PHY-6795

MAGNETOHYDRODYNAMIQUE ASTROPHYSIQUE

Notes de cours

Paul Charbonneau
Département de Physique
Université de Montréal

Septembre 2008

AVERTISSEMENT

Les notes qui suivent ont été originellement préparées dans le cadre d'un cours gradué en Physique Solaire enseigné à trois reprises par Tom Bogdan (HAO/NCAR) et moi-même à l'Université du Colorado à Boulder, sous les sigles APAS7500 et ASTR7500. Les quatre chapitres présentés ici devraient éventuellement faire partie d'un ouvrage gradué sur la Physique Solaire, que Tom et moi espérons bien finir par publier un de ces jours... Comme vous le constaterez rapidement, tous les chapitres n'en sont pas au même stade de "perfectionnement"; Tom et moi apprécierions grandement tout commentaires et suggestions que vous pourriez nous faire quant au contenu et à la présentation des sujets couverts ici.

Remercier: Rony Keppens, Gregg Wade, Fausto, David Galloway, TJB and KBM

Paul Charbonneau
Montréal, septembre 2008

Contents

Part I

Introduction

Chapter 1

Magnetohydrodynamics {chap:mhd}

To sum it all up in a single sentence, the **magnetohydrodynamics** (hereafter MHD) is concerned with the behavior of electrically conducting but globally neutral fluids flowing at non-relativistic speeds and obeying Ohm's Law. Before we dive into MHD proper, it would be wise to clarify what we mean by "fluid" (§1.1), and review the fundamental physical laws governing the flow of unmagnetized fluid, i.e., classical hydrodynamics (§1.2). We then introduce magnetic fields into the fluid picture (§§1.3—1.8), and close in §1.10 with useful mathematical tidbits.

1.1 The fluid approximation {sec:fluid}

1.1.1 Matter as a continuum {ssec:continuum}

It did take some two thousand years to figure it out, but we now know that Democritus was right after all: matter is composed of small, microscopic "atomic" constituents. Yet on our daily macroscopic scale, things sure look smooth and continuous. Under what cicumstances can an assemblage of a great many microscopic elements be treated as a continuum? The primary constraint is that there be a good **separation for scales** between the "microscopic" and "macroscopic".

Consider the situationd depicted on Figure 1.1, corresponding at an amorphous substance (spatially random distribution of microscopic constituents). Denote by λ the mean interparticle distance, and by L the macroscopic scale of the system; we now seek to construct macroscopic variables defining fluid characteristics at the macroscopic scales. For example, if we are dealing with an assemblage of particles of mass m, then the **density** (ρ) associated with a cartesian volume element of linear dimensions l centered at position $\mathbf{x}$ would be given by something like:

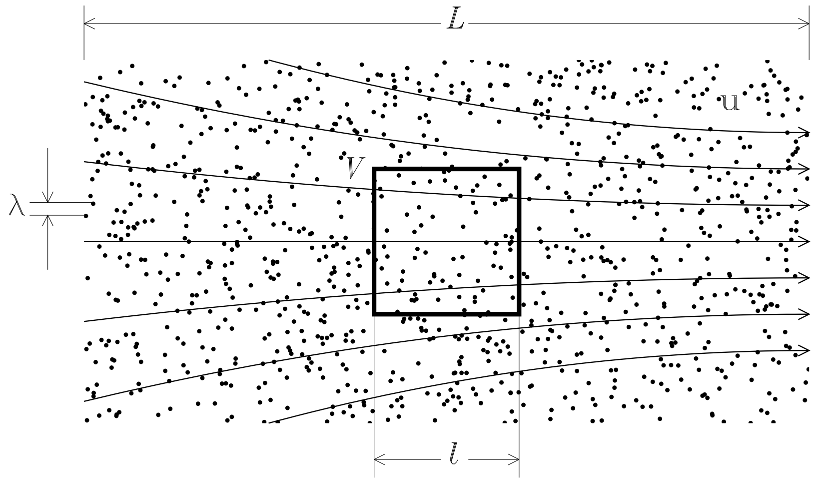

Figure 1.1: {fig:continuum} Microscopic view of a fluid. In general the velocity of microscopic constituent is comprised of two parts: a randomly-oriented thermal velocity, and a systematic drift velocity, which, on the macroscopic scale amounts to what we call a flow **u**. A fluid representation is possible if the mean inter-particle distance λ is much smaller than the global length scale L.

$$\rho(\mathbf{x}) = \frac{1}{l^3}\sum_k m_k \ , \qquad [\mathrm{km\,m^{-3}}] \ , \tag{1.1}$$ {eq:density}

where the sum runs over all particles contained within the volume element. One often hears or reads that for a continuum representation to hold, it is only necessary that the density be "large". But large with respect to what? For the above expression to yield a well-defined quantity, in the sense that the numerical value of ρ so computed does not depend sensitively on the size and location of the volume element, or on time if the particles are moving, it is essential that a great many particles be contained within the element. Moreover, if we want to be writing differential equations describing the evolution of ρ, the volume element better be infinitesimal, in the sense that it is much smaller that the macroscopic length scale over which global variables sucvh as ρ may vary. These two requirements translate in the double inequality:

$$\lambda \ll l \ll L \ . \tag{1.2}$$ {eq:scales}

Because the astrophysical systems and flows that will be the focus of our attention throughout these notes span a very wide range of macroscopic sizes, the continuum/fluid representation will turn out to hold in circumstances where the density is in fact minuscule, as you can verify for yourself upon perusing the collection of astrophysical systems and flows listed in Table 1.1.1 below[1]

Table 1
Properties of some astrophysical objects and flows

{tab:scales}

System/flow	ρ [g/cm^3]	λ [cm]	L	σ	u [km/s]	R$_m$
Solar interior	0.1	10^{-8}	10^5 km		0.1	10^{10}
Solar atmosphere	10^{-7}	10^{-6}			1	10^{11}
Solar corona		10^{-4}	10^5 km		10	10^{12}
Solar wind (1 AU)	10^{-24}	0.6	10^5 km		300	10^{12}
Molecular cloud	10^{-23}	0.1	10 ly		100	
Interstellar medium	10^{-24}	1	1000 ly		100	

1.1.2 Solid versus fluid

Most continuous media can be divided into two broad categories, namely **solids** and **fluids**. The latter does not just include the usual "liquids" of the vernacular, but also gases and plasmas. Physically, the distinction is made on the basis of a medium's response to an applied **stress**, as illustrated on Figure 1.2. A volume element of some continuous substance is subjected to a shear stresses, i.e., two force acting tangentially and in opposite directions on two of its parallel bounding surface (black arrows). A **solid** will immediately generate a restoring force (white arrows), ultimately due to electrostatic interactions between its microscopic constituents, and vigorously resist deformation (try shearing a brick held between the palms of your hands!). The solid will rapidly reach a new equilibrium state characterized by a finite deformation, and will relax equally rapidly to its initial state once the external stress vanishes. A **fluid**, on the other hand, can offer no resistance to the applied stress, at least in the initial stages of the deformation[2].

[1]All density-related estimate assume a gas of fully ionized Hydrogen ($\mu = 0.5$) for the Sun, of neutral Hydrogen for the interstellar medium ($\mu = 1$), and molecular Hydrogen ($\mu = 2$) for molecular clouds. The length scale listed for the solar wind is the size of Earth's magnetosphere, and that for the interstellar medium is the thickness of the galactic (stellar) disk. Velocity estimates correspond to large convectgive cells (solar interior), granulation (photosphere), solar wind speed (corona and solar wind), and turbulence (molecular clouds and interstellar medium). All these numbers (especially the turbulent velocity estimates) are very rough, and moreover rounded to the nearest factor of ten.

[2]We will return in due time to what happens once contiguous fluid elements have attained different, finite velocities. In short, the restoring force is often proportional to the velocity **gradient** produced by the action of the shear.

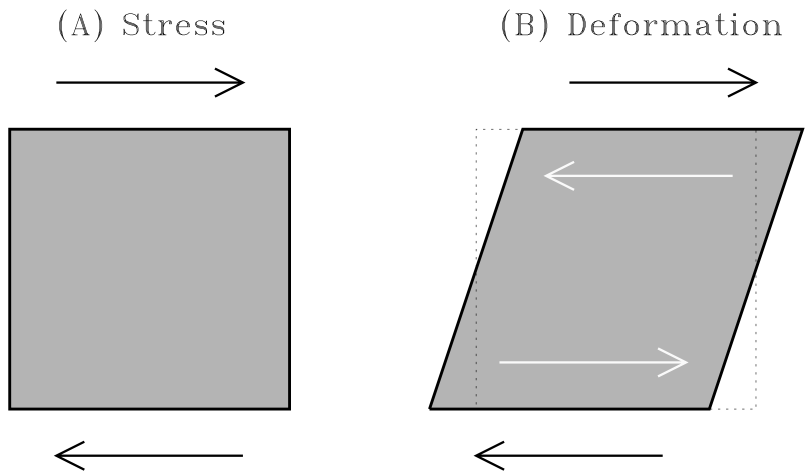

Figure 1.2: {fig:stress} Deformation of a mass element in response to a stress pattern producing an horizontal shear (black arrows). A solid will rapidly reach an equilibrium where internal stresses (white arrows) produced by the deformation will equilibrate the applied shear. A fluid cannot generate internal stresses, and so will be increasingly deformed for as long as the external shear is applied.

1.2 Essentials of hydrodynamics {sec:hydro}

The governing principles of classical hydrodynamics are the same as those of classical mechanics, transposed to continuous media: conservation of mass, linear momentum, angular momentum and energy. The fact that these principles must now be applied to spatially extended volume elements (which may well be infinitesimal, but they are still finite!) introduces some significant complications, mostly with regards to the manner in which forces act. Let's start with the easiest of our conservation statements, that for mass, as it exemplifies very well the manner in which conservation laws are formulated in moving fluids.

1.2.1 Mass: the continuity equation {ssec:mcons}

Consider the situation depicted on Figure 1.3, namely that of an arbitrarily shaped surface S fixed in space and enclosing a volume V embedded in a fluid of density $\rho(\mathbf{x})$ moving with velocity $\mathbf{u}(\mathbf{x})$. The mass flux associated the flow across the (closed) surface is

$$\Phi = \oint_S \rho \mathbf{u} \cdot \hat{\mathbf{n}} \mathrm{d}S \ , \qquad [\mathrm{gm\,s^{-1}}] \tag{1.3}$$ {eq:mcons1}

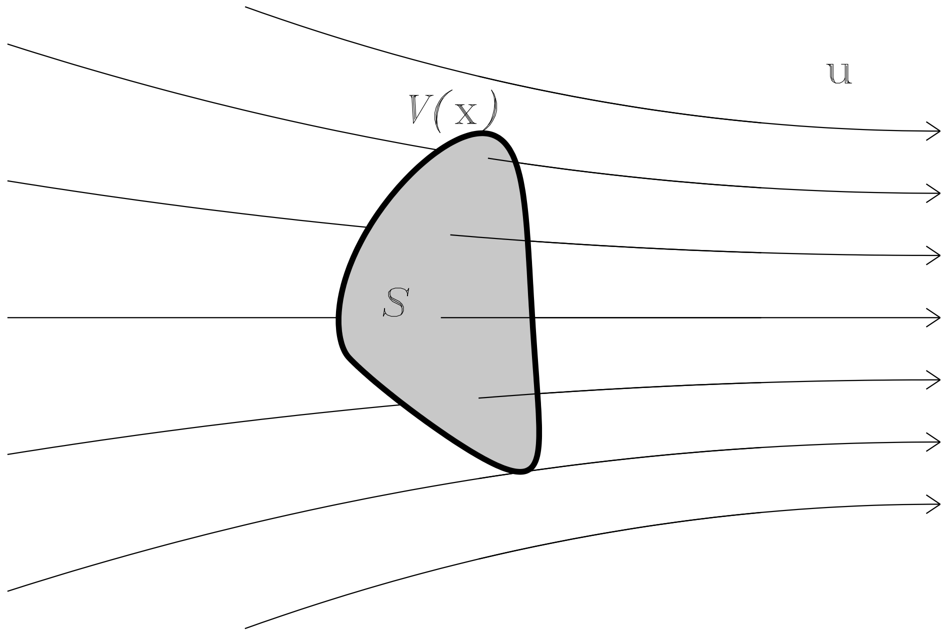

Figure 1.3: {fig:mcons1} An arbitrarily shaped volume element V bounded by a closed surface S, both fixed in space, and traversed by a flow $\mathbf{u}$.

where $\hat{\mathbf{n}}$ is a unit vector everywhere perpendicular to the surface, and by convention oriented towards the exterior. The mass of fluid contained within V is simply

$$M = \int_V \rho \, \mathrm{d}V \ . \qquad [\mathrm{kg}] \tag{1.4}$$ {eq:mcons3}

This quantity will evidently vary if the mass flux given by eq. (1.3) is non-zero

$$\frac{\partial M}{\partial t} = -\Phi \ , \tag{1.5}$$ {eq:mcons4}

here the minus sign is a direct consequence of the exterior orientation of $\hat{\mathbf{n}}$. Inserting eq. (1.3) and eq. (1.4) into (1.5) and applying the divergence theorem to the RHS of the resulting expression yields:

$$\frac{\partial}{\partial t}\int_V \rho \mathrm{d}V = -\int_V \nabla\cdot(\rho\mathbf{u})\mathrm{d}V \ . \tag{1.6}$$ {eq:mcons6}

Because V is fixed in space, the $\partial/\partial t$ et $\int_V$ operators commute, so that

$$\int_V \left[\frac{\partial \rho}{\partial t} + \nabla\cdot(\rho\mathbf{u})\right]\mathrm{d}V = 0 \ . \tag{1.7}$$ {eq:mcons7}

Because V is completely arbitrary, in general this can only be sarisfied provided that

$$\boxed{\frac{\partial\rho}{\partial t} + \nabla\cdot(\rho\mathbf{u}) = 0} \ . \tag{1.8}$$ {eq:mcons}

This expresses mass conservation in differential form, and is known in hydrodynamics as the **continuity equation**.

Incompressible fluids have constants densities, so that in this limiting case the continuity equation reduces to

$$\nabla\cdot\mathbf{u} = 0 \ , \qquad \text{[incompressible]}. \tag{1.9}$$ {eq:incompress}

Water is perhaps the most common example of an effectively incompressible fluid (under the vast majority of naturally occuring conditions anyway). The gaseous nature of most astrophysical fluids may lead you to think that incompressibility is likely to be a pretty lousy approximation in cases of interest in this course. It turns out that the incompressible approximation can lead to a pretty good approximation of the behavior of compressible fluids provided that the flow's Mach number (ratio of flow speed to sound speed) is much smaller than unity.

1.2.2 The D/Dt operator

{ssec:doper}

Suppose we want to compute the time variation of some physical quantity (Z, say) at some fixed location $\mathbf{x}_0$ in a flow $\mathbf{u}(\mathbf{x})$. In doing so we must take into account the fact that Z is in general both an explicit and implicit function of time, because the volume element "containing" Z is moving with the fluid, i.e., $Z \to Z(t, \mathbf{x}(t))$. We therefore need to use the chain rule and write:

$$\frac{\mathrm{d}Z}{\mathrm{d}t} = \frac{\partial Z}{\partial t} + \frac{\partial Z}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial Z}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial Z}{\partial z}\frac{\partial z}{\partial t} \ . \tag{1.10}$$ {eq:doper2}

Noting that $\mathbf{u} = \mathrm{d}\mathbf{x}/\mathrm{d}t$, this becomes

$$\frac{\mathrm{d}Z}{\mathrm{d}t} = \frac{\partial Z}{\partial t} = \frac{\partial Z}{\partial t} + \frac{\partial Z}{\partial x}u_x + \frac{\partial Z}{\partial y}u_y + \frac{\partial Z}{\partial z}u_z = \frac{\partial Z}{\partial t} + (\mathbf{u}\cdot\nabla)Z \ . \tag{1.11}$$ {eq:doper3}

This corresponds to the time variation of Z *following the fluid element as it is carried by the flow.* It is a very special kind of derivative in hydrodynamics, known as the Lagrangian derivative, which will represented by the operator:

$$\boxed{\frac{\mathrm{D}}{\mathrm{D}t} \equiv \frac{\partial}{\partial t} + (\mathbf{u}\cdot\nabla)} \ . \tag{1.12}$$ {eq:doper6}

Note in particular that the Lagrangian derivative of $\mathbf{u}$ yields the acceleration of a fluid element:

$$\mathbf{a} = \frac{\mathrm{D}\mathbf{u}}{\mathrm{D}t} \ , \tag{1.13}$$ {eq:doper7}

a notion that will soon come very handy when we'll write $F = ma$ for a fluid.

A **material surface** is defined as an ensemble of points that defining a surface, all moving along with the flow. Therefore, in a local frame of reference co-moving with any infinitesimal element of a material surface, $\mathbf{u}' = 0$. The distinction between material surfaces, as opposed to surfaces fixed in space such as on eq. (1.3), has crucial consequences with respect to the commuting properties of temporal and spatial differential operators. In the latter case $\int_V$ commutes with $\partial/\partial t$, whereas for material surfaces and volume elements it is D/Dt that commutes with $\int_V$ (and $\oint_S$, etc.).

1.2.3 Linear momentum: the Navier-Stokes equations {ssec:entropy}

A force $\mathbf{F}$ acting on a point-object of mass m is easy to deal with; it simply procuces an acceleration $\mathbf{a} = \mathbf{F}/m$ in the same direction as the force (sounds simple but it still took the genius of Newton to figure it out...). In the presence of a force acting on the surface of a spatially extended fluid element, the resulting fluid acceleration will depend on both the orientation of the force and the surface. We therefore define the net force $\mathbf{t}$ in terms of a **stress tensor**:

$$\mathbf{t}_x = \hat{\mathbf{e}}_x s_{xx} + \hat{\mathbf{e}}_y s_{xy} + \hat{\mathbf{e}}_z s_{xz} \ , \tag{1.14}$$ {???}

$$\mathbf{t}_y = \hat{\mathbf{e}}_x s_{yx} + \hat{\mathbf{e}}_y s_{yy} + \hat{\mathbf{e}}_z s_{yz} \ , \tag{1.15}$$ {???}

$$\mathbf{t}_z = \hat{\mathbf{e}}_x s_{zx} + \hat{\mathbf{e}}_y s_{zy} + \hat{\mathbf{e}}_z s_{zz} \ , \tag{1.16}$$ {???}

where "s_{xy}" denotes the force per unit area acting in the y-direction on a surface perpendicular to the x-direction, $\mathbf{t}_x$ is the net force acting on the surfaces perpendicular to the x-direction, and similarly for the other components. Consider now a unit vector perpendicular to a surface arbitrarily oriented in space:

$$\hat{\mathbf{n}} = \hat{\mathbf{e}}_x n_x + \hat{\mathbf{e}}_y n_y + \hat{\mathbf{e}}_z n_z \ , \qquad n_x^2 + n_y^2 + n_z^2 = 1 \ . \tag{1.17}$$ {???}

The net force along this direction is simply

$$\mathbf{t}_{\hat{\mathbf{n}}} = (\hat{\mathbf{n}} \cdot \hat{\mathbf{e}}_x)\mathbf{t}_x + (\hat{\mathbf{n}} \cdot \hat{\mathbf{e}}_y)\mathbf{t}_y + (\hat{\mathbf{n}} \cdot \hat{\mathbf{e}}_z)\mathbf{t}_z \qquad = \hat{\mathbf{n}} \cdot \mathbf{s} \ . \tag{1.18}$$ {???}

We can now use the Lagrangian acceleration to write the equivalent of "$F = ma$" for the fluid element:

$$\frac{\mathrm{D}}{\mathrm{D}t}\int_V \rho \mathbf{u} \mathrm{d}V = \oint_S \mathbf{s} \cdot \hat{\mathbf{n}} \mathrm{d}S \ . \qquad (1.19) \quad \{???\}$$

We now pull the same tricks as in §1.2.1: use the divergence theorem to turn the surface integral into a volume integral, commute the temporal derivative and volume integral on the RHS, invoke the arbitrariness of the actual integration volume V, and finally make use of the fact that $D\rho/Dt = 0$ as per the continuity equation ($\nabla \cdot \mathbf{u} = 0$???), to obtain the differential equation for $\mathbf{u}$:

$$\rho \frac{\mathrm{D}\mathbf{u}}{\mathrm{D}t} = \nabla \cdot \mathbf{s} \ . \qquad (1.20) \quad \{\texttt{eq:NS0}\}$$

We now define the **pression** as the isotropic part of the force acting perpendicularly on the volume's surfaces, and separate it explicitly from the stress tensor:

$$\mathbf{s} = -p\mathbf{I} + \boldsymbol{\tau} \ , \qquad (1.21) \quad \{???\}$$

where $\mathbf{I}$ is the identity tensor, and the minus sign arises from the convention that pressures acts on the bounding surface towards the interior of the volume element, and $\boldsymbol{\tau}$ will presently become the **viscous stress tensor**. Since $\nabla \cdot (p\mathbf{I}) = \nabla p$, eq. (1.20) becomes

$$\boxed{\frac{\mathrm{D}\mathbf{u}}{\mathrm{D}t} = -\frac{1}{\rho}\nabla p + \frac{1}{\rho}\nabla \cdot \boldsymbol{\tau}} \qquad (1.22) \quad \{\texttt{eq:Navier-Stokes}\}$$

The next step is to obtained expressions for the components of the tensor $\boldsymbol{\tau}$. The viscous force, which is what $\boldsymbol{\tau}$ stands for, can be viewed as a form of friction acting between contiguous laminae of fluid moving with different velocities, so that we expect it to be proportional to velocity *derivatives*. Consider now the following decomposition of a velocity gradient:

$$\frac{\partial u_k}{\partial x_l} = \underbrace{\frac{1}{2}\left(\frac{\partial u_k}{\partial x_l} + \frac{\partial u_l}{\partial x_k}\right)}_{D_{kl}} + \underbrace{\frac{1}{2}\left(\frac{\partial u_k}{\partial x_l} - \frac{\partial u_l}{\partial x_k}\right)}_{\Omega_{kl}} \ . \qquad (1.23) \quad \{\texttt{eq:vstress10}\}$$

The first term on the RHS is a **pure shear**, and is described by the (symmetric) **deformations tensor** D_{kl}; the second is a **pure rotation** , and is described by the antisymmetric **vorticity tensor** Ω_{kl}. It can be shown that the latter causes no deformation of the fluid element, *therefore the viscous force can only involve* D_{kl}. A **Newtonian fluid** is one for which the (tensorial) relation between $\boldsymbol{\tau}$ and D_{kl} is linear:

$$\tau_{ij} = f_{ij}(D_{kl}) \ , \qquad i,j,k,l = (1,2,3) \equiv (x,y,z) \qquad (1.24) \quad \{\texttt{eq:Newtonfl}\}$$

The next step is to invoke the invariance of the physical laws embodied in eq. (1.24) under rotation of the coordinate axes. The mathematics is rather tedious, but at the end of the day you end up with:

$$\begin{aligned}
\tau_{xx} &= 2\mu D_{xx} + (\mu_\vartheta - \tfrac{2}{3}\mu)(D_{xx} + D_{yy} + D_{zz}) && (1.25) \quad \{???\}\\
\tau_{yy} &= 2\mu D_{yy} + (\mu_\vartheta - \tfrac{2}{3}\mu)(D_{xx} + D_{yy} + D_{zz}) && (1.26) \quad \{???\}\\
\tau_{zz} &= 2\mu D_{zz} + (\mu_\vartheta - \tfrac{2}{3}\mu)(D_{xx} + D_{yy} + D_{zz}) && (1.27) \quad \{???\}\\
\tau_{xy} &= 2\mu D_{xy} && (1.28) \quad \{???\}\\
\tau_{yz} &= 2\mu D_{yz} && (1.29) \quad \{???\}\\
\tau_{zx} &= 2\mu D_{zx} && (1.30) \quad \{\texttt{eq:transf}\}
\end{aligned}$$

where μ and μ_ϑ are the coefficients **dynamical viscosity** and **bulk viscosity**, respectively. Is is often convenient to define a coefficent of **kinematic viscosity** as

$$\nu = \frac{\mu}{\rho} , \qquad [\mathrm{m^2\,s^{-1}}] . \qquad (1.31) \quad \{???\}$$

In an incompressible flow, the terms multiplying μ_ϑ vanish and it is possible to rewrite the Navier-Stokes equation in the simpler form:

$$\frac{\mathrm{D}\mathbf{u}}{\mathrm{D}t} = -\frac{1}{\rho}\nabla p + \nu\nabla^2\mathbf{u} . \qquad [\text{incompressible}] \qquad (1.32) \quad \{\texttt{eq:NSincompress}\}$$

Note here the presence of a Laplacian operator acting on a **vector** quantity (here $\mathbf{u}$); this is only equivalent to the Laplacian acting on the scalar components of $\mathbf{u}$ in the special case of cartesian coordinates.

Incompressible or not, the behavior of viscous flows will often hinge on the relative importance of the advective and dissipative terms in The Navier-Stokes equation:

$$\rho(\mathbf{u}\cdot\nabla)\mathbf{u} , \qquad \nabla\cdot\boldsymbol{\tau} ; \qquad (1.33) \quad \{\texttt{eq:NR2}\}$$

Introducing characteristic length scales u_0, L, ρ_0 and ν_0, dimensional analysis yields:

$$\rho_0\frac{u_0^2}{L} , \qquad \frac{1}{L}\rho_0\nu_0\frac{u_0}{L} \qquad (1.34) \quad \{\texttt{eq:NR3}\}$$

where we made use of the fact that tyhe viscous stress tensor has dimensions $\mu\times D_{ik}$, with $\mu = \rho\nu$ and the deformation tensor D_{ik} has dimension of velocity per unti length (cf. éq. 1.23). The ratio of these two terms is a dimensionless quantity called the **Reynolds Number**:

$$\boxed{\mathrm{Re} = \frac{u_0 L}{\nu_0}} . \qquad (1.35) \quad \{\texttt{eq:NR}\}$$

This measures the importance of viscous forces versus fluid inertia. It is a dimensionless key parameter in hydrodynamics, as it effectively controls fundamental processes such as the transition to turbulence, as well as more mundane matters such as boundary layer thicknesses.

A few words on boundary conditions; in the presence of viscosity, the flow speed must vanish wherever the fluid is in contact with a rigid surface S:

$$\mathbf{u}(\mathbf{x}) = 0 \ , \qquad \mathbf{x} \in S \tag{1.36}$$ {eq:NSBC}

this remains true even in the limit where the viscosity is vanishingly small. For a **free surface** (e.g., the surface of a fluid sphere floating in a vacuum), the normal components of both the flow speed and viscous stress must vanish instead:

$$\mathbf{u} \cdot \hat{\mathbf{n}}(\mathbf{x}) = 0 \ , \qquad \boldsymbol{\tau} \cdot \hat{\mathbf{n}} = 0 \ , \qquad \mathbf{x} \in S \tag{1.37}$$ {eq:NSBCfs}

1.2.4 Angular momentum: the vorticity equation {ssec:vort}

The "rotation" and "angular momentum" of a fluid system cannot simply be reduced to simple scalars such as angular velocity and moment of inertia, because the application of a torque to a fluid element can alter not just the rotation rate, but also its shape and mass distribution. A more usuful measure of "rotation" is the **circulation** Γ about some closed contour γ embedded in and moving with the fluid:

$$\Gamma(t) = \oint_\gamma \mathbf{u}(\mathbf{x}, t) \cdot \mathrm{d}\boldsymbol{\ell} = \int_S (\nabla \times \mathbf{u}) \cdot \hat{\mathbf{n}} \, \mathrm{d}S = \int_S \boldsymbol{\omega} \cdot \hat{\mathbf{n}} \, \mathrm{d}S \ . \tag{1.38}$$ {eq:vort1}

where the second equality follows from Stokes' theorem, and the third from the definition of **vorticity**:

$$\boxed{\boldsymbol{\omega} = \nabla \times \mathbf{u} \ .} \tag{1.39}$$ {eq:vortdef}

Thinking about flows in terms of vorticity $\boldsymbol{\omega}$ rather than speed $\mathbf{u}$ can be useful because of **Kelvin's theorem**, which states that the circulation Γ along a any closed loop $/gamma$ advected by the moving fluid is a conserved quantity:

$$\frac{\mathrm{D}\Gamma}{\mathrm{D}t} = 0 \ . \tag{1.40}$$ {eq:kelvin11}

Applying again Stokes' theorem yields the equivalent expression

$$\boxed{\frac{\mathrm{D}}{\mathrm{D}t} \int_S \boldsymbol{\omega} \cdot \hat{\mathbf{n}} \, \mathrm{d}S = 0 \ ,} \tag{1.41}$$ {eq:kelvin12}

stating that the flux of vorticity across any material surface S bounded by γ is also a conserved quantity, both in fact being integral expressions of angular momentum conservation.

An evolution equation for $\boldsymbol{\omega}$ can be obtained via the Navier-Stokes equation, in a particularly illuminating manner in the case of an incompressible fluid ($\nabla \cdot \mathbf{u} = 0$ with constant kinematic viscosity ν, in which case eq. (1.32) can be rewritten as

$$\frac{\mathrm{D}\mathbf{u}}{\mathrm{D}t} == -\nabla\left(\frac{p}{\rho}+\Phi\right) - \nu\nabla\times(\nabla\times\mathbf{u})\ , \qquad \text{[incompressible]} \qquad (1.42)$$ {eq:vorteq1}

where it was assumed that gravity can be expressed as the gradient of a (gravitational) potential. Taking the curl on each side of this expression then yields:

$$\nabla\times\left(\frac{\partial\mathbf{u}}{\partial t}\right) + \nabla\times(\mathbf{u}\cdot\nabla\mathbf{u}) = \underbrace{\nabla\times\left[\nabla\left(\frac{p}{\rho}+\Phi\right)\right]}_{=0} - \nu\nabla\times\nabla\times(\nabla\times\mathbf{u}) \qquad (1.43)$$ {eq:vorteq3}

then, commuting the time derivative with $\nabla\times$ and making judicious of some vector identities to develop the second term on the LHS, remembering also that $\nabla\cdot\boldsymbol{\omega} = 0$, eventually leads to:

$$\boxed{\frac{\mathrm{D}\boldsymbol{\omega}}{\mathrm{D}t} - \boldsymbol{\omega}\cdot\nabla\mathbf{u} = \nu\nabla^2\boldsymbol{\omega}\ , \qquad \text{[incompressible]}} \qquad (1.44)$$ {eq:vorteq}

This is the **vorticity equation**, expressing in differential form the conservation of the fluid's angular momentum.

1.2.5 Energy: the entropy equation

{sec:NSentropy}

Omitting to begin with the energy dissipated in heat by viscous friction, the usual accounting of energy flow into and out of a volume element V fixed in space leads to the following differential equation expressing energy conservation:

$$\frac{\partial}{\partial t}\left[\rho\left(\frac{u^2}{2}+\mathcal{E}\right)\right] = -\nabla\cdot\left[\rho\,\mathbf{u}\left(\frac{u^2}{2}+\mathcal{E}+\frac{p}{\rho}\right) - \chi\nabla T\right]\ , \qquad (1.45)$$ {eq:difec3}

where $\mathcal{E}$ is the fluid's internal energy. The p/ρ term on the RHS embodies the work done against the pressure force upon flowing into the (Eulerian) volume element, and $\chi\nabla T$ is the heat flux in or out of the fluid element, with χ the coefficient of thermal conductivity (units ...). Introducing the first law of thermodynamics allows to rewrite this expression into the more useful form

$$\boxed{\rho T \frac{\mathrm{D}S}{\mathrm{D}t} = \nabla \cdot (\chi \nabla T)} \, . \tag{1.46}$$ {eq:difec4}

which states that any change in the entropy S as one follows a fluid element (LHS) can only be due to heat flowing out of or into the domain by conduction (RHS). For incompressible fluids eq. (1.46) can be written

$$\rho c_p \left(\frac{\partial T}{\partial t} + \mathbf{u} \cdot \nabla T \right) = \nabla \cdot (\chi \nabla T) \, . \qquad \text{[incompressible]} \tag{1.47}$$ {???}

where

$$c_p = T \left(\frac{\partial S}{\partial T} \right)_p \, , \tag{1.48}$$ {???}

is the heat capacity at constant pressure.

While this is seldom an important factor in astrophysical flows, in general we must add to the RHS of eq. (1.46) the heat produced by viscous dissipation (and later, as we shall see later, by Ohmic dissipation). This is given by the so-called **viscous dissipation function**:

$$\phi_\nu = \frac{\mu}{2} \left(\frac{\partial u_i}{\partial x_k} + \frac{\partial u_k}{\partial x_i} - \frac{2}{3} \delta_{ik} \frac{\partial u_s}{\partial x_s} \right)^2 + \mu_\vartheta \left(\frac{\partial u_s}{\partial x_s} \right)^2 \tag{1.49}$$ {eq:viscenrg3c}

where summation over repeated indices is implied here. Note that since ϕ_ν is positive definite, its presence on the RHS of eq. (1.46) can only increase the fluid element's entropy, which makes perfect sense since friction, which is what viscosity is for fluids, is an irreversible process.

1.3 The magnetohydrodynamical induction equation {sec:MHD}

Our task is now to generalize the governing equations of hydrodynamics to include the effects of the electric and magnetic fields, and to obtain evolution equations for these two physical quantities. Keep in mind that electrical charge neutrality, as required by MHD, does not imply that the fluid's microscopic constituents are themselves neutral, but rather that positive and negative electrical charges are present in equal numbers in any fluid element.

The starting point, you guess it I hope, is Maxwell's celebrated equations:

$$\nabla \cdot \mathbf{E} = \frac{\rho_e}{\varepsilon_0} \, , \qquad \text{[Gauss' Law]} \tag{1.50}$$ {eq:max1}

$$\nabla \cdot \mathbf{B} = \frac{\rho_e}{\varepsilon_0} \ , \qquad \text{[Anonymous]} \tag{1.51}$$ {eq:max2}

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \ , \qquad \text{[Faraday's Law]} \tag{1.52}$$ {eq:max3}

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \mu_0 \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t} \ , \qquad \text{[Ampere/Maxwell's Law)]} \tag{1.53}$$ {eq:max4}

where ρ_e is the electrical charge density (units), $\mathbf{J}$ is the electrical current density (units). The permittivity ε_0 ($= XXX$ in vacuum) and magnetic permeability μ_0 ($= 4\pi \times 10^{-7}$ Units in vacuum) can be considered as constants in what follows, since we will not be dealing with polarisable or ferromagnetic substances.

The first step is (with all due respect to the man) to do away altogether with Maxwell's displacement current in eq. (1.53). This can be justified if the fluid flow is non-relativistic and there are no batteries around being turned on or off, two rather sweeping statement that will be substantiated in due time in §1.5. For the time being we just revert to the original form of Ampère's Law:

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J} \ . \tag{1.54}$$ {eq:Ampere}

In general, the application of an electrical field $\mathbf{E}$ across an electrically conducting substance will generate an electrical current density $\mathbf{J}$. Ohm's Law postulates that the relationship between $\mathbf{J}$ and $\mathbf{E}$ is linear:

$$\mathbf{J}' = \sigma \mathbf{E}' \ , \tag{1.55}$$ {eq:Ohmprime}

where σ is the electrical conductivity (units XXX). Here the prime ("$'$") is added to emphasize that Ohm's Law is expected to hold in a conducting substance **at rest**. In the context of a fluid moving with velocity $\mathbf{u}$ (relativistic or not), eq. (1.55) can only be expected to hold in a reference frame comoving with the fluid. So we need to transform eq. (1.55) to the laboratory (rest) frame. In the non-relativitic limit ($u/c \ll 1$, implying $\gamma \to 1$), the usual Lorentz transformation for the electrical current density simplifies to $\mathbf{J}' = \mathbf{J}$, and that for the electric field to $\mathbf{E}' = \mathbf{E} + \mathbf{u} \times \mathbf{B}$, so that Ohm's Law takes on the form

$$\mathbf{J} = \sigma(\mathbf{E} + \mathbf{u} \times \mathbf{B}) \ . \tag{1.56}$$ {eq:Ohm}

or, making use of the pre-Maxwellian form of Ampère's Law and reorganizing the terms:

$$\mathbf{E} = -\mathbf{u} \times \mathbf{B} + \frac{1}{\mu_0 \sigma} (\nabla \times \mathbf{B}) \ . \tag{1.57}$$ {eq:fullE}

We now insert this expression for the electric field into Faraday's Law (1.52) to obtain the very famous **magnetohydrodynamical induction equation**:

$$\boxed{\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{u} \times \mathbf{B} - \eta \nabla \times \mathbf{B})\ .} \tag{1.58}$$ {eq:induction}

where $\eta = 1/(\mu_0 \sigma)$ is the **magnetic diffusivity** (units $\mathrm{m}^2\,\mathrm{s}^{-1}$). The first term on the RHS represents the inductive action of fluid flowing across a magnetic field, while the second term represents dissipation of the electrical currents sustaining the field.

Keep in mind that any solution of eq. (1.58) must also satisfy eq. (1.51) at all times. It can be easily shown (try it!) that if $\nabla \cdot \mathbf{B} = 0$ at some initial time, the form of eq. (1.58) guarantees that zero divergence will be maintained at all subsequent times[3]

1.4 Scaling analysis {sec:scaling}

The evolution of a magnetic field under the action of a prescribed flow $\mathbf{u}$ will depend greatly on whether or not the inductive term on the RHS of eq. (1.58) dominates the diffusive term. Under what conditions will this be the case? We seek a first (tentative) answer to this question by performing a dimensional analysis of eq. (1.58); this involves replacing the temporal derivative by $\frac{1}{\tau}$ and the spatial derivatives by $1/\ell$, where τ and ℓ are time and length scales that suitably characterizes the variations of *both* $\mathbf{u}$ *and* $\mathbf{B}$:

$$\frac{\mathbf{B}}{\tau} = \frac{u_0 \mathbf{B}}{\ell} + \frac{\eta \mathbf{B}}{\ell^2}\ , \tag{1.59}$$ {eq:scalanal}

where B and u_0 are a "typical" values for the flow velocity and magnetic field strength over the domain of interest. The ratio of the first to second term on the RHS of eq. (1.59) is a dimensionless quantity known as the **magnetic Reynolds number**:

$$\mathrm{R}_m = \frac{u_0 \ell}{\eta}\ , \tag{1.60}$$ {eq:Rm}

which measures the relative importance of induction versus dissipation *over length scales of order* ℓ. Note that R_m dos not depend on the magnetic field strength, a direct consequence of the linearity (in $\mathbf{B}$) of the MHD induction equation. Our scaling analysis simply says that in the limit $\mathrm{R}_m \gg 1$, induction by the flow dominates the evolution of $\mathbf{B}$, while in the opposite limit of

[3]This is true under exact arithmetic; if numerical solutions to eq. (1.58) are sought, care must be taken to ensure $\nabla \cdot \mathbf{B} = 0$ as the solution is advanced in time.

$\mathrm{R}_m \ll 1$, induction makes a negligible contribution and $\mathbf{B}$ simply decays away under the influence of Ohmic dissipation. One may anticipate great simplifications of magnetohydrodynamics if we operate in either of these limits. If $\mathrm{R}_m \ll 1$, only the first term is retained on the RHS of eq. (1.59), which leads immediately to

$$\tau = \frac{\ell^2}{\eta} \;, \tag{1.61}$$ {eq:tdiff}

a quantity known as the **magnetic diffusion time**. It measures the time taken for a magnetic field contained in a volume of typical linear dimension ℓ to dissipate and/or diffusively leak out of the volume. Now, for most astrophysical objects, this timescale turns out to be quite large, indeed often larger than the age of the universe! (see Table 1.1). This is not so much because astrophysical plasmas are such incredibly good electrical conductors, but rather because astrophysical objects tend to be very large.

The opposite limit $\mathrm{R}_m \gg 1$, defines the **ideal MHD** limit. Then it is the second term that is retained on on the RHS of eq. (1.59), so that

$$\tau = \ell/u_0 \;, \tag{1.62}$$ {eq:tturnover}

corresponding to the **turnover time** associated with the flow $\mathbf{u}$.

From a purely mathematical point of view, taking the limit $\mathrm{R}_m \to \infty$ of the MHD induction equation is problematic, because the order of the highest spatial derivatives decreases by one. This situation is similar to the behavior of viscous flows at very high Reynolds number: solutions to eq. (1.58) with $\eta \to 0$ in general **do not** smoothly tend towards solutions obtained for $\eta = 0$.

A few final words of warning before we proceed. The distinction between the two physical regimes $\mathrm{R}_m \ll 1$ and $\mathrm{R}_m \gg 1$ is meaningful as long as one can define a suitable R_m for the flow as a whole, which, in turn, requires one to estimate, *a priori*, a length scale ℓ that adequately characterizes the evolving magnetic field at all time and throughout spatial domain of interest. As we proceed it will become clear that this is not always straightforward, or even possible. Likewise, the scaling analysis does away entirely with the geometrical aspects of the problem by substituting $u_0 B$ for $\mathbf{u} \times \mathbf{B}$; yet there are situations (e.g. a field-aligned flow) where even a very large $\mathbf{u}$ has no inductive effect whatsoever, in which case the induction equations assumes the mathematical form

$$\frac{\partial \mathbf{B}}{\partial t} = -\nabla \times (\eta \nabla \times \mathbf{B}) \;, \tag{1.63}$$ {eq:Bdiff}

even though R_m may be very large, and $\mathbf{B}$ evolves on the (long) magnetic diffusion timescale (1.61) rather than on the (short) turnover time τ_u.

1.5 The Lorentz force {sec:magF}

Getting to eq. (1.58) was pretty easy (because we summarily swept the displacement current under the rug), but it represents only half (in fact the easy half) of our task; we must now investigate the effect of the magnetic field on the flow $\mathbf{u}$; and this, it turns out, is the tricky part of the MHD approximation.

You will certainly recall that the **Lorentz force** acting on an electrically charge particle moving at velocity $\mathbf{u}$ in a region of space permeated by electric and magnetic fields is given by

$$\mathbf{f} = q(\mathbf{E} + \mathbf{u} \times \mathbf{B}) \ , \qquad [\mathrm{N}] \ . \tag{1.64}$$ {eq:lorentz1}

where q is the electrical charge. Consider now a volume element ΔV containing a great many such particles; in the continuum limit, the total force per unit volume ($\mathbf{F}$) acting on the volume element will be the sum of the forces acting on each individual charged constituents divided by the volume element:

$$\begin{aligned} \mathbf{F} = \frac{1}{\Delta V}\sum_k \mathbf{f}_k &= \sum_k q_k(\mathbf{E} + \mathbf{u_k} \times \mathbf{B}) \\ &= \left(\frac{1}{\Delta V}\sum_k q_k\right)\mathbf{E} + \left(\frac{1}{\Delta V}\sum_k q_k \mathbf{u_k}\right) \times \mathbf{B} \\ &= \rho_e \mathbf{E} + \mathbf{J} \times \mathbf{B} \ , \qquad [\mathrm{N\,m^{-3}}] \ . \end{aligned} \tag{1.65}$$ {eq:lorentz2}

where the last equality follows from the usual definition of charge density and electrical current density. At this point you might be tempted to eliminate the term proportional to $\mathbf{E}$, on the grounds that in MHD we are dealing with a globally neutral plasma, meaning $\rho_e = 0$, therefore $\rho_e \mathbf{E} \equiv 0$ and that's the end if it. That would be way too easy...

Let's begin by taking the divergence on both side of the generalized form of Ohm's Law (eq. (1.56)). We then make use of Gauss's Law (eq. (1.51)) to get rid of the $\nabla \cdot \mathbf{E}$ term, and of the charge conservation Law

$$\frac{\partial \rho_e}{\partial t} + \nabla \cdot \mathbf{J} = 0 \tag{1.66}$$ {eq:qcons}

to get rid of the $\nabla \cdot \mathbf{J}$ term. The end result of all this physico-algebraeical juggling is the following expression:

$$\frac{\partial \rho_e}{\partial t} + \frac{\rho_e}{(\varepsilon_0/\sigma)} + \sigma \nabla \cdot (\mathbf{u} \times \mathbf{B}) = 0 \ ; \tag{1.67}$$ {eq:indE}

The combination ε_0/σ has units of time, and is called the **charge relaxation time**, henceforth denoted τ_e. It is the timescale on which charge separation takes place in a conductor if an electric field is suddenly turned on. For most conductors, this a very small number, of order 10^{-18} s !! This is because the electrical field reacts to the motion of electric charges at the speed of light (in the substance under consideration, which is slower than in a vacuum but still mighty fast). Indeed, in a conducting fluid at rest ($\mathbf{u} = 0$) the above expression integrates readily to

$$\rho_e(t) = \rho_e(0) \exp(-t/\tau_e) \ , \tag{1.68}$$ {eq:electrorelax}

thus the name "relaxation time" for τ_e.

Now let us consider the case of a slowly moving fluid, in the sense that it is moving on a timescale much larger than τ_e; this means that the induced electrical field will vary on a similar timescale (at best), and therefore the time derivative of ρ_e can be neglected in comparison to the ρ_e/τ_e term in eq. (1.67), leading to

$$\rho_e = \varepsilon_0 \nabla \cdot (\mathbf{u} \times \mathbf{B}) \ . \tag{1.69}$$ {eq:rhoemhd}

This indicates that a finite charge density can be sustained inside a *moving* conducting fluid. The associated electrostatic force per unit volume, $\rho_e \mathbf{E}$, is definitely non-zero but turns out to much *smaller* than the magnetic force. Indeed, a dimensional analysis yields of eq. (1.65) using eq. (1.69) to estimate $\mathbf{E}$ gives:

$$\rho_e \mathbf{E} \sim \left(\frac{\varepsilon_0 u B}{\ell}\right)\left(\frac{J}{\sigma}\right) \sim \left(\frac{u\tau_e}{\ell}\right) JB \ , \tag{1.70}$$ {???}

$$\mathbf{J} \times \mathbf{B} \sim JB \ , \tag{1.71}$$ {???}

where Ohm's Law was used to express $\mathbf{E}$ in terms of $\mathbf{J}$, and once again ℓ is a typical length scale characterizing the variations of the flow and magnetic field. The ratio of electrostatic to magnetic force is thus of order $u\tau_e/\ell$. Now $\tau_e \ll 1$ to start with, and for non-relativistic fluid motion we can expect that the flow's turnover time ℓ/u is much larger than the crossing time for an electromagnetic disturbance $\sim \tau_e$; both effects conspire to render the electrostatic force absolutely minuscule compared to the magnetic force, so that eq. (1.65) becomes

$$\mathbf{F} = \mathbf{J} \times \mathbf{B} \ , \qquad \text{[MHD approximation]} \ . \tag{1.72}$$ {eq:FMHD}

and this must be added to the RHS of the Navier-Stokes equation (1.22)... with a $1/\rho$ prefactor for units to come out right.

Now, getting back to this business of having dropped the displacement current in the full Maxwellian form of Ampère's Law (eq. (1.53)); it can now be all justified on the grounds that the time derivative of the charge density can be neglected in the non-relativistic limit. Indeed, to be consistent the charge conservation equation (1.66) now reduces to

$$\nabla \cdot \mathbf{J} = 0 \; ; \tag{1.73}$$ {???}

taking the divergence on both sides of eq. (1.53) then leads to

$$\nabla \cdot \mathbf{J} = -\varepsilon_0 \nabla \cdot \left(\frac{\partial \mathbf{E}}{\partial t} \right) = \varepsilon_0 \frac{\partial}{\partial t} (\nabla \cdot \mathbf{E}) = \frac{\partial \rho_e}{\partial t} \; ; \tag{1.74}$$ {???}

this demonstrates that dropping the time derivative of the charge density is equivalent to neglecting Maxwell's displacement current in eq. (1.53). To sum up, provided we exclude very rapid transient events (such as the turning on of a battery, or any such process which would generate a large $\partial \rho_e / \partial t$), under the MHD aproximation the following statements all hold true:

- The fluid motions are non-relativistic;
- The electrostatic force can be neglected as compared to the magnetic force;
- Maxwell's displacement current can be neglected.

1.6 MHD waves {sec:MHDwaves}

Although it looks innocuous enough, the magnetic force in the MHD approximation has some rather complex consequences for fluid flows, as we will have ample occasions to verify throughout this course. One particularly intricate aspects relates to the types of **waves** that can be supported in a magnetized fluid; in a classical unmagnetized fluid, one deals primarily with sound waves (pressure acting as a restoring force), and gravity waves (gravity actring as restoring force). It turns out that the Lorentz force introduces not one, but really two additional restoring forces.

Making judicious use of eqs. (1.51) and (1.54), together with some classical vector identities, eq. (1.72) can be rewritten as

$$\mathbf{F} = \frac{1}{8\pi} \nabla (B^2) + (\mathbf{B} \cdot \nabla) \mathbf{B} \; , \tag{1.75}$$ {eq:FMHD2}

where $B^2 \equiv \mathbf{B} \cdot \mathbf{B}$. The first term on the RHS is the **magnetic pressure**, and the second the **magnetic tension**. The general idea is illustrated on Figure 1.4. Fluctuations in magnetic pressure can propagate as a longitudinal wave, much as a sound wave, as depicted on Fig. 1.4A. In fact, two such

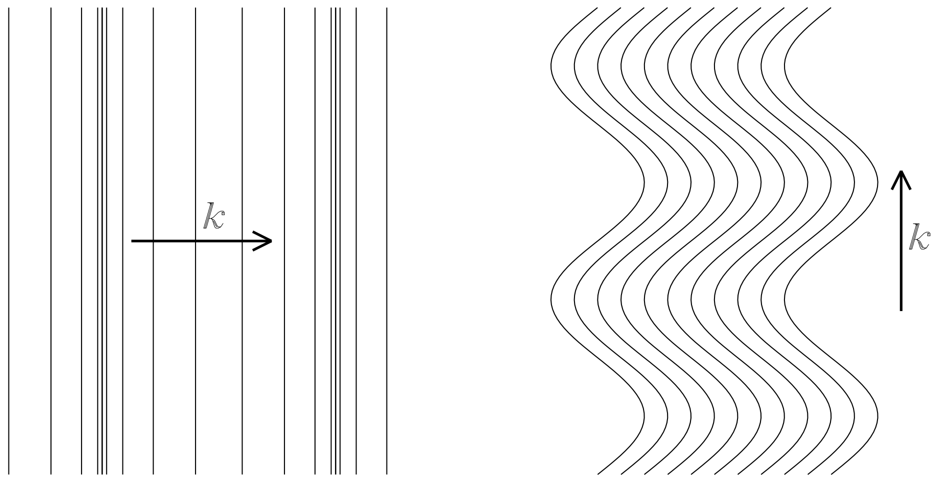

Figure 1.4: {fig:MHDwaves} The two fundamental MHD wave modes in a uniform background magnetic field: (A) magnetosonic mode, and (B) Alfvén mode. The wave vector **k** is indicated as a thick arrow, and highlights the fact that the magnetosonic mode is a longitudinal wave, while the Alfvén mode is a transverse wave. In the presence of plasma, the magnetosonic mode breaks into two submodes, according to the phasing between the magnetic pressure and gas pressure perturbations (see text).

magnetosonic waves modes actually, according to whether the magnetic pressure fluctuation is in phase with the gas pressure fluctuation (the so-called fast mode), or out of phase (the slow mode). In addition, magnetic tension can produce a restoring force that allows the propagation of wave-on-a-string-like transverse waves, known as **Alfvén waves**, as illustrated on Fig. 1.4B.

1.7 Magnetic energy {sec:magE}

Consider the expression resulting from dotting of **B** into the induction equation (1.58), integrating over the spatial domain (V) under consideration, and making judicious use of various well-known vector identities and of Gauss' theorem:

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_V \frac{\mathbf{B}^2}{8\pi}\mathrm{d}V = -\int_S (\mathbf{S}\cdot\mathbf{n})\,\mathrm{d}S - \int_V (\mathbf{u}\cdot\mathbf{L})\,\mathrm{d}V - \int_V \sigma_e^{-1}\mathbf{J}^2\,\mathrm{d}V\ , \qquad (1.76) \quad \{\texttt{eq:emag}\}$$

where **E** is the electric field, and **n** is a outward-directed unit vector normal to the boundary surface S. The vector quantities **S**, **L** and **J** are the Poynting flux, Lorentz force and current density, respectively. Recall that in in the

MHD limit we have, using cgs units,

$$\mathbf{S} = \frac{c}{4\pi}\mathbf{E} \times \mathbf{B} \ , \tag{1.77}$$ {E2.111}

$$\mathbf{L} = \frac{1}{4\pi}(\nabla \times \mathbf{B}) \times \mathbf{B} \ , \tag{1.78}$$ {E2.112}

$$\frac{4\pi}{c}\mathbf{J} = \nabla \times \mathbf{B} \ . \tag{1.79}$$ {E2.113}

We also made use of the fact that in the MHD approximation, the net current $\mathbf{J}$ is expressed as the sum of the conduction and induction currents:

$$\mathbf{J} = \sigma_e(\mathbf{E} + \frac{1}{c}\mathbf{u} \times \mathbf{B}) \ . \tag{1.80}$$ {E2.217c}

Examine now the three terms on the RHS of eq. (1.76); the first is the Poynting flux component into the domain, integrated over the domain boundaries, i.e., the flux of electromagnetic energy in (integrand < 0) or out (integrand > 0) of the domain. This term evidently vanishes in the absence of applied magnetic or electric fields on the boundaries. The second is the work done by the Lorentz force ($\mathbf{L}$) on the flow. In general this term can be either positive or negative, although in the dynamo context we are interested in situations where the magnetic field is amplified by the flow, i.e., the flow transfers energy to the magnetic field ($\mathbf{u} \cdot \mathbf{L} < 0$)[4]. The third term is evidently always negative, and represents the rate of energy loss due to Ohmic dissipation. Equations (1.76) then naturally leads to interpret the quantity $\mathbf{B}^2/8\pi$ as the magnetic energy density, since the LHS of eq. (1.76) is clearly the rate of change of the total **magnetic energy** ($\mathcal{E}_\mathrm{B}$) within the domain:

$$\mathcal{E}_\mathrm{B} = \frac{1}{8\pi}\int_V \mathbf{B}^2 \mathrm{d}V \tag{1.81}$$ {E2.217}

1.8 Magnetic flux freezing and Alfvén's theorem {sec:freeze}

Let us return to Faraday's Law, in the form given by the third Maxwell equation:

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \ . \tag{1.82}$$ {???}

[4]The $\mathbf{u} \cdot \mathbf{L} > 0$ case is quite relevant to the design of magnetic pumps for electrically conducting fluids. This has received quite a bit of attention in light of the use of liquid sodium to cool the core of nuclear reactors.

