\eta_H < 0$ otherwise. As μ is essentially a measure of the collisionality and mass of the charge carrier, it indicates that the sign of the Hall effect depends on the nature of the charge carriers. In the case where the positive and negative charge carriers have identical masses and γ so that $\mu_- = -\mu_+$, the Hall effect vanishes.

In the case of an electron–ion plasma, we have $|\mu_e| \simeq m_i/m_e |\mu_i| \gg |\mu_i|$. Hence, $\eta_O \propto |\mu_e|^{-1}$ and $\eta_A \propto |\mu_i|$, which justifies the usual statement that Ohmic diffusion is a result of electron–neutral collisions and ambipolar diffusion to ion–neutral collisions. We also have $\eta_H = |\mu_e| \eta_O$ and $\eta_A = |\mu_e \mu_i| \eta_O$. Hence, we can distinguish three regimes depending on the Hall parameter of the ion and electrons:

(i) $1 < \mu_i < |\mu_e|$ in which case $\eta_A > \eta_H > \eta_O$ and the regime is predominantly ambipolar;
(ii) $\mu_i < 1 < |\mu_e|$ in which case $\eta_H > (\eta_A, \eta_O)$ known as the Hall regime;
(iii) $\mu_i < |\mu_e| < 1$ where $\eta_O > \eta_H > \eta_A$ and which is dominated by Ohmic diffusion.

This allows us to delimit the Ohmic, Hall and ambipolar regime as a function of the neutral density and the field intensity (figure 9). As can be seen, the midplane of PPDs is expected to lie mostly in the Hall regime and possibly in the ambipolar regime in the outer-most parts of the disc.

A word of caution though: the physical nature of the Hall effect is different from the Ohmic and ambipolar counterparts (the Hall effect is dispersive, but not diffusive because $\mathbf{J} \times \mathbf{B} \cdot \mathbf{B} = 0$). Being in the Ohmic- or ambipolar-dominated regime does not automatically imply that the Hall effect is dynamically unimportant.

Finally, we obtain the usual expressions for the diffusivities in the electron–ion case:

$$
\eta_O = \frac{c^2 \gamma_{en} m_n m_e}{4\pi e^2} \frac{1}{\xi}
$$
$$
= 2.3 \times 10^{16} \left(\frac{\xi}{10^{-13}} \right)^{-1} \text{ cm}^2 \text{ s}^{-1}
$$

(3.31)

$$
\eta_H = \frac{cB}{4\pi e n_e}
$$
$$
= 5.0 \times 10^{17} \left(\frac{\xi}{10^{-13}} \right)^{-1} \left(\frac{B}{1 \text{ G}} \right) \left(\frac{n_n}{10^{14} \text{ cm}^{-3}} \right)^{-1} \text{ cm}^2 \text{ s}^{-1}
$$

(3.32)
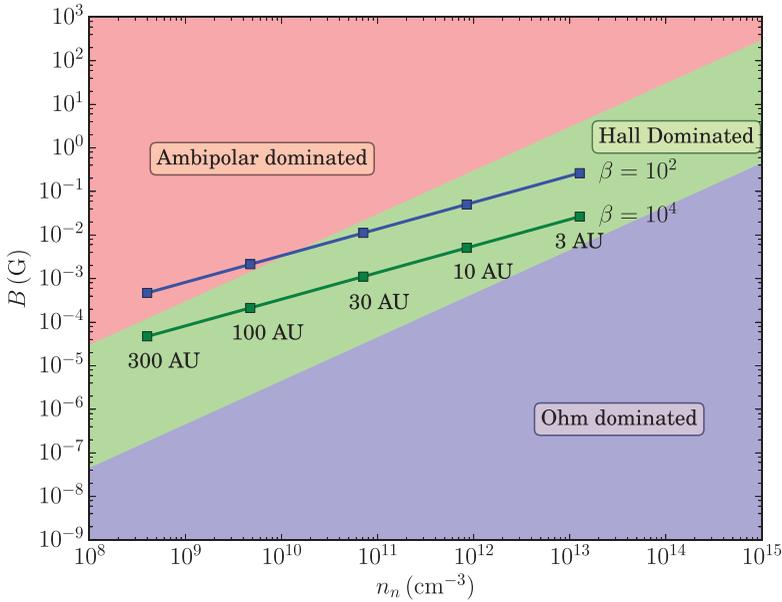
FIGURE 9. Non-ideal regimes as a function of the neutral density and magnetic field intensity, computed for an electron–ion plasma and assuming $T < 100$ K. The blue and green lines correspond to the typical values of a PPD midplane, for various plasma β parameters.

$$\eta_A = \frac{B^2}{4\pi\gamma_{in}\rho\rho_i}$$

$$= 1.6 \times 10^{16} \left(\frac{\xi}{10^{-13}}\right)^{-1} \left(\frac{B}{1\ \text{G}}\right)^2 \left(\frac{n_n}{10^{14}\ \text{cm}^{-3}}\right)^{-2}\ \text{cm}^2\ \text{s}^{-1}. \qquad (3.33)$$

These values can be compared with diffusivities of everyday material such as iron ($\eta = 8 \times 10^2$ cm^2 s^{-1}), demineralised water ($\eta = 1.4 \times 10^{15}$ cm^2 s^{-1}) and dry air ($\eta = 1.6 \times 10^{24}$ cm^2 s^{-1}). Even though one might wrongfully conclude from this that MHD effects are irrelevant, the time scales ($\sim$1 year) and length scales ($\sim$1 AU) are also much larger than conventional everyday experiments. This illustrates the fact that dimensionless numbers should be compared and not dimensional quantities. As we show, one obtains magnetic Reynolds numbers of $O(1)$, which put these flows in a regime comparable with liquid sodium experiments on Earth.

3.4.2. *Dimensionless numbers and application to disc models*

It is customary to define dimensionless numbers in association with non-ideal effects in order to quantify their relative importance in the induction equation. First, one can define Elsasser numbers

$$\Lambda_{O,H,A} \equiv \frac{V_A^2}{\Omega\eta_{O,H,A}}, \qquad (3.34)$$

where V_A is the Alfvén speed and Ω is the rotation rate of the system. Note, however, that $\Lambda_O \propto B^2$ and $\Lambda_H \propto B$, which makes these numbers less useful when it comes to predicting the saturation of MHD instabilities because B is *a priori* unknown. It is therefore useful to define two additional dimensionless numbers, the magnetic Reynolds number and the
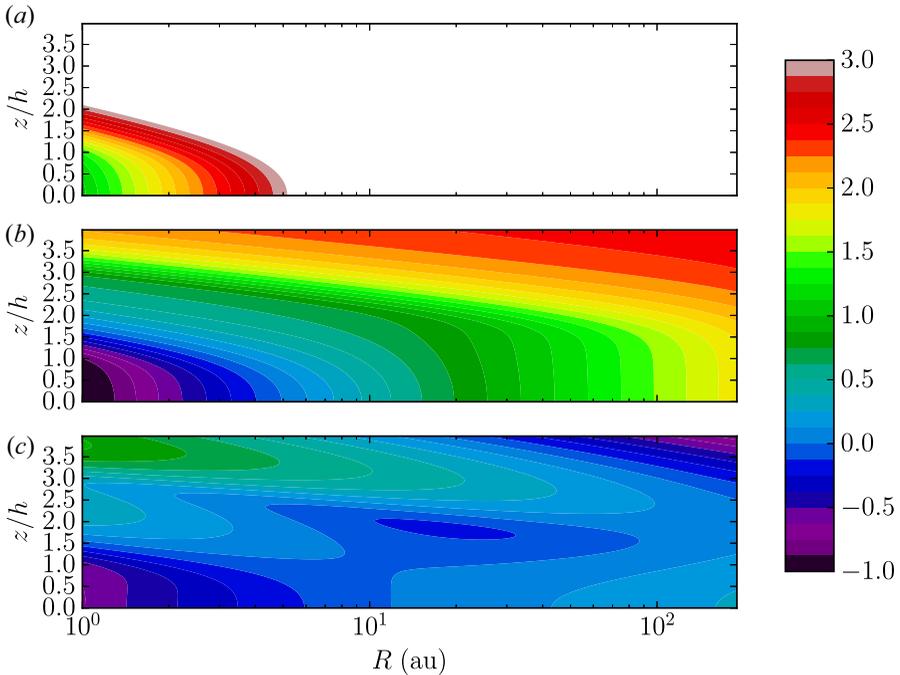
FIGURE 10. Magnetic Reynolds number Rm (*a*), Hall Lundquist number $\mathcal{L}_H$ (*b*) and ambipolar Elsasser number Λ_A (*c*) in our disc model (§ 2.1), using a simple ion–electron approximation with a metal-free chemistry. White values are $> 10^3$.

Hall Lundquist number

$$\left.\begin{aligned} Rm &\equiv \frac{\Omega H^2}{\eta_O}, \\ \mathcal{L}_H &\equiv \frac{V_A H}{\eta_H}. \end{aligned}\right\} \tag{3.35}$$

These two numbers do not depend on the field strength (at least in the two-species plasma case), and they turn out to be excellent saturation predictors in the non-linear regime of the MRI. We show in figure 10 the dimensionless numbers resulting from our grain-free metal-free ionisation model. As it can be seen, $Rm < 10^3$ only in the innermost regions of the disc. This is the region which was historically defined as the 'dead zone' (Gammie 1996). In addition, we find $10^{-1} < \mathcal{L}_H < 10$ in most of the disc midplane with a sharp increase at the disc surface whereas $\Lambda_A \simeq 1$ in most of the disc.

3.4.3. *The role of grains*

When it comes to the conductivity of PPDs, grains play essentially two roles.

(i) By capturing free electrons, they become predominantly negatively charged, and they increase the recombination rate with ions thanks to their large cross-section and reaction rates at the grain surface. The end product is generally a reduced ionisation fraction, possibly by several orders of magnitude (see § 2.5.3).

(ii) Owing to their high inertia, charged grains enter into the conductivity tensor as a very low Hall parameter species. In this case, the scaling laws obtained for the

two species case do not hold anymore. The abundance of charged grains, therefore, changes the diffusion regime in which the system lies.

To illustrate how the diffusivity depends on the presence of grains, we present in figure 11 an example of diffusivity computation in plasmas of different compositions (these compositions are identical to those discussed in § 2.5.3). As can be seen, the addition of $0.1\,\mu$m size grains to the system has a dramatic effect on the diffusivities: all of the dimensionless numbers decrease by several orders of magnitude close to the midplane, whereas ambipolar diffusion becomes stronger than Hall in the case with grains. Overall, ambipolar diffusion increases by 10^4 whereas Hall and Ohmic increase by 10^2 compared with the fiducial metal-free case. The Hall effect also changes sign at the disc surface. This arises because of the presence of negatively charged grains, which contribute to the disc conductivity tensor by reducing the 'effective' Hall parameter of negative charge carriers (electrons + grains−). In the end, the Hall conductivity becomes dominated by ions when they become sufficiently abundant: at the disc surface.

In the grain-free case, the presence of metal atoms tends to decrease the diffusivities by typically one to two orders of magnitude. However, let us point out that the effect of metal atoms disappears once grains are sufficiently abundant (Ilgner & Nelson 2006). Our grain-free metal-rich model is, therefore, a best-case scenario for the ionisation fraction and the diffusivities.

The simplified grain model we have used is by no mean the final answer to this question. However, it demonstrates the strong effect of grains on the dynamics of the plasma. The intensity of this effect also depends on the grain size and grain abundance, a lower abundance or larger grain size leading to a smaller effect (Ilgner & Nelson 2006; Salmeron & Wardle 2008). Here, we have purposely chosen very small grains with an interstellar abundance to illustrate a worst-case scenario for the ionisation fraction. Finally, if one assumes polycyclic aromatic hydrocarbons (PAHs) are present in the gas phase, they then behave as very small grains, capturing all of the floating electrons and also affecting significantly the amplitude of non-ideal effects (Bai 2011*b*).

3.4.4. *Conclusion on non-ideal MHD effects*

Overall, there is no general consensus on the quantitative strength of non-ideal MHD effects in the outer part ($R > 1$ AU) of PPDs. It is clear that these effects are qualitatively very important though, and that these objects are far from the ideal MHD regime. Let us summarise here the source of uncertainty and their implication for the strength of non-ideal effects.

Ionisation rate: Because CRs can be both shielded by the stellar wind or 'locally' produced in shocks surrounding the forming star, there is tremendous uncertainty of six orders of magnitude on the ionisation rate due to CRs (see the discussion in § 2.5.1). As this mechanism is the main source of ionisation in the disc below two scale heights, this implies a three orders of magnitude uncertainty in the ionisation fraction ξ and similarly on the diffusivity coefficients. Ionisation as a result of X-rays is also subject to caution because the X-ray flux coming from the star is largely variable, leading to order of magnitude fluctuations of the ionisation fraction close to the disc surface.

Disc structure: The disc structure is a fundamental parameter that determines the penetration depth of ionising radiations, but also the recombination rate. Denser disc models, such as the MMSN, tend to have lower ionisation fractions and larger diffusion. We have shown that by comparing a theoretical MMSN disc model with a model favoured by observations (§ 2.1), one can change the ionisation fraction by two orders of magnitude
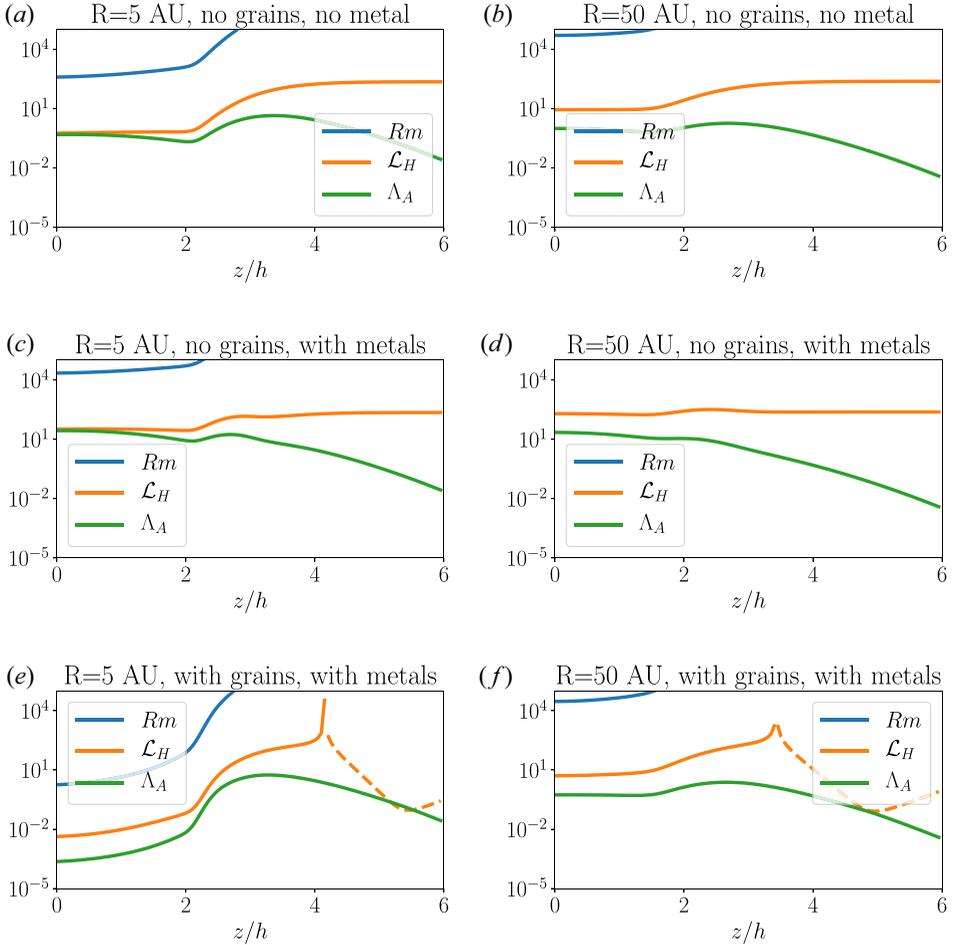
FIGURE 11. Dimensionless diffusivity for three different compositions: row 1 (*a,b*), no grains, no metals (identical to figure 10); row 2 (*c,d*), no grains with $[M] = 10^{-8}$; row 3 (*e,f*), with $a = 0.1\,\mu$m grains and metal atoms. The first column corresponds to $R = 5$ AU and the second column $R = 50$ AU. Dashed lines correspond to negative diffusivities for the Hall effect.

(see § 2.5.3). As the gas column density profile is largely unknown for $R \sim 10$ AU, one is forced to use the gas column density as a free parameter.

Grains: As shown previously, grains affect both the ionisation fraction and the dependence of the diffusivities on ξ. Overall, grains tend to reduce the ionisation fraction by several orders of magnitude (typically two to three). Owing to the change in composition (grains become the dominant charge carrier close to the midplane), diffusivities increase by two to four orders of magnitude in our worst-case scenario, compared with the metal-free case. In addition, the Hall diffusivity η_H can be reversed. The effect of grains naturally depends on the assumed grain size and abundance. It is usually found that grain size $a > 1\,\mu$m does not affect the conductivity tensor too much (Salmeron & Wardle 2008) and that a significant depletion of small grains also reduces their effect (Ilgner & Nelson 2006). All of these calculations assume all of the grains have the same size, which most presumably overestimates the abundance of grains and their effect on the conductivity tensor. More realistic grain size distribution, including more complex chemical reaction networks (e.g. Thi *et al.* 2019) tend to obtain diffusivities that

are within an order of magnitude of the diffusivities discussed in our grain-free metal-free scenarios.

Finally, the physics at the grain surface is poorly understood, which gives a lot of freedom to chemical models. For instance, some authors assume a fixed sticking coefficients of electrons and ions on dust grains as we do (Sano *et al.* 2000; Wardle 2007), whereas others include dependences as a function of the grain size, charge and temperature (Ilgner & Nelson 2006; Bai 2011*a*); note, however, that the dependency of the sticking coefficient on the grain charge differs significantly between these authors. This can have an additional order of magnitude effect on the resulting diffusivities.

Overall, one is forced to conclude that the conductivity tensor of PPDs is plagued by uncertainties and that no chemical/ionisation/grain model is better than the other. Given the previous discussion, the uncertainty on the diffusion coefficient is at least ± 3 orders of magnitude, which has dramatic effects on the dynamical behaviour of these objects. Until more constraints are obtained for these coefficients, theoreticians are forced to explore in a more or less systematic manner the parameter space of the conductivity tensor.

Owing to these uncertainties, we focus in the following on the 'intermediate' case of a metal-free grain-free case that we discussed in § 3.4.2 and for which diffusivities are given in figure 10.

PART TWO: Disc dynamics: global and local views

4. Introduction

4.1. *Motivations*

Explaining accretion in discs is a long-standing problem of modern astrophysics. Even though angular momentum transport equations have been known for a long time, the road to quantifying the level of stress in various discs has been paved with unforeseen difficulties. The main idea that has been followed since the pioneering work of Shakura & Sunyaev (1973) is that accretion discs are somehow turbulent, and this turbulence generates a radial stress. The key is then to relate this radial stress to the other large-scale quantities such as the disc surface density Σ and thickness H, the rotation rate Ω, the diffusivities and the magnetic field strength. This approach is, in essence, very similar to the mixing length theory of convection, except that in the disc case, one does not transport heat but angular momentum. In discs, it is called the α-disc theory.

Here, we present the basic concepts behind accretion and the α-disc theory. Then, we introduce the magneto-rotational instability (MRI), which is probably the most promising instability to explain the origin of accretion in astrophysical discs. Finally, we apply the MRI in the context of PPDs, taking into account non-ideal MHD effects.

4.2. *Disc equilibrium*

A PPD is typically made of gas (and possibly dust) orbiting a young stellar object of mass M. Here, we assume that the gravity of the orbiting gas onto itself (self-gravity) is negligible. This is not necessarily true in very massive discs or in the outer parts of young class 0 objects. Under these assumptions, the gravitational potential is simply that of the central object and the equilibrium may simply be written as

$$\left. \begin{array}{l} 0 = -\dfrac{1}{\rho}\dfrac{\partial P}{\partial R} - \partial_R \psi + \Omega^2 R, \\[2mm] 0 = -\dfrac{1}{\rho}\dfrac{\partial P}{\partial z} - \partial_z \psi, \end{array} \right\} \tag{4.1}$$

where (R, z) are cylindrical coordinates and Ω is the angular velocity of the flow, which we assume only depends on R and $\psi = -GM/(R^2 + z^2)^{1/2}$ is the cylindrical potential. A useful quantity will be the Keplerian frequency, which corresponds to the orbital frequency of a test particle on a circular orbit at radius R:

$$\Omega_K(R) = \sqrt{\frac{GM}{R^3}}. \tag{4.2}$$

In order to simplify the computation, let us assume that the disc is locally isothermal:[7] $T(R)$. Under these assumptions, the sound speed may be written

$$c_s \equiv \sqrt{\frac{P}{\rho}} = \sqrt{\frac{kT}{\mu}}, \tag{4.3}$$

where k is Boltzmann's constant and μ is the mean molecular mass. As the disc is locally isothermal, c_s only depends on R, as the temperature does.

We start with the vertical equilibrium, which we consider close to the disc midplane ($z \ll R$) because we assume the disc is thin:

$$c_s^2 \partial_z \log \rho = -\frac{GMz}{(R^2 + z^2)^{3/2}}$$
$$\simeq z\Omega_K^2 + O(z^3), \tag{4.4}$$

where we have assumed $z \ll R$. We deduce from this the vertical density profile

$$\rho = \rho_0(r) \exp\left(-\frac{z^2}{2H^2}\right), \tag{4.5}$$

where we have defined the disc scale height

$$H \equiv c_s/\Omega_K. \tag{4.6}$$

The thin disc approximation $H \ll R$ implies that the disc is cold, or in other words that $c_s \ll R\Omega_K$.

In the radial direction, we first have to compare the radial pressure gradient with the gravitational potential

$$0 = \underbrace{-\frac{1}{\rho}\frac{\partial P}{\partial R}}_{\sim c_s^2/R} - \underbrace{\partial_R \psi}_{\sim \Omega_K^2 R} + \Omega^2 R. \tag{4.7}$$

The pressure gradient is $(H/R)^2$ smaller than the gravitational potential and can be neglected in the thin disc approximation. This means that the disc is, to a very good approximation, a Keplerian disc $\Omega = \Omega_K$. Note, however, that local (i.e. on radial length scales of the order of H) pressure variations may exist leading to measurable deviation from the Keplerian rotation. These variations are typically responsible for zonal flows and local pressure maxima.

---

[7]This assumption is approximately valid because PPDs are passively irradiated. In turbulent discs, recent 3D relativistic magnetohydrodynamics (RMHD) simulations show a vertical temperature profile very close to isothermal, e.g. Flock *et al.* (2013).

4.3. *Accretion theory*

The energetics of MHD-driven discs has been discussed extensively by Balbus, Gammie & Hawley (1994), Balbus & Hawley (1998) and Balbus & Papaloizou (1999). Here, we revisit this question, and include the possibility of wind-driven accretion in the system. The accretion of mass in astrophysical discs is described by the equation of mass, angular momentum and mechanical energy conservation equations:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \rho \boldsymbol{u} = 0 \tag{4.8}$$

$$\frac{\partial R \rho u_\phi}{\partial t} + \nabla \cdot \left[R \rho u_\phi \boldsymbol{u} - R \frac{B_\phi \boldsymbol{B}}{4\pi} \right] = 0 \tag{4.9}$$

$$\frac{\partial \left(\frac{1}{2} \rho u^2 + \rho \psi + \frac{B^2}{8\pi} \right)}{\partial t} + \nabla \cdot \left[\left(\frac{1}{2} \rho u^2 + \rho \psi + P + \frac{B^2}{4\pi} \right) \boldsymbol{u} - \frac{\boldsymbol{u} \cdot \boldsymbol{B}}{4\pi} \boldsymbol{B} - \frac{\boldsymbol{\mathcal{E}}_{\mathrm{NI}} \times \boldsymbol{B}}{4\pi} \right]$$
$$= P \nabla \cdot \boldsymbol{u} + \frac{\boldsymbol{\mathcal{E}}_{\mathrm{NI}} \cdot \boldsymbol{J}}{c}, \tag{4.10}$$

where $\boldsymbol{\mathcal{E}}_{\mathrm{NI}}$ are electromotive forces owing to non-ideal effects. Note that molecular viscosity is usually negligible in these equations as it is several orders of magnitude smaller than non-ideal MHD effects. One notable exception is naturally when non-ideal MHD effects are absent, such as in ideal-MHD flows or in purely hydrodynamic flows subject to turbulence and/or spiral density waves. In these cases, viscosity becomes non-negligible in the energy equation because of the formation of small-scale structures, either through a direct turbulent cascade, or thanks to shocks. In any case, this viscosity then leads to an additional definite negative source term in the energy equation, which transforms mechanical energy into heat. The energy flux and angular momentum flux terms associated with viscosity are always negligible for practical applications.

In order to capture the dynamics of the disc, we separate the gravitational potential ψ as a midplane potential Ψ and a deviation as one moves away from the disc midplane Φ:

$$\psi = \Psi(R) + \Phi(R, z). \tag{4.11}$$

We also separate the mean rotational motion of the disc from its deviations (not necessarily small):

$$u_r = v_r; \quad u_\phi = \Omega R + v_\phi; \quad u_z = v_z, \tag{4.12a–c}$$

where we only assume that Ω satisfies the radial equilibrium in the disc midplane

$$\Omega^2 R = \partial_R \Psi. \tag{4.13}$$

Under these assumptions, it is possible to rewrite the angular momentum conservation as

$$\frac{\partial R \rho v_\phi}{\partial t} + \rho \boldsymbol{u} \cdot \nabla (\Omega R^2) + \nabla \cdot \left[R \rho v_\phi \boldsymbol{u} - R \frac{B_\phi \boldsymbol{B}}{4\pi} \right] = 0, \tag{4.14}$$

where we have used the continuity equation to eliminate the terms proportional to ΩR^2. A similar procedure can be followed for the energy equation, which can be written as

$$
\left(\frac{1}{2}\Omega^2 R^2 + \Psi\right)\left[\frac{\partial\rho}{\partial t} + \nabla\cdot\rho\boldsymbol{u}\right] + \rho\boldsymbol{u}\cdot\nabla\left(\frac{1}{2}\Omega^2 R^2 + \Psi\right)
$$

$$
+ \Omega\left(\frac{\partial R\rho v_\phi}{\partial t} + \nabla\cdot\left[R\rho v_\phi\boldsymbol{u} - R\frac{B_\phi\boldsymbol{B}}{4\pi}\right]\right) + \left[R\rho v_\phi\boldsymbol{u} - R\frac{B_\phi\boldsymbol{B}}{4\pi}\right]\cdot\nabla\Omega
$$

$$
+ \frac{\partial\left(\frac{1}{2}\rho v^2 + \rho\Phi + \frac{B^2}{8\pi}\right)}{\partial t} + \nabla\cdot\left[\left(\frac{1}{2}\rho v^2 + \rho\Phi + P + \frac{B^2}{4\pi}\right)\boldsymbol{v}\right.
$$

$$
\left. - \frac{\boldsymbol{v}\cdot\boldsymbol{B}}{4\pi}\boldsymbol{B} - \frac{\mathcal{E}_{\mathrm{NI}}\times\boldsymbol{B}}{4\pi}\right] = P\nabla\cdot\boldsymbol{u} + \frac{\mathcal{E}_{\mathrm{NI}}\cdot\boldsymbol{J}}{c}. \tag{4.15}
$$

We recognise the mass conservation equation in the first line, and the angular momentum conservation equation in the second line. Substituting (4.8) and (4.14) into the previous equation allows us to recast energy conservation as

$$
\rho\boldsymbol{u}\cdot[\nabla\Psi - \Omega^2 R\nabla R] + \left[R\rho v_\phi\boldsymbol{u} - R\frac{B_\phi\boldsymbol{B}}{4\pi}\right]\cdot\nabla\Omega
$$

$$
+ \frac{\partial\frac{1}{2}\rho v^2 + \rho\Phi + \frac{B^2}{8\pi}}{\partial t} + \nabla\cdot\left[\left(\frac{1}{2}\rho v^2 + \rho\Phi + P + \frac{B^2}{4\pi}\right)\boldsymbol{v}\right.
$$

$$
\left. - \frac{\boldsymbol{v}\cdot\boldsymbol{B}}{4\pi}\boldsymbol{B} - \frac{\mathcal{E}_{\mathrm{NI}}\times\boldsymbol{B}}{4\pi}\right] = P\nabla\cdot\boldsymbol{u} + \frac{\mathcal{E}_{\mathrm{NI}}\cdot\boldsymbol{J}}{c}, \tag{4.16}
$$

where we recognise the radial equilibrium in the first term, which can be cancelled out. Hence, we obtain an energy equation for the velocity fluctuations, which reads

$$
\frac{\partial\left(\frac{1}{2}\rho v^2 + \rho\Phi + \frac{B^2}{8\pi}\right)}{\partial t} + \nabla\cdot\left[\left(\frac{1}{2}\rho v^2 + \rho\Phi + P + \frac{B^2}{4\pi}\right)\boldsymbol{v} - \frac{\boldsymbol{v}\cdot\boldsymbol{B}}{4\pi}\boldsymbol{B} - \frac{\mathcal{E}_{\mathrm{NI}}\times\boldsymbol{B}}{4\pi}\right]
$$

$$
= P\nabla\cdot\boldsymbol{u} - \left[R\rho v_\phi\boldsymbol{u} - R\frac{B_\phi\boldsymbol{B}}{4\pi}\right]\cdot\nabla\Omega + \frac{\mathcal{E}_{\mathrm{NI}}\cdot\boldsymbol{J}}{c}. \tag{4.17}
$$

4.3.1. *Averaged equations*

In order to compute averaged conservation equations, we define an azimuthal average as

$$
\langle Q\rangle = \frac{1}{2\pi}\int \mathrm{d}\phi\, Q \tag{4.18}
$$

and a vertical integration of the azimuthal average

$$
\bar{Q} = \int_{z=-h}^{z=+h} \mathrm{d}z\langle Q\rangle, \tag{4.19}
$$

so that the continuity equation (4.8) reads

$$
\frac{\partial\Sigma}{\partial t} + \frac{1}{R}\frac{\partial}{\partial R}R\overline{\rho u_r} + \left[\langle\rho v_z\rangle\right]_{z=-h}^{+h} = 0, \tag{4.20}
$$

where $\Sigma \equiv \bar{\rho}$ is the gas surface density.

The equation of angular momentum conservation (4.14) can be recast using the same averaging procedure (4.19) defined above to obtain an equation relating the mass accretion rate $\overline{\rho v_r}$ as to the radial and surface stresses

$$\overline{\rho v_r}\frac{\partial}{\partial R}\Omega R^2 + \frac{1}{R}\frac{\partial}{\partial R}R^2 \underbrace{\left[\overline{\rho v_\phi v_r} - \frac{\overline{B_\phi B_r}}{4\pi}\right]}_{\text{Radial stress}} + \underbrace{\left[R\langle\rho v_\phi v_z\rangle - R\frac{\langle B_\phi B_z\rangle}{4\pi}\right]^{+h}_{z=-h}}_{\text{Surface stress}} = 0, \quad (4.21)$$

where we have assumed that $v \ll \Omega R$, which allows us to neglect the remaining time derivative.

This demonstrates the close relationship between the accretion rate and the transport of angular momentum by the stresses. Angular momentum can be transported outward in the disc by the radial stress, or evacuated from the disc by a torque applied at the disc surface, as for example when a magnetised wind is present.

This link between accretion and stress can also be seen by averaging of the mechanical energy equation (4.17):

$$\partial_t\overline{\mathcal{E}_m} + \frac{1}{R}\frac{\partial}{\partial R}R\overline{\mathcal{F}_{m,R}} + \left[\langle\mathcal{F}_{m,z}\rangle\right]^{+h}_{z=-h} = \overline{P\boldsymbol{\nabla}\cdot\boldsymbol{v}} - \underbrace{\left[\overline{\rho v_\phi v_R} - \frac{\overline{B_\phi B_R}}{4\pi}\right]\frac{\mathrm{d}\Omega}{\mathrm{d}\log R}}_{\text{Radial stress source term}} + \frac{\overline{\mathcal{E}_{\mathrm{NI}}\cdot\boldsymbol{J}}}{c},$$

$$(4.22)$$

where we have the mechanical energy of the fluctuations

$$\mathcal{E}_m = \frac{1}{2}\rho v^2 + \rho\Phi + \frac{B^2}{8\pi} \qquad (4.23)$$

and its associated energy flux

$$\left(\mathcal{E}_m + P + \frac{B^2}{8\pi}\right)\boldsymbol{v} - \frac{\boldsymbol{v}\cdot\boldsymbol{B}}{4\pi}\boldsymbol{B} - \frac{\mathcal{E}_{\mathrm{NI}}\times\boldsymbol{B}}{4\pi}. \qquad (4.24)$$

This energy equation demonstrates a very important fact: unless one assumes that the energy flux locally deposits energy (which implies that a source of energy is externally provided to the disc), then the only term that can balance diffusive (and viscous, when applicable) losses is the radial stress source term, which appears as a source term in the conservation of mechanical energy. Diffusive (and viscous) source terms being necessarily negative definite, we have

$$\left[\overline{\rho v_\phi v_R} - \frac{\overline{B_\phi B_R}}{4\pi}\right]\frac{\mathrm{d}\Omega}{\mathrm{d}\log R} < 0. \qquad (4.25)$$

As this term balances losses (which convert mechanical energy into heat), it is also equal to the local heating rate of the disc is we assume the fluctuations are statistically steady (as in a saturated turbulent state) and no energy escapes via the vertical energy flux.[8] Note that the surface stress does not appear as a source term, as it does not lead to any local heating, despite driving accretion. That's one of the key difference between radially-driven and vertically-driven accretion.

---

[8]When an outflow is present, one finds that the vertical energy flux $\mathcal{F}_{m,z}$ extracts energy from the disc, which implies that heating is actually smaller than the radial stress source term.

4.4. α *disc theory*

This theory assumes no wind is present at the disc surface. In order to solve the long-term evolution of the disc, one needs to express the radial stress

$$\overline{W_{r\phi}} = \overline{\rho v_\phi v_r} - \frac{\overline{B_\phi B_r}}{4\pi}, \tag{4.26}$$

as a function of vertically averaged quantities such as Σ or $\bar{P}$. Historically, and based on a purely dimensional argument (Shakura & Sunyaev 1973), it is usually assumed that

$$\overline{W_{r\phi}} = \alpha\bar{P}, \tag{4.27}$$

where α is a dimensionless constant. Physically, it can, however, be justified as a mixing length theory: let us consider turbulent velocity fluctuations v in a thin disc. The fluctuations are confined in the disc thickness H with a forcing frequency Ω_K (these two quantities are the only length and frequency accessible to an ideal system). Hence, we expect $v = \theta H \Omega_K$ where θ is a dimensionless constant, of order unity. Therefore, $\overline{W_{r\phi}} = \theta^2 \overline{\rho v^2} = \theta^2 \rho H^2 \Omega_K^2$. Using (4.6), one obtains $W_{r\phi} = \theta^2 \rho c_s^2 = \theta^2 \bar{P}$. Hence, thanks to the vertical equilibrium of a thin disc, the prescription of Shakura & Sunyaev (1973) shows up as a mixing length theory with a length H, a frequency Ω_K and $\alpha = \theta^2$.

Interestingly, because $H = c_s/\Omega_K$, θ is actually a measure of the Mach number of the flow $\theta = v/c_s$. If the turbulence was strongly supersonic, then strong shocks would appear, dissipating rapidly turbulent fluctuations until they become subsonic. For this reason, and in the absence of any supersonic excitation, turbulence is expected to be essentially subsonic with $\theta \lesssim 1$ and, therefore, $\alpha < 1$.

This prescription may be seen as a viscous theory. Indeed, the α-disc prescription leads to $\overline{W_{r\phi}} = \alpha\bar{P} = \alpha\Sigma c_s H\Omega_K$. As $R\,d\Omega_K/dR = -3/2\Omega_K$, the stress can be recast as

$$\overline{W_{r\phi}} = -\frac{2}{3}\nu_t\Sigma\frac{d\Omega}{d\log R}, \tag{4.28}$$

where we have defined an effective viscosity $\nu_t = \alpha c_s H$. Here, we clearly recognise the usual $R - \phi$ component of the viscous stress in the Navier–Stokes equations.

Plugging the α prescription into (4.21) and neglecting surface (wind) contribution leads to

$$\overline{\rho v_r} = -\frac{1}{R\partial_R(\Omega_K R^2)}\frac{\partial}{\partial R}R^2\alpha c_s^2\Sigma. \tag{4.29}$$

This allows us to express the mass accretion rate $\dot{M} \equiv -2\pi R\overline{\rho v_r}$ as

$$\dot{M} = \frac{4\pi}{R\Omega_K}\frac{\partial}{\partial R}R^2\alpha c_s^2\Sigma. \tag{4.30}$$

We can then use mass conservation (4.20) to obtain an equation for Σ

$$\frac{\partial\Sigma}{\partial t} = \frac{1}{R}\frac{\partial}{\partial R}\left[\frac{1}{\partial_R(\Omega_K R^2)}\frac{\partial}{\partial R}R^2\alpha c_s^2\Sigma\right], \tag{4.31}$$

which essentially constitutes a diffusion equation for the surface density. The diffusion timescale associated to accretion can be estimated using $c_s = \Omega_K H$. One finds

$$\tau_{\mathrm{visc}}^{-1} \sim \alpha\Omega_K\left(\frac{H}{R}\right)^2 \ll \Omega_K. \tag{4.32}$$

Accretion therefore occurs on timescales much longer than the orbital timescale in thin discs. This is usually a problem for simulations trying to capture the phenomenon of accretion. However, it allows us to separate accretion from dynamics occurring at the local orbital frequency, by stating that accretion is essentially non-existent on this timescale.

4.5. *α–υ disc theory*

This theory is identical to the alpha disc theory for the radial stress part, but it also includes a contribution from the surface term, owing to a hypothetical wind. To do so, let us define

$$W_{z\phi} = \rho v_\phi v_z - \frac{B_\phi B_z}{4\pi}, \tag{4.33}$$

and in a way similar to the α prescription, we assume

$$\left[\langle W_{z\phi} \rangle \right]_{z=-h}^{+h} = \upsilon P_{\mathrm{mid}}, \tag{4.34}$$

where P_{mid} is the midplane pressure of the disc. Using the same procedure as for the α disc, we can express the mass accretion rate as a function of α and υ

$$\dot{M} = \frac{4\pi}{R\Omega_K} \left[\underbrace{\frac{\partial}{\partial R} R^2 \alpha \overline{P}}_{\mathrm{radial}} + \underbrace{R^2 \upsilon P_{\mathrm{mid}}}_{\mathrm{vertical}} \right]. \tag{4.35}$$

The comparison between the α term and the υ term is revealing as it compares the role played by the radial and vertical stresses. One can assume that in first approximation $\overline{P} \simeq P_{\mathrm{mid}} H$ so that the vertical contribution is $R/H(\upsilon/\alpha)$ times larger than the radial one. This implies, in particular in thin discs where $R/H \gg 1$, that magnetised winds can easily be the dominant source of accretion.