Project now each side of this expression onto a unit vector normal to some surface S fixed in space and bounded by a closed countour γ, integrate over S, and apply Stokes' theorem to the LHS:

$$\int_S (\nabla \times \mathbf{E}) \cdot \hat{\mathbf{n}} \mathrm{d}S = \oint_\gamma \mathbf{E} \cdot \mathrm{d}\ell = -\int_S \left(\frac{\partial \mathbf{B}}{\partial t} \right) \cdot \hat{\mathbf{n}} \mathrm{d}S \ , \qquad (1.83) \quad \{???\}$$

So far the surface S remains completely arbitrary. If it is fixed in space, then we get the usual integral form of Faraday's Law:

$$\oint_\gamma \mathbf{E} \cdot \mathrm{d}\ell = -\frac{\partial}{\partial t} \int_S \mathbf{B} \cdot \hat{\mathbf{n}} \mathrm{d}S \ , \qquad (1.84) \quad \{\texttt{eq:almost}\}$$

with the LHS corresponding to the electromotive force, and the integral on the RHS the magnetic flux. If we now assume instead that the surface S is a material surface moving with the fluid, then (1) we must substitute the Lagrangian operator D/D for the partial derivative on the RHS of eq. (1.84); and (2) we are allowed to invoke Ohm's Law to eliminate $\mathbf{E}$ on the RHS since any point of the (material) contour is by definition co-moving with the fluid:

$$\frac{1}{\sigma} \oint_\gamma \mathbf{J} \cdot \mathrm{d}\ell = -\frac{\partial}{\partial t} \int_S \mathbf{B} \cdot \hat{\mathbf{n}} \mathrm{d}S \ , \qquad (1.85) \quad \{???\}$$

Now, obviously, in the limit of infinite conductivity we have

$$\frac{\mathrm{D}}{\mathrm{D}t} \int_S \mathbf{B} \cdot \hat{\mathbf{n}} \mathrm{d}S = 0 \ . \qquad (1.86) \quad \{\texttt{eq:alfven}\}$$

This states that in the ideal MHD limit $\sigma \to \infty$, the magnetic flux threading any (open) surface is a conserved quantity as the surface is advected (and possibly deformed) by the flow. This results is known as **Alfvén's theorem**. Note in particular that in the limit of an infinitisemal surface pierced by "only one" fieldline, Alfvén's theorem is equivalent to saying that magnetic fieldline must move in the same way as fluid elements; it is customary to stay that the magnetic field is "frozen" into the fluid. In this manner it behaves just like vorticity in the inviscid limit $\nu \to 0$. And like in the case of vorticity, sheared flow can amplify magnetic fields by stretching, a subject we will investigate in all great gory details in Part III of these class notes.

1.9 Magnetic helicity {sec:helicity}

1.10 Mathematical representations of magnetic fields {sec:mathB}

1.10.1 Pseudo-vectors and soleinodal vectors {ssec:vectors}

It is worth distinguishing between real vectors (also called axial vectors) and **pseudo-vectors**, the latter class including the magnetic field vector. Real vectors remain invariant upon inversion of the (3D) coordinates about the origin, i.e., $\mathbf{x} \to -\mathbf{x}$, hereafter thinking in cartesian coordinates to ease the discussion. This will leave the "physical" direction in space of a true vector (like a velocity $\mathbf{u}$) unchanged, since both the coordinate unit vectors and the components of the velocity will change sign:

$$\mathbf{u}' = (-u_x)(-\hat{\mathbf{e}}_x) + (-u_y)(-\hat{\mathbf{e}}_y) + (-u_z)(-\hat{\mathbf{e}}_z) = \mathbf{u} \ . \qquad (1.87)$$ {eq:vflip}

However, in terms of vector products, curl operators, orientation of surfaces and so on, the coordinate inversion will take us from a right-handed coordinate system to a left-handed system. This implies that a vector like the magnetic field must remain invariant under coordinate inversion. This can be appreciate by considering the expression for the magnetic force acting on a charge q moving at velocity $\mathbf{u}$ in a magnetic field $\mathbf{B}$:

$$\mathbf{f} = q\,\mathbf{u} \times \mathbf{B} \ ; \qquad (1.88)$$ {eq:magF}

we just argued that the components of $\mathbf{f}$ and $\mathbf{u}$ will change sign under coordinate inversion; therefore the magnetic field components must *not* change sign under coordinate inversion, for eq. (1.88) to remain valid (physical laws do not care about our coordinate conventions!). One must conclude that upon coordinate inversion, the direction of a vector field such as $\mathbf{B}$ immediately flips! So the Earth's north magnetic pole instantly becomes the south magnetic pole[5]. Weird behavior for a vector, which is why such vectors inherit the prefix "pseudo".

Pseudo or not, there are numrous vectors fields of physical interest out there that have the property that their divergence vanishes; the magnetic field is evidently such a vector field, as per our second Maxwell equation (1.51). Any vector field $\mathbf{G}$, say) satisfying $\nabla \cdot \mathbf{G} = 0$ is called a **solenoidal vector**.

Soleinodal vectors have a very interesting property related to the conservation of their flux across material surfaces transported and deformed by a

[5]Does this mean that your compass needle will instantly rotate by 180 degrees? Think about that one a bit...

flow field **u**. They can be shown to satisfy the following kinematic theorem:

$$\frac{\mathrm{D}}{\mathrm{D}t}\int_{S_m} \mathrm{G}\cdot\hat{\mathbf{n}}\mathrm{d}S = \int_{S_m}\left[\frac{\partial \mathbf{G}}{\partial t} - \nabla\times(\mathbf{u}\times\mathbf{G})\right]\cdot\hat{\mathbf{n}}\mathrm{d}S\ . \qquad (1.89)$$ {eq:kintheorem}

This is simply saying that the net variation of the flux (LHS) can be due either to intrinsic time-variation of the vector field (first term in the square brackets on the RHS) or to deformation of the material surface S_m by the flow **u**.

Note that we could have arrived at Alfvén's theorem (§1.8) starting from this kinematic theorem for solenoidal vector fields, as applied to **B**:

$$\frac{\mathrm{D}}{\mathrm{D}t}\int_{S_m} \mathrm{B}\cdot\hat{\mathbf{n}}\mathrm{d}S = \int_{S_m}\left[\frac{\partial \mathbf{B}}{\partial t} - \nabla\times(\mathbf{u}\times\mathbf{B})\right]\cdot\hat{\mathbf{n}}\mathrm{d}S \qquad (1.90)$$ {eq:kinth2}

Obviously, the quantity within square brackets on the RHS will vanish as per our MHD induction equation written in the ideal limit $\eta \to 0$, which gets us directly to eq. (1.86). You will recall, of course, that Ohm's Law is indeed already embodied in the MHD induction equation, so this is really getting to the same result by two mathematically distinct but physically equivalent paths.

1.10.2 The vector potential {ssec:A}

It will often prove useful to work with the MHD induction equation written in terms of a vector potential **A** (units T m) such that $\mathbf{B} = \nabla\times\mathbf{A}$. Equation (6.12) is then readily integrated to

$$\frac{\partial \mathbf{A}}{\partial t} = \mathbf{u}\times(\nabla\times\mathbf{A}) - \eta\nabla\times(\nabla\times\mathbf{A}) + \nabla\Phi\ , \qquad (1.91)$$ {E2.02}

where, in "uncurling" the induction equation we may append the gradient of a scalar function to the RHS, with no effect on **B**. This additional term may contribute to the electric field **E**, however, and so Φ is conveniently regarded as the electrostatic potential[6]. Clearly, any solution of eq. (1.91) identically satisfies the solenoidal constraint $\nabla\cdot\mathbf{B} = 0$.

1.10.3 Axisymmetric magnetic fields {ssec:axiB}

In many astrophysical situations to be encountered in subsequent chapters we will facing astrophysical magnetofluid systems that show symmetry about an axis, in fact usually a rotational axis. **axisymmetric** with respect to the

[6]In most (but not all!) situations dealt with in the following pages, Φ can (and will) be set to zero without objectionable consequences.

rotation axis ($\partial/\partial\phi = 0$). Likewise, the sun's differential rotation and meridional circulation, as inferred from surface measurements and helioseismology, are also very closely axisymmetric on the largest spatial scales. In spherical polar coordinates (r, θ, ϕ), the most general axisymmetric ($\partial/\partial\phi = 0$) magnetic field and flow can be written as

$$\mathbf{u}(r,\theta,t) = \rho^{-1}\nabla \times (\Psi(r,\theta,t)\hat{\mathbf{e}}_\phi) + \varpi\Omega(r,\theta,t)\hat{\mathbf{e}}_\phi \qquad (1.92) \quad \{\texttt{E2.70a}\}$$

$$\mathbf{B}(r,\theta,t) = \nabla \times A(r,\theta,t)\hat{\mathbf{e}}_\phi + B(r,\theta,t)\hat{\mathbf{e}}_\phi \qquad (1.93) \quad \{\texttt{E2.70b}\}$$

where $\varpi = r\sin\theta$. Here the vector potential A and stream function Ψ define the **poloidal** components of the field and flow, i.e., the component contained in meridional (r, θ) planes. The azimuthal component B is often called the **toroidal field**, and Ω is the angular velocity (units rad s^{-1}). Evidently eqs. (1.92)—(1.93) satisfies the constraints $\nabla \cdot (\rho\mathbf{u}) = 0$ (mass conservation in a steady flow) and $\nabla \cdot \mathbf{B} = 0$ by construction.

A practical advantage of this so-called mixed representation is that it allows the separation of the (vector) MHD induction equation into *two* components for the 2D scalar fields A and B:

$$\frac{\partial}{\partial t}(\varpi A) + \mathbf{u}_p \cdot \nabla(\varpi A) = \varpi\eta\left(\nabla^2 - \frac{1}{\varpi^2}\right) A \ , \qquad (1.94) \quad \{\texttt{E2.10a}\}$$

$$\frac{\partial}{\partial t}\left(\frac{B}{\varpi}\right) + \mathbf{u}_p \cdot \nabla\left(\frac{B}{\varpi}\right) = \frac{\eta}{\varpi}\left(\nabla^2 - \frac{1}{\varpi^2}\right) B + \frac{1}{\varpi}(\nabla\eta) \times (B\hat{\mathbf{e}}_\phi)$$

$$-\left(\frac{B}{\varpi}\right)\nabla \cdot \mathbf{u}_p + \mathbf{B}_p \cdot \nabla\Omega \ , \qquad (1.95) \quad \{\texttt{E2.10b}\}$$

where $\mathbf{B}_p$ and $\mathbf{u}_p$ are notational shortcuts for the poloidal field and meridional flow. Notice that the vector potential A evolves in a manner entirely independent of the toroidal field B, the latter being conspicuously absent on the RHS of eq. (1.94). This is not true of the toroidal field B, which is well aware of the poloidal field's presence via the $\nabla\Omega$ shearing term.

On numerous occasions in this and subsequent chapters we will seek solutions to eqs. (1.94)—(1.95) inside a sphere (radius R) of magnetized fluid; in the "exterior" $r > R$ there is only vacuum, which implies vanishing electric currents. In practice we will need to match whatever solution we compute in $r < R$ to a current-free solution in $r > R$; such a solution must satisfy

$$\frac{4\pi}{c}\mathbf{J} = \nabla \times \mathbf{B} = 0 \ . \qquad (1.96) \quad \{\texttt{E2.05}\}$$

For an axisymmetric system eq. (1.96) then translates into

$$\left(\nabla^2 - \frac{1}{\varpi^2}\right) A(r,\theta,t) = 0 \ , \qquad (1.97) \quad \{\texttt{E2.9a}\}$$

$$B(r,\theta,t) = 0 \ . \qquad (1.98) \quad \{\texttt{E2.9b}\}$$

Solutions to eq. (1.97) have the general form

$$A(r,\theta,t) = \sum_{l=1}^{\infty} a_l \left(\frac{R}{r}\right)^{l+1} P_l^1(\cos\theta) \qquad r > R \ , \qquad (1.99) \quad \{\texttt{E2.9c}\}$$

where the P_l^1 are the associated Legendre functions of order 1 and l is a positive integer. Solutions to eqs. (1.94)–(1.95) computed within the sphere must then be smoothly matched to eqs. (1.98)—(1.99) in the exterior. In particular, the vector potential A must be continuous up to its first derivative normal to the surface, so that the magnetic field component tangential to the surface remains continuous across $r = R$. Regularity of the magnetic field on the symmetry axis ($\theta = 0$) requires that we set $B = 0$ there. Without any loss of generality, we can also set $A = 0$ on the axis.

1.10.4 Force-free magnetic fields {ssec:ffB}

In many astrophysical systems, the magnetic field dominates the dynamics and energetics of the system. Left to itself, such a system would tend to evolve to a force-free state described by

$$\mathbf{F} = \mathbf{J} \times \mathbf{B} = 0 \ . \qquad (1.100) \quad \{???\}$$

Broadly speaking, this can be achieved in two physically distinct ways (excluding the trivial solution $\mathbf{B} = 0$). The first is of $\mathbf{J} = 0$ throughout the system. Then Ampère's Law becomes $\nabla \times \mathbf{B} = 0$, which means that, as with the electric field in electrostatic, $\mathbf{B}$ can be expressed as the gradient of a potential. Such a magnetic field is called a **potential field** Upon substitution into Ampère's Law yields a Laplace-type problem:

$$\mathbf{B} = \nabla\Phi \ , \qquad \nabla^2\mathbf{B} = 0 \ , \qquad [\text{Potential field}] \ . \qquad (1.101) \quad \{\texttt{eq:potB}\}$$

Alternately, a system including a non-zero current density can still be force free, provided the currents flow everywhere parallel to the magnetic field, i.e.,

$$\nabla \times \mathbf{B} = \alpha\mathbf{B} \ . \qquad (1.102) \quad \{\texttt{eq:alphaB}\}$$

where α need not necessarily be a constant, i.e., it can vary from one fieldline to another, vary in space, and even depend on the (local) value of $\mathbf{B}$. Imagine now a situation where, in some domain (for example, the exterior of a star), we are provided with a boundary condition on $\mathbf{B}$ and the task is to construct a force-free field. Adopting the potential field *anzatz* can lead to very different reconstructions that if we adopt instead eq. (1.102), given that in the latter case one is free to specify any elextric current distribution within the domain, as long as $\mathbf{J}$ remains parallel to $\mathbf{B}$.

A very important result in this context is known as **Aly's Theorem**; it states that in a semi-infinite domain with $\mathbf{B}_\perp$ imposed at the boundary and $\mathbf{B} \to 0$ as $\mathbf{x} \to \infty$, the (unique) potential field solution satisfying the boundary conditions has a magnetic energy that is lower than any of the (multiple) solutions of eq. (1.102) that satisfy the same boundary conditions, even with complete freedom to specify $\alpha(\mathbf{x})$ within the domain. This poses a strict limit to the amount of magnetic energy stored into a system that can actually be tapped into to power whichever astrophysically interesting phenomenon.

Problems:

1. Obtain the charge conservation equation (1.66) by following the general logic used in §1.2.1 to obtain the continuity equation (1.8).
2. Fill in the missing mathematical steps leading to eq. (1.67)
3. Fill in the missing mathematical steps leading to eq. (1.75)
4. Obtain equations (1.94) and (1.95) by substitution of eqs. (1.92) and (1.95) into the MHD induction equation (1.58. Hint: the induction equation is a vector equation; terms "oriented" in the ϕ-direction must cancel one another independently of terms oriented perpendicular to the ϕ-direction.
5. This problem lets you dig a bit deeper in the concept of magnetic energy (§1.7).

   (a) Starting from the induction equation, fill in the missing mathematical steps leading to eq. (1.76).

   (b) Show that in the absence of induction (meaning $\mathbf{u} = 0$), a force-free magnetic contained in a domain V will always decay.

   (c) Making use of eq. (1.80), obtain an expression involving $\mathbf{B}$ and $\mathbf{u}$ but not $\mathbf{E}$, for the Poynting flux component normal to the boundary S enclosing an electrically conducting fluid. Give a physical interpretation for each term in the resulting mathematical expression.

6. This problem lets you explore some astrophysical implications of the flux freezing constraints.

   (a) Assume that the Sun has formed from the spherically symmetric collapse of an initially spnerical gas cloud of radius XXX and threaded by a large-scale galactic magnetic field of strength 10^{-XX} G. Under the assumption of flux freezing, what shoulds then be the strength of the internal solar magnetic field? Is this a reasonable number?

   (b) The large-scale poloidal field of the Sun is actually of order 10 G at high latitudes. Under the same assumptions as in (a), compute the strength of the magnetic field expected in a White Dwarf ($R_{\rm WD}/R_\odot = 0.01$). Is this a reasonable number?

   (c) What would you think is the primary problem(s) with such estimates?

Bibliography:

There are a great many books available on classical hydrodynamics. The following are my own top-three personal favorites:

Tritton, D.J., *Physical Fluid Dynamics*, 2nd ed., Oxford University Press (1988),

Acheson, D.J., *Elementary Fluid Dynamics*, Clarendon Press (1990)

Landau, L., et Lifschitz, E., *Mécanique des Fluides*, Éditions MIR (2^e édition 1986).

If you are looking for an introduction to the topic targeted at physics-trained readers, you may want to work through the first six chapters of my class notes for PHY-3140, the 2007 version of which being still available on the Web:

`http://www.astro.umontreal.ca/p̃aulchar/phy3140/notes07.pdf`

(I'm of course working on a revised version for the winter of 2009). If you need a refresher on undergraduate electromagnetism, you should go back to

Griffith, D.J., *Introduction to Electrodynamics*, 3rd ed., Prentice Hall (1999).

At the graduate level, the standard reference remains

Jackson, J.D., *Classical Electrodynamics*, 2nd ed., John Wiley & Sons (1975)

who does devote a chapter to magnetohydrodynamics, including a discussion of magnetic wave modes. My personal favorite on magnetohydrodynamics is:

Davidson, P.A., *An Introduction to Magnetohydrodynamics*, Cambridge University Press (2001),

Sections 1.5 and 1.8 are strongly inspired by Davidson's own presentation of the subject. He also presents an illuminating proof of the kinematic theorem embodied in eq. (1.89). The following textbook is also well worth consulting:

Goedbloed, H., & Poedts, S., *Principles of Magnetohydrodynamics*, Cambridge University Press (2004).

These authors put greater emphasis on MHD waves, shocks, and on the intersection of MHD and plasma physics. For those seeking even more focus on plasma physics aspects, I would recommend:

Kulsrud, R.M., *Plasma Physics for Astrophysics*, Princeton University Press (2005).

On Aly's theorem, see

Aly, J.-J. 1984, *Astrophys. J.*, **283**, 349,
Aly, J.-J. 1991, *Astrophys. J.*, **375**, L61,
Smith, D.F., & Low, B.C. 1993, *Astrophys. J.*, **410**, 412.

Chapter 2

Magnetic fields in astrophysics

{chap:ApB}

By now you may think you have landed in some sort of deranged combined crash course on fluid mechanics, electromagnetism... and vector algebra! To dispel this idea we now return closer to our subject matter, by briefly documenting the omnipresence of magnetic fields throughout the universe (§2.1), pondering as to the conspicuous absence of electric fields (§2.2), and considering the ultimate origin of magnetic fields (§2.3).

2.1 A zoo of astrophysical magnetic fields

{sec:zoo}

2.1.1 Earth's magnetic field

{ssec:earth}

Natural magnetism (in technical parlance, ferromagnetism) is known at least since Antiquity, but it took the monumental treatise *De Magnete*, published in 1600 by William Gilbert (1544-1603), to really drive home the point that the Earth is one huge spherically-shaped bar magnet. Gilbert arrived at this conclusion from comparing the known behavior of compass needless to what he observed around a bar magnet carved into a sphere (see Figure 2.1). A medical doctor by training, in his book Gilbert also debunked many semi-occult beliefs about the behavior of magnetic objects and their influence on the human body and psyche.

To a good first approximation, the Earth's magnetic field has the form of a dipole approximately aligned with the Earth's rotation axis, with an average surface field strength of 0.5G. Geologic evidence has shown that the Earth's magnetic field is not steady, but flips polarities between the N and S hemisphere, these reversals being rapid (on geological timescales; they last some 10,000yr), are irregularly spaced, and punctuating much longer epochs of more or less stable field configuration, lasting a few 10^5 yr on average. At the present Earth's dipole moment is $M_\oplus = 8.1 \times 10^{22}$ A m^2. Paleomagnetic studies indicate that $M_\oplus$ has been declining rather rapidly over the past few

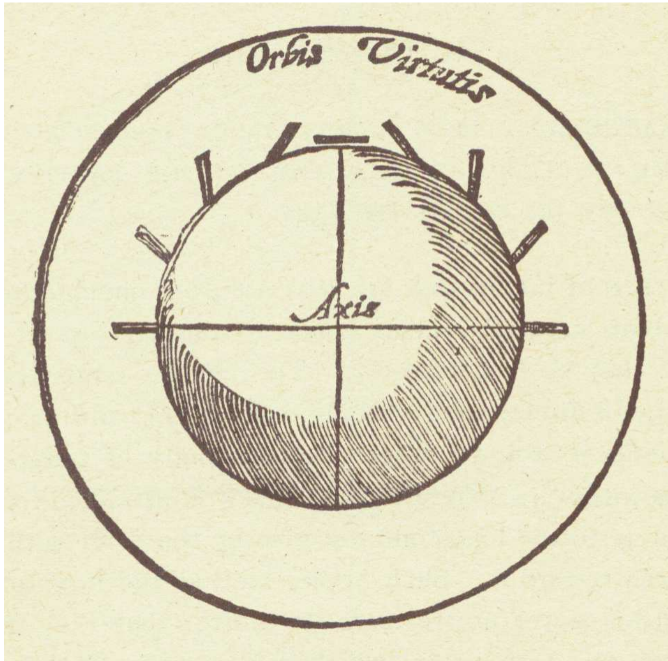

Figure 2.1: {fig:Gilbert}). Drawing in William Gilbert *De Magnete*, written in 1600. Gilbert polished a magnet in the form of a sphere, and could show that the pattern of inclination of the magnetic needle of as compass placed at various locations around the sphere was identical to what had already been observed by long-distance navigators and travellers of the sixteenth century.

1000 yr, suggesting that we may be heading for a polarity reversal sometimes in the next few 1000 yr if the current trend persists.

Because the Earth's crust and troposphere are such lousy electrical conductors, the presence of the geomagnetic field is seldom felt in our daily life (and is ever more fading from popular consciousness with the replacement of magnetic compasses by GPS). In the Earth's ionosphere, however, the geomagnetic field is quite significant, and it interaction with the solar wind (to be encountered in part II of these notes) is what defines the **magnetosphere**, which happens to shield us from a lot of high energy particles often accelerated as a side effect of solar eruptive events (more on those shortly!). The impact of solar ejecta on the magnetosphere triggers geomagnetic storms. Their most spectacular manifestation being auroral emission, but the induced electric fields can pose threats to technological infrastructures such as power lines and pipelines.

2.1.2 Other solar system planets

{ssec:planets}

Magnetic fields have been measured on most solar system planets by various space probes and landers. Table 2.1 lists some of the salient characteristics of planetary magnetic fields. Venus is the only planet in which no sign of a large-scale magnetic field has ever been detected (Pluto remains *Terra incognita* as far as megnetic fields go). Given what is known of planetary internal structure, only in a few cases (Mercury, Mars) can the magnetic field be assumed to arise from ferromagnetism, in other words a "frozen-in" relic of the formation of the solar system. For all other planets, a dynamo mechanism (part III of this course) must be invoked.

Table 2.1
Planetary magnetic fields and related data

Planet	Radius [km]	Spin period [hr]	Dipole $M/M_\oplus$	Incl.[deg.]
Mercury	2400	1406	5×10^{-4}	+14.0
Venus	6100	5832	$< 10^{-5}$	N/A
Earth	6378	24.0	1	+11.3
Mars	3400	24.7	3×10^{-4}	N/A
Jupiter	71400	9.9	20000	-9.6
Saturn	60300	10.7	600	0
Uranus	25600	17.2	50	-59
Neptune	24800	16.	25	-47

The symmetry axis of the dipolar component of most planetary magnetic fields is usually inclined significantly with respect to the rotation axis, Saturn being an interesting exception to which we shall return in due time. Table 2.1 also illustrates a noteworthy trend, namely the tendency for magnetic fields to become stronger with increasing rotation rate, Mars being here the outstanding exception.

Because they have magnetic fields, solar system planets (again Venus excepted) also have magnetospheres, whose presence is beautifully confirmed by observations of auroral emission in the ultraviolet (see Figure 2.2). Jupiter's magnetosphere is particularly interesting. Besides being the biggest "object" in the solar system (Sun included), it can interact with ionized plasma ejected by volcanic eruptions on Jupiter's moon Io to drive intense electrical current systems by a dynamo process not at all unlike those we will investigate in part III of this course.

Figure 2.2: {`fig:jupsat`} Auroral emission observed on Jupiter and Saturn by the ultraviolet camera on the Hubble Space Telescope. Public domain images courtesy of NASA.

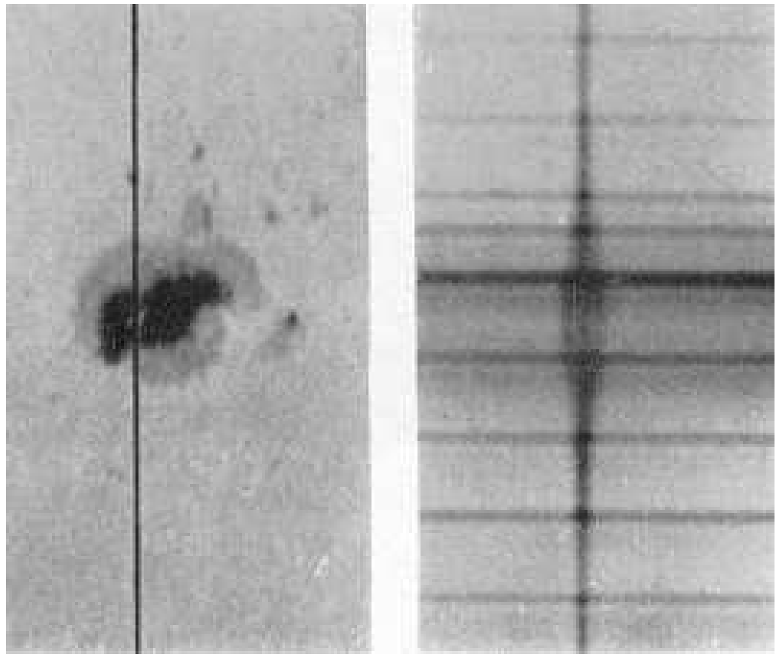

Figure 2.3: {F1.10} The magnetically-induced Zeeman splitting in the spectrum of a sunspot. Reproduced from the 1919 paper by G.E. Hale, F. Ellerman, S.B. Nicholson, and A.H. Joy (in *The Astrophysical Journal*, vol. **49**, pps. 153-178).

2.1.3 The Sun {ssec:solarB}

The Sun is the first astronomical object (Earth excluded) in which a magnetic field was detected, through the epoch-making work of George Ellery Hale (1868-1938) and collaborators, in the opening decades of the twentieth century. In 1907-1908, by measuring the Zeeman splitting in magnetically sensitive lines in the spectra of sunspots and detecting the polarization of the split spectral components (see Fig. 2.3), Hale provided the first unambiguous and quantitative demonstration that sunspots are the seat of strong **magnetic fields**. Not only was this the first detection of a magnetic field outside the Earth, but the inferred magnetic field strength, 3000 Gauss, turned out over a thousand times greater than the Earth's own magnetic field. It was subsequently realized that the pressure provided by such strong magnetic field would also lead naturally to the lower temperatures observed within the sunspots, as compared to the photosphere.

The solar surface magnetic field outside of sunspots, although of much weaker strength, is accessible to direct observations, usually by measuring Zeeman broadening of spectral lines, or the degree of linear or circular po-

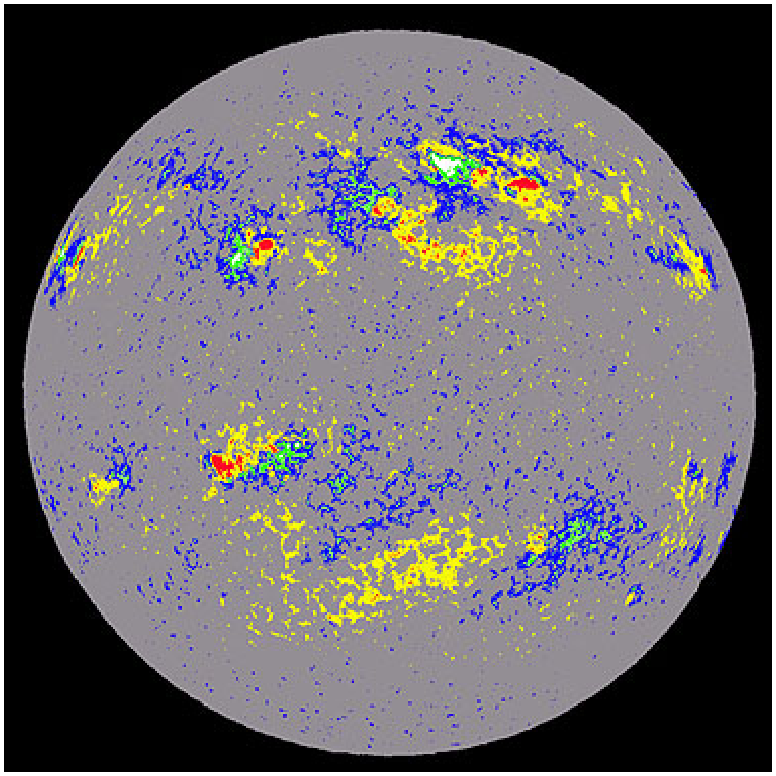

Figure 2.4: {fig:magnetogram} A full-disk solar magnetogram of the sun, showing the coincidence of strong magnetic fields with sunspots, but also the presence of magnetic fields essentially everywhere in the solar photosphere. Note the tilt of two large sunspot pairs in the S-hemisphere. Data courtesy of Jack Harvey, NSO.

larisation of light emitted from the solar photosphere. The first magnetic maps (**magnetograms** of the solar disk were obtained in the late 1950's by the father-and-son team of Harold D. Babcock (1882-1986) and Horace W. Babcock (1912-2003), and were little more than photographs of stacks of a few dozen horizontal scans of the solar disk displayed on an oscilloscope. Figure 2.4 is a modern equivalent in pixel form, with the color scale coding the strength of the normal component of the magnetic field (gray, $\mathbf{B} \lesssim 10\,\mathrm{G}$; yellow to red, positive normal field; blue to green, negative; peaking around 4kG in both cases). The stronger fields coincide with sunspots, but hefty fields of a few hundred Gauss can be found within and around groups of sunspots, as well as in the form of small clumps anywhere else in the photosphere. Far from taking the form of a large-scale, smooth diffuse field as on the Earth, the solar photospheric magnetic field is very fragmented and topologically complex, and shows up concentrated in small magnetized regions separated by field-free plasma. This dichotomy persists down to the smallest spatial scales than can be resolved with current observational techniques. It owes much to the fact that the outer 30% in radius of the Sun is a fluid in a strongly turbulent state.

Because a fraction the solar magnetic field extends into the corona, and because it is dynamically significant there, the equilibrium structure of the corona ends up being defined by a balance between three forces: gravity, plasma pressure, and the Lorentz force. As the photospheric magnetic field inexorably evolves as a result of advection by flows and flux emergence, this equilibrium is eventually lost, leading to rapid and often spectacular disruptions of coronal structures. The associated phenomena are grouped under the general name of **solar activity**, and include phenomena as diverse as flares (Fig. 2.5) and coronal mass ejections (Fig. 2.6). The sun's magnetic field is in fact the primary energy source for the majority of such coronal transients. Saturated as we have become with spectacular images and movies from space-borne solar observing instruments, it is perhaps worth recalling that it took the best part of the twentieth century to establish the causal link between these phenomena and the solar magnetic field, and that is is really only in the mid-1970's, with the X-Ray imager and coronagraph onboard NASA's Skylab , that the coronal *terra incognita* began to be systematically explored.

There is much, much more to be said about the solar magnetic field, its spatiotemporal evolution, and its dynamical impact on the sun's photosphere and extended outer atmosphere. The most prominent temporal variations are certainly those associated with the **solar magnetic activity cycle**, which modulates, on an approximately 11-yr timescale, nearly every solar observable: coronal structures, sunspot coverage, polar field strength, radio emission, irradiance, UV and X-Ray emission, and so on, as well as the frequency of solar eruptive events (flares, coronal mass ejections, eruptive prominences, etc.). We will come back to all of this in due time, but for now we leave the solar system to continue our grand tour of astrophysical magnetic fields...

2.1.4 Sun-like stars {ssec:sunlike}

The disk of solar-type stars other than the sun cannot be spatially resolved, and so direct observation of starspots is not possible, although rotational modulations of the luminosity associated with starspot darkening most certainly can. Direct measurements of magnetic polarisation of starlight is difficult as well, unless the field has a strong large-scale component, otherwise the polarisation associated with regions of opposite polarities —e.g., starspor pairs— cancel out when integrated over the solar disk. Most evidence

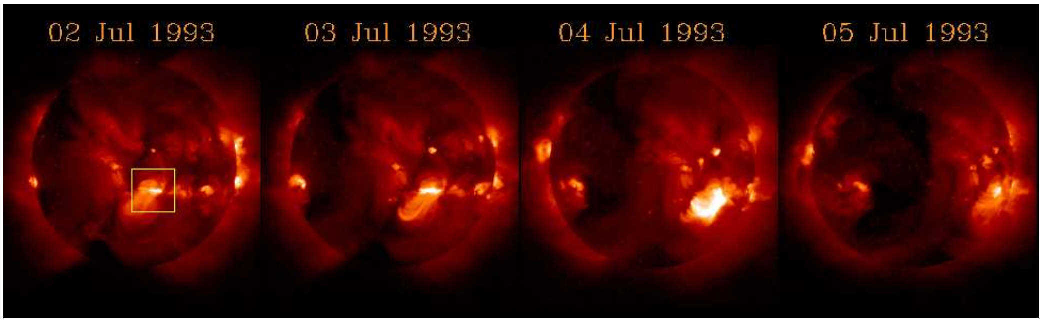

Figure 2.5: {F1.12} A solar flare, as seen in soft X-rays by the satellite YOHKOH. A large flare such as this one can liberate up to 10^{33} erg of thermal energy in the corona over a few minutes. The bulk of that energy goes into local plasma heating and copious emission of short-wavelength radiation. Non-flaring emission of soft X-ray usually coincides with sunspots and active regions. Note also the diffuse, low level coronal X-Ray emission.

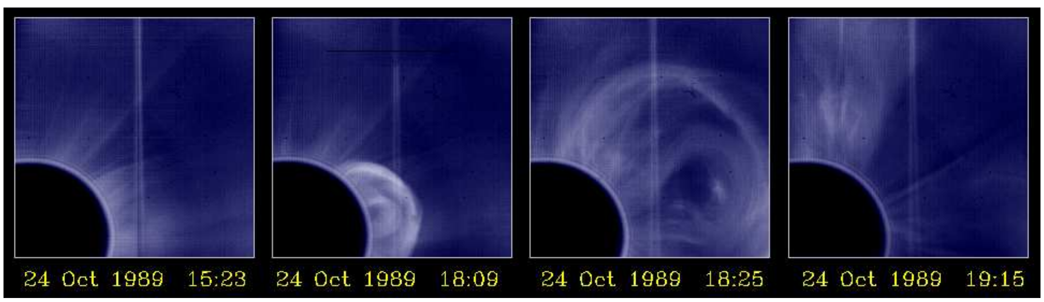

Figure 2.6: {F1.13} A coronal mass ejection (CME), as seen in polarized white light by the coronagraph onboard the Solar Maximum Mission satellite. Large CMEs such as this one can eject up to a few 10^9 tons of ionized plasma at speeds exceeding $10^3\,\mathrm{km\ s}^{-1}$. The occulting disk of the coronograph, on the lower left, has a projected radius of $1.25\,R_\odot$.

for the presence of magnetic fields on such stars is thus indirect, yet extremely compelling, as it covers a wide range of phenomena visible on the sun, such as spectral lines, rotational modulation of luminosity due to the passage of large starspots, flares, radio bursts, and variability in magnetically-sensitive spectral lines on a wide range of timescales.

Figure 2.7 shows a time series of X-Ray emission obtained by the ROSAT satellite, that looks very much like time series of disk-integrated flux observed by the the Earth-orbiting GOES satellites when a solar flare is taking place. The most likely interpretation of Fig. 2.7 is that ROSAT had the good fortune to catch a solar-type star just as it was producing a large flare.

Another magnetic field-related stellar observables that is particularly noteworthy is the emission in the cores of the H and K lines of CaII (396.8nm and

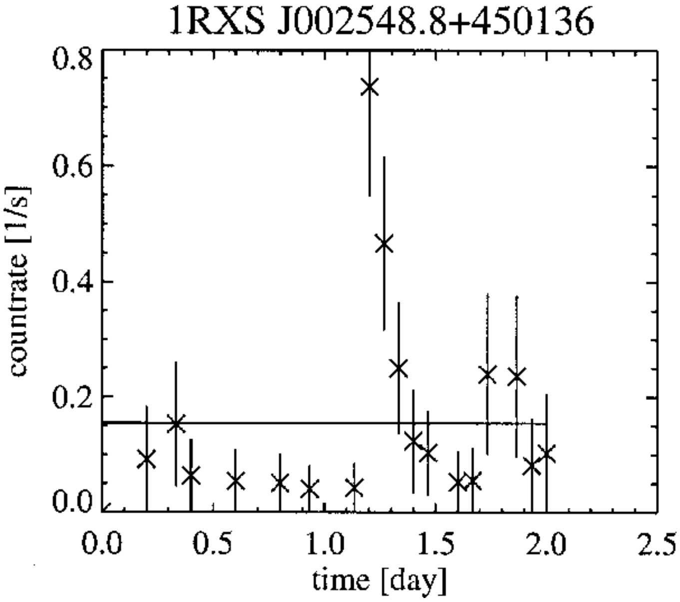

Figure 2.7: {F1.12b} X-Ray emission from a stellar source, as observed from the ROSAT satellite. The rapid rise (minutes) and slower decay (many hours) is similar to what is observed in disk-integrated X-Ray detections of solar flares. Figure reproduced from Fuhrmeister & Schmitt 2003, A&A **403**, 247-260 [Figure 4].

393.4nm, respectively). On the Sun, this emission is known to be associated with non-radiative heating of the upper atmosphere, and is known to scale well with the local photospheric magnetic flux. Starting back in 1968 at Mt Wilson Observatory, Olin C. Wilson (XXXX-YYYY) began measuring the CaII H+K flux in a sample of solar-type stars, a laborious task that was later picked up by a brave group of undeterrable associates and followers, whose collective labor has produced a 40 year long archive of CaII emission time series for no less than 111 stars in the spectral type range F2-M2, on or near the main-sequence.

Figure 2.8 shows a few sample time series of the so-called Calcium index S, mesuring the ratio of core emission intensity in the H and K lines to that of the neighbouring continuum. Some stars show solar-like cycles, others have irregular CaII emission, some show long term trends and others can only be dubbed "flatliners". Among cyclic stars, it was shown that a relatively tight parametric relationship exists between the cycle period ($P_{\rm cyc}$) and rotation period ($P_{\rm rot}$):

$$P_{\rm cyc} \propto \left(\frac{P_{\rm rot}}{\tau_c}\right)^{1.25} , \qquad (2.1) \quad \{\texttt{eq:pcycrot}\}$$

with τ_c being the convective turnover time estimated from mixing length theory of convection. The quantity within parenthesis is related to the so-called Rossby Number, measuring the influence of the Coriolis force on a flow, here convection. As we shall see in due time, such a link between rotation, convection and cycle period is indeed expected from dynamo theory. Later studies have shown that eq. (2.1) is probably an oversimplification, and will return to these remarkable data in part III of the course, when we construct dynamo models for the sun and stars.

The important conclusion here is that the Sun is not some weird oddball: indirect observational evidence for magnetic fields has been found on *every* late-type main-sequence star observed with sufficient sensitivity. Moreover, evidence for solar-like magnetic activity in late-type stars stops rather abruptly around spectral type F0-F2 on the main-sequence, which, according to current stellar structural models, coincides with the disappearance of significant surface convection zones.

2.1.5 Early-type stars

Although most main-sequence stars seem to have gone "magnetically quiet" on the hot side of the dividing line at F0-F2, extant observations suggest a true dichotomy with regards to stellar magnetism in intermediate-mass stars: most A and B stars (around 95%) on or near the main-sequence have no measurable magnetic field, but nearly all those who do combine strong, large-scale magnetic fields, steady on decadal timescales at least with slow rotation and pronounced photospheric abundance anomalies. As we will see later in this course, the presence of a strong, large-scale photospheric magnetic field (ot whatever origin) favors angular momentum loss, and therefore slow rotation; and a strong magnetic field and low rotation favor atmospheric stability, giving full leeway for chemical separation to operate and alter photospheric abundances.

Figure 2.10 shows a particularly well-studied examplar, namely the chemically peculiar star 49Cam. The field strength is high, the magnetic topology quite complex, with the idea of a strongly inclined dipole, historically the common interpretation for Ap stars magnetic fields, being at best a *very* rough approximation.

It is an intriguing fact that the few chemically-normal, (relatively) rapidly rotating early-type stars on which magnetic fields have been detected all sit in the early-B range of spectral types and belong to the βCep sub-class (and include the prototype star βCep itself). However, indirect evidence for photospheric magnetism in O and B star has been accumulating steadily, be it as emission of hard radiation above and beyond what shock dissipation can provide, channelling of stellar winds, and spectral variability. Ongoing

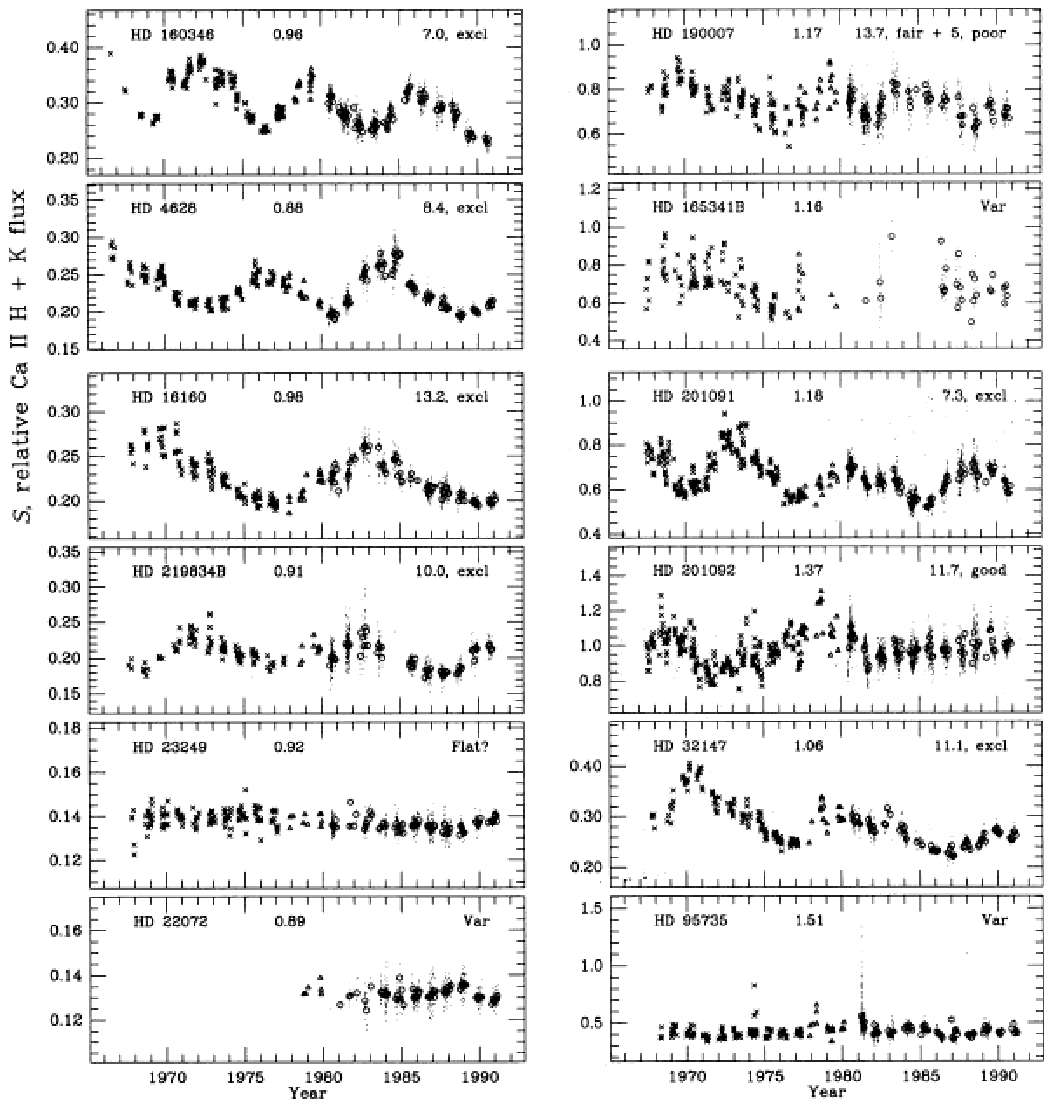

Figure 2.8: {`fig:stellcyc`} Calcium emission index in a small subsample of the Mt Wilson dataset, showing the variety of CaII emission patterns: cycles, non-cyclic irregular emission, long term trend, and constant emission. On such plots, the sun would have a mean emission level $\langle S_\odot \rangle = 0.179$, with a min/max range of about 0.04. Figure cropped from a much larger Figure in Baliunas *et al.* 1995, ApJ,438, 269 [Figure 1g].

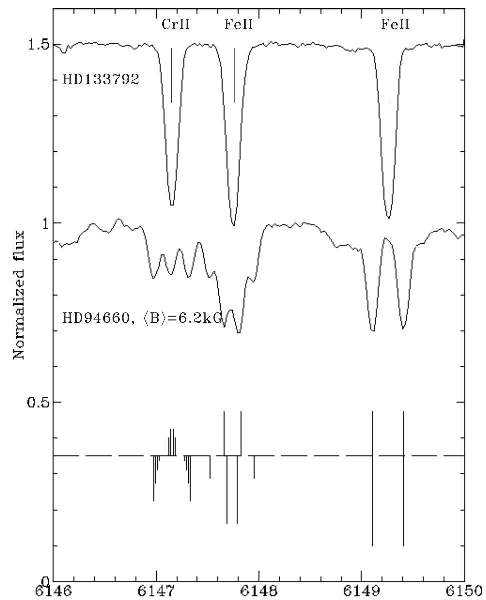

Figure 2.9: {fig:modB} Zeeman splitting of magnetically-sensitive absorption line in the spectrum of the Ap star HD94660. The inferred mean field strength for this star is $<\mathbf{B}>=$ kG. The top trace is that of a typical unmagnetized star of similar spectral type. The horizontal axis is the wavelength, measured in Å. Figure reproduced from the Mathys *et al.* (19XX) paper cited in the bibliography, with a few labels added.

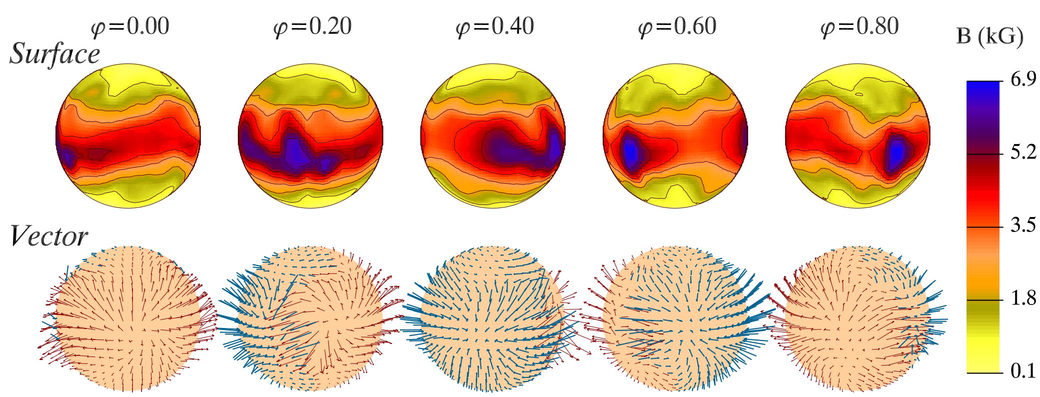

Figure 2.10: {fig:49cam} The surface magnetic field on the Ap star 49Cam, as reconstructed for various rotational phases (φ) by magnetic Doppler imaging. The top row shows the net field strength, and the bottom row the orientation of the surface magnetic field vector. Plot courtesy of J. Silvester and G. Wade, RMC/Kingston.

spectropolarimetric campains targeting massive stars will hopefully provide more data for theoreticians/modellers to chew on in ucoming years.

2.1.6 Pre- and post-main-sequence stars

As with main-squence late-type stars, abundant evidence for magnetic fields in pre- and post-main sequence stars of spectral types later than F has now been accumulating, mostly again in the form of stellar analogs to well-observed solar phenomena: X-Ray and EUV emission, flaring, spectral variability, rotational modulation by starspots, and so on. More recently magnetic Doppler imaging has been used to reconstruct the surface magnetic field of some pre-main-sequence stars in the TTauri evolutionary phase. Whether TTauri or giants, all these stars have low surface temperature and thick convection zones, so observations of magnetic activity indicators similar to what is observed in late-type main-sequence stars points once again to the importance of convection zones of significant radial extent below the photosphere. Indeed, there seems to exist a rather clear-cut, slightly inclined dividing line bisecting the upper part of the HR diagram (main-sequence and up in luminosity), on the right side (low $T_{\rm eff}$) of which evidence of magnetic activity is ubiquitous. Things get messy again with very cool supergiants, with signs of magnetic activity disppearing across various not quite coincident dividing lines, depending on the indicator chosen (X-Ray emission, non-thermal emission lines, etc).

With classical TTauri stars, additional complications also come from the presence of an accretion disk, itself most likely magnetized and perhaps even the site of magntic field generation by dynamo action, and perhaps even magnetically coupled to its central star. Such a coupling has been invoked to explain the (relatively) low rotation rates of TTauri stars, which after all are contracting and accreting large amounts of mass —and angular momentum— from their disk, and should therefore spin up far more than is observed. Indeed, without angular momentum loss mediated by magnetic fields in the early stages of star formation, it is quite likely that stars could simply not eliminate enough angular momentum to form at all!

In hot post-main sequence stars, the observational situation is not well documented or understood. It is a remarkable fact that magnetic fields have been detected in *all* sdO and sdB hot subdwarfs for which a serious attempt has been made. The evolutionary status of these objects is not well-understood, but they most likely represent what used to be the inner core of giants prior to the episode of strong mass loss that accompanies the transition to the horizontal branch. Detection of kG-strength magnetic fields in such stars is strong evodence for the existence of magnetic fields in the deep interior of their main-sequence progenitors.

2.1.7 Compact objects

Magnetic fields in isolated white dwarfs have been detected through circular polarisation measurements in the wings of strong spectral lines, usually Balmer lines in the so-called "DA" white dwarfs showing Hydrogen lines in their photospheres. Actual Zeeman splitting is only detected in the most strongly magnetized objects ($\gtrsim$ a few 10^6 G). Inferred field strengths range from a few tens of kG up to a whopping 10^9 G, with the overall incidence of magnetism standing at a few percent. However, these techniques are only sensitive to large-scale magnetic fields, still producing a net polarisation signal when integrated over the stellar disk, and so the true incidence of magnetism in white dwarfs may actually be significantly higher.

Inferred magnetic field strengths in neutron stars range from 10^8 to 10^{15} G, Neutron stars magnetic fields are of course most readily detected via the pulsar phenomenon, most likely arising from slight misalignement of the magnetic axis with respect to the rotation axis of the (very rapidly rotating) neutron star. It is quite striking that the highest strengths of large-scale magnetic fields in main-sequence stars (a few tens of kG in Ap stars), in white dwarfs ($\sim 10^9$ G) and in the most strongly magnetized neutron stars ($\sim 10^{15}$ G) all amount to similar surface magnetic fluxes, lending support to the idea that these high field strengths can be understood from simple flux-freezing arguments. There is also observational evidence that actual magnetic field evolution is taking place as pulsars age, but this remains very slippery territory, both from the modelling and observational points of view.

Observationally, very little is known about black holes except that there is quite possibly one at the center of our galaxy, so you won't be surprised to hear that even less is known about black hole magnetic fields. One should perhaps just point out that solutions to the field equations of general relativity for rotating, electrically charged black holes do exist, which is a good start towards magnetic fields production. Evidence to date is limited to energetic phenomena interpreted in terms of magnetic channelling of material onto the black hole. But beyond that, at the present time there is only religious fervor.

2.1.8 Galaxies

Magnetic fields in the diffuse, low-density interstellar gas is most readily detected through synchrotron radiation emitted by relavistic charged particles spiralling along magnetic fieldlines. This technique is succesfull not only within the Milky Way, but also for other galaxies. Other means of detection, for the time being limited to the Milky Way, include the polarisation of optical starlight by elongated (i.e., non-spherical) dust grains aligning themselves perpendicularly to magnetic fieldlines; these aligned dust grains also sometimes emit detectable polarized infrared radiation. Finally, for relatively strong fields Zeeman splitting of spectral lines in the radio domain has also been measured. As with stars, magnetic fields seem to be ubiquitous features in pretty much all galaxies.

The galactic magnetic field in the solar neighbourhood has a strength of about 6μG, up to a few tens of μG near galactic center. This is indeed typical of spiral galaxies, which show field strengths in the range 5–15μG, up to some 30μG in high density regions of spiral arms. The strongest large-scale galactic magnetic fields so far measured have strength reaching 100μG, and have been found in starburst galaxies. While this may seem quite low values, such field strengths have important consequences for star formation, the distribution of cosmic rays, and equilibrating the interstellar medium against gravity.

Given that most stars appear to be magnetized to some degrees, and that many stars tend lose mass (some by blowing up!), it is perhaps not surprising to detect magnetic field in the galactic interstellar medium. What is surprising is that this magnetic field tends to be organized on large spatial scales, commensurate in fact with galactic dimensions. An example is shown on Figure 2.11, showing radio intensity isocontours and polarisation vectors superimposed on an optical image of the spiral galaxy M51. Such large-scale, spatially well-organized magnetic fields are most likely produced by a dynamo mechanism, not at all dissimilar to that responsible for the presence of magnetic fields in many stars, including the Sun. We will return to the dynamo origin of galactic fields in the very last chapter of these notes.

2.2 Why B and not E? {sec:BvE}

Even the very brief survey of astrophysical magnetic fields of the preceding section should have made it clear that there are magnetic fields of all sizes and shapes pretty much everywhere we look in the known universe. Yet electric fields are conspicuously absents[1]. Why is that? You might think, looking at Maxwell's equations (1.50)–(1.53) that **E** and **B** appear therein on apparently equal footing, leaving nothing to allow us to anticipate the observed astrophysical preponderance of magnetic fields over electrical fields.

Well, think again. The crucial difference between **E** and **B** in Maxwell's equations is not the fields themselves, but in their *sources*. The Universe may be largely empty (you've heard that one before), but the fact is that is contains a whopping number of elecrically charged particles of various sorts (free electrons, ionized atoms or molecules, photoelectrically charged dust

[1]Excepts in some very remarkable transient phenomena, some of which to be looked into in part IV of this course.

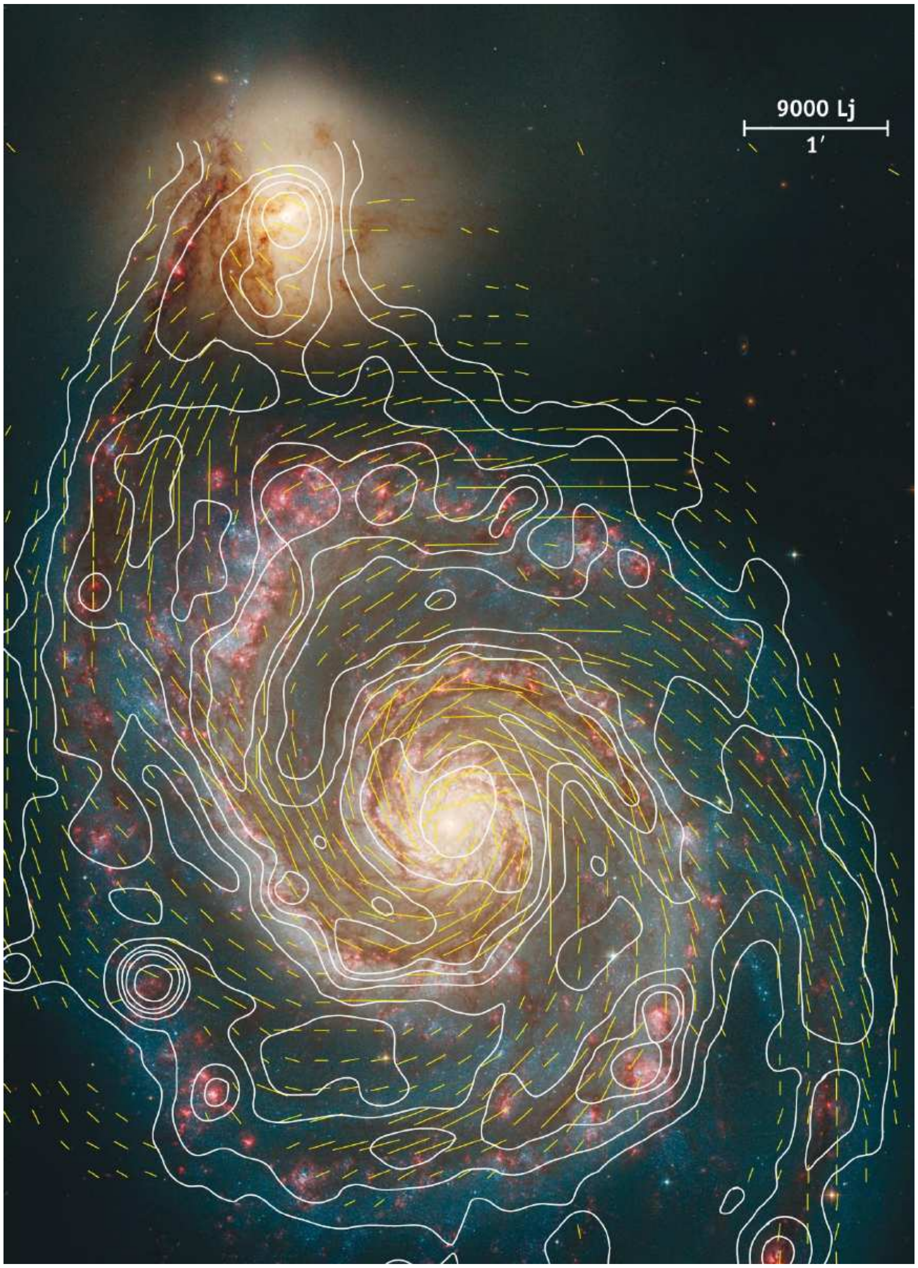

Figure 2.11: {fig:M51} Optical image (Hubble) with overlaid isocontours of radio emission intensity at $\lambda = 6\,\mathrm{cm}$ (in white) and polarisation orientation (orange line segments, both from VLA observations). Note the large-scale organization of the magnetic field, following the optical spiral structure. Image downloaded from the *Scholarpedia* article by Rainer Beck cited in the bibliography at the end if this chapter.

grains, etc). If a large-scale electric field were suddenly to be turned on, all these charges will do the honorable thing, which is to separate along the electric field direction until the secondary electric field so produced cancels the externally applied electric field, at which point charge separation ceases. Moreover, the low densities of most astrophysical plasmas leads to very large mean-free paths for microscopic constituents, leading in turn to fairly good electrical conductivities and very short electrostatic relaxation times τ_e (see eq. (1.68)). In other words, astrophysical electric fields, if and whenever they appear, get shortcircuited mighty fast.

Not so with magnetic fields. For starters, as far as anyone can tell there are no magnetic monopoles out there (well, maybe just one, of primordial origin... more on thi shortly), so shortcircuiting the magnetic field by monopole separation is out of the question. Magnetic fields, left to themselves, will simply decay as the electrical currents that support them (remember Ampère's Law) suffer good ol'Ohmic dissipation. We already obtained a timescale for this process given by eq. (1.61), and we already noted, on the basis of the compilation presented in Table 1.1, that this timescale is extremely large, often exceeding the age of the universe. Once magnetic fields are produced, by whatever means, they stick around for a long, long time.

2.3 Origin of astrophysical magnetic fields {sec:origB}

So, there are magnetic fields all over the place in the Universe. How did they originate? If we stick to MHD, then we immediately hit a Big Problem, arising from the linearity of the MHD induction equation (1.58): if $\mathbf{B} = 0$ at some time t_0 then $\mathbf{B} = 0$ at all subsequent times $t > t_0$, a problem that persists unabated as t_0 is pushed all the way back to the Big Bang.

In part III of this course we will see that astrophysical flows are actually quite apt at amplifying magnetic fields, so what we are after here is a very small "seed field" to start up the process. (Kulsrud REF) Cheap and easy explanations along the line of an original seed magnetic field being a primordial relic of the Big Bang need not concern us here. Nor is early-universe ferromagnetism a viable option, since permanent magnets require an externally-applied magnetic field to become magnetized in the first place. Interestingly, the two options that are deemed viable stand at the opposite ends of the physical exotism scale: magnetic monopoles... and batteries. Let's briefly discuss these in turn[2].

[2]Most of the remainder of this section was written by T.J. Bogdan as part of an earlier version of these class notes.

2.3.1 Magnetic monopoles {ssec:Dirac}

P.A.M. Dirac (1931) pointed out that there is nothing to prevent there from being magnetic monopoles so long as the magnetic charge on a particle is some integer multiple of $g \equiv hc/(4\pi e) \approx 69e$, where h is Planck's constant, and e is the fundamental electric charge. With just *one* magnetic monopole in the universe we have our basic seed field. In the early 1970's, G. t'Hooft and A.M. Polyakov argued that the spontaneous symmetry-breaking of the Grand Unified (field) Theory Lagrangian, which occurs very early in the formation of the universe at $k_B T \approx 10^{15}$ GeV, would produce a *lot* of $m_g \approx 10^{16}$ GeV/c^2 magnetic monopoles.[3] So many in fact that inflationary cosmology was invented in part to deal with this embarrassment of riches and to leave about one monopole within each subdomain of the inflated universe(s). But again, we only need one monopole to produce our seed field, so the realist stops there.

2.3.2 Batteries {ssec:Biermann}

Leaving magnetic monopoles aside, we should inquire about more pedestrian means to create seed magnetic fields. Since it could be that t'Hooft and Polyakov got the wrong Lagrangian, GUT's will be superseded by something else, etc. So it would be nice to have a fall back mechanism to generate a seed magnetic field that relies on basic physics that we know functions sensibly at least in our part of the universe. To this end, we carry on with our (re)derivation of the induction equation. Recall that the next step toward MHD from Maxwell required stipulating Ohm's law,

$$\mathbf{j}_e = \sigma_e \Big[\mathbf{E} + \frac{1}{c}\mathbf{U} \times \mathbf{B}\Big] + \mathbf{J}_{\rm mech} + \cdots . \tag{2.2}$$ {???}

If we keep only the very first term on the RHS of equation (1.6), *and* drop the displacement current in equation (1.5), then we get back to the induction equation (1.1). If we avail ourselves of neither of these opportunities then we have,

$$\Big\{1 + \frac{\eta}{c^2}\frac{\partial}{\partial t}\Big\}\frac{\partial \mathbf{B}}{\partial t} = \nabla \times \Big(\mathbf{U} \times \mathbf{B} - \eta \nabla \times \mathbf{B} + \frac{4\pi\eta}{c}\mathbf{J}_{\rm mech}\Big) + \cdots \tag{2.3}$$ {???}

instead. Notice that our only hope for creating $\mathbf{B}$ out of nothing (so to speak) is the '$\mathbf{J}_{\rm mech} + \cdots$' part of Ohm's law. Retaining the displacement current gives us no advantage.

The $\mathbf{J}_{\rm mech}$ term represents our ability to mechanically grab a hold of electric charges and force currents to flow.[4] In the dense interior of a conducting

[3]You might find it amusing to figure out how many grams that works out to be!

[4]The other term on the RHS of Ohm's law is the collective current response of the dielectric plasma to an imposed electric field $\mathbf{E} + (\mathbf{U} \times \mathbf{B})/c$ in the comoving frame.

star, plasma kinetic theory permits one to write down a prescription for this "battery" contribution to the total electric current density,

$$\mathbf{J}_{\rm mech} = \frac{\sigma_e}{e n_e}\Big[\nabla p_e - \frac{1}{c}\mathbf{j}\times\mathbf{B}\Big] \tag{2.4}$$ {???}

where p_e is the contribution of the electrons alone to the thermal pressure. For a completely ionized pure hydrogen plasma, p_e is just half of the total gas pressure, and $n_e = \rho/m_p$, and so,

$$\mathbf{J}_{\rm mech} = \frac{\sigma_e m_p}{2 e \rho}\Big[\nabla p - \frac{2}{c}\mathbf{j}\times\mathbf{B}\Big]\ . \tag{2.5}$$ {???}

Now the second term on the RHS of equation (1.9) does not do us any good since it carries a factor of $\mathbf{B}$, so the whole plan rests upon the first term generating a seed magnetic field. For a spherically symmetric star, we know from hydrostatic equilibrium that $\nabla\Phi = (\nabla p)/\rho$, and so the product $\eta\mathbf{J}_{\rm mech} \propto \nabla\Phi$. Which does not do us any good because of the presence of the curl operator on the RHS of equation (1.58)!

Back to the drawing board. The battery mechanism failed because $(\nabla p)/\rho$ ended up being the gradient of a scalar function. How can we get around this constraint? Well magnetic fields for one thing would certainly mitigate this situation, but the whole point of the exercise is to try to create magnetic fields out of nothing, so that is not an option. Another possibility is rotation. If the star is rotating, then there is a centrifugal force per unit density of $\varpi\Omega^2\hat{\mathbf{e}}_\varpi$ which adds to $\nabla\Phi$ and which leads to the generation of a seed magnetic field. This process of the centrifugal force driving a flow of electrons relative to the ions was first pointed out by L. Biermann (1950) and is called the *Biermann battery.*

In fact *any* process that can produce a relative motion between the ions and electrons is a potential battery mechanism, and a possible candidate for creating seed magnetic fields. For example, consider a rotating proto-galaxy, where the outer portions of the proto-galaxy move at a speed $U = R\Omega$ relative to the frame in which the microwave background is isotropic. The Thomson scattering of the microwave photons by the electrons results in the so-called *Compton drag effect*, which causes the electrons to counter-rotate with respect to the ions The net result is an azimuthal current which generates a poloidal magnetic field.

Of course, if you bother to put typical numbers in these various examples you will find that you don't really generate very *much* magnetic field. But generating a lot of field is not the point, that's what we are planning to do with the $\mathbf{u}\times\mathbf{B}$ term in our MHD induction equation. The basic idea to take away from this section is that invoking weird, unproven physics is not necessary.

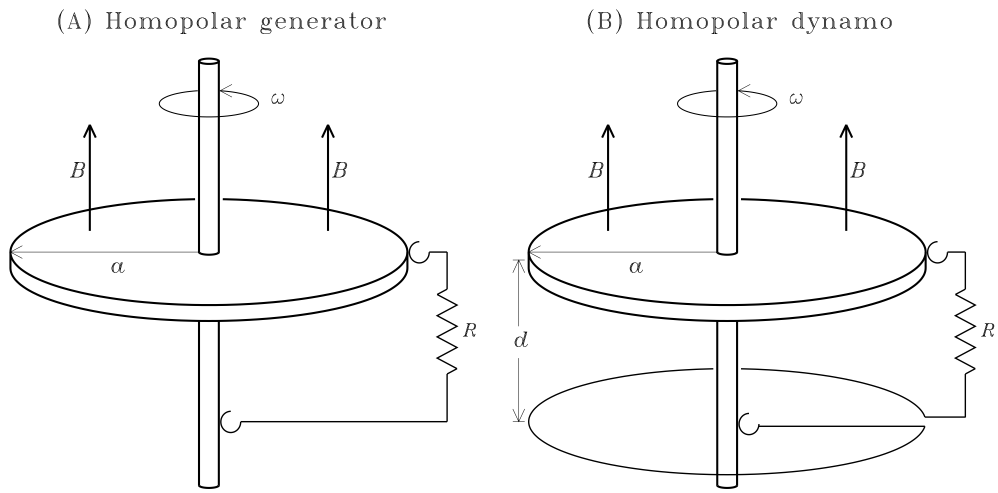

Figure 2.12: {fig:homopolar} A homopolar generator (A) versus a homopolar dynamo (B). An external magnetic field B is applied across a rotating conducting disk, producing an electromotive force that drives a radial current, a wire connecting the edge of the disk to the axle, forming a circuit of resistance R. The only difference between the two electro-mechanical devices illustrated here is that in the latter case, the wire completing the circuit by connecting on the axle is wrapped into a loop in a plane parallel to the disk, so that a secondary vertical magnetic field is produced (see text).

2.4 A simple dynamo {sec:homopolar}

So, astrophysical "batteries" can provide a seed magnetic field on which the inductive action of a flow can, at least in principle, further amplify the field. We shall see in part III that this is indeed possible, although not at all trivial. For now it is we will only consider the following simple example, which illustrates nicely how the idea of amplyfying magnetic field by moving electrical charges across the magnetic field is not so mysterious as one may initially think.

One of the many practical invention of Michael Faraday was a DC electric current generator based on the rotation of a conducting metallic disk threaded by an external magnetic field. Figure 2.12(A) illustrates the basic design: a circular disk of radius a mounted on an axle, rotating at angular velocity ω through the agency of some external mechanical force (like Faraday turning a crank). A vertical magnetic field is imposed across the disk. Electrical charges in the disk will feel the usual Lorentz force $\mathbf{F} = q\mathbf{u} \times \mathbf{B}$ where, (initially) $\mathbf{u}$ is just the motion imposed by the rotation of the disk. Working in cylindrical coordinates (s, ϕ, z) one can write

$$\mathbf{u} = (\omega s)\hat{\mathbf{e}}_\phi \ , \tag{2.6}$$ {eq:homo2a}

$$\mathbf{B} = B_0 \hat{\mathbf{e}}_z \ . \tag{2.7}$$ {eq:homo2b}

so that

$$\mathbf{F} = (q\omega s B_0)\hat{\mathbf{e}}_s \ . \tag{2.8}$$ {eq:homo3}

Now consider the circuit formed by connecting the edge of the disk to the base of the axle via frictionless sliding contacts. With the lower part of the circuit away from the imposed magnetic field, the only portion of the circuit where the magnetic force acts on the charges is within the disk, amounting to an electromotive force

$$\mathcal{E} = \oint_{\text{circuit}} \left(\frac{\mathbf{F}}{q}\right) \cdot \mathrm{d}\boldsymbol{\ell} = \int_0^a \omega B_0 s \mathrm{d}s = \frac{\omega B_0 a^2}{2} \ . \tag{2.9}$$ {eq:homo4}

Neglecting for the time being the self-inductance of the circuit, the current flowing through the resistor is simply given by $I = \mathcal{E}/R$. This device is called a **homopolar generator**.

There is a subtle modification to this setup that can turn this generator into a **homopolar dynamo**, namely a device that converts mechanical energy into self-amplifying electrical currents and magnetic fields. Instead of simply connecting the resistor straight to the axle as on 2.12(A), the wire is wrapped around the axle in a loop lying in a plane parallel to the disk, and then connected to the axle, as shown on 2.12(B). Use your right-hand rule to convince yourself that this current loop will now produce a secondary magnetic field B_* that will superpose itself on the external field B_0. The magnetic flux through the disk associated with this secondary field will be proportional to the current flowing in the wire loop, the proportionality constant being defined as the inductance (M):

$$MI = \Phi = \pi a^2 B_* \ , \tag{2.10}$$ {eq:homo5}

where the second equality comes from assuming that the secondary field is vertical and constant acros the disk; but what really matters here is that $B^* \propto I$ since the geometry is fixed. We now write an equation for the electrical current, this time taking into consideration the counter-electromotive force associated with self-inductance of the circuit:

$$\mathcal{E} - L\frac{\mathrm{d}I}{\mathrm{d}t} = RI \tag{2.11}$$ {eq:homo6}

where L is the coefficient of self-inductance, and the current I is now a function of time. Substituting eqs. (2.9) and (2.10) into this expression, leads to

$$L\frac{\mathrm{d}I}{\mathrm{d}t} = \frac{\omega a^2}{2}\left(B_0 + \frac{MI}{\pi a^2}\right) - RI \tag{2.12}$$ {eq:homo7}

indicating that the current –and thus the magnetic field– will grow provided that initially,

$$\frac{\omega a^2 B_0}{2} > RI \ . \tag{2.13}$$ {eq:homo8}

which it certainly will at first since $I = 0$ at $t = 0$. There will eventually come a time (t_*) when the secondary magnetic field will be comparable in strength to the externally applied field B_0, at which point we may as well "disconnect" B_0; eq. (2.12) then becomes

$$L\frac{\mathrm{d}I}{\mathrm{d}t} = \left(\frac{\omega M}{2\pi} - R\right) I \tag{2.14}$$ {eq:homo9}

which integrates to

$$I(t) = I(t_*) \exp\left[\frac{1}{L}\left(\frac{\omega M}{2\pi} - R\right) t\right] \ . \tag{2.15}$$ {eq:homo10}

indicating that the current —and magnetic field field— will grow provided the externally-imposed angular velocity exceeds a critical value:

$$\omega > \omega_c = \frac{2\pi R}{M} \ . \tag{2.16}$$ {eq:homo11}

This is not a (dreaded) case of perpetual motion, or creating energy out of nothing, or anything like that. The energy content of the growing magnetic field ultimately comes from the biceps of the poor bastard working ever harder and harder to turn the crank and keep the angular velocity ω at a constant value, as you'll get to verify in one of the problem at the end of this chapter in the simpler context of the homopolar generator.

There are many features of this dynamo system worth nothing, and which all find their equivalent in the MHD dynamos to be studied in part III of this course:

1. There exist a critical angular velocity that must be reached for the self-inductance to beat Ohmic dissipation in the resistor, leading to an exponential growth of the magnetic field; below this critical value, the field decays away exponentially once the initial field B_0 is removed.

2. Not all circuits connecting the edge of the disk to the axle will operate in this way; if we suddenly reverse the rotation of the disk, or wrap the wire the other way around the axle, the magnetic field produced by the loop will *oppose* the applied field;

3. The externally applied magnetic field B_0 is only needed as a *seed field* to initiate the amplification process.

4. The homopolar dynamo is really nothing more than a device turning mechanical energy onto electromagnetic energy, more specifically magnetic energy.

Problems:

1. homopolar generator: compute work done against magnetic force by externally applied torque; verify that it is equal to the energy dissipated in the resistor.

Bibliography:

For good introduction on planetary magnetic fields, see

Bagenal, F. 1992, *Ann. Rev. Earth & Planet. Sci.*, **20**, 289-328

augmented by data extracted from the Goedbloed & Poedts textbook cited in the preceding chapter. The following (relatively) recent conference proceedings is devoted to documenting the ubiquitous presence of magnetic fields throughout the Hertzsprung-Russell diagram:

Mathys, G., Solanki, S.K., & Wickramasinghe, D.T. (eds.), *Magnetic fields in the Hertzsprunbg-Russell diagram*, ASP Conf. Series **248** (2001).

See in particular the review papers by XXX (Ap stars), Campbell (binary stars), Valenti & Johns-Krull (cool stars), Schmidt (white dwarfs) Reisenegger (neutron stars), as well as the introductory overview paper by Mestel. On magnetic field detection in hot subdwarfs, see

O'Toole, S.J., Jordan, S., Friedrich, S., & Heber, U. 2005, *Astron. Ap.*, **437**, 227.

On galactic magnetic fields, see the following article (and references therein):

Beck, R. 2007, *Scholarpedia*, **2**(8), 2411
`http://www.scholarpedia.org/articles/Galactic_magnetic_fields`.

Two good, recent textbooks on Solar Physics at the higher undergraduate level are:

Foukal, P., 1990, *Solar Astrophysics*. John Wiley & Sons,
Stix, M., 2001, *The Sun: an introduction*, 2$^{\text{nd}}$ ed., Springer.

You can also consult the ever-being-enlarged Web site "Great Moments in the History of Solar Physics", from which a lot of the material in §§1.1 through 1.4 was precipitated *in extremis*:

http : //www.astro.umontreal.ca/ ∼ paulchar/history.html.

On the telescopic re-discovery of sunspots in the seventeenth century, and ensuing debates over priority and physical interpretation, see

Mitchell, W.M. 1916, "The history of the discovery of the solar spots", in *Popular Astronomy*, **24**, 22-ff,

Shea, W.R. 1970, "Galileo, Scheiner, and the interpretation of Sunspots", *Isis*, **61**, 498-519,

Van Helden, A. 1996, "Galileo and Scheiner on sunspots", in *Proc. Am. Phil. Soc.*, **140**, 358-396,

and especially Galileo's original writings on the topic:

Galileo, G. 1613, *Letters on Sunspots* [in S. Drake (trans.) 1957, *Ideas and Opinions of Galileo*, Doubleday].

Hale's original papers on sunspots are still well worth reading. The two key papers are:

Hale, G.E. 1908, "On the probable existence of a magnetic field in sunspots", *The Astrophysical Journal*, **28**, 315-343,

Hale, G.E., Ellerman, F., Nicholson, S.B., and Joy, A.H. 1919, *The Astrophysical Journal*, **49**, 153-178.

Figure 2.8 was taken from the following paper, still today one of the more cogent exposition of the Mt Wilson CaII project and data:

Baliunas, S.L., Donahue, R.A., Soon, W.H., and 24 co-authors, ApJ, 438, 269,

For a nice, recent review of magnetic field detections in early-type stars, see

Wade, G.A. 2005, in *Element stratification in stars: 40 years of atomic diffusion*, eds. G. Alecian, O. Richard & S. Vauclair, EAS Publication Series, **17**, 227-237.

For some "light" reading on magnetic monopoles in field theory and astrophysics, try,

Dirac, P.A.M. 1931, Proc. R. Soc. Lond. A, 133, 60

Parker, E.N. 1970, ApJ, 160, 383

t'Hooft, G. 1974, Nucl. Phys. B, 79, 276

Polyakov, A.M. 1975, Sov. Phys. JETP, 41, 988

Cabrera, B. 1982, Phys. Rev. Lett., 48, 1378

Kolb, E.W., & Turner, M.S. 1990, *The Early Universe*, (New York: Addison-Wesley), §7.6.

For more on Biermann's battery, see

Biermann, L. 1950, Zeits. f. Naturforsch. A, 5, 65,

Roxburgh, I.W. 1966, MNRAS, 132, 201,
Chakrabarti, S.K., Rosner, R., & Vainshtein, S.I. 1994, Nature 368, 434.

Part II

Magnetized stellar winds

Chapter 3

The solar wind {chap:solwind}

A fool sees not the same tree that a wise man sees
No bird soars too high if he soars with his own wings
If the fool would persist in his folly he would become wise
What is now proved was once only imagin'd

William Blake
The Marriage of Heaven and Hell (1793)

Like rotation and magnetic fields, mass loss is rather ubiquitous across the Herztsprung-Russel diagram. Some stars lose mass in an episodic, often spectacular manner, but most do so in a more calmly, via a **wind** emanating from their surface. Many different physical mechanisms can power a wind, and guess what, magnetic fields often plays an important part in in many of them, as we will explore in the following two chapters. But first we need to establish our baseline wind theory, pertaning to unmagnetized, thermally-driven winds, and towards this goal the Sun is the best starting point, because its wind can be sampled and measured *in situ* by Earth-orbiting satellites.

3.1 Solar and stellar coronae and winds {sec:coronae}

3.1.1 The solar corona {ssec:solcorona}

The story of the solar wind is intimately tied to that of the **solar corona**. The corona being spectacularly visible at times of solar eclipses (see Figure 3.1), we can safely assume that it was first observed a very long time ago by some hairy neanderthal with smelly armpits and questionable table manners. Its first unambiguous description (of the corona, not the neanderthal) is due to the Byzantine chronicler Leo Diaconus (ca. 950-994) who, after witnessing the 22 December 968 solar eclipse, reports:

> ”...at the fourth hour of the day ... darkness covered the Earth and all the brightest stars shone forth. And is was possible to see the disk of the Sun, dull and unlit, and a dim and feeble glow like a narrow band shining in a circle around the edge of the disk.”.

Only by the early decades of the eighteenth century had most astronomers finally convinced themselves that the corona was part of the sun, rather than the moon. The actual name “corona” was coined even later, in 1806, by the Spanish astronomer José Joachin de Ferrer. By the nineteenth century it had become a rite of passage for solar physicists to travel to faraway corners of the Earth to observe solar eclipses, a tradition still very much alive today. Despite rapid advances in spectroscopic and photographic techniques, the physical nature of the corona remained a mystery until the development of the coronagraph by Bernard Lyot (1897-1952) in the early 1930's allowed systematic studies of the corona outside eclipses. By the late 1930's, mostly through the laboratory work of of Walter Grotrian (1890-1954) and Bengt Edlén (1906-1993), the solar corona was recognized as being composed of very hot (1–2×10^6 K) ionized gas. The key in reaching that conclusion was the realization that many of the hitherto unidentified lines seen in coronal spectra were not due to chemical elements unknown on Earth, as believed for a while in the nineteenth century, but rather belonged to high ionization stages of common elements, notably Iron and Nickel. The mechanism(s) through which the corona can be heated to such high temperatures remains, to this day, one of the grand unsolved problems of solar physics.

The peculiar flame-like structures so prominently visible on eclipse photographs such as Fig. 3.1, called **helmet streamers**, are produced by large-scale loop-like magnetic structures emanating from the solar photosphere and trapping the ionized coronal plasma (flux-freezing, remember...). This leads to overdensities in magnetically closed regions of the corona, leading to enhanced Thompson scattering of sunlight, and thus enhanced brightness. The shape of the solar corona varies according to the distribution of photospheric magnetic fields (viz. Fig. 2.4).

Like if a few million degrees K wasn't hot enough already, the corona harbors even hotter plasma, at temperatures sometimes reaching 10 million degrees during transient events called **flares**. We'll have a lot more to say on this topic in part IV of these notes. The take-home message, at this point, is that there is a hot corona out there, and that it is structured at all spatial scales by the solar magnetic field.

3.1.2 The solar wind {ssec:solwind}

The existence of an outflow of matter from the Sun was suggested at the end of the nineteenth century by the Norwegian physicist Kristian Birkeland

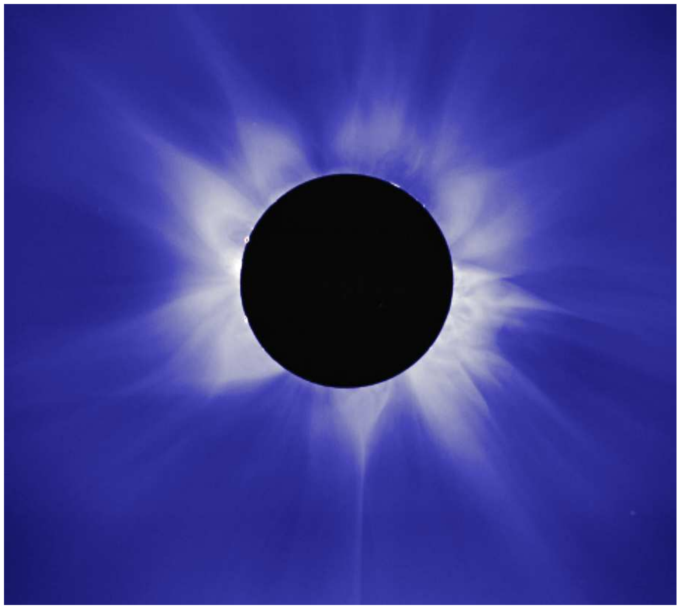

Figure 3.1: {`fig:ecl80`} Total solar eclipse of 16 February 1980, essentially at the maximum phase of the solar activity cycle. Coronal brightness is due to Thompson scattering of sunlight by free electrons, so that on such images brightness is proportional to plasma density. The elongated spiky structures are called helmet streamers, and correspond to regions of closed magnetic fields trapping plasma, eventually pulled open and stretched radially by the solar wind a solar radius or so above the photosphere. Image courtesy of A. Stanger, High Altitude Observatory.

(1867-1917), as an explanation for geomagnetic storms (in particular auroral emission) and zodiacal light. Indeed, by 1899 Birkeland had convinced himself (but unfortunately not a great many others) that interplanetary space was filled with electrically charged particles streaming way from the Sun. The idea did not catch on at the time, but was brought back to the fore half a century later by Ludwig Biermann (1907-1986), as an explanation for the different orientations of neutral and ionized components of cometary tails.

The first quantitative, physical model of what we now call the **solar wind** was proposed in 1958 by Eugene Parker, and led to the surprising prediction that the solar wind should have supersonic speed at the Earth's orbit. This was spectacularly confirmed by the first *in situ* measurements carried out by the Earth-orbiting satellites Lunik 2 (1960), Explorer 10 (1961), and Mariner 2 (1962). Later generations of space probes have now measured solar wind properties out to the far reaches of the solar system (in particular *Pioneer* and *Voyager*), as well as close to the Sun and away from the ecliptic plane

(*Ulysses*).

The physical properties of the solar wind vary significantly on a broad range of timescales; as one can verify from the data summarized in the first columns of Table 3.1 below, at 1 AU fluctuations about the mean are quite large. These large fluctuations are not due to measurements errors. Examination of the distributions of deviations about the mean yields not a Gaussian, but rather a bimodal distributions, indicating that the solar wind exists in to distinct modes, dubbed "low-speed streams" and "high-speed streams". Separating the data in two groups then leads to much smaller deviations about the mean (rightmost columns on Table 3.1). It is now understood that low-speed streams originate from regions of the corona where the magnetic field is mostly closed, while high-speed streams originate from coronal holes, where magnetic fieldlines extend from the solar surface all the way out into the solar system. This was spectacularly demonstrated by the measurements carried out by the space probe *Ulysses* near solar activity minimum, when the solar corona assumes a dipolar shape, with large coronal holes spanning the high heliospheric latitudes in both the Northern and Southern solar hemispheres (see Figure 3.2).

Table 3.1
Observed properties of the solar wind in the ecliptic plane at 1 AU

{tab:solwind}

Quantity	Average	Low-speed	High-speed
N [cm^{-3}]	8.7±6.6 (76%)	11.9±4.5 (38%)	3.9±0.6 (15%)
v [km s^{-1}]	468±116 (25%)	327±15 (5%)	702±32 (5%)
Nv [10^8 cm^{-2} s^{-1}]	3.8±2.4 (63%)	3.9±1.5 (38%)	2.7±0.4 (15%)
ϕ_v (degrees)	−0.6±2.6 (430%)	+1.6±1.5 (94%)	−1.3±0.4 (31%)
T_p (10^5 K)	1.2±0.9 (75%)	0.34±0.15 (44%)	2.3±0.3 (13%)
T_e (10^5 K)	1.4±0.4 (29%)	1.3±0.3 (20%)	1.0±0.1 (8%)
T_α (10^5 K)	5.8±5.0 (86%)	1.1±0.8 (68%)	14.2±3.0 (21%)

The last three lines of Table 3.1 list the (kinetic) temperatures inferred for protons, electrons and He nucleii, the most abundant constituents in the solar wind plasma. These are kinetic temperatures, obtained by measuring particle speeds u and setting

$$kT = \frac{1}{2}mu^2 \ . \tag{3.1}$$ {???}

The fact that kinetic temperatures turn out considerably different for protons and Helium nucleii indicates that the plasma is no longer collision-dominated, meaning we are approaching the limit of our fluid approximation.

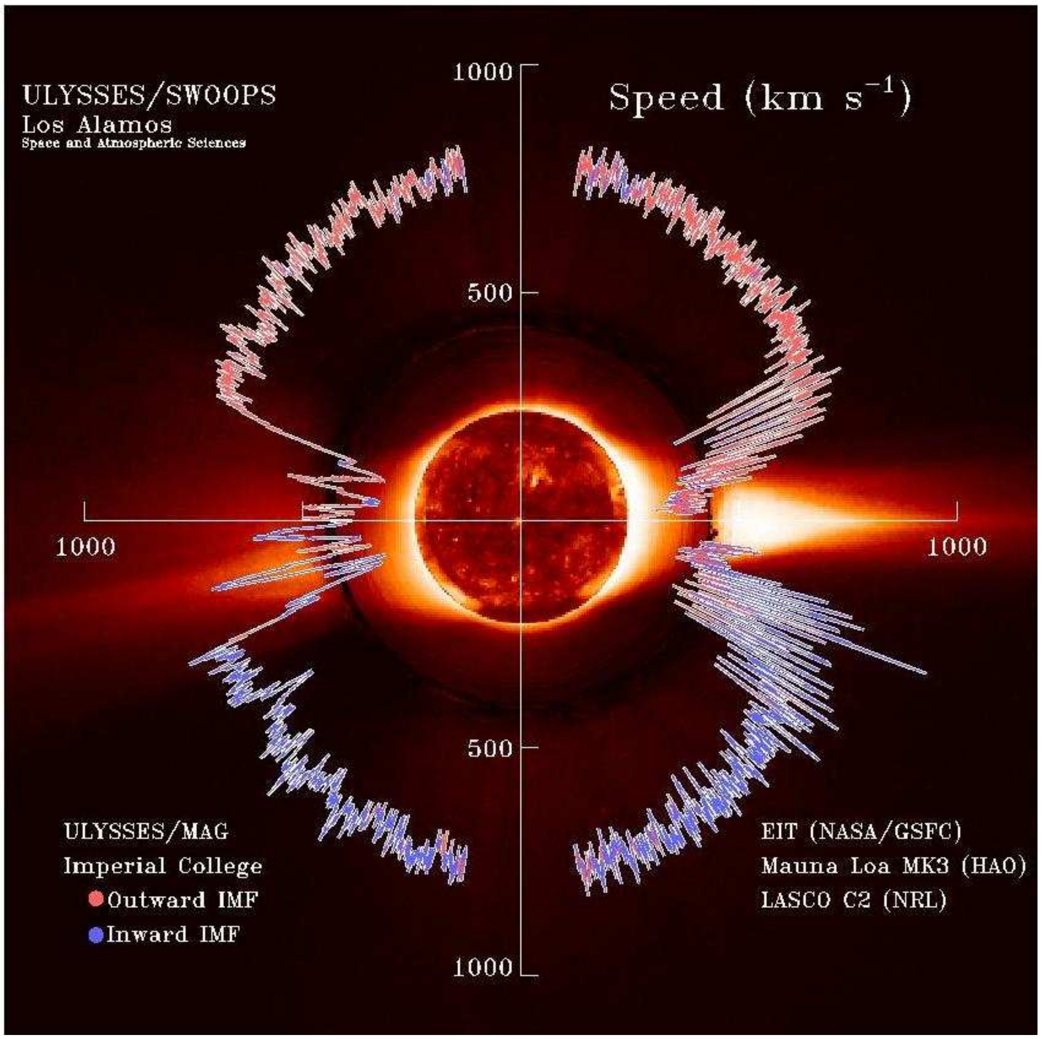

Figure 3.2: {`fig:ulysses`} Image of the solar corona, on which is superposed a polar coordinate plot of the solar wind speed as measured approximately at 1.5 AU by the space probe *Ulysses*. The colors blue/red code the sign of the radial component of the magnetic field measured in the wind. This coronal/wind configuration is typical of activity minimum conditions, with the large-scale coronal magnetic field assuming a dipolar configuration, with a more or less axisymmetric helmet streamer belt straddling the solar equator. The faster wind component emanates from polar coronal holes, where magnetic fieldlines stretch directly out into interplanetary space.

On very short timescales (seconds to minutes), there exist a wide spectrum of fluctuations in all wind variables (flow speed, magnetic field strength and orientation, density, etc.). Based on the type of correlations determined between these various fluctuating variables, a good case can be made that they correspond to a superposition of various type of magnetic and magnetosonic waves (briefly discussed in §1.5). The only type of waves for which a good case can be made for a solar origin are Alfvén waves, and these in fact can have a significant influence on wind dynamics, a topic to be revisited in due time. This is because numerous physical processes could generate waves in the expanding wind itself, and in the low density environment of the expanding solar wind wave-particles are guaranteed to alter the properties of any outgoing wave superimposed on the background flow. On the other hand, the power-law form of the fluctuation spectra is suggestive of turbulence, and an equally good case can be made that MHD turbulence should develop in the solar wind, even if the wind outflow is purely laminar at the coronal base.

3.2 Hydrostatic Corona Model {SHcor}

Since it is an observational fact that there *is* a hot corona out there, our task is now to construct a model allowing us to interpret these observations in a quantitative and coherent way. We start with a simple model, which is almost always a good idea. We assume that the corona is static ($\mathbf{u} = 0$), in a steady-state ($\partial/\partial t = 0$), spherically symmetric ($\partial/\partial\theta = 0$, $\partial/\partial\phi = 0$, $\partial/\partial r \to \mathrm{d}/\mathrm{d}r$), and unmagnetized ($\mathbf{B} = 0$). We construct a solution above a reference radius r_0, at which the density (ρ_0) and temperature (T_0) are assumed known. We also assume that the corona is composed only of fully ionized hydrogen ($m = m_p = 1.67 \times 10^{-24}$ gm, $\mu = 0.5$) obeying the equation of state for a perfect gas.

The r-component of the equations of motion becomes a statement of hydrostatic balance:

$$\frac{\mathrm{d}p}{\mathrm{d}r} = -\rho\frac{GM}{r^2}, \qquad (3.2) \quad \{E3.1.1\}$$

where we have assumed a spherically symmetric gravitational potential $\Phi = -GM/r$. This says nothing more that the (outward-directed) pressure gradient balances exactly the (inward-directed) gravitational acceleration, a particularly simple form of force balance. Assume now that a polytropic relationship exists between the pressure and density:

$$\frac{p}{p_0} = \left(\frac{\rho}{\rho_0}\right)^{\alpha}, \qquad 1 \leq \alpha \leq 5/3 \qquad (3.3) \quad \{E3.1.6\}$$

Using for conciseness the definition of the base polytropic sound speed $c_{s0}^2 = \alpha p/\rho = \alpha k T_0/\mu m$ for a perfect gas, eq. (3.2) now becomes

$$c_{s0}^2 \left(\frac{\rho}{\rho_0}\right)^{\alpha-1} \mathrm{d}\rho = -\rho \frac{GM}{r^2} \mathrm{d}r, \qquad (3.4) \quad \{\texttt{E3.1.7}\}$$

which is readily integrated to yield an expression for the density profile

$$\frac{\rho(r)}{\rho_0} = \left[1 - \frac{(\alpha-1)GM}{r_0 c_{s0}^2}\left(1 - \frac{r_0}{r}\right)\right]^{1/(\alpha-1)}, \qquad (3.5) \quad \{\texttt{E3.1.8}\}$$

from which the pressure profile is immediately obtained via eq. (3.3):

$$\frac{p(r)}{p_0} = \left[1 - \frac{(\alpha-1)GM}{r_0 c_{s0}^2}\left(1 - \frac{r_0}{r}\right)\right]^{\alpha/(\alpha-1)}, \qquad (3.6) \quad \{\texttt{E3.1.9}\}$$

and the temperature profile from the equation of state:

$$\frac{T(r)}{T_0} = \left[1 - \frac{(\alpha-1)GM}{r_0 c_{s0}^2}\left(1 - \frac{r_0}{r}\right)\right]. \qquad (3.7) \quad \{\texttt{E3.1.10}\}$$

Examination of these expressions reveals that there may be combinations of T_0 and α values that yield zero pressure and density at a *finite* value of r. Obviously, this occurs whenever

$$c_{s0}^2 < (\alpha - 1)GM/r_0, \qquad (3.8) \quad \{\texttt{E3.1.11}\}$$

with

$$\frac{r_{\rm top}}{r_0} = \left(1 - \frac{r_0 c_{s0}^2}{(\alpha-1)GM}\right)^{-1} \qquad (3.9) \quad \{\texttt{E3.1.12}\}$$

being the maximum radial extent of the polytropic atmosphere. Eqs. (3.5) through (3.7) describe a static polytropic atmosphere occupying the volume $r_0 \leq r \leq r_{\rm top}$. For $r > r_{\rm top}$ there is only mathematical vacuum, something Nature abhors, or so Aristotle used to claim. What if T_0 is too large for eq. (3.8) to be satisfied ? Figure 3.3 illustrates a series of polytropic solar coronal models, for $T_0 = 1.5 \times 10^6\,$K, $r_0 = 1.15\,R$, fully ionized hydrogen, and various values of α (for this adopted value of T_0 and for solar parameters, satisfying eq. (3.8) requires $\alpha > 1.1765$). It looks like the solutions that violate eq. (3.8) extend to infinity with non-vanishing pressures and densities. From eq. (3.6) one immediately obtains

$$p_\infty \equiv \lim_{r\to\infty} \frac{p}{p_0} = \left[1 - \frac{(\alpha-1)GM}{r_0 c_{s0}^2}\right]^{\alpha/(\alpha-1)}, \qquad (3.10) \quad \{\texttt{E3.1.16b}\}$$

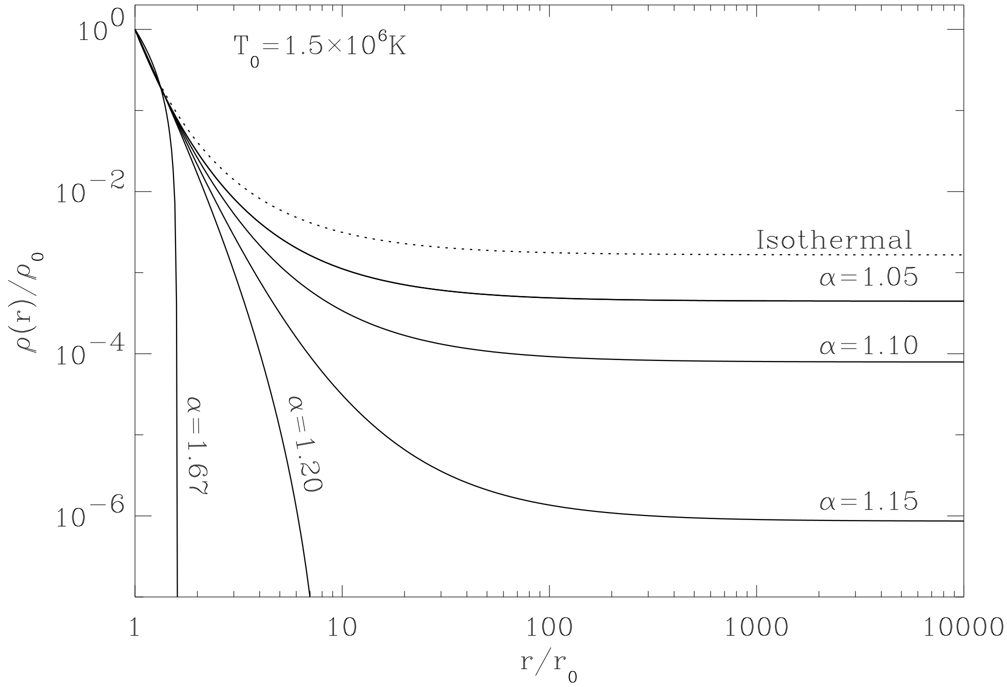

Figure 3.3: {`F3.1`} Density profiles for a few polytropic static coronal model with $T_0 = 1.5 \times 10^6$ K. and $r_0 = 1,15\,R$. Note the asymptotically constant densities as $r \to \infty$ for $\alpha < 1.1765$.

and similar expressions (with different exponents) for the asymptotic density and temperature. For the parameter values used on Fig. 3.3 and $N_0 = \rho_0/(\mu m_p) = 10^8\,\mathrm{cm}^{-3}$, one obtains $\rho_\infty = 10^4\,\mathrm{cm}^{-3}$, $p_\infty = 8\times10^{-7}\,$dyne cm^{-2}, and $T_\infty = 6 \times 10^5$ K. These values are much larger than anything the interstellar medium has to offer. In the solar galactic neighborhood, typical densities and temperatures are believed to be $N_{\rm ism} = 1\,\mathrm{cm}^{-3}$ and $T_{\rm ism} = 100$ K, so that $p \sim 10^{-14}$ dyne cm^{-2}. All of these are insufficient by orders of magnitude[1]. Something is deeply wrong. Given the assumptions made in constructing our simplistic model, three avenues are open to “save” our model:

1. Abandon the hypothesis of a steady-state ($\partial/\partial t = 0$) corona,

2. Work on the energetics to produce a corona with a different asymptotic temperature profile,

[1]Actually, a realistic estimate of the *total* pressure in the interstellar medium should take into consideration the contribution of the interstellar magnetic field. Far from being negligible, magnetic pressure can provide $\sim 10^{-12}$ dyne cm^{-2} for $\|\mathbf{B}\|_{\rm ism} \sim \mathbf{10^{-5}}$ G. But this is still insufficient to equilibrate our hot hydrostatic corona.

3. Abandon the hypothesis of a static ($\mathbf{v} = 0$) corona.

Possibility (1) flies in the face of observations, at least as far as the larger spatial scales are concerned. Early efforts (and efforts to come, as per see problems 1.1 and 1.2...) were mostly directed along avenue (2). Yet avenue (3) proved to be the right one.

3.3 Polytropic winds {sec:Swind}

In this section we will construct a simple, yet reasonably realistic, solar wind model, which will turn out to do a surprisingly good job at reproducing a lot of the large-scale flow properties of the real solar wind. The same underlying physical mechanism turns out to be responsible for the winds emanating from the atmospheres of the polar terrestrial ionosphere, of the atmosphere of other late-type stars, and from the galactic halo. So pay attention to this one.

3.3.1 The Parker Solution {ssec:Parker}

We follow the initial approach of E.N. Parker, in seeking steady state ($\partial/\partial t = 0$) solutions that are spherically symmetric ($\partial/\partial\theta = 0$, $\partial/\partial\phi = 0$). This also implies $u_\theta = 0$, $u_\phi = 0$ (think about it a bit). We *assume* that the star is surrounded by a hot corona (temperature $\sim 10^6$ K), as in §3.2 extending outward from a reference radius $r = r_0$ where the base temperature (T_0) and density (ρ_0) are assumed known. We seek a wind solution in the domain $r \in [r_0, \infty]$. We consider an inviscid ($\tau = 0$), unmagnetized ($\mathbf{B} = 0$) plasma. We will also limit ourselves to a *single fluid* model. That is, we consider a wind composed exclusively of fully ionized Hydrogen where charge neutrality always holds down to the smallest spatial scales considered[2]. This implies that the proton-electron mixture can be treated as a single fluid, with each particle having a mass μm_p, with $\mu = 0.5$. Once again we make the further simplifying assumption that the flow is *polytropic*, i.e., the pressure and density are assumed to be related by a relation of the form

$$\left(\frac{p}{p_0}\right) = \left(\frac{\rho}{\rho_0}\right)^\alpha , \tag{3.11}$$ {E3.3.4}

or, equivalently,

$$\frac{\mathrm{d}}{\mathrm{d}r}\left(\frac{p}{\rho^\alpha}\right) = 0 , \tag{3.12}$$ {E3.3.4b}

[2]Does that also mean that protons and electrons must have identical bulk velocities? Think about that one a bit.

with α constant and specified *a priori* (cf. §1.3). This implies that the sound speed varies with heliocentric radius as

$$c_s^2(r) = c_{s0}^2 \left(\frac{\rho}{\rho_0}\right)^{\alpha-1} , \qquad (3.13) \quad \{\texttt{E3.3.5}\}$$

where $c_{s0}^2 = \alpha p_0/\rho_0$ is the sound speed at the reference radius. In view of the spherical symmetry assumption, the mass conservation equation reduces to

$$\frac{1}{r^2}\frac{\partial}{\partial r}(r^2 \rho u_r) = 0 \ , \qquad (3.14) \quad \{\texttt{E3.3.1}\}$$

which integrates directly to

$$\rho r^2 u_r = \text{const.} \qquad (3.15) \quad \{\texttt{E3.3.2}\}$$

We only have to deal with the r-component of the equations of motion:

$$\rho u_r \frac{\partial u_r}{\partial r} = -\rho \frac{GM}{r^2} - \frac{\partial p}{\partial r}, \qquad (3.16) \quad \{\texttt{E3.3.3}\}$$

assuming again a spherically symmetric gravitational potential $\Phi = -GM/r$. Upon making use of eqs. (3.11) and (3.13), equation (3.16) can be manipulated into the form

$$\frac{\partial u_r}{\partial r} = \frac{u_r}{r}\left[\frac{2c_s^2 - GM/r}{u_r^2 - c_s^2}\right] . \qquad (3.17) \quad \{\texttt{E3.3.6}\}$$

Now, the denominator of eq. (3.17) vanishes when the flow speed becomes equal to the *local* sound speed. This means that the numerator must simultaneoulsy vanish to avoid the appearance of (unwanted) infinite accelerations. The radius r_s at which this occurs is called the *sonic point*, and is located at

$$r_s = \left(\frac{1}{c_{s0}^2}\right)^{2/(5-3\alpha)} \left(\frac{GM}{2}\right)^{(\alpha+1)/(5-3\alpha)} \left(\frac{1}{u_{r0} r_0^2}\right)^{2(\alpha-1)/(5-3\alpha)} , \qquad (3.18) \quad \{\texttt{E3.3.8}\}$$

where u_{r0} is the base flow speed. At the sonic point we also have

$$u_{rs} = c_s(r_s) = \left(\frac{GM}{2r_s}\right)^{1/2} . \qquad (3.19) \quad \{\texttt{E3.3.9}\}$$

Now, eq. (3.16) can be rewritten as

$$\frac{\partial}{\partial r}\left[\frac{u_r^2}{2} + \frac{c_s^2}{\alpha - 1} - \frac{GM}{r}\right], \qquad (3.20) \quad \{\texttt{E3.3.7a}\}$$

which immediately integrates to

$$\frac{u_r^2}{2}+\frac{c_s^2}{\alpha-1}-\frac{GM}{r}=E \qquad (3.21) \quad \{E3.3.7\}$$

Equation (3.21) is the *Bernoulli equation*, and the integration constant E is the energy per unit mass in the flow. The Bernoulli equation contains the essence of solar wind acceleration: thermal energy of the gas ($c_s^2/(\alpha-1)$) gets converted to gravitational potential energy (GM/r) and flow kinetic energy ($u_r^2/2$), while the total energy is conserved (as it should!). Since the sound speed c_s can be expressed entirely as a function of r and u_r via the mass concervation equation, a solution $u_r(r)$ is then *any* functional $u_r(r)$ satisfying eq. (3.21), for *any* given value of E. But how do we pick an appropriate value for this quantity?

3.3.2 Computing a solution {SSWsol}

The key in constructing a wind solution is to realize that any non-singular transsonic solution *must* pass through the sonic point. Let us then begin by writing down expressions for E evaluated at the base of the flow and at the sonic point:

$$E(r_0,u_{r0})=\frac{u_{r0}^2}{2}-\frac{GM}{r_0}+\frac{c_{s0}^2}{\alpha-1}, \qquad (3.22) \quad \{E3.3.9b\}$$

$$E(r_s,u_{rs})=-\frac{3GM}{4r_s}+\frac{c_{s0}^2}{\alpha-1}\left(\frac{u_{r0}r_0^2}{\sqrt{GM/2r_s}r_s^2}\right)^{\alpha-1}, \qquad (3.23) \quad \{E3.3.9c\}$$

where we made good use of eq. (3.19). Equating The RHSs of these two expressions yields a nonlinear rootfinding problem for u_{r0}, which can be written schematically as

$$E(r_s,u_{rs})-E(r_0,u_{r0})=0. \qquad (3.24) \quad \{E3.3.10\}$$

This root finding problem is not particularly easy, in view of the fact that the sonic point r_s itself a nonlinear function of u_{r0} (as per eq. (3.18). The *bisection method* (see Box N.1) is a simple, robust, and easy to code algorithm that works fine here. The solution of eq. (3.24) yields the base flow speed u_{r0} for the transsonic solution, which then allows to compute r_s and E_s. Once u_{r0} is known, computing $u_r(r)$ proceeds by solving a new nonlinear rootfinding problem for u_r defined by setting $E(r,u_r)-E(r_0,u_{r0})=0$, with r ($>r_0$) given and r_s now known via eq. (3.18). At this juncture note also that the location of the sonic point is entirely determined by the assumed base sound

Refer to Numerical box N.1: the bisection method speed c_{s0} i.e., by the coronal base temperature T_0, and polytropic index α; equally important, u_{r0} is *not* an input parameter of the solution.

Box N.1: The bisection method
Given a nonlinear function of one variable $f(x)$ and a range $[x_1, x_2]$ in which we seek a root x_r such that x_r is within within a (predetermined) absolute accuracy ε of the true root. The following is a pseudocode for the bisection method:

while $\delta \geq \varepsilon$ **do**
    $x_m = (x_1 + x_2)/2$
    if $f(x_2) \times f(x_m) \geq 0$ **then**
        $x_2 = x_m$
    else
        $x_1 = x_m$
    endif
    $\delta = x_2 - x_1$
enddo

The nice thing about the bisection method is that it is *guaranteed* to find the root *if* a root indeed exists in the (user-specified) interval $[x_1, x_2]$. Its chief drawback is that it converges linearly to the root, which, if high accuracy is required, can take a lot of iterations.

Chapter 9 of the book *Numerical Recipes* (see bibliography) contains a very accessible introduction to the numerical solution of nonlinear rootfinding problems, including the bisection method. You will also learn therein about a neat trick called *bracketing* that ensures that the initial bisection interval $[x_1, x_2]$ always contains at least one root (if any root at all exists in $x \in [-\infty, \infty]$).

Should we put relevant references at the end of boxes?

With E_s now a fixed quantity, what happens for solutions that start off with different values of u_{r0} and c_{s0}, subjected to the constraint $E(r_0, u_{r0}) = E_s$? Figure 3.4 shows the family of solutions obtained in this manner. There are in fact *two* transsonic solutions (thicker lines) that cross at the sonic point. The accelerating solution is the one we are after for the solar wind. The deccelerating solution has a lower base temperature ($T_0 = 8.7 \times 10^5\,\mathrm{K}$), to compensate for its much higher base flow speed ($u_{r0} = 477.7\,\mathrm{km\,s^{-1}}$). The two transsonic solutions partition the $[r, u_r]$ plane in four distinct regions. Region I correspond to solutions that are supersonic everywhere including at the coronal base; in the solar context, such solutions, as well as the deccelerating transonic solution, conflict with the lack of significant blueshift observed in coronal spectral lines. Regions II and IV do not contain outflow solutions. This leaves the accelerating transsonic solution and solutions in region III as possible valid outflow solutions for the solar wind.