In addition, using (4.35) in the continuity equation, the vertical stress term shows up as a first-order radial derivative of Σ (=advection) whereas the radial term appears as a second-order derivative as in the usual alpha disc theory. For this reason, wind-driven discs cannot be treated as viscous discs, because the wind component appears as an advective term in the surface density evolution.

4.6. *Beyond the α prescription*

The α disc model is useful as a starting point to characterise the evolution of discs. However, it is not based on first principles, and it would be desirable to compute the turbulent stress $W_{r\phi}$ directly from the equations of motion for the gas.

This is, however, a rather complicated task that often implies using numerical tools, as the equations of motion cannot, in general, be solved analytically. As the disc is thin, and turbulence, in the α disc theory, is supposed to be confined by the scale height $H \ll R$, one can start by using this scale separation to look only at what is happening at the scale H, leaving the global scale (R) apart. This is the idea of local models, often called 'shearing box' models, following Hawley, Gammie & Balbus (1995).

5. Local models

5.1. *The Hill's approximation*

The Hill's approximation is a local view of the dynamics of an orbiting system, which was initially used by Hill (1878) to model the libration motions of the moon along its orbit.
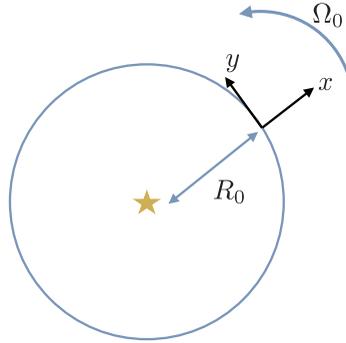
FIGURE 12. Rotating frame on a circular orbit at R_0.

It has been used more recently as an efficient tool to model the dynamics of gas or stars in gravitating systems (e.g. Goldreich & Lynden-Bell 1965) and it was later implemented numerically in the so-called 'shearing-box' by Hawley *et al.* (1995). In this model, one considers the dynamics of the flow around an equilibrium point R_0, which is rotating with the disc at the angular velocity $\Omega_O \equiv \Omega_K(R_0)$. We define a Cartesian frame (x, y, z), attached to this point so that x is aligned with the radius, y with the azimuth and z is aligned with the vertical direction (figure 12).

In this frame, the system follows the usual single-fluid equations of motion (MHD). As it is rotating, we have, in addition, a Coriolis force and a centrifugal force, so that the equations of motion read

$$\partial_t \rho + \nabla \cdot \rho \boldsymbol{v} = 0, \tag{5.1}$$

$$\partial_t \boldsymbol{v} + \boldsymbol{v} \cdot \nabla \boldsymbol{v} = -\frac{1}{\rho} \nabla P + \frac{\boldsymbol{J} \times \boldsymbol{B}}{\rho c} - 2\Omega_0 \boldsymbol{e}_z \times \boldsymbol{v} + \Omega_0 R^2 \boldsymbol{e}_R - \nabla \psi, \tag{5.2}$$

$$\partial_t P + \boldsymbol{v} \cdot \nabla P = -\gamma \nabla \cdot \boldsymbol{v}, \tag{5.3}$$

$$\partial_t \boldsymbol{B} = \nabla \times (\boldsymbol{v} \times \boldsymbol{B} + c\boldsymbol{E}_{\mathrm{NI}}), \tag{5.4}$$

where ψ is the gravitational potential, $\boldsymbol{E}_{\mathrm{NI}}$ is the non-ideal electromotive force and $R \equiv \sqrt{(R_0 + x)^2 + y^2}$ is the cylindrical radius. We also assume the gas follows an ideal equation of state with first adiabatic exponent γ. As is well known, the centrifugal force derives from a potential of the form $\psi_c = -\Omega_0^2 R^2 / 2$. The effective potential (gravitational plus centrifugal) in the corotating frame therefore reads

$$\psi_{\mathrm{eff}} = -\frac{GM}{\left((R_0 + x)^2 + y^2 + z^2\right)^{1/2}} - \frac{1}{2}\Omega_0^2 \left((R_0 + x)^2 + y^2\right). \tag{5.5}$$

Hill's model focuses on a 'small' (i.e. of the order of the disc scale height in the case of a gaseous disc) region around the fiducial point R_0. We therefore expand the effective potential around this point, assuming $x \sim y \sim z \lesssim H$ to obtain Hill's potential

$$\psi_{\mathrm{eff,Hill}} = \Omega_0^2 \left[-\frac{3R_0^2}{2} - \frac{3}{2}x^2 + \frac{1}{2}z^2 + O\left(\frac{H^3}{R_0}\right) \right]. \tag{5.6}$$

This effective potential has been truncated at the first non-trivial order. It is, however, interesting to note that it does not depend on y and that is does not contain any cross
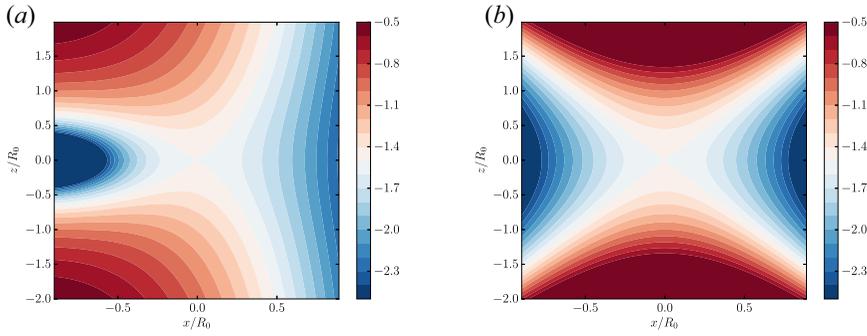
FIGURE 13. Effective potential in (*a*) the rotating frame ψ_{eff} and (*b*) its Hill's approximation $\psi_{\text{eff,Hill}}$. Note the $x \to -x$ symmetry of Hill's potential, as well as the different asymptotic behaviour.

term such as xy or xz. It is also independent from R_0 (apart from the constant term). This simplicity in the effective potential is what makes this model so useful for analytical and numerical computation. Any higher-order expansion will include curvature terms such as x/R_0 dependences and cross-dependences, making calculations much more tedious.

Let us emphasise already at this stage that Hill's potential is not adapted to global phenomenon. This can be seen by comparing the iso-potentials of ψ_{eff} and $\psi_{\text{eff,Hill}}$ (figure 13). Hill's approximation is found to be symmetrical in $x \to -x$, implying that one does not know where the centre of attraction is located (both $x \to -\infty$ and $x \to +\infty$ are technically valid). Moreover, the neutral iso-potential $\psi(x, z) = -3\Omega_0^2 R_0^2/2$ has an asymptote for $z \to +\infty$ at $x = R_0(\sqrt{3} - 1)$, which is absent in Hill's approximation. This asymptote is key for outflows to be ejected to $z \to \infty$, and is the main reason why local models always produce outflows which depend on the location of the z boundary conditions (see § 9.2).

Finally, it can be shown that in the more general case of a central gravity of the form $\psi_{\text{grav}} = r^{-2(q-1)}$, the equilibrium rotation profile is $\Omega \propto R^{-q}$ and Hill's potential reads $\psi_{\text{eff,Hill}} = \Omega_0^2 R_0^2 (qx^2 + z^2)/2 + \text{constant}$. In the case of a gravitational potential from a central point mass, we simply have $q = 3/2$.

The equations of motion in Hill's approximation finally reads

$$\partial_t \rho + \nabla \cdot \rho \boldsymbol{v} = 0, \tag{5.7}$$

$$\partial_t \boldsymbol{v} + \boldsymbol{v} \cdot \nabla \boldsymbol{v} = -\frac{1}{\rho} \nabla P + \frac{\boldsymbol{J} \times \boldsymbol{B}}{\rho c} - 2\Omega_0 \boldsymbol{e}_z \times \boldsymbol{v} + \Omega_0^2 (2qx\boldsymbol{e}_x - z\boldsymbol{e}_z) \tag{5.8}$$

$$\partial_t P + \boldsymbol{v} \cdot \nabla P = -\gamma \nabla \cdot \boldsymbol{v} \tag{5.9}$$

$$\partial_t \boldsymbol{B} = \nabla \times (\boldsymbol{v} \times \boldsymbol{B} + c\boldsymbol{E}_{\text{NI}}), \tag{5.10}$$

where we have differentiated Hill's potential to obtain the tidal acceleration $\Omega_0^2 (qx\boldsymbol{e}_x - z\boldsymbol{e}_z)$. This set of equation admits a simple steady solution $\boldsymbol{V}_0$ for the velocity field by simply balancing the Coriolis and tidal forces in the x direction:

$$\boldsymbol{V}_0 = -q\Omega_0 x\boldsymbol{e}_y. \tag{5.11}$$

This velocity field represents a constant radial shear. It corresponds to the local representation of the Keplerian flow, which is not in solid body rotation. As Hill's approximation only retains the first terms of the effective potential, the shear in Hill's

model is linear with x and does not depend on z. It is sometimes useful to work with the deviations from this velocity field. Let us define $w \equiv v - V_0$ (where w is not necessarily small compared with v), for which the equations of motion read

$$D_t\rho + \nabla \cdot \rho w = 0, \tag{5.12}$$

$$D_t w + w \cdot \nabla w = -\frac{1}{\rho}\nabla P + \frac{J \times B}{\rho c} - 2\Omega_0 e_z \times w + q\Omega_0 w_x e_y - \Omega_0^2 z e_z \tag{5.13}$$

$$D_t P + w \cdot \nabla P = -\gamma \nabla \cdot w \tag{5.14}$$

$$D_t B = -q\Omega_0 B_x e_y + \nabla \times (w \times B + cE_{NI}), \tag{5.15}$$

where we have defined the comoving derivative $D_t \equiv \partial_t - q\Omega x \partial_y$. It is worth noting that this last set of equations does not present any x dependency, except in the comoving derivative. This property is the key allowing the definition of a numerical 'shearing box', which we present later. In exchange, new source terms have appeared when going from v to w: the 'lift up' effect $q\Omega_0 w_x e_y$, which is essentially the advection of the mean Keplerian flow by radial motions, and the 'Ω effect' of dynamo theory (actually due to shear, not rotation) $-q\Omega_0 B_x e_y$. As we show later, these terms are key elements to the local physics of accretion discs. A local equivalent of the angular momentum equation can also be derived from (5.13). By defining $\mathcal{L} = w_y + (2 - q)\Omega_0 x$, we obtain the conservation equation

$$D_t\mathcal{L} + w \cdot \nabla\mathcal{L} = -\frac{1}{\rho}\partial_y P + \frac{[J \times B]_y}{\rho c}, \tag{5.16}$$

where the Coriolis and tidal terms have vanished in the definition of the local angular momentum.

5.2. *The shearing box model*

5.2.1. *Introduction*

The shearing box model formally comes in two forms, the 'large' shearing box (LSB, also known as the stratified shearing box) model which is essentially a numerical equivalent of Hill's approximation in a finite size numerical domain, and the 'small' shearing box (SSB, also known as the unstratified shearing box), which is a local approximation in Hill's approximation. In essence, the SSB model assumes that one zooms on a region close to the disc midplane so that the box size $L \ll H$. In this case, vertical gravity can be neglected in (5.13) and because one expects $w \sim \Omega L$, the flow is strongly subsonic so that an incompressible approximation[9] can be made in place of (5.12). The asymptotic of these two models are discussed in details by Umurhan & Regev (2004). Here, we mostly use the LSB, keeping in mind that an incompressible approximation is possible to study the basic effects of the shear.

5.2.2. *Boundary conditions*

The shearing box model makes use of Hill's approximation in the simplest possible numerical setup: a periodic box. Because Hill's approximation is local, the shearing box model is also local and should satisfy the asymptotic rules shown above. In particular, for a box of size L, we should have $L \sim H \ll R_0$. However, the presence of the comoving derivative makes things a bit more difficult. The fields ρ, w, etc. are advected everywhere at the azimuthal velocity $-q\Omega x e_y$. It is, therefore, physically inconsistent to assume

---

[9]Note, however, that in most of the unstratified shearing box models published to date, compressibility is retained in the equation of motion.
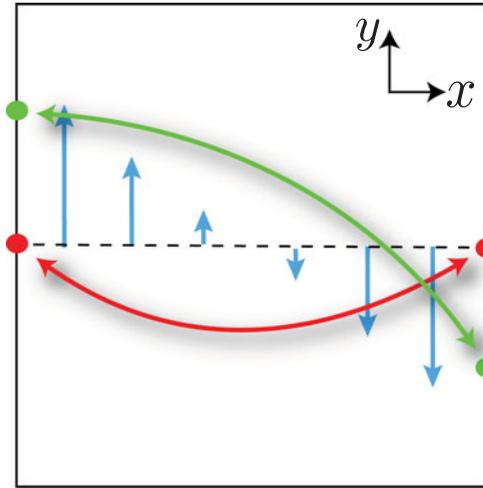
FIGURE 14. Radial boundary conditions in a shearing box. The background shear is represented in blue. At $t = 0$ the boundary conditions are strictly periodic (red). At $t > 0$, the periodic boundary conditions are shifted in time (green), according to the advection by the mean flow.

$Q(-L_x/2, y, z) = Q(L_x/2, y, z)$ for any quantity Q as one would do with x periodic boundary conditions in a box of size L_x. In order to take into account the constant shear in the boundary conditions, one therefore enforces periodic boundary condition in a Lagrangian view known as 'shearing sheet': $Q(-L_x/2, y, z) = Q(L_x/2, y + q\Omega L_x t, z)$, which can be represented graphically as in figure 14.

In the azimuthal direction, periodic boundary conditions are the most natural choice. In the vertical direction, however, no obvious choice comes to mind. Depending on the problem at hand, one can use outflow, free-slip or even periodic boundary conditions.

5.2.3. *Stress and accretion measurement*

One of the goals of the shearing box approach is to measure directly the turbulent stress in the global dynamical equations (4.21). One therefore needs to quantify

$$W_{xy} = \overline{\rho w_x w_y} - \frac{\overline{B_x B_y}}{4\pi} \tag{5.17}$$

which is Hill's equivalent of the radial stress term in (4.21). Despite a non-zero radial stress, the shearing box model does not exhibit any accretion owing to this component. This is because the shearing box model is a local asymptotic expansion. The terms leading to accretion are higher-order terms in H/R_0, which have been neglected, and which would break the in–out symmetry of Hill's potential.

5.2.4. *Conserved quantities*

The shearing box model is very useful in the sense that it allows one to have a tight control on the conserved quantities of the flow: vertical and radial magnetic flux, momentum and energy are all conserved thanks to the relatively simple boundary conditions used in the horizontal direction. We thus define

$$\langle \cdot \rangle \equiv \iiint \mathrm{d}x\, \mathrm{d}y\, \mathrm{d}z. \tag{5.18}$$

Mass, momentum and flux conservation equations eventually read

$$\partial_t\langle\rho\rangle + \left[\rho v_z\right]_{z=z_b} = 0 \tag{5.19}$$

$$\partial_t\langle\rho\boldsymbol{w}\rangle + \left[\rho\boldsymbol{w}w_z + \left(P + \frac{B^2}{8\pi}\right)\boldsymbol{e}_z - \frac{1}{4\pi}\boldsymbol{B}B_z\right]_{z=z_b} = -2\Omega_0\boldsymbol{e}_z \times \langle\rho\boldsymbol{w}\rangle$$
$$+ q\Omega_0\langle\rho w_x\rangle\boldsymbol{e}_y - \Omega_0^2\langle\rho z\rangle\boldsymbol{e}_z \tag{5.20}$$

$$\partial_t\langle\boldsymbol{B}\rangle + \left[w_z\boldsymbol{B} - B_z\boldsymbol{w}\right]_{z=z_b} = -q\Omega_0\langle B_x\rangle\boldsymbol{e}_y + \text{non-ideal terms.} \tag{5.21}$$

Interestingly, the momentum equation exhibits source terms connected to the effective potential and the Coriolis force. The vertical magnetic flux is found to be exactly conserved whereas the horizontal flux is not necessarily conserved. Horizontal flux can escape the box through the vertical boundary or, alternatively, can be amplified thanks to the Ω-effect that shows up as a source term of toroidal field. These simple conservation laws allows one to control very carefully the box physics. Moreover, because the vertical flux is conserved, one can classify shearing-box setup as a function of the average vertical flux. It is also possible to do this for the toroidal component in the y direction, though conservation can be violated by flux escape at the boundary.

Energetics of the shearing box may be found by dotting the momentum equation with $\boldsymbol{w}$ and the induction equation with $\boldsymbol{B}$. One eventually obtains an equation for the mechanical energy (kinetic plus magnetic) in the box

$$\partial_t\mathcal{E}_{\text{Mech}} + \nabla\cdot\mathcal{F}_{\text{Mech}} = P\nabla\cdot\boldsymbol{w} + \boldsymbol{E}_{\text{NI}}\cdot\boldsymbol{J} + q\Omega_0\left(\rho w_x w_y - \frac{B_x B_y}{4\pi}\right) - \Omega_0^2\rho w_z z, \tag{5.22}$$

where

$$\begin{rcases} \mathcal{E}_{\text{Mech}} = \frac{1}{2}\rho w^2 + \frac{1}{8\pi}B^2, \\ \mathcal{F}_{\text{Mech}} = \frac{1}{2}\rho w^2\boldsymbol{w} + \frac{1}{4\pi}\left(B^2\boldsymbol{w} - (\boldsymbol{w}\cdot\boldsymbol{B})\boldsymbol{B} - c\boldsymbol{E}_{\text{NI}}\times\boldsymbol{B}\right) + P\boldsymbol{w}. \end{rcases} \tag{5.23}$$

This set of equations is the local equivalent of (4.17). Several comments can be made on the energetics. First, we find an energy flux made of three contributions, kinetic energy, Poynting flux (split into advective, magnetic and non-ideal contributions) and a pressure term. Second, we find several source/sink terms.

(i) *PdV* work. This term is usually small when thermal effects are negligible, but can become important in thermally driven winds for example.

(ii) Non-ideal effects. For Ohmic and ambipolar diffusion, $\boldsymbol{E}_{\text{NI}} \sim -\boldsymbol{J}$, hence for these two terms, $\boldsymbol{E}_{\text{Ohm/ambipolar}}\cdot\boldsymbol{J} < 0$. Meanwhile, the Hall effect has no contribution to this term because $\boldsymbol{J}\times\boldsymbol{B}\cdot\boldsymbol{J} = 0$. Overall, the non-ideal term therefore appears as a sink of mechanical energy, as expected. Energy is dissipated into heat.

(iii) Radial stress. We recover the radial stress term of the global angular momentum equation (4.21), which appears as a source term here.

(iv) Potential work. The last term denotes the work done by the vertical gravity on the gas, which can become important in ejecting disc models. In this case, this term is negative.

(v) Viscous terms (not shown). When the flow is turbulent, viscous terms can become important at the small scale. In this case, they show up as an additional sink term in the mechanical energy equation.

Overall, unless strong thermal effects are present, the only source of mechanical energy in this system is the radial stress term, which is therefore positive definite, as already pointed out in the global version of this equation. The dynamics of the disc will therefore be dictated by how this source term is balanced by the various loss/flux terms in the energetics.

Let us finally point out that in a shearing box, the energy flux $\mathcal{F}_{\mathrm{Mech}}$ is periodic/shear periodic. If we average the energy equation, we find that the only relevant flux component is that escaping through the vertical boundaries, as for the mass and momentum fluxes.

6. The linear MRI in local models

6.1. *Lagrangian analysis*

6.1.1. *Linear hydrodynamic stability*

Let us now consider a particle evolving in Hill's effective potential under the influence of the effective gravity and the Coriolis force. The particle is initially at rest at $(x, y) = 0$. Magnetic fields are neglected in this first approach. The equation of motion for the fluid particle may be written

$$\frac{\mathrm{d}^2 x}{\mathrm{d}t^2} = 2q\Omega_0^2 x + 2\Omega_0 \frac{\mathrm{d}y}{\mathrm{d}t}, \tag{6.1}$$

$$\frac{\mathrm{d}^2 y}{\mathrm{d}t^2} = -2\Omega_0 \frac{\mathrm{d}x}{\mathrm{d}t}, \tag{6.2}$$

$$\frac{\mathrm{d}^2 z}{\mathrm{d}t^2} = -\Omega_0^2 z. \tag{6.3}$$

We first note that the vertical and horizontal equations of motion are separable. In the vertical direction, it describes oscillations of the fluid particle around the midplane at frequency Ω_0.

In the horizontal direction, the equations describes epicycles. To show it, let us first integrate (6.2):

$$\mathcal{L} = \frac{\mathrm{d}y}{\mathrm{d}t} + 2\Omega_0 x, \tag{6.4}$$

where $\mathcal{L}$ is the local angular momentum of the particle, as defined in (5.16). Our particle being initially in equilibrium at $(x, y) = 0$, it has $\mathcal{L} = 0$ and we can write the radial equation of motion as

$$\frac{\mathrm{d}^2 x}{\mathrm{d}t^2} = -2\Omega_0^2 (2 - q)x \tag{6.5}$$

hence, our effective gravitational potential, which was initially unstable ($\partial_x \psi_{\mathrm{eff, Hill}} < 0$), is stabilised thanks to the conservation of angular momentum, provided that $q < 2$ ($q = 3/2$ for astrophysical discs). The oscillations described by this particle have a frequency

$$\omega^2 = 2\Omega_0^2 (2 - q) \equiv \kappa^2. \tag{6.6}$$

This characteristic frequency is named epicyclic frequency. In the particular case of a Keplerian disc ($q = 3/2$), we find $\kappa^2 = \Omega^2$, i.e. the epicyclic frequency coincides with the
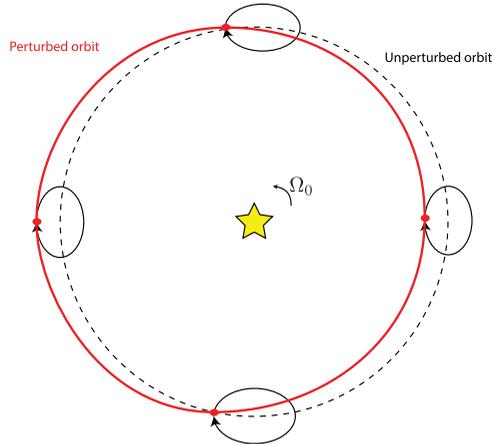
FIGURE 15. Epicyclic oscillations of a fluid particle orbiting a point mass resulting in an closed elliptic orbit.

orbital frequency. As a result, orbits are closed, a well-known property of the two-body problem (e.g. figure 15).

This shows that at the linear level, pure Keplerian flows are stable. This does not mean that non-linear (or subcritical) instabilities cannot exist in these flows, given that the shear is a natural reservoir of free energy to trigger instabilities. Such non-linear instabilities are well known to develop in non-rotating sheared flows and in pipe flows. This question of subcritical instabilities is at the origin of many experiments, numerical simulations and theoretical developments. Today, there seems to be a consensus on the fact that pure Keplerian flows appear to be stable for Reynolds numbers up to a few million. However, thermal effects, such as heating and vertical stratification, are known to affect this picture, leading to thermally driven linear and non-linear instabilities. This is a whole field of research, which is not covered here. The interested reader may consult Fromang & Lesur (2019) for a more detailed overview of this topic.

6.1.2. *Linear MHD stability*

As we have shown, PPDs are hydrodynamically stable at the linear level. In MHD, however, things start to become a bit more interesting (see also Balbus & Hawley 1998 for a similar treatment). Let us embed our disc in an external and constant magnetic field B_0, which we assume is vertical. Assuming we still consider infinitesimal displacements around the equilibrium position of the fluid particles, the velocities are infinitely small, and the induction equation for magnetic fluctuations δb reads

$$\frac{\partial \delta b}{\partial t} = B_0 \frac{\partial v}{\partial z}. \tag{6.7}$$

Clearly, the stability will now depend on how we move the particles with respect to each other. Let us consider a set of particles initially at $(x, y) = 0$ and let us perturb these particles with a vertical harmonic perturbation

$$x = x_0 \exp(\mathrm{i}kz), \tag{6.8}$$

the resulting magnetic perturbation can be obtained by integrating the induction equation with respect to time:

$$\delta b = \mathrm{i}kB_0 x. \tag{6.9}$$

In order to model how the field affects the dynamics, we have to include the Lorentz force $\boldsymbol{F}_L$ in the equation of motion. In the horizontal direction, only the magnetic tension term $\boldsymbol{B} \cdot \nabla \boldsymbol{B}$ appears, so we have

$$
\begin{aligned}
\frac{\boldsymbol{F}_L}{\rho} &= \frac{\boldsymbol{B}_0 \cdot \nabla \delta \boldsymbol{b}}{4\pi \rho} \\
&= -\frac{k^2 B_0^2}{4\pi \rho} \boldsymbol{x} \\
&= -V_A^2 k^2 \boldsymbol{x},
\end{aligned} \tag{6.10}
$$

where V_A is the Alfvén speed. The horizontal equations of motion are therefore reduced to

$$
\left.
\begin{aligned}
\frac{\mathrm{d}^2 x}{\mathrm{d}t^2} &= 2q\Omega_0^2 x + 2\Omega_0 \frac{\mathrm{d}y}{\mathrm{d}t} - V_A^2 k^2 x, \\
\frac{\mathrm{d}^2 y}{\mathrm{d}t^2} &= -2\Omega_0 \frac{\mathrm{d}x}{\mathrm{d}t} - V_A^2 k^2 y,
\end{aligned}
\right\} \tag{6.11}
$$

where it is clear that the magnetic forces are acting as a restoring force (hence, the usual representation of a spring for the Lorentz force). Note also that angular momentum conservation is now broken by the azimuthal tension force. It is this effect that leads to an instability.

To show this, let us assume $\boldsymbol{x} = \boldsymbol{x} \exp(\sigma t)$. The equations of motion lead to the following eigenvalue problem

$$
\left.
\begin{aligned}
(\sigma^2 + V_A^2 k^2)x &= 2q\Omega_0^2 x + 2\Omega_0 \sigma y, \\
(\sigma^2 + V_A^2 k^2)y &= -2\Omega_0 \sigma x,
\end{aligned}
\right\} \tag{6.12}
$$

which allows us to obtain the dispersion relation

$$
(\sigma^2 + V_A^2 k^2)^2 - 2q\Omega_0^2(\sigma^2 + V_A^2 k^2) + 4\Omega_0^2 \sigma^2 = 0, \tag{6.13}
$$

where we recover epicyclic oscillations when $V_A = 0$ with $\sigma^2 = -2\Omega_0^2(2-q) = -\kappa^2$ and pure Alfvénic oscillations when $\Omega_0 = 0$ with $\sigma^2 = -V_A^2 k^2$. Expanding this dispersion relation leads to

$$
\sigma^4 + \sigma^2 \left(\kappa^2 + 2V_A^2 k^2\right) + V_A^2 k^2 \left(V_A^2 k^2 - 2q\Omega_0^2\right) = 0. \tag{6.14}
$$

This dispersion relation describes a linear instability when σ^2 is positive, i.e. when

$$
V_A^2 k^2 - 2q\Omega_0^2 < 0. \tag{6.15}
$$

This instability is the MRI. It appears when the magnetic tension force is not too strong, as suggested by (6.15). It is possible to solve the full dispersion analytically to obtain the eigenvalues (see figure 16). When $V_A k < \sqrt{2q}\Omega_0$, positive eigenvalues are found, which are the signature of the MRI. The maximum growth rates are obtained for $V_A k = \sqrt{2q}\Omega_0/2$ with $\sigma_{\max} = q\Omega/2$. Above the limit (6.15), the unstable branch becomes a stable Alfvén wave, which shows that the MRI is mostly an Alfvénic perturbation. In addition to this pair of branches, we find a pair of epicyclic modes which are stable for all kV_A (figure 16), and have a non-zero frequency for $V_A = 0$. If compressibility is added, the degeneracy between slow magnetosonic and Alfvén waves is lifted. In this case, it can be shown that the MRI arises from the slow magnetosonic mode (Balbus & Hawley 1992).
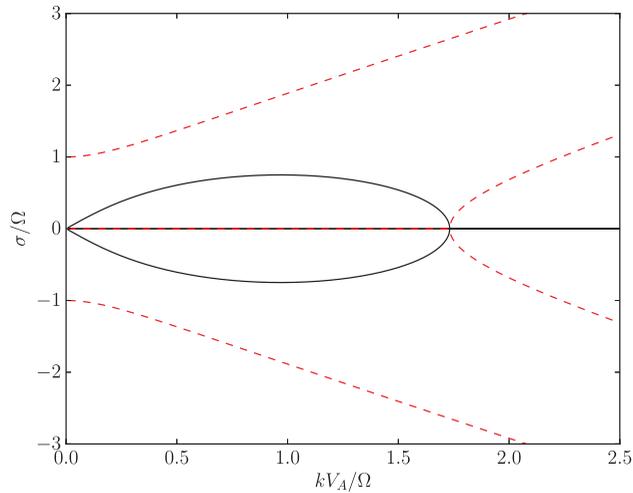
FIGURE 16. Real part (black) and imaginary part (red dashed line) of the solutions of (6.14) with $q = 3/2$. The MRI appears for weak enough fields $V_A k < \sqrt{3}\Omega_0$.
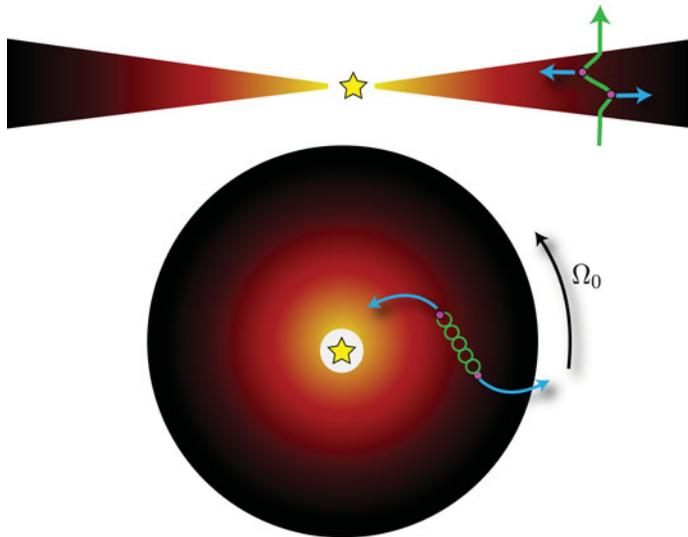
FIGURE 17. Physical representation of the MRI mechanism (see the text).

The physical interpretation of the MRI is straightforward: consider two fluid particles attached to a vertical field line and assume we slightly move these particles radially. First, they will start an epicyclic motion and drift azimuthally (figure 17). As they drift away, the azimuthal magnetic tension will act as a spring bringing back the particles together, slowing down the inner particle and accelerating the outer particle. This results in a loss of angular momentum for the inner particle, which falls further down, and conversely for the outer particle. This mechanism can only work if the radial magnetic tension is sufficiently weak; otherwise, the particles return to their initial point resulting in an Alfvénic oscillation. It is this radial component of the Lorentz force that is the stabilising agent of the MRI.
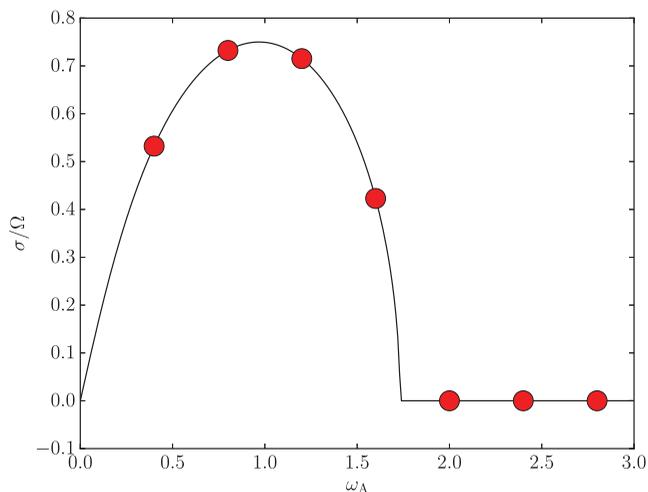
FIGURE 18. MRI growth rate as a function of the Alfvén frequency ω_A (black). Quantised modes accessible to a disc model with $V_{Az} = 0.2\,\Omega H$ are shown in red with increasing n from left to right. Only four modes are MRI-unstable in this example (see the text).

6.2. *Application of the MRI to local models*

In real disc models, the disc is characterised by a mean vertical and potentially a mean azimuthal magnetic field, as in the linear analysis. However, the vertical wavelength of the perturbation cannot be larger than the disc scale height.[10] In other words, $k_{z,\min} \simeq 1/H$, which implies that any disc model threaded by a vertical field has a minimal Alfvén frequency $\omega_A \equiv k_z V_{Az}$. In addition, the vertical wavenumbers accessible to a disc are quantised because of the limited vertical extension (see § 6.4.6), hence $k_z = n k_{z,\min}$ and ω_A is also quantised. As an example, we show in figure 18 the modes accessible to a disc threaded by a vertical field with $V_{Az} = 0.2\,\Omega H$. In this particular example, only the four smallest n are MRI-unstable, the most unstable mode being $n = 2$.

This quantisation of MRI modes also indicates that the MRI can be stabilised for sufficiently strong fields. Indeed, as the field strength increases, quantised unstable modes drift to the right of figure 18 because $k_{z,\min}$ increases. The MRI is entirely stabilised when the lowest n enters the stable regime, i.e. when $V_{Az} > \sqrt{2q}\Omega/k_{z,\min}$. With $k_{z,\min} \simeq 1/H$ this implies in Keplerian discs

$$V_{Az} \gtrsim \sqrt{3}\Omega H \to \text{Stability}. \tag{6.16}$$

For this reason, the MRI is often seen as a 'weak field' instability, even though technically, this is more a result of the geometrically thin disc approximation. Note also that this criterion is not an effect of compressibility, as is sometimes thought. Indeed, because $\Omega H = c_s$, the previous criterion implies that the disc stabilises for vertical field strength above equipartition. This is, however, just a coincidence resulting from the vertical equilibrium in a geometrically thin disc. The MRI also exists well above equipartition in a slightly modified form, provided that large enough wavelengths are allowed (Kim & Ostriker 2000).

---

[10]We return more quantitatively to this point in § 6.4.6.

6.3. *Non-axisymmetric MRI*

We start from (5.12)–(5.15) in which we assume the disc is threaded by a mean field having a vertical and azimuthal components: $B = B_{0,y}e_y + B_{0,z}e_z$. In this subsection, we relax the axisymmetry hypothesis used previously, but we still assume the flow is incompressible, and neglect vertical gravity and stratification. We then consider the following set of equations:

$$\partial_t w + w \cdot \nabla w - q\Omega x \partial_y w = -\nabla \Pi + \frac{B \cdot \nabla B}{4\pi\rho} + 2\Omega w_y e_x - (2-q)\Omega w_x e_y, \quad (6.17)$$

$$\partial_t B + w \cdot \nabla B - q\Omega x \partial_y B = B \cdot \nabla w - q\Omega B_x e_y, \quad (6.18)$$

$$\nabla \cdot w = 0. \quad (6.19)$$

We are going to linearize this system, assuming the field can be decomposed as $B = B_0 + b$. The presence of the advection term $q\Omega x \partial_y$, however, leads to some technical difficulties as it involves an explicit spatial dependency whenever modes are non-axisymmetric. We therefore follow Kelvin (1880) and Craik & Criminale (1986) using decomposition into time-dependent 'waves',[11] which, for any quantity $\tilde{X}$, assumes as spatial decomposition

$$\tilde{X} = X(t) \exp(ik(t) \cdot x). \quad (6.20)$$

Using this decomposition, it is easy to show that

$$\partial_t \tilde{X} - q\Omega x \partial_y \tilde{X} = \dot{X} + iX \left(\dot{k} \cdot x - q\Omega x k_y \right). \quad (6.21)$$

As it is the only term that shows this explicit x dependency in the equations of motion, and because these equations are assumed to be valid for all x, we are forced to conclude that the term in parentheses cancels out:

$$\dot{k}_x - q\Omega k_y = 0; \quad \dot{k}_y = 0; \quad \dot{k}_z = 0. \quad (6.22a\text{--}c)$$

Without loss of generality, we can therefore assume a decomposition into 'shearing waves', defined as

$$k(t) = k_0 + q\Omega t k_y e_x \quad (6.23)$$

which is solution to the set of equations described previously. Using this shearing wave decomposition, we obtain

$$\dot{w} = -ik\Pi + i\frac{k \cdot B_0}{4\pi\rho}b + 2\Omega w_y e_x - (2-q)\Omega w_x e_y, \quad (6.24)$$

$$\dot{b} = i(k \cdot B_0)w - q\Omega b_x e_y, \quad (6.25)$$

$$k \cdot w = 0, \quad (6.26)$$

where we have dropped the explicit time dependency of k and the ˜ symbols for simplicity. We next take the time derivative of (6.26):

$$q\Omega k_y w_x + k \cdot \dot{w} = 0. \quad (6.27)$$

---

[11]Although these solutions are often called 'waves', these are not normal modes of the physical system, but they are just a convenient way to decompose solutions in a sheared flow.

This allows us to express the generalised pressure

$$\Pi = -\frac{i}{k^2} \left(2\Omega w_y k_x - 2\Omega(1-q)w_x k_y \right).$$ (6.28)

Thus, finally, the equations of motion read

$$\dot{w} = i\frac{k \cdot B_0}{4\pi\rho}b + 2\Omega w_y(1-g_{xx})e_x + 2(1-q)\Omega w_x g_{xy}e_x$$
$$- q\Omega w_x g_{yy}e_y - (2-q)\Omega w_x(1-g_{yy})e_y - 2\Omega w_y g_{xy}e_y$$
$$+ 2(1-q)\Omega w_x g_{yz}e_z - 2\Omega w_y g_{xz}e_z,$$ (6.29)

where we have introduced $g_{ij} = k_i k_j / k^2$.

It is not possible to go any further without making any approximation. Indeed, although by construction $k \cdot B_0$ does not have any time dependency, the pressure factors g_{ij} do have one, so that a standard normal mode decomposition is prone to failure. It is possible to numerically integrate these equations as a function of time (see, e.g., Balbus & Hawley 1992). This always leads to transiently growing solutions, i.e. perturbations that only grow for a finite time. To understand why this is always the case, let us continue our analysis using a first-order Wentzel–Kramers–Brillouin (WKB) approximation.