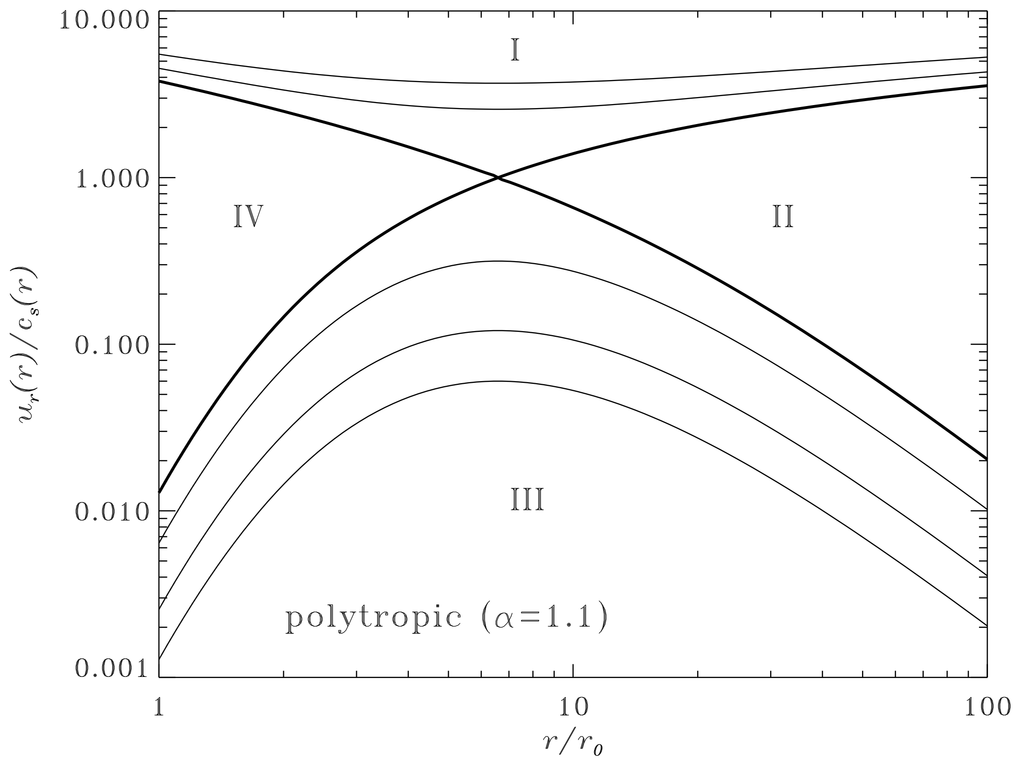

Figure 3.4: {FTopo} Some solutions to equation (3.21). The thick lines are the two transsonic solution satisfying eq. (3.24). The accelerating transsonic solution, to be identified with the solar wind, has a base flow speed $u_{r0} = 2.12\mathrm{km\,s^{-1}}$, sound speed $c_{s0} = 165.1\,\mathrm{km\,s^{-1}}$, with the sonic point located at $r_s/r_0 = 6.59$. The thin lines are solutions for other values of E ($\neq E_s$), and are labeled in units of E_s.

Once $u_r(r)$ is known, it is straightforward to obtain expressions for the density, pressure, and temperature profiles:

$$\frac{\rho(r)}{\rho_0} = \left[1 - \frac{(\alpha-1)GM}{r_0 c_{s0}^2}\left(1 - \frac{r_0}{r}\right) - \frac{(\alpha-1)}{2c_{s0}^2}\left(u_r^2 - u_{r0}^2\right)\right]^{1/(\alpha-1)}, \quad (3.25) \quad \texttt{\{E33.ws1a\}}$$

$$\frac{p(r)}{p_0} = \left[1 - \frac{(\alpha-1)GM}{r_0 c_{s0}^2}\left(1 - \frac{r_0}{r}\right) - \frac{(\alpha-1)}{2c_{s0}^2}\left(u_r^2 - u_{r0}^2\right)\right]^{\alpha/(\alpha-1)}, \quad (3.26) \quad \texttt{\{E33.ws1b\}}$$

$$\frac{T(r)}{T_0} = \left[1 - \frac{(\alpha-1)GM}{r_0 c_{s0}^2}\left(1 - \frac{r_0}{r}\right) - \frac{(\alpha-1)}{2c_{s0}^2}\left(u_r^2 - u_{r0}^2\right)\right]. \quad (3.27) \quad \texttt{\{E33.ws1c\}}$$

Note that these expressions are valid for either the transsonic or class-III solutions. The latter evidently have $\lim_{r\to\infty} u_r \to 0$ (see Fig. 3.4), so that

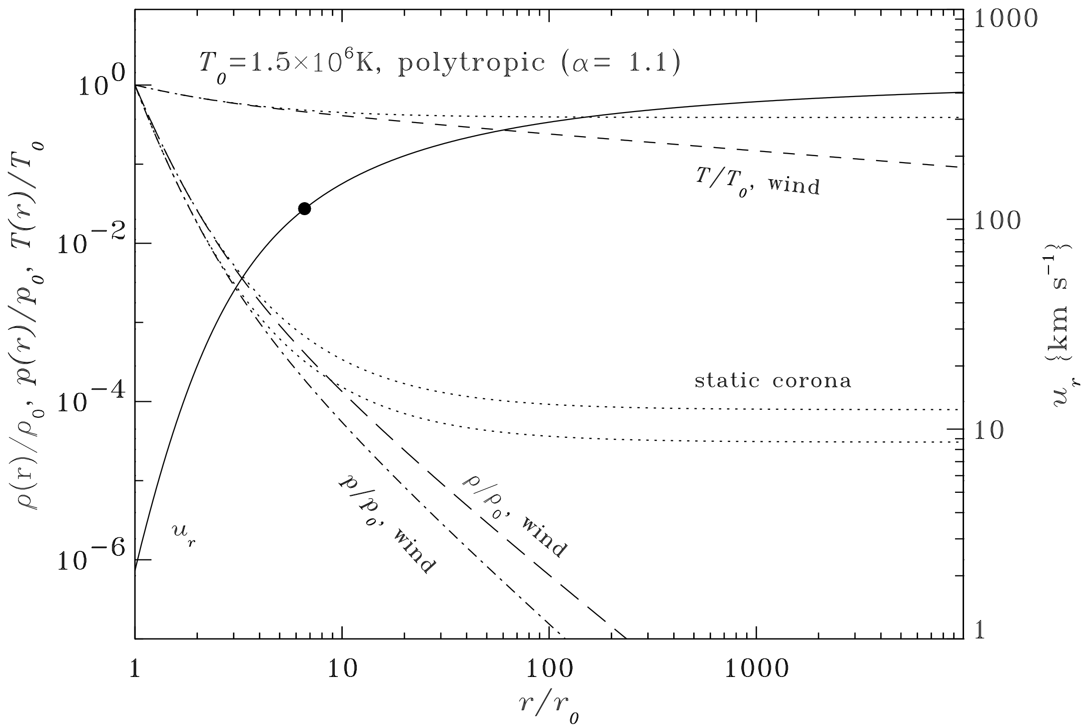

Figure 3.5: {F3.Wsol} The full wind solution corresponding to the accelerating transsonic solution of Fig. 3.1. Dotted lines correspond to a polytropic static coronal model with identical α and T_0. The solid dot indicates the location of the sonic point.

asymptotically, eq. (3.26) becomes identical to eq. (3.10), obtained for a *static* corona! The class-III solutions thus suffer from the same shortcoming: an asymptotic pressure much too high to match that of the interstellar medium, and so can be ruled out.

This leaves us with a single possible outflow solution, namely the accelerating transsonic solutions, which we hereafter refer to as the "wind solution"[3]. Figure 3.5 illustrates the variations with radial distance of the density, pressure and temperature for the transsonic wind solution of Fig. 3.4, together with the corresponding profiles for a $\alpha = 1.1$ polytropic static corona of identical base temperature (dotted lines). Within the sonic point the structure of the solution is very much like that of a static atmosphere, while for $r > r_s$ the solutions differ markedly, reflecting the dynamical effect of the outflow.

[3]Traditionally, class-III solutions have been dubbed "solar breeze", since the flow speed they predict at the Earth's orbit is much smaller than for the wind solution

3.3.3 Mass loss {ssec:mloss}

One important consequence of the existence of a wind is that it carries away mass from the star. Under the assumption of spherical symmetry used here, the mass loss rate is

$$\dot{M} = 4\pi r_0^2 \rho u_{r0} \ , \qquad [\mathrm{gm\,s^{-1}}] \ . \tag{3.28}$$ {eq:solmassloss}

For the solar-type solution considered here, $\dot{M} = 10^{-14}\, M_\odot\, \mathrm{yr}^{-1}$, so that over its lifetime the Sun would lose a mere 10^{-4} fraction of its total mass, assuming that this mass loss rate has remained constant since the Sun's arrival on the ZAMS; as we shall see later, there are good reasons to believe that the ZAMS mass loss rate may have been substantially higher.

3.3.4 Asymptotic behavior and existence of wind solutions {SSAsymp}

To analyze the asymptotic behavior of the wind solution we do something undoubtedly familiar by now: we equate $E(r, u_r)$ evaluated at r_0 and in the limit $r \to \infty$:

$$\frac{u_{r0}^2}{2} - \frac{GM}{r_0} + \frac{c_{s0}^2}{\alpha - 1} = \lim_{r\to\infty} \left[\frac{u_r^2}{2} - \frac{GM}{r} + \frac{c_{s0}^2}{\alpha - 1} \left(\frac{u_{r0} r_0^2}{u_r r^2} \right)^{\alpha - 1} \right] . \tag{3.29}$$ {E33.As1}

Now, what the wind solution does is convert all thermal energy in excess of what is needed to climb out of the Sun's gravitational potential well into bulk flow kinetic energy. This implies $u_r \gg c_s$ asymptotically. Furthermore, we also have $\lim_{r\to\infty} u_r \gg u_{r0}$ and $u_{r0} \ll c_{s0}$, so that eq. (3.29) readily yields

$$\lim_{r\to\infty} u_r \equiv u_{r\infty} = \left(\frac{2c_{s0}^2}{\alpha - 1} - \frac{2GM}{r_0} \right)^{1/2} . \tag{3.30}$$ {E33.As2}

indicating that the flow speed becomes constant at large r.

Clearly all wind solutions must have $u_{r\infty} > 0$ for finite r, so that we have the constraint[4]:

$$\frac{c_{s0}^2}{\alpha - 1} - \frac{GM}{r_0} \geq 0. \tag{3.31}$$ {E33.As3}

For $\alpha = 1.1$, this requires that $T_0 \gtrsim 9 \times 10^5\,\mathrm{K}$. An accelerating transsonic solution also requires $\mathrm{d}u_r/\mathrm{d}r > 0$ near the base of the flow. Going back to eq. (3.17), this implies

$$2c_{s0}^2 - \frac{GM}{r_0} < 0, \tag{3.32}$$ {E33.As4}

[4]Why is that so ? What's wrong with a steady, spherically symmetric wind accelerating in the corona and then, at some large distance, deccelerating again until everything grinds to a full stop ?

requiring that $T_0 \lesssim 5 \times 10^6$ K. So a transsonic wind can only exist for a base temperature in the range

$$\left(\frac{\alpha-1}{\alpha}\right)\frac{GM\mu m_p}{kr_0} \leq T_0 < \left(\frac{1}{2\alpha}\right)\frac{GM\mu m_p}{kr_0} \tag{3.33}$$ {E33.As5}

What do these two bounds on T_0 correspond to physically? The lower bound is simply the criterion for the existence of a gravitationally bound atmosphere, which we encountered already in §3.2, in fact. The upper bound is trickier to interpret. It represents the temperature above which steady, transsonic wind solutions no longer exist. If this criterion were to be violated (e.g. by a sudden increase in base temperature), the whole atmosphere would "explode" outward in a very time-dependent manner. Indeed, we will revisit this issue in part IV of this course.

There is something else that is extremely important that can be extracted from eq. (3.33); if there is to be a finite temperature interval over which it is to be satisfy, then we *must* have $\alpha > 3/2$. Otherwise both criteria cannot be satisfied simultaneously. We therefore have the additional constraint $1 \leq \alpha \leq 3/2$, independently of the assumed base temperature T_0.

Figure 3.6 illustrates, in the $[T_0, \alpha]$ plane, the region in parameter space where steady, transsonic solutions are allowed. The thermodynamically allowed bounds on α ($\leq \alpha \leq 5/3$ for a perfect monoatomic gas) restrict solutions to the region located below the dotted line. Equation (3.31) (finite asymptotic flow velocity) restricts solution to the right of the dash-dotted line. Equation (3.32) (subsonic, accelerating flow at r_0) restricts solutions to the leff of the dashed line. So our allowed region is that labeled "II". Region I is that of steady hydrostatic coronae of finite radial extent, discussed in §3.2. In region III no steady wind-type solution is possible.

3.3.5 Energetics {SSEbal}

As discussed in Appendix **??**, buried deep in the polytropic approximation (i.e., eq. [3.11] with α a constant specified *a priori*) is a very specific energy source/sink functional form. We now have pretty good solution, in terms of its asymptotic behavior, etc., but must now ask ourselves whether or not this solution involves distributions of energy sources/sinks that are even mildly reasonable.

Now our solutions, in general, will not satisfy the energy equation (which we didn't solve for anyway, having effectively *replaced* it by the polytropic approximation). But we can turn the issue around and use our solution to determine what sources/sinks *should* appear on the RHS of the energy equation in order for our solution to satisfy it. Neglecting thermal conduction and limiting ourselved to steady-state ($\partial/\partial t = 0$) systems, the energy equation

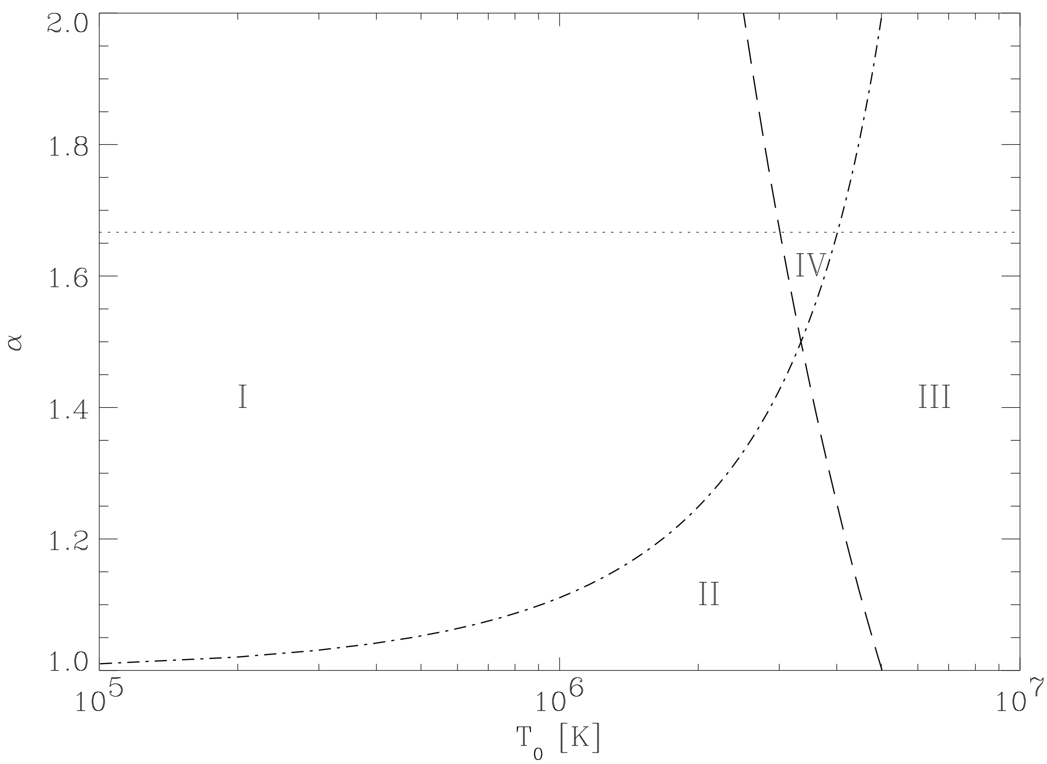

Figure 3.6: {`fig:aTplane`} Allowed region in the $[T_0, \alpha]$ plane for the existence of steady, transsonic polytropic wind solutions. The region $\alpha > 5/3$ is thermodynamically excluded. Region II is the allowed region, as defined by eq. (3.33) (see text).

can be manipulated into the form

$$\nabla \cdot \left[\rho \mathbf{u} \left(\frac{1}{2}\|\mathbf{u}\|^2 + \frac{3}{2}\frac{p}{\rho}\right)\right] + \nabla \cdot (p\mathbf{u}) - \rho \mathbf{u} \cdot \nabla V = s(r), \qquad (3.34)$$ {`E33.eb2`}

where $s(r)$ is our extraneous source/sink term (which has units of erg s^{-1} cm^{-3}), artificially added on the RHS. Direct substitution of our polytropic solutions on the LHS of this expressions allows to calculate directly the functional form of the heating term $s(r)$ so that the energy equation is now satisfied by construction. Figure 3.3.5 shows the resulting $s(r)$, for a sequence of solution having $T_0 = 1.5 \times 10^6\,\mathrm{K}$ and different values of α. The *total* energy input associated with our source is

$$S(\alpha, T_0) = 4\pi \int_{r_0}^{\infty} s(\alpha, T_0; r) r^2 \mathrm{d}r. \qquad [\mathrm{erg\ s}^{-1}] \qquad (3.35)$$ {`E33.eb3`}

Carrying out this integral yields $S = 1.1 \times 10^{28}$, 2.3×10^{27}, and 8.7×10^{26} erg s^{-1} for $\alpha = 1.05$, 1.1, and 1.15 respectively; in all cases, this is less than 10^{-5}

{F3.Eb}

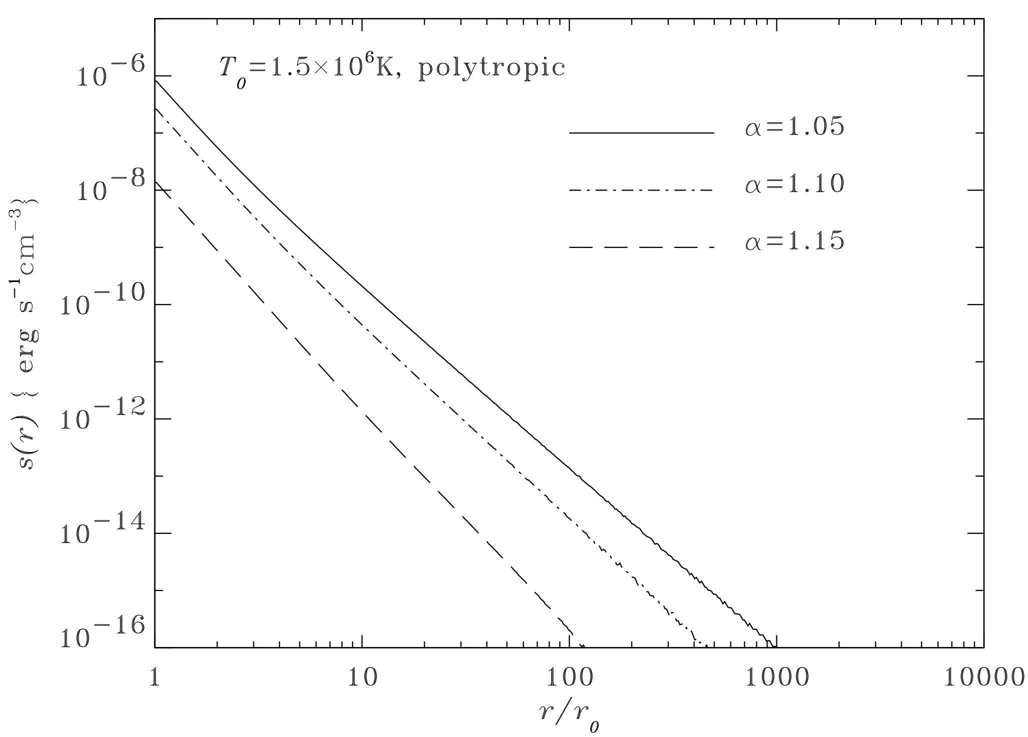

Figure 3.7: Energy input implicit in the polytropic wind solutions, for a few values of α. The base temperature is $T_0 = 1.5 \times 10^6$ K in all cases.

of the solar luminosity, a fortunate state of affairs. Likewise, it is reassuring that the heating term peaks at the coronal base and decreases rapidlyt outwards, since the heating ultimately originates near the solar surface.

3.3.6 Comparison with the Solar Wind

{SSRealsw}

Time to compare our polytropic solutions to the real solar wind. Flow properties at 1 AU for the solution of Fig. 3.5 are listed in Table 3.2 below. Compare this to Table 3.1, in particular to the flow properties of low speed streams. Pretty amazing; our model values are within the observed fluctuations for the flow speed, *and* particle number density. We are off by a whopping factor of 10 on the temperature (the proton temperature should be the meaningful one to compare to in the context of our single-fluid model), but the fact that the observed temperatures for protons, electrons and Helium nucleii differ by large factors (cf. Table 2.1) is telling us (very loudly) something about the breakdown of our single fluid approximation.

Table 3.2
Parker's solar wind solution

r	u_r [km s^{-1}]	N [cm^{-3}]	T [K]
r_0	2.1	10^8	1.5×10^6
r_s	113	4×10^4	6.8×10^5
$r_\oplus$	315	16	3.1×10^5
$10\, r_\oplus$	377	0.12	1.9×10^5

This leaves unexplained the higher speeds and lower densities observed in high speed streams. Within the framework of the thermally-driven models discussed here, a large increase in the asymptotic flow velocity can only be generated by increasing the base temperature. This being generally ruled out by observations, a number of authors have attempted to "speed up" the solar wind at 1 AU. One way of doing so is by introducing additional sources of energy and/or momentum at various distances from the base of the corona. This can be mediated by outward propagating acoustic and/or magnetoacoustic and/or Alfvén waves. An important and very robust result in that context is that

- Adding momentum or energy in the subsonic ($r < r_s$) region increases the overall mass flux, but not the flow speed at 1 AU.

- Adding momentum or energy in the supersonic ($r > r_s$) region increases the flow speed at 1AU, but not the overall mass flux.

Nice and fine, but how do we achieve that? Guess what, magnetic fields can do the trick, both in indirect and direct ways. This is the focus of the following two chapters.

Problems:

1. Obtain eqs. (3.18)—(3.19)

2. Obtain eqs. (3.25)—(3.27)

3. Assuming that the Sun's present mass loss rate has remained constant since its arrival on the ZAMS, calculate by how much the Sun-Earth distance has varied over the past 4.5 Gyr.

4. The purpose of this problem is to get you to construct a coronal model that is more realistic, energetically speaking, than the isothermal and

polytropic models discussed in §3.2. Your starting point is the assumption that *thermal conduction* dominates the energy transport in the corona. For a static and steady-state corona, the energy equation then reduces to

$$\nabla \cdot (\chi \nabla T) = 0.$$

where χ is the coefficient of thermal conductivity. In a low density, high temperature plasma of fully ionized hydrogen, an approximate (yet fairly accurate) expression for κ is

$$\chi(T) = \chi_0 T^{5/2},$$

where $\chi_0 \simeq 8 \times 10^{-7}$ erg cm^{-1} s^{-1} K$^{-7/2}$. So your task is the following:

(a) Obtain expressions for $T(r)$, $\rho(r)$, and $p(r)$, and plot these as a function of r for a few values of T_0 in the range $10^6 \leq T_0 \leq 5 \times 10^6$ K.

(b) Obtain asymptotic ($r \to \infty$) expression for $T(r)$, $\rho(r)$ and $p(r)$, and calculate these asymptotic values for the solutions you obtained in (a)

(c) Compare and contrast your results in (b) with the corresponding results for the polytropic coronae discussed at the end of this chapter.

(d) What is the energy input (erg s^{-1}) require to maintain the corona in its assumed steady state, given the outward transport of energy? Can you think of other important coronal energy "sinks"?

5. This problem further explores possible "fixes" for our static coronal models.

   (a) Determine how fast the temperature profile would have to fall with distance for the pressure to vanish at infinity in a static corona.

   (b) What should be the coronal temperature for a static, *isothermal* corona to by dynamically balanced by the pressure in the interstellar medium?

6. Code in the pseudocode for the bisection method, as given in the text, to reconstruct (and plot) a full polytropic wind solution (i.e. $u_r(r)$, $\rho(r)$, $p(r)$ and $T(r)$). Keeping the polytropic index fixed at $\alpha = 1.1$, examine how the sonic point location, base flow speed, and wind properties at 1 AU vary with base temperature, in the range $10^6 \leq T_0 \leq 2 \times 10^6$ K. And please do provide a listing of your code.

7. This problem lets you construct an *isothermal* solar wind solution. Upon examination of the expressions we obtained for the polytropic model, one rapidly sees that simply setting $\alpha = 1$ leads to divergence, so you actually need to start from scratch;

   (a) Using the definition of the isothermal sound speed $a^2 = p/\rho$, obtain the isothermal equivalents of eqs. (3.17) and (3.21).

   (b) Obtain an expression for the location of the sonic point in terms of a and other model input quantities.

   (c) Construct a transsonic wind solution for $T_0 = 1.5 \times 10^6$ K; compare its base flow speed, sonic point location, and speed and densities at Earth's orbit with the corresponding quantities for the polytropic solution of §3.3.

   (d) Obtain an expression for the asymptotic flow speed, i.e., the isothermal equivalent of eq. (3.30). How can you explain your (presumably surprising) result?

8. Using the procedure outlined in the text, construct a numerical solution corresponding to a class III polytropic solution (i.e., subsonic for all r, cf. Fig. 3.1; use also $\alpha = 1.1$ and $T_0 = 1.5 \times 10^5$ K). Provide plots of the flow speed, density, pressure and temperature as a function of r. Examine the asymptotic ($r \to \infty$) behavior of your solution, and discuss its physical relevance.

Bibliography:

The first reasonable models for hydrostatic coronae were constructed by

Chapman, S. 1957, Smithsonian Contrib. Ap., 2, 1.

Evidence for the existence of ionized gas (static or not) in the interplanetary environment was first infered on the basis of the measured polarization of zodiacal light. The classic reference is

Behr, A., & Siedentopf, H. 1953, *Zeitschrift Ap.*, **32**, 19.

On circumstantial evidence for the present of a continuous outflow from the Sun, based on comet tail acceleration, see

Biermann, L. 1951, *Zeitschrift Ap.*, **29**, 274.

Parker's original paper on the topic is a great lesson in how to do good physics starting from simple assumptions, and should not be missed:

Parker, E.N. 1958, *Astrophys. J.*, **128**, 664

The Web site of the *Ulysses* mission is also well worth looking at:

`http://ulysses.jpl.nasa.gov`

The effect of additional energy/momentum sources in the corona on the solar wind is investigated in

Leer, E., & Holzer, T.E. 1980, JGR, 85, 4681

For a thorough, recent review of MHD waves and wave-particle interaction in the solar corona and wind, see

Marsch, E. 2006, Liv. Rev. Sol. Phys., 3, 1
`http://solarphysics.livingreviews.org/Articles/lrsp-2006-1`

The following is a relatively recent and noteworthy monograph on stellar winds:

Lamers, H.G.L.M., et Cassinelli, J.P., *Introduction to Stellar Winds*, Cambridge University Press (1999).

Chapter 4

Magnetic confinement of winds {chap:confwind}

Science is experiment; science is trying things.
It is trying each possible alternative in turn,
intelligently and systematically;
and throwing away what won't work, and accepting what will,
no matter how it goes against our prejudices.

Jacob Bronowski
A Sense of the Future (1948)

4.1 Magnetic fields in the solar corona

Up to now we have considered the dynamics of the coronal plasma independently of the coronal magnetic field, which we had earlier argued (§3.1.1) is a dominant structural agent in the corona. In this chapter we begin to flirt with the interaction of magnetic fields and plasma flow, by considering its two limiting cases, when either the plasma or magnetic field entirely dominates the force balance.

The interaction between the coronal plasma and magnetic field owes a lot of its complexity to the fact that the two are tightly coupled under typical coronal conditions. A central concept is that of *flux-freezing* (§1.8); in a highly conducting plasma in a steady-state, plasma can only flow along magnetic fieldlines, i.e., $\mathbf{u} \times \mathbf{B} = 0$. The induction equation (in the $\eta \to 0$ limit, as per the high electrical conductivity) then yields $\partial \mathbf{B}/\partial t = 0$, which of course is precisely what is required for a steady-state to be maintained.

For a corona composed of fully ionized hydrogen, we have

$$\sigma_e \simeq 2 \times 10^7 T^{3/2} \qquad [\mathrm{s}^{-1}], \tag{4.1}$$ {???}

so that

$$\eta = \frac{c^2}{4\pi\sigma_e} \simeq 3.6 \times 10^{12} T^{-3/2} \qquad [\mathrm{cm}^2\,\mathrm{s}^{-1}], \tag{4.2}$$ {???}

which for $T \sim 10^6$ K yields $\eta \sim 10^4\,\mathrm{cm}^2\,\mathrm{s}^{-1}$. Assuming $\ell \sim R$ and $\|\mathbf{u}\| \sim 1\,\mathrm{km\,s}^{-1}$, the magnetic Reynolds number is $\sim 10^{12}$. That the Reynolds number is so large here is not so much that the diffusivity is extremely small in the solar corona (it is not), but rather that the dimensions involved, of order R, are so large. Consequently, care must be exercised when applying the flux-freezing approximation in modelling structures evolving on small spatial scales.

4.2 The plasma-β {Ss341}

We must first ask under which circumstances the coronal dynamics is dominated by either the plasma or magnetic field. This is far from a trivial question. A useful quantity, called the *plasma-β*, is defined as the ratio of gas pressure to magnetic pressure:

$$\beta = \frac{8\pi p}{\|\mathbf{B}\|^2}\ , \tag{4.3}$$ {E3.pbeta}

which is basically equivalent to the ratio of thermal energy to magnetic energy. In a first approximation, when $\beta \gg 1$ the flow drags along the magnetic fieldlines, while for $\beta \ll 1$ the magnetic field either traps the plasma, or constrains it to flow along magnetic fieldlines. In which regime is the solar corona?

Computing β in the presence of a supersonic wind is complicated by the fact that the kinetic energy of the flow must also be taken into consideration, i.e., we must replace p by $p+\rho u_r^2/2$ in eq. (4.3). Figure 4.1 plots the variations with heliocentric distance of the plasma-β, for the polytropic wind solution of §3.3, and assuming either monopolar (dashed line) or dipolar (solid line) magnetic field falloffs with heliocentric distance. As argued in the preceeding chapter, the solar minimum corona lies somewhere in between these two limits.

Of course in juxtaposing in this way a radial outflow with a dipolar magnetic field, flow and field are not parallel as required by the flux-freezing constraint, but for estimating the plasma-β the procedure is justified[1]. The conclusion to be drawn from Figure 4.1 is clear: in the low corona $\beta \ll 1$, so that the magnetic field constrains plasma motions, while beyond the sonic

[1]Note however than in the case of the monopolar field, this is in fact a valid solution, which moreover is nowhere as silly as one might imagine (more on this shortly).

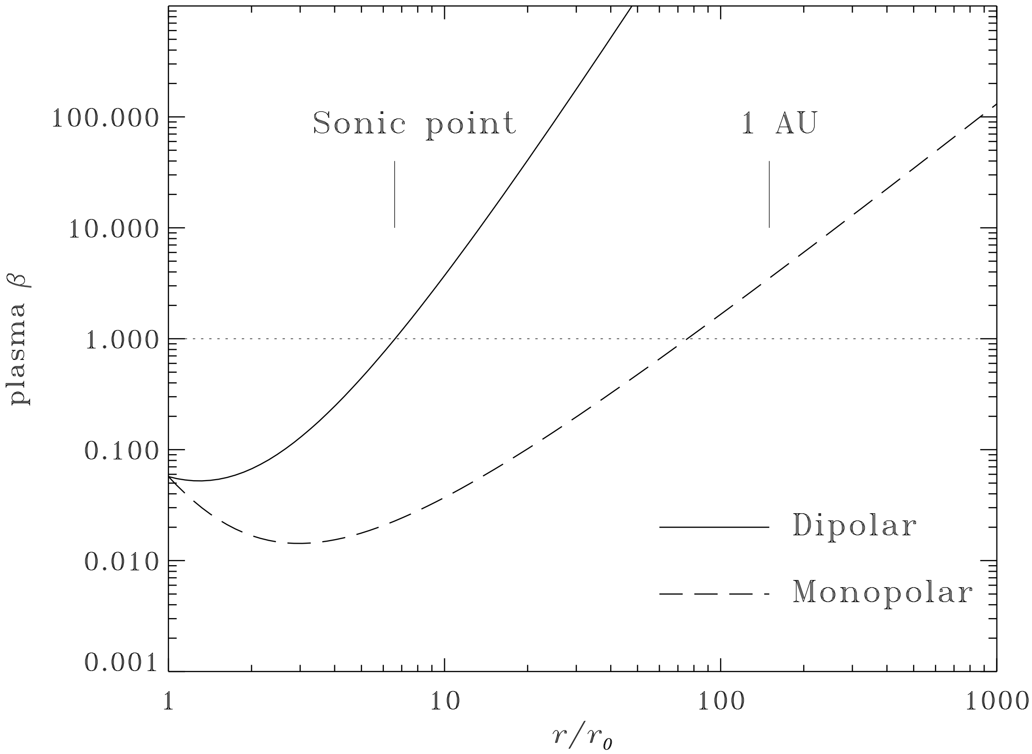

Figure 4.1: {F3pbeta} Variation of the plasma-β in the solar corona. The plasma energy is computed from the polytropic wind solution of Fig. 3.5, and the magnetic energy assuming coronal base magnetic field of 10 G, with either a monopolar ($1/r^2$) or dipolar ($1/r^3$) falloff. In either case the corona is field-dominated ($\beta \ll 1$) below the sonic point, and plasma-dominated ($\beta \gg 1$) at the Earth's orbit and beyond.

point the high-β plasma deforms the magnetic field until $\mathbf{u}$ and $\mathbf{B}$ are parallel. For a radial outflow, assuming a radial field is not so silly after all. At intermediate heliocentric distances we have $\beta \sim 1$, and the dynamics reflects the full complexity of the flow-field interaction... which is the central topic of the next chapter.

We now turn to two interesting aspects of solar wind dynamics that materialize in the extreme regimes $\beta \gg 1$ and $\beta \ll 1$.

4.3 The $\beta = 0$ case: magnetostatic solutions {sec:magnetostat

As a prelude to our study of magnetic confinement of stellar winds, we first consider a steady-state ($\partial/\partial t = 0$) situation where the dynamics is completely controlled by the magnetic field, i.e., $\beta = 0$. If this is the case, then the problem reduces to finding a force-free field under prescribed boundary conditions at the base of the corona. We'll make the task harder by requiring that these force-free solutions look "solar-like", in the sense that they are compatible with what one would infer from coronal images such as on Fig. 3.2: an axisymmetric ($\partial/\partial\phi = 0$) dipole-like corona, with open fiedlines over the polar caps, and a streamer belt straddling the equator, with closed fieldlines low down, stretched open more or less radially above a certain height above the coronal base.

This may seem like a tall order, but it turns out that someone has already done the hard work for us, that someone being Boon-Chye Low. He constructed a family of partially-open force-free axisymmetric magnetotatic solutions, by judicious insertion of force-free current sheets into otherwise potential (i.e., current-free) solutions. Working in spherical polar coordinates (r, θ, ϕ), the starting point of the model is the specification of an axisymmetric magnetic field $\mathbf{B}(r, \theta)$ in terms of an axisymmetric **stream function** $Z(r, \theta)$ via:

$$\mathbf{B}(r,\theta) = \frac{f_B B_0}{r \sin\theta} \left[\frac{1}{r} \frac{\partial Z}{\partial \theta} \hat{\mathbf{e}}_r - \frac{\partial Z}{\partial r} \hat{\mathbf{e}}_\theta \right] . \qquad (4.4) \quad \texttt{\{eq:low0\}}$$

Note that this expression will identically satisfy $\nabla \cdot \mathbf{B} = 0$, as Z can be interpreted as the z-component of a vector potential such that $\mathbf{B} = \nabla \times (Z\hat{\mathbf{e}}_\phi)$. This is really nothing fancier than the poloidal part of the toroidal/poloidal decomposition of axisymmetric magnetic fields already encountered in §1.10.3). Under this representation, Z is constant on each axisymmetric flux surface, and the value of Z can be used to label distinct such surfaces[2]. The stream function itself is constructed from two contributions:

$$Z(r, \theta; a_1, a_2) = Z_1(r, \theta; a_1) + Z_1(r, \theta; a_2) \qquad (4.5) \quad \texttt{\{eq:low1\}}$$

where a_1 and a_2 are scale parameters, and

$$Z(r, \theta; a) = r(1 - v^2) \left[(1 + u^2) \mathrm{atan} \left(\frac{1}{u} \right) - u \right] - \frac{\pi a^2}{2} \frac{\sin^2\theta}{r} + 2a\eta , \qquad (4.6) \quad \texttt{\{eq:low2\}}$$

with

$$u^2 = -\frac{1}{2} \left(1 - \frac{a^2}{r^2} \right) + \frac{1}{2} \left[\left(1 - \frac{a^2}{r^2} \right)^2 + \frac{4a^2}{r^2} \cos^2\theta \right]^{1/2} , \qquad (4.7) \quad \texttt{\{eq:low3a\}}$$

$$v^2 = -\frac{1}{2} \left(\frac{a^2}{r^2} - 1 \right) + \frac{1}{2} \left[\left(\frac{a^2}{r^2} - 1 \right)^2 + \frac{4a^2}{r^2} \cos^2\theta \right]^{1/2} , \qquad (4.8) \quad \texttt{\{eq:low3b\}}$$

[2]This means that plotting magnetic fieldline in a meridional plane amounts to plotting contours of constant Z. Very useful property for plotting purposes!

$$u^2 = -\frac{1}{2}\left(\frac{r^2}{a^2}-1\right)+\frac{1}{2}\left[\left(\frac{r^2}{a^2}-1\right)^2+\frac{4r^2}{a^2}\cos^2\theta\right]^{1/2} , \qquad (4.9)$$ {eq:low3c}

A single parameter of force-free solutions can be constructed by posing a fixed relationship between the scale parameters a_1 and a_2, e.g., $a_1 = a/2$ and $a_2 = a$. Figure 4.2 shows a sequence of four such magnetostatic solutions, for increasing values of a. These are indeed good qualitative representations of the inferred coronal magnetic at timesi of minimum activity (cf.Fig. 3.2).

Although this is a somewhat sadistic exercise in differential calculus, you should be able to verify that these solutions are current-free ($\nabla \times \mathbf{B} = 0$) everywhere except beyond a certain radius in the equatorial plane, with a measuring the radial equatorial extent of the closed region. However, $\mathbf{B} = 0$ there, so that the field configuration is indeed force-free ($\mathbf{J} \times \mathbf{B} = 0$). While this may all seem rather artificial, but we'll see in the following chapter that fully dynamically consistent MHD axisymmetric wind solutions do look a lot like this.

How about the plasma? If we further specify that we are operating in the ideal MHD limit, then any plasma "added" a posteriori to this solution will behave very differently according to whether it is added in the magnetically closed or open region of the solutions; with plasma constrained to flow along magnetic fieldlines, in a steady-state the plasma within the closed region can only remain in hydrostatic equilibrium, while in the open regions it can in principle flow out to infinity along the magnetic fieldlines (more on this very shortly).

This duality in plasma behavior forms the basis of the **minimal energy corona conjecture** put forth some years ago by Arthur J. Hundhausen. His reasoning runs as follows: In the complete absence of plasma, the magnetic field should relax to the potential state (force-free and current-free) compatible with the lower boundary conditions on $\mathbf{B}$, as guaranteed by Aly's theorem (§1.10.4). Now, if you have done problem XXX you already know that the transsonic Parker-type wind solution is a minimal energy state for all possible outflow solutions, chiefly because of the large densities characterizing the fully subsonic "solar breeze" solutions (class III solutions on Fig. 3.4, and from there it is but a small step to show that the corresponding polytropic steady corona solution (§3.2) has even higher energy. So now, back to the sequence of magnetostatic solutions depicted on Fig. 4.2. As fas as the plasma is concerned, the energy minimizing solution should have $a \to 0$, leading to radial fieldlines everywhere, and thus Parker-type spherically-symmetric outflow. But as far as $\mathbf{B}$ goes, the minimal energy state would here be a dipole, with $a \to \infty$. In other words, plasma energy is a *increasing* function of the parameter a, while magnetic energy is a *decreasing* function of a. From there it is but a small step to conclude that a partially open configuration, with

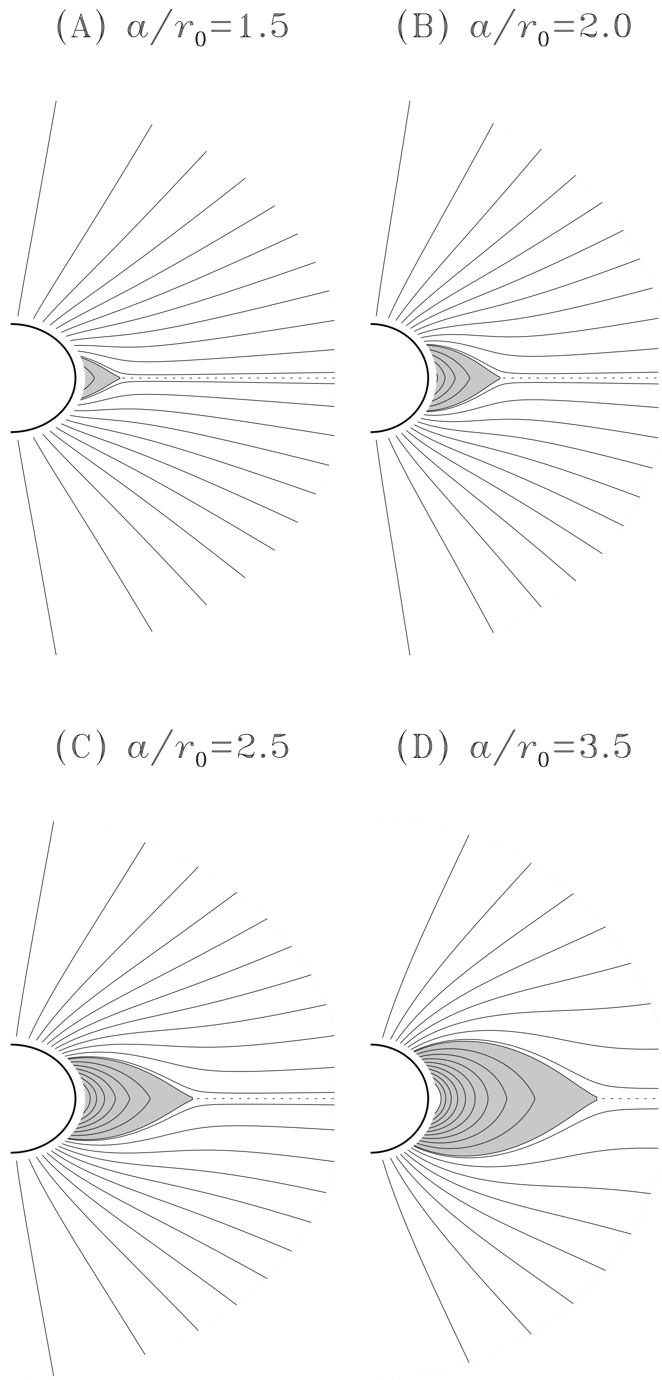

Figure 4.2: {`fig:magstat`} Four magnetostatic solution as defined by eqs. (4.4)—(4.9), for increasing values of the scale parameter a, as labeled. Shaded areas correspond to regions of the corona threaded by closed field-lines, i.e., with both footpoint anchored at the coronal base. The dotted line indicates the location and extent of the equatorial current sheet associated with these solutions.

equatorial current sheet and all, represents the configuration that minimizes the *total* energy, plasma plus magnetic. This is the essence of Hundhausen's minimal energy corona conjecture.

Of course a minimal energy state is what one would expect from any closed physical system left to himself long enough to relax, but the the solar corona is anything but closed, and a number of mechanisms force it on a variety of timescales. Yet the opening and closing of magnetic arcades that accompany coronal mass ejection can indeed be interpreted as a forced destabilization opening the arcade and releasing excess plasma trapped therein, with the subsequent closing corresponding to the return to the (quasi-)steady minimal energy state. Those interested in further exploring these intringuing ideas will find pointers to the relevant literature in the bibliography at the end of this chapter.

4.4 The $\beta \ll 1$ limit: magnetic flow tubes {Ss343}

We already alluded to the idea that plasma could flow as a wind directed along magnetic fieldlines; the purpose of this section is to examine this channelling process in somewhat greater quantitative detail. Figure 4.4 shows fieldlines corresponding to the $a = XXX$ magnetostatic solution discussed in §4.3, plotted in a single meridional quadrant. With plasma constrained to flow along magnetic fieldlines and base conditions (temperature, density, etc.) independent of latitude, the first thing to note is that in the region of the corona threaded by fieldlines that have both footpoints anchored on the boundary, we *must* have $\mathbf{u} = 0$ to satisfy mass conservation. Indeed this has nothing to do with the field being force-free or not, it is a direct consequence of flux-freezing in a steady-state. If the magnetic field is force-free in the closed region, the force balance therein then reverts to the simple statement of hydrostatic balance encountered earlier in §3.2; if not, eq. (3.14) must be modified to include the Lorentz force, but the solution remains static.

The situation is quite different in the open region, threaded by fieldlines that extend to infinity. A wind outflow along fieldlines is now possible, but the outflow is no longer radial. Consider a narrow *flow tube* defined by two adjacent fieldlines in the open region of the magnetostatic solution (thick lines on Figure 4.4). Define a coordinate s measuring distance along the line $\mathbf{s}$ oriented along the central axis of the flow tube (dashed line on Figure 4.4). The cross-section $A(s)$ along the coordinate line is readily constructed from the known form of the magnetostatic solution.

The s-component of the equations of motion is then

$$u_s \frac{\partial u_s}{\partial s} = -\frac{1}{\rho}\frac{\partial p}{\partial s} - \frac{GM}{r^2}(\hat{\mathbf{e}}_s \cdot \hat{\mathbf{e}}_r) \qquad (4.10)$$ {EFt.1}

{F3.ftube}

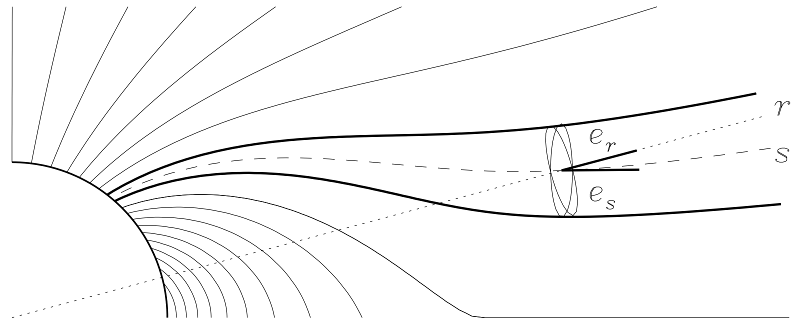

Figure 4.3: Outflow confined by a magnetic flux tube. The coordinate s (dotted line) is oriented along the central axis of the flow tube, the boundaries of which are indicated by thicker lines.

where $\hat{\mathbf{e}}_s$ is a unit vector along the coordinate line $\mathbf{s}$. Since $\hat{\mathbf{e}}_s$ is everywhere perpendicular to the tube cross section, we have $A(r) = A(s)(\hat{\mathbf{e}}_r \cdot \hat{\mathbf{e}}_s)$ and $\partial s/\partial r \equiv \hat{\mathbf{e}}_s \cdot \hat{\mathbf{e}}_r$, eq. (4.10) becomes

$$u_s \frac{\partial u_s}{\partial r} = -\frac{1}{\rho}\frac{\partial p}{\partial r} - \frac{GM}{r^2}, \qquad (4.11)$$ {EFt.2}

which means that u_s obeys an equation strictly equivalent to the r-component of the momentum equation considered in §3.3.1; a very remarkable result indeed! The only difference with the spherically symmetric solution obtained earlier is that the mass conservation statement now takes the form

$$\frac{\partial}{\partial r}\left[\rho u_r A(r)\right] = 0, \qquad (4.12)$$ {???}

where in general $A(r) \neq 1/r^2$.

4.5 Generalized polytropic wind solutions {sec:genwind}

The fact that in the low-β regime magnetic fields will "rigidly" channel wind-type outflows means that the wind acceleration is akin to a nozzle flow, with magnetic flux surfaces playing the role of the nozzle's rigid boundaries, with the area expansion factor $A(s)$ acting as the nozzle's cross-section. Now, in a polytropic flow, we already saw that the Bernoulli constant, corresponding to the total energy per unit mass, is given by:

$$E = \frac{u_r}{2} + \frac{c_s^2}{\alpha - 1} - \frac{GM}{r} \ . \qquad (4.13)$$ {eq:nozzle1}

You'll recall (hopefully) that the three terms on the RHS are, from left ro right: the flow's kinetic energy, the plasma's thermal energy, and gravitational potential energy, all per unit mass. The most any nozzle can do, starting from a fluid at rest in the "combustion chamber" (here the coronal base) is to convert all of the plasma's original thermal energy into outflow kinetic energy; here this limiting velocity is given by something like:

$$u_\infty = \frac{2c_{s0}^2}{\alpha - 1} - \frac{GM}{r_0} \ . \qquad (4.14) \quad \{\texttt{eq:nozzle2}\}$$

where as before c_{s0} is the sound speed at r_0. But how does this work out in practice? Working once again through the mathematical steps we encountered in the case of Parker's spherically symmetric polytropic wind solution, it can be shown that for an arbitrary expansion factor $A(r)$, the r-momentum equation can be written in the following general form:

$$\frac{M^2-1}{2M^2}\frac{\mathrm{d}M^2}{\mathrm{d}r} = \left[1+\left(\frac{\alpha-1}{2}\right)\right]\left[\frac{1}{A}\frac{\mathrm{d}A}{\mathrm{d}r} - \frac{1}{2}\left(\frac{\alpha+1}{\alpha-1}\right)\frac{GM/r^2}{(E+GM/r)}\right]$$

$$= \frac{1}{2}\left(\frac{\alpha+1}{\alpha-1}\right)\left[1+\left(\frac{\alpha-1}{2}\right)\right]\frac{1}{g}\frac{\mathrm{d}g}{\mathrm{d}r} \ , \qquad (4.15) \quad \{\texttt{eq:k-h1}\}$$

where E is the Bernoulli constant of eq. (4.13), M is the Mach number

$$M(r) = \frac{u_r}{c_s} \ , \qquad (4.16) \quad \{\texttt{eq:k-h2}\}$$

and the function g is given by

$$g(r) = A^{2(\alpha-1)/(\alpha+1)}\left(E + \frac{GM}{r}\right) \ . \qquad (4.17) \quad \{\texttt{eq:k-h3}\}$$

The mathematics are more complex, but this is really the same general idea as with the spherically-symmetric Parker polytropic wind solution of §3.3. In particular, solutions to eq. (4.15) include sonic critical points that must be crossed by wind solution in order to avoid infinite accelerations. What is novel here is that for expansion factors with fast divergence, more than one critical points can exist in the flow. You get to explore this aspect of the problem in Problem XXX below.

An integral form of eq. (4.15) can also be obtained:

$$M^{4/(\alpha+1)} + \left(\frac{2}{\alpha-1}\right)M^{-2(\alpha-1)/(\alpha+1)} =$$

$$\frac{g(r)}{g_0}\left[M_0^{4/(\alpha+1)} + \left(\frac{2}{\alpha-1}\right)\right]M_0^{-2(\alpha-1)/(\alpha+1)} \qquad (4.18) \quad \{\texttt{eq:k-h4}\}$$

with $g_0 \equiv g(r_0)$ and $M_0 \equiv M(r_0)$. This is really nothing more than the Bernoulli equation (4.13) written in terms of the Mach number. This form is useful for reconstructing full wind solution, since it amounts to yet another root-finding problem for M (at fixed r).

4.6 The $\beta \gg 1$ limit: The Parker spiral {ssec:spiral}

Consider now the other opposite, extreme limit of $\beta \gg 1$, in which the magnetic field is passively advected by the wind outflow. More specifically, assume a steady-state situation where

1. Flux-freezing is effectively enforced,
2. Magnetic stresses are neglected in the force balance,
3. The poloidal part of the magnetic field is purely radial in the equatorial plane, with the field strength known at the reference radius.

In view of (3), the constraint $\nabla \cdot \mathbf{B} = 0$ is readily integrated to

$$B_r(r) = B_{r0}\left(\frac{r_0}{r}\right)^2 . \qquad (4.19) \quad \{E34.1\}$$

For an average surface field $B_0 \sim 10\,\mathrm{G}$, eq. (4.19) yields $B_r \simeq 2.5 \times 10^{-6}\,\mathrm{G}$ at the Earth's orbit, which is not that far from the observed average magnetic field at 1AU. In view of (1), the flow streamlines coincide with magnetic fieldlines. In the absence of rotation, the Parker solution is immediately applicable. Consider now the introduction of rotation, at a rate Ω such that centrifugal effects do not affect significantly the force balance in the r-direction. In a frame co-rotating with the Sun, the wind still flows along the magnetic fieldlines. But in a stationary frame, the total velocity is now

$$\mathbf{u} = \mathbf{u}' + \Omega\, r\, \hat{\mathbf{e}}_\phi, \qquad (4.20) \quad \{E34.2\}$$

where primed quantities refer to quantities evaluated in the co-rotating frame. In general, the magnetic fieldlines are thus defined by the spiral

$$r = (u_r/\Omega_\odot)(\phi - \phi_0)\ , \qquad (4.21) \quad \{eq:spiral\}$$

with the r and ϕ-components of the magnetic field given by

$$B_r(r,\theta,\phi) = B_r(r_0,\theta,\phi - r\Omega_\odot/u_r)\left(\frac{r_0}{r}\right)^2 , \qquad (4.22) \quad \{E34.3a\}$$

$$B_\phi(r,\theta,\phi) = B_r(r_0,\theta,\phi - r\Omega_\odot/u_r)\left(\frac{r_0\Omega_\odot}{u_r}\right)\left(\frac{r_0}{r}\right) . \qquad (4.23) \quad \{E34.3c\}$$

Figure 4.4 shows, in the equatorial plane, the magnetic fieldlines defined by eqs. (4.22)–(4.23), with the dashed circle corresponding to the Earth's orbit[3]. The angle between a magnetic fieldline and the Sun-Earth radial segment is:

$$\phi_B = \arctan\left(\frac{B_\phi}{B_r}\right) = \arctan\left(\frac{r\Omega_\odot}{u_r}\right) , \qquad (4.24) \quad \{E34.4\}$$

[3] On Fig. 4.4, is the Sun rotating clockwise or counterclockwise?

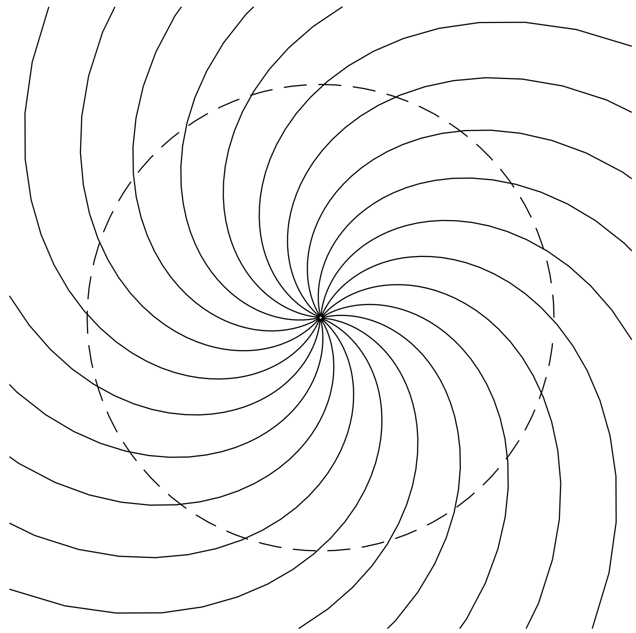

Figure 4.4: {`F34.1`} The spiral drawn by the solar magnetic field, as it is advected outward by the solar wind. Solid lines corresponds to magnetic fieldlines, and the circular dashed line to the Earth's orbit. The wind itself flows radially outward from the Sun (located at the center of the spiral, of course).

which at 1 AU gives the rather large value $\phi_B \simeq 55^o$, which in fact compares favorably with observations. The net wind velocity at 1 AU, on the other hand, is essentially radial[4].

Now, remember the equatorial current sheet that characterized the partially open magnetostatic solutions considered in §4.3? Well this has been detected also through situ solar wind observations at 1AU. One of the most intriguing aspect of early space-borne solar wind measurements was the semi-regular polarity flips of the magnetic field carried by the wind. It was soon realized that this could be traced to the fact that the "neutral line" $B_r = 0$ at the solar surface does not coincide exactly with the equatorial circle, but is often deformed into a wavy line crossing back and forth across hemispheres. This warping is maintained in the corona and solar wind, so that observations made in the ecliptic see alternately above and below the equatorial current sheet leading to apparent polarity flip as solar rotation carries this pattern

[4]Confusion on the horizon. Didn't we argue that in the flux-freezing limit, the gas could only flow parallel to the fieldlines? Shouldn't we then have $\arctan(v_\phi/v_r) \simeq 55^o$ also? How do you explain this?

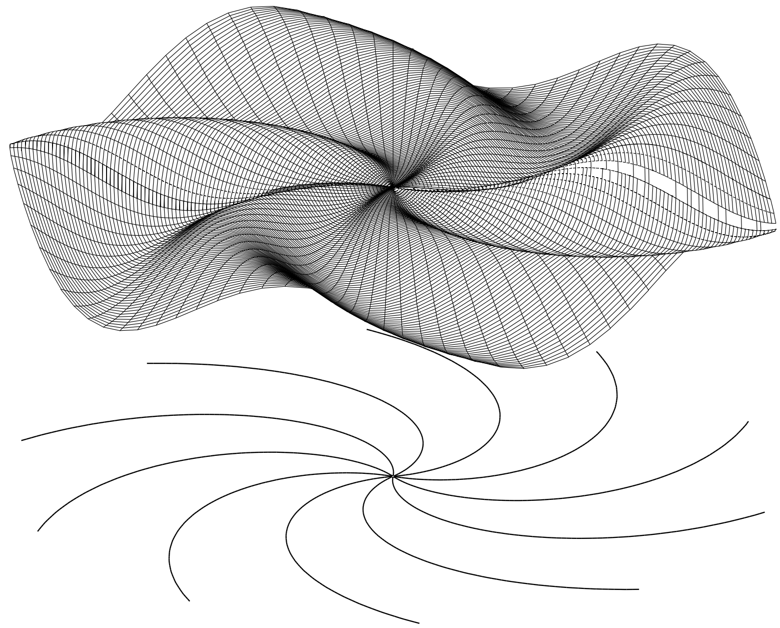

Figure 4.5: {`fig:warp`} Warped spiral caused by the radial dragging of the solar magnetic field by the solar wind emanating from the rotating Sun. Any break of equatorial symmetry is imprinted on the expanding magnetic field, leading to a warping of the equatorial current sheet beyond the point where the magnetic arcades are is opened up by the wind. For an observed in the equatorial plane, this leads to apparent polarity flips of the magnetic field measured in the solar wind, as one alternately "looks" above and below the current sheet.

along. The basic shape of this warped spiral is illustrated in cartoon form on Figure 4.5. The wavy equatorial current sheet has even been compared to the tutu of a not-so-bashful ballerina!

Problems:

1. Work out the nozzle cross-section variation with r that would produce the same flow acceleration as in the solar wind.

2. Obtain eq. (4.14)

3. Verify that the introduction of a spherical expansion factor $A = r^2$ in eqs. (4.15)—(4.15) brings you back to eq. (3.17).

4. The aim of this (and the following) problem is to construct a polytropic wind solution for non-radial expansion factors. Consider (and, while

you're at it, plot) the following function

$$f(r; R_e) = \left(\frac{1}{r^2} + \frac{R_e}{r^3}\right)^{-1}$$

for $R = 1, 2$ and 10. Clearly we have $\lim_{r\to 0} f(r) \propto r^3$, and $\lim_{r\to\infty} f(r) \propto r^2$, with the "turning point" occurring at $r \sim R_e$. Qualitatively, this is the kind of behavior we would get from constructing expansion factors in the polar regions of the configurations shown on Fig. 4.2. Using this expansion factor, construct a few polytropic solutions with $T_0 = 1.5 \times 10^6$ K (e.g., consider $\alpha = 1.05$, 1.1, and 1.15) and compare the resulting flow speed and densities at 1 AU with the solar wind data for high speed streams (cf. Table 3.1). Can you generate high speed stream in this way?

5. In the lower corona, the expansion factor associated with coronal hole scales much faster than r^2 at the base of the corona. This can be modeled with the following expansion factor:

$$\frac{A(r)}{A(r_0)} = \left(\frac{r}{r_0}\right)^2 f(r)$$

with the parametric function

$$f(r) = \frac{f_m \exp((r - r_0)/\sigma) + f_1}{\exp((r - r_0)/\sigma) + 1}$$

where

$$f_1 = 1 - (f_m - 1)\exp((r_0 - r_1)\sigma) \ .$$

The quantity r_1 determines where the expansion is most pronounced, the width parameter σ controls the width of the interval in r where rapid expansion taeks place, and f_m is the asymptotic expansion factor. In what follow you may set $r_1/r_0 = 1.5$ and $\sigma/r_0 = 0.1$. Use a polytropic index $\alpha = 1.1$ and a base temperature $T_0 = 2 \times 10^6$ K, somewhat higher than the value used in the preceding chapter but in fact more appropriate for coronal holes. As usual, assume a perfect gas composed of fully ionized Hydrogen.

(a) Set the Bernoulli constant at $E =$, and calculate the base sound speed c_{s0}, the base flow speed u_{r0} and the base Mach number M_0.

(b) Equation (4.18) being a Bernoulli-type equation, a quantity E^* defined as $E^* = \text{LHS} - \text{RHS}$ is directly related to the Bernoulli constant, and therefore is a constant of the flow; using your value of M_0 calculated in (a), plot, in the $[r, M]$ plane, contours of constant $E^*(r, M)$, first for a solution with purely radial expansion factor. Verify that the resulting "topology" is the same as that depicted on Fig. (3.4).

(c) Now construct outflow topologies for two rapidly diverging geometries, namely $f_m = 3$ (slow divergence) and $f_m = 12$ (fast divergence).

(d) How many critical point do you have in the "topologies" uncovered in (c)? Does the wind solution cross all of them?

If you need inspiration, consult the Kopp & Holzer paper listed in the bibliography below.

Bibliography:

The magnetostatic solutions described in §4.3 are taken from

Low, B.C. 1986, *Astrophys. J.*, **310**, 953-965,

but there is a huge literature on analytic magnetostatic coronal solutions, with or without plasma contribution to the force balance. The mathematicaly courageous wishing to look into the matter should start with

Tsinganos, K., & Low, B.C. 1989, *Astrophys. J.*, **342**, 1028.

On the minimal energy corona conjecture, see

Charbonneau, P., & Hundhausen, A.J. 1996, *Sol. Phys.*, **165**, 237-256,

and references therein. The behavior of polytropic wind solutions with rapidly diverging expansion factors is discussed in detail by

Kopp, R.A., & Holzer, T.A. 1976, *Sol. Phys.*, **49**, 43-56,

from which §4.5 is largely inspired.

Chapter 5

Magnetic driving of winds {chap:drivwind}

We now seek to obtain wind solutions incorporating in a dynamically consistent manner the dynamical interaction between flows and magnetic fields. To do so in a mathematically tractable manner, we first compromise at the level of geometrical realism in §5.1, to obtain the justly famous Weber-Davis wind solution. Despites its geometrical restrictions, this will prove very useful in looking at the interesting problem of stellar angular momentum loss and spin-down (§5.3). We then close the chapter —and this part of the course— by looking at the possible contribution of Alfvén waves in accelerating stellar winds beyond what purely thermal driving can achieve (§5.4).

5.1 The Weber-Davis MHD wind solution {sec:WDsol}

The general geometrical setup is the same as that used in obtaining non-rotating, unmagnetized polytropic wind solutions in §3.3. We consider steady ($\partial/\partial t = 0$) spherically symmetric ($\partial/\partial\theta = \partial/\partial\phi = 0$) outflow from a star rotating at angular velocity Ω and characterized by a known surface radial component of the magnetic field B_{r0}. As before we consider the coronal base temperature $T(r_0) \equiv T_0$ as known, and do away with the energy equation by assuming a polytropic relationship between pressure and density. But here it gets a tad fishy; we will seek ouflow solutions restricted to the equatorial plane, where we impose $\mathbf{B}_\theta = 0$. This may smell of monopolar magnetic fields, but this is really what we also did before when constructing the Parker spiral in §4.6, and the discussion of §4.3, (see in particular Fig. 4.2) indicates that a solar radius or so above the photosphere, this is a fair representation of the interplanetary magnetic field during solar minimum conditions.

A bit of reflection should convince you that we now need five input quantities need to define a Weber-Davis wind model (as opposed to three for the Parker wind solution of §3.3.1):

1. the polytropic index α;
2. something measuring coronal temperature, for which we'll use the base sound speed $c_{s0} = \sqrt{\alpha p_0/\rho_0}$ at the reference radius r_0;
3. something measuring the strength of gravity; it will prove convenient to use the dimensionless ratio (γ) of the gravitational escape speed u_G ($= \sqrt{2GM/r_0}$) to the base sound speed;
4. something measuring the rotation rate, for which we can use the dimensionless parameter $\zeta = \Omega r_0/c_{s0}$;
5. something measuring the strength of the radial magnetic ield component at r_0, for which we can use the dimensionless parameter $\beta = A_{r0}/c_{s0}$;

We group these into a solution input vector:

$$\mathbf{z} = (\alpha, c_{s0}, \gamma, \zeta, \beta) \ . \tag{5.1}$$ {eq:WD20}

Let's get going. As usual, the symmetry properties imposed *a priori* on our wind solution let to significant simplification of the governing fluid equations. Mass continuity remains what is was for the Parker wind solution of §3.3):

$$\frac{1}{r^2}\frac{\partial}{\partial r}(r^2 \rho u_r) = 0 \ , \tag{5.2}$$ {eq:WD1}

while the r and ϕ-components of the momentum equation become

$$\rho\left(u_r\frac{\partial u_r}{\partial r} - \frac{u_\phi^2}{r}\right) = -\rho\frac{\partial \Phi}{\partial r} - \frac{\partial p}{\partial r} - \frac{B_\phi}{4\pi r}\frac{\partial}{\partial r}(rB_\phi) \ , \tag{5.3}$$ {eq:WD2a}

$$\rho\left(u_r\frac{\partial u_\phi}{\partial r} - \frac{u_r u_\phi}{r}\right) = \frac{B_r}{r}\frac{\partial}{\partial r}(rB_\phi) \ . \tag{5.4}$$ {eq:WD2b}

The ϕ-component of the induction equation reduces to

$$\frac{1}{r}\frac{\partial}{\partial r}(r u_r B_\phi - r u_\phi B_r) = 0 \ , \tag{5.5}$$ {eq:WD3}

while the equation for the r-component is trivially satisfied (i.e., $0 = 0$!). An equation for B_r is obtained instead via the magnetic flux conservation constraint $\nabla \cdot \mathbf{B} = 0$, which here reduces to:

$$\frac{1}{r^2}\frac{\partial}{\partial r}(r^2 B_r) = 0 \ . \tag{5.6}$$ {eq:WD4}

Equations (5.2), (5.6), and (5.5) integrate directly to

$$r^2 \rho u_r = C_1 \ , \tag{5.7}$$ {eq:WD5a}

$$r^2 B_r = C_2 \ , \tag{5.8}$$ {eq:WD5b}

$$r(u_r B_\phi - u_\phi B_r) = C_3 \ , \tag{5.9}$$ {eq:WD5c}

where C_1, C_2 and C_3 are integration constants. The first two correspond respectively to the mass and magnetic flux associate with the wind. To evaluate C_3 we transform to a reference frame co-rotating with the Sun:

$$u_\phi \rightarrow u'_\phi + \Omega r \ , \tag{5.10}$$ {eq:W6}

where the prime indicates evaluation in the co-rotating frame. Note that this (non-relativistic) transformation leaves the radial components of **u** and **B** unaffected. In that frame **B** is stationary, and since we are working under the flux-freezing approximation **u** and **B** must be parallel:

$$\frac{u'_r}{u'_\phi} = \frac{B'_r}{B'_\phi} \ . \tag{5.11}$$ {eq:WD6}

Since $B_r = B'_r$, eq. (5.9) yields

$$C_3 = -\Omega r^2 B_r \ , \tag{5.12}$$ {eq:WD7}

so that

$$B_\phi = \frac{B_r}{u_r}(u_\phi - \Omega r) \ . \tag{5.13}$$ {eq:WD8}

Now, eq. (5.4) can obviously be rewritten as

$$\frac{\partial}{\partial r}(r u_\phi) = \frac{B_r}{4\pi\rho u_r}\frac{\partial}{\partial r}(r B_\phi) \ ; \tag{5.14}$$ {eq:WD9}

but in view of eqs. (5.7) and (5.8), we have $B_r/4\pi\rho u_r = C_2/4\pi C_1$, i.e., a constant! Which means that eq. (5.14) integrates immediately to

$$r u_\phi - -\frac{r B_\phi B_r}{4\pi\rho u_r} = L \ , \tag{5.15}$$ {eq:WD10}

where L is yet another integration constant. It has a well-defined physical meaning, as it corresponds to the total angular momentum carried away by the wind, which is made up of two contributions: the specific angular momentum of the expanding fluid (first term on LHS), and the torque density associated with magnetic tension (remember that the magnetic field is being dragged away by the wind outflow!)

The results of all this algebraic juggling, without giving us a full wind solution, still allow us to draw some interesting conclusions regarding the behavior of the outflow. First we rewrite eqs. (5.13) and eqs. (5.15) in terms of the components of the **Alfvén velocity**[1]

$$A_r = \frac{B_r}{\sqrt{4\pi\rho}} \ , \qquad A_\phi = \frac{B_\phi}{\sqrt{4\pi\rho}} \ , \tag{5.16}$$ {eq:WD11}

leading to

$$A_\phi = \frac{A_r}{u_r}(u_\phi - \Omega r) \ , \tag{5.17}$$ {eq:WD12a}

$$u_\phi = \frac{L}{r} + \frac{A_r A_\phi}{u_r} \ . \tag{5.18}$$ {eq:WD12b}

Substituting now for A_ϕ in eq. (5.18) and making good use of eqs. (5.16) and eqs. (5.17) yield, after some straightforward algebra:

$$u_\phi = \Omega r \frac{(u_r^2 L/\Omega r^2) - A_r^2}{u_r^2 - A_r^2} \ . \tag{5.19}$$ {eq:WD13}

Look at the denominator of this expression; clearly, if the radial flow velocity ever becomes to the radial Alfvén speed, we are in trouble... unless the numerator also happens to vanish. We can save the day in this way provided we set

$$L = \Omega r_A^2 \ , \tag{5.20}$$ {eq:WD14}

where r_A is the **Alfvén radius** where $u_r = A_r$. Now, remember that L is the total angular momentum carried away by the wind, *including* the torque density provided by magnetic tension. Equation (5.20) states that this is equal to the angular momentum that would be carried away by an unmagnetized wind flowing strictly radially, and co-rotating with the solar/stellar surface out to radius r_A. We are going to get a lot of mileage from this remarkable result later on. But let's first try to get a full wind solution. Go back to the r-component of the equation of motion (eq. (5.3)); use eq. (5.13) to eliminate B_ϕ in the last term on the RHS; then use eq. (5.13) to eliminate

[1] Please do not confuse the "A" here with components of the magnetic vector potential...

the B_ϕ derivative multiplying u_ϕ (but leave the one multiplying Ω alone!). Somewhat tedious algebra eventually leads to

$$\frac{\partial}{\partial r}\left[\frac{1}{2}(u_r^2+u_\phi^2)-\frac{GM}{r^2}+\frac{c_{s0}^2}{\alpha-1}\left(\frac{\rho}{\rho_0}\right)^{\alpha-1}-\frac{r\Omega A_r A_\phi}{u_r}\right]=0. \qquad (5.21)$$ {eq:WD15}

where the magnetic field components are again expressed in terms of their corresponding Alfvén speed components, and the polytropic approximation was used to deal with the pressure gradient term. This indicates that the quantity within the square brackets must be a constant[2]. This is again a Bernoulli-type statement for the flow, expressing conservation of energy, and as before we will denote the quantity in square brackets by E.

Obtaining a full solution (i.e., u_r, $u_\phi(r)$, etc.) is now a much more complicated procedure. The starting point is the manipulation of eq. (5.3) into the form:

$$\frac{\partial u_r}{\partial r}=\left(\frac{u_r}{r}\right)\frac{(u_r^2-A_r^2)(2c_s^2+u_\phi^2-GM/r)+2u_r u_\phi A_r A_\phi}{(u_r^2-A_r^2)(u_r-c_s^2)-u_r^2A_\phi^2)} \qquad (5.22)$$ {eq:WD16}

which involves some straightforward but tedious algebraic juggling. Now, that denominator looks like trouble once again. It actually vanishes whenever the radial flow speed u_r becomes equal to the phase speed of either the slow or fast magnetosonic wave mode[3], which in general occurs at distinct radial distances denoted r_s and r_f in what follows. Denote now by N and D the numerator and denominator on the RHS of eq. (5.22); to avoid divergence at r_s or r_f we require that

$$N(r_f,u_f)=0\ , \qquad (5.23)$$ {eq:WD17}

$$D(r_f,u_f)=0\ , \qquad (5.24)$$ {???}

$$N(r_s,u_s)=0\ , \qquad (5.25)$$ {???}

$$D(r_s,u_s)=0\ , \qquad (5.26)$$ {???}

complemented by the requirement that solutions running through the two critical points should also be characterized by the same value of the Bernoulli

[2]If eq. (5.21) doesn't look at least a bit familiar, go back and read chapter 3, before proceeding, because you're already in trouble enough.

[3]These correspond to sound-like longitudinal waves for which the sum of gas and magnetic pressures act as a restoring force; if both are in (out of) phase, the magnetosonic wave is fast (slow).

constant[4]:

$$E(r_f, u_f) = E(r_0, u_{r0}) \ , \qquad (5.27) \quad \{\texttt{eq:WD18}\}$$

$$E(r_s, u_s) = E(r_0, u_{r0}) \ . \qquad (5.28) \quad \{\texttt{???}\}$$

These expressions represent a set of six coupled nonlinear algebraic equations that must be solved simultaneously for a "solution vector"

$$\mathbf{w} = (u_{r0}, u_{\phi 0}, r_s, u_s, r_f, u_f) \ . \qquad (5.29) \quad \{\texttt{eq:WD19}\}$$

Well, we can find reassurance in the fact that we have as many equations as we have unknowns, but the fact remains that solving this nonlinear algebraic system is A BEAR of a root finding problem. It can be turned into a (somewhat easier) optimization problem, by seeking solutions that minimize the sum of the squared N's, D's and E's, but even then you better have a pretty good initial guess for the solution vector to start a conjugate gradient or whatever, because the 6-dimensional search space is very multimodal. But it can be done; and if you do it for the "solar" input vector

$$\mathbf{z}_\odot = (1.1, 165., 0.01415, 3.495, 3.688) \ , \qquad (5.30) \quad \{\texttt{eq:zsol}\}$$

you find a "solar" solution vector

$$\mathbf{w}_\odot = (0.0123, 0.0140, 6.60, 0.676, 29.5, 1.378) \ , \qquad (5.31) \quad \{\texttt{eq:wsol}\}$$

where the flow speed are expressed in fractions of the base sound speed c_{s0}, and the critical point radii in units of the reference radius r_0. Reconstructing as full solution is a lengthy but straightforward process which involves the following sequential steps:

1. Construct $A_r(r, u_r)$; this is a function of B_r and ρ, B_r is only a function of r as pr eq. (5.8), and ρ of u_r and r via eq. (5.2).

2. With A_r known, construct $u_\phi(r, u_r)$ via eq. (5.19), evaluating the constant L at r_0:

$$L = r_0 u_{\phi 0} \left[1 - \left(\frac{A_{r0}}{u_{r0}} \right)^2 \right] + \Omega r_0^2 \left(\frac{A_{r0}}{u_{r0}} \right)^2 \ . \qquad (5.32) \quad \{\texttt{eq:Lsol}\}$$

[4]Hold on now, didn't we say a little while back that the wind also had to go through the Alfvén point, to avoid a blowup of the azimuthal velocity, as per eq. (5.19)? Well it turns out that in the Weber-Davis-type wind models, any solution going through the slow and fast magnetosonic points (r_s, u_s), (r_f, u_f) *automatically* goes through the Alfvén point (r_A, u_{rA}). Skeptics should either get a life, or consult Goldreich & Julian 1970, ApJ, 160, 971.

3. B_ϕ (and thus A_ϕ) can now be constructed using eq. (5.17).

This gives us all the needed pieces to express the Bernoulli constant E (cf. eq. (5.21)) in terms of r and u_r only. Setting then

$$E(r_0, u_{r0}) = E(r, u_r) \tag{5.33}$$ {eq:Lsol2}

brings us back to a one-dimensional root finding problem, which we've handled before. Once we have u_r, $\rho(r)$ follows immediately from eq. (5.2). Knowing B_r/B_{r0} from eq. (5.8), u_ϕ is evaluated using eq. (5.19), and finally A_ϕ via eq. (5.17), AND THAT'S FINALLY IT!

The resulting solar solution is plotted on Figure 5.1, with some strategic numbers listed in Table 5.3. The purely hydrodynamical components of the solution having a counterpart in the unmagnetized, non-rotating solar wind solution obtained in §3.3, i.e. u_r, ρ, and $T(r)$, look an awful lot similar to Parker's solution. The notable difference is that the flow is no longer purely radial but now has a non-vanishing ϕ-component (as measured by the flow's pitch angle $\phi_v \equiv \mathrm{atan}(u_\phi/u_r)$), quite important near r_0 but falling off rapidly with increasing radial distance. Yet the value computed at Earth's orbit is in agreement with *in situ* mesurements. That's certainly something worth celebrating.

Table 5.3
Weber-Davis solar wind solution

r	u_r [km s^{-1}]	N [cm^{-3}]	T [K]	ϕ_v [deg]	ϕ_B [deg]
r_0	2.0	10^8	1.5×10^6	48.7	-0.59
r_s	115	37400	6.8×10^5	4.01	-3.95
r_f	231	891	4.7×10^5	2.00	-15.7
$r_\oplus$	319	17	3.1×10^5	0.53	-54.5
$10\, r_\oplus$	380	0.14	1.9×10^5	0.06	-81.0

Nonetheless, after all this work, it is almost disappointing how little our solar WD wind solution differs from its non-rotating, unmagnetized counterpart of §3.3. This is due, to a large extent, to the relatively low rotation rate of the Sun, and to its relatively low magnetic field strength (refering here to the global-scale, diffuse coronal magnetic field, not that immediately overlying sunspots and active regions). But in other parameter regimes the differences become striking indeed. Consider for example the WD solution depicted on Figure 5.2; this corresponds to a Sun-like star, still with a $T_0 = 1.5 \times 10^6\, K$ corona and polytropic index $\alpha = 1.1$ as before, but now rotating at 25 times the solar angulare velocity, and with a surface field strenth also increased by a factor 25. The flow speed at large distances now

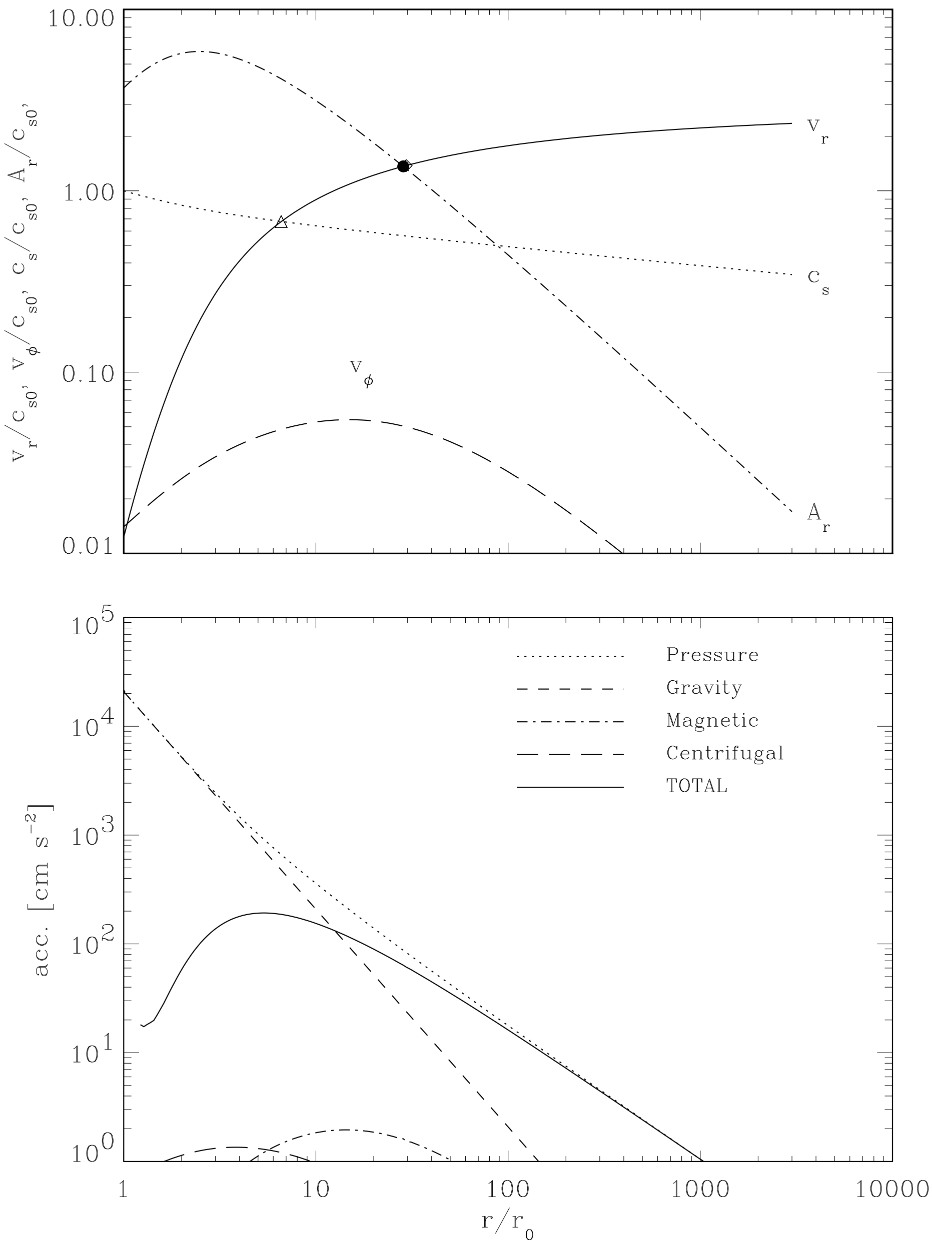

Figure 5.1: {fig:WDsol} The full solution for the WD wind model with solar-like input parameters. The top panels shows the wind properties, with the slow magnetosonic point indicated by a triangle, the Alfvén point by a solid dot, and the fast magnetosonic point by a diamond (here almost coincident with the Alfvénic critical point). The bottom panel shows details of the force balance in the wind.

exceeds the local sound speed by nearly two orders of magnitude, meaning a Mach 100 flow! The azimuthal velocity u_ϕ now exceeds the radial flow speed inside the slow magnetosonic point, and remains comparable to it out to the Alfvén point. These dramatic difference can be traced to the centrifugal and magnetic contributions to the force balance (bottom panel). While the flow remains mostly thermally-driven near the base of the corona, within a few r_0 the centrifugal and magnetic accelerations become comparable to the pressure gradient term, and completely dominate the dynamics thereafter. You may verify that the asymptotic flow speed is now given by

$$\lim_{r\to\infty} \equiv u_{r\infty} \simeq \left(\frac{\Omega^2 r_0^4 B_{r0}^2}{\dot{M}}\right)^{1/3} , \qquad (5.34) \quad \{\texttt{eq:WDstar1}\}$$

with $\dot{M} = 4\pi\rho_0 r_0^2 u_{r0}$ as in §3.3.

Figure 5.3 shows the variations with distance of the two contributions to angular momentum loss in the WD wind solutions of Figs. 5.1 and 5.2. Note how the rapidly rotating, strongly magnetized wind carries away a lot more angular momentum than in the solar solution. In view of eq. (5.15) one may have expected a factor of 25 coming from the Ω dependency, but the Alfvén point also moves outwards by a factor of nearly four (cf. Figs. 5.1 and 5.2, solid dots on top panels). In fact, it is (relatively) easy to show that in the limit of weak centrifugal and magnetic driving,

$$r_A \simeq \frac{r_0^2 B_{r0}}{\sqrt{\dot{M} u_{r\infty}}} , \qquad (5.35) \quad \{\texttt{eq:WDstar2}\}$$

as opposed to

$$r_A \simeq \sqrt{\frac{3}{2}\frac{u_{r\infty}}{\Omega}} \qquad (5.36) \quad \{\texttt{eq:WDstar3}\}$$

for the rapidly rotating, strongly magnetized wind solution of Fig. 5.2.

5.2 Numerical models of rotating MHD winds {sec:WDnum}

The Weber-Davis solution of §5.1 is applicable only in the equatorial plane; but could we not "project" it on conical surface of decreasing opening angle to reconstruct an axisymmetric solution in a full meridional $[r, \theta]$ plane? As you get to verify in Problem XXX below, this leads to an unbalanced latitudinal gradient of magnetic pressure. Moreover, the monopolar magnetic configuration of the Weber-Davis solution should cerainly be improved upon.

In obtaining fully two-dimensional Weber-Davis-like wind solutions there is no recourse but a approach that is numerical from the onset. We will now

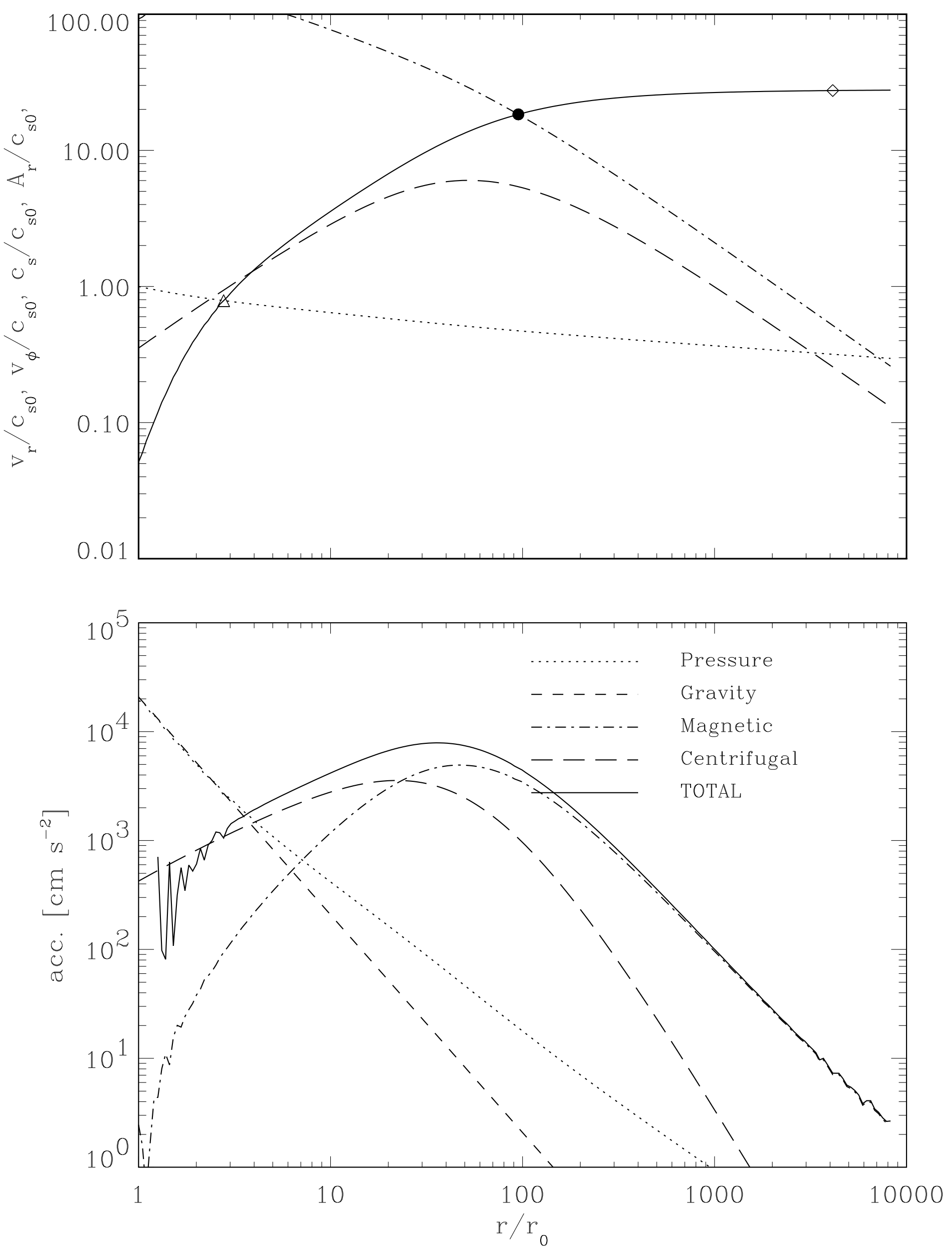

Figure 5.2: {`fig:WDstar`} Similar to Fig. 5.1, but now for a solar-like star that is rapidly rotating ($\Omega = 25\Omega_\odot$) and strongly magnetized ($B_{r0} = 25B_{r0,\odot}$), representative of a young solar-type star. Note how the magnetic tension force is now the primary contributor to the wind's acceleration at large distances.

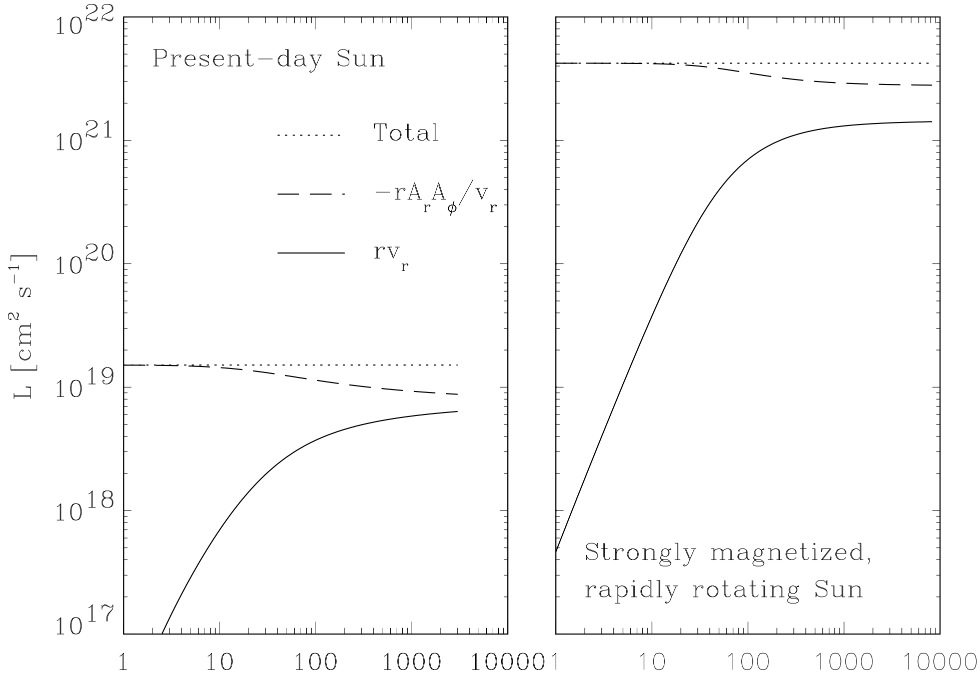

Figure 5.3: {fig:WDamloss} Contributions to the total angular momentum carried away by the wind for the two WD solutions depicted on Figs. 5.1 (left) and 5.2 (right). The solid line is the angular momentum per unit mass, and the dashed line the torque density produced by magnetic tension in the deformed magnetic field. The sum of these two contributions (dotted line), the total specific angular momentum, is a constant of motion, as per eq. (5.15).

look in to such numerical solutions, computed a few years ago by R. Keppens and H. Goedbloed (see bibliography). These solutions are particularly interesting because they are Weber-Davis-like in a number of ways: steady ($\partial/\partial t = 0$), axisymmetric ($\partial/\partial\phi$) and polytropic ($\alpha = 1.13$), and computed in the ideal MHD limit. The magnetic configuration they simulate is qualitatively similar to the magnetostatic solution depicted on Fig. 4.2, in that in contains a closed-fieldline region symetrically straddling the equator, and open fieldline regions over the pole. In addition to the magnetic field strenth and rotation parameter, a third parameter is now introduced to set the latitudinal extent of the closed field region (often called "dead zone" because **u** must vanish therein, due to the flux-freeezing constraint imposed by ideal MHD). The solutions are obtained as a time-dependent relaxation problem, starting with a purely hydrodynamical rotating wind solution, and a magnetic field patched up as a combination of a dipole for the closed region, and split monopole for the open regions. The solution is then integrated forward in time until a steady-state is attained.

The top panel on Figure 5.4 shows a solar-like solution, with a 2G polar field strength and a closed region extending $\pm 30°$ on either side of the equator. Note first how the wind outflow is directed along the magnetic fieldlines, as it must in a steady state as per the flux-freezing constraint characterizing ideal MHD. At mid-latitude, the solution shows many similarities to the Weber-Davis solution of Fig. 5.1. The slow magnetosonic surface is well within the Alfvén critical surface, and the latter very very nearly coincide with the fast magnetosonic surface. (these were all critical points in the 1D WD solution of §5.1, cf. the triangle, diamond and solid dot on Fig. 5.1). At low latitudes, the effect of the closed field region alters the flow quite significantly, although beyond $10\,R_\odot$ or so the wind speed is comparable to that at high latitudes. The wind density, however, is larger by about a factor of three. Close examination of the Figure will reveal that the outflow speed has a poleward-directed latitudinal component, which turns out to be very well fitted by a $\sin(2\theta)$ dependency at heights much larger than the radial extent of the closed field region.

As can be seen on the bottom panel of Figure 5.4, doubling the field strength and latitudinal extent of the magnetically closed region leaves these basic solution characteristics unaltered. Not surprisingly, away from the closed region the solution is characterized by a greater degree of spherical symmetry, which is what is to be expected in a split-monopole configuration where the field is better able to channel the flow without being distorted. Indeed, the shape of poloidal fieldlines (solid lines) show a striking resemblance to those characterizing the magnetostatic solutions considered in §4.3 (cf. Fig. 4.2).

Figure 5.5 depicts yet another wind solution, this time for a sun rotating at 20 times its present rate, but maintaining the same surface magnetic field configuration and strength as on Fig. 5.4. The impact of this high rotation rate on the wind is substantial in many ways. As on the Weber-Davis solution of Fig. 5.2, the fast magnetosonic surface is now well-separated from the Alfvén surface. Rotation leading to tighter winding of the magnetic field, the latitudinal magnetic pressure gradient associated with the strong toroidal field component leads to a collimation of the wind towards the poles. In addition, efficient magnetocentrifugal driving at low latitudes leads to enhanced densities in and near the equatorial plane.

5.3 Stellar spin-down {sec:spindown}

Even though it is primarily thermally-driven, the solar-like WD solution of Fig. 5.1 is losing far more angular momentum than in the absence of magnetic fields, as per eq. (5.20). This, it turns out, goes a long way in explaining

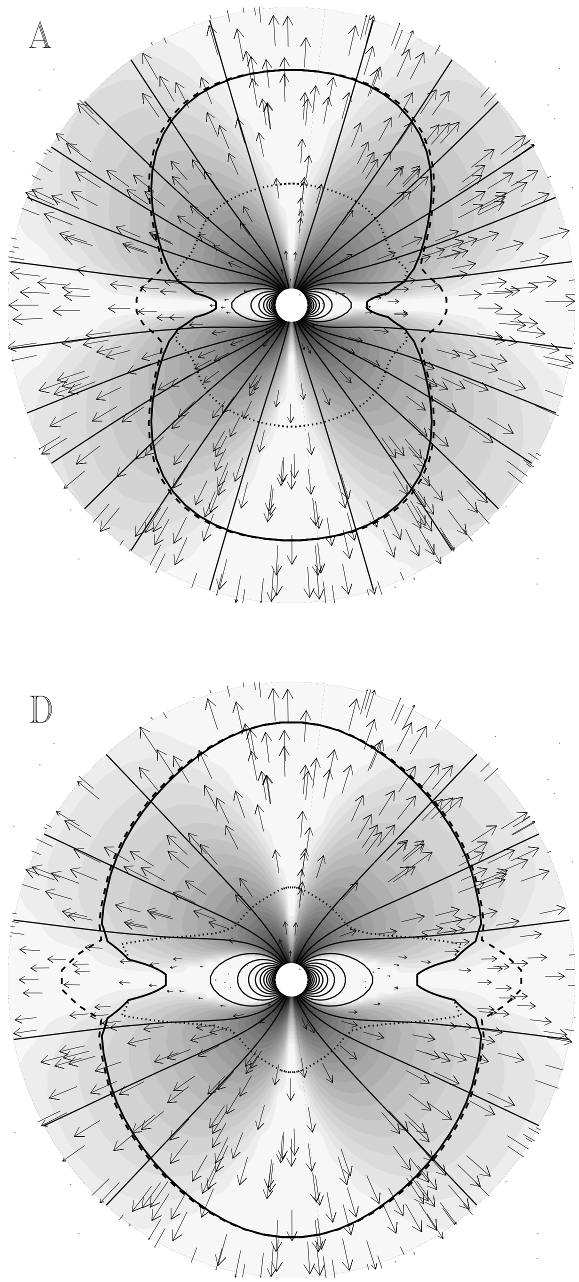

Figure 5.4: {`fig:numWD1`} Axisymmetric 2D MHD model of the solar wind. The flow field is indicated as vector, the poloidal magnetic fieldlines by solid lines, and the gray scale encodes the strength of the toroidal magnetic component. The dotted, dashed and thick solid lines are respectively the slow and fast magnetosonic surfaces, and the Alfvén surface. The top solution is for present-day solar parameters, while the bottom solution pertains to a strongly magnetized sun (see text). Graphics courtesy of Rony Keppens (U. Leuven).

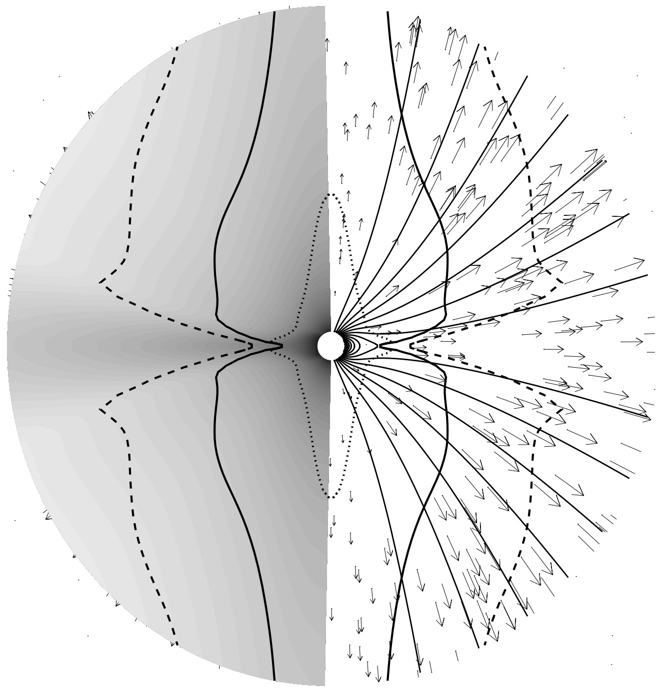

Figure 5.5: {fig:numWD3} Similar to Fig. 5.4A, but for a sun rotating at 20 times its present rotation rate. The grayscale on the left now encodes the density, rather then the toroidal field strength. Note the poleward collimation of the wind into a polar jet, and the density excess in the equatorial plane. Graphics courtesy of Rony Keppens (U. Leuven).

the very peculiar distribution of observed stellar rotational velocities on the main-sequence.

5.3.1 Stellar rotation: the observational picture {ssec:rotstar}

The rotation of a star other than the Sun as first measured serendipitously at the beginning of this century by F. Schlesinger, in the brighter component of the eclipsing binary δLibrae at occultation. Subsequent determinations of rotation rates for single stars relied on the Doppler broadening of spectral lines, as originally suggested by W.W. Abney in 1877, but first succesfully executed much later, in 1929, by C.T. Elvey. For a single star, this projected rotational velocity ($v \sin i$) yields only a lower limit on the true equatorial rotation rate, as the angle i between the line of sight and the star's rotation axis is generally unknown.

As increasingly sensitive spectroscopic determinations of $v \sin i$ for a growing sample of single stars accumulated, the existence of systematic differences between the average rotation rates for late-type versus early-type stars was soon noted. Figure 5.6 below is a reproduction of a diagram put together by R. Kraft in 1967, showing the distribution in a HR diagram of $v \sin i$'s measured in a sample of field stars. As one runs down the main sequence, there occurs a sharp drop in $v \sin i$ starting around spectral type F5. Slow rotation is the rule on the cool side of this so-called rotational dividing line, while on the hot side rapid rotation is common. Kraft went on to show that under the assumption of solid-body rotation, in the interval $1.5 \lesssim M/M_\odot \lesssim 20$ observed rotation rates are consistent with a power-law dependence betwen stellar angular momentum (J) and mass (M) of the form

$$J \propto M^{1.57} \ . \tag{5.37}$$ {eq:Kraft}

It was already understood then that the decrease in the moment of inertia of stars associated with their contraction towards the main-sequence could easily account for ZAMS equatorial rotational velocities of a few hundreds of kilometers per second, so that the anomaly in Kraft's diagram was in fact with the slowly-rotating low-mass stars. Rather than some strongly mass-dependent process (such as proto-early-type-stars diverting a substantial fraction of their spin angular momentum into planetary orbital angular momentum, for example), the favored interpretation back then was already that late-type stars somehow lose angular momentum on the main-sequence, i.e., they undergo spin-down.

Spectacular evidence for such main-sequence spin-down was provided in the short, now classical 1972 paper by Andy Skumanich[5]. Figure 5.7, reproduced from this paper, illustrate the gradual decrease of average rotation rates for late-type stars in a few (young) open clusters of known ages.

Later observations focusing on young open clusters such as αPersei and the Pleiades have revealed that main-sequence spin-down for late-type stars is very swift, with the bulk of it completed in the first few 100 Myr after arrival on the ZAMS.

5.3.2 The Skumanich square root law {ssec:skulaw}

In case you haven't seen it coming yet, our WD wind models provide us with some of the key physical pieces required to understand main-sequence spin-down. Towards this goal, the most important result obtained in §5.1 is eq. (5.20), stating that the total angular momentum per unit mass (L) carried away by the wind is equal to that which would be carried away by an unmagnetized wind remaining in a state of strict co-rotation out to the

[5] Still today affectionately know to his HAO colleagues as Doctor Slamdunk

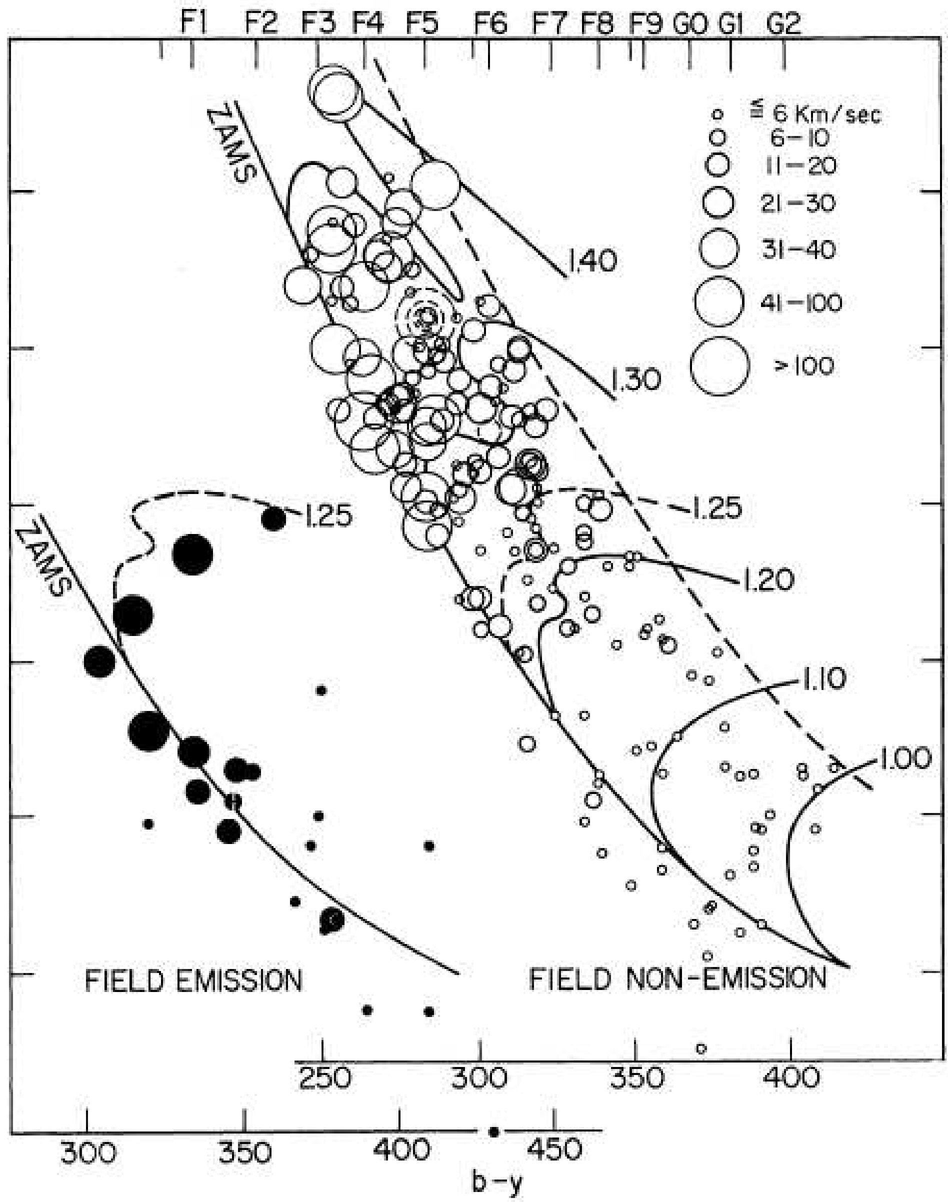

Figure 5.6: {fig:Kraft} Distribution of projected rotational velocities ($v \sin i$) for main-sequence stars, plotted in an observational HR diagram. Luminosity increases verticaly upwards, and effective temperature horizontally leftward. Astronmical spectral types are listed along the upper axis. Solid lines are stellar evolutionay tracks, labeled according to mass in solar units. These tracks, particularly for $M/M_{\odot} \gtrsim 1.2$, are now somewhat obsolete. Diagram reproduced from Kraft, R. 1967, ApJ, 150, 551 (Figure 1, p. 558).

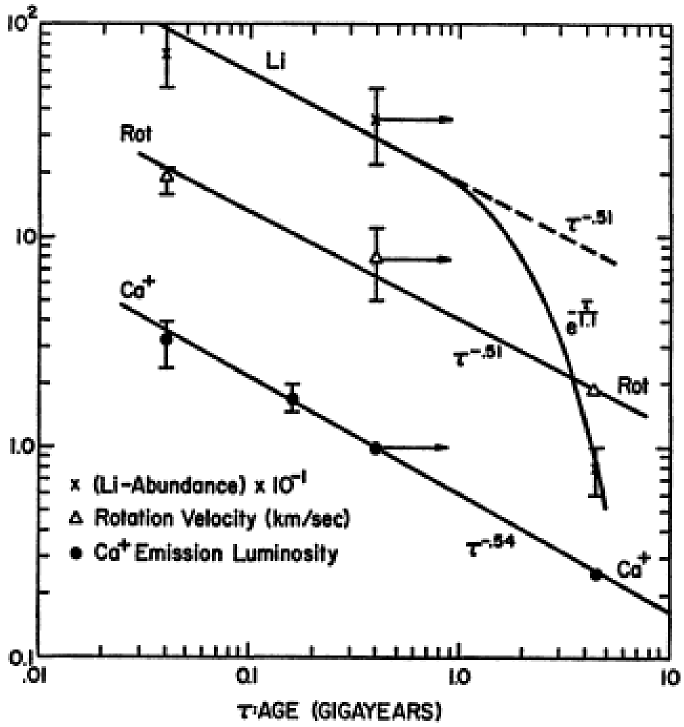

Figure 5.7: {fig:sku} Main-sequence temporal evolution of rotation rates, Calcium emission and Lithium abundance in solar-type stars. Diagram reproduced from Skumanich, A. 1972, ApJ, 171, 565 (Figure 1, p. 566).