To obtain a dispersion relation, one needs to assume that k is 'almost' steady, i.e. $k_y \ll k_x$ so that $d \log k/dt \ll \Omega$. This limit is often described as a strongly leading or strongly trailing wave. This limit implies that we can neglect all the g_{yj} terms in the previous expansion. Assuming $X = \tilde{X} \exp[\sigma t + ik(t) \cdot x)]$, we then obtain

$$\sigma w = i\frac{k \cdot B_0}{4\pi\rho}b + 2\Omega v_y(1-g_{xx})e_x - (2-q)\Omega v_x e_y - 2\Omega v_y g_{xz}e_z,$$ (6.30)

$$\sigma b = i(k \cdot B_0)v - q\Omega b_x e_y.$$ (6.31)

These equations clearly exhibit an Alfvén mode in the z direction, with $\sigma = \pm i(k \cdot V_{A,0})$ and $V_{A,0} \equiv B_0/\sqrt{4\pi\rho}$ as one would expect. We can then solve independently the horizontal problem to obtain the dispersion relation:

$$\sigma^4 + \sigma^2(2\omega_A^2 + 2(2-q)\Omega^2 g_{zz}) + \omega_A^2(\omega_A^2 - 2q\Omega^2 g_{zz})$$ (6.32)

where we have defined the Alfvén frequency $\omega_A = k \cdot V_{A,0}$. This equation describes both the traditional MRI mode and the non-axisymmetric MRI, as its close resemblance with (6.14) suggests. It should, however, be noted that in the $k_z/k \to 0$ limit, the instability is lost because $g_{zz} \to 0$ and the last term of the dispersion relation, responsible for the MRI, vanishes. Physically, this happens because the pressure gradient balances the Coriolis force in the x component of the equation of motion.

Now, it should be kept in mind that this dispersion relation is derived for a shearing wave whose k is slowly evolving with time. As $t \to \infty$, one then expects $|k_x| \to \infty$ and therefore $g_{zz} \to 0$. Hence, by nature, shearing waves automatically quench the growth of the MRI as they evolve. For this reason, non-axisymmetric structures, even without any dissipation, are necessarily transiently growing solution, and therefore never lead to a genuine linear instability. That being said, non-axisymmetry is required when only a large-scale toroidal field is available in the system, because one needs $k \cdot B_0 \neq 0$. Therefore, there is no 'toroidal field MRI' as is sometimes presented in the literature. There is just a transient growth in the linear phase, which can be described as a temporary MRI

in the WKB approximation, but only a non-linear feedback can re-excite new shearing waves to keep increasing the energy of the fluctuations.

It should be stressed that the absence of any linear non-axisymmetric instability is observed only in the local limit (i.e. shearing box). If one considers a global disc, including curvature and radial boundaries, then a genuine linear instability can be recovered (e.g. Curry & Pudritz 1996), with properties similar to that found in the WKB approximation presented previously.

6.4. *MRI in non-ideal MHD*

6.4.1. *Historical background*

The linear MRI in the non-ideal MHD regime has been explored by many researchers. After the discovery of the MRI in the disc context by Balbus & Hawley (1991), it was soon realised that PPDs, but also discs in cataclysmic variables could be in the non-ideal MHD regime, casting doubts on the applicability of this instability to these objects. Blaes & Balbus (1994) were the first to consider this problem, by working out the ambipolar-dominated MRI in the two-fluid approximation. Ohmic diffusion was first considered by Jin (1996) in unstratified models and Sano & Miyama (1999) in stratified discs, which led to the dead zone model of PPDs (Gammie 1996). Later, Wardle (1999) considered the three non-ideal MHD effects in simple axial geometry whereas Desch 2004 considered also oblique modes and toroidal fields. Balbus & Terquem (2001) isolated the physics of the Hall-MRI and Kunz (2008) demonstrated that one of the Hall-MRI branches was actually a new instability: the HSI. Finally, ambipolar diffusion was revisited in the single-fluid approximation by Kunz & Balbus (2004), demonstrating the existence of oblique ambipolar modes and their origin.

In the following, we revisit and discuss each of these effects with a unified notation and geometry. Note, however, that our approach and dispersion relation is formally identical to that of Desch (2004).

6.4.2. *Linearised equations*

As in the cases described previously, we start from (5.12)–(5.15) in which we assume the disc is threaded by a mean field having a vertical and azimuthal components: $\boldsymbol{B} = B_{0,y}\boldsymbol{e}_y + B_{0,z}\boldsymbol{e}_z$. In the following, we consider small axisymmetric[12] perturbations of the equilibrium. We seek solutions of the form $\boldsymbol{w} = \boldsymbol{u}\exp[\boldsymbol{k}\cdot\boldsymbol{x} + \sigma t]$ and $\boldsymbol{B} = \boldsymbol{B}_0 + \boldsymbol{b}\exp[\boldsymbol{k}\cdot\boldsymbol{x} + \sigma t]$, where $\boldsymbol{k} = k_x\boldsymbol{e}_x + k_z\boldsymbol{e}_z$ is the wavenumber and σ is the linear growth rate of the instability. We moreover neglect vertical stratification and vertical gravity and assume the flow is incompressible, which implies that we consider the SSB approximation. As we show later, this is enough to capture most of the physics relevant to the problem. We explore in § 6.4.6 the effect of the vertical stratification on linear modes.

Under these assumptions, the linearised equations read

$$\sigma\boldsymbol{u} = -\mathrm{i}k\Pi + \mathrm{i}\frac{\boldsymbol{k}\cdot\boldsymbol{B}_0}{4\pi\rho}\boldsymbol{b} + 2\Omega u_y\boldsymbol{e}_x - (2-q)\Omega u_x\boldsymbol{e}_y, \qquad (6.33)$$

$$\sigma\boldsymbol{b} = \mathrm{i}(\boldsymbol{k}\cdot\boldsymbol{B}_0)\boldsymbol{u} - q\Omega b_x\boldsymbol{e}_y - \eta_O k^2\boldsymbol{b} + \eta_H\left(\boldsymbol{k}\cdot\hat{\boldsymbol{B}}_0\right)\boldsymbol{k}\times\boldsymbol{b},$$
$$- \eta_A\left[\left(\boldsymbol{k}\cdot\hat{\boldsymbol{B}}_0\right)^2\boldsymbol{b} - \left(\boldsymbol{b}\cdot\hat{\boldsymbol{B}}_0\right)\left(\left[\boldsymbol{k}\cdot\hat{\boldsymbol{B}}_0\right]\boldsymbol{k} - k^2\hat{\boldsymbol{B}}_0\right)\right], \qquad (6.34)$$

---

[12]In Hill's approximation, non-axisymmetric linear perturbations only lead to transiently growing solutions, which ultimately decay (Balbus & Hawley 1992). These transient 'modes' are, however, important for the MRI dynamo once non-linear feedback is taken into account (Lesur & Ogilvie 2008*a*).

$$k \cdot u = 0, \tag{6.35}$$

$$k \cdot b = 0. \tag{6.36}$$

The expression of the non-ideal terms can be interpreted in the following way. First, Ohmic diffusion acts as a pure linear damping operator, as expected. The Hall term is proportional to $k \times b$, which means it rotates the magnetic perturbation around the k direction, keeping its norm constant. Note that the direction of rotation is given by the sign of η_H, which shows that the handedness given by the Hall effect is connected directly to the microphysics of the plasma (see § 3.4.1). Finally, ambipolar diffusion involves an anisotropic diffusion term that we discuss in the following.

The solenoidal conditions can be used to eliminate u_z and b_z from the equations in favour of the horizontal components, leading to a fourth-order problem

$$\sigma u_x = \mathrm{i} \frac{k \cdot B_0}{4\pi\rho} b_x + 2\Omega \frac{k_z^2}{k^2} u_y, \tag{6.37}$$

$$\sigma u_y = \mathrm{i} \frac{k \cdot B_0}{4\pi\rho} b_y - (2-q)\Omega u_x, \tag{6.38}$$

$$\sigma b_x = \mathrm{i}(k \cdot B_0)u_x - \eta_O k^2 b_x - \eta_H \left(k \cdot \hat{B}_0\right) k_z b_y - \frac{\eta_A}{B_0^2}\left(k^2 B_{0,z}^2 b_x - k_x B_{0,y}\left(k \cdot B_0\right) b_y\right), \tag{6.39}$$

$$\sigma b_y = \mathrm{i}(k \cdot B_0)u_y - q\Omega b_x - \eta_O k^2 b_y + \eta_H \left(k \cdot \hat{B}_0\right) \frac{k^2}{k_z} b_x$$
$$- \frac{\eta_A}{B_0^2}\left(\left[(k \cdot B_0)^2 + k^2 B_{0,y}^2\right] b_y - \frac{k^2}{k_z^2} k_x B_{0,y}(k \cdot B_0) b_x\right). \tag{6.40}$$

The role played by ambipolar diffusion here is a bit more self-explanatory. We observe that the diagonal terms in the second pair of equations are always negative definite, hence ambipolar diffusion is really acting as a diffusion term on the diagonal components with an amplitude controlled by the magnitude, but also the orientation of B_0. However, there are also off-diagonal terms proportional to $k_x B_{0,y}$. As we show, these terms can lead to oblique unstable modes (see also Kunz & Balbus 2004).

We next solve the set of equations for σ described previously and look for unstable eigenvalues. We follow Pandey & Wardle (2012) and first obtain an equation for the velocity fluctuations

$$u = \frac{\mathrm{i} k \cdot B_0}{4\pi\rho(\sigma^2 + \kappa^2)} \begin{pmatrix} \sigma & 2\Omega \\ -(2-q)\Omega & \sigma \end{pmatrix} b, \tag{6.41}$$

so that the induction equation becomes, in matrix form,

$$\left[\begin{pmatrix} \sigma + \eta_O k^2 + \tau_A k^2 V_{Az}^2 & \ell_H \omega_A k_z - \tau_A k_x V_{Ay}\omega_A \\ q\Omega - \dfrac{k^2}{k_z^2}\left(\ell_H \omega_A k_z + \tau_A k_x V_{Ay}\omega_A\right) & \sigma + \eta_O k^2 + \tau_A(\omega_A^2 + k^2 V_{Ay}^2) \end{pmatrix}\right.$$
$$\left.+ \frac{\omega_A^2}{\sigma^2 + \dfrac{k_z^2}{k^2}\kappa^2} \begin{pmatrix} \sigma & 2\Omega \dfrac{k_z^2}{k^2} \\ -(2-q)\Omega & \sigma \end{pmatrix}\right] b = 0 \tag{6.42}$$

where we have introduced the Alvén speed $V_A = B_0/(4\pi\rho)^{1/2}$, the Alfvén frequency $\omega_A \equiv \mathbf{k} \cdot \mathbf{V}_A$, the Hall length $\ell_H \equiv \eta_H/V_A$, the ambipolar time $\tau_A \equiv \eta_A/V_A^2$ and the epicyclic frequency $\kappa^2 \equiv 2\Omega^2(2-q)$.

After a long but straightforward calculation, one eventually obtains the dispersion relation which can be written

$$\sigma^4 + \mathcal{C}_3\sigma^3 + \mathcal{C}_2\sigma^2 + \mathcal{C}_1\sigma + \mathcal{C}_0 = 0 \tag{6.43}$$

with

$$\mathcal{C}_3 = 2\eta_O k^2 + \tau_A(k^2 V_A^2 + \omega_A^2), \tag{6.44}$$

$$\mathcal{C}_2 = \frac{k_z^2}{k^2}\kappa^2 + 2\omega_A^2 + \eta_O^2 k^4 + \tau_A^2\omega_A^2 k^2 V_A^2 + q\Omega\tau_A\omega_A k_x V_{Ay} + \ell_H k_z \omega_A \left(\frac{k^2}{k_z^2}\ell_H k_z \omega_A - q\Omega\right), \tag{6.45}$$

$$\mathcal{C}_1 = \mathcal{C}_1\left(\omega_A^2 + \frac{k_z^2}{k^2}\kappa^2\right), \tag{6.46}$$

$$\mathcal{C}_0 = \omega_A^2\left(\underbrace{\omega_A^2 - 2q\Omega^2\frac{k_z^2}{k^2}}_{\text{MRI}}\right) + \kappa^2 k_z^2 k^2 \eta_O^2 + \ell_H\omega_A k_z\left(\underbrace{(4-q)\Omega\omega_A^2}_{\text{ion-cyclotron instability}} - \underbrace{q\Omega\frac{k_z^2}{k^2}\kappa^2 + \ell_H\omega_A k_z\kappa^2}_{\text{HSI}}\right),$$

$$+ \kappa^2\tau_A^2\omega_A^2 k_z^2 V_A^2 + \underbrace{q\Omega\tau_A k_x V_{Ay}\omega_A\left(\frac{k_z^2}{k^2}\kappa^2 + \omega_A^2\right)}_{\text{Oblique ambipolar modes}}. \tag{6.47}$$

The stability of this linear system can be analysed in the vicinity of $\sigma = 0$. In this case, a necessary and sufficient condition for instability is $\mathcal{C}_0 < 0$. This allows us to identify three sources of instability: the usual MRI, the ion-cyclotron instability, the Hall-shear instability (HSI) and the term at the origin of Oblique ambipolar modes, which is not a genuine instability branch.

Before exploring the non-ideal regime, let us point out that in the ideal MHD limit, this dispersion relation shows that modes with $k_x \neq 0$ have a lower growth rate than $k_x = 0$ modes. As a result, $k_x = 0$ modes are always the most unstable eigenmodes of the system. These modes are often called 'channel modes' as they do not have any horizontal spatial dependency. They are also exact non-linear solutions of the full MHD equations (Goodman & Xu 1994). For this reason, they are very robust and they often show up in the non-linear regime, as we show later.

6.4.3. *Ohmic diffusion*

The effect of Ohmic diffusion on the stability of the MRI is physically very intuitive: it stabilises MRI modes, starting from the largest ω_A of the system (see figure 19). The stability condition in the presence of Ohmic diffusion is deduced from the condition $\mathcal{C}_0 = 0$ and reads

$$\frac{2q\Omega^2}{k^2 V_{Az}^2} - 1 < \kappa^2\frac{k^2}{k_z^2}\frac{\eta_O^2}{V_{Az}^4} \rightarrow \text{Stability}. \tag{6.48}$$

Clearly, the modes with $k_x = 0$ are the last to be stabilised when one increases Ohmic diffusion. Let us focus on this case and assume the flow is Keplerian so that $\kappa = \Omega$
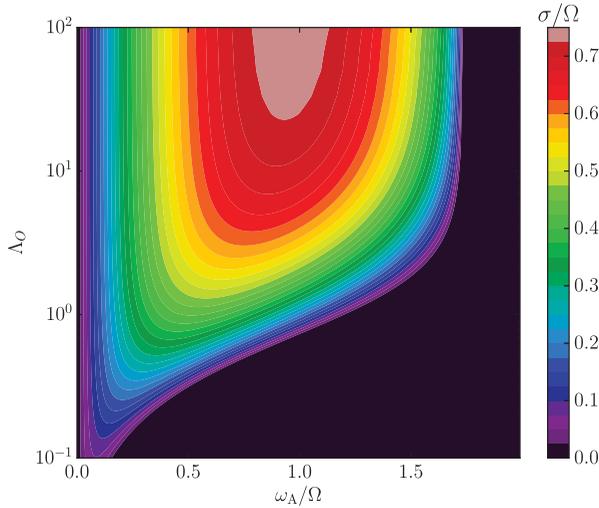
FIGURE 19. MRI growth rate as a function of the Ohmic Elsasser number Λ_O for a Keplerian disc ($q = 3/2$) with $k_x = 0$. Note the damping of the most unstable mode for $\Lambda_O = 1/\sqrt{3}$ and the survival of low growth rate modes in the limit $\omega_A \to 0$, $\Lambda_O \to 0$.

and $q = 3/2$. The stability of the most unstable ideal mode is often considered as a proxy for the stability of the flow. This mode has $\omega_A = \sqrt{3}\Omega/2$ so that the inequality reduces to

$$\Lambda_O^2 < \frac{1}{3} \to \text{Stability of the most unstable ideal MRI mode.} \tag{6.49}$$

Note, however, that this does not imply that all of the modes available to the disc are stable, and that the disc is stable. A more constraining criterion results from this and requires that the mode with the largest length scale is MRI stable, i.e. that

$$\frac{3\Omega^2}{k_{z,\text{min}}^2 V_{Az}^2} - 1 < \Lambda_O^{-2} \to \text{General disc stability.} \tag{6.50}$$

6.4.4. *Hall effect*

Hall-driven linear waves: The Hall effect is known to be at the origin of new linear waves. These waves can be captured by letting $\Omega \to 0$ and neglecting Ohmic and ambipolar diffusion. In this case the dispersion relation (6.43) gives

$$0 = \sigma^4 + \sigma^2 \left(2\omega_A^2 + \ell_H^2 \omega_A^2 k^2\right) + \omega_A^4$$
$$= \left(\sigma^2 + i\sigma \ell_H \omega_A k + \omega_A^2\right)\left(\sigma^2 - i\sigma \ell_H \omega_A k + \omega_A^2\right). \tag{6.51}$$

We recognise two waves with frequency $\omega \equiv i\sigma$ given by

$$\omega = \omega_A \left[\pm \frac{\ell_H k}{2} + \sqrt{\frac{\ell_H^2 k^2}{4} + 1} \right], \tag{6.52}$$

where '+' waves are known as whistlers or electron-cyclotron modes, whereas '−' waves are ion-cyclotron modes. The whistler frequency $\omega_H = \omega_A \ell_H k$ increases as k^2 in the limit

$k \to \infty$ whereas the ion-cyclotron frequency tends to a constant $\omega_{\mathrm{IC}} = \omega_A/(\ell_H k)$. Hence, these two waves behave very differently at the small scale. For positive ℓ_H, whistlers are right-handed polarised wave whereas ion-cyclotron are left-handed. Physically, whistlers are essentially an oscillation of the electrons fluid (or of the lightest charged particle), leaving all of the other components of the plasma unaffected. Whistler and ion-cyclotron waves become standard right and left-handed circularly polarised Alfvén wave in the limit $k \to 0$.

HSI: The HSI[13] is a new branch of instability (Kunz 2008), which is often confused with the traditional MRI, despite its different physical origin. It is essentially an instability of whistler waves under the action of shear.

To capture the HSI, one can let $\omega_A \to 0$ while keeping $\ell_H \omega_A > 0$. This 'low magnetisation' limit allows one to decouple the ions from the electrons, as is evident from (6.42). Neglecting Ohmic and ambipolar diffusion, one obtains the following dispersion relation

$$\left(\sigma^2 + \frac{k_z^2}{k^2} \kappa^2 \right) \left[\sigma^2 + \ell_H \omega_A k_z \left(\frac{k^2}{k_z^2} \ell_H k_z \omega_A - q\Omega \right) \right] = 0, \qquad (6.53)$$

which exhibits a linear instability when

$$\ell_H V_{Az} k_z^2 \left(k^2 \ell_H V_{Az} - q\Omega \right) < 0 \to \text{HSI unstable}. \qquad (6.54)$$

Interestingly, the HSI shows up only when $q\ell_H V_{Az} > 0$ or, in other words, when the vertical field points in the same direction as the rotation axis in Keplerian discs,[14] assuming $\ell_H > 0$. When the whistler frequency becomes too large ($k^2 \ell_H V_{Az} > q\Omega$), the instability disappears. For a given k_z, the most unstable mode has $k_x = 0$, hence $k = k_z$. For this reason, the HSI often shows up as channel-like mode in simulations, in a similar way to the MRI. Last, the maximum growth rate is identical to the MRI $\sigma_{\max} = q\Omega/2$ and is obtained for $k^2 \ell_H V_{Az} = q\Omega/2$.

Physically, this instability is a result of sheared whistler waves. If we look at the disc from the top with the vertical field pointing towards us, the magnetic perturbation of a whistler wave will tend to rotate counter-clockwise (figure 20). If the vertical field is positive (i.e. aligned with the rotation axis), the Keplerian shear is going to stretch the perturbation in the opposite direction, amplifying the toroidal field from the radial field and feeding the instability. In contrast, if the vertical field is anti-aligned with the rotation axis, the direction of rotation of whistler perturbations is that of the shear, resulting in a damped whistler wave.

Ion-cyclotron instability: The ion-cyclotron instability[15] is more difficult to isolate compared with the HSI because ion inertia cannot be neglected in this case. However, it is still possible to filter out whistler waves by letting $k_z \to \infty$ and keeping constant $k_x \lesssim k_z$. In this case, the whistler wave frequency ω_H becomes infinite whereas the cyclotron frequency remains finite. As we are looking for a finite growth rate, we assume σ is finite when $k \to \infty$. Neglecting ambipolar and Ohmic diffusion and keeping only $O(k^4)$ terms, the dispersion relation (6.43) becomes (see also Simon *et al.* 2015)

$$\sigma^2 \ell_H^2 k^2 \omega_A^2 + \left[\omega_A^2 + (2-q)\Omega \ell_H \omega_A k_z \right] \left[\omega_A^2 + 2\Omega \ell_H \omega_A k_z \right] = 0. \qquad (6.55)$$

---

[13]This instability is also named 'diffusive instability' (DI) by Pandey & Wardle (2012) to emphasise that rotation is unimportant for this instability, in contrast to the MRI.

[14]Physically, it is not the rotation but the background shear of the flow $S \equiv \partial_x V_y$ that matters. The general criterion is therefore $SV_{Az} < 0$.

[15]This instability is also named 'diffusive MRI' (DMRI) by Pandey & Wardle (2012).
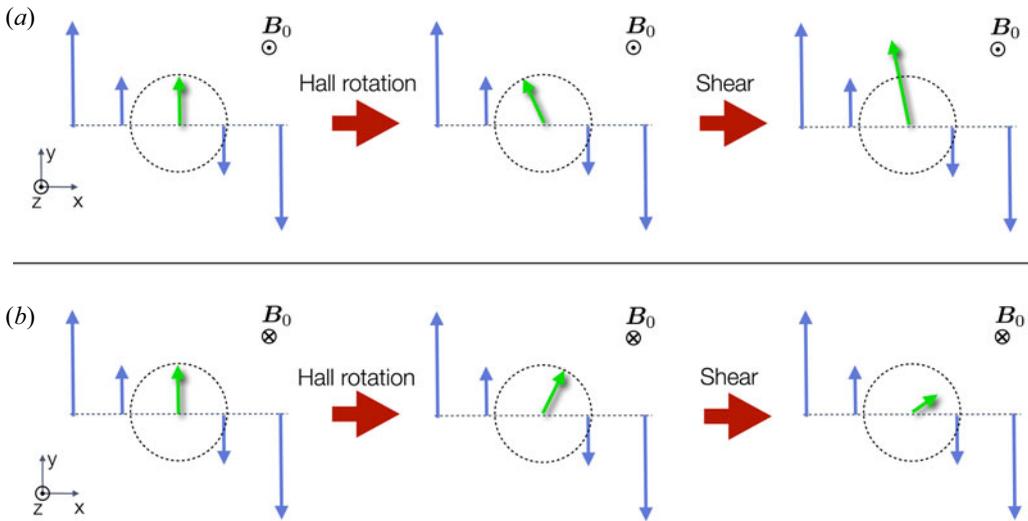
FIGURE 20. Physical principle of the HSI. The magnetic perturbation (in green) is rotated (*a*) clockwise or (*b*) counter-clockwise by the Hall effect depending on the polarity of the mean field B_0. When $B_0 > 0$, the rotated perturbation is amplified by the shear (in blue) while it is damped by the shear when $B_0 < 0$.

This relation clearly describes ion-cyclotron modes because in the limit $\Omega \to 0$, we recover the ion-cyclotron frequency $\sigma = \pm i\omega_{CI}$. It describes unstable modes, resulting from the interaction between ion-cyclotron waves and epicyclic motions, provided that

$$-\frac{V_{Az}^2}{(2-q)\Omega} < \ell_H V_{Az} < -\frac{V_{Az}^2}{2\Omega} \to \text{Ion-cyclotron unstable.} \qquad (6.56)$$

This inequality brings up two remarks: (i) the ion-cyclotron instability appears for anti-aligned field configurations, i.e. field configuration opposite to the HSI; (ii) this instability does not vanish in the limit $k_z \to \infty$. Therefore, there is no small-scale quenching similar to the MRI and the HSI, but there is a field strength limit.

General case: In the general case, one cannot compute the growth rate of the Hall-MRI easily. To illustrate the general growth rate of the Hall-MRI, we present in figure 21 the growth resulting from (6.43) in the Hall-only case. To quantify the intensity of the Hall effect, we have defined a Lundquist number based on the vertical wavenumber

$$\mathcal{L}_{\mathcal{H}}^* = \frac{1}{\ell_H k_z} = \frac{V_A}{\eta_H k_z}. \qquad (6.57)$$

We have chosen the most unstable modes $k_x = 0$ and assumed $k_z > 0$ so that $\omega_A < 0$ corresponds to B_{0z} anti-aligned with Ω. As can be seen in figure 21, we recover the MRI in the limit $\mathcal{L}_{\mathcal{H}} \to \infty$, which gives identical growth rates under the symmetry $\omega_A \to -\omega_A$. As $\mathcal{L}_{\mathcal{H}}^*$ decreases and the Hall effect increases, this symmetry is broken.

For $\omega_A > 0$ (aligned field configuration), the MRI becomes the HSI, and the optimum growth rate starts to move to lower ω_A. This is expected from our analysis because the maximum growth rate of the HSI is found for $\omega_A/\Omega \simeq \mathcal{L}_H^*/2$. We show in white filled contours the stability limit of the HSI obtained from (6.54). This limit matches the full dispersion relation for $\mathcal{L}_H^* \lesssim 0.3$.
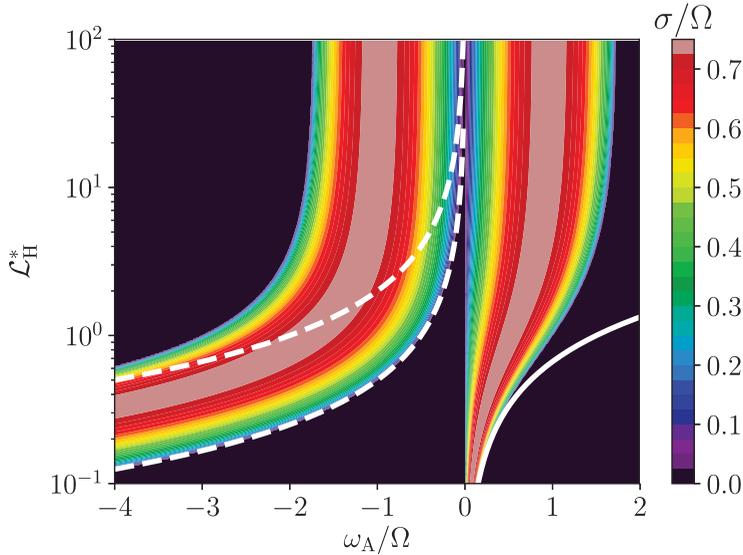
FIGURE 21. Hall-MRI growth rate as a function of the modified Hall Lundquist number $\mathcal{L}_{\mathcal{H}}^* = (\ell_H k_z)^{-1}$ for a Keplerian disc ($q = 3/2$) with $k_x = 0$. We have assumed $k_z > 0$ so that $\omega_A < 0$ corresponds to the anti-aligned case $V_{Az} < 0$. The white plain line corresponds to the HSI stability limit (6.54) and the white dashed lines to the ion-cyclotron stability limits (6.56).

For $\omega_A < 0$ (anti-aligned field configuration), the MRI becomes the ion-cyclotron instability and the optimum growth rate moves to higher $|\omega_A|$. The stability contours of the ion-cyclotron instability (6.56) are shown as dashed white lines. As for the HSI, they match the full dispersion relation for $\mathcal{L}_H^* \lesssim 0.3$.

Effect of Ohmic diffusion on the Hall-MRI: As shown previously, Ohmic diffusion tends to damp MRI modes, starting from the largest ω_A and moving to lower ω_A as Λ_O decreases. On the other hand, the Hall effect creates two distinct branches from the MRI, depending on the field alignment configuration. As Ohmic diffusion suppresses first high $|\omega_A|$ modes, the ion-cyclotron instability will be the first to be stabilised. On the other hand, because the HSI is living at lower $|\omega_A|$ compared with the MRI, it will be less affected by Ohmic diffusion than the MRI. This effect is illustrated in figure 22, demonstrating that HSI modes survive more easily to Ohmic diffusion, even when ideal MRI modes are suppressed.

We can go further and estimate the stability limit of the HSI in the limit $\omega_A \to 0$ while keeping $\ell_H \omega_A > 0$. In this limit, the dispersion relation of the HSI including Ohmic diffusion reads

$$\left(\sigma^2 + \frac{k_z^2}{k^2}\kappa^2\right)\left[(\sigma + \eta_O k^2)^2 + \ell_H \omega_A k_z\left(\frac{k^2}{k_z^2}\ell_H k_z \omega_A - q\Omega\right)\right] = 0, \qquad (6.58)$$

from which we can deduce a general instability criterion. For the most unstable HSI mode, this criterion reads

$$\Lambda_O^2 < \frac{1}{2}\Lambda_H^2 \rightarrow \text{Stability of the most unstable HSI mode}, \qquad (6.59)$$
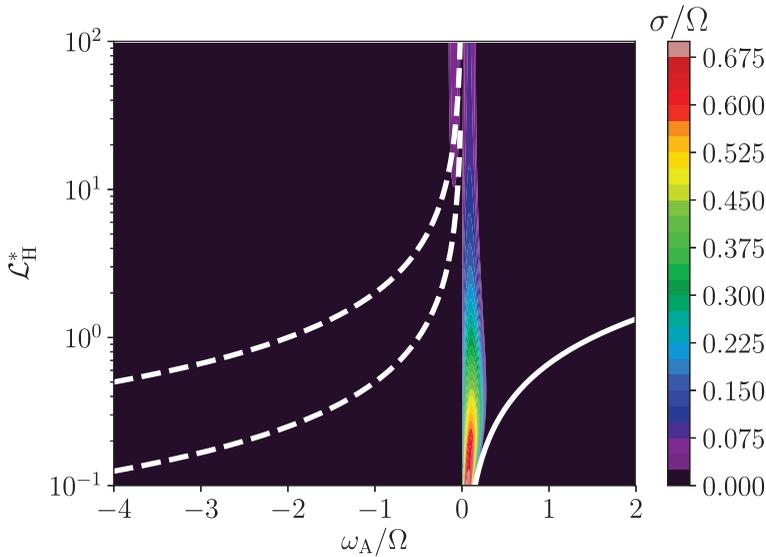
FIGURE 22. Same as figure 21 but including Ohmic diffusion with $\Lambda_O = 0.1$. Note the complete stabilisation of the ion-cyclotron branch and the strong damping of ideal MRI modes. Only HSI modes subsist at low $\mathcal{L}_H^*$ with growth rates comparable with the ideal MRI case.

where Λ_H is the Hall Elsasser number (see §3.4.2). This expression can be compared directly with the Ohmic-only criterion (6.49), and shows that, as expected, the HSI can make the system unstable even when $\Lambda_O \ll 1$, provided that $\Lambda_H \ll 1$ as well.

For this reason, some researchers have proposed that the MRI could be 'resuscitated' in regions having a strong Ohmic diffusion thanks to the presence of the Hall effect, which could be dominant in some parts of the disc (Wardle & Salmeron 2012). The physical interpretation of this effect is relatively simple. In the case of the MRI, the maximum growth rate is found for $\omega_A \sim \Omega$, in other words, the Alfvénic and rotation frequencies match. In the HSI case, it is the whistler and rotation frequencies $\omega_H \sim \Omega$ that have to match. In the limit of strong Hall effect, $\omega_H = \ell_H k_z \omega_A$. Therefore, when $\ell_H kz \gg 1$, the scale at which the whistler frequency matches the rotation frequency is much larger than the scale at which the Alfvén frequency matches the rotation frequency. In other words, the optimum HSI mode has a much larger wavelength than its MRI counterpart. For this reason, the HSI is less sensitive to diffusion than the MRI, because diffusion first damps small-scale modes.

Although these statements are true in the linear regime, the non-linear saturation of the HSI is not guaranteed to be similar to that of the MRI. Effectively, the linear analysis is unable to predict the turbulent angular momentum transport one could obtain from the HSI. As we show in §8.4, the HSI in the non-linear regime is indeed full of surprises.

6.4.5. *Ambipolar diffusion*

The linear ambipolar diffusion is a non-diagonal operator as shown in (6.42). The non-diagonal terms are non-zero only when $k_x V_{Ay} \neq 0$, i.e. when both non-axial wavevectors and guide fields are considered. Otherwise, ambipolar diffusion acts as a usual diffusion operator by damping magnetic perturbation. Note, however, that even in this case, it does not act as a scalar diffusion, unless $k_x = 0$ and $V_{Ay} = 0$.

Diagonal case $(k_x V_{Ay} = 0)$: In this case, the stability criterion is very similar to that of Ohmic diffusion. By requiring $\mathcal{C}_0 > 0$, one obtains

$$\frac{2q\Omega^2}{k^2 V_{Az}^2} - 1 < \kappa^2 \tau_A^2 \left(\frac{V_A}{V_{Az}}\right)^2 \rightarrow \text{Stability.} \tag{6.60}$$

This criterion shows that, as for Ohmic diffusion, the most unstable modes have $k_x = 0$. Moreover, it shows that the stability condition is affected by the toroidal field strength, a stronger V_{Ay} leading to a more stable system. As for Ohmic resistivity, we can derive a criterion for the stability of the most unstable ideal MRI mode in Keplerian flows, which reads

$$\Lambda_A^2 < \frac{1}{3}\left(\frac{V_A}{V_{Az}}\right)^2 \rightarrow \text{Stability of the most unstable ideal MRI mode,} \tag{6.61}$$

and a criterion for the stability of all the modes available smaller than the minimum vertical wavenumber $k_{z,\min}$:

$$\frac{3\Omega^2}{k_{z,\min}^2 V_{Az}^2} - 1 < \Lambda_A^{-2}\left(\frac{V_A}{V_{Az}}\right)^2 \rightarrow \text{General disc stability.} \tag{6.62}$$

This shows that the MRI can be stabilised even for $\Lambda_A > 1$ provided that the toroidal field dominates over the poloidal field. The field topology is therefore a key point for the stability under the action of ambipolar diffusion. The physical reason for this is relatively simple. The MRI mechanism is only sensitive to the magnetic tension, which is due to B_z in the absence of non-axisymmetric perturbations. In contrast, ambipolar diffusion increases with the total field strength. Hence, an increase in V_{Ay} results in an increase of the effective diffusion η_A, whereas the MRI feedback loop is left unperturbed.

Non-diagonal case $(k_x V_{Ay} \neq 0)$: In this case, the non-diagonal terms of ambipolar diffusion can act as a positive feedback loop in the induction equation. The criterion for this positive feedback is easily obtained from the dispersion relation (6.43): $q\Omega k_x k_z V_{Ay} V_{Az} < 0$. An illustration of the effect of this term on the general stability property is shown in figure 23 where we have assumed $V_{Ay} = V_{Az}$, $q = 3/2$ and $\Lambda_A = 0.4$. We observe that oblique modes $(k_x \neq 0)$ tend to be more unstable than axial modes when $\Lambda_A < 1$. It can be shown that oblique modes are always unstable, albeit with a vanishing growth rate, in the limit $\Lambda_A \rightarrow 0$ provided that k_x/k_z is sufficiently large (Kunz & Balbus 2004).

6.4.6. *Vertical stratification*

All of the results described previously were computed ignoring vertical stratification. It should, however, be pointed out that because discs are vertically stratified, unstratified results should be taken with care. Let us emphasise here a few important results regarding the effect of stratification on MRI modes. In this section, we assume that the disc is vertically isothermal (see § 4.2) and is only threaded by a vertical field (no toroidal field). We quantify the intensity of the imposed field with the plasma beta parameter in the midplane:

$$\beta_{\text{mid}} \equiv \frac{8\pi P_{\text{mid}}}{B^2} = \frac{2c_s^2}{V_{Az}^2(z=0)}. \tag{6.63}$$

Ideal MRI: The ideal MRI case with an isothermal disc is presented in details by Latter, Fromang & Gressel (2010). We therefore just summarise here the main conclusions.
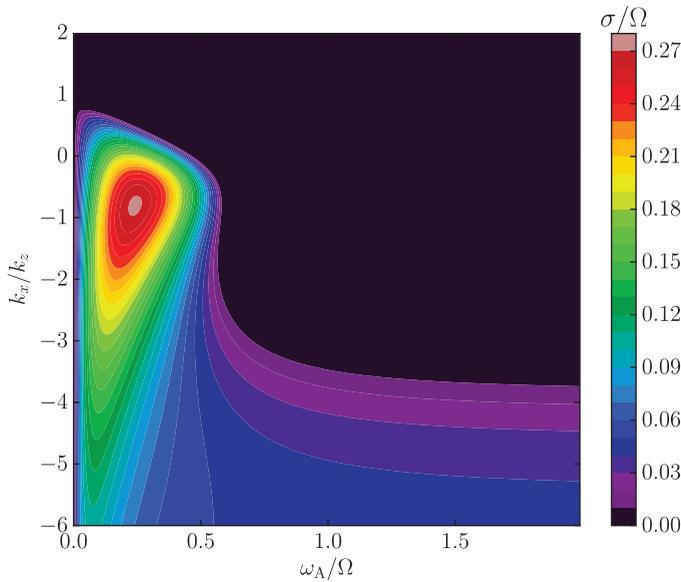
FIGURE 23. MRI growth rate with ambipolar diffusion and $\Lambda_A = 0.4$. We have chosen $V_{Ay} = V_{Az}$ and $k_x \neq 0$ to illustrate the effect of non-diagonal ambipolar terms. The most unstable mode in this case is found for $k_x < 0$ and a weakly unstable branch exists for $k_x/k_z \simeq 3.5$ in the limit $\omega_A \to \infty$. We name these modes 'oblique ambipolar modes'. Note, however, their low relative growth rate.

First, stratified eigenvalues (growth rates) satisfy the same dispersion relation (6.14) as non-stratified modes with $V_A = V_A(z = 0)$. Vertical wavenumbers k are quantised, but eigenmodes are not simple harmonic functions in the z direction. Latter *et al.* (2010) have shown that the first eigenmodes have $k_n H = 1.1584, 2.0796, 2.9829, 3.8798, \ldots$ for $n = 1 \ldots 4$. Taking k_1 as the lowest wavenumber accessible to the system, the MRI is stable for all possible eigenmodes in Keplerian discs ($q = 3/2$) if $V_{Az}(z = 0) > \sqrt{3}\Omega/k_1$, i.e. if

$$\beta_{\mathrm{mid}} < \frac{2(k_1 H)^2}{3} \simeq 0.89 \to \text{Stability}. \tag{6.64}$$

This value can vary slightly depending on the boundary conditions for the perturbation as $z \to \infty$. Overall, it is safe to assume that the MRI exists provided that $\beta_{\mathrm{mid}} \gtrsim 1$.

Second, eigenmodes come in two symmetries: odd symmetry modes (which corresponds to odd n) have odd $v_{x,y}(z)$, even $B_{x,y}(z)$ and exhibit a maximal magnetic perturbation at the midplane and no velocity perturbation at this location. Even symmetry modes (even n), on the other hand, have a velocity jet in the midplane and no magnetic perturbation at $z = 0$ (see figure 24). The MRI does not choose specifically any symmetry in these local stratified models: both even and odd n follow the same dispersion relation.

Third, the fastest growing modes tend to be more oscillatory in the disc midplane at higher β_{mid}. This is because qualitatively, one expects $kV_{Az} \sim \Omega$ for the fastest growing mode, hence in the midplane, one obtains $k_z H \sim \sqrt{\beta_{\mathrm{mid}}}$. An example of such a mode is shown in figure 25.