Alfvén radius r_A:

$$L = \Omega r_A^2 \ . \tag{5.38}$$ {eq:amloss2}

To obtain the net angular momentum loss, we just need to multiply L by the wind's mass flux. However, eq. (5.38) holds only in the equatorial plane, where the WD solution is computed. We need to construct an equivalent expression for the whole sphere, which is not simply $4\pi\Omega r_A^2$. Remember that what matters for angular momentum extraction is the component of the flow moving away perpendicularly to the rotation axis. The WD model can be "stretched" to the whole sphere by assuming that a whole spherical shell is co-rotating out to r_A; this means replacing eq. (5.38) by:

$$L_{\rm sph} = \frac{2}{3}\Omega r_A^2 \ , \tag{5.39}$$ {eq:amloss3}

where the factor 2/3 simply arises from the moment of inertia. The angular momentum loss rate then follows directly:

$$\frac{{\rm d}J}{{\rm d}t} = \dot{M} \times L_{\rm sph} = -4\pi\rho_A r_A^2 u_{rA}\left(\frac{2}{3}\Omega r_A^2\right) \ . \tag{5.40}$$ {eq:lossrate1}

Now, at the Alfvén radius we have $u_{rA} = A_{rA}$, with $B_{rA}^2 = 4\pi\rho_A A_{rA}^2$. Moreover, conservation of magnetic flux implies $r_0^2 B_{r0} = r_A^2 B_{rA}$. Putting all this into eq. (5.40) leads to

$$\frac{\mathrm{d}J}{\mathrm{d}t} = \dot{M} \times L_{\rm sph} = -\frac{2}{3} B_{r0}^2 r_0^4 \Omega A_{rA}^{-1} \ . \qquad (5.41)$$ {eq:lossrate2}

Now, for rotating magnetized winds that are mostly thermally driven (as on Fig. 5.1), we have $A_{rA} \sim c_s$ to within a factor of two or so. If the coronal temperature is held fixed, this means that the angular momentum loss rate is only a function of the rotation rate and surface magnetic field strength. Both are known for the Sun, but how about the "young Sun" of 4.5 Gyr ago? If stars of one solar mass in αPersei or the Pleiades are representative of the ZAMS Sun, then its rotation could have been anywhere between 5 and 100 times its present value. How about its surface field strength? In later chapters of these notes we will encounter various lines of arguments, both observational and theoretical, indicating that it should increase with increasing rotation rate. Some of the dynamo models we will construct in chapter 10 would "predict" $B_{r0} \propto \Omega$. If this is the case, and for a fixed moment of inertia on the main-sequence (a very good approximation, for a change...), then eq. (5.40) would lead to

$$\frac{\mathrm{d}\Omega}{\mathrm{d}t} \propto \Omega^3 \ , \qquad (5.42)$$ {eq:sku1}

which readily integrates to

$$\frac{1}{\Omega^2(t)} - \frac{1}{\Omega^2(t_0)} \propto t - t_0 \ , \qquad (5.43)$$ {eq:sku2}

where t_0 is the time of arrival on the ZAMS (or shortly thereabouts). In the asymptotic limit $t \gg t_0$, $\Omega \ll \Omega(t_0)$, this becomes

$$\Omega(t) \propto t^{-1/2} \ , \qquad (5.44)$$ {eq:sku3}

which, how about that, is precisely the power-law relationship inferred observationally by Skumanich (cf. Fig. 5.7). Looks like we're in business!

5.3.3 The spindown of late-type stars {ssec:MSspindown}

The missing proportionality constant in eq. (5.44) is of course readily computed from our Weber-Davis solution; in fact we did nearly all the work already in ariving at eq. (5.41), the missing element being the expression of stellar angular momentum in terms of a star's angular velocity distribution.

If, for the time being, we assume that stars rotate as rigid bodies, then we have

$$J = I_* \Omega_* \ , \tag{5.45}$$ {eq:spd1}

and dimensional analysis yields the following expression for the **spin-down timescale**:

$$\tau_{\rm sp} = \frac{1}{I_* \Omega_*} \left(\frac{{\rm d}J}{{\rm d}t} \right)^{-1} \ . \tag{5.46}$$ {eq:spd2}

All we are missing are the stellar moments of inertia, which are readily computed if we have stellar structural models on hand. The third column of Table 5.4 list the resulting spin-down timecales, for ZAMS stellar models between 0.8 and 1.2 $M_\odot$. In all cases it as assumed that the ZAMS rotation period is one day, and the radial surface magnetic field strength is 50 G, reasonable numbers to the extent we can tell from observations and models of stellar formation and pre-main-sequence evolution.

Table 5.4
ZAMS spindown timescales for late-type stars

$M/M_\odot$	$R/R_\odot$	$I_*[10^{53}]$	$\tau_{J,*}$ [Myr]	$I_E[10^{53}]$	$\tau_{J,E}$ [Myr]
0.8	0.703	4.41	810	1.025	188
0.9	0.784	5.50	604	0.979	107
1.0	0.882	6.75	396	0.833	48.9
1.2	1.131	9.02	133	0.139	2.05

The spin-down timescales are of order 10^8 and 10^9 yr, which is nicely smaller than the solar age, but a factor of ten longer than the spin-down timecales inferred from $v \sin i$ determinations in young stellar clusters. Observations do offer an interesting hint, in that after arriving on the main-sequence, more massive stars seem to spin down *faster* than less massive, even though their moment of inertia is larger (second column of Table 5.4).

The favored escape from this quandary is to assume that the torque applied by the wind to the photospheric layers is not transmitted throughout the whole star, but (at first anyway) only to its convective envelope, where the vigorous turbulent thermally-driven convective fluid motions are expected to redistribute momentum on the convective turnover time, of the order of a month for convection in solar-type stars. Now, the thickness of the convection decreases rapidly as mass increases, leading to a decrease of the moment of inertia of main-sequence convective envelop with increasing mass (see fifth column in 5.4. This then leads to spin-dowm times (last column in Table 5.4) that (a) are in much better agreement with observationally-inferred values (2) *decrease* with increasing mass. It all fits together!

In late type stars spun down by a wind-mediated surface torque, many physical processes can exchange angular momentum between the convective envelope and underlying radiative core. Indeed, helioseismology has shown that the angular velocity of the solar core is comparable to that of its convective envelope, implying that whatever dynamical coupling is taking place between the core and envelope acts on a timescale much smaller than the solar age (but still significantly longer that the ZAMS spindown timescales, otherwise we're in trouble again). It turns out that internal magnetic fields can do the trick, and remain at this writing the most physically viable explanation for the rotation rate of the solar radiative core. To substantiate this claim would take us too deep inside the sun, but references listed in the bibliography to this chapter provide good entry points into this area of research. Time to get back up into the wind and see what we can do about those famous high-speed streams...

5.4 Wind driving by Alfvén waves {sec:wavewind}

In the solar photosphere, the plasma-β is high enough that magnetic fieldlines get continuously displaced by convective fluid motions. Vertical displacements will generate magnetosonic waves, which are expected to shock and dissipate before they reach the corona. Horizontal displacements of magnetic fieldlines, on the other hand, will propagate upwards into the corona in the form of Alfvén waves. These, it turns out, can have a significant dynamical impact on wind-like outflows, and this is what we'll look into in this section.

The physical/geometrical setup we consider here closely resembles that of the Weber-Davis solution of §5.1, i.e., working in spherical polar coordinates we solve the steady ($\partial/\partial t = 0$) axisymmetric ($\partial/\partial\phi = 0$) wind equations in the equatorial plane of the star, assuming a radial reference magnetic field. The two important differences are: (1) rotation is neglected, and (2) we consider an isothermal, rather than polytropic wind, otherwise the mathematics really gets too messy.

The key in formulating the wave-wind model is to assume that the total flow and magnetic field can be written as

$$\mathbf{u}(r,t) = u_r(r)\hat{\mathbf{e}}_r + \delta u(r,t)\hat{\mathbf{e}}_\phi \ , \qquad (5.47)$$ {eq:aw1a}

$$\mathbf{B}(r,t) = B_r(r)\hat{\mathbf{e}}_r + \delta B(r,t)\hat{\mathbf{e}}_\phi \ , \qquad (5.48)$$ {eq:aw1b}

where u_r, B_r define the large-scale wind outflow, and the two leftmost terms correspond to a transverse wave travelling in the r-direction and "oscillating" in the ϕ-direction; that latter choice is entirely arbitrary, but will facilitate the mathematical developments to follow. As with any wave, the time averages of the local wave contribution to the flow and field vanish:

$$\langle \delta u \rangle = 0 \ , \qquad \langle \delta B \rangle = 0 \ . \tag{5.49}$$ {eq:aw1c}

5.4.1 The magnetic force exerted by Alfvén waves

Looking at the momentum equation, you should be able to convince yourself that the contribution to the force per unit volume (f_w) associated with the wave component is given by:

$$f_w = \left(\rho(\delta\mathbf{u} \cdot \nabla)\delta\mathbf{u} + \frac{1}{4\pi}(\nabla \times \delta\mathbf{B}) \times \delta\mathbf{B} \right)_r \ . \tag{5.50}$$ {eq:aw2}

For the assumptions embodied in eqs. (5.47)–(5.49), time averaging of this expression over a wave period yields

$$\langle f_w \rangle = \frac{\rho \, \langle \delta u^2 \rangle}{r} - \frac{\langle \delta B^2 \rangle}{4\pi r} - \frac{\mathrm{d}}{\mathrm{d}r}\left(\frac{\langle \delta B^2 \rangle}{8\pi} \right) \ . \tag{5.51}$$ {eq:aw3}

There are thus two contribution to the wave-induced force: a centrifugal force associated with the wave displacement in the ϕ-direction (first term on RHS of eq. (5.51), and a Lorentz force that can be broken into tension and magnetic pressure gradient contributions[6].

But how to we compute δu and δB? Simply by solving the ϕ-components of the momentum and induction equations, which here reduce to

$$\frac{\partial}{\partial t}\delta u_r + \frac{u}{r}\frac{\partial}{\partial r}(r\delta u) = \frac{B_r}{4\pi\rho r}\frac{\partial}{\partial r}(r\delta B) \ , \tag{5.52}$$ {eq:aw4a}

$$\frac{\partial}{\partial t}\delta B = \frac{1}{r}\frac{\partial}{\partial r}\left(r(B_r\delta u - u_r\delta B)\right) \ , \tag{5.53}$$ {eq:aw4b}

with B_r and u_r given by the "wind" part of the governing equation; these take on the usual form for a steady, spherically-symmetric outflow (cf. §3.3), except that now the isothermality assumption leads to

$$u_r\frac{\mathrm{d}u_r}{\mathrm{d}r} = -\frac{a^2}{r} - \frac{GM}{r^2} + \frac{\langle f_w \rangle}{\rho} \ , \tag{5.54}$$ {eq:aw5a}

[6]As you get to verify in problem XXX below, the spherical geometry is essential here in producing a non-zero time-averaged wave force.

where $a = \sqrt{kT/m}$ is the **isothermal sound speed** for a perfect gas. As in the Webed-Davis case, the constraints of mass and magnetic flux conservation lead to

$$\frac{\rho(r)}{\rho_0} = \left(\frac{r_0}{r}\right)^2 \left(\frac{u_{r0}}{u_r}\right) \ , \tag{5.55}$$ {eq:aw5b}

$$\frac{B_r(r)}{B_{r0}} = \left(\frac{r_0}{r}\right)^2 \ , \tag{5.56}$$ {eq:aw5c}

So, in principle all is well: with $B_r(r)$ known from (5.56) and provided all needed physical quantities are specified at the coronal base r_0, we have here a set of four coupled equations (namely (5.52), (5.53), (5.54), and (5.55)) for the four unknown functionals u_r, ρ, δu and δB. However, the strong nonlinear coupling mediated by eq. (5.51) is not easy to deal with in general, so we need to introduce an additional approximation into the model.

5.4.2 The Wave force in the WKB approximation {ssec:wfWKB}

The heart of the so-called WKB approximation is to assume that the wavelength λ of the propagating Alfvén wave is much smaller than the length scale ℓ over which the background flow is varying. In such cases one can expand the wave amplitudes δu and δB as

$$\delta u(r,t) = \left[\delta u_1(r) + \epsilon\delta u_2(r) + \epsilon^2\delta u_3(r) + ...\right] \exp(i[\psi(r) - \omega t]) \tag{5.57}$$ {eq:aw6}

with a similar expression characterizing δB, $\varepsilon = \lambda/\ell = 2\pi/k\ell$ is a small parameter, and $k(r) = \mathrm{d}\psi/\mathrm{d}r$ is the radius-dependent wavenumber. Inserting these expression into eqs. (5.52)—(5.53), one then equates all terms of similar power in ϵ. To lowest order this yields

$$\omega = k(u_r + A_r) \ , \tag{5.58}$$ {eq:aw7a}

$$\delta u_1 = \pm\frac{\delta B}{\sqrt{4\pi\rho}} \ , \tag{5.59}$$ {eq:aw7b}

with the minus sign retained in what follows, since it corresponds to the outward propagating waves. Substituting these expression in the first order equations leads to a differential equation for δB_1, which (it can be shown...) integrates to

$$\delta B(r) = \delta B_0 \left(\frac{M_{A0}}{M_A}\right)^{1/2} \left(\frac{1+M_{A0}}{1+M_A}\right)^{1/2} \tag{5.60}$$ {eq:aw7c}

where

$$M_A \equiv \frac{u_r}{A_r} = M_{A0}\left(\frac{\rho_0}{\rho}\right)^{1/2} \tag{5.61}$$ {eq:aw8}

is the **Alfvénic Mach Number**, the Alfvén speed A_r here being the component associated with the radial magnetic field component, and the subscript "1" has been dropped for clarity. The corresponding expression for $\delta u(r)$ follows directly from eq. (5.59). Substituting all this back into eq. (5.51) for the time-averaged wave force, one eventually arrives at

$$\langle f_w \rangle = -\frac{\mathrm{d}}{\mathrm{d}r}\left(\frac{\langle \delta B^2 \rangle}{8\pi}\right) = \frac{\langle \epsilon_w \rangle}{4}\left(\frac{1+3M_A}{1+M_A}\right)\left(\frac{2}{r}+\frac{1}{u_r}\frac{\mathrm{d}u_r}{\mathrm{d}r}\right) \tag{5.62}$$ {eq:aw11}

where

$$\langle \epsilon_w \rangle = \frac{\langle \delta B^2 \rangle}{4\pi} = \frac{\langle \delta B_0^2 \rangle}{4\pi}\left(\frac{M_{A0}}{M_A}\right)\left(\frac{1+M_{A0}}{1+M_A}\right)^2 \tag{5.63}$$ {eq:aw10}

is the wave energy density. The RHS of eqs. (5.62)—(5.63) now involve only properties of the large-scale outflow, so in principle we can proceed with confidence.

5.4.3 Obtaining wind solutions

Getting a complete wind solution once again is done numerically. Substituting the expression for the wave force obtained above into the r-component of the momentum equation (5.54) leads, after a fair bit of algebraic juggling, to

$$\left[u_r^2 - a^2 - \frac{\langle \epsilon_w \rangle}{4\rho}\left(\frac{1+3M_A}{1+M_A}\right)\right]\frac{r}{u_r}\frac{\mathrm{d}u_u}{\mathrm{d}r} = \left[a^2 - \frac{GM}{r} + \frac{\langle \epsilon_w \rangle}{2\rho}\left(\frac{1+3M_A}{1+M_A}\right)\right] \tag{5.64}$$ {eq:awsol1}

where a is the isothermal sound speed. This equation is best treated as a initial value problem for u_r, of the general form:

$$\frac{\mathrm{d}u_r}{\mathrm{d}r} = g(r)\ . \tag{5.65}$$ {???}

Assuming a starting guess for the base flow speed u_{r0} (for example that of the pure isothermal solution), eq. (5.64 is integrated forward in r using some suitable ODE integration scheme (see Appendix XXX). The problem is that the solution must go through a sonic critical point. If the starting guess is wrong, as one integrates forward in r there will come a point where the solution will diverge (infinite acceleration). The starting guess must then be adjusted (upwards or downwards depending on how divergene occurs, and the process repeated until one finds a solution that shoots smoothly through the critical point and sails away ever on and on until you reach a value of r that is a good enough approximation of infinity for practical purpose (we'll settle for $100\,r_0$ in what follows). Then it is a simple matter to reconstruct ρ via eq. (5.55), then δB via (5.60), and finally δu via (5.59).

5.4.4 Some representative solar solutions {ssec:awsolar}

We first consider the effect of Alfvén wave driving on solar-type outflows. Accordingly, the reference parameters for our reference isothermal solution (without waves) are chosen to produce wind characteristics at 1U commensurate with low-speed streams; we thus set $T_0 = 10^6\,K$ leading to an isothermal sound speed $a = XXX\,\mathrm{km\ s^{-1}}$ and $N = 2\times 10^7\,\mathrm{cm^{-3}}$. This flow has a base speed $u_{r0} = 1.19\,\mathrm{km\ s^{-1}}$ becomes supersonic at $r_s/r_0 = XXX$, and reaches $\simeq 450\,\mathrm{km\ s^{-1}}$ at 1AU. We set the base radial magnetic field at $B_{r0} = 1\,\mathrm{G}$, and set the size of the magnetic pertubation at the base via the parameter

$$\alpha = \left(\frac{\delta B_0}{B_0}\right)^2 , \qquad (5.66)$$ {eq:driveamp}

measuring the base ratio of magnetic energy density in the wave to that in the background magnetic field.

Wind profiles are shown on Fig. 5.8 for three values of the parameter α, with the corresponding profiles of wave amplitudes plotted on Fig. 5.9. Several features of these wind solutions are noteworthy. The wind speed is an increasing function of Alfvén wave amplitude (not surprisingly), but the increase is proportionaly greater at the base of the wind (from $4.85\mathrm{km\,s^{-1}}$ at $\alpha = 0.01$ up to $37.35\mathrm{km\,s^{-1}}$ at $\alpha = 0.1$) as compared to the wind speeds at large distances (from $888\mathrm{km\,s^{-1}}$ to $1005\mathrm{km\,s^{-1}}$ at $100\,r_0$). As mentioned in the preceding chapter, this behavior is characteristic of situations where additional momentum occurs primarily within the sonic point, located at $r_A/r_0 = 13.53$, 8.54 for $\alpha = 0.01$, 0.03 and 0.1, respectively. The wave amplitude decrease rapidly with distance, reflecting the $1/r^2$ behavior of B_r but also the fact that wave energy is being transfered to the wind. The increase of $\langle\delta u^2\rangle$ in the first ten r_0 or so is a direct reflection of the rapid decrease of the density with r. This causes the inertia of the fluid to decrease faster than the wave's restoring force, leading to an increase in $\langle\delta u^2\rangle$ even though $\langle\delta B^2\rangle$ falls off rapidly (cf. eq. 5.59). Indeed it can be easily shown that in the sub-Alfvénic portion of the wind the wave amplitude scale as $\langle\delta u^2\rangle^{1/2} \sim \rho^{1/4}$ and $\langle\delta B^2\rangle^{1/2} \sim \rho^{-1/4}$ while in the superAlfvénic portion of the wind, where the Alfv'én speed is nearly constant, the amplitudes scale as $\langle\delta u^2\rangle^{1/2} \sim r^{-1/2}$ and $\langle\delta B^2\rangle^{1/2} \sim r^{-3/2}$.

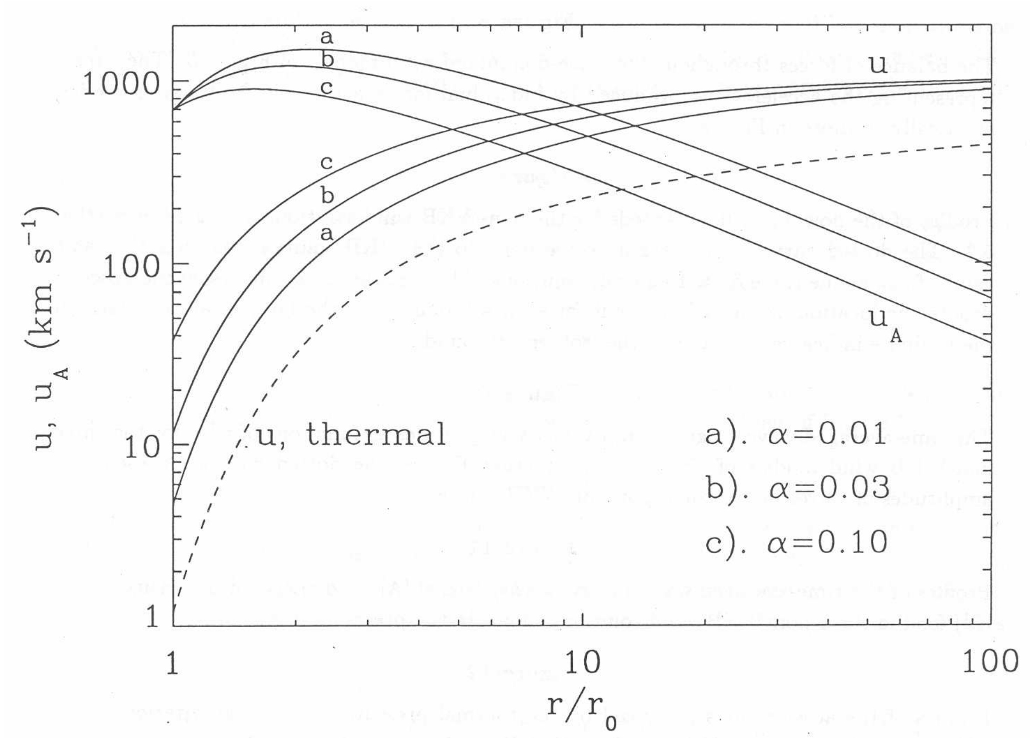

Figure 5.8: {fig:aw1} Radial profiles of the flow and Alfvén speed (here denoted u_A) for wave-driven isothermal wind models in the WKB approximation. The three solutions depicted here are obtained for three distinct values of the forcing amplitude parameter $\alpha = (\delta B_0/B_0)^2$, as labeled. The dashed curve is the flow profile for an equivalent isothermal solution without any wave driving. Reproduced from the MacGregor & Charbonneau 1994 book chapter cited in the bibliography.

Figure 5.10 depicts details of the force balance in the $\alpha = 0.01$ and 0.1 solutions. As with the Parker wind solution considered in §3.3, near the base the flow is in near-hydrostatic equilibrium, with the Alfvén wave force contributing little even at $\alpha = 0.1$. However the wave force rapidly starts to dominate the dynamics at larger distances, exceeding the thermal pressure force beyond the Alfvén point.

You shouldn't be too impressed by the $\sim 1000\,\mathrm{km\ s}^{-1}$ asymptotic speeds of our wave-enhanced wind solutions. Even though this is largely sufficient to account for high-speed streams, in fact the isothermality assumptions guarantees that the asymptotic flow speed tends to... infinity! (Haven'y you done Problem XXX already?). What *is* noteworthy is that beyond the sonic point, the wind solution with WKB Alfvén waves has a flow speed a factor of about two larger than the reference wave-free isothermal solution,

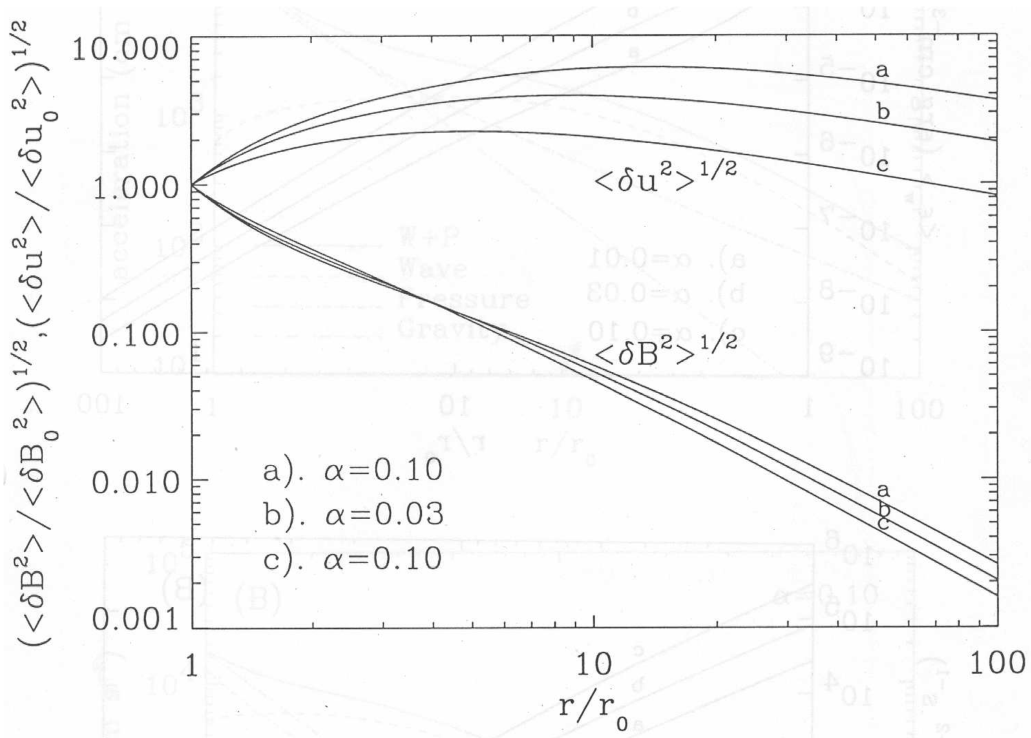

Figure 5.9: {fig:aw2} Time averaged wave velocity and magnetic amplitudes as a function of r/r_0, for the solutions of Fig. 5.8. In all cases amplitudes are normalized to their value at the reference radius r_0 On each panel the solid curve is the sum of the wave and thermal pressure gradient accelerations. Reproduced from the MacGregor & Charbonneau 1994 book chapter cited in the bibliography.

which *is* the speedup factor suggested by Table 3.1. And this results, it turns out, does carry over to polytropic version of the model.

5.4.5 Wave-driven winds {ssec:awwind}

There many classes of non-solar late-type stars that show evidence for wind-like outflows, yet their inferred coronal temperatures are too low for sustaining a (mostly) thermally-driven wind. These include

All the results discussed so far are predicated on the use of the WKD approximation in computing the force exterted by the Alfvén waves on the flow. This is expected to be a good approximation provided the wave period is much shorter than the advective transit time on the wind over a distance over which background properties of the flow (in particular the Alfvén speed) vary significantly. For the solar-type solutions considered in §5.4.4, it can be

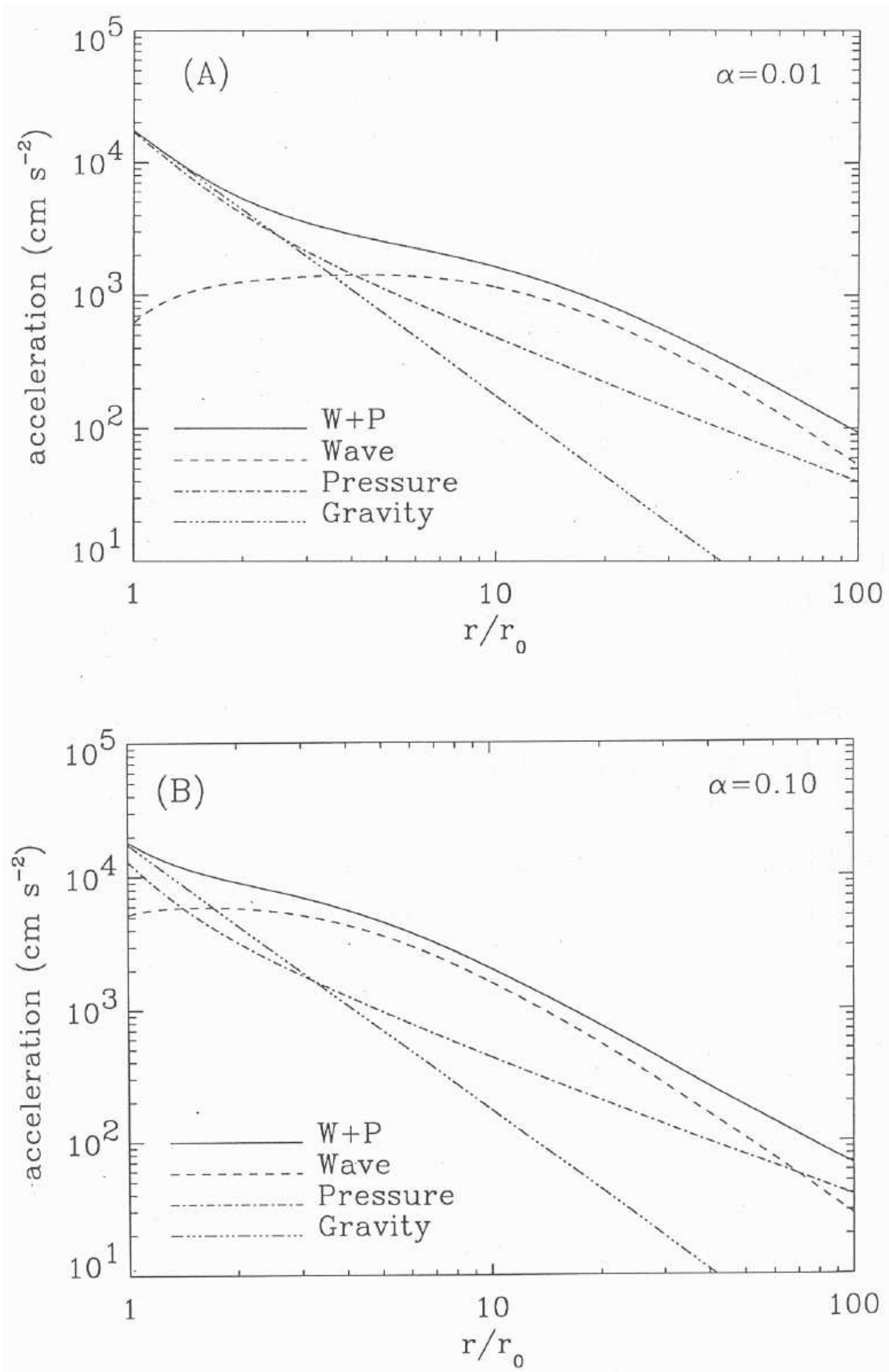

Figure 5.10: {fig:aw3} Force balance in the (A) $\alpha = 0.01$ and (B) $\alpha = 0.1$ solutions of Figs. 5.8–5.9. On each panel the solid curve is the sum of the wave and thermal pressure gradient accelerations. Reproduced from the MacGregor & Charbonneau 1994 book chapter cited in the bibliography.

verified that waves with periods larger than about 10 minutes will violate the WKB constraints near the base of the flow, where the gradient in Alfvén speed is substantial (cf. 5.8). Now, ten days is about the turnover time for solar photospheric granules, so wave power in this period range may well be significant. It turns out that relaxing the WKB approximation has little impact on solar-type solutions, but large differences do materialize in wind models where the wave force is the primarily driver. If you wish to look deeper into this aspect of the problem, see the references listed in the bibliography to this section.

Problems:

1. Repeat derivation of $\langle f_w \rangle$ for homogeneous cartesian flow.
2. Work out the missing mathematical steps leading to eq. (5.19).
3. Make the (bold) assumption that the Weber-Davis solution obtained in §5.1 remains valid outside of the equatorial plane. Demonstrate that if this is the case, there exists an unbalanced force term in the θ-component of the momentum equation. Discuss in qualitative terms (no actual calculations) how the wind solution would be altered.
4. Assuming that stars arrive on the zero-age main-sequence in state of solid-body rotation, calculate their subsequent rotational evolution on the main-sequence. Plot Ω as a function of time, for stars of 0.8 and $1.2\,M_\odot$, and ZAMS rotation rates of 200, 100, 50 and $10\,\mathrm{km\,s^{-1}}$, under the following two assumptions regarding internal angular momentum disribution:

   (a) The stars rotate as rigid bodies throughout main-sequence evolution

   (b) Only the outer convective envelope is spun-down by the wind-mediated torque.

   Your starting point is eq. (5.41), with the additional "dynamo" assumption $B_{r0} \propto \Omega$ already encountered in deriving Skumanich's square root law, and the moment of inertia data listed in Table 5.4. How does the assumption made regarding internal angular momentum redistribution affect the spreads in rotation rates at age 100Myr? 1Gyr?
5. Work out the missing mathematical steps leading to eq. (5.64).

Bibliography:

The extension of the Sun's magnetic field into interplanetary space by the solar wind is clearly described in

Parker, E.N. 1963, *Interplanetary Dynamical Processes* (New York: John Wiley), chap. 10.

The original paper by Weber and Davis is

Weber, E.J., & Davis, L. Jr. 1967, ApJ, 148, 217.

but the presentation of §5.1 follows mostly

Belcher, J.W., & MacGregor, K.B. 1976, ApJ, 210, 498.

The confrontation of the WD wind model with the observed solar wind is discussed critically in

Pizzo, V., Schwenn, R., Marsch, E., Rosenbauer, H., Mülhaüser, K.-H., & Neubauer, F.M. 1983, ApJ, 271, 335.

If you are interested in application of the WD model to protostars or hot stars, see respectively:

Hartmann, L., & MacGregor, K.B. 1982, ApJ, 259, 180.
Friend, D.B., & MacGregor, K.B. 1984, ApJ, 282, 591.

The following are a small selection of noteworthy (very) early papers on stellar rotation, including those alluded to in §5.3.1:

Vogel, H.C. 1872, Astron. Nachr., 78, 248,
Abney, W. de W. 1877, MNRAS, 37, 278,
Schlesinger, F. 1909, Pub. Allegheny Obs., 1, 123,
Elvey, C.T. 1929, ApJ 70, 141.

The idea that magnetized outflows can lead to stellar angular momentum loss can be traced to

Schatzman, E. 1962, Ann. Astrophys., 25, 18.

In this context, another important early paper is

Mestel, L. 1968, MNRAS, 138, 359.

The first theoretical derivation of Skumanich's square root relation in the context of $\alpha\Omega$ dynamo theory is due to

Durney, B. 1972, in *Solar Wind*, eds. C.P. Sonett, P.J. Coleman, & L.M. Wilcox (Washington: NASA), p. 282.

There is a huge literature out there on the rotational evolution of late-type stars on or near the main-sequence; on the observational front, including the case for core-envelope decoupling, see:

Stauffer, J.R., & Hartmann, L.W. 1986, ApJ, 318, 337,
Soderblom, D.R., Stauffer, J.R., MacGregor, K.B., & Jones, B.F. 1993, ApJ, 409, 624,
A recent review?

while a good feel for the range of modelling approaches, assumptions (and debates!) can be obtained from taking a look at

MacGregor, K.B., & Brenner, M. 1991, ApJ, 376, 204,
Charbonneau, P., & MacGregor, K.B. 1993, ApJ, 417, 762,
Chaboyer, B., Demarque, P., & Pinsonneault, M. 1995, ApJ, 441, 865,
Keppens, R., MacGregor, K.B., & Charbonneau, P. 1995, A&A, 294, 469.
Talon, S.

The content of §5.2 is largely based on the following two papers:

Keppens, R., & Goedbloed, J.P. 1999, *Astron. Astrophys.*, **343**, 251-260,
Keppens, R., & Goedbloed, J.P. 2000, *Astrophys. J.*, **530**, 1036-1048,

The content of section 5.4 is, to a large extent, adapted from the following book chapter:

MacGregor, K.B., & Charbonneau, P. 1994, in *Cosmic winds and the heliosphere*, ed. J.R. Jokipii, C.P. Sonett, & M.S. Giampapa, Tucson: University of Arizone Press, 327-ff.

with Figures 5.8, 5.9 and 5.9 in fact digitized straight out of these pages. The many “it can be shown” in §5.4.2, implicit or explicit, are substantiated in

Belcher, J.W. 1971, ApJ, 168, 509,
Hollweg, J.V. 1973, JGR, 78, 3643.

On the non-WKB generalization of these models, see

MacGregor, K.B., & Charbonneau, P. 1994, ApJ, 430, 387,
Charbonneau, P., & MacGregor, K.B. 1995, ApJ, 454, 901,

and references therein. On the observational front, a very recent and spectacular breakthrough is the direct observation of Alfvén waves in the solar corona; see

Tomczyk, S. Science .

Part III

Astrophysical Dynamos

Chapter 6

The solar cycle as a dynamo {chap:solcycle}

If the sun did not have a magnetic field, it would be as boring a star as most astronomers believe it to be.

Attributed variously to E.N. Parker or R.B. Leighton

The aim of this chapter is to introduce observational aspects of the solar magnetic activity cycle that have most direct relevance as constraints to the dynamo mechanism and models that will occupy us in this third part of the course. By the time we're done, you will hopefully begin to appreciate the fact that the statement cited above is definitely not an understatement!

Once again we turn to the sun as an exemplar of astrophysical magnetohydrodynamics, this time with regards to dynamo action. As with the wind models discussed in part II, this is not because the solar dynamo is more simple or complicated or interesting than other astrophysical duynamos, but simply because ot is the dynamo for which we have the most observational information. Even more so than the geodynamo in the Earth,s core, in fact, because as we shall see, the dynamo-powered solar magnetic cycle operates on a timescale commensurate with the human lifespan, rather than glacial or geological.

6.1 The solar cycle {sec:obssolcycle}

6.1.1 Sunspots {ssec:sunspots}

Until the beginning of the twentieth century, the story of the solar activity cycle is coincident with the story of **sunspots**. As their name suggest, sunspots look like dark blemishes on the solar disk, but the vast majority are too small to be readily visible without a telescope. Only the largest sunspots can be visible to the naked-eye under suitable viewing conditions, for example when the sun is partially obscured by clouds or mist, particularly at sunrise or

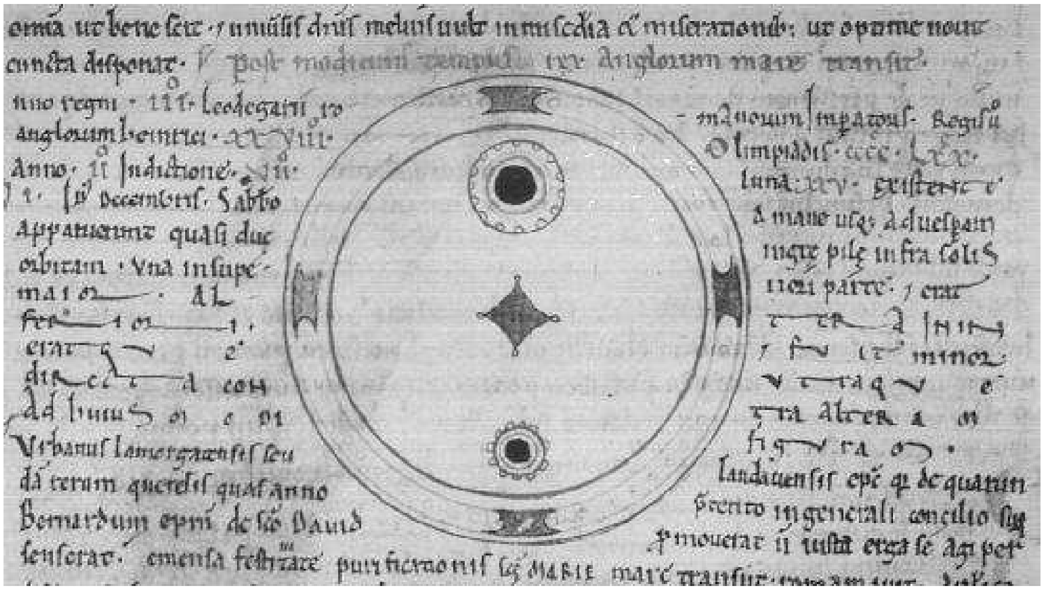

Figure 6.1: {F1.1} Sunspot drawing in the Chronicles of John of Worcester, twelfth century. Notice the depiction of the penumbra around each spot. Reproduced from R.W. Southern, *Medieval Humanism*, Harper & Row 1970, [Plate VII].

sunset. Numerous such sighting exist in the historical records, starting with Theophrastus (374-287 B.C.) in the fourth century B.C. However, by far the most extensive pre-telescopic records are found in the far east, especially in the official records of the Chinese imperial courts, starting in 165 B.C.

Figure 6.1 represents, to the best of our knowledge, the first surviving sunspot drawing, from a sighting on Saturday, 8 December 1128. The drawing is found in the Chronicles of John of Worcester, one of the many monks who contributed to the Worcester Chronicles. The accompanying text translates to something like:

> "...from morning to evening, appeared something like two black circles within the disk of the Sun, the one in the upper part being bigger, the other in the lower part smaller. As shown on the drawing." (trans. A. Van Helden)

The facts that the Worcester monks could apparently distinguish the umbrae and penumbrae of the sunspots they observed suggests that these spots must have been exceptionally large.

A fascinating pre-telescopic sunspot sighting is certainly that of 28 May 1607 by none other than Johannes Kepler (1571-1630). Kepler had been observing the sun for over a month using his *camera obscura* projection technique, basically a pinhole camera. He was hoping to detect a transit of mercury across the solar disk, as predicted by extant planetary ephemerides, and was well aware of the latter's deficiencies. But on May 28 he did noticed a small black spot on the solar disk, and concluded that he was indeed seeing Mercury in transit. It did not take long before he came to realize his mistake.

In the first decade of the seventeenth century, four astronomers more or less simultaneously turned the newly invented telescope toward the Sun, and noted the existence of sunspots. They were Johann Goldsmid (1587-1616, a.k.a. Fabricius) in Holland, Thomas Harriot (1560-1621) in England, Galileo Galilei (1564-1642) in Italy, and the Jesuit Christoph Scheiner (1575-1650) in Germany. Fabricius was the first to publish his results in 1611, and to correctly interpret the apparent motion of sunspots in terms of axial rotation of the Sun. Like Harriot, Fabricius and his father (the then-well-known astronomer David Fabricius) first observed sunspots directly through their telescope shortly after sunrise or before sunset. The harrowing account of their observations is worth quoting: (excerpt from the translation in the paper by W.M. Mitchell cited below):

> "... Having adjusted the telescope, we allowed the sun's rays to enter it, at first from the edge only, gradually approaching the center, until our eyes were accustomed to the force of the rays and we could observe the whole body of the sun. We then saw more distinctly and surely the things I have described [sunspots]. Meanwhile clouds interfered, and also the sun hastening to the meridian destroyed our hopes of longer observations; for indeed it was to be feared that an indiscreet examination of a lower sun would cause great injury to the eyes, for even the weaker rays of the setting or rising sun often inflame the eye with a strange redness, which may last for two days, not without affecting the appearance of objects."

Galileo and Scheiner, however, were the most active in using sunspots to attempt to infer physical properties of the Sun (Figs. 6.3, 6.4). To Galileo belongs the credit of making a convincing case that sunspots are indeed features of the solar surface, as opposed to intra-Mercurial planets (Scheiner's original position).

6.1.2 The sunspot cycle {ssec:schwabe}

Early sunspots observers noted the curious fact that sunspots rarely appear outside of a latitudinal band of about $\pm 30°$ centered about the solar equator, but otherwise failed to discover any clear pattern in the appearance and disappearance of sunspots. In 1826, the German amateur astronomer Samuel

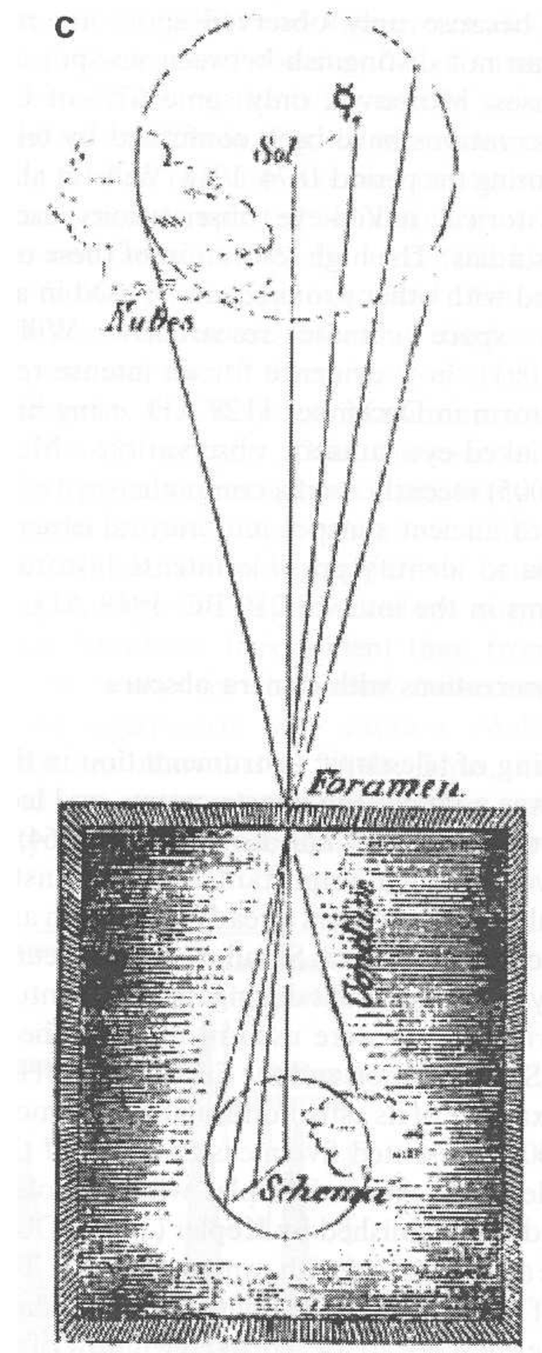

Figure 6.2: {`fig:keplerspot`} Naked-eye observation of a sunspot on 18 May 1607 by Johannes Kepler. Observing the sun intermittently on a cloudy day, Kepler could only make a few observations, and concluded he had had the good fortune of catching the planet Mercury in transit across the solar disk. Diagram reproduced from Vaquero, J.M. 2007, *Adv. Sp. Res.*, **40**, 929 [Fig. 2].

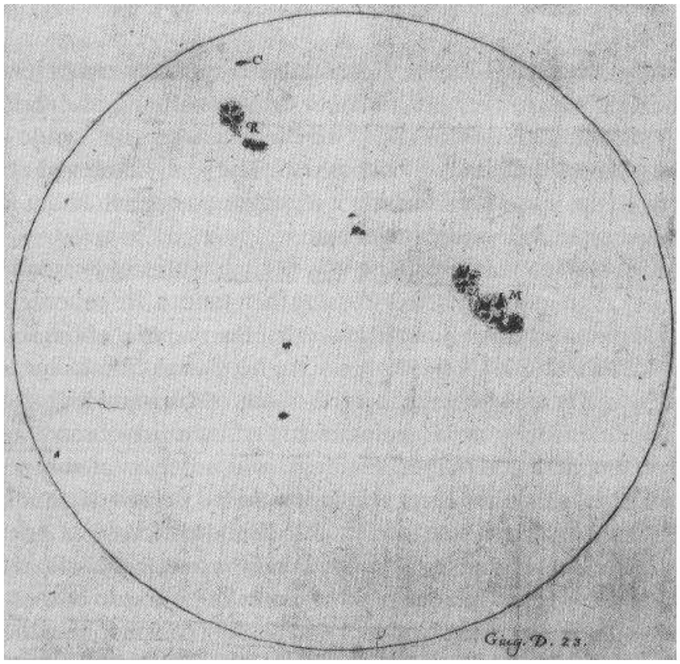

Figure 6.3: {`F1.2`} Reproduction of one of Galileo's sunspot drawings for 23 June, 1612. The umbrae/penumbrae structure is clearly visible here.

Heinrich Schwabe (1789-1875) set himself about the task of discovering intra-mercurial planets, whose existence had been conjectured for centuries. Like many before him, Schwabe realized that his best chances of detecting such planets lay with the observation of the apparent shadows that they would cast upon crossing the visible solar disk during conjunction; the primary difficulty with this research program was the ever-present danger of confusing such planets with small sunspots. Accordingly, Schwabe began recording very meticulously the position of any sunspot visible on the solar disk on any day that weather would permit solar observation. In 1843, after 17 years of observations, Schwabe had not found a single intra-mercurial planet, but had discovered something else of great importance: the cyclic increase and decrease with time of the *average* number of sunspot visible on the Sun, with a period that Schwabe originally estimated to be 10 years.

As Schwabe's discovery of the sunspot cycle gained recognition, the question immediately arose as to whether the cycle could be traced farther in the past on the basis of extant sunspot observations. In this endeavour the most active researcher was without doubt the Swiss astronomer Rudolf Wolf (1816-1893). Faced with the daunting task of comparing sunspot observations carried out by many different astronomers using various instruments and observing techniques, Wolf defined a **relative sunspot number** (r) as

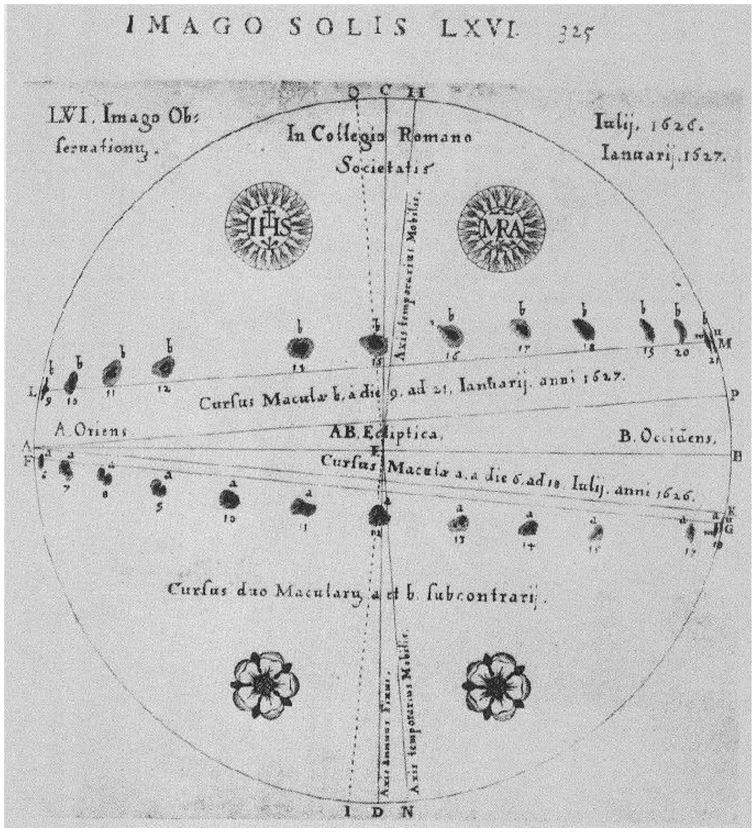

Figure 6.4: {F1.3} Solar rotation as inferred from the drift of sunspots. Diagram from Scheiner's *Rosa Ursina*, actually meant to illustrate the variations of the Sun's apparent rotation axis (as seen from Earth) in the course of the year.

follows:

$$r = k(f + 10g) \ , \tag{6.1}$$ {???}

where g is the number of sunspots groups visible on the solar disk, f is the number of individual sunspots (including those distinguishable within groups), and k is a correction factor that varies from one observer to the next (with $k = 1$ for Wolf's own observations, by definition). This definition is still in used today, but r is now usually called the Wolf (or Zürich) sunspot number. Wolf succeeded in reliably reconstructing the variations in sunspot number as far as the The 1755–1766 cycle, which has has since been known conventionally as "Cycle 1", with all subsequent cycles numbered consecutively thereafter; at this writing (August 2004), we are in the descending phase of cycle 23.

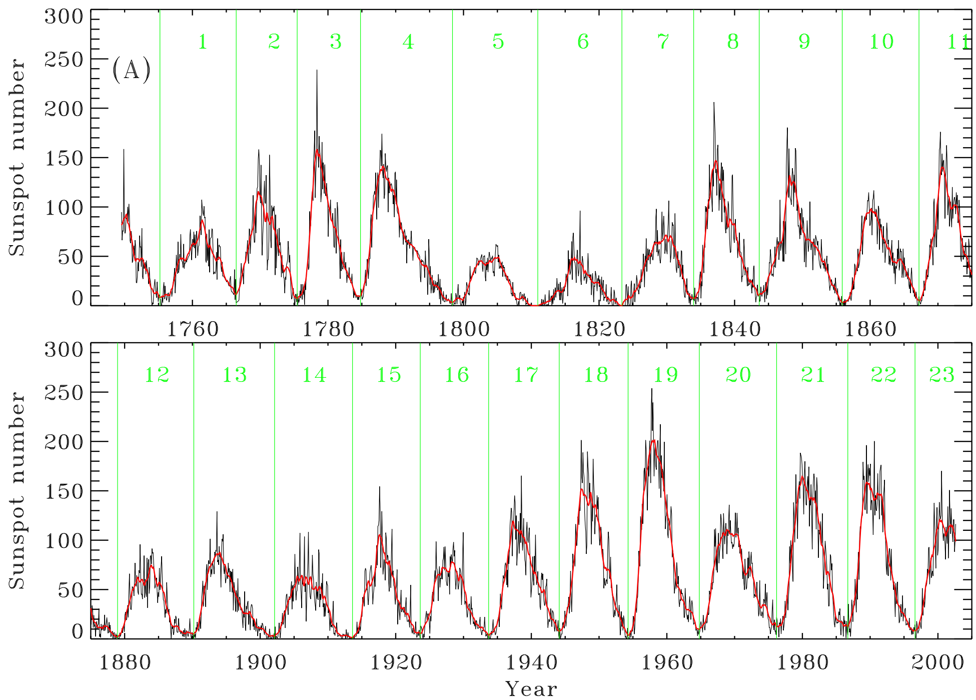

Figure 6.5: {F1.9} Two time series of the celebrated Wolf Sunspot Number. The thin black line is the monthly-averaged sunspot number, and the thick red line a 13-month running mean thereof. These and other related data are publicly available at the Solar International Data Center in Brussels, Belgium (http://sidc.oma.be).

Figure 6.5 shows two time series of the relative sunspot number. The first (thin black line) is the monthly-averaged value of r as a function of time, and the thick red line is a 13-month running mean of the same. Note how the amplitude, duration and even shape of sunspots cycles can vary substantially from one cycle to the next. The period, in particular, ranges from 9 (cycle 2) to 14 years (cycle 4), although it remains costumary to speak of the "11-year cycle".

6.1.3 The Waldmaier and Gnevyshev-Ohl Rules {ssec:wald}

Starting with Wolf himself, the sunspot number time series (monthly, monthly smoothed, yearly, etc) has been analyzed to death in every possible manner known to statistics, nonlinear dynamics, and numerology[1]. Many otherwise serious and respectable people engaged in this type of work seem to forget

[1]Two colleagues of mine, both world-renowned expertsi in the analysis of time series, have independently fremarked to me that the sunspot number time series are quite possibly

that the definition of the sunspot number is largely arbitrary, and its link to the real dynamical quantity, the sun's magnetic internal field, uncertain at best.

Of the various patterns uncovered in the sunspot number time series, some actually appear to robust, in that they do not depend too much on the manner the analysis is being carried out, and are also found in other indicators of solar activity; to the point in fact that they have been upraded to the status of empirical "Rule". We'll consider here only the two most convincing ones.

The Waldmaier rule refers to the fact that an anticorrelation seems to exist between cycle amplitude and duration. Starting for example from the time series of smoothed monthly sunspot number (red line on Figure 6.5, it is straightforward to assign to each cycle n a peak amplitude A_n and a duration T_n, the latter being simply the time interval between the two minima bracketing a given cycle. Figure 6.6A shows a correlation plot of these two quantities, which are characterized by a linear correlation coefficient of $r = -0.7$, which is definitely large enough to merit attention. The anticorrelation is intriguing, because one might have (naively) expected that high amplitude cycles should also last longer, but in fact the opposite seems to hold.