Ohmic and ambipolar diffusion: As shown in the non-stratified section, Ohmic and ambipolar diffusion tend to suppress the MRI when the Elsasser numbers are less than 1. Typically, this situation occurs in the densest regions of the disc (see figure 10). We show
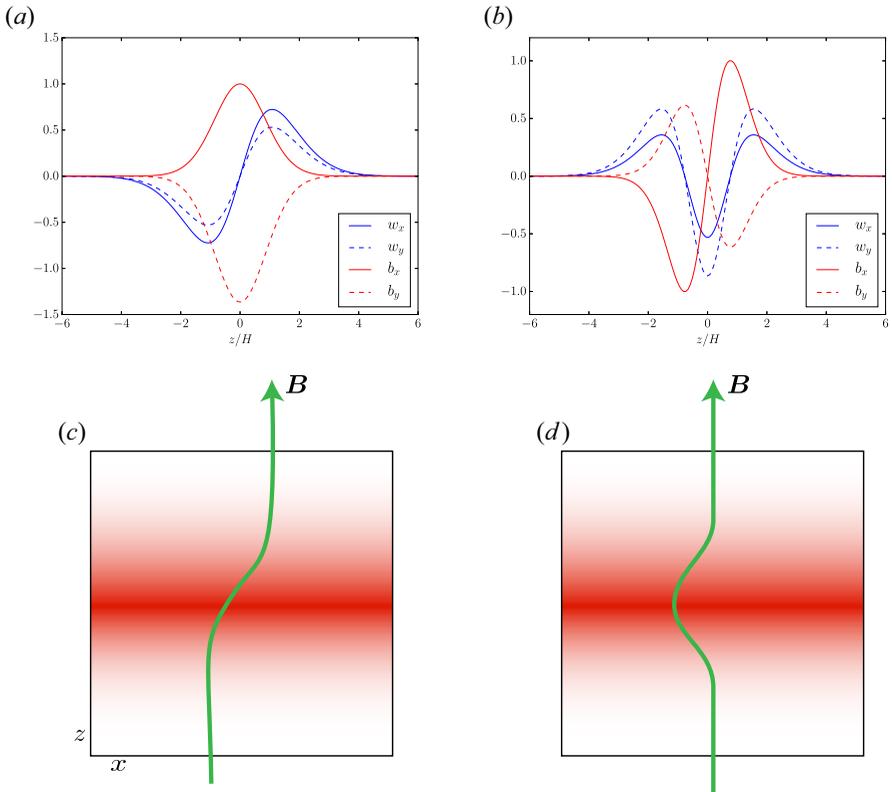
FIGURE 24. (*a,b*) MRI eigenmodes computed at $\beta_{\mathrm{mid}} = 5$ for $n = 1$ (left, odd mode, $\sigma = 0.72\Omega$) and $n = 2$ (right, even mode $\sigma = 0.67\Omega$). (*c,d*) Schematic representation of the field perturbation owing to odd (left) and even (right) modes.
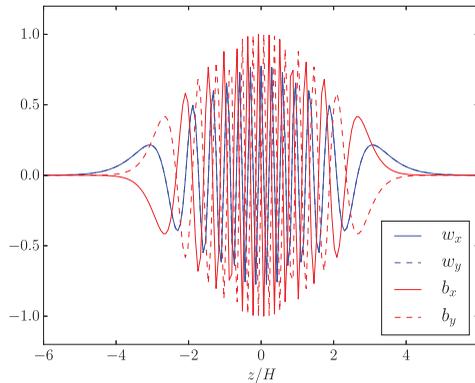
FIGURE 25. Fastest growing MRI eigenmode ($\sigma = 0.75\Omega$) in a stratified model for $\beta_{\mathrm{mid}} = 10^3$. Note the strongly oscillating behaviour close to the midplane.

in figure 26 the most unstable MRI eigenmode with $\beta = 10^3$ resulting from our fiducial metal-free mode at $R = 1$ AU. As expected, the MRI is strongly suppressed where $\Lambda_O < 1$. Ambipolar diffusion tends to reduce the overall growth rate but does not affect the shape of the eigenmode significantly. This is a perfect linear illustration of the historical 'layered
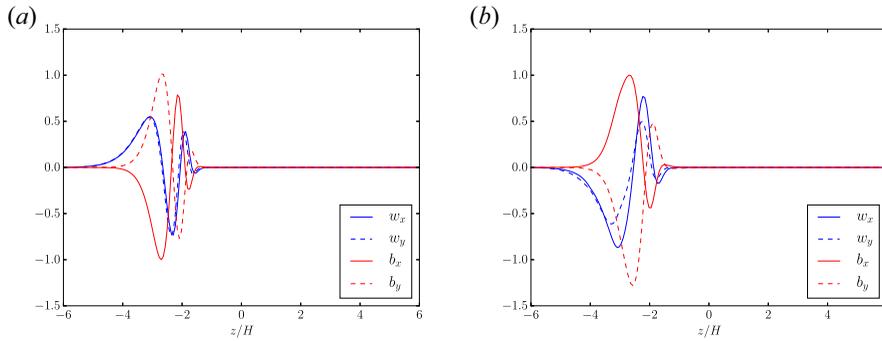
FIGURE 26. Fastest growing MRI eigenmode in a stratified model for $\beta_{\mathrm{mid}} = 10^3$ using the diffusivity profiles from figure 10 at $R = 1$ AU. (*a*) Including Ohmic diffusion only ($\sigma = 0.72\Omega$) and (*b*) including Ohmic and ambipolar diffusion ($\sigma = 0.56\Omega$). Note the lack of perturbations in the disc midplane due to Ohmic diffusion and partially to ambipolar diffusion. Equivalent eigenmodes with the same growth rates are found on the $z > 0$ side of the disc.

accretion' paradigm where the MRI still survives at the surface of the disc, leaving the disc midplane up to one to two scale-heights essentially magnetically dead (Gammie 1996).

Note that when the linear eigenmodes do not propagate in the disc, as is the case here, there are no well-defined 'odd' and 'even' modes. Instead, there are two families of modes localised either at the top or the bottom side of the disc, with identical growth rates. It is, in principle, possible to combine these top and bottom modes to reconstruct even and odd modes, but here we have chosen here to stress the lack of propagation through the disc of the perturbation.

Hall effect: As discussed previously, the Hall effect can potentially revive dead zones when the field is aligned with the rotation axis, thanks to the HSI branch of the Hall-MRI. This effect is illustrated in figure 27 where the MRI eigenmode now propagates in the midplane resulting in a fully active disc column at $R = 1$ AU when $V_{Az}\Omega > 0$. In this case, however, the growing perturbation is mostly magnetic and velocity perturbations are mostly absent in the midplane of the eigenmode. This is because the HSI is an unstable whistler mode, which is essentially an electronic perturbation leaving the ions (and the neutrals) unperturbed. When the field is anti-aligned, we recover results similar to the purely diffusive case, albeit with a slightly reduced growth rate.

7. The helicoidal MRI

The helicoidal MRI (HMRI) was first identified by Hollerbach & Rüdiger (2005) using a spectral analysis of a rotating Taylor–Couette flow. This instability is known to work for arbitrarily low magnetic Reynolds numbers in rotating sheared flows, implying that it could work both in liquid sodium experiments and in weakly ionised astrophysical discs such as PPDs. For the sake of completeness, let us show here the origin of this instability and discuss its application to Keplerian discs.

In contrast to the non-axisymmetric MRI (which is fully local), the HMRI is tightly linked to the presence of curvature in the physics of the system, hence it cannot be captured in Hill's approximation (even though a WKB analysis can capture it in the global geometry; see Kirillov, Stefani & Fukumoto (2014) and the following discussion). Therefore, let us return to the evolution equations in global geometry.
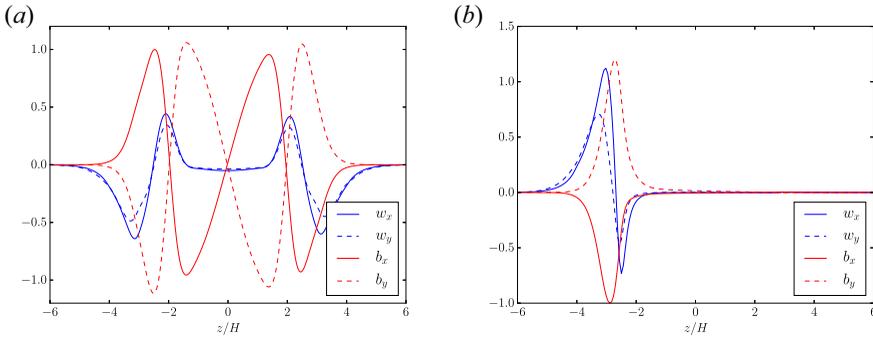
FIGURE 27. Fastest growing MRI eigenmode in a stratified model for $\beta_{\mathrm{mid}} = 10^3$ using the diffusivity profiles from figure 10 at $R = 1$ AU including Ohmic, Hall and ambipolar diffusion: (*a*) assuming $V_{Az}\Omega > 0$ ($\sigma = 0.65\Omega$) and (*b*) assuming $V_{Az}\Omega < 0$ ($\sigma = 0.51\Omega$). The 'dead zone' is now subject to long-wavelength perturbations when $V_{Az}\Omega > 0$, as a result of the HSI.

7.1. *Full set of equations in cylindrical geometry*

We consider the motion of a conductive fluid in cylindrical coordinates. We denote the velocity of the fluid **u** and the magnetic field **B**. We consider the inviscid, incompressible and ideal equations of MHD, which reads, component by component,

$$
\partial_t u_R + u_R \partial_R u_R + \frac{u_\phi}{R} \partial_\phi u_R - \frac{u_\phi^2}{R} + u_z \partial_z u_R
$$
$$
= -\frac{1}{\rho_0} \partial_R \left(P + \frac{B^2}{8\pi} \right) + \frac{1}{4\pi\rho_0} B_R \partial_R B_R + \frac{1}{4\pi\rho_0} \frac{B_\phi}{R} \partial_\phi B_R - \frac{1}{4\pi\rho_0} \frac{B_\phi^2}{R} + \frac{1}{4\pi\rho_0} B_z \partial_z B_R \tag{7.1}
$$

$$
\partial_t u_\phi + u_R \partial_R u_\phi + \frac{u_\phi}{R} \partial_\phi u_\phi + \frac{u_R u_\phi}{R} + u_z \partial_z u_\phi
$$
$$
= -\frac{1}{\rho_0} \frac{1}{R} \partial_\phi \left(P + \frac{B^2}{8\pi} \right) + \frac{1}{4\pi\rho_0} B_R \partial_R B_\phi + \frac{1}{4\pi\rho_0} \frac{B_\phi}{R} \partial_\phi B_\phi + \frac{1}{4\pi\rho_0} \frac{B_R B_\phi}{R} + B_z \partial_z B_\phi \tag{7.2}
$$

$$
\partial_t u_z + u_R \partial_R u_z + \frac{u_\phi}{R} \partial_\phi u_z + u_z \partial_z u_z
$$
$$
= -\frac{1}{\rho_0} \partial_z \left(P + \frac{B^2}{8\pi} \right) + \frac{1}{4\pi\rho_0} B_R \partial_R B_z + \frac{1}{4\pi\rho_0} \frac{B_\phi}{R} \partial_\phi B_z + \frac{1}{4\pi\rho_0} B_z \partial_z B_z \tag{7.3}
$$

for the equation of motion, which we combine into the induction equation,

$$
\partial_t B_R + u_R \partial_R B_R + \frac{u_\phi}{R} \partial_\phi B_R + u_z \partial_z B_R = B_R \partial_R u_R + \frac{B_\phi}{R} \partial_\phi u_R + B_z \partial_z u_R \tag{7.4}
$$

$$
\partial_t B_\phi + u_R \partial_R B_\phi + \frac{u_\phi}{R} \partial_\phi B_\phi + u_z \partial_z B_\phi + \frac{u_\phi B_R}{R} = B_R \partial_R u_\phi + \frac{B_\phi}{R} \partial_\phi u_\phi + B_z \partial_z u_\phi + \frac{u_R B_\phi}{R} \tag{7.5}
$$

$$\partial_t B_z + u_R \partial_R B_z + \frac{u_\phi}{R} \partial_\phi B_z + u_z \partial_z B_z = B_R \partial_R u_z + \frac{B_\phi}{R} \partial_\phi u_z + B_z \partial_z u_z, \tag{7.6}$$

and the continuity equation,

$$\frac{1}{R} \partial_R R u_R + \frac{1}{R} \partial_\phi u_\phi + \partial_z u_z = 0. \tag{7.7}$$

7.2. Linearisation

We linearise the equations with respect to a background flow

$$\boldsymbol{u}_0 \equiv R\Omega(R)\boldsymbol{e}_\phi, \tag{7.8}$$

where $\Omega(R)$ is the angular velocity profile of the mean flow (arbitrary, for the moment). In addition, we consider the case with a mean magnetic field defined as

$$\boldsymbol{B}_0 = B_{\phi,0}(R)\boldsymbol{e}_\phi + B_{z,0}\boldsymbol{e}_z \tag{7.9}$$

so the only spatial dependency is in the R direction for the toroidal component of the field. In addition, we are going to assume that the flow is axisymmetric, so that we can cancel ∂_ϕ derivatives.

The velocity and magnetic fields are then expanded as

$$\left. \begin{aligned} \boldsymbol{u} &= \boldsymbol{u}_0 + \boldsymbol{v}, \\ \boldsymbol{B} &= \boldsymbol{B}_0 + \boldsymbol{b}, \end{aligned} \right\} \tag{7.10}$$

where deviations are assumed to be infinitely small compared with the means. The linearised equations of motion eventually read

$$\left. \begin{aligned} \partial_t v_R &= -\partial_R \Pi - \frac{2B_{\phi,0}b_\phi}{4\pi\rho_0 R} + \frac{1}{4\pi\rho_0} B_{z,0} \partial_z b_R + 2\Omega v_\phi \\ \partial_t v_\phi &= \frac{1}{4\pi\rho_0} b_R \frac{1}{R} \partial_R R B_{\phi,0} + \frac{1}{4\pi\rho_0} B_{z,0} \partial_z b_\phi - (2\Omega + R\partial_R\Omega)v_R \\ \partial_t v_z &= -\partial_z \Pi + \frac{1}{4\pi\rho_0} B_{z,0} \partial_z b_z \end{aligned} \right\} \tag{7.11}$$

whereas the induction equation reads

$$\left. \begin{aligned} \partial_t b_R &= B_{z,0} \partial_z v_R, \\ \partial_t b_\phi &= b_R R\partial_R\Omega + B_{z,0} \partial_z v_\phi - v_R R\partial_R \frac{B_{\phi,0}}{R}, \\ \partial_t b_z &= B_{0,z} \partial_z v_z. \end{aligned} \right\} \tag{7.12}$$

In order to make the notations more concise and consistent with the usual MRI derivation, we define the following coefficients

$$\left. \begin{aligned} q &\equiv -\frac{\mathrm{d}\log\Omega}{\mathrm{d}\log R}, \\ p &\equiv -\frac{\mathrm{d}\log B_{\phi,0}}{\mathrm{d}\log R}. \end{aligned} \right\} \tag{7.13}$$

It should be noted that the particular case $p = 1$ corresponds to a case where no axial current is present in the system. This is the reference case considered by Hollerbach & Rüdiger (2005)

Using this notation, the equations eventually read

$$
\left.
\begin{aligned}
\partial_t v_R &= -\partial_R \Pi - \frac{2 B_{\phi,0} b_\phi}{4\pi \rho_0 R} + \frac{1}{4\pi \rho_0} B_{z,0} \partial_z b_R + 2\Omega v_\phi \\[2mm]
\partial_t v_\phi &= -(p-1)\frac{B_{\phi,0}}{4\pi \rho_0 R} b_R + \frac{1}{4\pi \rho_0} B_{z,0} \partial_z b_\phi - \Omega(2-q)v_R \\[2mm]
\partial_t v_z &= -\partial_z \Pi + \frac{1}{4\pi \rho_0} B_{z,0} \partial_z b_z
\end{aligned}
\right\}
\tag{7.14}
$$

whereas the induction equation reads

$$
\left.
\begin{aligned}
\partial_t b_R &= B_{z,0} \partial_z v_R, \\[2mm]
\partial_t b_\phi &= -q\Omega b_R + B_{z,0} \partial_z v_\phi + (p+1)\frac{B_{\phi,0}}{R} v_R, \\[2mm]
\partial_t b_z &= B_{0,z} \partial_z v_z.
\end{aligned}
\right\}
\tag{7.15}
$$

7.3. *WKB approximation and pressure*

We are going to look for a local solution, i.e. a solution with fast variation compared with R. We therefore focus on a tiny region around a fiducial radius R_0 defining $x \equiv R - R_0$, and expand the solution as

$$
\boldsymbol{v}, \boldsymbol{b} \propto \exp\left[\sigma t + \mathrm{i}(k_x x + k_z z)\right],
\tag{7.16}
$$

thus we obtain

$$
\left.
\begin{aligned}
\sigma v_R &= -\mathrm{i}k_R \Pi - 2\omega_{A\phi} b_\phi + \mathrm{i}\omega_{Az} b_R + 2\Omega v_\phi, \\
\sigma v_\phi &= -(p-1)\omega_{A\phi} b_R + \mathrm{i}\omega_{Az} b_\phi - \Omega(2-q)v_R, \\
\sigma v_z &= -\mathrm{i}k_z \Pi + \mathrm{i}\omega_{Az} b_z,
\end{aligned}
\right\}
\tag{7.17}
$$

whereas the induction equation reads

$$
\left.
\begin{aligned}
\sigma b_R &= \mathrm{i}\omega_{Az} v_R, \\
\sigma b_\phi &= -q\Omega b_R + \mathrm{i}\omega_{A\phi} v_\phi + (p+1)\omega_{A\phi} v_R, \\
\sigma b_z &= \mathrm{i}\omega_{Az} v_z,
\end{aligned}
\right\}
\tag{7.18}
$$

where we have defined the Alfvén frequency $\omega_{Az} = k_z B_{0,z}/\sqrt{4\pi \rho_0}$ and the toroidal Alfvén frequency $\omega_{A\phi} = B_{\phi,0}/(R_0\sqrt{4\pi \rho_0})$. It is obvious from here that the vertical equation of motion and induction are just fed by the horizontal problem, but they do not have any feedback in the horizontal plane. We therefore consider only the horizontal equations without loss of generality. In order to solve for the pressure, we dot the equation of motion

by k to obtain an equation for Π. This gives

$$\Pi = \frac{2ik_R\omega_{A\phi}b_\phi}{k^2} - 2i\Omega v_\phi \frac{k_R}{k^2}. \tag{7.19}$$

This eventually leads to the following linear problem

$$\begin{pmatrix} -\sigma & 2\Omega g_{zz} & i\omega_{Az} & -2\omega_{A\phi}g_{zz} \\ -(2-q)\Omega & -\sigma & -(p-1)\omega_{A\phi} & i\omega_{Az} \\ i\omega_{Az} & 0 & -\sigma & 0 \\ (p+1)\omega_{A\phi} & i\omega_{Az} & -q\Omega & -\sigma \end{pmatrix} \begin{pmatrix} v_r \\ v_\phi \\ b_r \\ b_\phi \end{pmatrix} = 0, \tag{7.20}$$

where $g_{zz} = k_z^2/k^2$. This has non-trivial roots provided that the matrix determinant cancels out. This condition leads to a dispersion relation on σ:

$$\sigma^4 + \sigma^2 \left[2\omega_{Az}^2 + \kappa^2 g_{zz} + 2(p+1)\omega_{A\phi}^2 g_{zz} \right] - 8i\sigma\,\omega_{Az}\omega_{A\phi}\Omega g_{zz}$$
$$+ \omega_{Az}^2 \left[\omega_{Az}^2 - 2q\Omega^2 g_{zz} + 2(p-1)\omega_{A\phi}^2 g_{zz} \right] = 0. \tag{7.21}$$

We recognise the usual MRI-related form of the dispersion relation in the last term. The case $p = 1$, which corresponds to the HMRI initially derived by Hollerbach & Rüdiger (2005), therefore does not affect the original MRI criterion. As we show later, the HMRI driving term is actually the linear term in σ.

7.4. *HMRI with resistivity*

Adding resistivity η means that the linearised induction equations are modified as follows:

$$\left.\begin{aligned} \sigma b_R &= i\omega_{Az}v_R - \eta k^2 b_R, \\ \sigma b_\phi &= -q\Omega b_R + i\omega_{A\phi}v_\phi + (p+1)\omega_{A\phi}v_R - \eta k^2 b_\phi, \\ \sigma b_z &= i\omega_{Az}v_z - \eta k^2 b_z. \end{aligned}\right\} \tag{7.22}$$

The resulting linear system is then transformed into

$$\begin{pmatrix} -\sigma & 2\Omega g_{zz} & i\omega_{Az} & -2\omega_{A\phi}g_{zz} \\ -(2-q)\Omega & -\sigma & -(p-1)\omega_{A\phi} & i\omega_{Az} \\ i\omega_{Az} & 0 & -\sigma - \eta k^2 & 0 \\ (p+1)\omega_{A\phi} & i\omega_{Az} & -q\Omega & -\sigma - \eta k^2 \end{pmatrix} \begin{pmatrix} v_r \\ v_\phi \\ b_r \\ b_\phi \end{pmatrix} = 0, \tag{7.23}$$

which results in the following dispersion relation

$$\sigma^2(\sigma + \eta k^2)^2 + \sigma(\sigma + \eta k^2) \left[2\omega_{Az}^2 + 2(p+1)\omega_{A\phi}^2 g_{zz} \right] + (\sigma + \eta k^2)^2 \kappa^2 g_{zz}$$
$$- i\omega_{Az}\omega_{A\phi}\Omega g_{zz} \left[(8-2q)(\sigma + \eta k^2) + 2q\sigma \right]$$
$$+ \omega_{Az}^2 \left[\omega_{Az}^2 - 2q\Omega^2 g_{zz} + 2(p-1)\omega_{A\phi}^2 g_{zz} \right] = 0. \tag{7.24}$$

This can be combined into a standard fourth-order polynomial in σ:

$$\sigma^4 + 2\sigma^3\eta k^2 + \sigma^2 \left[2\omega_{Az}^2 + \kappa^2 g_{zz} + 2(p+1)\omega_{A\phi}^2 g_{zz} + \eta^2 k^4 \right]$$
$$- \sigma \left[8i\omega_{Az}\omega_{A\phi}\Omega g_{zz} - 2\eta k^2 \left(\omega_{Az}^2 + \kappa^2 g_{zz} \right) \right]$$
$$+ \omega_{Az}^2 \left[\omega_{Az}^2 - 2q\Omega^2 g_{zz} + 2(p-1)\omega_{A\phi}^2 g_{zz} \right]$$
$$- (8-2q)i\omega_{Az}\omega_{A\phi}\Omega g_{zz}\eta k^2 + \eta^2 k^4 \kappa^2 g_{zz} = 0, \tag{7.25}$$

which is formally equivalent to Kirillov *et al.* (2014, appendix B). Although this dispersion relation contains both the MRI and the HMRI in the resistive regime, it is difficult to disentangle the two instabilities. We therefore follow Liu *et al.* (2006) and consider the limit where the resistivity is much larger than the other terms.

7.5. *Inductionless limit*

In the inductionless limit, we neglect magnetic induction by considering roots $\sigma \ll \eta k^2$. We therefore introduce a small parameter $\varepsilon = 1/\eta k^2$, and expand the dispersion relation at first order in ε. This suppresses the strongly damped magnetic modes, having $\sigma \sim -\eta k^2$, which are primarily damped Alfvén waves and on which the usual MRI lives. For this reason, the inductionless limit allows us to suppress (stabilised) MRI modes. Following this limit, we get a second-order dispersion relation:

$$\sigma^2 + \sigma\varepsilon \left[2\omega_{Az}^2 + 2(p+1)\omega_{A\phi}^2 g_{zz}\right] + \kappa^2 g_{zz} - 2i\varepsilon\omega_{Az}\omega_{A\phi}\Omega g_{zz}(4-q) = 0. \tag{7.26}$$

The roots are simple to obtain and give at first order in ε:

$$\sigma = \pm i\kappa g_{zz}^{1/2} + \varepsilon\left[\pm\frac{1}{\kappa}\omega_{Az}\omega_{A\phi}\Omega g_{zz}^{1/2}(4-q) - \omega_{Az}^2 - (p+1)\omega_{A\phi}^2 g_{zz}\right]. \tag{7.27}$$

As can be seen, this growth rate describes a small deviation from pure epicyclic oscillations. Therefore, in contrast to the MRI, which is an instability of the (slow) Alvén branch, the HMRI is an overstability of the epicyclic branch. This explains, in part, its survival in the limit of small magnetic Reynolds numbers.

The instability clearly arises for the $+$ sign of the roots, when

$$\frac{1}{\kappa}\omega_{Az}\omega_{A\phi}\Omega g_{zz}^{1/2}(4-q) - \omega_{Az}^2 - (p+1)\omega_{A\phi}^2 g_{zz} > 0 \rightarrow \text{Instability}. \tag{7.28}$$

As this condition is on a second-order polynomial on $\omega_{A\phi}$, it is strictly equivalent to requiring that the discriminant of the polynomial is positive, i.e. that

$$\frac{(4-q)^2}{2(2-q)} - 4(p+1) > 0. \tag{7.29}$$

Assuming $\kappa^2 > 0$, this is again equivalent to asking that

$$q^2 + 8qp - 16p > 0 \tag{7.30}$$

which is (again) a second-order polynomial in q, which we ought to resolve. The stability condition is then simply

$$q > 4(\sqrt{p(1+p)} - p) \quad \text{or} \quad q < -4(\sqrt{p(1+p)} + p) \rightarrow \text{Instability}. \tag{7.31}$$

In the current-free configuration ($p = 1$), we recover the so-called Liu limit (Liu *et al.* 2006) $q > 4(\sqrt{2} - 1) \simeq 1.657$, i.e. that a current-free toroidal field is stable in the Keplerian regime. It is the main reason why this instability has been mostly neglected in the astrophysical context.

However, if one allows for an axial current in the system ($p \neq 1$), it is possible to recover the instability for Keplerian rotation profiles. This can be deduced from (7.30):

$$p < \frac{q^2}{8(2-q)} \rightarrow \text{Instability}. \tag{7.32}$$

In the Keplerian regime, we therefore require $p < 9/16 = 0.5625$.

The maximum growth rates are obtained from the mean of the two roots of (7.28), which is simply

$$\omega_{A\phi,\max} = \frac{1}{2\kappa}\omega_{Az}\Omega g_{zz}^{1/2}(4-q).$$ (7.33)

The maximum growth rate is then given by

$$\sigma_{\max} = i\kappa g_{zz}^{1/2} + \frac{V_{Az}^2}{\eta}\left[g_{zz}\frac{(4-q)^2}{2(2-q)(p+1)} - 1\right],$$ (7.34)

which turns out to be spatially scale-free. From these results, we can deduce a few important results for Keplerian discs with $q = 3/2$. First, one finds that the growth rate in units of the local orbital frequency scales like the Ohmic Elsasser number, i.e. $\mathrm{Im}(\sigma_{\max})/\Omega \simeq \Lambda_O$. This implies relatively low growth rates in PPDs, unless they are hosting a dynamically strong vertical field (with $V_{Az} \sim \Omega H$ where H is the disc thickness). In addition, these 'optimum' growth rates can only be reached for $\omega_{A\phi} \sim \omega_{Az}$ (from (7.33)). Combining these constraints, we find that $V_{A\phi} \sim R\Omega Hk$. Given that, by construction, $Hk > 1$ (because the vertical wavelength needs to fit in the disc), this shows that the azimuthal Alfvén velocity needs to be of the order of the Keplerian velocity (or larger). This is a very strong azimuthal field, well above equipartition. Among other things, such a strong field implies that the disc is no longer Keplerian as the radial tension and magnetic pressure forces are of the order of the central gravity.

For these reasons, the role played by the HMRI in the context of PPDs has been mostly ignored, as it lives in a parameter regime probably distant from that of real systems, which are known to be in Keplerian rotation, and hence for which $\omega_{A\phi} \ll \Omega$.

PART THREE: Non-linear saturation of the MRI

The saturation of the MRI has been mostly studied in local shearing box simulations. As the seminal paper by Hawley *et al.* (1995), several studies have been dedicated to non-ideal effects and the role they play in the MRI saturation.

Let us first distinguish unstratified and vertically stratified shearing box models. In the unstratified model, the box is periodic in the vertical direction, making the system much simpler to analyse. It corresponds to the SSB model of Umurhan & Regev (2004) (see also § 5.2). When vertical stratification is included, the flow is usually allowed to escape through the vertical boundary conditions, potentially leading to outflows. Here, we first focus on unstratified models before moving to stratified models.

It is also important to separate simulations with and without a mean field. Although the MRI (the linear instability) does require a mean field to exist, it has been shown that in the non-linear regime, MHD turbulence exists without any mean field. We call this case the 'MRI dynamo'. Although this was initially a mere curiosity, it turned out to be the fiducial configuration for many stratified and even global simulations owing to the technical difficulties associated with simulations with a mean field.

In all of these models, one uses box averages, defined for a quantity Q by

$$\langle Q \rangle = \iiint d^3\boldsymbol{x}\, Q.$$ (7.35)

One of the key elements is then to quantify the α parameter from the turbulent stress, which is defined as

$$\alpha \equiv \frac{1}{\langle P \rangle}\left\langle \rho w_x w_y - \frac{B_x B_y}{4\pi}\right\rangle,$$ (7.36)

which is also sometimes averaged in time.

8. Unstratified models

Except in rare circumstances, unstratified models are periodic in the vertical direction. In this case, the energy conservation equation (5.22) reads

$$\partial_t \langle \mathcal{E}_{\text{Mech}} \rangle = \langle P \boldsymbol{\nabla} \cdot \boldsymbol{w} \rangle + \langle \boldsymbol{E}_{\text{NI}} \cdot \boldsymbol{J} \rangle + q\Omega_0 \left\langle \rho w_x w_y - \frac{B_x B_y}{4\pi} \right\rangle. \tag{8.1}$$

In numerical experiments, because MRI turbulence is subsonic, the contribution of the $P\,\mathrm{d}V$ term is small whereas the x–y stress term is definite positive (it is the term driving accretion). Therefore, a quasi-steady turbulent state for which $\mathcal{E}_{\text{mech}} \sim$ constant necessarily implies that non-ideal effects (or viscous effects, which have been ignored here) are non-negligible. This statement can be understood in terms of Kolmogorov's turbulent cascade argument. In the energy equation, the stress is a source of mechanical energy. In the cascade argument, it corresponds to the energy injection rate. For the cascade to be in a steady state, some dissipative effect must enter the picture at some scale to dissipate what has been injected at the beginning of the cascade. One concludes from this argument that there is no such thing as ideal MRI turbulence.

Here, we keep the terminology 'ideal MRI' in cases where dissipation is sufficiently small to be negligible on large scales $\Lambda_{O,H,A} \gg 1$, or when no physical dissipation is considered explicitly in the numerical method (an approach referred to as 'implicit large Eddy simulations' (ILES)). Nevertheless, numerical dissipation is always actively playing a role in these models, which is not without consequences, as we shall show.

Configurations with a mean field are often characterised by the plasma β parameter of the mean field, as defined by

$$\beta_{\text{mean}} \equiv \frac{8\pi\langle P \rangle}{\langle B \rangle^2}. \tag{8.2}$$

This should not be confused with the turbulent plasma β parameter

$$\langle \beta \rangle \equiv \left\langle \frac{8\pi P}{B^2} \right\rangle, \tag{8.3}$$

which is not a control parameter of the physical system because usually $B \gg \langle B \rangle$.

In principle, the vertical box extension L_z does not necessarily match the pressure scale height $H = c_s/\Omega$ because vertical stratification is ignored in these models. However, in most of the simulations published today, it is the case, so that the two scales can be identified. This allows us to obtain an alternative expression for β_{mean}, which can be useful to interpret numerical simulations

$$\beta_{\text{mean}} \equiv \frac{2\Omega^2 L_z^2}{V_{Az}^2}. \tag{8.4}$$

It is this expression for β that is used in incompressible simulations.

8.1. *Ideal MHD*

8.1.1. *With a mean field*

The first historical models of MRI turbulence (Hawley *et al.* 1995) tested both mean vertical and toroidal fields and were computed in the ideal MHD framework. MRI unstable modes grow, break up presumably thanks to secondary instabilities that are essentially Kelvin–Helmholtz-type modes (Goodman & Xu 1994; Latter, Lesaffre & Balbus 2009) and end up in fully developed MHD turbulence.

In the mean vertical field case, Hawley *et al.* (1995) found that and $\alpha \simeq 6.7 \beta_{\text{mean}}^{-1/2}$. However, it was later pointed out by Bodo *et al.* (2008) that, in the presence of a mean vertical field, the box aspect ratio used by Hawley *et al.* (1995) was prone to recurrent channel mode solutions, which was absent in wider boxes (more elongated in x). This implies that α and $\langle \beta \rangle$ are overestimated by a factor of two compared with wider boxes (Bodo *et al.* 2008). We therefore correct (Hawley *et al.* 1995) for this and obtain the following scaling law in the range $400 < \beta_{\text{mean}} < 5 \times 10^4$:

$$\left.\begin{aligned} \alpha &\simeq 0.61 \langle \beta \rangle^{-1}, \\ \alpha &\simeq 3.3 \beta_{\text{mean}}^{-1/2}, \end{aligned}\right\} \tag{8.5}$$

a scaling which has also been verified in incompressible simulations (Longaretti & Lesur 2010). The extrapolation down to $\beta \to 1$ suggests that $\alpha \gtrsim 1$ could be reached. However, it is very difficult to explore this limit numerically because the MRI in this case becomes very strong and channel modes never break up into developed turbulence (Hawley *et al.* 1995). Lesur & Longaretti (2007) explored this limit in the incompressible regime and found that turbulence was in this case very intermittent with long 'quiet' episodes of linear growth followed by strong and sudden bursts of turbulence. In this situation, a constant α is probably not a good model of the disc physics.

In the mean toroidal field case, the scaling deduced from Hawley *et al.* (1995) in the range $2 < \beta_{\text{mean}} < 1200$ gives lower α values:

$$\left.\begin{aligned} \alpha &\simeq 0.51 \langle \beta \rangle^{-1}, \\ \alpha &\simeq 0.35 \beta_{\text{mean}}^{-1/2}. \end{aligned}\right\} \tag{8.6}$$

8.1.2. *Zero mean field: the 'MRI dynamo'*

The existence of an MRI-driven dynamo was first demonstrated by Hawley, Gammie & Balbus (1996). In this configuration, no external field is imposed and turbulence regenerates a field on which the MRI can grow, feeding back turbulent motions. Owing to this feedback loop, the system needs a finite-amplitude perturbation to sustain turbulence (Rincon *et al.* 2008), implying that the instability is in this case subcritical. The dynamo feedback loop has been the subject of intense studies after the discovery of dynamo cycles (Lesur & Ogilvie 2008*b*), both based on a quasi-linear theory of the toroidal MRI (Lesur & Ogilvie 2008*a*) and on a dynamical system approach (Herault *et al.* 2011; Riols *et al.* 2013).

According to Hawley *et al.* (1996), the MRI dynamo yields $\alpha \sim 0.01$. However, it was later realised that the value of α in ideal MHD simulations (where no physical dissipation is introduced) depends on the numerical resolution (Fromang & Papaloizou 2007), with $\alpha \propto 1/N$ where N is the number of grid points in one direction.[16] This problem of numerical convergence (because numerics do not seem to be converging as one increases the resolution) implies that the estimation for α from zero net flux simulation is intrinsically flawed. Obviously, this is most probably a result of numerical dissipation (the only dissipation channel of these simulations), which is not acting as a real physical dissipation operator. If one introduces explicit viscosity ν and resistivity η_O in the system, α then converges to a finite value, which seems to depend only on $Pm \equiv \nu / \eta_O$ (Fromang 2010) and no longer depend on the resolution. This, however, leads to another complication known as the '*Pm* effect'.

---

[16]The scaling of α on N actually depends on the order of the spatial reconstruction scheme as shown by Bodo *et al.* (2011).

8.1.3. *The Pm effect*

The *Pm* effect shows up in quasi-ideal simulations both with and without a mean vertical field. When one introduces only a small amount of viscosity ν and resistivity η_O, the large-scale linear MRI modes are largely unaffected. For this reason, this regime corresponds to a 'quasi ideal-MHD' regime. However, it turns out that the saturation level depends on the magnetic Prandtl number. A lot of literature has been devoted to this effect (Fromang *et al.* 2007; Lesur & Longaretti 2007). In simulations with a mean field (Lesur & Longaretti 2007; Simon & Hawley 2009), the *Pm* effect results in an increase of α when *Pm* increases. In the zero mean field case, the effect is even stronger because the MRI dynamo, and therefore MHD turbulence, simply disappears when $Pm \lesssim 1$ (Fromang *et al.* 2007; Walker & Boldyrev 2017), giving $\alpha = 0$ in that regime.

This effect shows up even in situations where the Reynolds and Elsasser numbers are much larger than one, indicating that linear or quasi-linear theory cannot explain it (Longaretti & Lesur 2010). Instead, it has been proposed that non-local energy transfers in spectral space (Lesur & Longaretti 2011) could be responsible for this effect. If true, because non-locality is necessarily bounded (Aluie & Eyink 2010), the *Pm* effect should disappear as $(\eta_O, \nu) \to 0$. An alternative viewpoint is to separate the small 'turbulent' scales from the large scales where the dynamo mechanism is presumably lying. By carefully computing the energy exchange between the scales, one finds that the small scales act as an effective viscosity when $Pm < 1$ and tend to damp the large-scale mechanism (Riols *et al.* 2013). This, however, is not fully satisfactory as it does not prove that the MRI dynamo is non-existent in the limit $Pm \to 0$ and $Rm \to \infty$.

Despite numerous efforts, recent results indicate that the *Pm* effect is still very much alive in simulations, up to $Rm = 5 \times 10^4$ and $\Lambda_O = O(100)$ (Potter & Balbus 2017). This effect is still a very open question regarding the MRI saturation level in the 'quasi-ideal' regime.

Note, finally, that the *Pm* effect is relevant only for the innermost parts of PPDs where $Rm \gg 1$ is expected. For the regions above 1 AU, large-scale physics is dominated by magnetic diffusion and the *Pm* effect is not relevant.

8.2. *Ohmic diffusion*

In non-stratified shearing box simulations with a net vertical flux, the presence of at least one MRI unstable mode is given by (6.50). Using the expression for β_{mean}, and assuming that $k_{z,\mathrm{min}} = 2\pi/L_z$, the stability condition of a non-stratified shearing box reads

$$Rm > \frac{\beta_{\mathrm{mean}}}{\sqrt{\dfrac{3}{2\pi^2}\beta_{\mathrm{mean}} - 4}} \to \text{Instability.} \qquad (8.7)$$

The effect of a strong Ohmic diffusion on the non-linear saturation of the MRI was first explored by Sano, Inutsuka & Miyama (1998) in two dimensions and Fleming, Stone & Hawley (2000) in three dimensions. In the case with a mean field, it is found that when $\Lambda_O \lesssim 1$ (or, equivalently, $Rm \lesssim \beta_{\mathrm{mean}}$) and (8.7) is verified so that the system is MRI-unstable, MRI turbulence is affected: α becomes lower than in the ideal case, and turbulence becomes intermittent, with periods of linear growth followed by rapid decay owing to reconnection events. As in the low-β ideal case, it is unclear whether this regime can be modelled with a constant α coefficient.