Another intriguing unspot cycle amplitude pattern is known as the Gnevyshev-Ohl rule, and is illustrated on Fig. 6.6B. Cycle amplitude A_n are plotted as solid dots, versus cycle number n. For reasons that will become clear shortly, odd-numbered cycles (according to Wolf's numbering convention) have been plotted in orange, and even-numbered cycles in red. Compute now a 1-2-1 running mean of cycle amplitude, i.e.,

$$< A_n >= \frac{1}{4}(A_{n-1} + 2A_n + A_{n+1}) \ , \qquad n = 2, 3, ... \qquad (6.2)$$ {eq:121}

The resulting time series for $< A_n >$ is plotted as a thick blue line on Fig. 6.6B; notice now how most odd-numbered cycles lie *above* the running mean curve, while even-numbered cycles usually lie below. In fact, from cycle 9 to 21 inclusive, the pattern has held true without interruption. As we will see in due time, both the Waldmaier and Gnevyshev Rules pose quite a challenge to most current solar dynamo models.

6.1.4 The butterfly diagram {ssec:butterfly}

To the striking cyclic pattern uncovered by Schwabe was soon added an equally striking *spatial* regularity. In 1858, G. Spörer (1822-1895) and R.C. Carrington (1826-1875) independently pointed out that sunspots are observed at

---

the "natural" time series having produced the largest number of research journal pages per byte of actual data!

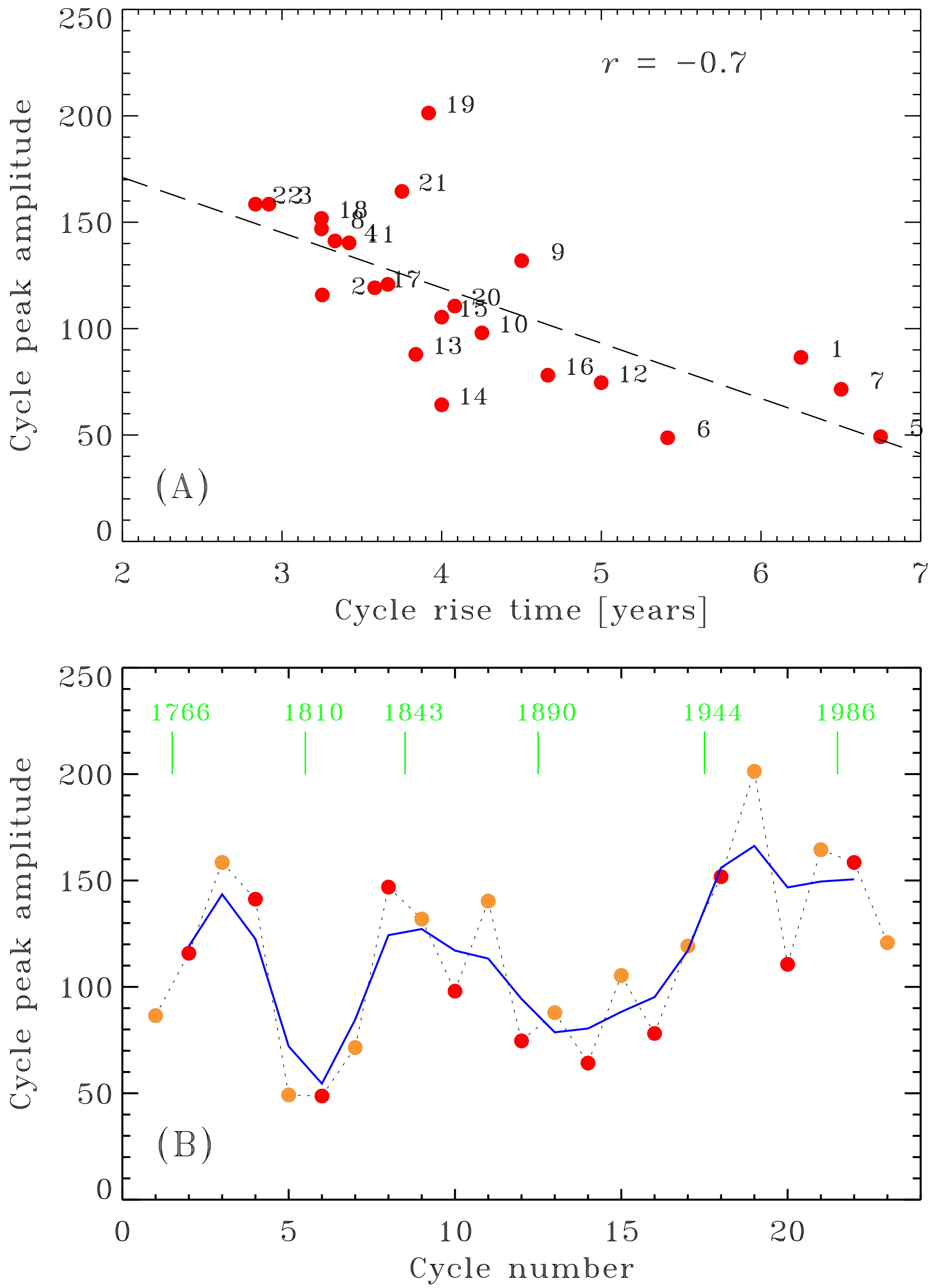

Figure 6.6: {`fig:wald`} (A) The anticorrelation between cycle rise time and amplitude, known as the Waldmaier Rule. A similar correlation, although weaker, characterizes cycle amplitudes and durations; (B) The Gnevyshev-Ohl Rule. Under Wolf's numbering convention, the odd-numbered cycles (orange dots) are more often found above the running mean (blue line) than even-numbered cycles (red dots), a pattern that held true uninterrupted from cycle 9 to 21 inclusively.

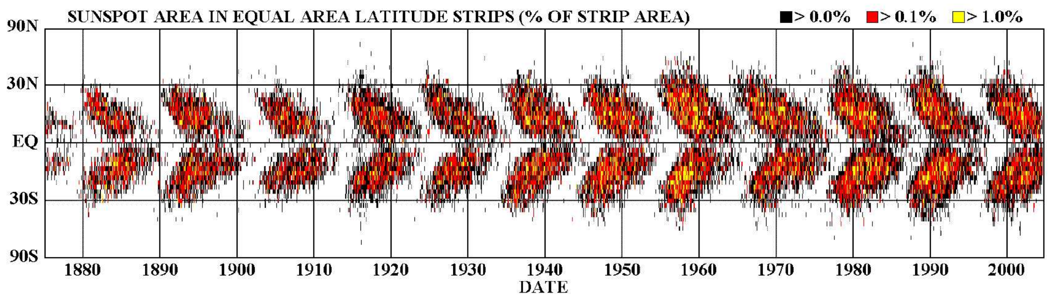

Figure 6.7: {F1.7} A sunspot butterfly diagram, showing the equatorward migration of sunspot latitudes in the course of each cycle. The sunspot number peaks about midway though the equatorward migration Data and graphics courtesy of David Hathaway, NASA/MSFC.

relatively high ($\sim 40°$) heliocentric latitudes at the beginning of a sunspot cycle, but are seen at lower and lower latitudes as the cycle proceeds, until at the end of the cycle they are seen mostly near the equator, at which time spots announcing the onset of the next cycle begin to appear again at $\sim 40°$ latitude. This is illustrated on Figure 6.7, in the form of a **butterfly diagram** for the time period 1875—2003. The construction of sunspot butterfly diagrams was first carried out by the husbnd-and-iwfe team of Annie and E. Walter Maunder in 1904, and proceeds as follows: one begins by laying a coordinate grid on, for example, a solar white light or calcium image, with, as in the case of geographic coordinates on Earth, the rotation axis defining the North—South vector. The visible solar disk is then divided in latitudinal strips of constant projected area, and for each such strip the percentage of the area covered by sunspots and/or active regions is calculated and color coded. This defines a one-dimensional (vertical) array describing the average sunspot coverage at one time. By repeating this procedure at constant time intervals and stacking the arrays one besides the other, one obtains a two-dimensional image of average sunspot coverage as a function of heliospheric latitude (vertical axis) and time (horizontal axis).

The absence of sunspots at high latitudes ($\gtrsim 40°$) at any time during the cycle, and the equatorward drift of the sunspot distribution as the cycle proceeds from maximum to minimum, are both particularly striking on such a diagram. Note how the latitudinal distribution of sunspots is never exactly the same, and how for certain cycles (for example the 1965—1976 cycle) there exists a pronounced North–South asymmetry in the hemispheric distributions. Note also how, at solar minima, spots from each new cycle begin to appear at mid-latitudes while spots from the preceding cycle can still be seen near the equator, and how sunspots are almost never observed within a few degrees in latitude of the equator. Sunspot maximum (1991, 1980, 1969,...) occurs about midway along each butterfly, when sunspot coverage is maximal at about 15 degrees latitude.

6.1.5 Hale's polarity laws {ssec:hale}

The study of sunspots and their 11-year cycle was finally put on a firm physical footing by the epoch-making work of George Ellery Hale (1868-1938) In the decade following their groundbreaking discovery of sunspot magnetic fields, Hale and collaborators went on to establish what are now known as **Hale's polarity laws**:

1. At any given time, the polarities of the leading spots of sunspot pairs are the same in a given solar hemisphere;

2. At any given time, the polarities of the leading spots of sunspot pairs are opposite in the N and S hemispheres;

3. Sunspot polarities reverse in each hemisphere from one 11-yr sunspot cycle to the next;

(see Fig. 6.8). The most straightforward interpretation of this common opposite polarity grouping is that we are seeing the surface manifestation of a large-scale **toroidal field** residing somewhere below the photosphere, having risen upwards and pierced the photosphere in the form of a so-called "Ω–loop" (see Figure 6.9).

Because the flux rope can be expected to expand as it rises buoyantly through the convective envelope, neither the size or magnetic field strength of sunspots can be assumed to be identical to that of the underlying toroidal flux ropes. However, if the rope maintains its cohesion throughout the rise and emergence processes then its **magnetic flux** is a conserved quantity:

$$\Phi_B = \int B_\phi \cdot \mathrm{d}S \tag{6.3}$$ {???}

Observations indicate that for sunspots $\Phi_B \sim 10^{21}$—10^{23} Mx (Mx≡G cm^2), with 10^{22} Mx a representative value for a "typical" sunspot.

Hale's polarity rules, interpreted in terms of this "model" of sunspots, imply that the toroidal component of the solar internal magnetic field is antisymmetric about the equator[2], and evolves cyclically on a $\simeq$ 22 year timescale. So, from a physical —rather than botanical— standpoint, the

[2]which, as we shall see in chapter 2 below, is what one would expect from the kinematic shearing of a dipolar magnetic field by axisymmetric differential rotation.

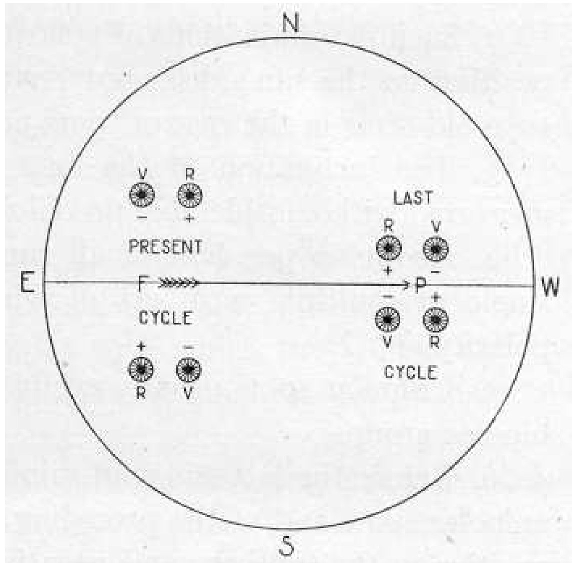

Figure 6.8: {F1.11} A diagram taken from the 1919 paper by G.E. Hale, F. Ellerman, S.B. Nicholson, and A.H. Joy (in *The Astrophysical Journal*, vol. **49**, pps. 153-178), illustrating Hale's polarity laws. This presented solid evidence for the existence of a well-organized large-scale magnetic field in the solar interior, which cyclically changes polarity approximately every 11 years.

true length of the solar cycle is not 11 years, but rather 22 years. Yet astronomers are creatures of tradition, and solar astronomers are no exception; nearly a century after Hale's discovery of the sunspot polarity law, it remains customary to speak of the "11 year solar cycle".

Hale and collaborators also showed that the line segment joining two members of a sunspot pair shows a systematic tilt angle ($\tilde{\theta}$) with respect to the East-West direction, the sunspot farther ahead (in the direction of solar rotation) being closer to the equator. Although there exists considerable variations in observed tilt angles, statistically the magnitude of the tilt increases with increasing heliocentric latitude (see Fig. 6.10). This is known as Joy's Law, after the poor bastard who was told by his boss G.E. Hale to go back and measure the tilts of all sunspot pairs to be found on the sunspot drawings of R.C. Carrington and G. Spörer.

Least-squares fits to observations yield a parametric representation of the form:

$$\sin\tilde{\theta} = 0.48\cos(\theta) + 0.03 \tag{6.4}$$ {???}

where θ is the colatitude angle. This pattern plays an important role in some of the solar cycle models to be considered in later chapters. This is because

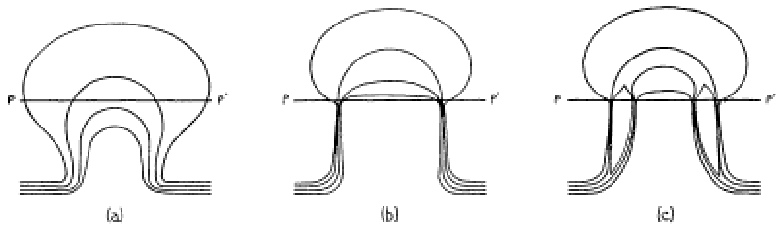

Figure 6.9: {F1.11b} Schematic representation of a sunspot pair as the manifestation of an underlying toroidal flux rope having risen through the photosphere. The magnetic fields impedes convective energy transport, so that cooling leads to a collapse of the magnetic field into two sunspots of opposite polarities. Reproduced from the 1955 paper by E.N. Parker's in *The Astrophysical Journal*, vol. **121**, pps. 491-507 [Figure 2, p. 496].

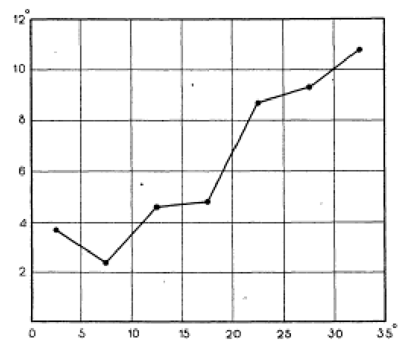

Figure 6.10: {F1.11c} Variation of the average tilt angle $\tilde{\theta}$ (ordinate) with respect to heliocentric latitude (abcissa, both in degrees). Reproduced from Hale, Ellerman, Nicholson & Joy 1919, *Astrophys. J.*, **49**, 153 [Figure 5, p. 168].

the existence of a finite, systematic tilt implies a net dipole moment, which can contribute to the net solar **poloidal field**.

All these regularities carry a very important message; no matter how vigorous convective fluid motions may be in the convective envelope, they are not vigorous enough to completely disrupt the solar internal toroidal magnetic field.

6.1.6 Modeling the buoyant rise of magnetic flux ropes {ssec:risetime}

In translating the cartoon of Figure 6.9 into a quantitative physical model, we have a number of issues that need to be clarified. Perhaps the most pressing are: (1) to identify the region(s) of the solar interior from which the flux ropes originate, and (2) to estimate the time required for a magnetic flux tube to rise through the convective envelope[3]. It turns out that both questions are very much related.

Consider a toroidal flux tube of diameter a and mean field strength B, embedded in a convective envelope with pressure and density profiles $p(r), \rho(r)$ and scale height $h = kT/\mu m_p g$, where g is the gravitational acceleration and $a \ll h$. Lateral pressure equilibrium demands that

$$p(r) = p_i(r) + \frac{B^2}{8\pi} \ , \qquad (6.5) \quad \{\texttt{E1.12}\}$$

where ρ_i is the mean density within the tube, and the second term on the RHS is the **magnetic pressure**. Clearly eq. (6.5) can only be satisfied provided that $\rho_i < \rho(r)$, i.e., the flux tube is evacuated. Assume now that the temperature is the same inside and outside the tube; this implies $\rho/\rho_i = p_i/p$, so that

$$\rho(r) - \rho_i(r) = \frac{\rho(r)B^2(r)}{8\pi p(r)} \ . \qquad (6.6) \quad \{\texttt{E1.13}\}$$

The (radial) buoyancy force per unit length along the tube is then

$$F = \pi a^2 g(\rho - \rho_i) \ . \qquad (6.7) \quad \{\texttt{E1.14}\}$$

As a consequence of this buoyancy, the tube is accelerated upwards and begins to rise towards the surface. If thermal equilibrium is maintained between the tube and its surroundings, the only force left to equilibrate the buoyancy force is the aerodynamic drag:

$$F_D = \frac{C_D}{2}\rho u^2 a \ , \qquad (6.8) \quad \{\texttt{E1.15}\}$$

where u is the rise velocity of the tube, and the coefficient of aerodynamic drag C_D is a number of order unity for low viscosity subsonic flows. Equating eqs. (6.7) and (6.8) yields, after making use of the definition for the scale height h, the following expression for the terminal rise velocity:

$$u^2 = \frac{B^2}{4\pi\rho}\left(\frac{\pi a}{C_D h}\right) \ . \qquad (6.9) \quad \{\texttt{E1.16}\}$$

The rise time τ for a magnetic flux tube starting at a depth r_0 within the envelope is then approximately

$$\tau \simeq \frac{R_\odot - r_0}{u} \ . \qquad (6.10) \quad \{\texttt{E1.17}\}$$

[3]What follows is largely inspired from the 1975 paper by E.N. Parker cited in the bibliography at the end of this chapter

If you start plugging in numbers in the above expressions (you get to do just that in problem 1.3 below!), you soon come to the conclusion that for flux ropes of strengths $\gtrsim 10^3\,$G and magnetic fluxes $\gtrsim 10^{21}\,$Mx released at $r/R_\odot = 0.8$, the rise time is well under one year. More elaborate calculations, taking into account heat exchange between the tube and its surroundings as well as the slightly superadiabatic stratification of the envelope yield even shorter rise times. As we shall see in later chapter, amplification of the solar magnetic field by the dynamo process requires that the field remains in its generating region for a few years before sufficiently high field strengths are produced. This has led to the conclusion that the solar magnetic field is stored —maybe even produced— not in the convective envelope proper, but rather immediately below it.

The issue of *storage* of the magnetic flux ropes below the core-envelope interface is far from trivial. Basically, the flux ropes are subject to non-axisymmetric instabilities that take the form of growing waves of low azimuthal wavenumbers. The growth time for the instability turns out to depend rather sensitively on the thermodynamic structure (namely, the degree of subadiabaticity) of the storage layers. The bibliography at the end of this chapter lists a few good papers concerned with such stability analyses. Such calculations indicate that magnetic flux tubes of strength 60—160 kG can be stored immediately beneath the core-envelope interface for time periods of a few years, with the growth time for the instability decreasing rapidly with increasing field strength.

Considerable efforts have also gone into making more realistic models of the rise of thin magnetic flux tubes through the solar convective envelope. Such models include all kinds of reasonable things like rotation, non-axisymmetric perturbations, storage below the core-envelope interface, etc. While this represents a considerable improvement, mathematically flux tubes are still treated as structureless, flux-carrying material lines and so these kinds of calculations cannot properly take into account the interaction of the tube with the surrounding turbulent fluid motions. With this caveat in mind, thin flux tube modeling has produced the following two important results:

1. The flux ropes rise essentially radially if they have a field strength $B \gtrsim 60$—$100\,$kG; otherwise the Coriolis force deflects the rising flux tubes to high latitudes.

2. The flux ropes emerge without any tilt for $B \gtrsim 10^6\,$G, and with tilts compatible with Joy's Law for fields strengths in the range 60—160 kG.

Now, this is *great* stuff: the observed emergence of sunspots at low heliocentric latitudes puts a *lower* limit on the strength of the participating flux ropes; Joy's Law, on the other hand, translate into an *upper* limit on the field strength. One concludes that the toroidal flux ropes must have magnetic field strengths in the rather narrow range

$$60 \lesssim B \lesssim 160\,\mathrm{kG}\ . \tag{6.11}$$ {???}

The basic physical mechanism underlying these two remarkable results is the same: if the rise time of the flux ropes is much smaller than the solar rotation period, the Coriolis force has a strong influence. It is the Coriolis force that deflects the rising flux ropes to high latitudes, and gives rise to the twist that, upon emergence, manifests itself as Joy's Law. If the field is strong enough for the rise time to be much shorter than the rotation period, then the rising flux rope does not "feel" the rotation, rises radially, and emerges without a tilt.

6.1.7 Poloidal field reversals

While the surface magnetic field on Fig. 2.4 may look like a total mess, on long timescales a well-defined spatiotemporal pattern once again emerges. Figure 6.11 is once again a synoptic (time-latitude) diagram of the radial magnetic field component (averaged in longitude on the visible disk) covering three sunspot cycles. Superimposed on the diagram are vertical line segments indicating latitudes where sunspots are observed. New poloidal field first shows up at mid-latitudes (e.g., 1977), a year or two after the new cycle sunspots have begun to appear at high latitudes, and then migrates to higher and possibly lower latitudes in the course of the cycle. The situation is greatly complicated by the active region fields, which make a very strong contribution to the line-of-sight magnetograms at low heliocentric latitudes. Furthermore, the tilt of active regions amounts to a net dipole moment, which is carried to higher latitudes by the poleward surface meridional flow (more on this later) following the decay of the active regions. This poleward transport is clearly visible on Fig. 6.11, in the form of elongated, inclined stripes extending from mid to high latitudes. Whether this transport of poloidal field contributes to —or even dominates— the evolution of the high latitude poloidal field remains an open question.

At high heliocentric latitude ($\gtrsim 50°$) there exist a cleaner pattern of polarity changes occurring on the solar cycle period. For example, polarity reversal occurs in 1980, at solar maximum. During the 1976—1986 cycle the toroidal field was negative in the N-hemisphere; taken at face value, Figure 6.11 then indicates that the high latitude poloidal field lags the toroidal field by a phase interval $\Delta\varphi \simeq \pi/2$.

A different observational tracer that yields similar results is the count of **polar faculae**, concentrated regions of relatively strong magnetic field often

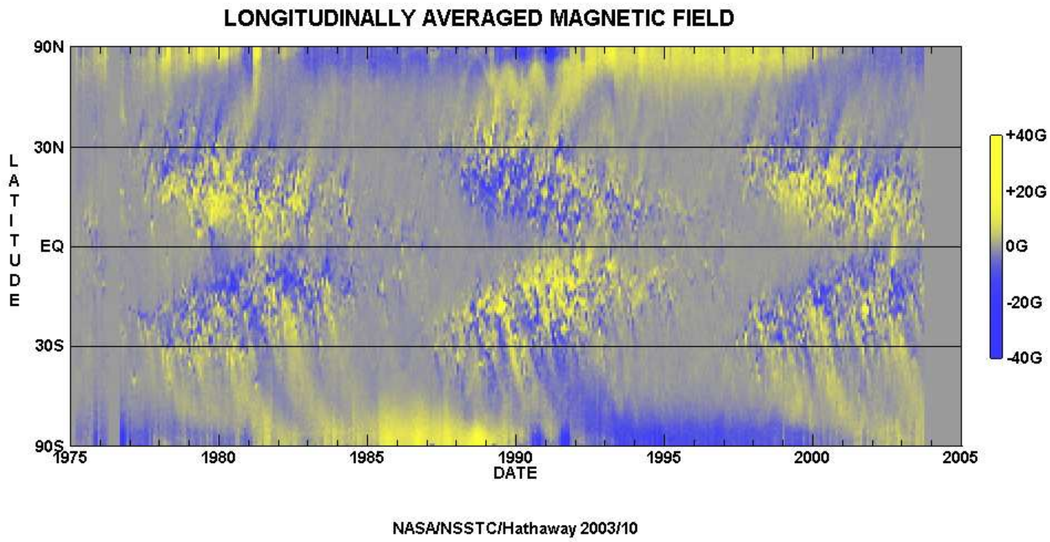

Figure 6.11: {F1.8} A synoptic magnetogram covering the last three sunspot cycles. The radial component is azimuthally averaged over a solar rotation, and the resulting latitudinal strips stacked one against the other for successive rotations. Data and graphics courtesy of David Hathaway, NASA/MSFC.

seen at high latitudes. Under the assumption that the structure of the faculae themselves is independent of the phase in the solar cycle, their number at any given time gives a measure of the overall poloidal field strength, once calibrated against magnetograms. Reliable polar faculae records exists for nearly one hundred years, allowing to reconstruct the solar cycle evolution of the large-scale solar poloidal field back to the beginning of the twentieth century.

6.1.8 The Maunder Minimum {ssec:maunder}

One final, peculiar feature associated with the solar cycle needs to be discussed, because of its implications for dynamo modelling. The historical reconstructions began by Wolf have been pushed as far back as the invention of the telescope in the opening decade of the seventeenth century, which marks the beginning of regular sunspot monitoring by astronomers. One such full reconstruction, starting in 1610, is shown on Figure 6.12 (bottom panel). While observations are a tad patchy from 1610 to 1640, coverage is actually quite good beyond this date. The lack of sunspots in the period 1645-1715 is therefore not due to lack of data, but represents a phase of strongly suppressed solar activity now known as the **Maunder Minimum**, after the solar astronomer E.W. Maunder, who, following the pioneering historical investigations of Gustav Spörer, was most active and steadfast in investigating

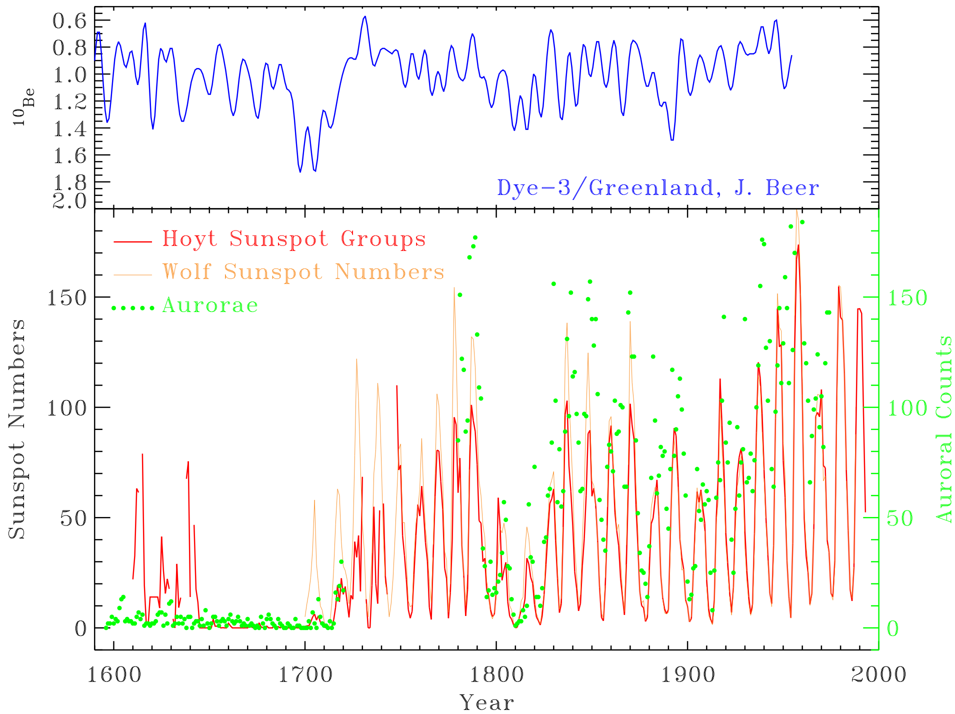

Figure 6.12: {F1.16} The Maunder minimum, as seen through cosmogenic radioisotopes (top panel) and sunspot and auroral counts (bottom panel). The thick red line is the so-called Group Sunspot Number, a reconstruction similar to Wolf's (thin orange line) but deemed more reliable in the eighteenth century because it relies exclusively on the more easily observable sunspot groups. Beryllium 10 data courtesy of J. Beer, EAWAG/Zürich.

the dearth of sunspot sightings by astronomers active in the second half of the seventeenth century. The documented occurrence of exceptionally cold winters throughout Europe during those years may be causally related to reduced solar activity, although this remains a topic of controversy.

That this is not just a matter of failing to form sunspots is confirmed by historical reconstructions of auroral counts, which are also strongly reduced during the Maunder Minimum (cf. Fig. 6.12). On the other hand, cosmogenic radioisotopes such as ^{10}Be, whose production frequency is known to be modulated by the frequency of solar eruptive phenomena, continue to show a cyclic pattern throughout the Maunder minimum (Fig. 6.12, top panel), indicating that the cycle had actually not come to a complete standstill.

The cosmogenic isotope record also indicated that episodes of markedly reduced solar activity occurred in 1282-1342 (Wolf minimum) and 1416-1534 (Spörer minimum), and that solar activity was significantly above its modern average in the time period 1100-1250 (dubbed Medieval Maximum by

Min/Max aficionados).
Solanki/Usoskin reconstruction

6.1.9 Cyclic modulation of solar activity {ssec:cycact}

The solar cycle also modulates the solar luminosity, the sun being about 0.15% brighter at sunspot maximum than at minimum (see Fig. 6.13, top panel). The emission of short-wavelength, non-thermal emission also varies in phase with the solar cycle; in the far-ultraviolet regions of the solar spectrum ($\lambda \lesssim 120\,\mathrm{nm}$), variations by $\sim 100\%$ are observed between solar minimum and maximum, with corresponding variations by over a factor of ten in the X-Ray domain. The sun's radio emission, indicative non-thermal acceleration of electrons in the lower corona, also follows the sunspot cycle quite closely (see Fig. 6.13, third panel).

6.1.10 Summary of solar cycle characteristics

For convenience, let's now collect a short list of fundamental observational features that a physical model of the solar magnetic cycle should reproduce (omitting for the time being anything related to amplitude fluctuation):

1. A large-scale magnetic field, axisymmetric to a good approximation and antisymmetric about the solar equator;
2. A cyclic variation of this large-scale magnetic field, characterized by polarity reversals with a $\sim 20\,\mathrm{yr}$ oscillation period;
3. An internal toroidal field of strength $\sim 10^4$—$10^5\,\mathrm{G}$, concentrated at low solar latitudes ($\lesssim 45°$, say), and migrating equatorward in the course of the cycle with minimal spatiotemporal overlap between successive cycles;
4. A large-scale surface poloidal field of a few tens of Gauss, migrating poleward in the course of the cycle, and reversing polarity at sunspot maximum.
5. Hemispheric helicity pattern ...

6.2 The astrophysical dynamo problem(s) {sec:dynprob}

Copper wires and sliding contacts being a rather sparse commodity in the universe, we must now figure out to apply the general idea of a dynamo

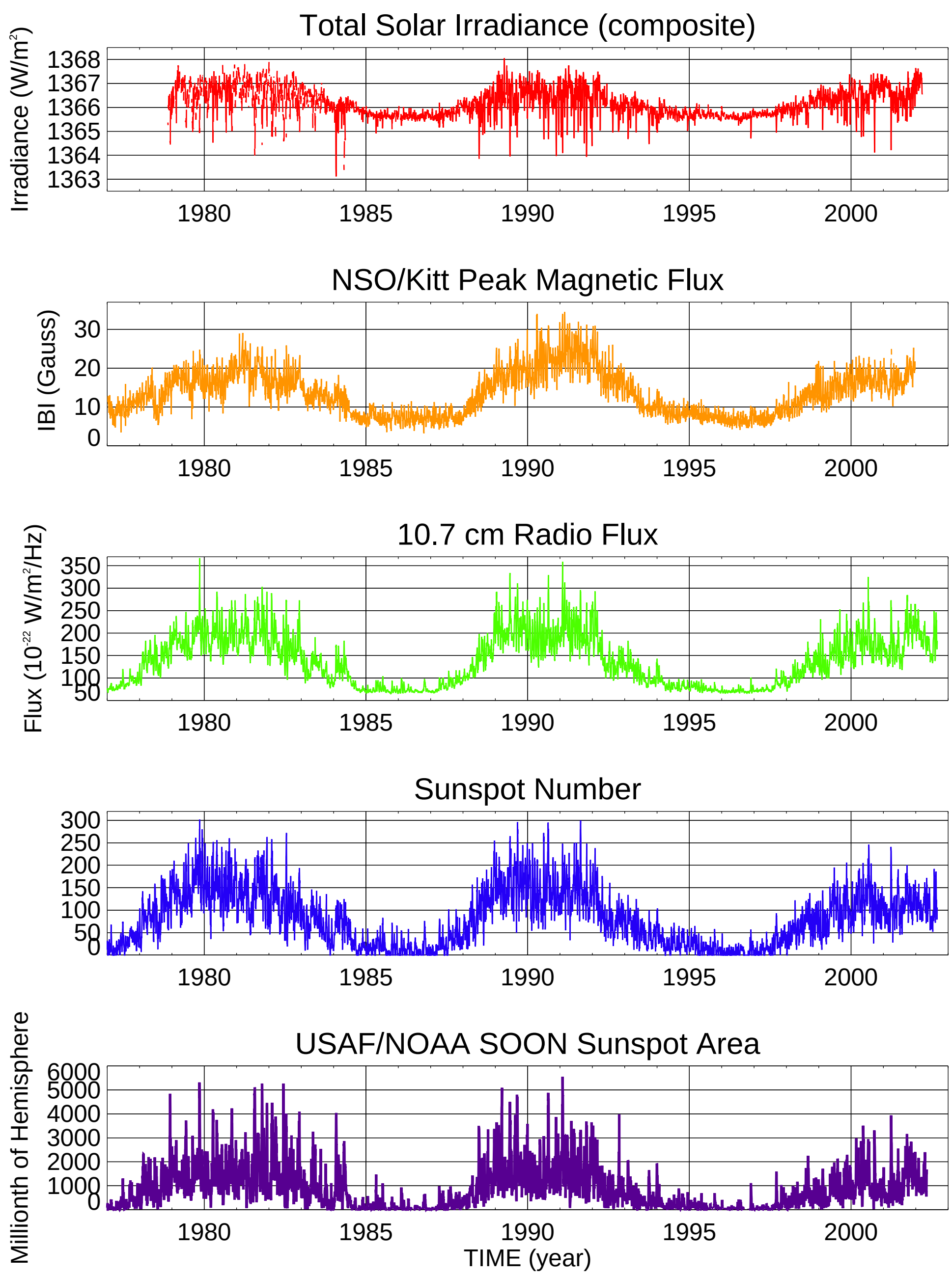

Figure 6.13: {`F1.15`} Variation of various solar activity indicators with the solar cycle. The NSO/Kitt Peak magnetic flux includes the contribution of magnetic fields outside of sunspots. The 10.7cm radio flux is a measure of non-thermal processes in the lower corona. Data and graphics courtesy of Giuliana DeToma, HAO/NCAR.

to astrophysical fluids. In the MHD limit, our hope lies evidently with the induction term $\nabla \times (\mathbf{u} \times \mathbf{B})$ in the induction equation

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{u} \times \mathbf{B}) - \nabla \times (\eta \nabla \times \mathbf{B}) \ , \tag{6.12}$$ {E2.01}

Remember that there are no true source term in eq. (6.12); if $\mathbf{B} = 0$ at some t_0, then $\mathbf{B} = 0$ for all subsequent $t > t_0$. We must therefore that some seed field exists to start up the dynamo process, just as in the homopolar dynamo we just looked into. As we saw in §2.3, there exist viable candidates to produce this seed field, most notably battery mechanism associated with mechanical separation of electric charges.

In its simplest form, the **dynamo problem** consists in finding a flow field $\mathbf{u}$ that can sustain a magnetic field against Ohmic dissipation. We must distinguish **kinematic dynamo**, where the flow field $\mathbf{u}$ is considered given *a priori* and constructed without any regards for its underlying dynamics, from what can only be called (for lack of a generally agreed-upon terminology) the **full dynamo problem**, in which the flow $\mathbf{u}$ results from a solution of the full set of MHD equations (§**??**), including the backreaction of the magnetic field on the flow via the Lorentz force term $\mathbf{J} \times \mathbf{B}$ on the RHS of the Navier-Stokes equation.

The kinematic regime carries the immense practical advantage that the induction equation then becomes truly linear in $\mathbf{B}$, and the dynamo problem reduces to finding a (smooth) flow field $\mathbf{u}$ that has the requisite topological properties to lead to field amplification. In the following chapters we will concentrate mostly on this kinematic regime, but will occasionally touch upon the much more difficult dynamical problem, mostly via direct numerical simulation of the full set of MHD equations.

As we'll see in the following chapter, there are flows that can amplify a magnetic field during a transient time interval, after which $\mathbf{B}$ decays again. So we tighted our definition of the dynamo problem by demanding that a flow be a dynamo if it can lead to $\mathcal{E}_{\rm B} > 0$ for times much larger than all relevant advective and diffusive timescales of the problem. To make things even harder, we'll add the additional condition that no electromagnetic energy be supplied across the domain boundaries S, i.e., $\mathbf{S} \cdot \mathbf{n} = 0$ in eq. (1.76). It is readily shown that this latter condition is satisfied if either (1) $\mathbf{B} = 0$ on the boundary, or (2) the components normal to the boundaries of $\mathbf{U}$, $\mathbf{B}$, and $\mathbf{L}$ all vanish on S (do problem XXX!).

The **solar dynamo problem** can be tackled either in kinematic or fully dynamical form. The aim there is to reproduce observed spatiotemporal patterns of magnetic field evolution, a minimum list of features having already been listed at the end of §6.1. As will become obvious in the following chapter, even this basic short list is a pretty tall order. Yet, from solar irradiance variations and their possible influence on Earth's climate to space weather prediction, it all begins with the solar cycle. Keep this in mind as we now start to dig into the mathematical and physical intricacies of magnetic field generation in electrically conducting fluids. We'll seem to venture pretty far away from the sun and stars at times, but stick to it and you'll see it all fitting together at the end. And now, into the abyss...

Problems:

1. Estimate the solar rotation period from the apparent motion on the sunspots drawing reproduced on Fig. 6.4. What are the primary difficulties in carrying out this kind of analysis?

2. The sunspot number time series reproduced on Fig. 6.5 are almost certainly the most intensively studied time series in All Of Astrophysics, as measured by the number of published research papers per data point. So you need to try your hand at crunching it a little bit. First, go to the SIDC's Web Page:

   `http://sidc.oma.be`

   click on "Sunspot archive & graphics", and grab the dataset for the 13-month running mean of the monthly sunspot number (red line on Fig. 6.5 herein). Then,

   (a) Measure the duration of each cycle, and compute the mean sunspot cycle period;

   (b) Measure the peak and integrated (i.e., area-under-the-curve) cycle amplitude; do these two measures of cycle amplitude correlate well?

   (c) Measure the rise time, i.e., the time elapsed from start of a cycle to its peak. Does this correlate to anything you have extracted so far (cycle duration, amplitude, etc.)?

   (d) Do a lag analysis by looking for correlation between the amplitude of one cycle, and that of the preceeding cycle; that of two cycles ago; three cycles ago, etc. Do you find any significant correlation for some lag?

   (e) Finally, calculate a power spectrum of the time series. Do you find significant peaks at periods other than $\sim 11\,\mathrm{yr}$?

3. This problem has you quantify and reflect upon some of the statements made in §6.1.6.

(a) Fill in all missing mathematical steps leading to eq. (6.10).

(b) Compute and plot curves showing the variations of the rise time with assumed magnetic field strength, for four fixed values of the magnetic flux: $\log \Phi = 20$, 21, 22 and 23. In all cases you may assume that the participating flux tube are released at a depth $r_0/R_\odot = 0.80$ within the convective envelope, where the density and temperature assume values $\rho = 4\,\mathrm{g\ cm^{-3}}$ and $T = 5 \times 10^6\,\mathrm{K}$. Remember that fixing the magnetic flux implies a relationship between a and B.

(c) Make a list of all the assumptions having entered this little derivation; which are the most/least reasonable ones?

Bibliography:

Hale's original papers on sunspots are still well worth reading. The two key papers are:

Hale, G.E. 1908, "On the probable existence of a magnetic field in sunspots", *The Astrophysical Journal*, **28**, 315-343,

Hale, G.E., Ellerman, F., Nicholson, S.B., and Joy, A.H. 1919, *The Astrophysical Journal*, **49**, 153-178.

On the Maunder minimum, see

Eddy, J. A., 1976, *Science*, **192**, 1189-1202,

Eddy, J. A., 1983, *Solar Phys.*, **89**, 195-207,

Ribes, J. C., and Nesme-Ribes, E. 1993, *Astron. Ap.*, **276**, 549-563.

and on cosmogenic radioisotopes:

Beer, J. 2000, *Sp. Sci. Rev.*, **94**, 53-66.

The study of rising toroidal flux ropes a proxy for the emergence of the solar internal toroidal field in the form of sunspot pairs is a topic that has exploded in the past 15 years or so. Among the many noteworthy contributions in this field , we recommend the following as starting points:

Moreno-Insertis, F. 1986, A&A 166, 291,

Choudhuri, A.R., & Gilman, P.A. 1987, *Astrophys. J.*, **316**, 788,

Fan, Y., Fisher, G.W., & DeLuca, E.E. 1993, *Astrophys. J.*, **405**, 390,

D'Silva, S., & Choudhuri, A.R. 1993, A&A 272, 621,

Caligari, P., Moreno-Insertis, F., & Schüssler, M. 1995, *Astrophys. J.*, **441**, 886.

Important earlier papers on the topic are:

Parker, E.N. 1955, *Astrophys. J.*, **122**, 293,
Parker, E.N. 1975, *Astrophys. J.*, **198**, 205,
Schüssler, M. 1977, A&A 56, 439,
Moreno-Insertis, F. 1983, A&A 122, 241,

The thin flux tube approximation used in most of these calculations is usually credited to

Spruit, H.C. 1981, A&A 98, 155.

Considerable effort is currently being put into doing away with the thin flux tube approximation, in order to see which of the above results remains robust once the flux tube is no longer treated as a one-dimensional object. This is a rapidly moving field, so for the latest see the following recent on-line review by Yuhong Fan:

`http://www.livingreviews.org/lrsp-2004-1`.

as well as these now-classics:

Longcope, D.W., Fisher, G.W., & Arendt, S. 1996, *Astrophys. J.*, **464**, 999,
Emonet, T., & Moreno-Insertis, F. 1998, *Astrophys. J.*, **492**, 804,
Fan, Y., Zweibel, E., and Lantz, S.R. 1998, *Astrophys. J.*, **502**, 968.

On the storage and stability of toroidal flux ropes below the solar convective envelope, see

Ferriz-Mas, A., & Schüssler, M. 1994, *Astrophys. J.*, **433**, 852,
Ferriz-Mas, A. 1996, *Astrophys. J.*, **458**, 802.

The amount and variety of solar data (numbers, images, movies) available online is quite simply staggering. Use your Web-surfing skills to locate the Web sites of NASA's Marshall Space Flight Center, of the Solar Stanford Center, of the High Altitude Observatory, of NASA's Goddard Space Flight Center, of SOHO and TRACE, to name but a few.

Chapter 7

Decay and Amplification of Magnetic Fields {chap:evolB}

It's not whether a thing is hard to understand.
It's whether, once understood, it makes any sense.

Hans Zinsser
Rats, Lice and History (1934)

We now begin our long journey towards astrophysical dynamos. It is a road long and hard and, n'en déplaise à Nick Cave, we would like to avoid too many falling by the side. Consequently this chapter will for the most part concentrate on a series (relatively) simple model problems illustrating the myriad of manners in which a flow and a magnetic field can interact. We will first consider the purely resistive decay of magnetic fields (§7.1), then examine various circumstances under which stretching by a flow can amplify a magnetic field (§7.2), and then examine some important subtleties of this process in the context of some (relatively) simple 2D and 3D flows (§§7.2 and 7.3). The chapter close with some so-called anti-dynamo theorems (§7.4), which will shed light on results from previous sections and indicate the way towards true magnetohydrodynamical dynamo action, which, I may as well admit it at the onset, we will first encounter only in the next chapter. Some of the material contained in this chapter may feel pretty far remote from the realm of astrophysics at times, but please do stick to it because the physical insight (hopefully) developed in the following sections will prove essential to pretty much everything that will come next.

7.1 Resistive decays of magnetic fields {sec:bdifuz}

Before we try to come up with flows leading to field amplification and dynamo action, we better understand the enemy, namely magnetic field decay by

Ohmic dissipation. Consequently, we first consider the evolution of magnetic fields in a conducting fluid, in the absence of any fluid motion (or, more generally, in the $\mathrm{R}_m \ll 1$ limit). The induction equation then reduces to

$$\frac{\partial \mathbf{B}}{\partial t} = -\nabla \times (\eta \nabla \times \mathbf{B}) = \eta \nabla^2 \mathbf{B} - (\nabla \eta) \times (\nabla \times \mathbf{B}) \ . \qquad (7.1) \quad \{\texttt{E2.12}\}$$

Were it not that we are dealing here with a vector —as opposed to scalar— quantity, for constant η this would look just like a simple heat diffusion equation, with η playing the role of thermal diffusivity. Before attempting to formally solve this equation, let us first obtain an order-of-magnitude estimate for the timescale τ_η over which a magnetic field $\mathbf{B}$ with typical length scale ℓ can be expected to decay. Replacing $\partial/\partial t$ by and $1/\tau_\eta$ and ∇^2 by $1/\ell^2$, eq. (7.1) yields the **diffusion time**:

$$\tau_\eta \sim \frac{\ell^2}{\eta} \ . \qquad (7.2) \quad \{\texttt{E2.13}\}$$

Now, for conditions typical of the solar interior we have $\eta \sim 10^3\,\mathrm{cm}^2\,\mathrm{s}^{-1}$, so that the diffusive time scale for a large-scale field pervading the solar interior ($\ell \sim R_\odot \simeq 7 \times 10^{10}$ cm) is $\tau \sim 10^{11}\,$yr, i.e., longer than the main-sequence lifetime of the Sun! Note that this is due primarily to the large spatial scale of the system, as opposed to an exceedingly low diffusivity; the solar interior is a much better electrical conductor than pure copper at room temperature ($\eta \simeq 10^5\,\mathrm{cm}^2\,\mathrm{s}^{-1}$, so that a magnetic field within a one meter wide sphere of copper would diffusively decay on a timescale $\tau \sim 0.1\,$s!). The *existence* of a solar magnetic field is then not really surprising; any large-scale fossil field present in the Sun's interior upon its arrival on the ZAMS would still be there today at almost its initial strength. The challenge in modeling the solar magnetic field is to reproduce the peculiarities of its spatial and temporal variations, in particular the cyclic variation of its large-scale component on a $\sim 22\,$yr timescale. But we are getting ahead of ourselves here. Back to simple resistive decay.

7.1.1 Reformulation as an eigenvalue problem

Let us now seek specific solutions for a few situations of solar interest, and (hopefully) verify our estimate of 10^{11} yr for the decay time of a fossil solar magnetic field. We are free to work directly with the magnetic induction equation for $\mathbf{B}$ (eq. (6.12)), or the "uncurled" equation for the vector potential $\mathbf{A}$ (eq. (1.91)). Choosing here the latter route, the magnetic and electric fields are obtained from the relations

$$\mathbf{B} = \nabla \times \mathbf{A} \ , \qquad \mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{A}}{\partial t} \ . \qquad (7.3) \quad \{\texttt{???}\}$$

The first point to notice is that the coefficients that appear in eq. (1.3) are independent of time, and so it is profitable to seek a separable solution of the form,

$$\mathbf{A} = e^{\lambda t}\mathbf{A}_\lambda(\mathbf{r}) \ , \quad \Phi = e^{\lambda t}\Phi_\lambda(\mathbf{r}) \ . \tag{7.4}$$ {???}

The decay rate, λ, is then determined by the eigenvalue problem,

$$\lambda\mathbf{A}_\lambda + \eta\nabla\times(\nabla\times\mathbf{A}_\lambda) = c\nabla\Phi_\lambda \ , \tag{7.5}$$ {E2.tom1}

along with some appropriate boundary conditions that we shall presently get to. We are still carrying the electrostatic potential Φ along just to keep matters as general as possible, but we shall make every effort to rid ourselves of this encumberance as soon as the opportunity presents itself.

The LHS of eq. (7.5) is the vector-Helmholtz equation which arises routinely in the description of electromagnetic wave propagation problems.[1] Therefore we should take advantage of the hard work others have done in order to make our present task easy. The elegant way to proceed is to define three vector operators which act upon scalar functions of $\mathbf{r}$ according to the prescriptions,

$$\mathbf{T} = -\hat{\mathbf{e}}_r\times\nabla \ , \qquad \mathbf{P} = -\nabla\times(\mathbf{r}\times\nabla) \ , \qquad \mathbf{L} = \nabla \ , \tag{7.6}$$ {???}

and generate *toroidal*, *poloidal*, and *longitudinal* vector fields, respectively.[2] We now construct $\mathbf{A}$ from these operators and three scalar functions according to,

$$\mathbf{A}_\lambda = r\mathbf{T}[\alpha_\lambda] + \mathbf{P}[\beta_\lambda] + \mathbf{L}[\gamma_\lambda] \ . \tag{7.7}$$ {???}

The benefit of all this is that the three vector operators have very nice transformation properties under the action of the curl operator,

$$\nabla\times r\mathbf{T} = \mathbf{P} \ , \qquad \nabla\times\mathbf{P} = -r\mathbf{T}\nabla^2 \ , \qquad \nabla\times\mathbf{L} = 0 \ , \tag{7.8}$$ {???}

where $\nabla\cdot\mathbf{L} = \nabla^2$ is the Laplacian, and $\nabla\cdot\mathbf{T} = \nabla\cdot\mathbf{P} = 0$.

It is now straightforward bookkeeping to substitute this representation for $\mathbf{A}_\lambda$ into eq. (7.5), collect similar looking terms, and arrive at the following set of uncoupled equations,

$$\lambda\alpha_\lambda = \nabla^2\alpha_\lambda \ , \tag{7.9}$$ {???}

$$\lambda\beta_\lambda = \nabla^2\beta_\lambda \ , \tag{7.10}$$ {???}

$$\lambda\gamma_\lambda = c\Phi_\lambda \ , \tag{7.11}$$ {???}

[1] Verify that Maxwell's equations in vacuum reduce to eq. (7.5) with $\eta = 1$ and $\lambda = -\omega^2/c^2$.

[2] The prescription presented here is for spherical coordinates. For other coordinate systems one replaces $\mathbf{r}$ and $\hat{\mathbf{e}}_r$ by the relevant vectors.

provided η is at worst only function of the radius r. The first two of these expressions are identical to the *scalar* Helmholtz equation encountered in the study of stellar oscillations. We recall that the spherical harmonics are the canonical angular functions that span the surface of a sphere. And so we may write either α_λ or β_λ as the product,

$$f_\lambda(r) Y_{lm}(\Omega) \; , \tag{7.12}$$ {E2.tom4}

for any non-negative integer l. The remaining unknown function and the much-anticipated eigenvalue λ are determined by the resulting ODE,

$$\Big[\frac{1}{r^2}\frac{\mathrm{d}}{\mathrm{d}r} r^2 \frac{\mathrm{d}}{\mathrm{d}r} - \frac{l(l+1)}{r^2} + \frac{\lambda}{\eta(r)}\Big] f_\lambda(r) = 0 \; . \tag{7.13}$$ {E2.tom2}

By virtue of the second term on the LHS of this equation, $r = 0$ is a singular point of this ODE and accordingly the non-analytic of the two linearly-independent solutions about this point must be discarded to maintain a sensible physical solution. The freedom to choose λ is necessary to force the remaining analytic solution to satisfy a prescribed boundary condition at the surface of the star ($r = R$). The nature of this boundary condition depends sensitively on the vector character of the decaying magnetic field.

7.1.2 Poloidal field decay

A poloidal magnetic field is generated by the $\alpha_\lambda(\mathbf{r})$ function. Hence, if we set $\beta_\lambda = 0$ we obtain,

$$\mathbf{B}_\lambda = \mathbf{P}[\alpha_\lambda] \; , \mathbf{E}_\lambda = -\frac{\lambda}{c}\Big\{ r\mathbf{T}[\alpha_\lambda] + \mathbf{L}[\gamma_\lambda] \Big\} \; , \tag{7.14}$$ {???}

valid for $r \neq R$. In the vaccuum surrounding the star $\eta = \infty$ since no material currents are allowed to be present, and Maxwell's displacement current has also been neglected. In this region we have the familiar potential field with $\alpha_\lambda \propto (R/r)^{l+1} Y_{lm}(\Omega)$. Inside the star, $\eta \neq \infty$, and the radial dependence of α_λ follows from the eigenvalue ODE, eq. (7.13).

Examination of the components of the $\mathbf{P}$ operator indicates that α_λ must be continuous across the stellar surface, $r = R$, else $\mathbf{B}$ will not be defined there. This can be accommodated through the freedom to multiply the exterior potential field solution by an arbitrary constant. So λ is still undetermined.

The current density (and hence the electric field) are given by the curl of the magnetic field. For $\mathbf{E}$ to be well-defined on the surface $r = R$ as the appropriate limit of the interior and exterior solutions, $\mathbf{B}$ must be continuous across the stellar surface. Since both the interior and exterior solutions carry the common factor of $Y_{lm}(\Omega)$, this is achieved merely by having $\partial\alpha_\lambda/\partial r$ continuous across $r = R$. As λ is the only thing left at our disposal to make this happen, the eigenvalue, and the decay-rate, are thus so-determined.

To see how this plays out, assume $\eta = \eta_0$ is constant throughout the interior of the star. The appropriate radial dependence within the star is describes by a spherical Bessel function, i.e.,

$$\alpha_\lambda = Y_{lm}(\Omega) j_l(kr) \qquad r < R \tag{7.15}$$ {???}

$$\alpha_\lambda = Y_{lm}(\Omega) j_l(kR) \Big(\frac{R}{r}\Big)^{l+1} \qquad r > R \tag{7.16}$$ {???}

where $k^2 \equiv \lambda/\eta_0$. The continuity of the radial derivative is assured if

$$kR j_l'(kR) + (l+1) j_l(kR) = kR j_{l-1}(kR) = 0 \ , \tag{7.17}$$ {???}

and so one need only hunt for the zeros of a spherical Bessel function in order to determine the decay rate of a poloidal magnetic field! An $l = 1$ dipole calls for the positive zeros of $j_0(x) = \sin x/x$. These are simply integer multiples of π, thus

$$\lambda_n = \frac{\eta_0 \pi^2 n^2}{R^2} \ \text{, for } l = 1 \ , n = 1, 2, 3, ... \tag{7.18}$$ {???}

Notice the many possible *overtones* associated with $n \geq 2$. These decay more rapidly than the fundamental ($n = 1$), since the radial eigenfunctions possess $n - 1$ field reversals. For such overtones, the effective length scale to be used in the decay-time estimate is roughly the radial distance between the field reversals, or $\approx R/n$.