In the zero mean field case, Fleming *et al.* (2000) found that the MRI was disappearing below a critical Reynolds number $10^3 < Rm_c < 10^4$. Whether Rm_c depends on resolution, as the saturation level of the MRI dynamo does, is an open question.
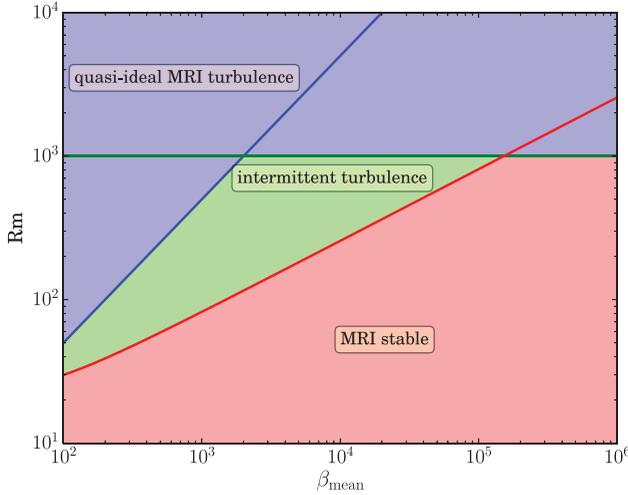
FIGURE 28. MRI turbulence regions as a function of *Rm* and β_{mean} in the Ohmic diffusion case. The red line corresponds to the linear stability criterion (8.7), the blue line to the limit $\Lambda_O = 1$ and the green line to the zero net field limit $Rm_c = 10^3$.

One can combine these results into a global map showing how turbulence saturates in simulations with and without a net flux (figure 28). We have assumed that in the limit $\beta_{\text{mean}} \to \infty$, the zero net flux subcritical threshold was to be considered instead of the linear stability limit. As a result, the region $\beta_{\text{mean}} > 10^5$ follows the zero mean field criterion, whereas $\beta_{\text{mean}} < 10^5$ shows a transition region between fully developed turbulence with $\Lambda_O > 1$ and a linearly stable flow, which we have named intermittent turbulence, in reference to Fleming *et al.* (2000). As the data is sparse, we do not have any clear estimate of the values α in the intermittent turbulence region.

8.3. *Ambipolar diffusion*

The role played by ambipolar diffusion on the saturation level of the MRI was first studied by Hawley & Stone (1998) in the two-fluid limit. They found that MRI turbulence was unaffected by ambipolar diffusion for $\Lambda_A \gtrsim 100$. Because PPDs are strongly collisional, the two-fluid approach is not very efficient because numerical time steps are limited by the collision timescale. For this reason, this problem was revisited using the single-fluid approach by Bai & Stone (2011). They found that MRI is progressively suppressed by ambipolar diffusion as Λ_A decreases.

Starting from Bai & Stone (2011) results, it is possible to create a simple phenomenology for MRI saturation under the effect of ambipolar diffusion. First, let us recall from (6.62) that the linear stability criterion with $k_{z,\text{min}} = 2\pi/L_z$ is given by

$$\left.\begin{array}{c} \Lambda_A > \Lambda_{A,\text{crit}} \dfrac{V_A}{V_{Az}} \to \text{Instability,} \\[2ex] \text{with} \quad \Lambda_{A,\text{crit}} \equiv \dfrac{1}{\sqrt{\dfrac{3}{8\pi^2}\beta_{\text{mean}} - 1}}. \end{array}\right\} \quad (8.8)$$

The linear stability of a box with a pure vertical field is therefore very similar to the Ohmic case. However, as the instability grows, $\langle V_A \rangle$ increases, which eventually leads
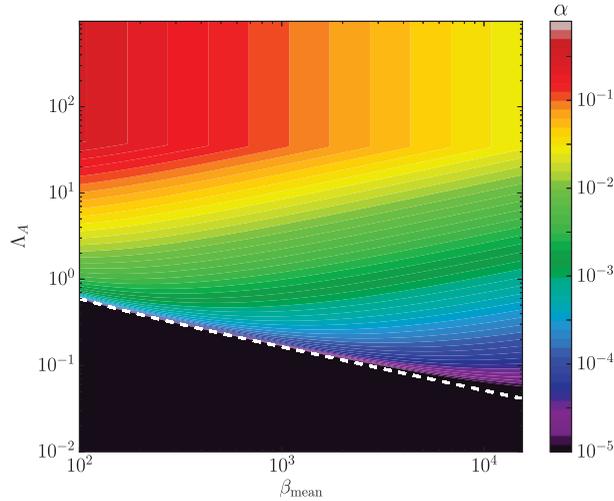
FIGURE 29. MRI turbulent transport deduced from (8.11) and (8.5) in the ambipolar-dominated regime with a mean vertical field. These estimates match the numerical values of Bai & Stone (2011) at $\pm 50\,\%$. The white dashed line corresponds to the marginal stability limit $\Lambda_A = \Lambda_{A,\text{crit}}$. The region below this line has $\alpha = 0$ in the net vertical field case, but can reach $\alpha \sim 10^{-4}$ when a mean toroidal field component is introduced, thanks to the presence of unstable oblique modes (see the text).

to the violation of the stability criterion. This self-suppression effect of MRI-turbulence allows us to deduce the saturation level by assuming that at saturation

$$\Lambda_A = \Lambda_{A,\text{crit}} \left(\frac{V_A}{V_{Az}} \right)_{\text{sat}}. \tag{8.9}$$

The ratio of Alfvén speeds can be evaluated by $\langle \beta \rangle$:

$$\left(\frac{V_A}{V_{Az}} \right)_{\text{sat}} \simeq \left(\frac{\beta_{\text{mean}}}{\langle \beta \rangle} + 1 \right)^{\delta}. \tag{8.10}$$

In principle, one would naïvely expect $\delta = 0.5$. However, we find that this estimate does not fit the numerical results of Bai & Stone (2011). This is because the large-scale field V_A is not necessarily proportional to the instantaneous $\langle \beta \rangle$ computed from the fluctuations at all scales. Instead, we therefore choose $\delta = 1$, which presents a better correlation to the available data. Noting that, similarly to the ideal case of Hawley *et al.* (1995), Bai & Stone (2011) found that $\langle \beta \rangle \simeq (2\alpha)^{-1}$, we obtain a simple estimate for α by combining the previous expression:

$$\alpha \simeq \frac{1}{2\beta_{\text{mean}}} \left(\frac{\Lambda_A}{\Lambda_{A,\text{crit}}} - 1 \right). \tag{8.11}$$

Of course, this estimates diverges as $\Lambda_A \to \infty$ because the saturation mechanism owing to ambipolar diffusion becomes non-existent. In this case, the ideal-MHD estimate (8.5) should be used instead. A sample of predicted α from this saturation estimation is given in figure 29. These estimates are within $50\,\%$ of the calculated value with a pure vertical field of Bai & Stone (2011) and can be used as a proxy to estimate the transport in ambipolar-dominated discs. This estimate also successfully recovers the ideal regime when $\Lambda_A \gtrsim 50$, as reported by numerical simulations.
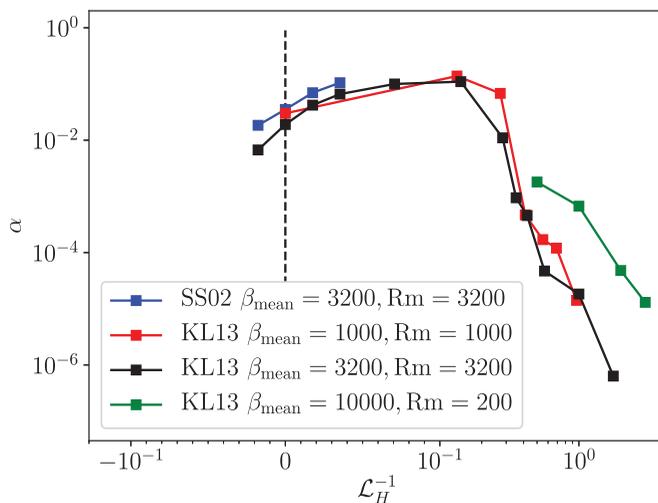
FIGURE 30. Evolution of the turbulent transport as a function of the intensity of the Hall effect in the mean vertical field case. Data from Sano & Stone (2002) (SS02) and Kunz & Lesur (2013) (KL13). Here $\mathcal{L}_H < 0$ corresponds to anti-aligned field configuration. Note that the KL13 $\beta_{\text{mean}} = 10^4$ case is linearly stable for $\mathcal{L}_H^{-1} = 0$ because of Ohmic diffusion, and exemplify the reactivation of the linear MRI under the action of Hall (see § 6.4.4). Note that the α values from Sano & Stone (2002) have been renormalised to match the definition of α in Kunz & Lesur (2013).

In the case of a pure azimuthal field, Bai & Stone (2011) have shown that no turbulence is sustained below $\Lambda_A \lesssim 3$, and α progressively drops to 0 from the ideal MHD value at $\Lambda_A = 100$.

Finally, in the case of a mixed mean vertical and azimuthal field, the subsistence of oblique modes (see § 6.4.5) even at $\Lambda_A \lesssim 0.1$ creates a weak transport with $\alpha \sim 3 \times 10^{-4}$ for $V_{Ay} \sim V_{Az}$ (Bai & Stone 2011). Above $\Lambda_A \sim 1$, the scaling (8.11) is approximately recovered as oblique modes become unimportant.

8.4. *Hall effect*

The effect of the Hall effect on the saturation level of MRI turbulence was first explored by Sano & Stone (2002) with a relatively weak Hall effect ($\mathcal{L}_H \gtrsim 20$). They found that with a mean vertical field, α increases with decreasing $\mathcal{L}_H$ in the aligned case, whereas α decreases in the anti-aligned case. However, PPDs are likely to have lower $\mathcal{L}_H$ than those studied by Sano & Stone (2002) (typically $\mathcal{L}_H \sim 1$ as in figure 10). For this reason, this problem was revisited by Kunz & Lesur (2013) with simulations in the $\mathcal{L}_H \sim 1$ regime. In the case with a mean vertical field aligned with the rotation axis, it is found that despite being violently unstable from the linear point of view owing to the HSI, the flow settles down into a quasi-laminar state for $\mathcal{L}_H \lesssim 5$, with negligible turbulent transport (figure 30). This 'low transport state' is characterised by a self-organised flow where the vertical field is concentrated in a narrow region in the x direction (which can be identified as a ring in global geometry). The mechanism behind self-organisation in Hall-MHD is detailed in § 8.5.

In the zero net-flux case, Sano & Stone (2002) found evidence that Rm_c decreases from around 10^4 in the ideal MHD regime possibly down to a few times 10^3 for $\mathcal{L}_H \simeq 20$. At the same time, α increases by a factor of a few compared with the ideal case. For stronger Hall effects ($\mathcal{L}_H \lesssim 5$), a low transport state similar to the case with net flux is observed.
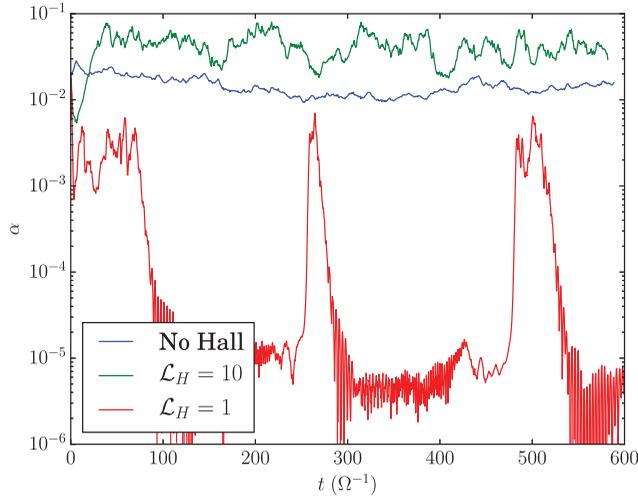
FIGURE 31. Turbulent transport α as a function of time in zero net flux simulations including the Hall effect. For a weak Hall effect ($\mathcal{L}_H \gtrsim 10$), α is larger than in the case without Hall effect by a factor ~ 3. When the Hall effect increases, the system enters the low transport state ($\mathcal{L}_H \lesssim 5$). In contrast to the net field case, it periodically switches back to a high transport state, resulting in bursts of α.

However, in this case, the system switches back to a turbulent state periodically, resulting in short bursts of turbulence in the system (figure 31).

Overall, it is found that despite a powerful large-scale instability, the MRI in the Hall-dominated regime does not result into an efficient turbulent transport of angular momentum in unstratified boxes. This surprising result, however, does not hold in stratified boxes, where the Hall effect effectively leads to an enhanced radial stress (see § 9.2.2). Hence, the low transport state of Kunz & Lesur (2013) is really a peculiarity of the unstratified setup.

8.5. *Self-organisation*

Self-organisation is a process by which a disorganised (i.e. turbulent) flow creates large-scale and long-lived structures. There are several examples of self-organisation in nature, the best-known being probably the great red spot of Jupiter, resulting from small-scale turbulent motions which cascade to large scales forming a giant anticyclone. Self-organisation is a spontaneous symmetry-breaking process: the system starts from a statistically homogenous state and ends up in a heterogeneous state with well-identified structures. As such, self-organisation should be distinguished from local instabilities such as Kelvin–Helmholtz or the Rossby wave instability (RWI; Lovelace *et al.* 1999) that result from a special location in the flow (e.g. vortensity extremum).

Although self-organisation was clearly pointed out as a key phenomenon in Hall-MHD by Kunz & Lesur (2013), this phenomenon (or a weaker version of it) has been observed in ideal MHD simulations of MRI turbulence by several authors since early 2000. It is also a very promising mechanism to explain some of the structures observed in the sub-millimetric range (§ 1.4). Let us overview the different mechanisms that have been proposed to explain this phenomenon.

8.5.1. *Ideal MHD*

Hawley (2001) and Steinacker & Papaloizou (2002) were among the first to note the formation of 'ring-like' structures in MRI simulations with net vertical flux.

The simulations are in these cases semi-global: vertical stratification is neglected whereas the radial curvature is retained, leading to a cylindrical setup. It is found that the net vertical flux is trapped in a low-density region, forming a gap. In these gaps, α can reach values as high as 1, consistently with the fact that these gaps correspond to low β_{mean} regions. Hawley (2001) proposed that this could be the signature of a viscous instability: a local density minimum results in a local decrease of β_{mean} (assuming the mean field is kept at its initial value). As $\alpha \propto \beta_{mean}^{-1/2}$ (see (8.5)), α increases in this region, which removes mass from the region because of angular momentum conservation. Unfortunately, this proposition has never been investigated further, leaving the origin of self-organisation in these simulations unexplained.

The same phenomenon was reported in local shearing box models with a mean vertical field by Bai & Stone (2014). They proposed that the non-diagonal components of the turbulent resistivity tensor (Lesur & Longaretti 2009) could be at the origin of the effect by acting as an 'effective negative diffusivity'. This explanation is, however, dubious because several authors have measured the turbulent resistivity tensor of MRI turbulence (Fromang & Stone 2009; Guan & Gammie 2009; Lesur & Longaretti 2009) and found that the effective resistivity was always positive. Strikingly, Bai & Stone (2014) simulations clearly show that despite having a globally concentrated field $B_z(x)$, the turbulent electromotive force $\mathcal{E} = \langle w \times B \rangle$ does not depend on x (Bai & Stone 2014, figure 3). Hence, the turbulent resistivity prescription $\mathcal{E} = \eta_{turb} \nabla \times \langle B \rangle$ likely breaks down altogether and should be replaced with a more elaborate closure scheme.

In simulations without net flux, Fromang & Nelson (2005) reported the spontaneous formation of giant anticyclones, though this could be a boundary condition artefact (Fromang, private communication). Johansen, Youdin & Klahr (2009) also reported the formation of 'pressure bumps' that are caused by large-scale fluctuations of α. In contrast to the simulations with a mean field vertical field, the features observed in the zero net field case are transient and only survive for a limited time, which depends on the box size. Johansen *et al.* (2009) proposed a model based on a stochastic α, which predicts long-lived axisymmetric structures in quasi-geostrophic equilibrium, as observed in their simulations.

8.5.2. *Hall-MHD*

The first mention of self-organisation in Hall-MHD appears in Kunz & Lesur (2013), where self-organisation has a dramatic effect on the saturation level of Hall-dominated MRI (see § 8.4). Self-organisation appears in an obvious manner by looking at the vertical field component of the flow (figure 32). Although in the ideal-MHD case, self-organisation appears as a 'second-order' effect on top of MRI turbulence, in the Hall-MHD case, self-organisation is the main saturation mechanism of the MRI. In other words, turbulence is mostly suppressed by self-organisation. Self-organisation shows up when $\mathcal{L}_H \lesssim 5$, and α essentially vanishes as a result of the lack of turbulence (figure 30). This surprising result also holds in the case with zero net flux or in the mixed case having both a mean azimuthal and vertical magnetic flux.

The origin of Hall-driven self-organisation can be tracked down to the induction equation in the presence of a Hall effect. Indeed, When the Hall length is constant, the induction equation reads

$$\partial_t B = \nabla \times (w \times B) + \ell_H \nabla \times \left(\nabla \cdot \frac{-BB}{4\pi} \right), \tag{8.12}$$

which highlights the role of the Maxwell stress in the induction equation. Guided by numerical simulations that show the appearance of a vertical magnetic field with variations
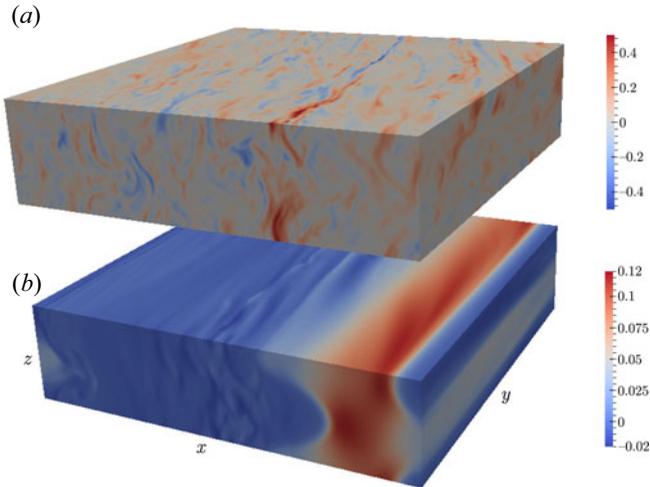
FIGURE 32. Vertical field component in snapshots of MRI turbulence. (*a*) Ideal-MHD simulation with $\beta_{\mathrm{mean}} = 3200$. (*b*) Hall-MRI simulation with $\beta_{\mathrm{mean}} = 3200$ and $\mathcal{L}_H = 1.75$. Figure from Kunz & Lesur (2013).

in the *x* direction, let us define an average

$$\bar{Q} = \iint \mathrm{d}y\,\mathrm{d}z Q \tag{8.13}$$

so that the induction equation for the 'mean' vertical field reads

$$\partial_t \overline{B_z} = \partial_x \left(\overline{w_z B_x - w_x B_z} \right) + \ell_H \partial_x^2 \frac{-\overline{B_x B_y}}{4\pi}, \tag{8.14}$$

where we recognised the radial Maxwell stress term $\mathcal{M}_{xy} = -B_x B_y / 4\pi$, also present in the angular momentum conservation equation (4.21). This demonstrates that in Hall-MHD, the transport of magnetic flux is tightly linked to the transport of mass. Owing to energetic constraints, $\mathcal{M}_{xy} > 0$ (see §5.2.4) in shear-driven instabilities/turbulence. Therefore, a concentration of magnetic field owing to the Hall effect is possible at local stress minimum. In a turbulent flow, short-lived stress minima occur randomly in the flow, and these minima tend to accumulate vertical magnetic flux according to the previous equation. When the Hall effect is strong enough, a local minimum can accumulate enough flux to become stable for the HSI. In this case, the flow becomes stable and the local turbulent stress vanishes, becoming a permanent minimum. This minimum continues to accumulate magnetic flux thanks to the remaining stress present on both sides until the flux outside of the minimum of stress becomes negative. At this point, the stress also vanishes in the regions $\overline{B_z} < 0$ and the systems settles down into a quasi-stationary state with very low-stress level (see also figure 33).

Although this process was initially identified in unstratified shearing box simulations, it was later unambiguously identified in cylindrical unstratified simulations (Béthune, Lesur & Ferreira 2016). However, vertically stratified simulations do not seem to exhibit this process, for reasons not yet identified to date (Lesur *et al.* 2014).

8.5.3. *Ambipolar diffusion*

Self-organisation owing to ambipolar diffusion was first mentioned by Bai & Stone (2014) in unstratified shearing boxes. It was also observed in cylindrical simulations
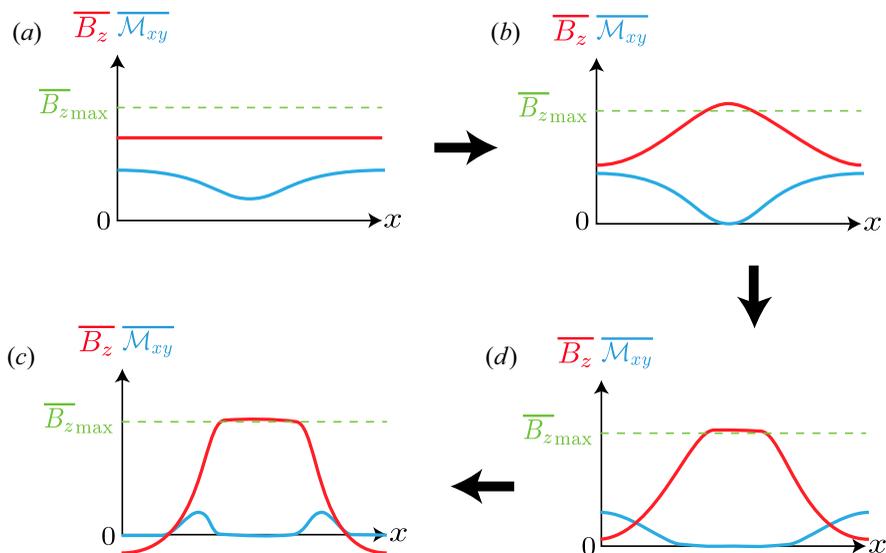
FIGURE 33. Hall self-organisation phenomenology. We start from a local fluctuation of the stress $\mathcal{M}_{xy}$ (*a*). This minimum creates a local maximum of $\overline{B_z}$, which overshoot the maximum $\overline{B_z}$ allowed by the HSI. Because the flow becomes locally stable, the stress vanishes (*b*). On the boundaries of this stable region, the Maxwell stress still transport magnetic flux towards the stable region, making it larger and emptying the rest of the domain (*d*). At some point, the total flux in the rest of the domain becomes negative, and it becomes HSI-stable. The stress therefore vanishes in this region as well, leaving only the interface with a minimal stress (*c*). Figure inspired by Kunz & Lesur (2013).

including ambipolar diffusion by Béthune *et al.* (2016), albeit at a very low level. However, whether ambipolar diffusion plays an active role in the self-organisation mechanism is an open question. Clearly, the 'strength' of self-organisation (quantified by the ratio $B_{z\max}/B_{z\min}$) in the ideal-MHD case is larger than the ambipolar diffusion case by a factor 10 (figure 2 in Bai & Stone 2014). In addition, the mechanism proposed for self-organisation only involves ideal-MHD terms (Bai & Stone 2014). Finally, the existence of zonal flows in this configuration (non-stratified, ambipolar diffusion dominated) largely depends on the box aspect ratio. Bai & Stone (2014) largely explored the situation with $L_x = L_y = 4H$. However, a choice of box with $L_y > L_x$ tends to break zonal flows (figure 34). Overall, it is very possible that self-organisation in ambipolar-dominated unstratified shearing boxes is a mere numerical artefact.

9. Stratified shearing boxes

Stratified models have been mostly explored in the zero mean field configuration. When a mean vertical field is included, simulations lead to 'high magnetic pressures that disrupt the vertical structure of the disk before the flow makes the transition to MHD turbulence' (quoting Stone *et al.* 1996 p. 659, § 3). Although the situation is less dramatic when more adapted boundary conditions are used (see below), this statement explains the lack of simulations with a net vertical field until recently. The situation with a mean toroidal field is less interesting since the toroidal component can usually escape through the vertical boundaries, eventually leading to a situation identical to that without a mean field.
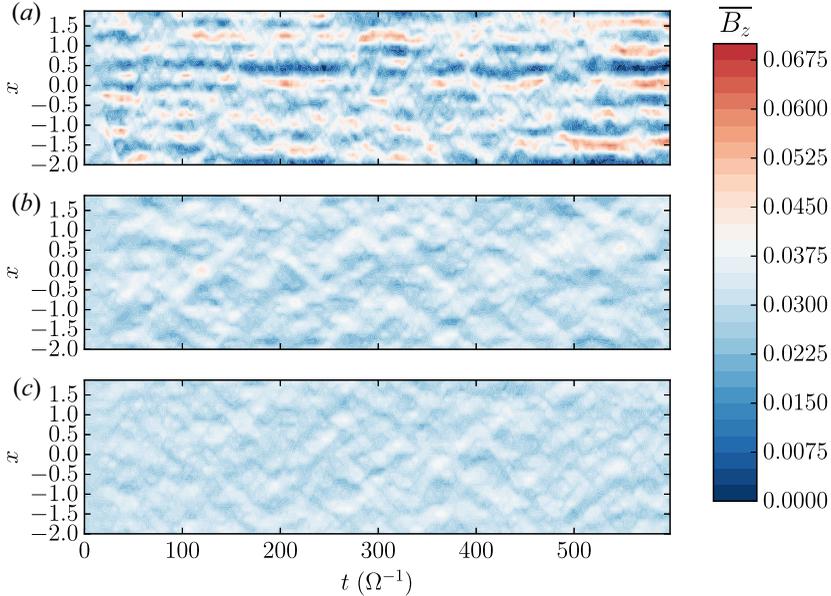
FIGURE 34. The y–z average of B_z as a function of time for non-stratified MRI simulations with $\Lambda_A = 3$ and $\beta_{\text{mean}} = 1000$: (a) $L_x = L_y = 4H$; (b) $L_x = 4H$, $L_y = 8H$; and (c) $L_x = 4H$, $L_y = 16H$. Note the disparity in zonal flows when $L_y > L_x$. All three simulations have been computed with a fixed resolution per scale height in the three spatial directions $n_{x,y,z} = 64$ pts/H using the Snoopy code (Lesur & Longaretti 2007).

As in the non-stratified case, we define the plasma β parameter of the mean vertical field threading the disc

$$\beta_{\text{mean}} \equiv \frac{8\pi \langle P \rangle_{z=0}}{\langle B_z \rangle_{z=0}^2}, \tag{9.1}$$

where averages are taken in the disc midplane. In the following, we use a definition of α using the box averaged pressure, i.e.

$$\alpha \equiv \frac{1}{\langle P \rangle} \left\langle \rho w_x w_y - \frac{B_x B_y}{4\pi} \right\rangle, \tag{9.2}$$

which does not depend on the vertical extension of the simulation domain (provided that all of the significant stress is contained in the box). However, some authors such as Stone and coworkers usually defines α from the vertically averaged stress and the midplane pressure $P_0 \equiv P(z = 0)$, which leads to predicted α values two or three times smaller than those obtained with the previous definition (Davis, Stone & Pessah 2010). Note also that Stone's definition yields α, which decreases as the box size increases. Indeed, the stress being concentrated in the region $z \lesssim 2H$, the vertically averaged stress decreases as the box size increases above $2H$. These differences in the definition of α should be kept in mind when comparing results from different groups.

9.1. *Zero mean field*

9.1.1. *Ideal MHD*

The zero mean field stratified case was first explored by Brandenburg *et al.* (1995) and Stone *et al.* (1996). They first noted the spontaneous appearance of a 'butterfly' diagram when looking at the space–time behaviour of the x–y average toroidal magnetic field
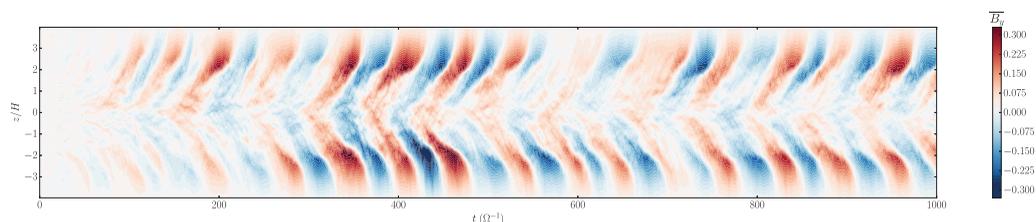
FIGURE 35. The x–y average of B_y ($\overline{B_y}$) as a function of time and z for zero vertical net flux stratified MRI simulations. Note the presence of a butterfly pattern indicating a periodic reversal of the toroidal field followed by its increase away from the disc midplane.

(figure 35; see also Davis *et al.* 2010 and Simon, Beckwith & Armitage 2012). This diagram shows quasi-periodic flip of the toroidal field, with a periodicity close to 10 local orbital periods. Whether this butterfly diagram leads to observational counterpart by modulating the turbulent transport is still debated (see, e.g., Hogg & Reynolds 2016 for an example). On the theoretical side, it is well reproduced by mean field models (Gressel 2010; Gressel & Pessah 2015), but we are still lacking a first principle theory for this dynamo.

In principle, once vertical stratification is included, one needs to specify both the vertical density and temperature profile as well as an equation of state. However, most of the literature has focused on isothermal simulations, in which the temperature is constant. In this case, it is found that

$$\alpha \simeq 0.02 \pm 0.01 \quad \text{(Stratified, isothermal, zero vertical flux)} \tag{9.3}$$

(e.g. Stone *et al.* 1996; Davis *et al.* 2010). In the outer region of PPDs, radiative transfer calculations tell us that the vertical temperature profile is approximately constant, so the isothermal approximation is probably a good model in this case.

Nevertheless, some researchers have explored non-isothermal models, such as Hirose and collaborators (Hirose, Krolik & Stone 2006; Hirose, Blaes & Krolik 2009; Hirose *et al.* 2014), Flaig and collaborators (Flaig, Kley & Kissmann 2010; Flaig *et al.* 2012) and Bodo and collaborators (Bodo *et al.* 2013, 2015). It is found that when the vertical profile becomes unstable for convection (i.e. it violates the Schwarzschild criterion), the turbulent transport of angular momentum can increase by up to an order of magnitude (Hirose *et al.* 2014), thanks to a mechanism which is yet to be fully elucidated.

Following the numerical convergence issue pointed out by Fromang & Papaloizou (2007) in unstratified simulations (see § 8.1), numerical convergence in stratified setups has also been tested. Initial explorations were limited by numerical resources to 128 points per scale height (Davis *et al.* 2010) and showed the convergence of α as a function of the number of grid points, which led to the conclusion that 'stratification saves the day'. However, this issue was tackled again with higher-resolution simulations: 200 points per scale height (Bodo *et al.* 2014) and 256 points per scale height (Ryan *et al.* 2017). These recent results show a weak dependence on the resolution with $\alpha \propto N^{-1/3}$ at the resolution explored, albeit with a different vertical boundary condition (Davis *et al.* (2010) used periodic boundary conditions while Ryan *et al.* (2017) used outflow boundary conditions). Overall, these results indicate that numerical convergence is also an issue in stratified models.

9.1.2. *Non-ideal MHD*

Realistic Ohmic diffusion profiles were introduced in zero net flux stratified simulations for the first time by Fleming & Stone (2003). These simulations exemplified (Gammie

1996) layered accretion model with a mid-plane dead zone and an active surface layer. The debate then crystallised on the thickness of the active layer, which, depending on the model, could lead to accretion rates compatible with observations.

This layered accretion paradigm, however, omits ambipolar diffusion and the Hall effect. Ambipolar diffusion is particularly important at low densities, right in the active layers of Gammie (1996). It was quickly realised that ambipolar diffusion could dramatically reduce the strength of MRI turbulence in this layer, and even suppress it, because $\Lambda_A \lesssim 1$ in this region (Perez-Becker & Chiang 2011*a*,*b*; Dzyurkevich *et al.* 2013). This is confirmed by direct numerical simulations including Ohmic and ambipolar diffusion but no Hall effect (Bai & Stone 2013*b*; Simon *et al.* 2013*b*). These simulations suggest dramatically low values for α, with $\alpha \simeq 3 \times 10^{-6}$ at 1 AU (Bai & Stone 2013*b*) to $\alpha \simeq 10^{-3}$ at 100 AU (Simon *et al.* 2013*b*).

These results imply accretion rates $\dot{M} \lesssim 10^{-10} M_\odot/\text{yr}$ in the region 1–30 AU, which are too low by at least two orders of magnitude compared with observational constraints. However, the zero mean field configuration is rather artificial. Indeed, PPDs are expected to form from a collapsing cloud that has dragged some fraction of its initial magnetic flux. Therefore, a non-negligible poloidal flux is likely to be present in these objects, which motivated the need for simulations including a mean poloidal field.

9.2. *Mean field and outflows*

9.2.1. *Ideal MHD*

The first viable shearing-box simulation of a vertically stratified disc with a mean vertical field was performed by Suzuki & Inutsuka (2009) and Suzuki, Muto & Inutsuka (2010), thanks to more robust numerical techniques compared with the initial attempts of Stone *et al.* (1996) and carefully designed vertical boundary conditions. Their simulations have a relatively weak field $\beta_{\text{mean}} \gtrsim 10^4$, but show significant deviations from the zero mean field scenario and, in particular, the presence of a strong outflow, quantified by the mass flux leaving the disc. In a box with $L_z = 8H$, they find a vertical mass flux 100 times larger with $\beta_{\text{mean}} = 10^4$ compared with the zero net field case (Suzuki & Inutsuka 2009).

This problem was revisited by a large number of authors, both in the strong field limit $\beta_{\text{mean}} \simeq 1$ (Moll 2012; Ogilvie 2012; Lesur, Ferreira & Ogilvie 2013, figure 36), and in the intermediate regime $\beta_{\text{mean}} = 10^3$–$10^4$ (Fromang *et al.* 2013; Bai & Stone 2013*a*). It was quickly demonstrated that these outflows were a variation of Blandford & Payne (1982) magneto-centrifugal paradigm (Fromang *et al.* 2013; Lesur *et al.* 2013) with a strong time-dependency. Despite this connection to well-known launching mechanisms, some properties of shearing box outflows are intrinsically flawed.

First, the mass loss rate depends on the vertical box extension (Fromang *et al.* 2013), taller boxes leading to lower mass loss rates. This surprising result is actually expected from the shape of the gravitational potential in a shearing box (§ 5.1), which is unbounded when $z \rightarrow \infty$. Some authors have proposed to fix this issue by adding higher-order terms to the vertical gravity force, making it possible to escape at $z = \pm\infty$ with a finite amount of mechanical energy (Suzuki *et al.* 2010; McNally & Pessah 2015). This approach, however, violates the conservative nature of gravitation, because the resulting force does not derive from a potential anymore. A more rational approach would be to include all of the third-order terms in Hill's approximation. This, however, also introduces radial curvature terms, making the shear-periodic boundary conditions unadapted.

Second, the geometry of the outflow is problematic. In principle, shearing boxes allow both vertically even and odd magnetic configuration (figure 37). At the linear level (see § 6.4.6), the two symmetries show similar growth rates and properties. However, in the non-linear regime, the odd symmetry is problematic for the physical interpretation of
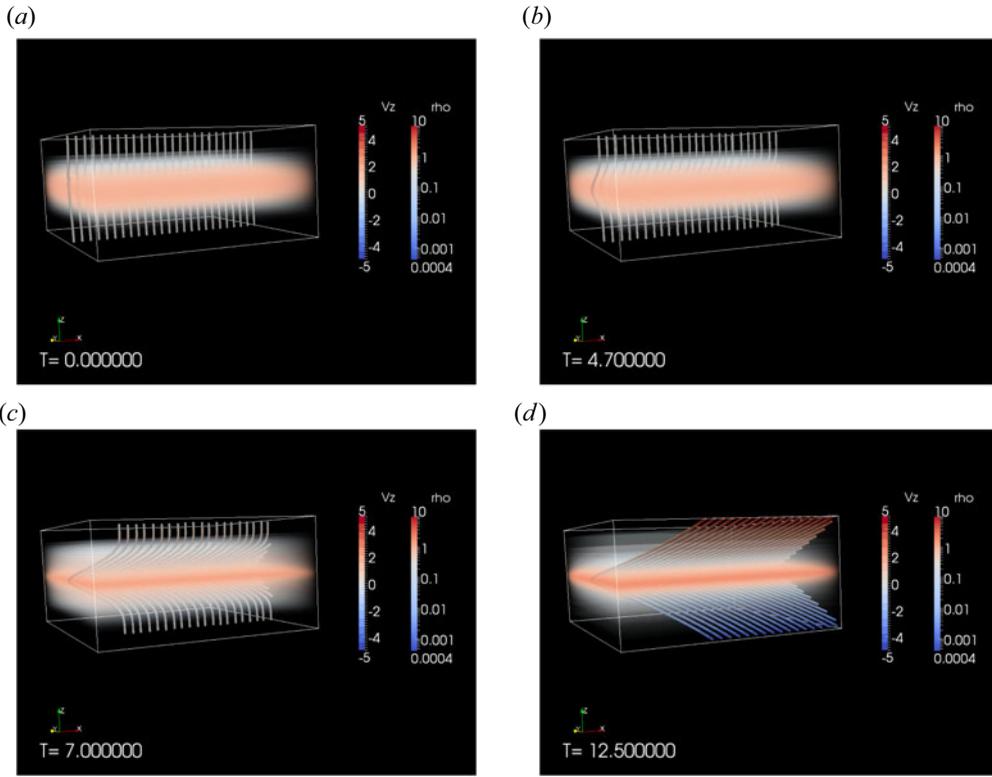
FIGURE 36. Growth and saturation of the MRI in the presence of a strong vertical field ($\beta_{\mathrm{mean}} = 10$) in ideal MHD. Tubes are magnetic field lines whereas gas density is represented with volume rendering. Note the initial growth of the linear mode in the disc midplane, which eventually saturates into a quasi-laminar outflow configuration similar to Blandford & Payne (1982) paradigm. From Lesur *et al.* (2013).
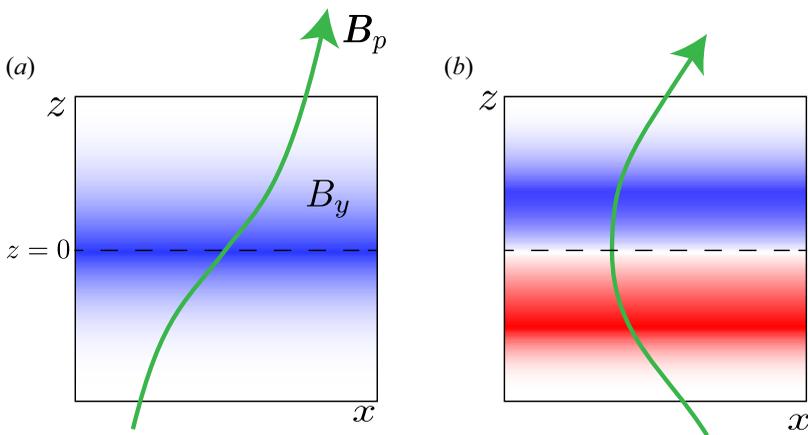
FIGURE 37. Symmetries of the outflow configuration allowed in a shearing box. Poloidal field lines are represented in green whereas the toroidal field component $B_y < 0$ is shown in blue and $B_y > 0$ in red: (*a*) odd symmetry configuration; (*b*) even symmetry configuration. The usual (Blandford & Payne 1982) picture corresponds to the even symmetry in a shearing box model.

the outflow. By simply looking at the poloidal field topology (figure 37), it is clear that the odd symmetry connects the bottom side of the disc to the central star, whereas the top side is connected to $R \to +\infty$. Physically, it means that no angular momentum is actually extracted from the disc: angular momentum comes from an unknown source at $z \to -\infty$ and flows through the disc up to $z \to +\infty$. The angular momentum conservation equation (4.21) clearly describes this situation: in the odd symmetry case, one has $B_y B_z(z) = B_y B_z(-z)$, so the surface stresses on both sides of the disc cancel out and the disc is not accreting any material.

Naturally, this situation is rather unrealistic and the only physical configuration is the even symmetry one (or at least, a configuration where $B_y B_z$ have a different sign on both sides of the disc). However, it turns out that shearing boxes tend to settle into the odd configuration when $\beta_{\text{mean}} \lesssim 10^3$ (Salvesen *et al.* 2016). The reason for this unexpected result is still debated. It is certainly related to the fact that the shearing box has 'too many symmetries', and does not differentiate $x \to \pm\infty$. Some authors have enforced the even symmetry manually (Lesur *et al.* 2013), which leads to physical outflow configurations. However, this is not satisfactory as the flow symmetry should be enforced by the global geometry of the disc and its surrounding, which is not captured in shearing box models.

Despite these difficulties with outflows, it is possible to measure the turbulent stress in these simulation. A systematic exploration of shearing box models with $\beta_{\text{mean}} \in [10, \infty]$ (Salvesen *et al.* 2016) shows that

$$\alpha = 10.1 \beta_{\text{mean}}^{-0.53} \quad \text{for } \beta_{\text{mean}} < 10^5 \tag{9.4}$$

with α recovering the zero net flux value when $\beta_{\text{mean}} > 10^5$. This scaling leads to α values about 3 times larger than the unstratified estimate (§ 8.1). As in the unstratified case, the level of turbulent stress depends on the magnetic Prandtl number, α increasing with increasing *Pm* (Fromang *et al.* 2013), so caution is still needed when using these scalings in phenomenological models.

The characterisation of winds in a shearing box is always a bit difficult because of the symmetry issues noted previously. In addition, the usual MHD outflow invariants (§ 11) are only defined in global geometry, making them inaccessible in local models (but see Lesur *et al.* (2013) for local equivalents of the global invariants). Despite these difficulties, one usually defines an outflow rate ζ and a torque parameter υ, which measure the mass and angular momentum evacuated by the magnetised wind, respectively.

It is customary to define the outflow rate as

$$\zeta \equiv \frac{\overline{\rho w_z}|_{\text{top}} - \overline{\rho w_z}|_{\text{bottom}}}{\rho_{\text{mid}} c_s}, \tag{9.5}$$

where 'top' and 'bottom' subscripts denote the top and bottom of the disc (which, in principle, can be chosen freely) and — denotes a horizontal averaging procedure. In a real (global) disc, the local domain is emptied by the wind, but it is also replenished by the divergence of the accretion flow, so that a steady state is locally achieve. As there is technically no accretion flow in shearing boxes, this replenishment is absent, leading to boxes that slowly lose mass. The typical timescale over which a box loses a significant fraction of it mass is given by $\tau_{\text{loss}} \sim (\zeta \Omega)^{-1}$. This implies that the shearing box model is really valid in the limit $\zeta \ll 1$.

The outflow rate is known to depend on several key parameters. First, it should be noted that the outflow rate depends on the box extension, both horizontal and vertical. In the horizontal plane, it seems that convergence is reached for $L_x, L_y \gtrsim 4H$

(Fromang *et al.* 2013). In the vertical direction, however, as the box gets higher, the outflow rate decreases. Fromang *et al.* (2013) and Lesur *et al.* (2013) showed that doubling the vertical extension of the box divides ζ by a factor of three for $\beta_{\text{mean}} = 10^4$ and by a factor of 1.6 for $\beta_{\text{mean}} = 10$. In any case, this indicates that the outflow rate does not converge in shearing boxes. This is because it is not possible to escape to infinity in the gravitational potential of Hill's approximation (see the discussion in § 5.1). Hence, the 'border' of the gravitational potential is set artificially by the location of the physical boundary in the z direction of the shearing box.

Assuming $L_z \simeq 10H$, the combination of the data published by Suzuki & Inutsuka (2009), Bai & Stone (2013*a*) and Fromang *et al.* (2013) leads to

$$\zeta \simeq 5 \times 10^{-5} + \frac{10}{\beta_{\text{mean}}} \quad \text{for } \beta \gtrsim 10^2. \tag{9.6}$$

This relation gets shallower for $\beta_{\text{mean}} < 10^2$ (Bai & Stone 2013*a*; Lesur *et al.* 2013), and eventually leads to a sharp decrease of ζ when $\beta \lesssim 1$ (see figure 9 in Lesur *et al.* 2013). It is also expected that the prefactor is a decreasing function of L_z and could trace the geometrical aspect ratio H/R of the disc, which is not captured in shearing boxes (Bai & Stone 2013*a*). However, this dependence is complex because it likely depends on the magnetisation. It should also be pointed out that this relation is close to that found in RMHD simulation of the MRI in the context of dwarf novae (Scepi *et al.* 2018). Hence, the disc thermodynamics do not seem to greatly affect the outflow rate. Nevertheless, this scaling should be taken with caution, and is probably only an upper bound to the real outflow rates one would obtain solving for the global problem.

The second parameter characterising the outflow is the angular momentum extracted by the wind, υ, defined as

$$\upsilon \equiv \frac{\overline{T_{yz}}|_{\text{top}} - \overline{T_{yz}}|_{\text{bottom}}}{P_{\text{mid}}}, \tag{9.7}$$

where we have defined the stress tensor $T_{yz} = \rho w_y w_z - B_y B_z / 4\pi$. This quantity directly enters the angular momentum conservation law (see § 4.5) so that, once it is known, one can automatically compute the accretion rate associated with the wind. Evaluating υ is, however, notoriously difficult in shearing boxes because its sign is not well-defined (see the symmetry discussion given previously). A naïve temporal averaging therefore leads to a negligible value of the angular momentum extracted by the wind. In order to circumvent this difficulty, several strategies have been used: Fromang *et al.* (2013) computed υ on a short time period, during which the polarity remains fixed, whereas Bai & Stone (2013*a*) and Scepi *et al.* (2018) computed the time-averaged absolute value of υ. Another difficulty lies in the altitude at which this quantity is evaluated. Bai & Stone (2013*a*) evaluated it at the box boundary, whereas Scepi *et al.* (2018) computed its maximum as a function of z. In all these cases, the results behave like the scaling inspired from Scepi *et al.* (2018):

$$\upsilon \simeq \frac{(4 \pm 3) \times 10}{\beta_{\text{mean}}} \left[\left(\frac{\beta_{\text{mean}}}{4.7 \times 10^4} \right)^2 + 1 \right]^{0.3}, \tag{9.8}$$

where the uncertainty is a result of the different methods of measurement found in the literature. As for ζ, this scaling probably gets shallower for $\beta_{\text{mean}} < 10^2$. However, this has never been properly evaluated in the literature.

It should be noted that for $\beta_{\text{mean}} > 1$, $\upsilon < \alpha$. Hence, the vertical stress is always weaker than the radial stress, by a factor $O(\beta_{\text{mean}}^{1/2})$. However, the respective contribution of these
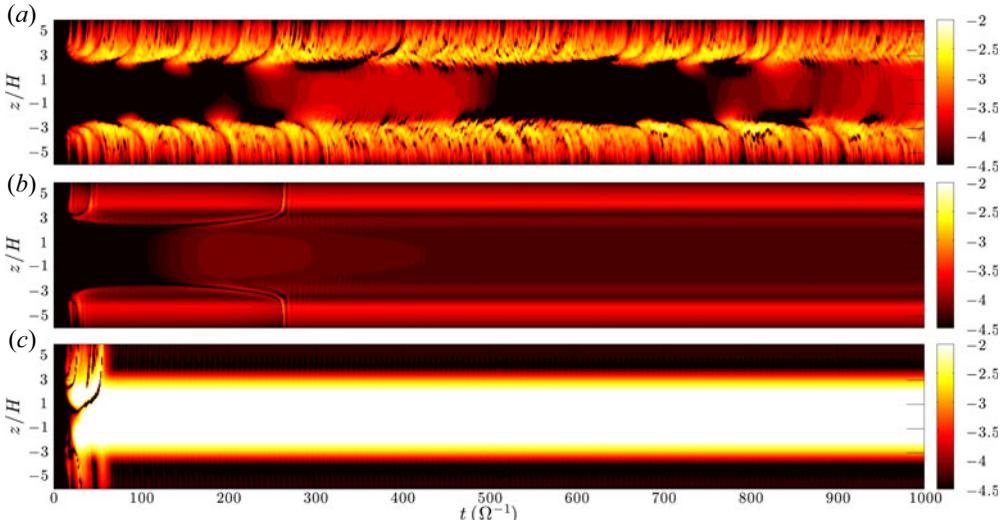
FIGURE 38. Maxwell stress M_{xy} averaged horizontally as a function of t and z in simulations with Ohmic diffusion only (a), Ohmic and ambipolar diffusion (b) and Ohmic, ambipolar and Hall effect (c). Simulations computed at 1 AU assuming a MMSN, with a mean vertical field $\beta_{\mathrm{mean}} = 10^5$. From Lesur *et al.* (2014).

two terms to the mass accretion rate in the disc is proportional to $(R/H)(\upsilon/\alpha)$ (see § 4.5), implying that the vertical stress is the dominant mass accretion mechanism whenever $\beta_{\mathrm{mean}} \lesssim (R/H)^2$ (Fromang *et al.* 2013; Bai & Stone 2013*a*).

9.2.2. *Non-ideal MHD*

Simulations with a mean vertical field first focused on the effect of Ohmic and ambipolar diffusion (Bai & Stone 2013*b*; Simon *et al.* 2013*a*), neglecting the Hall effect for technical reasons. It was found that the presence of a mean field allows the formation of a magnetised outflow at the disc surface ionised by far UV radiation, leading to accretion because of the angular momentum extracted by the outflow. Complete models including also the Hall effect (Bai 2014; Lesur *et al.* 2014; Bai 2015; Simon *et al.* 2015) showed similar outflows, with, in addition, the presence of a midplane laminar stress owing to the Hall effect (Lesur *et al.* 2014).

The difference between Ohmic only, Ohmic + ambipolar, and Ohmic + ambipolar + Hall effect is demonstrated in figure 38 for an MMSN disc at 1 AU. In the Ohmic-only case, one recovers the turbulent surface layer and a dead midplane. Adding ambipolar diffusion leads to a different picture where the disc surface becomes mostly laminar, with a weakly magnetised outflow. This might be surprising at first sight because the linear analysis shows little difference between the Ohmic and Ohmic+ambipolar cases regarding the localisation and growth rates of the eigenmodes (see figure 26). The difference between Ohmic and ambipolar cases is a result of the fact that $\eta_A \propto B^2$ whereas η_O does not depend on B. As eigenmodes grow, the horizontal field grows as well and η_A increases. This rapidly leads to the saturation of the eigenmode by diffusion, in a way similar to § 8.3. In the end, the perturbation never grows to the large-enough amplitudes required to break up in developed turbulence, as in the Ohmic case.

When the Hall effect is eventually added, the disc midplane is subject to a laminar stress in the region 1–10 AU, in addition to the weak surface outflow that subsists. The midplane stress is tightly linked to the field polarity: a negative mean field ($\boldsymbol{\Omega} \cdot \boldsymbol{B} < 0$)

leads to its disappearance. This laminar stress is a result of the HSI (Lesur *et al.* 2014), which only shows up for positive polarities (see the discussion in §§ 6.4.4 and 6.4.6). Hence, the Hall effect is indeed able to revive the MRI in dead zones provided that the mean field has a positive polarity, as already deduced from the linear theory. However, it saturates as a laminar stress. This effect is recovered up to $R \sim 30$ AU but disappears at 100 AU in MMSN models (Simon *et al.* 2015) because the Hall effect is much weaker in the outermost regions of the disc. The presence of this laminar stress suggests that the correlation lengths of the midplane structures are larger than the horizontal box size, which contradicts the spirit of the shearing box approximation and calls for global simulations to properly characterise this transport process.

The top/down symmetry in non-ideal MHD models is also problematic in simulations including non-ideal MHD effects. When $\beta_{\mathrm{mean}} < 10^5$, these simulations exhibit most of the time an odd symmetry with respect to the disc midplane, which implies that the outflow is not extracting any net angular momentum from the disc. There is today no clear explanation for this trend. It could be that the midplane current layer required by the odd symmetry configuration is expelled by the strong Ohmic or ambipolar diffusion (Bai & Stone 2013*b*), or that the HSI spontaneously saturates in an even configuration (Lesur *et al.* 2014), or a combination of these effects. Using similar techniques to that used in ideal-MHD shearing box, several groups have managed to measure the transport coefficient in non-ideal MHD simulations, computing the vertical stress on one side, or computing its absolute value. A representative summary of the results is given in figure 39.

Several trends can be seen in this figure. First and foremost, all of the transport coefficients are reduced by the non-ideal MHD effects (with the notable exception of one run from Lesur *et al.* (2014) with a very strong Hall effect at 1 AU). The most reduced coefficient is the radial angular momentum transport α, whereas the vertical (wind) transport υ appears to be the less affected. This difference in behaviour is the main reason why winds are today favoured in PPDs: the efficiency of wind-driven transport is less affected by non-ideal MHD effects than radial angular momentum transport. In addition to these remarks, one should note that the sensitivity to the field polarity is more pronounced at 1 AU than at 30 AU, as expected from the profile of dimensionless Hall number (figure 10). As already guessed in the previous discussion, aligned fields tend to have a larger α, ζ and υ. It is interesting to also note that, for simulations at 30 AU, the scaling with β_{mean} is similar to that found in ideal MHD, apart from a constant offset.

9.2.3. *Outflow-induced self-organisation*

There is much evidence of self-organisation in stratified shearing boxes. The first piece of evidence can be seen in the space–time diagrams of Simon & Armitage (2014). Bai (2015) then explicitly reported the spontaneous formation of self-organised flows, which is particularly strong for $\beta_{\mathrm{mean}} \lesssim 10^4$. The same process was later identified by Simon *et al.* (2018) and Riols & Lesur (2019). Self-organisation is clearly dissociated from the Hall effect in stratified flows, as it appears also in simulations where the Hall effect is absent (Simon & Armitage 2014; Simon *et al.* 2018; Riols & Lesur 2019).

This self-organisation shows again a very tight intricacy between the field and the gas: magnetic field lines tends to concentrate into 'gaps', leaving regions with a lot of mass but no poloidal field (figure 40). Several scenarios have been proposed to explain this effect. Bai & Stone (2014) proposed that this is a result of the anisotropy of the turbulent diffusivity tensor, following the unstratified box argument (§ 8.5.1).
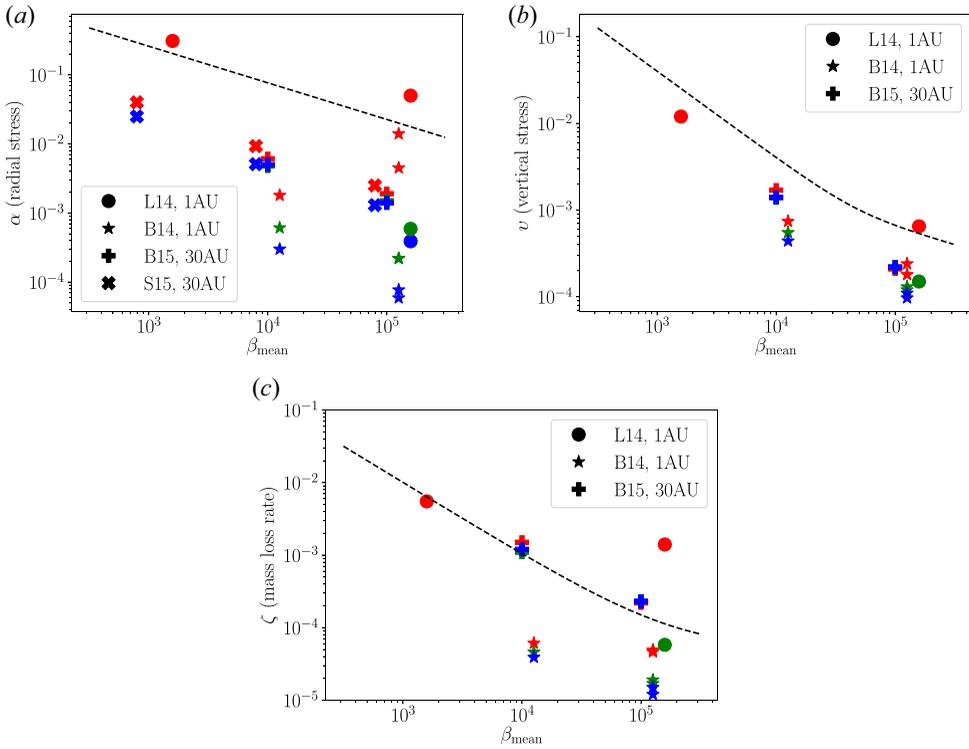
FIGURE 39. Measurements of transport coefficients in the literature, assuming ionisation structures at 1 AU and at 30 AU. The ideal MHD relations (9.4), (9.6) and (9.8) are shown in black dashed lines. We have used data from Lesur *et al.* (2014)=L14, Bai (2014)=B14, Bai (2015)=B15 and Simon *et al.* (2015)=S15. Simulation with the vertical field aligned with the rotation axis are in red, simulations with vertical field anti-aligned are in blue and simulations without Hall effect are in green. Note that Bai (2014) has two chemical models at 1 AU, with and without grains, hence the presence of two sets of points. Note that points on the same β_{mean} have been slightly shifted horizontally to improve readability.
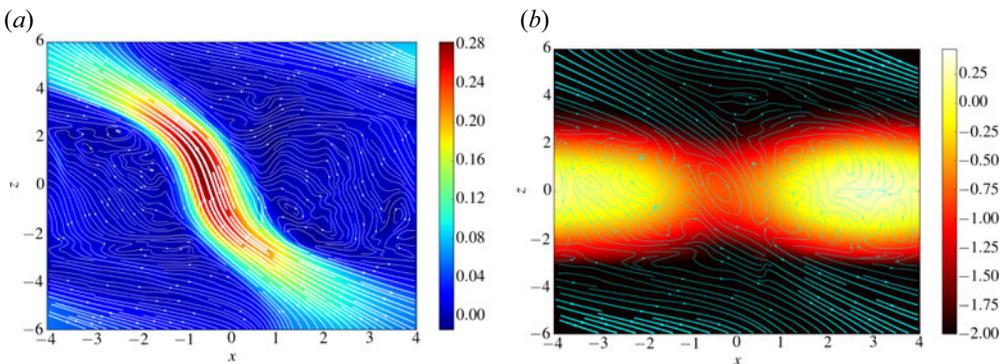
FIGURE 40. Shearing box simulation exhibiting self-organisation with ambipolar diffusion only. (*a*) Magnetic configuration, with poloidal field lines and colours showing the toroidal field component amplitude. (*b*) Density map (colour) and poloidal velocity streamlines. Note the poloidal field concentration in the region $(x, z) \simeq 0$, associated with a minimum of density (='gap'). Figure from Riols & Lesur (2019).
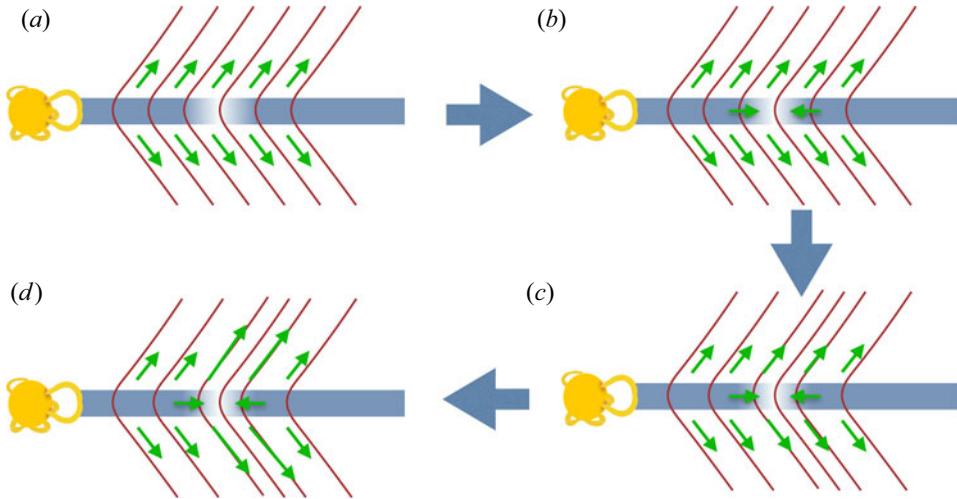
FIGURE 41. Self-organisation feedback loop proposed by Riols & Lesur (2019). Consider a small density deficit (*a*), this deficit induces a radially converging flow (*b*), which drags the poloidal field line inwards (*c*). The stronger field, leads to a more efficient local ejection index that empties the region even more (*d*).

An alternative viewpoint was proposed by Riols & Lesur (2019) who noted that, in the presence of vertical stratification, the average radial flow was converging towards the region of flux accumulation, and hence in the gap. This counter-intuitive finding implies that gaps are necessarily emptied by the outflow, and the radial flow is simply trying to replenish the gaps. This radial flow drags the poloidal field towards the gap (cf. figure 41), and is therefore a pure advection process, not a turbulent anti-diffusion effect as proposed by Bai & Stone (2014). The stronger field in the gap leads to a more efficient ejection, as observed in the ejection 'plume' (cf.figure 40), leading to a quasi-steady state from the disc viewpoint . This feedback loop turns out to be a linear instability of the wind-emitting disc, with predictable growth rate and optimum disturbance. Let us note that the linear instability criterion is simply

$$-\frac{\mathrm{d}\log\zeta}{\mathrm{d}\log\beta_{\mathrm{mean}}} > -\frac{\mathrm{d}\log\alpha}{\mathrm{d}\log\beta_{\mathrm{mean}}} \rightarrow \text{Instability,} \tag{9.9}$$

which essentially compares the time needed to empty the disc by the outflow to the time needed to refill the disc by 'viscous' diffusion (Riols & Lesur 2019). As shown previously, this criterion is generally satisfied in stratified shearing box models, so this 'wind instability' mechanism could potentially explain the self-organisation observed in stratified models.

10. Conclusion

In ideal MHD, the MRI is found to be a robust angular momentum transport mechanism, but in the absence of an external poloidal field, it fails to account for observed accretion rates if one includes the relevant non-ideal MHD effects.

When a mean field is added, magnetised outflows are found as a result of the saturation of the MRI. In ideal MHD, outflows are expected to be the dominant angular momentum transport mechanism when $\beta_{\mathrm{mean}} \lesssim (R/H)^2$. When non-ideal effects are introduced, however, the radial transport of angular momentum is reduced significantly,

whereas outflows are much less affected. As a result, when non-ideal MHD effects are strong, magnetised outflows are expected to become the dominant mechanism of angular momentum transport, even for relatively weak fields.

We note, however, that the shearing box approach cannot model these outflows properly because of several inconsistencies (top/down symmetry of the outflow, mass loss rate depends on boundary conditions). In addition, the presence of a laminar midplane stress owing to the HSI and the existence of a self-organisation mechanism associated with the outflow make the shearing box ill-suited to study the dynamics of these objects. The only way to connect our local understanding of the physics to the global dynamics of these objects is, therefore, to perform global simulations, in which outflows are modelled properly.

PART FOUR: Global models of PPDs

In this section, we revisit the physical concepts underlying ejection in accretion discs. We then use these concepts to interpret the most recent numerical models of PPDs that exhibit accretion and outflows.

11. Some outflow definitions and properties

Before exploring the connection between a weakly ionised disc and an outflow, let us underline a few important definitions and properties on outflows. To this end, let us work first in the ideal MHD approximation and assume for the moment that the outflow is stationary and axisymmetric. In a cylindrical frame (R, ϕ, z), the conservation of mass and momentum read

$$\nabla \cdot \rho \boldsymbol{u}_p = 0 \tag{11.1}$$

$$\rho \boldsymbol{u}_p \cdot \nabla u_R = \rho \Omega^2 r - \partial_R P + \frac{J_\phi B_z}{c} - \frac{J_z B_\phi}{c} - \rho \partial_R \psi \tag{11.2}$$

$$\frac{1}{R} \rho \boldsymbol{u}_p \cdot \nabla \Omega R^2 = \frac{1}{R} \nabla \cdot \left(R \frac{\boldsymbol{B}_p B_\phi}{4\pi} \right) \tag{11.3}$$

$$\rho \boldsymbol{u}_p \cdot \nabla u_z = -\partial_z P - \partial_z \left(\frac{B_\phi^2 + B_R^2}{8\pi} \right) + \frac{B_R \partial_R B_z}{4\pi} - \rho \partial_z \psi \tag{11.4}$$

where the index p denotes the poloidal components (R, z) of the vector fields, $\Omega \equiv u_\phi/R$ (Ω is not necessarily Keplerian) and $\psi = GM/(R^2 + z^2)^{1/2}$ is the gravitational potential, assumed to be solely a result of the central star. Equation (11.3) can be recast in conservative form to obtain an angular momentum conservation equation

$$\nabla \cdot \left(\rho \boldsymbol{u}_p \Omega R^2 - R \frac{\boldsymbol{B}_p B_\phi}{4\pi} \right) = 0, \tag{11.5}$$

which shows that the magnetic field can carry angular momentum in the Maxwell stress $B_p B_\phi$.

In addition to the equation of motion, one needs to solve the induction equation in the ideal regime in the poloidal and azimuthal directions

$$\nabla \times \left(\boldsymbol{u}_p \times \boldsymbol{B}_p \right) = 0, \tag{11.6}$$

$$\nabla \cdot \frac{1}{R} \left(\Omega R \boldsymbol{B}_p - B_\phi \boldsymbol{u}_p \right) = 0. \tag{11.7}$$

Finally, the magnetic field has to satisfy the solenoidal condition $\nabla \cdot \boldsymbol{B} = 0$. In an axisymmetric configuration this can be used to express the poloidal field components as a function of a scalar streamfunction a:

$$\boldsymbol{B}_p = \frac{1}{R}\nabla a \times \boldsymbol{e}_\phi. \tag{11.8}$$

By construction, the value of a is constant along a poloidal field line. Hence, we can label each poloidal field line with the value of its streamfunction. From these equations, it is possible to derive several physical constants of motion that are useful to interpret outflow solutions.

11.1. *Frozen-in condition*

The frozen-in condition is derived from mass conservation (11.1) and the poloidal induction equation (11.6). Let us first start with the vertical component of the induction equation, which reads

$$\frac{1}{R}\frac{\partial}{\partial R}R\left(u_z B_R - u_R B_z\right) = 0, \tag{11.9}$$

from which one deduces that

$$u_z B_R - u_R B_z = \frac{\beta(z)}{R}, \tag{11.10}$$

where β is an unknown scalar function (physically, β/R is the ϕ component of the electromotive force E_ϕ). Using the radial induction equation, we may see that $\partial_z \beta = 0$ so that β is a constant. To avoid any singularity for E_ϕ at $R = 0$, we are then forced to have $\beta = 0$ and hence $\boldsymbol{B}_p$ and $\boldsymbol{u}_p$ are parallel, i.e.

$$\boldsymbol{u}_p = \mu(R, z)\boldsymbol{B}_p. \tag{11.11}$$

We can inject this relation in the mass conservation equation to obtain

$$\boldsymbol{B}_p \cdot \nabla(\rho\mu) = 0, \tag{11.12}$$

which indicates that $\eta \equiv \rho\mu$ is a constant along magnetic field lines, i.e. $\eta = \eta(a)$. We can then recast (11.11) into

$$\boldsymbol{u}_p = \eta(a)\frac{\boldsymbol{B}_p}{4\pi\rho}, \tag{11.13}$$

which constitutes the frozen-in condition where η describes the amount of mass loaded along a poloidal field line. Here η is a direct measure of the gas density ρ_A at the Alfvén surface, which is the region defined by $\boldsymbol{u}_p = V_{A,p} = B_p/\sqrt{4\pi\rho}$:

$$\eta(a) = \sqrt{4\pi\rho_A(a)}. \tag{11.14}$$

Note regarding the applicability of the frozen-in condition in shearing boxes: when deriving the frozen-in condition, we have used a regularity condition at $R = 0$ for the electromotive force. Such a condition does not exist in the shearing box approximation so that E_ϕ can be non-zero in principle. This implies that in a stationary shearing box solution, the poloidal field and velocity are not necessarily parallel (Lesur *et al.* 2013). Physically, it means that field lines can be indefinitely dragged radially inward or outward, without affecting the stationarity condition, thanks to the assumed radial periodicity. This is physically impossible from a global point of view, but it illustrates once again the numerous drawbacks of the shearing box model when dealing with outflows.

11.2. *Magnetic surface rotation*

The rotation of magnetic surfaces is obtained by substituting the frozen-in condition (11.13) into the azimuthal component of the induction equation (11.7):

$$\boldsymbol{B}_p \cdot \boldsymbol{\nabla} \left(\Omega - \frac{B_\phi \eta(a)}{4\pi\rho R} \right) = 0, \tag{11.15}$$

which allows us to define a new invariant along field lines

$$\Omega^*(a) \equiv \Omega - \frac{\eta(a)}{4\pi\rho R} B_\phi = \Omega_A - \frac{V_{A\phi}(R_A)}{R_A}, \tag{11.16}$$

where the A indices denote quantities evaluated at the Alfvén surface and R_A is the Alfvén radius, where the outflow becomes super-Alfvénic. It can easily be checked that in a frame rotating at $\Omega^*(a)$, the total field is parallel to the total velocity. In other words, there is no induced electromotive force in the frame rotating at $\Omega^*(a)$. Therefore, Ω^* can be interpreted as the rotation speed of magnetic surfaces. By combining (11.13) and (11.16), we can express the total velocity as a function of the magnetic field,

$$\boldsymbol{u} = \frac{\eta(a)}{4\pi\rho} \boldsymbol{B} + R\Omega^*(a)\boldsymbol{e}_\phi. \tag{11.17}$$

This illustrates the fact that only the poloidal components of the field and the velocity are parallel in general. In the particular case where $\eta = 0$ (no motion along the poloidal field lines), we recover Ferraro's iso-rotation law $\Omega = \Omega^*(a)$.

11.3. *Angular momentum*

The angular momentum invariant is easily deduced from the angular momentum conservation equation (11.5):

$$\boldsymbol{B}_p \cdot \boldsymbol{\nabla} \left(\Omega R^2 - \frac{R B_\phi}{\eta(a)} \right) = 0 \tag{11.18}$$

from which we deduce the angular momentum invariant $\mathcal{L}(a)$:

$$\mathcal{L}(a) \equiv \Omega R^2 - \frac{R B_\phi}{\eta(a)}, \tag{11.19}$$

$$= \Omega^*(a) R_A^2, \tag{11.20}$$

where the last equality has been obtained at the Alfvén surface using (11.14). Hence, the amount of angular momentum transported by the outflow is an invariant made of two parts: the classical kinetic contribution and a magnetic part stored in B_ϕ.

11.4. *Bernoulli invariant*

The Bernoulli invariant is obtained from the scalar product of the poloidal equations of motion (11.2)–(11.4) with $\boldsymbol{u}$. One obtains

$$\boldsymbol{u}_p \cdot \boldsymbol{\nabla} \left[\frac{u^2}{2} + \psi_G \right] = -\boldsymbol{u}_p \cdot \frac{\boldsymbol{\nabla} P}{\rho} + \boldsymbol{u} \cdot \frac{\boldsymbol{J} \times \boldsymbol{B}}{\rho c}, \tag{11.21}$$

where $\psi_G = GM/(R^2 + z^2)^{1/2}$ is the gravitational potential. We then use (11.17) to obtain the work of the Lorentz force:

$$\boldsymbol{u} \cdot \frac{\boldsymbol{J} \times \boldsymbol{B}}{c} = R\Omega^*(a)\frac{\boldsymbol{J}_p \times \boldsymbol{B}_p}{\rho c}$$

$$= \Omega^*(a)\frac{\boldsymbol{B}_p \cdot \nabla\left(RB_\phi\right)}{4\pi\rho}$$

$$= \boldsymbol{u}_p \cdot \nabla\left(\frac{R\Omega^*(a)B_\phi}{\eta(a)}\right). \tag{11.22}$$

In addition, we may express the work of pressure forces using the enthalpy per unit mass $\mathcal{H}$

$$\boldsymbol{u}_p \cdot \frac{\nabla P}{\rho} = \boldsymbol{u}_p \cdot \nabla\mathcal{H} + \delta\mathcal{Q} \tag{11.23}$$

in which we have also considered the effect of an additional heating term owing to radiative heating/cooling $\delta\mathcal{Q}$. Putting all the terms together and integrating along one particular streamline a, we find that the quantity

$$\mathcal{B} \equiv \frac{u^2}{2} + \psi_G + \mathcal{H} - \frac{R\Omega^*(a)B_\phi}{\eta(a)} + \int_{\mathcal{S}(a)} \delta\mathcal{Q}\,\mathrm{d}s \tag{11.24}$$

is conserved along poloidal field lines and streamlines. Note that because the heating term is not a proper differential, it depends on the integral of the heating term along the chosen streamline $\mathcal{S}(a)$. Of course, for an outflow to escape up to $z \to \infty$, one needs $\mathcal{B} > 0$ on the streamlines (assuming $\mathcal{H} > 0$). In the midplane of a Keplerian disc, one has $u^2/2 + \psi_G = -v_K^2/2$ where v_K is the Keplerian velocity, so that additional ingredients are required to produce an outflow. Two extreme situations can be identified.

Thermal winds: Here magnetic effects are neglected and ejection is possible because the disc is hot (large initial enthalpy) or because a lot of heating is applied along the streamlines (named photo-evaporation in the PPDs community).

Cold MHD winds: Here thermal effects are neglected and the toroidal field acts as an energy reservoir to launch the outflow.

11.5. *Dimensionless numbers characterising an outflow*

Based on the MHD invariants, it is possible to define a series of dimensionless numbers that characterise an outflow streamline unambiguously. These numbers use the physical properties at the base of the outflow. Let us therefore write Ω_0 as the rotation rate at the base of the outflow and R_0 as its cylindrical radius (one typically has $\Omega_0 = \Omega_K(R_0)$), and B_0 as the poloidal field strength threading the disc at the location where the outflow is launched. We then define

$$\left.\begin{aligned} \kappa &\equiv \eta\frac{\Omega_0 R_0}{B_0}, \\ \omega^* &\equiv \frac{\Omega^*}{\Omega_0}, \\ \lambda &\equiv \frac{\mathcal{L}}{\Omega_0 R_0^2} = \omega\left(\frac{R_A}{R_0}\right)^2, \\ e &\equiv \frac{\mathcal{B}}{\Omega_0^2 R_0^2/2}. \end{aligned}\right\} \tag{11.25}$$

The Bernoulli invariant can be easily expressed in terms of the other invariants, using the dimensionless rotation rate $\omega \equiv \Omega/\Omega_0$:

$$e = \frac{u_p^2}{\Omega_0^2 R_0^2} + \omega^2 \frac{R^2}{R_0^2} - \frac{2R_0}{\sqrt{R^2 + z^2}} + 2\omega^* \left(\lambda - \omega \frac{R^2}{R_0^2} \right) + \theta, \qquad (11.26)$$

where $\theta = 2 \left[\mathcal{H} + \int_{a=\text{cste}} \delta \mathcal{Q} \, ds \right] / \Omega_0^2 R_0^2$ is the dimensionless thermal energy content of the flow. This expression clearly demonstrates the contribution of λ to the energy content of the flow. This parameter, often called the magnetic level arm, is of key importance, as can be seen if one computes its value at the outflow base, assuming $u_p \ll \Omega_0 R_0$, $\omega = 1$ and $R = R_0$:

$$e \simeq 2\omega^* (\lambda - 1) + \theta - 1. \qquad (11.27)$$

Obviously, for the outflow to be able to propagate up to infinity, one needs $e > 0$. We recover here the two extreme limits discussed previously: purely thermal winds, which have $\lambda = 1$ and require $\theta > 1$; or cold MHD winds with $\theta = 0$, which need

$$2\omega^* (\lambda - 1) > 1. \qquad (11.28)$$

For all practical applications of MHD outflows, one has $\omega^* \simeq 1$ to a very good approximation. This implies that an outflow can exist only if

$$\lambda > \tfrac{3}{2}. \qquad (11.29)$$

11.6. *Ejection efficiency*

The existence of an outflow is tightly linked to the process of accretion happening inside the disc because the energy of the wind is obtained from the accretion power of the disc. To understand this connection, let us first define the accretion rate of the disc

$$\dot{M}_{\text{acc}}(R) \equiv - \int_{z^-}^{z^+} dz \int_0^{2\pi} R \, d\phi \rho u_r, \qquad (11.30)$$

where the integration is performed on the box vertical extension, defined by z^- and z^+. It is also useful to define the outflow rate of the wind between the inner radius of the disc R_{in} and the radius under consideration

$$\dot{M}_{\text{wind}}(R) \equiv \int_{R_{\text{in}}}^R dR \int_0^{2\pi} R \, d\phi [\rho u_z]_{z^-}^{z^+}. \qquad (11.31)$$

These two quantities are connected by the continuity equation (4.20):

$$2\pi \frac{\partial \Sigma}{\partial t} + \frac{1}{R} \frac{\partial}{\partial R} \left(\dot{M}_{\text{wind}} - \dot{M}_{\text{acc}} \right) = 0. \qquad (11.32)$$

At this stage, it is useful to introduce the ejection efficiency index

$$\xi \equiv \frac{d \log \dot{M}_{\text{acc}}}{d \log R} \qquad (11.33)$$

$$= \frac{1}{\dot{M}_{\text{acc}}} \frac{d\dot{M}_{\text{wind}}}{d \log R}, \qquad (11.34)$$

which quantifies what fraction of the mass is being lost in the wind, the second line being obtained from the continuity equation, assuming stationarity. As expected, $\xi = 0$ corresponds to a situation without any wind.

One can also relate the accretion rate to the ejection rate using the angular momentum conservation equation (4.21) as

$$\dot{M}_{\mathrm{acc}} = \frac{4\pi}{\Omega_K} \left[\underbrace{\frac{1}{R}\frac{\partial}{\partial R}R^2\overline{W}_{R\phi}}_{\tau_R} + \underbrace{R\langle W_{z\phi}\rangle_{z^-}^{z^+}}_{\tau_z} \right], \tag{11.35}$$

where we have defined the stress tensor $W_{i\phi} \equiv \rho v_i v_\phi - B_i B_\phi / 4\pi$. This allows us to define the radial and vertical contribution to the accretion of the disc $\tau_{R\phi}$ and $\tau_{z\phi}$. It is then useful to introduce the ratio of these two quantities

$$\Lambda \equiv \frac{\tau_z}{\tau_r} \tag{11.36}$$

so that the accretion rate is simply

$$\dot{M}_{\mathrm{acc}} = \frac{4\pi}{\Omega_K}\tau_z\frac{1+\Lambda}{\Lambda}. \tag{11.37}$$

It is then possible to relate the vertical torque $\tau_{z\phi}$ to the mass ejection rate by noting that the torque is evaluated high above the disc midplane, so that the kinetic contribution to the stress is negligible, which implies $\tau_{z\phi} \simeq -[RB_\phi B_z/4\pi]_{z^-}^{z^+} \simeq -RB_0 B_\phi(z^+)/2\pi$ where the second equality assumes the outflow is top/down symmetric and that the poloidal field strength does not vary much between the midplane and the disc surface. It is then simple to show that the vertical torque is directly connected to the MHD invariants

$$\tau_z \simeq 2(\Omega^* R_A^2 - \Omega_0 R_0^2)[\rho u_z](z_+). \tag{11.38}$$

The vertical mass flux being directly related to the radial derivative of the outflow rate, we can express the accretion rate as

$$\dot{M}_{\mathrm{acc}} \simeq \frac{2}{R_0^2\Omega_0}\frac{1+\Lambda}{\Lambda}(\Omega^* R_A^2 - \Omega_0 R_0^2)\frac{\mathrm{d}\dot{M}_{\mathrm{wind}}}{\mathrm{d}\log R}. \tag{11.39}$$

From which we obtain an expression for the mass ejection index

$$\xi \simeq \frac{\Lambda}{2(\Lambda+1)}\frac{1}{\lambda-1}. \tag{11.40}$$

This relation reveals several key features of MHD outflows. First, it shows that large level arms λ are associated with small ejection indices. Interestingly, this result does not depend on the radial contribution to the angular momentum budget Λ. Second, the energy constraint (11.29) imposes an upper bound on ξ in cold MHD winds: $\xi \lesssim \Lambda/(\Lambda+1) \lesssim 1$. Note that this relation allows for outflows approximately as massive as the mass accretion rate, but it does not allow for outflows vastly more massive than this. Outflows with $\xi \gg 1$ therefore necessarily require some thermal energy driving to escape the potential well, as one would expect.

11.7. *Connection to shearing box simulations*

In the previous part, we had to introduce several local quantities to characterise outflows in local shearing box models. As pointed out, however, shearing boxes lack global constraints, which implies that some of the solutions are likely unphysical. As a first step, it is therefore useful to relate these local quantities to global MHD invariants to test the domain of validity of shearing box solutions. In this subsection, we make the assumption that the global outflow is top/down symmetric, so that the invariants are the same on both sides of the disc.

We first start with the outflow rate ζ (defined in (9.5)), which can be easily connected to the mass loading parameter κ

$$\kappa = \frac{1}{4}\frac{R}{H}\beta_{\text{mean}}\zeta. \tag{11.41}$$

The magnetic level arm can also be related to υ (defined in (9.7)), provided that we neglect the kinetic contribution to the vertical stress, which is valid if we choose the disc upper boundary to lie high enough above the midplane

$$\lambda = 1 + \frac{H}{R}\frac{\upsilon}{\zeta}. \tag{11.42}$$

The energetic constraint (11.29) therefore leads to a new constraint in shearing boxes: $2\zeta/\upsilon < H/R$. This constraint cannot be satisfied by the shearing box scalings (9.6) and (9.8) when $H/R \lesssim 0.2$. Hence, shearing box wind solutions always eject too much mass with too little energy to escape the global potential well of discs with realistic aspect ratios. In other words, if one puts a shearing box wind solution in a global disc configuration, the ejected material should fall back onto the disc.

Finally, we can relate the stress rate Λ to α and υ. For this, let us assume that the disc surface density follows a power law: $\Sigma = \Sigma_0(R/R_0)^{-p}$, that α and H/R are constant with radius, and that the disc is vertically isothermal. Under these assumptions, the contributions to the mass accretion are

$$\left.\begin{aligned} \tau_R &= \alpha\Omega_0^2 R_0^2 \left(\frac{H}{R}\right)^2 (1-p)\,\Sigma \\ \tau_z &= \frac{1}{\sqrt{2\pi}}\upsilon\Omega_0^2 R_0^2 \frac{H}{R}\Sigma \end{aligned}\right\} \tag{11.43}$$

so that the ratio simply reads

$$\Lambda = \frac{1}{\sqrt{2\pi}(1-p)}\frac{\upsilon}{\alpha}\frac{R}{H}. \tag{11.44}$$

The scalings (9.6) and (9.8) then suggest that $\Lambda \to 0$ when $\beta_{\text{mean}} \to \infty$. Combining these results with (11.40), this implies that the ejection index ξ tends to decrease as $\beta_{\text{mean}} \to \infty$, as one would expect.

12. Outflow phenomenology

The launching of an outflow is tightly linked to physics of the accreted material because the outflow energy eventually comes from the accretion power of the disc. Let us divide the overall structure into a 'disc region' and an 'outflow' region (figure 42) and describe the physical processes in each region.
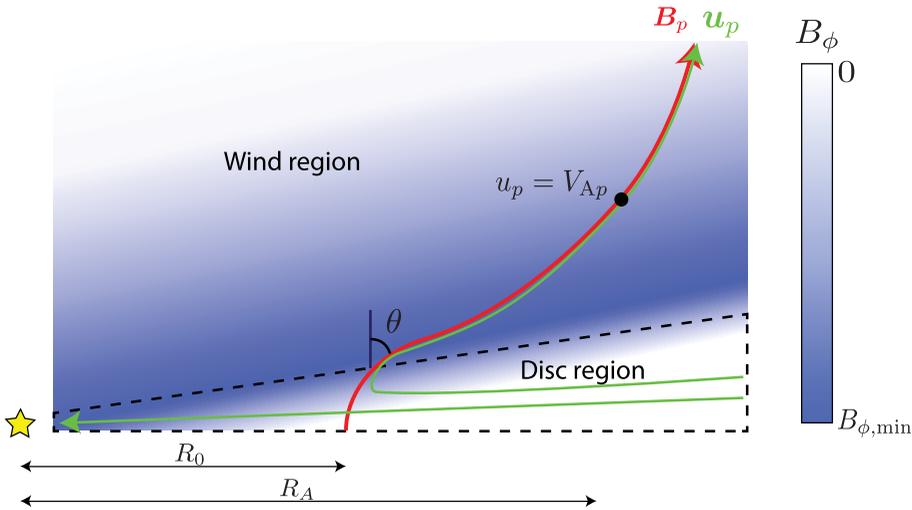
FIGURE 42. Global disc–wind interaction scheme. We distinguish a disc region in dashed lines from the outflow region. The poloidal field line is represented in red and the poloidal streamlines in green. In addition, the toroidal field component is shown in blue in a frame corotating at $\Omega_K(R_0)$.

12.1. *Disc region*

In the disc region, the flow is mostly accreted, whereas the field lines are stationary. The fact that poloidal field and streamlines are not parallel implies that non-ideal MHD effects are necessarily present in the disc. Historically, these non-ideal effects have been treated in two ways: (i) assuming that the disc ionisation fraction is extremely low as in PPDs, so that ambipolar and Ohmic diffusion are very large (Konigl 1989; Wardle & Konigl 1993); or (ii) assuming that the disc was turbulent, the turbulence leading to a 'turbulent resistivity' modelled as a non-isotropic diffusivity tensor (Ferreira & Pelletier 1993).

The toroidal component of the magnetic field B_ϕ is the key ingredient of the interaction between the disc and the outflow. In the disc, the toroidal field is produced by the shearing of the radial field by the differential rotation of the disc $\partial_t B_\phi \simeq B_R R \partial_R \Omega$. This effect is actually the main energy source for the outflow, which converts shear energy into magnetic energy stored in B_ϕ at the disc surface. It is often assumed that outflows are top-down symmetric to satisfy the global symmetries of the system, so that $B_R(z=0) = 0$ and $B_\phi(z=0) = 0$ follows. The bending of poloidal field lines in the disc implies $\partial_z B_R > 0$ for $z > 0$ and hence $\partial_z B_\phi < 0$ up to the disc surface. The actual value of B_ϕ depends on the competition between the shearing of B_R and the magnetic diffusion which damps the shear amplification. Overall, one can estimate B_ϕ^+ at the disc surface by balancing shear and diffusion

$$B_\phi^+ \sim B_z \tan(\theta) \frac{\Omega h^2}{\eta}, \tag{12.1}$$

where h is the disc scale height, η is the magnetic diffusivity and θ is the inclination of the poloidal field at the disc surface (see figure 42). Accretion of the disc material naturally follows from the profile of $B_\phi(z)$. The toroidal field is responsible for an azimuthal magnetic tension force $B_z \partial_z B_\phi / 4\pi$, which slows down the rotating material and leads

to accretion. We can deduce the accretion rate from (4.21):

$$\rho u_R \sim \frac{B_\phi^+ B_z}{4\pi\Omega h} \sim \frac{B_z^2 \tan(\theta)h}{4\pi\eta}, \tag{12.2}$$

where ρ is measured in the middle of the accretion flow (which, in general, is the disc midplane, but can also be off-midplane for dissymmetric outflows). This accretion is physically a result of a transfer of angular momentum from the accreted material to the toroidal field. In the end, the angular momentum is stored in B_ϕ^+ and is eventually used to launch the outflow.

The outflow base has to be replenished by the disc. Hence, a net positive vertical acceleration is required in the disc region to push material upward, even if at modest velocities. A careful examination of (11.4) allows us to isolate the role played by each term in the vertical acceleration: the magnetic pressure is necessarily directed downward because B_R^2 and B_ϕ^2 are both increasing functions of z. Gravity is also directed downward. The magnetic tension term is usually small (it involves a radial derivative whereas the other terms involve vertical derivatives, which are larger by a factor of R_0/h), and typically directed downward if we assume the most natural situation with B_z decreasing with radius. Hence, the only term leading to an upward acceleration is the vertical thermal pressure gradient. The role played by the thermal pressure at the base of the outflow has often been missed. However, it has some important consequences. For instance, B_ϕ^+ cannot reach arbitrarily large values that would otherwise prevent thermal pressure from pushing materials to the wind base and vertically squeeze the disc. Ejection therefore requires $(B_\phi^+)^2/8\pi \lesssim P^+$ (Ferreira 1997).

A summary of the physics of the disc region is as follows.

   (i) The region is non-ideal as the gas has to be allowed to stream through poloidal field lines which are stationary.
  (ii) The toroidal field at the disc surface stores the energy and angular momentum taken from the accreted material.
 (iii) The upward motion needed to replenish the outflow base is a result of the vertical thermal pressure gradient. This sets a limit to the amount of toroidal field that can be stored at the disc surface because magnetic pressure prevents this upward motion.

12.2. *Outflow region*

In order to understand the dynamics of the outflow from its base, it is convenient to introduce the Alfvénic Mach number ξ defined as

$$\xi \equiv \frac{u_p}{V_{A,p}} = \eta\sqrt{\frac{4\pi}{\rho}}, \tag{12.3}$$

where the second equality is deduced from (11.13). In the outflow, we can expect a continuous acceleration of the flow, so that ξ is an increasing function as one moves along a poloidal streamline, with $\xi \sim 0$ at the outflow base. We can eliminate B_ϕ in favour of the angular velocity by combining the magnetic surface rotation invariant (11.16) and the angular momentum invariant (11.19) to obtain

$$\Omega = \Omega^* \left(\frac{1 - \xi^2(R_A/R)^2}{1 - \xi^2} \right). \tag{12.4}$$

This expression looks singular for $\xi = 1$, which corresponds to the Alfvén point. However, at this particular point, $R = R_A$ so that Ω is actually smooth across this point. Second, in

the limit of low Mach numbers $\xi \ll 1$, which corresponds to the base of the outflow, we can expand this expression to obtain

$$\Omega \simeq \Omega^* \left[1 - \xi^2 \left(\frac{R_A^2}{R^2} - 1 \right) + O(\xi^4) \right]. \tag{12.5}$$

Hence, the outflow is rotating at a constant angular velocity up to the point where $\xi^2 \simeq 1/(R_A^2/R^2 - 1)$. Physically, the angular momentum stored in B_ϕ^+ by the disc is progressively used to accelerate the flow via the azimuthal magnetic tension force, leading to the apparent solid rotation profile. This works until most of the toroidal field has been used and the angular momentum is all in kinetic form. On the opposite limit $\xi \to \infty$, we find $\Omega \simeq \Omega^* R_A/R^2$, i.e. a constant angular momentum rotation profile, as expected.

The vertical acceleration of the outflow results from angular momentum conservation. As demonstrated previously, angular momentum stored in B_ϕ is converted into kinetic angular momentum. Hence, B_ϕ decreases along the streamline. This leads to a magnetic pressure force $\partial_z B_\phi^2$ directed upward in (11.4) and, hence, a vertical acceleration of the outflow. As the outflow bends toward the vertical axis, B_R decreases as well, leading to an additional magnetic pressure term $\partial_z B_R^2$ accelerating the outflow vertically. Overall, the vertical acceleration is a magnetic pressure effect owing to the decrease of B_ϕ and B_R along the streamlines.

One can use the Bernoulli invariant to characterise the topology of the outflow close to the wind base. For simplicity and following Blandford & Payne (1982), let us assume that the wind base is located at $(R_0, z = 0)$. We consider a fluid particle, initially following a Keplerian rotation orbit at $(R_0, z = 0)$ and we assume this particle follows the streamline of the outflow so that at a later time, the particle is located at $(R_0 + \delta R, \delta z)$. As we focus on the launching region of the outflow, we assume that $\Omega \simeq \Omega^*$. During this displacement, the Bernoulli invariant should be conserved, so we should have

$$-\frac{1}{2}(\Omega^* R_0)^2 + \psi_G(R_0, 0) > -\frac{1}{2}[\Omega^*(R_0 + \delta R)]^2 + \psi_G(R_0 + \delta R, \delta z)$$

$$\to \frac{1}{2}(\Omega^* R_0)^2 \left(-3 \left(\frac{\delta R}{R_0} \right)^2 + \left(\frac{\delta z}{R_0} \right)^2 \right) < 0, \tag{12.6}$$

where the inequality comes from the assumption that $u_p(R_0 + \delta R, \delta z) > u_p(R_0, 0)$ (i.e. the flow accelerates). It can be converted into a criterion on θ: $\tan \theta > 1/\sqrt{3}$ or $\theta > 30°$. This well-known criterion, initially derived by Blandford & Payne (1982), is valid provided that thermal effects (enthalpy, heating) are negligible, i.e. in cold winds. This critical angle is usually interpreted in the framework of the 'magneto-centrifugal acceleration' of Blandford & Payne (1982): magnetic field lines are assumed to behave as rigid poloidal wires on which fluid particles are drifting ('bead on a wire' analogy). If the field lines are inclined sufficiently, the particles can overcome the vertical gravity and are accelerated by the centrifugal force.

This simple physical interpretation based on the Bernoulli invariant close to the wind base is very useful as a first approach to outflow physics, but it should not be taken too strictly as it leads to several misconceptions about the very nature of outflows. Let us underline the main physical differences between the magneto-centrifugal acceleration picture and the processes at work in an outflow.

(i) Magnetic field lines do not behave as rigid poloidal wires. As shown previously, the magnetic configuration has a strong toroidal field at the wind base B_ϕ^+, which indicates that the field is wound up in this region.

(ii) The constant angular velocity approximation is only valid close to the wind base (and far from the Alfvén radius). It is a result of the toroidal field tension, which accelerates the flow azimuthally, hence the necessity of a wound field configuration. Winds having $R_A/R_0 \sim 1$ will therefore experience a very small region of solid rotation (if any).

(iii) As discussed previously, vertical acceleration is always a magnetic pressure effect that does not rely on gravity, centrifugal acceleration or magnetic tension.

As proposed by Contopoulos & Lovelace (1994) and Ferreira (1997), it is therefore preferable to qualify outflows as 'magnetically driven' instead of 'centrifugally driven'.

Let us summarise here the physics of a cold outflow close to the launching region up to the Alfvén point.

(i) The flow is accelerated azimuthally by the magnetic tension owing to B_ϕ. This results in a transfer of angular momentum from B_ϕ to the gas as it accelerates.

(ii) As B_ϕ^2 (and, possibly, B_R^2) decreases with z, the flow is accelerated vertically by magnetic pressure.

(iii) Close to the launching point and far from the Alfvén radius, the outflow is approximately in solid rotation.

(iv) In a cold wind, energy conservation at the wind base implies $\theta > 30°$.

13. Global numerical models

Many numerical models of PPDs have been published in the literature. Here, we focus on models tackling the effect of non-ideal MHD effects and winds on the dynamics of a disc, because we focus on the outer parts of PPDs. We therefore exclude 'ideal MHD' models and simulations without a large-scale magnetic field.

The first model to investigate this regime was published by Gressel *et al.* (2015). However, the limited vertical extension of this work (typically four scale heights) makes the interpretation of outflow properties difficult. Therefore, we focus on models with a larger vertical extension such as those published by Béthune, Lesur & Ferreira (2017) and Bai (2017).

13.1. *Global topology*

One of the main problems with shearing-box models is the presence of an odd symmetry for the solution, leading to difficulties to interpret the role played by the outflow. The use of global models avoids this problem because, in this case, the gravitational potential is not symmetrical with respect to r. Many models seem to converge towards dissymmetric outflows, i.e. outflows that are neither even or odd, but are essentially odd for $-3H < z < 3H$ and exhibit a strong current sheet on one side of the disc. In this configuration, the outflow is dissymmetric, and more mass ejected from one side of the disc than the other (with ratios of $\dot{M}_w$ reaching a factor of a few). This kind of solution was found both with only Ohmic and ambipolar diffusion (Gressel *et al.* 2015) and with all three non-ideal effects (Bai 2017; Béthune *et al.* 2017).

The presence of an odd symmetry in the midplane region is reminiscent of the shearing box solutions described in § 9.2. However, this symmetry is not verified far away from the midplane as global effects enter the scene. Gressel *et al.* (2015) have proposed that the global field structure was playing a significant role in shaping the outflow. Indeed, by choosing initially a field configuration with $\partial_R B_z > 0$, it is possible to produce outflows directed inwards (towards the star). It has therefore been proposed that the global radial magnetic pressure gradient was responsible for the symmetry breaking observed in
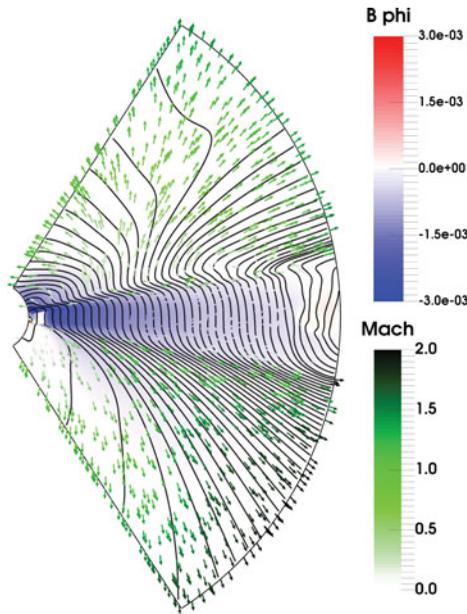
FIGURE 43. Global simulation for a wind in a PPD which exhibits a dissymmetric outflow. Black lines are poloidal magnetic field lines, green arrows represent the poloidal velocity and the background colour traces the azimuthal field B_ϕ. Close to the midplane, the configuration has an odd symmetry, as found in shearing box models and a current sheet is found at $z \sim 3H$. Figure from Béthune *et al.* (2017).

these models. In addition, the vertical rotation profile (and therefore the vertical temperature structure) might be playing a role by shearing the poloidal field (Gressel *et al.* 2015). The question of the origin of the global outflow configuration therefore remains open (figure 43).

Outflows are not always dissymmetric. Indeed, both Béthune *et al.* (2017) and Bai (2017) have reported symmetric (even) outflow configurations. These symmetric configurations seem to be found mostly when B_z is weak enough ($\beta_{\mathrm{mid}} \gtrsim 10^4$) (Bai 2017; Béthune *et al.* 2017) and for anti-aligned cases ($B_z \Omega < 0$). The sensitivity of the outflow configuration on the field polarity suggests that the HSI is partly responsible for the outflow configuration. However, there is no one-to-one correspondence between the field alignment/strength and the outflow configuration, and the choice of initial conditions also seems to be playing a non-negligible role. If true, it might be desirable to start from a magnetic configuration as close as possible to the configuration expected from core collapse calculations (Bai 2017).

13.2. *Accretion*

The question of the engine driving accretion can be directly addressed in global simulations. Indeed, one can measure individually each term in the angular momentum conservation equation

$$\dot{M}_{\mathrm{acc}} = \frac{4\pi}{R\Omega_K} \left\{ \underbrace{\frac{\partial}{\partial R} R^2 \left[\overline{\rho v_\phi v_r} - \frac{\overline{B_\phi B_r}}{4\pi} \right]}_{\tau_r} + \underbrace{\left[R^2 \rho v_\phi v_z - R^2 \frac{B_\phi B_z}{4\pi} \right]_{z=-h}^{+h}}_{\tau_z} \right\}, \qquad (13.1)$$
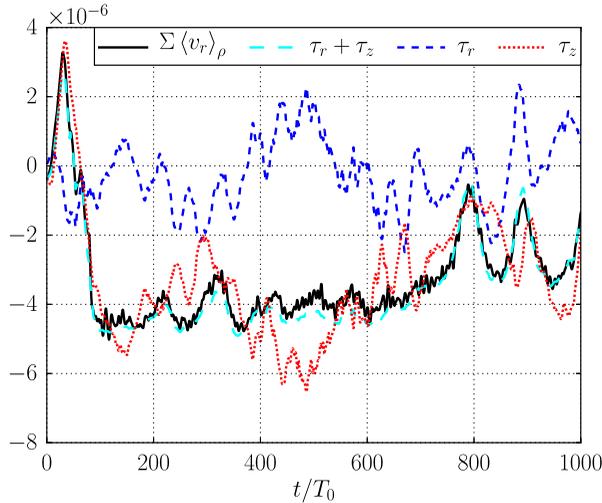
FIGURE 44. Measured accretion rate as a function of time in a non-ideal model. The accretion rate and torque contributions have been average radially. Most of the accretion is a result of the wind (τ_z term) whereas the radial stress does not seem to contribute to the angular momentum budget. Figure from Béthune *et al.* 2017.

where we have defined the mass accretion rate $\dot{M}_{\rm acc} = -2\pi R \overline{\rho v_r}$ and the radial and vertical torques $\tau_{r,z}$. An example of such a measure is given in figure 44. We find that accretion is mostly a result of the wind, the surface torque being the main contribution to angular momentum extraction in the disc.

Note, however, that the fact that the radial torque is negligible does not imply that the radial stress is also negligible. As can be seen from (13.1), one can cancel the radial torque if the radial stress is proportional to $1/R^2$. In Béthune *et al.* (2017), a strong laminar radial stress is indeed present in the disc, with effective α values reaching a few times 10^{-2}. However, because of the surface density profile, the net torque exerted on the disc mostly cancels out. Similar behaviours have been obtained by Bai (2017), with disc regions exhibiting positive and negative torques.

Overall, it is clear that in these models, the wind torque is playing a significant if not dominant role in the mass accretion rate. The radial torque, on the other hand, is less straightforward as it depends on the initial conditions chosen for the model. Accretion, decretion or both can be obtained from the radial torque, despite the presence of a relatively strong positive laminar radial stress when the field is aligned with the vertical rotation axis, as in shearing box models (see § 9.2.2).

Despite the uncertainties, these models predict mass accretion rates $\dot{M}_{\rm acc} \sim 10^{-8} - 10^{-7} M_\odot/\rm yr$. However, even these values should be interpreted with care, as they are usually measured in the middle of the simulation domain, typically 5–20 AU from the central star. If an outflow is indeed present and carrying mass away, the mass accretion rate onto the star can be significantly smaller than that derived in the bulk of the disc. From the definition of the ejection efficiency (11.33), we have

$$\dot{M}_{\rm acc}(R) = \dot{M}_{\rm acc}(R_0) \left(\frac{R}{R_0} \right)^\xi. \tag{13.2}$$

Large ejection efficiencies ($\xi = O(1)$) such as that found in recent global models therefore lead to dramatically reduced mass accretion rates at the inner radius of the disc.

Unless one assumes an ionisation rate much higher than that expected in these objects (see § 3.4), the flow is mostly laminar, with a very low time-dependency. Hence, the radial stress measured in these models is not the usual turbulent stress found in ideal simulations, but really a purely magnetic term with no velocity counterpart. This implies that dust grains present in the disc will be less subject to turbulent fluctuations and will therefore settle towards the disc midplane more rapidly.

13.3. *Ejection and mass loss rate*

The outflow is not only responsible for carrying angular momentum away from the disc, but it also contributes significantly to the mass loss of the disc. All of the simulations published up to now find that the mass loss rate in the outflow is, broadly speaking, comparable with or even larger than the mass accretion rate in the disc ($\dot{M}_{\rm w} \gtrsim 10^{-8}$–$10^{-7}\,M_\odot/$yr), which implies $\xi \sim 1$.

The mass outflow rate is tightly connected to the amount of flux threading the disc, with $\dot{M}_{\rm w} \propto \beta_{\rm mid}^{-1/2}$ (Béthune *et al.* 2017), indicating that the mass flux is proportional to the magnetic flux threading the disc. Interestingly, similar scalings are obtained in non-ideal shearing box models (Bai & Stone 2013*b*; Lesur *et al.* 2014) whereas steeper dependences are found in ideal shearing box models (Suzuki & Inutsuka 2009; Bai & Stone 2013*a*).

The engine driving ejection can be isolated first by looking at the magnetic level arm of these outflows. Most of the models published up to now find level arms $\lambda < 2$. This is coherent with the very high mass loss rates found in these simulations (high ξ, see (11.40)). Although it is, in principle, possible to obtain cold outflows with low λ and high ξ in discs threaded by a weak field ($\beta_{\rm mean} \gg 1$, see for instance Jacquemin-Ide, Ferreira & Lesur 2019), some of the outflows published to date in PPDs have $\lambda < 3/2$ (e.g. Béthune *et al.* 2017), which violate the cold MHD wind constraint (11.29). Hence, these outflows are not purely magnetically driven.

This conclusion can also be reached by analysing directly the Bernoulli function of the outflow (e.g. figure 45). Such an analysis shows that thermal effects (enthalpy and heating terms) both contribute significantly to the energetics of the outflow (Béthune *et al.* 2017). Still, magnetic effects are clearly not negligible at the base of the outflow, where magnetic pressure helps pushing the flow upward.

For these reasons, these outflows have been labelled 'magneto-thermal'. This kind of outflow has already been identified in self-similar solutions by Casse & Ferreira (2000). Compared with historical cold wind solutions, they are (obviously) warmer, denser and slower. They reach high ξ values (typically $\xi > 0.1$) and have moderate λ. Of course, the fact that they extract angular momentum from the disc and that the initial acceleration is a result of magnetic effects implies that they are not purely thermal.

13.4. *Self-organisation*

Self-organisation was unambiguously identified in global simulations by Béthune *et al.* (2017) for simulations with $\beta_{\rm mid} \lesssim 10^3$ (e.g. figure 46), but it is absent from the models of Gressel *et al.* (2015) and Bai (2017), who only considered $\beta \gtrsim 10^4$. It is most of the time found in simulations exhibiting dissymmetric wind configurations, but does not seem to prefer one given field polarity, as would be expected from Hall-driven self-organisation. A careful examination of the flow shows that the poloidal field lines are concentrated in low-density regions. Hence, magnetic effects are playing a very important role in the mechanism.

It is possible to identify which process is responsible for self-organisation by looking closely at the non-ideal induction equation. It is then found that ambipolar diffusion
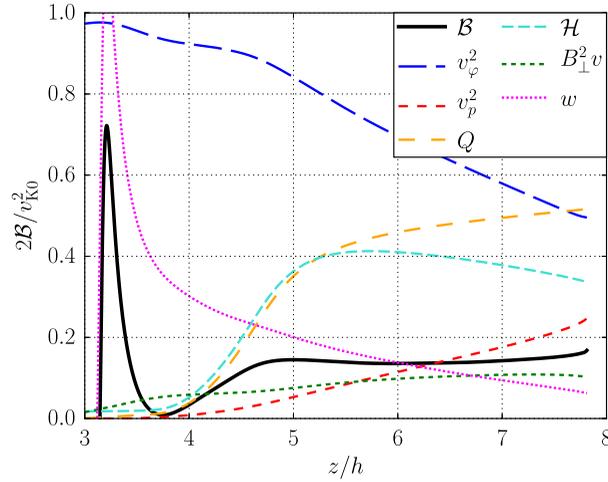
FIGURE 45. Bernoulli invariant $\mathcal{B}$ measured in an outflow driven from a non-ideal PPD. Magnetic contribution are $B_\perp^2 v$ and w, thermal contributions comes from the enthalpy $\mathcal{H}$ and external heating $\mathcal{Q}$ whereas kinetic energy terms are represented by v_ϕ^2 and v_p^2. Magnetic terms contribute significantly close to the launching point, whereas thermal energy becomes important higher up in the outflow. The ideal MHD region starts for $z \gtrsim 5h$. Figure from Béthune *et al.* (2017).
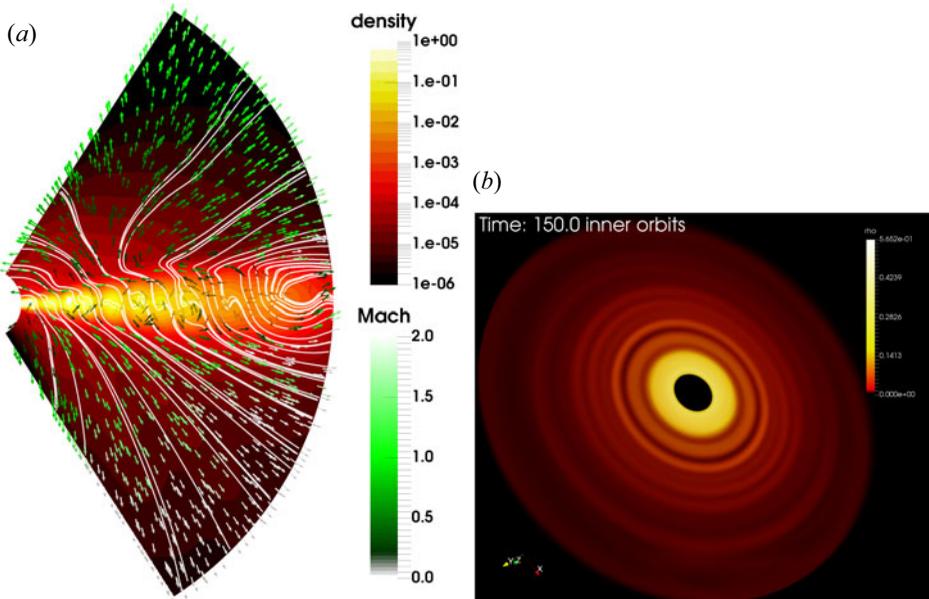

FIGURE 46. (*a*) Self-organisation in a simulation with $\beta_{\mathrm{mid}} = 10^2$ computed in 2.5 dimensions. The density is represented in colormap whereas magnetic field lines are in white lines and velocity field is shown in green arrows. Note that field lines are accumulated in regions of reduced density in the midplane. Figure from Béthune *et al.* (2017). (*b*) Volume rendering of a similar model, this time computed in the full three dimensions. Note that the flow remains axisymmetric.

is the only term responsible for the accumulation of magnetic flux in narrow regions, whereas Ohm and Hall effects are both diffusing the field away (Béthune *et al.* 2017). Hence, despite the presence of a rather strong Hall effect in these simulations, it is not the Hall-driven self-organisation that is at work in these models, but ambipolar-driven self-organisation. In essence, the mechanism seems to be similar to that driving self-organisation in stratified shearing box models subject to ambipolar diffusion only (Bai 2015). The local configuration found in the global simulations is indeed identical to the configuration found in shearing boxes (see figure 40), making the shearing box model a valuable tool to understand self-organisation in this regime.

Unfortunately, there is today no general theory predicting in which situation self-organisation is occurring nor what are the general properties of the structures that are formed.

14. Conclusions

In conclusion, the global modelling of the weakly ionised part of PPDs ($R \gtrsim 1$ AU) is still in its infancy. The inclusion of non-ideal MHD effects confirms that the disc midplane is mostly laminar, whereas the disc is still accreting thanks to magnetised outflows. Owing to the low magnetisations used in these models ($\beta_{\mathrm{mid}} \gtrsim 10^2$–$10^3$), the outflow is not purely magnetically driven as external heating and thermal pressure are found to contribute to the global energy budget. This implies that it is difficult to draw systematic conclusions regarding the accretion rate and mass loss rate, as these can depend both on the ionisation structure but also on the thermodynamics of the disc wind, both of which are plagued by huge uncertainties.

It would be tempting to consider stronger field models with $\beta = O(1)$ which are known to lead to historical 'cold wind' models (Ferreira 1997). However, such models would also have larger average accretion velocities which are typically sonic. If one wants to keep an accretion rate approximately compatible with observation, such a model would imply a disc surface density reduced by at least one order of magnitude (most likely two), which would be incompatible with the disc masses inferred from observations. Hence, the fact that the disc is massive, with an average accretion velocity much smaller than the speed of sound implies that if outflows exist in these regions, they must be due to a weak field: $\beta_{\mathrm{mean}} \gg 1$.

Self-organisation is also a very promising mechanism to explain some of the observables in PPDs. However, the theoretical background for these features remains limited. The fact that they are seen both in global and local simulations is encouraging, but, clearly, a detailed theoretical work is needed before any satisfactory prediction can be made and tested against observational data.

PART FIVE: Summary and future directions

15. Summary

These are exciting times for PPD modelling and planet formation theory in general. Indeed, we now start to have direct observational constraints for these astrophysical objects, ranging from large resolved structures, such as rings and non-axisymmetric bumps, to turbulent velocity dispersion measurements. Even magnetic field strength and topology are now beginning to be probed at the disc scale, giving more constraints to models.

The plasma in these discs is, however, relatively cold and therefore weakly ionised, reaching an ionisation fraction of $\xi \sim 10^{-13}$ in the disc midplane around 1 AU. This implies that non-ideal MHD effects (Ohmic, ambipolar and Hall effect) are essential to

obtaining a proper description of the plasma. The amplitude of these effects, however, is poorly constraint. Because ionisation is mostly non-thermal, the amplitude of these non-ideal effects depends on the details of the ionisation sources, the disc structure and the plasma composition (especially the abundance of tiny dust grains). Overall, there is an uncertainty of several orders of magnitude on these effects, implying that very detailed models including complex reaction networks are likely unnecessary at this stage because the input parameters (disc composition and environment) are largely unknown.

In order to explain the observed accretion rates in these discs, and because angular momentum is a conserved quantity, one needs to find a way to remove the disc angular momentum, either by transporting it radially outwards in the disc bulk, or by transporting it vertically away in a magnetised wind. Although radial transport has been historically favoured thanks to its simplicity and elegance (the well-known α disc model), vertical transport is now believed to be key in several astrophysical objects because of its high efficiency.

The most favoured mechanism to explain angular momentum transport in discs is the MRI, a linear, ideal MHD instability found in rotating sheared flows. This instability has been the subject of intense studies since the early 1990s and it is known that it is strongly affected by the non-ideal effects present in PPDs. Most notably, it is suppressed in the regions where Ohmic and ambipolar diffusion are strong, and it gives a new branch, known as the HSI, when the Hall effect is dominant, in the case where the poloidal field is aligned with the rotation axis.

In the non-linear regime, the MRI behaviour strongly depends on the presence and strength of a mean vertical field threading the disc. Historically, most of the simulations published until early 2010 were in a regime without a mean field, commonly known as the 'MRI dynamo'. In ideal MHD, this regime is known to produced vigorous 3D turbulence and radial transport of angular momentum, but in non-ideal MHD, turbulence is suppressed, leading to a laminar flow, no angular momentum transport and no accretion. This 'dead-zone' problem is circumvented by considering a mean vertical field threading the disc. Doing so, the MRI can indeed be revived in the upper layers, as expected from the linear analysis, but it then saturates into magnetised outflows, and the flow remains mostly laminar. In this situation, angular momentum is transported, mostly in the vertical direction, so accretion is saved, but its physical description then becomes fairly different from that of an α disc.

This connection between the MRI and magnetised outflow in discs threaded by a mean field was only realised during the past 10 years. In the presence of outflows, local models are insufficient since the dynamics of outflow is dictated by the global geometry of the system. In the case of PPDs, it is found that accretion driven by magnetised winds are compatible with observed accretion rates for relatively weak mean fields, $\beta_{\mathrm{mean}} = O(10^4)$, which are compatible with the upper bounds on the field strength from observations. Because the outflow is in a weakly magnetised regime, it is usually found that the mass loss rate can be of the order of, or even larger than, the mass accretion rate measured at the inner radius of the disc (this result being perfectly consistent with mass, angular momentum and energy conservation). In addition, the outflow is highly sensitive to thermal effects, which contribute significantly to the flow energetics, as is found in many models. Hence, these outflows have been called 'magneto-thermal'.

16. Perspectives

Research on this topic is now following several paths. First, the fact that thermal effects can play a significant role implies that they should be modelled accurately. Several groups are now working actively on this problem (e.g. Wang, Bai & Goodman 2019; Gressel

et al. 2020). It should be realised that thermal driving is a very complicated problem, as it involves heating by X-rays and UV photons, in addition to cooling, mostly by molecular and atomic lines. The computation of thermal processes therefore rely on complex chemical networks, coupled to radiative transfer codes, which are all computationally very intensive. Eventually, one hope is to find a way to simplify this physics using prescribed heating and cooling functions tested on complete models. This would allow a more systematic exploration of the long-term impact of thermodynamics on these systems, and make a connection to the winds observed in the sub-millimetric range.

A second question is the dynamical evolution of the mean magnetic field threading the disc. As shown previously, this mean field is key for magnetised outflows. It is strongly suspected that such a field should be present, as a direct result of the disc formation process, which relies on the collapse of a magnetised molecular cloud. During the collapse, a fraction of the magnetic field is trapped in the forming disc, and then plays the role of the mean field for outflows. However, once the disc is formed, it would be desirable to describe how this mean field evolves with time. It could be advected inwards by the accretion flow, leading to a strongly magnetised inner region, or it could inversely diffuse outwards because of non-ideal MHD effects. At the time of writing, this question is not settled, even qualitatively. Numerical models suggest that the field is diffusing outwards (Bai & Stone 2017; Gressel *et al.* 2020) whereas analytical models suggest inwards transport (Leung & Ogilvie 2019). If magnetised outflows are the dominant mechanism of accretion, then it is essential to address quantitatively this question in order to be able to model the long-term dynamics of these discs, because mass and magnetic flux are tightly linked. This flux transport can be at the origin of complex dynamics in the disc, such as time variability or even eruptions, which are also observed in these systems.

A third axis of research is the effect of this dynamics on planet formation, from the dynamics and growth of dust grains to the migration of giant planets. Most of the literature published to date rely on the α disc paradigm, which itself assumes that the disc is turbulent. However, if accretion is driven by magnetised winds in a mostly laminar flow, this framework has to be revised. Indeed, the lack of turbulence affects the dynamics of large grains ($\gtrsim 10\,\mu$m): vertical and radial settling, coagulation and disruption efficiency, etc., are all strongly modified. In addition, the fact that accretion is driven by surface stress, and not by turbulence is also going to reshape planet migration. It is not clear yet how type I and type II migration processes react to this shift in accretion paradigm, but the first attempts at including the wind stress in 2D planet migration numerical models already show a very significant effect (Kimmig, Dullemond & Kley 2020). Even the long-term evolution of the disc is quite different from that of a viscous disc, as viscous spreading is absent for a wind-driven disc. These are only a few example, but it shows that many things which were thought to be well established are now standing on wobbly foundations.

Editor Alex Schekochihin thanks the referees for their advice in evaluating this article.

Declaration of interest

The author reports no conflict of interest.

REFERENCES

ALUIE, H. & EYINK, G. L. 2010 Scale locality of magnetohydrodynamic turbulence. *Phys. Rev. Lett.* **104** (8), 081101.

ANDREWS, S. M., HUANG, J., PÉREZ, L. M., ISELLA, A., DULLEMOND, C. P., KURTOVIC, N. T., GUZMÁN, V. V., CARPENTER, J. M., WILNER, D. J., ZHANG, S., *et al.* 2018 The Disk Substructures at High Angular Resolution Project (DSHARP). I. Motivation, sample, calibration, and overview. *Astrophys. J. Lett.* **869** (2), L41.

ANDREWS, S. M., ROSENFELD, K. A., KRAUS, A. L. & WILNER, D. J. 2013 The mass dependence between protoplanetary disks and their stellar hosts. *Astrophys. J.* **771** (2), 129.

ANDREWS, S. M., WILNER, D. J., HUGHES, A. M., QI, C. & DULLEMOND, C. P. 2009 Protoplanetary disk structures in ophiuchus. *Astrophys. J.* **700**, 1502–1523.

ANDREWS, S. M., WILNER, D. J., ZHU, Z., BIRNSTIEL, T., CARPENTER, J. M., PÉREZ, L. M., BAI, X.-N., ÖBERG, K. I., HUGHES, A. M., ISELLA, A., *et al.* 2016 Ringed substructure and a gap at 1 au in the nearest protoplanetary disk. *Astrophys. J. Lett.* **820** (2), L40.

BAI, X.-N. 2011*a* Magnetorotational-instability-driven accretion in protoplanetary disks. *Astrophys. J.* **739**, 50.

BAI, X.-N. 2011*b* The role of tiny grains on the accretion process in protoplanetary disks. *Astrophys. J.* **739**, 51.

BAI, X.-N. 2014 Hall-effect-controlled gas dynamics in protoplanetary disks. I. Wind solutions at the inner disk. *Astrophys. J.* **791**, 137.

BAI, X.-N. 2015 Hall effect controlled gas dynamics in protoplanetary disks. II. Full 3D simulations toward the outer disk. *Astrophys. J.* **798**, 84.

BAI, X.-N. 2017 Global simulations of the inner regions of protoplanetary disks with comprehensive disk microphysics. *Astrophys. J.* **845**, 75.

BAI, X.-N. & GOODMAN, J. 2009 Heat and dust in active layers of protostellar disks. *Astrophys. J.* **701**, 737–755.

BAI, X.-N. & STONE, J. M. 2011 Effect of ambipolar diffusion on the nonlinear evolution of magnetorotational instability in weakly ionized disks. *Astrophys. J.* **736**, 144.

BAI, X.-N. & STONE, J. M. 2013*a* Local study of accretion disks with a strong vertical magnetic field: magnetorotational instability and disk outflow. *Astrophys. J.* **767**, 30.

BAI, X.-N. & STONE, J. M. 2013*b* Wind-driven accretion in protoplanetary disks. I. Suppression of the magnetorotational instability and launching of the magnetocentrifugal wind. *Astrophys. J.* **769**, 76.

BAI, X.-N. & STONE, J. M. 2014 Magnetic flux concentration and zonal flows in magnetorotational instability turbulence. *Astrophys. J.* **796**, 31.

BAI, X.-N. & STONE, J. M. 2017 Hall effect-mediated magnetic flux transport in protoplanetary disks. *Astrophys. J.* **836** (1), 46.

BALBUS, S. A., GAMMIE, C. F. & HAWLEY, J. F. 1994 Fluctuations, dissipation and turbulence in accretion discs. *Mon. Not. R. Astron. Soc.* **271**, 197.

BALBUS, S. A. & HAWLEY, J. F. 1991 A powerful local shear instability in weakly magnetized disks. I – linear analysis. II – nonlinear evolution. *Astrophys. J.* **376**, 214–233.

BALBUS, S. A. & HAWLEY, J. F. 1992 A powerful local shear instability in weakly magnetized disks. IV. Nonaxisymmetric perturbations. *Astrophys. J.* **400**, 610–621.

BALBUS, S. A. & HAWLEY, J. F. 1998 Instability, turbulence, and enhanced transport in accretion disks. *Rev. Mod. Phys.* **70**, 1–53.

BALBUS, S. A. & PAPALOIZOU, J. C. B. 1999 On the dynamical foundations of α disks. *Astrophys. J.* **521** (2), 650–658.

BALBUS, S. A. & TERQUEM, C. 2001 Linear analysis of the Hall effect in protostellar disks. *Astrophys. J.* **552**, 235–247.

BENISTY, M., JUHASZ, A., BOCCALETTI, A., AVENHAUS, H., MILLI, J., THALMANN, C., DOMINIK, C., PINILLA, P., BUENZLI, E., POHL, A., *et al.* 2015 Asymmetric features in the protoplanetary disk MWC 758. *Astron. Astrophys.* **578**, L6.

BÉTHUNE, W., LESUR, G. & FERREIRA, J. 2016 Self-organisation in protoplanetary discs. Global, non-stratified Hall-MHD simulations. *Astron. Astrophys.* **589**, A87.

BÉTHUNE, W., LESUR, G. & FERREIRA, J. 2017 Global simulations of protoplanetary disks with net magnetic flux. I. Non-ideal MHD case. *Astron. Astrophys.* **600**, A75.

BIERSTEKER, J. B., WEISS, B. P., HEINISCH, P., HERČIK, D., GLASSMEIER, K.-H & AUSTER, H.-U. 2019 Implications of philae magnetometry measurements at comet 67P/Churyumov-Gerasimenko for the nebular field of the outer solar system. *Astrophys. J.* **875** (1), 39.

BLAES, O. M. & BALBUS, S. A. 1994 Local shear instabilities in weakly ionized, weakly magnetized disks. *Astrophys. J.* **421**, 163–177.

BLANDFORD, R. D. & PAYNE, D. G. 1982 Hydromagnetic flows from accretion discs and the production of radio jets. *Mon. Not. R. Astron. Soc.* **199**, 883–903.

BODO, G., CATTANEO, F., FERRARI, A., MIGNONE, A. & ROSSI, P. 2011 Symmetries, scaling laws, and convergence in shearing-box simulations of magneto-rotational instability driven turbulence. *Astrophys. J.* **739** (2), 82.

BODO, G., CATTANEO, F., MIGNONE, A. & ROSSI, P. 2013 Fully convective magnetorotational turbulence in stratified shearing boxes. *Astrophys. J. Lett.* **771**, L23.

BODO, G., CATTANEO, F., MIGNONE, A. & ROSSI, P. 2014 On the convergence of magnetorotational turbulence in stratified isothermal shearing boxes. *Astrophys. J. Lett.* **787**, L13.

BODO, G., CATTANEO, F., MIGNONE, A. & ROSSI, P. 2015 Fully convective magneto-rotational turbulence in large aspect-ratio shearing boxes. *Astrophys. J.* **799**, 20.

BODO, G., MIGNONE, A., CATTANEO, F., ROSSI, P. & FERRARI, A. 2008 Aspect ratio dependence in magnetorotational instability shearing box simulations. *Astron. Astrophys.* **487**, 1–5.

BRANDENBURG, A., NORDLUND, A., STEIN, R. F. & TORKELSSON, U. 1995 Dynamo-generated turbulence and large-scale magnetic fields in a Keplerian shear flow. *Astrophys. J.* **446**, 741.

BURROWS, C. J., STAPELFELDT, K. R., WATSON, A. M., KRIST, J. E., BALLESTER, G. E., CLARKE, J. T., CRISP, D., GALLAGHER, J. S. III, Griffiths, R. E., *et al.* 1996 Hubble space telescope observations of the disk and jet of HH 30. *Astrophys. J.* **473**, 437.

CASSE, F. & FERREIRA, J. 2000 Magnetized accretion-ejection structures. V. Effects of entropy generation inside the disc. *Astron. Astrophys.* **361**, 1178–1190.

CHO, J. & LAZARIAN, A. 2007 Grain alignment and polarized emission from magnetized T Tauri disks. *Astrophys. J.* **669**, 1085–1097.

CLEEVES, L. I., ADAMS, F. C. & BERGIN, E. A. 2013 Exclusion of cosmic rays in protoplanetary disks: stellar and magnetic effects. *Astrophys. J.* **772**, 5.

CLEEVES, L. I., BERGIN, E. A., QI, C., ADAMS, F. C. & ÖBERG, K. I. 2015 Constraining the x-ray and cosmic-ray ionization chemistry of the TW Hya protoplanetary disk: evidence for a sub-interstellar cosmic-ray rate. *Astrophys. J.* **799**, 204.

CONTOPOULOS, J. & LOVELACE, R. V. E. 1994 Magnetically driven jets and winds: exact solutions. *Astrophys. J.* **429**, 139–152.

CRAIK, A. D. D. & CRIMINALE, W. O. 1986 Evolution of wavelike disturbances in shear flows: a class of exact solutions of the Navier–Stokes equations. *Proc. R. Soc. Lond.* A **406**, 13–26.

CURRY, C. & PUDRITZ, R. E. 1996 On the global stability of magnetized accretion discs – III. Non-axisymmetric modes. *Mon. Not. R. Astron. Soc.* **281** (1), 119–136.

D'ALESSIO, P., CANTÖ, J., CALVET, N. & LIZANO, S. 1998 Accretion disks around young objects. I. The detailed vertical structure. *Astrophys. J.* **500** (1), 411–427.

DAVIS, S. W., STONE, J. M. & PESSAH, M. E. 2010 Sustained magnetorotational turbulence in local simulations of stratified disks with zero net magnetic flux. *Astrophys. J.* **713**, 52–65.

DESCH, S. J. 2004 Linear analysis of the magnetorotational instability, including ambipolar diffusion, with application to protoplanetary disks. *Astrophys. J.* **608**, 509–525.

DONATI, J.-F., PALETOU, F., BOUVIER, J. & FERREIRA, J. 2005 Direct detection of a magnetic field in the innermost regions of an accretion disk. *Nature* **438**, 466.

DRAINE, B. T. 2011 *Physics of the Interstellar and Intergalactic Medium*. Princeton University Press.

DRAINE, B. T., ROBERGE, W. G. & DALGARNO, A. 1983 Magnetohydrodynamic shock waves in molecular clouds. *Astrophys. J.* **264**, 485–507.

DZYURKEVICH, N., TURNER, N. J., HENNING, T. & KLEY, W. 2013 Magnetized accretion and dead zones in protostellar disks. *Astrophys. J.* **765**, 114.

ELMEGREEN, B. G. 1979 Magnetic diffusion and ionization fractions in dense molecular clouds: the role of charged grains. *Astrophys. J.* **232**, 729–739.

FEDELE, D., VAN DEN ANCKER, M. E., HENNING, T., JAYAWARDHANA, R. & OLIVEIRA, J. M. 2010 Timescale of mass accretion in pre-main-sequence stars. *Astron. Astrophys.* **510**, A72.

FERREIRA, J. 1997 Magnetically-driven jets from Keplerian accretion discs. *Astron. Astrophys.* **319**, 340–359.

FERREIRA, J. & PELLETIER, G. 1993 Magnetized accretion-ejection structures. 1. General statements. *Astron. Astrophys.* **276**, 625.

FLAHERTY, K. M., HUGHES, A. M., ROSE, S. C., SIMON, J. B., QI, C., ANDREWS, S. M., KÓSPÁL, Á., WILNER, D. J., CHIANG, E., ARMITAGE, P. J., *et al.* 2017 A three-dimensional view of turbulence: constraints on turbulent motions in the HD 163296 protoplanetary disk using DCO$^+$. *Astrophys. J.* **843**, 150.

FLAHERTY, K. M., HUGHES, A. M., ROSENFELD, K. A., ANDREWS, S. M., CHIANG, E., SIMON, J. B., KERZNER, S. & WILNER, D. J. 2015 Weak turbulence in the HD 163296 protoplanetary disk revealed by ALMA CO observations. *Astrophys. J.* **813** (2), 99.

FLAIG, M., KLEY, W. & KISSMANN, R. 2010 Vertical structure and turbulent saturation level in fully radiative protoplanetary disc models. *Mon. Not. R. Astron. Soc.* **409**, 1297–1306.

FLAIG, M., RUOFF, P., KLEY, W. & KISSMANN, R. 2012 Global structure of magnetorotationally turbulent protoplanetary discs. *Mon. Not. R. Astron. Soc.* **420**, 2419–2428.

FLEMING, T. & STONE, J. M. 2003 Local magnetohydrodynamic models of layered accretion disks. *Astrophys. J.* **585**, 908–920.

FLEMING, T. P., STONE, J. M. & HAWLEY, J. F. 2000 The effect of resistivity on the nonlinear stage of the magnetorotational instability in accretion disks. *Astrophys. J.* **530**, 464–477.

FLOCK, M., FROMANG, S., GONZÁLEZ, M. & COMMERÇON, B. 2013 Radiation magnetohydrodynamics in global simulations of protoplanetary discs. *Astron. Astrophys.* **560**, A43.

FRANK, A., RAY, T. P., CABRIT, S., HARTIGAN, P., ARCE, H. G., BACCIOTTI, F., BALLY, J., BENISTY, M., EISLÖFFEL, J., GÜDEL, M., *et al.* 2014 Jets and outflows from star to cloud: observations confront theory. In *Protostars and Planets VI* (ed. H. Beuther, R. S. Klessen, C. P. Dullemond, & T. Henning), pp. 451–474. University of Arizona Press.

FROMANG, S. 2010 MHD simulations of the magnetorotational instability in a shearing box with zero net flux: the case *Pm* = 4. *Astron. Astrophys.* **514**, L5.

FROMANG, S., LATTER, H., LESUR, G. & OGILVIE, G. I. 2013 Local outflows from turbulent accretion disks. *Astron. Astrophys.* **552**, A71.

FROMANG, S. & LESUR, G. 2019 Angular momentum transport in accretion disks: a hydrodynamical perspective. In *EAS Publications Series* (ed. A. S. Brun, S. Mathis, C. Charbonnel & B. Dubrulle), vol. 82, pp. 391–413.

FROMANG, S. & NELSON, R. P. 2005 On the accumulation of solid bodies in global turbulent protoplanetary disc models. *Mon. Not. R. Astron. Soc.* **364**, L81–L85.

FROMANG, S. & PAPALOIZOU, J. 2007 MHD simulations of the magnetorotational instability in a shearing box with zero net flux. I. The issue of convergence. *Astron. Astrophys.* **476**, 1113–1122.

FROMANG, S., PAPALOIZOU, J., LESUR, G. & HEINEMANN, T. 2007 MHD simulations of the magnetorotational instability in a shearing box with zero net flux. II. The effect of transport coefficients. *Astron. Astrophys.* **476**, 1123–1132.

FROMANG, S. & STONE, J. M. 2009 Turbulent resistivity driven by the magnetorotational instability. *Astron. Astrophys.* **507**, 19–28.

FROMANG, S., TERQUEM, C. & BALBUS, S. A. 2002 The ionization fraction in α models of protoplanetary discs. *Mon. Not. R. Astron. Soc.* **329**, 18–28.

FU, R. R., WEISS, B. P., LIMA, E. A., HARRISON, R. J., BAI, X.-N., DESCH, S. J., EBEL, D. S., SUAVET, C., WANG, H., GLENN, D., *et al.* 2014 Solar nebula magnetic fields recorded in the Semarkona meteorite. *Science* **346** (6213), 1089–1092.

GAMMIE, C. F. 1996 Layered accretion in T Tauri disks. *Astrophys. J.* **457**, 355.

GINSKI, C., STOLKER, T., PINILLA, P., DOMINIK, C., BOCCALETTI, A., DE BOER, J., BENISTY, M., BILLER, B., FELDT, M., GARUFI, A., *et al.* 2016 Direct detection of scattered light gaps in the transitional disk around HD 97048 with VLT/SPHERE. arXiv:1609.04027.

GOLDREICH, P. & LYNDEN-BELL, D. 1965 II. Spiral arms as sheared gravitational instabilities. *Mon. Not. R. Astron. Soc.* **130**, 125.

GOODMAN, J. & XU, G. 1994 Parasitic instabilities in magnetized, differentially rotating disks. *Astrophys. J.* **432**, 213–223.

GRESSEL, O. 2010 A mean-field approach to the propagation of field patterns in stratified magnetorotational turbulence. *Mon. Not. R. Astron. Soc.* **405**, 41–48.

GRESSEL, O. & PESSAH, M. E. 2015 Characterizing the mean-field dynamo in turbulent accretion disks. *Astrophys. J.* **810**, 59.

GRESSEL, O., RAMSEY, J. P., BRINCH, C., NELSON, R. P., TURNER, N. J. & BRUDERER, S. 2020 Global hydromagnetic simulations of protoplanetary disks with stellar irradiation and simplified thermochemistry. arXiv:2005.03431.

GRESSEL, O., TURNER, N. J., NELSON, R. P. & MCNALLY, C. P. 2015 Global simulations of protoplanetary disks with ohmic resistivity and ambipolar diffusion. *Astrophys. J.* **801**, 84.

GUAN, X. & GAMMIE, C. F. 2009 The turbulent magnetic prandtl number of MHD turbulence in disks. *Astrophys. J.* **697**, 1901–1906.

HAWLEY, J. F. 2001 Global magnetohydrodynamic simulations of cylindrical Keplerian disks. *Astrophys. J.* **554**, 534–547.

HAWLEY, J. F., GAMMIE, C. F. & BALBUS, S. A. 1995 Local three-dimensional magnetohydrodynamic simulations of accretion disks. *Astrophys. J.* **440**, 742.

HAWLEY, J. F., GAMMIE, C. F. & BALBUS, S. A. 1996 Local three-dimensional simulations of an accretion disk hydromagnetic dynamo. *Astrophys. J.* **464**, 690.

HAWLEY, J. F. & STONE, J. M. 1998 Nonlinear evolution of the magnetorotational instability in ion-neutral disks. *Astrophys. J.* **501**, 758–771.

HERAULT, J., RINCON, F., COSSU, C., LESUR, G., OGILVIE, G. I. & LONGARETTI, P.-Y. 2011 Periodic magnetorotational dynamo action as a prototype of nonlinear magnetic-field generation in shear flows. *Phys. Rev.* E **84** (3), 036321.

HILL, G. W. 1878 Researches in the lunar theory. *Am. J. Maths.* **1**, 5–26.

HIROSE, S., BLAES, O. & KROLIK, J. H. 2009 Turbulent stresses in local simulations of radiation-dominated accretion disks, and the possibility of the Lightman–Eardley instability. *Astrophys. J.* **704**, 781–788.

HIROSE, S., BLAES, O., KROLIK, J. H., COLEMAN, M. S. B. & SANO, T. 2014 Convection causes enhanced magnetic turbulence in accretion disks in outburst. *Astrophys. J.* **787**, 1.

HIROSE, S., KROLIK, J. H. & STONE, J. M. 2006 Vertical structure of gas pressure-dominated accretion disks with local dissipation of turbulence and radiative transport. *Astrophys. J.* **640**, 901–917.

HOGG, J. D. & REYNOLDS, C. S. 2016 Testing the propagating fluctuations model with a long, global accretion disk simulation. *Astrophys. J.* **826**, 40.

HOLLERBACH, R. & RÜDIGER, G. 2005 New type of magnetorotational instability in cylindrical Taylor–Couette flow. *Phys. Rev. Lett.* **95** (12), 124501.

IGEA, J. & GLASSGOLD, A. E. 1999 X-ray ionization of the disks of young stellar objects. *Astrophys. J.* **518**, 848–858.

ILGNER, M. 2012 Grain charging in protoplanetary discs. *Astron. Astrophys.* **538**, A124.

ILGNER, M. & NELSON, R. P. 2006 On the ionisation fraction in protoplanetary disks. I. Comparing different reaction networks. *Astron. Astrophys.* **445**, 205–222.

ILGNER, M. & NELSON, R. P. 2008 Turbulent transport and its effect on the dead zone in protoplanetary discs. *Astron. Astrophys.* **483** (3), 815–830.

JACQUEMIN-IDE, J., FERREIRA, J. & LESUR, G. 2019 Magnetically driven jets and winds from weakly magnetized accretion discs. *Mon. Not. R. Astron. Soc.* **490** (3), 3112–3133.

JIN, L. 1996 Damping of the shear instability in magnetized disks by Ohmic diffusion. *Astrophys. J.* **457**, 798.

JOHANSEN, A., YOUDIN, A. & KLAHR, H. 2009 Zonal flows and long-lived axisymmetric pressure bumps in magnetorotational turbulence. *Astrophys. J.* **697**, 1269–1289.

KATAOKA, A., MUTO, T., MOMOSE, M., TSUKAGOSHI, T., FUKAGAWA, M., SHIBAI, H., HANAWA, T., MURAKAWA, K. & DULLEMOND, C. P. 2015 Millimeter-wave polarization of protoplanetary disks due to dust scattering. *Astrophys. J.* **809**, 78.

KELVIN, L. 1880 Vibrations of a columnar vortex. *Phil. Mag.* **10**, 155–68.

KIM, W.-T. & OSTRIKER, E. C. 2000 Magnetohydrodynamic instabilities in shearing, rotating, stratified winds and disks. *Astrophys. J.* **540**, 372–403.

KIMMIG, C. N., DULLEMOND, C. P. & KLEY, W. 2020 Effect of wind-driven accretion on planetary migration. *Astron. Astrophys.* **633**, A4.

KIRILLOV, O. N., STEFANI, F. & FUKUMOTO, Y. 2014 Local instabilities in magnetized rotational flows: a short-wavelength approach. *J. Fluid Mech.* **760**, 591–633.

KONIGL, A. 1989 Self-similar models of magnetized accretion disks. *Astrophys. J.* **342**, 208–223.

KUNZ, M. W. 2008 On the linear stability of weakly ionized, magnetized planar shear flows. *Mon. Not. R. Astron. Soc.* **385**, 1494–1510.

KUNZ, M. W. & BALBUS, S. A. 2004 Ambipolar diffusion in the magnetorotational instability. *Mon. Not. R. Astron. Soc.* **348**, 355–360.

KUNZ, M. W. & LESUR, G. 2013 Magnetic self-organization in Hall-dominated magnetorotational turbulence. *Mon. Not. R. Astron. Soc.* **434**, 2295–2312.

LATTER, H. N., FROMANG, S. & GRESSEL, O. 2010 MRI channel flows in vertically stratified models of accretion discs. *Mon. Not. R. Astron. Soc.* **406**, 848–862.

LATTER, H. N., LESAFFRE, P. & BALBUS, S. A. 2009 MRI channel flows and their parasites. *Mon. Not. R. Astron. Soc.* **394**, 715–729.

LESUR, G., FERREIRA, J. & OGILVIE, G. I. 2013 The magnetorotational instability as a jet launching mechanism. *Astron. Astrophys.* **550**, A61.

LESUR, G., KUNZ, M. W. & FROMANG, S. 2014 Thanatology in protoplanetary discs. The combined influence of Ohmic, Hall, and ambipolar diffusion on dead zones. *Astron. Astrophys.* **566**, A56.

LESUR, G. & LONGARETTI, P.-Y. 2007 Impact of dimensionless numbers on the efficiency of magnetorotational instability induced turbulent transport. *Mon. Not. R. Astron. Soc.* **378**, 1471–1480.

LESUR, G. & LONGARETTI, P.-Y. 2009 Turbulent resistivity evaluation in magnetorotational instability generated turbulence. *Astron. Astrophys.* **504**, 309–320.

LESUR, G. & LONGARETTI, P.-Y. 2011 Non-linear energy transfers in accretion discs MRI turbulence. I. Net vertical field case. *Astron. Astrophys.* **528**, A17.

LESUR, G. & OGILVIE, G. I. 2008*a* Localized magnetorotational instability and its role in the accretion disc dynamo. *Mon. Not. R. Astron. Soc.* **391**, 1437–1450.

LESUR, G. & OGILVIE, G. I. 2008*b* On self-sustained dynamo cycles in accretion discs. *Astron. Astrophys.* **488**, 451–461.

LEUNG, P. K. C. & OGILVIE, G. I. 2019 Local semi-analytic models of magnetic flux transport in protoplanetary discs. *Mon. Not. R. Astron. Soc.* **487** (4), 5155–5174.

LIU, W., GOODMAN, J., HERRON, I. & JI, H. 2006 Helical magnetorotational instability in magnetized Taylor–Couette flow. *Phys. Rev.* E **74** (5), 056302.

LONGARETTI, P.-Y. & LESUR, G. 2010 MRI-driven turbulent transport: the role of dissipation, channel modes and their parasites. *Astron. Astrophys.* **516**, A51.

LOUVET, F., DOUGADOS, C., CABRIT, S., MARDONES, D., MÉNARD, F., TABONE, B., PINTE, C. & DENT, W. R. F. 2018 The HH30 edge-on T Tauri star. A rotating and precessing monopolar outflow scrutinized by ALMA. *Astron. Astrophys.* **618**, A120.

LOVELACE, R. V. E., LI, H., COLGATE, S. A. & NELSON, A. F. 1999 Rossby wave instability of Keplerian accretion disks. *Astrophys. J.* **513**, 805–810.

VAN DER MAREL, N., VAN DISHOECK, E. F., BRUDERER, S., BIRNSTIEL, T., PINILLA, P., DULLEMOND, C. P., VAN KEMPEN, T. A., SCHMALZL, M., BROWN, J. M., HERCZEG, G. J., *et al.* 2013 A major asymmetric dust trap in a transition disk. *Science* **340**, 1199–1202.

MCNALLY, C. P. & PESSAH, M. E. 2015 On vertically global, horizontally local models for astrophysical disks. *Astrophys. J.* **811**, 121.

MOLL, R. 2012 Shearing box simulations of accretion disk winds. *Astron. Astrophys.* **548**, A76.

OGILVIE, G. I. 2012 Jet launching from accretion discs in the local approximation. *Mon. Not. R. Astron. Soc.* **423**, 1318–1324.

O'KEEFFE, W. & DOWNES, T. P. 2014 Multifluid simulations of the magnetorotational instability in protostellar discs. *Mon. Not. R. Astron. Soc.* **441**, 571–581.

OPPENHEIMER, M. & DALGARNO, A. 1974 The fractional ionization in dense interstellar clouds. *Astrophys. J.* **192**, 29–32.

PADOVANI, M., IVLEV, A. V., GALLI, D. & CASELLI, P. 2018 Cosmic-ray ionisation in circumstellar discs. *Astron. Astrophys.* **614**, A111.

PANDEY, B. P. & WARDLE, M. 2012 Magnetorotational instability in magnetic diffusion dominated accretion discs. *Mon. Not. R. Astron. Soc.* **423**, 222–235.

PARTNERSHIP, A., BROGAN, C. L., PEREZ, L. M., HUNTER, T. R., DENT, W. R. F., HALES, A. S., HILLS, R. E., CORDER, S., FOMALONT, E. B., VLAHAKIS, C., *et al.* 2015 The 2014 ALMA long baseline campaign: first results from high angular resolution observations toward the HL tau region. *Astrophys. J. Lett.* **808** (1), L3.

PÉREZ, L. M., CARPENTER, J. M., ANDREWS, S. M., RICCI, L., ISELLA, A., LINZ, H., SARGENT, A. I., WILNER, D. J., HENNING, T., DELLER, A. T., *et al.* 2016 Spiral density waves in a young protoplanetary disk. *Science* **353** (6307), 1519–1521.

PEREZ-BECKER, D. & CHIANG, E. 2011*a* Surface layer accretion in conventional and transitional disks driven by far-ultraviolet ionization. *Astrophys. J.* **735**, 8.

PEREZ-BECKER, D. & CHIANG, E. 2011*b* Surface layer accretion in transitional and conventional disks: from polycyclic aromatic hydrocarbons to planets. *Astrophys. J.* **727**, 2.

PINTE, C., DENT, W. R. F., MÉNARD, F., HALES, A., HILL, T., CORTES, P. & DE GREGORIO-MONSALVO, I. 2016 Dust and gas in the disk of HL Tauri: surface density, dust settling, and dust-to-gas ratio. *Astrophys. J.* **816**, 25.

POTTER, W. J. & BALBUS, S. A. 2017 Demonstration of a magnetic Prandtl number disc instability from first principles. arXiv:1704.02485.

RINCON, F., OGILVIE, G. I., PROCTOR, M. R. E. & COSSU, C. 2008 Subcritical dynamos in shear flows. *Astron. Nachr.* **329**, 750.

RIOLS, A. & LESUR, G. 2019 Spontaneous ring formation in wind-emitting accretion discs. *Astron. Astrophys.* **625**, A108.

RIOLS, A., RINCON, F., COSSU, C., LESUR, G., LONGARETTI, P.-Y., OGILVIE, G. I. & HERAULT, J. 2013 Global bifurcations to subcritical magnetorotational dynamo action in Keplerian shear flow. *J. Fluid Mech.* **731**, 1–45.

RIOLS, A., RINCON, F., COSSU, C., LESUR, G., OGILVIE, G. I. & LONGARETTI, P. Y. 2015 Dissipative effects on the sustainment of a magnetorotational dynamo in Keplerian shear flow. *Astron. Astrophys.* **575**, A14.

RODGERS-LEE, D., RAY, T. P. & DOWNES, T. P. 2016 Global multifluid simulations of the magnetorotational instability in radially stratified protoplanetary discs. *Mon. Not. R. Astron. Soc.* **463**, 134–145.

RYAN, B. R., GAMMIE, C. F., FROMANG, S. & KESTENER, P. 2017 Resolution dependence of magnetorotational turbulence in the isothermal stratified shearing box. *Astrophys. J.* **840**, 6.

SALMERON, R. & WARDLE, M. 2008 Magnetorotational instability in protoplanetary discs: the effect of dust grains. *Mon. Not. R. Astron. Soc.* **388**, 1223–1238.

SALVESEN, G., SIMON, J. B., ARMITAGE, P. J. & BEGELMAN, M. C. 2016 Accretion disc dynamo activity in local simulations spanning weak-to-strong net vertical magnetic flux regimes. *Mon. Not. R. Astron. Soc.* **457**, 857–874.

SANO, T., INUTSUKA, S.-I. & MIYAMA, S. M. 1998 A saturation mechanism of magnetorotational instability due to ohmic dissipation. *Astrophys. J. Lett.* **506** (1), L57–L60.

SANO, T. & MIYAMA, S. M. 1999 Magnetorotational instability in protoplanetary disks. I. On the global stability of weakly ionized disks with ohmic dissipation. *Astrophys. J.* **515** (2), 776–786.

SANO, T., MIYAMA, S. M., UMEBAYASHI, T. & NAKANO, T. 2000 Magnetorotational instability in protoplanetary disks. II. Ionization state and unstable regions. *Astrophys. J.* **543**, 486–501.

SANO, T. & STONE, J. M. 2002 The effect of the Hall term on the nonlinear evolution of the magnetorotational instability. II. Saturation level and critical magnetic Reynolds number. *Astrophys. J.* **577**, 534–553.

SCEPI, N., LESUR, G., DUBUS, G. & FLOCK, M. 2018 Turbulent and wind-driven accretion in dwarf novae threaded by a large-scale magnetic field. *Astron. Astrophys.* **620**, A49.

SHAKURA, N. I. & SUNYAEV, R. A. 1973 Black holes in binary systems. Observational appearance. *Astron. Astrophys.* **24**, 337–355.

SIMON, J. B. & ARMITAGE, P. J. 2014 Efficiency of particle trapping in the outer regions of protoplanetary disks. *Astrophys. J.* **784** (1), 15.

SIMON, J. B., BAI, X.-N., ARMITAGE, P. J., STONE, J. M. & BECKWITH, K. 2013a Turbulence in the outer regions of protoplanetary disks. II. Strong accretion driven by a vertical magnetic field. *Astrophys. J.* **775**, 73.

SIMON, J. B., BAI, X.-N., FLAHERTY, K. M. & HUGHES, A. M. 2018 Origin of weak turbulence in the outer regions of protoplanetary disks. *Astrophys. J.* **865** (1), 10.

SIMON, J. B., BAI, X.-N., STONE, J. M., ARMITAGE, P. J. & BECKWITH, K. 2013b Turbulence in the outer regions of protoplanetary disks. I. Weak accretion with no vertical magnetic flux. *Astrophys. J.* **764**, 66.

SIMON, J. B., BECKWITH, K. & ARMITAGE, P. J. 2012 Emergent mesoscale phenomena in magnetized accretion disc turbulence. *Mon. Not. R. Astron. Soc.* **422**, 2685–2700.

SIMON, J. B. & HAWLEY, J. F. 2009 Viscous and resistive effects on the magnetorotational instability with a net toroidal field. *Astrophys. J.* **707**, 833–843.

SIMON, J. B., LESUR, G., KUNZ, M. W. & ARMITAGE, P. J. 2015 Magnetically driven accretion in protoplanetary discs. *Mon. Not. R. Astron. Soc.* **454**, 1117–1131.

STEINACKER, A. & PAPALOIZOU, J. C. B. 2002 Three-dimensional magnetohydrodynamic simulations of an accretion disk with star-disk boundary layer. *Astrophys. J.* **571**, 413–428.

STEPHENS, I. W., LOONEY, L. W., KWON, W., FERNÁNDEZ-LÓPEZ, M., HUGHES, A. M., MUNDY, L. G., CRUTCHER, R. M., LI, Z.-Y. & RAO, R. 2014 Spatially resolved magnetic field structure in the disk of a T Tauri star. *Nature* **514**, 597–599.

STEPHENS, I. W., YANG, H., LI, Z.-Y., LOONEY, L. W., KATAOKA, A., KWON, W., FERNÁNDEZ-LÓPEZ, M., HULL, C. L. H., HUGHES, M., SEGURA-COX, D., *et al.* 2017 ALMA reveals transition of polarization pattern with wavelength in HL tau disk. *Astrophys. J.* **851**, 55.

STONE, J. M., HAWLEY, J. F., GAMMIE, C. F. & BALBUS, S. A. 1996 Three-dimensional magnetohydrodynamical simulations of vertically stratified accretion disks. *Astrophys. J.* **463**, 656.

SUZUKI, T. K. & INUTSUKA, S.-I. 2009 Disk winds driven by magnetorotational instability and dispersal of protoplanetary disks. *Astrophys. J. Lett.* **691**, L49–L54.

SUZUKI, T. K., MUTO, T. & INUTSUKA, S.-I. 2010 Protoplanetary disk winds via magnetorotational instability: formation of an inner hole and a crucial assist for planet formation. *Astrophys. J.* **718**, 1289–1304.

TEAGUE, R., BAE, J. & BERGIN, E. A. 2019 Meridional flows in the disk around a young star. *Nature* **574** (7778), 378–381.

THI, W. F., LESUR, G., WOITKE, P., KAMP, I., RAB, C. & CARMONA, A. 2019 Radiation thermo-chemical models of protoplanetary disks. Grain and polycyclic aromatic hydrocarbon charging. *Astron. Astrophys.* **632**, A44.

UMEBAYASHI, T. & NAKANO, T. 1980 Recombination of ions and electrons on grains and the ionization degree in dense interstellar clouds. *Publ. Astron. Soc. Jpn.* **32**, 405.

UMEBAYASHI, T. & NAKANO, T. 1981 Fluxes of energetic particles and the ionization rate in very dense interstellar clouds. *Publ. Astron. Soc. Jpn.* **33**, 617.

UMEBAYASHI, T. & NAKANO, T. 2009 Effects of radionuclides on the ionization state of protoplanetary disks and dense cloud cores. *Astrophys. J.* **690**, 69–81.

UMURHAN, O. M. & REGEV, O. 2004 Hydrodynamic stability of rotationally supported flows: linear and nonlinear 2D shearing box results. *Astron. Astrophys.* **427**, 855–872.

VENUTI, L., BOUVIER, J., FLACCOMIO, E., ALENCAR, S. H. P., IRWIN, J., STAUFFER, J. R., CODY, A. M., TEIXEIRA, P. S., SOUSA, A. P., MICELA, G., *et al.* 2014 Mapping accretion and its variability in the young open cluster NGC 2264: a study based on u-band photometry. *Astron. Astrophys*. **570**, A82.

VLEMMINGS, W. H. T., LANKHAAR, B., CAZZOLETTI, P., CECCOBELLO, C., DALL'OLIO, D., VAN DISHOECK, E. F., FACCHINI, S., HUMPHREYS, E. M. L., PERSSON, M. V., TESTI, L., *et al.* 2019 Stringent limits on the magnetic field strength in the disc of TW Hya. ALMA observations of CN polarisation. *Astron. Astrophys*. **624**, L7.

WALKER, J. & BOLDYREV, S. 2017 Magnetorotational dynamo action in the shearing box. arXiv:1704.08636.

WANG, L., BAI, X.-N. & GOODMAN, J. 2019 Global simulations of protoplanetary disk outflows with coupled non-ideal magnetohydrodynamics and consistent thermochemistry. *Astrophys. J*. **874** (1), 90.

WARDLE, M. 1999 The Balbus–Hawley instability in weakly ionized discs. *Mon. Not. R. Astron. Soc*. **307**, 849–856.

WARDLE, M. 2007 Magnetic fields in protoplanetary disks. *Astrophys. J. Suppl*. **311**, 35–45.

WARDLE, M. & KONIGL, A. 1993 The structure of protostellar accretion disks and the origin of bipolar flows. *Astrophys. J*. **410**, 218–238.

WARDLE, M. & NG, C. 1999 The conductivity of dense molecular gas. *Mon. Not. R. Astron. Soc*. **303**, 239–246.

WARDLE, M. & SALMERON, R. 2012 Hall diffusion and the magnetorotational instability in protoplanetary discs. *Mon. Not. R. Astron. Soc*. **422**, 2737–2755.

YANG, H., LI, Z.-Y., LOONEY, L. & STEPHENS, I. 2016 Inclination-induced polarization of scattered millimetre radiation from protoplanetary discs: the case of HL Tau. *Mon. Not. R. Astron. Soc*. **456**, 2794–2805.