Figure 7.1 (top row) shows the first three fundamental ($n = 1$) modes of angular degrees $l = 1, 2, 3$, corresponding to dipolar, quadrupolar, an hexapolar magnetic fields, as well as a few higher overtones for $l = 1, 2$ (bottom row). It is worth noting that the azimuthal quantum number, m, has no impact on computed decay rate. And last, but not least, the fossil field lifetime estimate provided by eq. (7.2) is just a little on the large side, by a factor of $\pi^2 \approx 10$, for a sun with constant diffusivity.

And what about γ_λ? Since everyone is continuous and well-defined there is no need for it, i.e., $\gamma_\lambda = \Phi_\lambda = 0$![3]

---

[3]In fact, eq. (7.13) is readily obtained by adopting the mixed poloidal/toroidal axisymmetric ($m = 0$) formulation of §1.10.3, and setting $B = 0$ and $A(r, \theta, t) = f_\lambda(r) Y_{l0}(\cos\theta) e^{\lambda t}$. But the formulation developed in this section remains of far greater applicability since it is not restricted to axisymmetric magnetic fields.

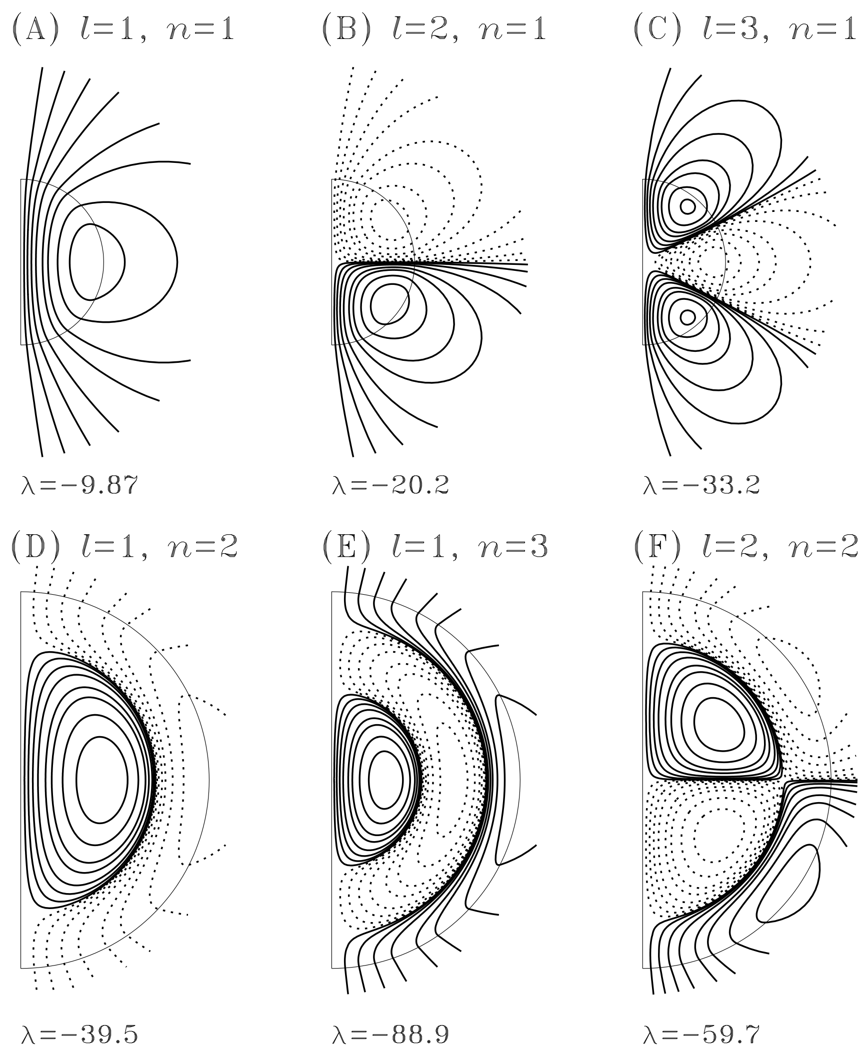

Figure 7.1: {F2.1} Six diffusive eigenmodes for a purely poloidal field pervading a sphere of constant magnetic diffusivity embedded in vacuum. The top row shows the three fundamental ($n = 1$) diffusive eigenmodes with smallest eigenvalues, i.e., largest decay times. They correspond to the well-known dipolar, quadrupolar, and hexapolar modes ($l = 1, 2$ and 3). The bottom row shows a few eigenmodes of higher radial overtones. Poloidal fieldlines are shown in a meridional plane, and the eigenvalues are given in units of the inverse diffusion time ($\tau^{-1} \sim \eta/R^2$).

7.1.3 Toroidal field decay

OK, now let's see how a toroidal magnetic field will decay. Now we can zero-out α_λ, giving

$$\mathbf{B}_\lambda = -r\mathbf{T}[\nabla^2\beta_\lambda]\ , \mathbf{E}_\lambda = -\frac{\lambda}{c}\Big\{\mathbf{P}[\beta_\lambda] + \mathbf{L}[\gamma_\lambda]\Big\}\ , \qquad (7.19) \quad \{???\}$$

again, valid for $r \neq R$. Everywhere *except* on the stellar surface, we can make good use of the fact that,

$$\nabla^2\beta_\lambda = \frac{\lambda}{\eta(r)}\beta_\lambda\ . \qquad (7.20) \quad \{???\}$$

In the surrounding vacuum, $\eta = \infty$, and so as before, $\beta_\lambda \propto (R/r)^{l+1}Y_{lm}(\Omega)$, for $r > R$. *However*, in this case, the consequence is that $\mathbf{B} = 0$ for $r > R$.

Now let's see what has to be continuous across the stellar surface. Since $\mathbf{B}$ is generated from the Laplacian of β_λ, for $\mathbf{B}$ to be well-defined on $r = R$ it is necessary that *both* β_λ *and* its radial derivative be continuous. Notice that this nets us *two* boundary conditions in one go. Of course, we must still be able to curl $\mathbf{B}$ safely to get $\mathbf{E}$. If the magnetic field vanishes in the vacuum, so too must it vanish at the stellar surface. This "third" boundary condition in fact determines the eigenvalue λ. Strictly speaking, this third requirement is

$$\lim_{r\to R^-}\nabla^2\beta_\lambda = 0\ . \qquad (7.21) \quad \{???\}$$

With these conditions in mind, we can set about the construction of β_λ. By keeping in mind that the magnetic field arises only from the toroidal vector field acting on the Laplacian of this function, and again specializing to the case of a constant diffusivity (η_0), by trial and error one (eventually) finds,

$$\beta_\lambda = Y_{lm}(\Omega)\Big[j_l(kr) - \frac{kR}{2l+1}j_l'(kR)\Big(\frac{r}{R}\Big)^l\Big] \qquad r < R \qquad (7.22) \quad \{???\}$$

$$\beta_\lambda = -Y_{lm}(\Omega)\frac{kR}{2l+1}j_l'(kR)\Big(\frac{r}{R}\Big)^{l+1} \qquad r > R \qquad (7.23) \quad \{???\}$$

where again, $k^2 = \lambda/\eta_0$. This convoluted result is worth an extra glance! The magnetic field is generated *only* by the term which carries the $j_l(kr)$. All the extra bits of potential field have been added to ensure the continuity of β_λ *and* its first radial derivative, *provided* that one carefully select k to satisfy the "third" boundary condition, viz.,

$$j_l(kR) = 0\ . \qquad (7.24) \quad \{???\}$$

In arranging all of these pieces to fall neatly into place, one discovers that β_λ does *not* vanish at the stellar surface. The decay rate λ^{-1} is again related to the zero of a spherical Bessel function—only of index l rather than $l-1$ as was found for the decay of the poloidal field. Hence, a dipole ($l=1$) toroidal magnetic field decays at precisely the same rate as a quadrupole ($l=2$) poloidal magnetic field (at least for constant diffusivity)! As before, the azimuthal quantum number m remains a non-issue. Looking up the expression for $j_1(x)$ in your favorite tome on special functions, the decay rate of a dipole toroidal field follows from the transcendental equation,

$$\tan kR = kR\ . \tag{7.25}$$ {???}

The smallest non-zero solution of this equation gives,

$$\lambda_1 = \frac{\eta_0(4.493409...)^2}{R^2}\ , l=1 \text{ toroidal and } l=2 \text{ poloidal.} \tag{7.26}$$ {???}

What about γ_λ? Well, in this case, it is *very* necessary! While we managed to get a continuous $\mathbf{B}$ to curl by adding in potential field contributions, the price is an additional electric field through the aegis of $\mathbf{P}[\beta_\lambda]$ contribution. The curl of $\mathbf{B}$ will *not* be proportional to $\mathbf{E}$ *unless* a non-trivial γ_λ is chosen to cancel the offending potential field contributions from the elaborately constructed β_λ. Luckily, and in fact by design, this action has no further impact on the magnetic field, and so the correct physical solution is finally arrived at.

It is worth a final remark to point out that the correct decay rate for a toroidal field would have been obtained merely from solving eq. (7.1) in the mixed representation by setting $A=0$ and forcing B to vanish at $r \gtrsim R$. But then, think of all the subtleties and fun that would have been missed!

7.1.4 Results for a magnetic diffusivity varying with depth {ssec:etavar}

We end this section by a brief examination of the diffusive decay of large-scale poloidal magnetic fields in the solar interior. The primary complication centers on the magnetic diffusivity, which is no longer constant throughout the domain, and turns out to be rather difficult to compute from first principles[4]. To begin with, the depth variations of the temperature and density in a solar model causes the magnetic diffusivity to increase from about $10^2\,\mathrm{cm^2\,s^{-1}}$ in the central core to $\sim 10^4\,\mathrm{cm^2\,s^{-1}}$ at the core-envelope interface. This already substantial variation is however dwarfed by the much larger increase in the

[4]See the bibliography at the end of this chapter for some references.

net magnetic diffusivity expected in the turbulent environment of the convective envelope. We will look into this in some detail in chapter 4, but for the time being let us simply take for granted that η is much larger in the envelope than in the core.

In order to examine the consequences of a strongly depth-dependent magnetic diffusivity on the diffusive eigenmodes, we consider a simplified situation whereby η assumes a constant value η_c in the core, a constant value η_e ($\gg \eta_c$) in the envelope, the transition occurring smoothly across a thin spherical layer coinciding with the core-envelope interface. Mathematically, such a variation can be expressed as

$$\eta(r) = \eta_c + \frac{\eta_e - \eta_c}{2}\left[1 + \mathrm{erf}\left(\frac{r - r_c}{w}\right)\right], \qquad (7.27)$$ {E2.127}

where $\mathrm{erf}(x)$ is the error function, r_c is the radius of the core-envelope interface, and w is the half-width of the transition layer.

We are still facing the 1D eigenvalue problem presented by eq. (7.13)! Expressing time in units of the diffusion time R^2/η_e based on the envelope diffusivity, we seek numerical solutions, subjected to the boundary conditions $f_\lambda(0) = 0$ and smooth matching to a potential field solution in $r/R > 1$, with the diffusivity ratio $\Delta\eta = \eta_c/\eta_e$ as a parameter of the model. Since we can make a reasonable guess at the eigenvalue on the basis of the diffusion time and adopted values of l and η_c ($\sim \pi^2 ln\Delta\eta$, for l and n not too large), inverse iteration is the technique of choice.

Figure 7.2 shows the radial eigenfunctions for the slowest decaying poloidal eigenmodes, with $r_c/R = 0.7$, $w/R = 0.05$ in eq. (7.27) and diffusivity contrasts $\Delta\eta = 1$ (constant diffusivity), 10^{-1} and 10^{-3}. The corresponding eigenvalues, in units of R^2/η_e, are $\lambda = -9.87$, -2.14 and -0.028. Clearly, the (global) decay time is regulated by the region of *smallest* diffusivity, since λ scales approximately as $(\Delta\eta)^{-1}$. Notice also how the eigenmodes are increasingly concentrated in the core region ($r/R \lesssim 0.7$) as $\Delta\eta$ decreases, i.e., they are "expelled" from the convective envelope. This is sometimes called the **diamagnetic effect** in the astrophysical literature. It has interesting consequences for models of the solar dynamo, and will be encountered again in later chapters.

The marked decrease of the diffusive decay time with increasing angular and radial degrees of the eigenmodes is a noteworthy result. It means that left to decay long enough, any arbitrarily complex magnetic field in the Sun or stars will eventually end up looking dipolar[5]. Conversely, a fluid flow acting as a dynamo in a sphere and trying to "beat" Ohmic dissipation can be expected to prefentially produce a magnetic field approximating diffusive

[5]Is this always true? Can you think of circumstances where this would not be the case?

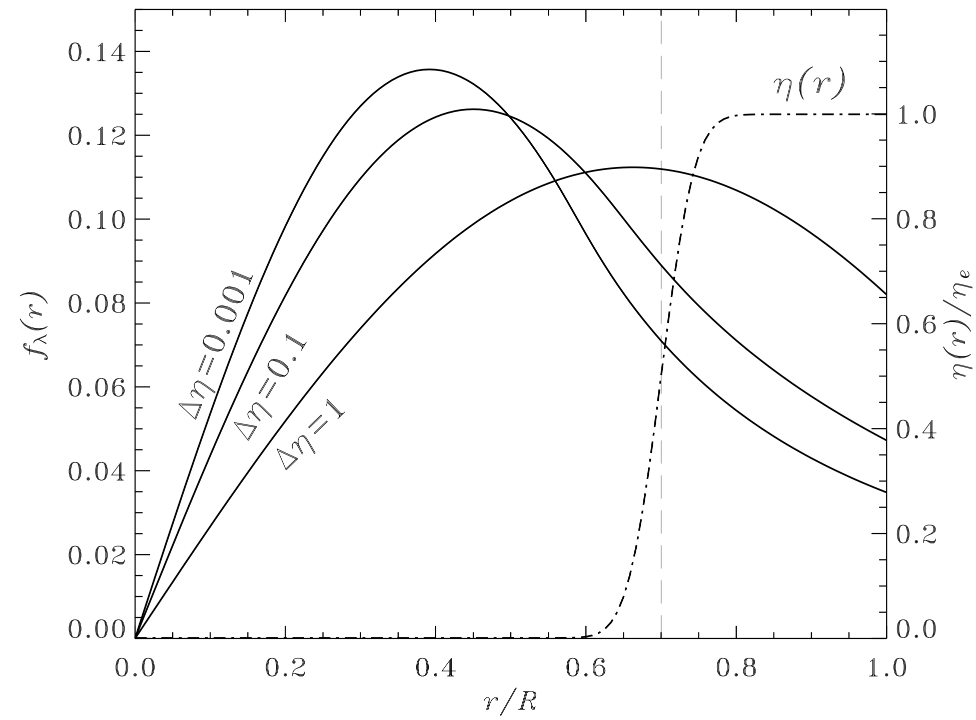

Figure 7.2: {F2.2} Radial eigenfunctions for the slowest decaying ($\ell = 1$) poloidal eigenmodes in a sphere embedded in a vacuum. The diffusivity computed using eq. (7.27) with $r_c/R = 0.7$, $w/R = 0.05$, and for three values of the core-to-envelope diffusivity ratio ($\Delta\eta$). The eigenvalues, in units of η_e/R^2, are $\lambda = -9.87$, -2.14 and -0.028 for $\Delta\eta = 1$, 0.1, and 10^{-3}, respectively. The diffusivity profile for $\Delta\eta = 10^{-3}$ is also plotted (dash-dotted line). The dashed line indicates the location of the core-envelope interface.

eigenmodes of low angular and radial degrees (or some combination thereof), since these are the least sensitive to Ohmic dissipation.

There exists classes of early-type main-sequence stars, i.e. stars hotter and more luminous than the Sun and without deep convective envelope, that are believed to contain strong, large-scale fossil magnetic fields left over from their contraction toward the main-sequence. The chemically peculiar Ap stars are the best studied class of such objects. Reconstruction of their surface magnetic field distribution suggests almost invariably that the fields are largely dominated by the dipole component, as one would have expected from the preceding discussion if the observed magnetic fields have been diffusively decaying for tens or hundreds of millions of years[6].

[6]Care is warranted in making such conclusions on the basis of stellar observations, as the current techniques used to infer the presence and structure of the surface fields, based

7.2 Magnetic field amplification by stretching and shearing

{sec:stretch}

Having now investigated in some the details the resistive decay of magnetic field, we turn to the other physical mechanism embodied in eq. (6.12): growth of the magnetic field in response to the inductive action of a flow **u**. We first take a quick look at field amplification in a few idealized model, and in the next section move on to a specific example using a "real" flow.

7.2.1 Hydrodynamical stretching and field amplification

Let's revert for a moment to the ideal MHD case ($\eta = 0$). The induction equation can then expressed as

$$\frac{\partial \mathbf{B}}{\partial t} + (\mathbf{u} \cdot \nabla)\mathbf{B} = \mathbf{B} \cdot \nabla \mathbf{u} \ , \tag{7.28}$$ {E2.204b}

where it was further assumed that the flow is incompressible ($\nabla \cdot \mathbf{u} = 0$). The LHS of eq. (7.28) is the Lagrangian derivative of **B**, expressing the time rate of change of **B** in a fluid element moving with the flow. The RHS expresses the fact that this rate of change is proportional to the local **shear** in the flow field. Shearing has the effect of **stretching** magnetic fieldlines, which is what leads to magnetic field amplification.

As a simple example, consider on Figure 7.3 a cylindrical fluid element of length L_1, threaded by a magnetic field parallel to the axis of the cylinder, imbedded in a perfectly conducting incompressible fluid and subjected to a stretching motion ($\partial u_x/\partial x > 0$) along its central axis such that its length increases to L_2. Mass conservation demands that $R_2/R_1 = \sqrt{L_1/L_2}$. Conservation of the magnetic flux ($= \pi R^2 B$) in turn leads to

$$\frac{B_2}{B_1} = \frac{L_2}{L_1} \ , \tag{7.29}$$ {E3.1bis}

i.e., the field strength is amplified in direct proportion to the level of stretching. This almost trivial result is in fact at the very heart of *any* magnetic field amplification in the magnetohydrodynamical context, and illustrates two crucial aspects of the mechanism: first, this works only if the fieldlines are frozen into the fluid, i.e., in the high-R_m regime. Second, mass conservation plays an essential role here; the stretching motion along the tube axis *must* be accompanied by a compressing fluid motion perpendicular to the axis

on Zeeman splitting and/or polarization of starlight, are significant biased towards the lower multipoles because the stellar surface remains spatially unresolved.

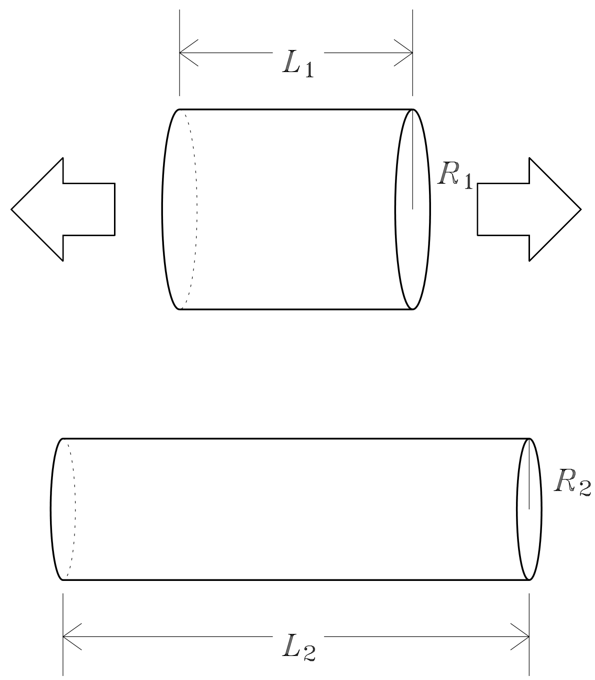

Figure 7.3: {F2.0} Stretching of a magnetized cylindrical fluid element by a diverging flow. The magnetic field is horizontal within the tube, has a strength B_1 originally, and B_2 after stretching. In the flux-freezing limit mass conservation within the tube requires its radius to decrease, which in turn leads to field amplification (see text).

if mass conservation is to be satisfied. It is this latter compressive motion, occurring perpendicular to the magnetic fieldlines forming the flux tube, that is ultimately responsible for field amplification; the horizontal motion occurs parallel to the magnetic fieldline, and so cannot in itself have any inductive effect as per eq. (6.12)[7]. The challenge, of course, is to realize this idealized scenario in practice, i.e., to find a flow which achieves the effect illustrated on Figure 7.3.

7.2.2 The Vainshtein & Zeldovich flux rope dynamo

As trivial as the above example may appear, it can form the basis of a dynamo. S. Vainshtein and Ya. B. Zeldovich have proposed one of the first and justly celebrated "cartoon" model for this idea, as illustrated on Figure

[7]Hold it now, how do you reconcile this statement with eq. (11.3), which indicates rather unambiguously that one can have $\partial\mathbf{B}/\partial t > 0$ with $\mathbf{B} = B_x\hat{\mathbf{e}}_x$ and $\mathbf{u} = u_x(x)\hat{\mathbf{e}}_x$?

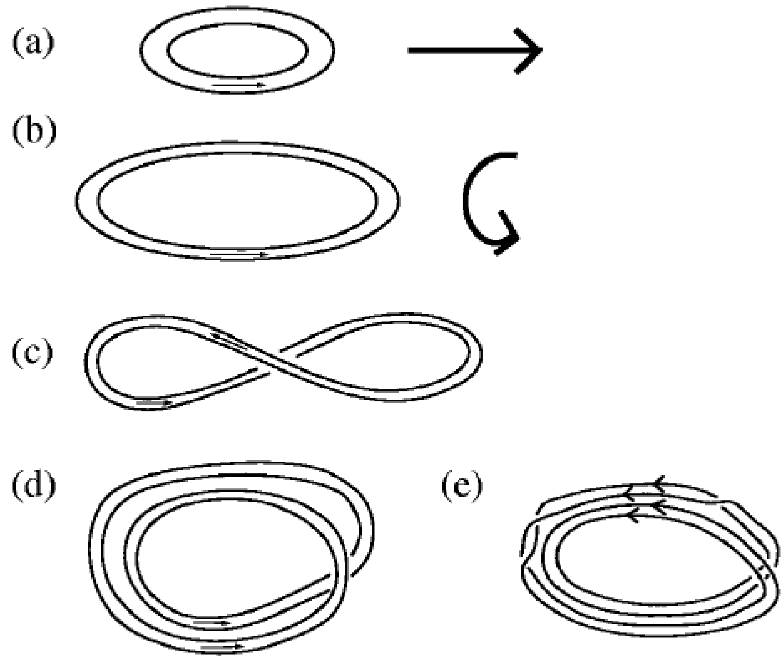

Figure 7.4: {F3.VZ} Cartoon of the Strech-Twist-Fold flux rope dynamo of Vainshtein & Zeldovich. A circular flux rope (a) is (b) stretched, (c) twisted, and (d) folded. Diagram (e) shows the resulting structure after another such step. Diagram digitized straight out of A.D. Gilbert's excellent dynamo review listed in the bibliography.

7.4. The steps are the following:

1. A circular rope of magnetic field is stretched to twice its length ($a \to b$). As we just learned, this doubles the magnetic field strength;
2. The rope is twisted by half a turn ($b \to c$);
3. One half of the rope is folded over the other half in such a way as to align the magnetic field of each half ($c \to d$).

Clearly, this so-called **stretch-twist-fold** sequence (hereafter STF) doubles the field strength while conserving the total cross-section of the original rope, so that the magnetic flux is also doubled. If the sequence is repeated n times, the magnetic field strength (and flux) is then amplified by a factor

$$\frac{B^n}{B_0} \propto 2^n = \exp(n \ln 2) \ , \qquad (7.30) \quad \{\texttt{E3.2}\}$$

with n playing the role of a (discrete) time-like variable, eq. (7.30) indicates an exponential growth of the magnetic field, with a growth rate $\sigma = \ln 2$. Rejoyce! This is our first dynamo!

A concept central to the STF dynamo —and other dynamos to be encountered later— is that of **constructive folding**. Note how essential the twisting step is to the STF dynamo: without it (or with an even number of twists), the magnetic field in each half of the folded rope would end up pointing in opposite direction, and would then add up to zero net flux, a case of **destructive folding**. We'll have a more to say on the STF dynamo in the following chapter; for now we switch gears to consider a mechanism of field amplification of more obvious astrophysical relevance.

7.2.3 Toroidal field production by differential rotation {SSdr}

A situation of great (astro)physical interest is the induction of a toroidal magnetic field via the shearing of a poloidal magnetic field threading a differentially rotating sphere of electrically conducting fluid. Assuming axisymmetry (i.e., the poloidal field and differential rotation share the same symmetry axis) and neglecting once again magnetic dissipation, the induction equation take on the reduced form[8]

$$\frac{\partial A}{\partial t} = 0 \ , \tag{7.31}$$ {E2.07a}

$$\frac{\partial B}{\partial t} = \varpi[\nabla \times (A\hat{\mathbf{e}}_\phi)] \cdot \nabla\Omega \ . \tag{7.32}$$ {E2.07b}

where we took advantage of the poloidal/toroidal separation discussed in §1.10.3. For a steady rotation profile, equation (7.32) integrates immediately to

$$B(r,\theta,t) = B(r,\theta,0) + \{\varpi[\nabla \times (A\hat{\mathbf{e}}_\phi)] \cdot \nabla\Omega\}\, t \ . \tag{7.33}$$ {E2.08}

Anywhere in the domain, the toroidal component of the magnetic field grows linearly in time, at a rate proportional to the net local shear and local poloidal field strength[9]. A toroidal magnetic component is being generated by *stretching* the initially purely poloidal fieldlines in the ϕ-direction; the magnitude of the poloidal magnetic component remains unaffected, as per eq. (7.31)!

Evidently computing B via eq. (7.33) requires a knowledge of the solar internal (differential) rotation profile $\Omega(r,\theta)$. Consider the following parametrization:

$$\Omega(r,\theta) = \Omega_C + \frac{\Omega_S(\theta) - \Omega_C}{2}\left[1 + \mathrm{erf}\left(\frac{r - r_C}{w}\right)\right] \ , \tag{7.34}$$ {E2.68a}

[8] Work it out!

[9] How long would it take for the solar differential rotation to shear a 1 G poloidal field into a 10^5 G toroidal field?

where

$$\Omega_S(\theta) = \Omega_{Eq}(1 - a_2 \cos^2\theta - a_4 \cos^4\theta) \tag{7.35}$$ {E2.68b}

is the surface latitudinal differential rotation. We will make repeated use of this parametrization in this and following and chapters, so let's look into it in some detail. Figure 7.5 shows a 2D helioseismic inversion of the solar internal rotation, together with the profile $\Omega(r, \theta)$ generated using the above expressions with parameter values $\Omega_C/2\pi = 432.8\,\mathrm{nHz}$, $\Omega_{Eq}/2\pi = 460.7\,\mathrm{nHz}$, $a_2 = 0.1264$, $a_4 = 0.1591$, $r_c = 0.713R$, and $w = 0.05R$. The degree of similarity with the "real" Sun is quite reasonable. Note in particular that both profiles are characterized by:

1. A convective envelope ($r \gtrsim r_c$) where the shear is purely latitudinal, with the equatorial region rotating faster than the poles;

2. A core ($r \lesssim r_c$) that rotates rigidly, at a rate equal to that of the surface mid-latitudes;

3. A smooth matching of the core and envelope rotation profiles occurring across a thin spherical layer coinciding with the core-envelope interface ($r = r_c$), so that strong radial shears of opposite signs exist in the polar and equatorial regions.

Figure 7.6 shows the distribution of toroidal magnetic field (part B) resulting from the shearing of pure dipole with field strength 1 G at $r/R = 0.7$ (part A, dotted lines) by the above solar-like differential rotation profile (part A, solid lines). This is nothing more that eq. (7.33) evaluated for $t = 10\,\mathrm{yr}$, with $B(r, \theta, 0) = 0$. Not surprisingly, the toroidal field is concentrated in the regions of large radial shear, at the core-envelope interface (dashed line). Note how the toroidal field distribution is *antisymmetric* about the equatorial plane, in agreement with Hale's polarity rules, and precisely what one would expect from the inductive action of a shear flow that is equatorially symmetric on a poloidal magnetic field that is itself antisymmetric about the equator.

Knowing the distributions of toroidal and poloidal fields on Figure 7.6 allows us to flirt a bit with dynamics, by computing the ϕ-component of the Lorentz force:

$$[\mathbf{F}_L]_\phi = \frac{1}{4\pi\varpi}\mathbf{B}_p \cdot \nabla(\varpi B) \ , \tag{7.36}$$ {E5.09}

The resulting spatial distribution of $[\mathbf{F}_L]_\phi$ is plotted on Figure 7.6C. Examine Fig. 7.6 carefully to convince yourself that the Lorentz force is such as to *oppose* the driving shear. This is an important and totally general property

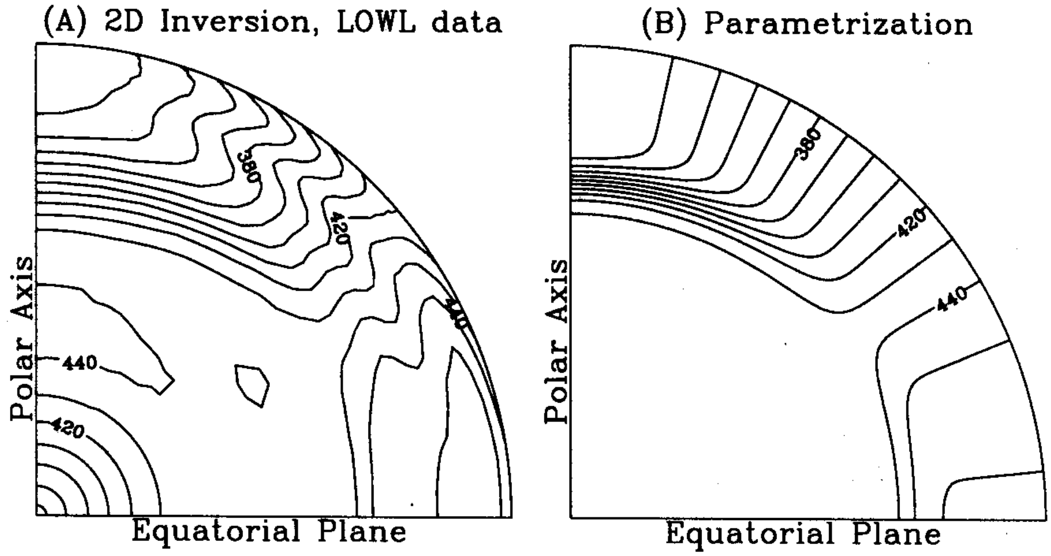

Figure 7.5: {F2.12} Regularized least-square inversion for the internal solar angular velocity, obtained with the LOWL 2-year frequency splitting dataset (left), and parametric representation obtained from eqs. (7.34) — (7.35) (right). The angular velocity is shown in a meridional quadrant, in the form of of angular *frequency*, in the range $340 \leq \Omega/2\pi \leq 460\,$nHz with 10 nHz spacing.

of interacting flows and magnetic fields: the Lorentz force tends to resist the hydrodynamical stretching responsible for field induction. The ultimate fate of the system depends on whether the Lorentz force become dynamically significant before the growth of the toroidal field is mitigated by resistive dissipation; in the solar interior the former situation is far more likely[10].

Clearly, the growing magnetic energy of the toroidal field is supplied by the kinetic energy of the rotational shearing motion (this is hidden the second term on the RHS of eq. (1.76)). In the solar case, this is an attractive field amplification mechanism, because the available supply of rotational kinetic energy is immense. But don't make the mistake of thinking that this is a dynamo! In obtaining eq. (7.33) we have completely neglected magnetic dissipation, and remember, the dynamo we are seeking are flows that can amplify and sustain a magnetic field *against* Ohmic dissipation. Nonetheless, shearing of a poloidal field by differential rotation will turn out to be a central component of *all* solar/stellar dynamo models constructed in later chapters. It is also largely responsible for the strong alignement of galactic magnetic

[10]How would you go about seeking a theoretical justification for this rather sweeping statement?

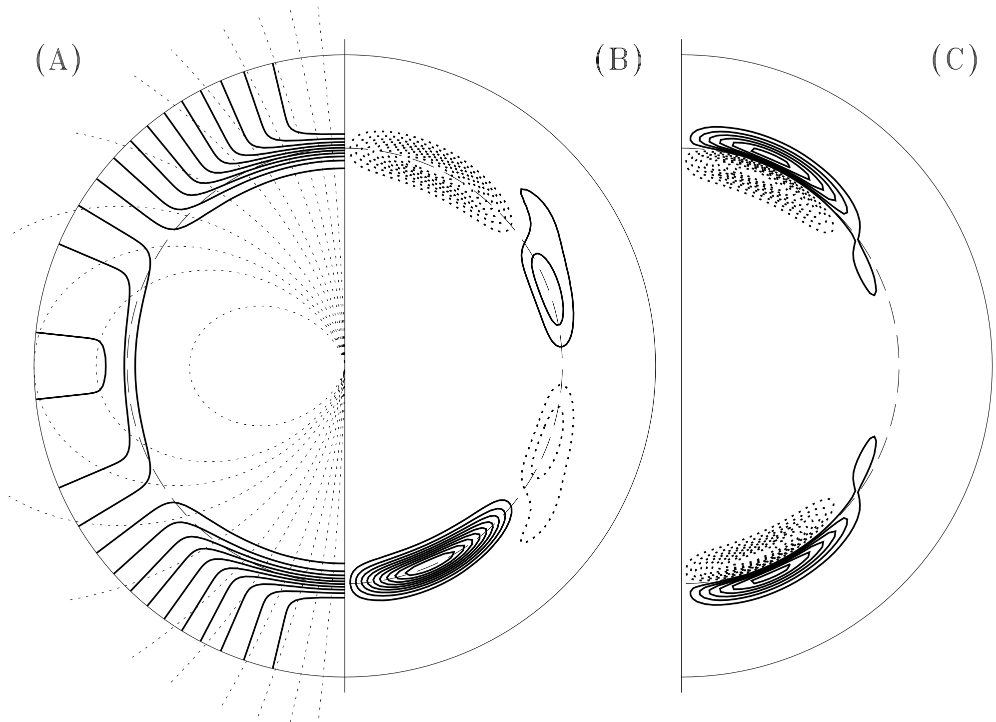

Figure 7.6: {F2.14} Shearing of a poloidal field into a toroidal component by a solar-like differential rotation profile. Part A shows isocontours of the rotation rate $\Omega(r,\theta)/2\pi$ (solid lines, contour spacing 10 nHz as on Fig. 7.5). The dotted lines are fieldlines for a pure dipole. The dashed line is the core-envelope interface at $r/R = 0.7$. Part B shows isocontours of the toroidal field, with solid (dotted) contours corresponding to positive (negative) B. The maximum toroidal field strength is about 2kG, and contour spacing is 0.2kG. Part C shows logarithmically spaced isocontours of the ϕ-component of the Lorentz force associated with the poloidal/toroidal fields of panels A and B.

fields with the direction of galactic rotation, as evidence e.g. on Fig. 2.11.

7.3 Magnetic field evolution in a cellular flow

{Sbadvec}

Having examined separately the resistive decay and hydrodynamical induction of magnetic field, we now turn to a situation where both processes operate simultaneously.

7.3.1 A cellular flow solution {Scell}

In Cartesian geometry, we consider the action of a steady, incompressible ($\nabla \cdot \mathbf{u} = 0$) two-dimensional flow

$$\mathbf{u}(x, y) = u_x(x, y)\hat{\mathbf{e}}_x + u_y(x, y)\hat{\mathbf{e}}_y \qquad (7.37) \quad \{E2.29\}$$

on a two-dimensional magnetic field

$$\mathbf{B}(x, y, t) = B_x(x, y, t)\hat{\mathbf{e}}_x + B_y(x, y, t)\hat{\mathbf{e}}_y \ . \qquad (7.38) \quad \{E2.210\}$$

Note that neither the flow nor the magnetic field have a z-component, and that their x and y-components are both independent of the z-coordinate. The flow is said to be **planar** because $u_z = 0$, and has an ignorable coordinate (i.e., translational symmetry) since $\partial/\partial z \equiv 0$ for all field and flow components. Such a magnetic field can be represented by the vector potential

$$\mathbf{A} = A(x, y, t)\hat{\mathbf{e}}_z \ , \qquad (7.39) \quad \{E2.211\}$$

where, as usual, $\mathbf{B} = \nabla \times \mathbf{A}$. Under this representation, lines of constant A in the $[x, y]$ plane coincide with magnetic fieldlines. The only non-trivial component of the induction equation (1.91) is its z-components, which takes the form

$$\frac{\partial A}{\partial t} + \mathbf{u} \cdot \nabla A = \eta \nabla^2 A \ . \qquad (7.40) \quad \{E2.212\}$$

This is a linear advection-diffusion equation, describing the transport of a passive scalar quantity A by a flow $\mathbf{u}$, and subject to diffusion, the magnitude of which being measured by η. In view of the symmetry and planar nature of the flow, it is convenient to write the 2-D flow field in terms of a stream function $\Psi(x, y)$:

$$\mathbf{u}(x, y) = u_0 \left(\frac{\partial \Psi}{\partial y}\hat{\mathbf{e}}_x - \frac{\partial \Psi}{\partial x}\hat{\mathbf{e}}_y \right) \ . \qquad (7.41) \quad \{E2.213\}$$

It is easily verified that any flow so defined will identically satisfy the condition $\nabla \cdot \mathbf{u} = 0$. As with eq. (7.39), a given numerical value of Ψ uniquely labels one streamline of the flow. Consider now the stream function

$$\Psi(x, y) = \frac{L}{4\pi} \left(1 - \cos\left(\frac{2\pi x}{L} \right) \right) \left(1 - \cos\left(\frac{2\pi y}{L} \right) \right) , \qquad x, y \in [0, L] \quad (7.42) \quad \{E2.214\}$$

This describes a counterclockwise cellular flow centered on $(x, y) = (L/2, L/2)$, as shown on Figure 7.7. The maximal velocity amplitude $\max\|\mathbf{u}\| = u_0$ is found along the streamline $\Psi = u_0 L/(2\pi)$, plotted as a thicker line on Figure 7.7. This streamline is well approximated by a circle of radius $L/4$, and its

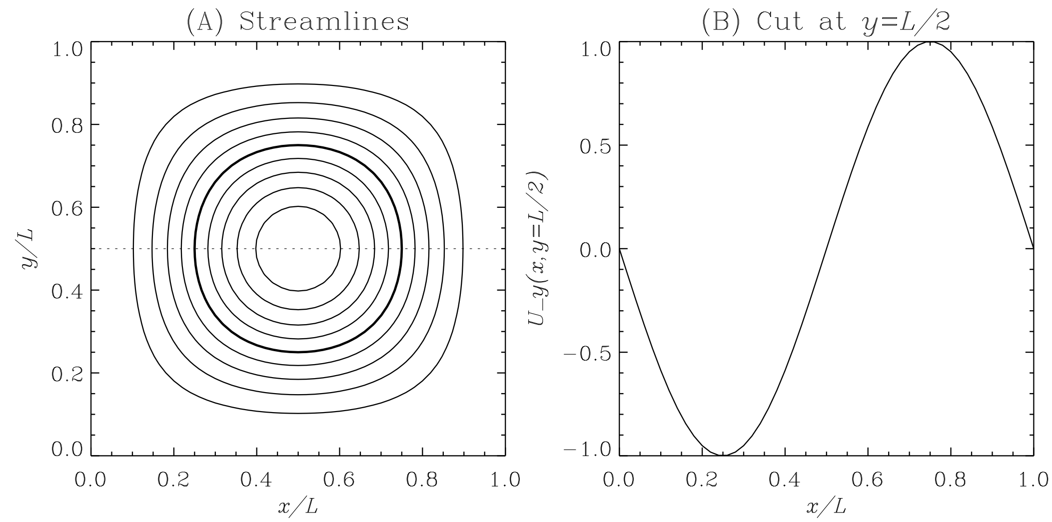

Figure 7.7: {F2.3} Counterclockwise cellular flow generated by the streamfunction given by eq. (7.42). Part (A) shows streamlines of the flow, with the thicker streamline corresponding to $\Psi = u_0 L/(2\pi)$, on which the flow attains its maximum speed u_0. Part (B) shows the profile of $u_y(x)$ along an horizontal cut at $y = 1/2$. A "typical" length scale for the flow is then $\sim L$.

streamwise circulation period turns out to be $1.065\pi L/2u_0$, quite close to what one would expect in the case of a perfectly circular streamline. In what follow this timescale is denoted τ_c and referred to as the **turnover time** of the flow. Note that both the normal and tangential components of the flow vanish on the boundaries $x = 0, L$ and $y = 0, L$. This implies that the domain boundary is itself a streamline ($\Psi = 0$, in fact), and that every streamline interior to the boundary closes upon itself within the spatial domain. These (simple) topological properties of the flow defined by eqs. (7.41) and (7.42) may seem largely irrelevant at this stage of our inquiries, but later chapters will reveal that they are in fact crucial to the dynamo problem.

We now investigate the inductive action of this flow by solving a nondimensional version of eq. (7.40), by expressing all lengths in units of L, and time in units of the **advection time** L/u_0, so that

$$\frac{\partial A}{\partial t} = -\frac{\partial \Psi}{\partial y}\frac{\partial A}{\partial x} + \frac{\partial \Psi}{\partial x}\frac{\partial A}{\partial y} + \frac{1}{\mathrm{R}_m}\left(\frac{\partial^2 A}{\partial x^2} + \frac{\partial^2 A}{\partial y^2}\right) \, , \qquad x, y \in [0, L] \, , \quad (7.43) \qquad \{\texttt{E2.215}\}$$

where $\mathrm{R}_m = u_0 L/\eta$ is the magnetic Reynolds number for this problem, and the corresponding diffusion time is then $\tau_\eta = \mathrm{R}_m$ in dimensionless units. Equation (7.43) is solved as an initial-boundary value problem in two spatial dimensions, with spatial and temporal derivatives both evaluated using second-order centered finite differences. All calculations described below start at $t = 0$ with an initially uniform, constant magnetic field $\mathbf{B} = B_0 \hat{\mathbf{e}}_x$, equivalent to:

$$A(x, y, 0) = B_0 \, y \; . \tag{7.44}$$ {E2.216a}

We consider a situation where the magnetic field normal to the boundaries is held fixed, which amounts to holding the vector potential fixed on the boundary[11]. Figure 7.8 shows the variation with time of the magnetic energy (eq. (1.81)), for four solutions having $\mathrm{R}_m = 10, 10^2, 10^3$ and 10^4. Figure 7.9 shows the evolving shape of the magnetic fieldlines in the $\mathrm{R}_m = 10^3$ solution at 9 successive epochs[12]. The solid dots are "floaters", namely Lagrangian markers moving along with the flow. At $t = 0$ all floaters are equidistant and located on the fieldline initially coinciding with the coordinate line $y/L = 0.5$, that (evolving) fieldline being plotted in the same color as the floaters on all panels. Figure 7.9 covers two turnover times.

At first, the magnetic energy increases quadratically in time. This is precisely what one would expect from the shearing action of the flow on the initial B_x-directed magnetic field, which leads to a growth of the B_y-component that is linear in time. However, for $t/\tau_{\mathrm{c}} \gtrsim 2$ the magnetic energy starts to decrease again and eventually ($t/\tau_{\mathrm{c}} \gg 1$) levels off to a constant value. To understand the origin of this behavior we need to turn to Figure 7.9 and examine the solutions in some detail.

The counterclockwise shearing action of the flow is quite obvious on Fig. 7.9 in the early phases of the evolution, leading to a rather pretty spiral pattern as magnetic fieldlines get wrapped around one another. Note that the distortion of magnetic fieldlines by the flow implies a great deal of *stretching* in the streamwise direction. This is most obvious upon noting that the distance between adjacent floaters increases monotonically in time. It is no accident that the floaters end up in the regions of maximum field amplification on frames 2—5; they are initially positioned on the fieldline coinciding with the line $y = L/2$, everywhere perpendicular to the shearing flow (see Fig. 7.7), which pretty much ensures maximal inductive effect, as per eq. (7.40). The fact that all floaters remain at first "attached" onto their original fieldline is what one would have expected from the fact that this is a relatively high-R_m solution, so that flux-freezing is effectively enforced. As the evolution proceeds, the magnetic field keeps building up in strength (as indicated by the color scale), but is increasingly confined to spiral "sheets" of decreasing thickness.

By the time we hit one turnover time (corresponding approximately to frame 5 on Fig. 7.9), it seems that we are making progress towards our goal of producing a dynamo; we have a flow field which, upon acting on a

[11] Can you figure that one out?

[12] An animation of this evolving solution can be viewed on the course Web Page.

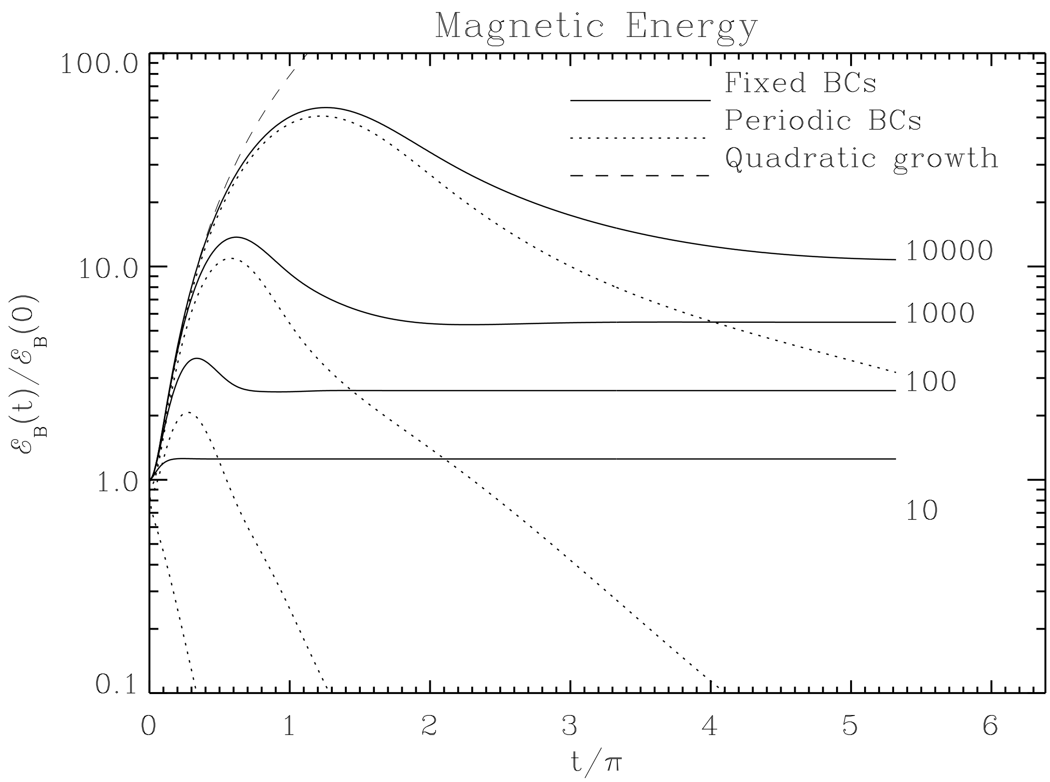

Figure 7.8: {F2.5} Evolution of the magnetic energy for solutions with increasing R_m. The solutions have been computed over 10 turnover times, at which point they are getting reasonably close to steady-state, at least as far as magnetic energy is concerned. One turnover time corresponds to $t/\pi = 0.532$.

preexisting magnetic field, has intensified the strength of that field, at least in some localized regions of the spatial domain. However, beyond $t \sim \tau_{\rm c}$ the sheets of magnetic fields are gradually disappearing, first near the center of the flow cell (frames 5—7), and later everywhere except close to the domain boundaries (frames 7—9). Notice also how, from frame 5 onward, the floaters are seen to "slip" off their original fieldlines. This means that flux-freezing no longer holds; in other words, diffusion is taking place. Yet, we evidently still have $t \ll \tau_\eta$ ($\equiv \mathrm{R}_m = 10^3$ here), which indicates that diffusion should not yet have had enough time to significantly affect the solution. What is going on here?

7.3.2 Flux expulsion {SSflexp}

The solution to this apparent dilemma lies with the realization that we have defined R_m in terms of the global length scale L characterizing the flow. This

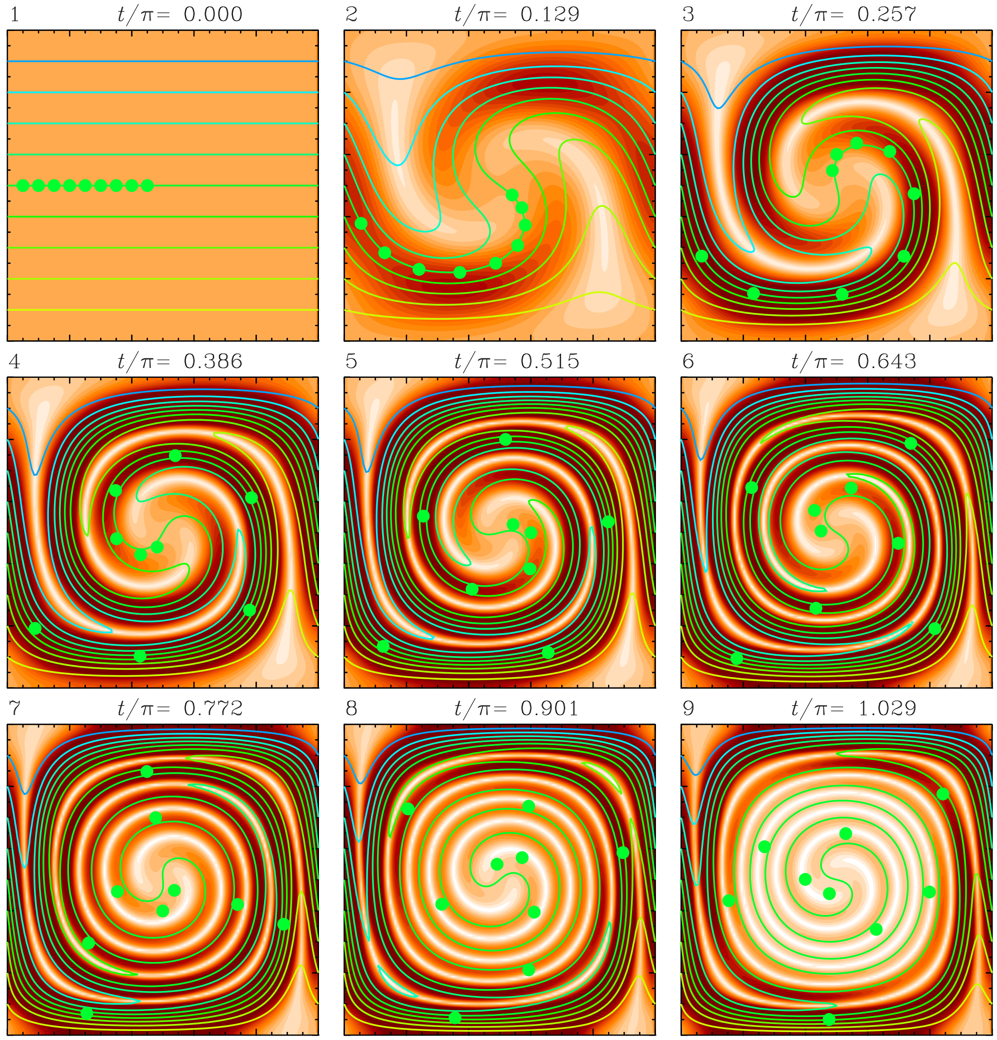

Figure 7.9: {`F2.4`} Solution to equation (7.43) starting from an initially horizontal magnetic field. The panels show the shape of the magnetic fieldlines at successive times. The color scale encodes the absolute strength of the magnetic field, i.e., $\sqrt{B_x^2 + B_y^2}$. The x- and y-axes are horizontal and vertical, respectively, and span the range $x, y \in [0, L]$. Time t is in units of L/u_0. The solid dots are “floaters”, i.e., Lagrangian marker passively advected by the flow. The magnetic Reynolds number is $\mathrm{R}_m = 10^3$.

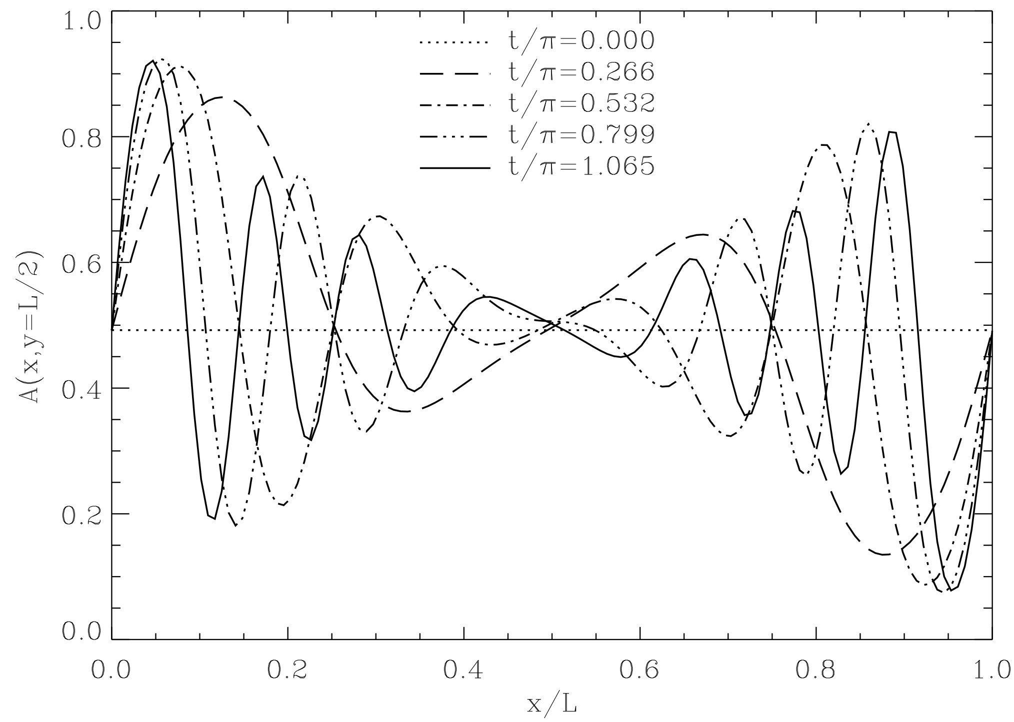

Figure 7.10: {**F2.7**} Cuts of a $\mathrm{R}_m = 10^4$ solution along the coordinate line $y = 0.5$, at successive times. Note how the "typical" length scale ℓ for the solution decreases with time, from $\ell/L \sim 0.25$ at $t/\pi = 0.269$, down to $\ell/L \sim 0.05$ after two turnover times ($t/\pi = 1.065$).

was a perfectly sensible thing to do on the basis of the flow configuration and initial condition on the magnetic field. However, as the evolution proceeds beyond $\sim \tau_\mathrm{c}$ the decreasing thickness of the magnetic field sheets means that the global length scale L is no longer an adequate measure of the "typical" length scale of the magnetic field, which is what is needed to estimate the diffusion time τ_η (see eq. (7.2)). Figure 7.10 shows a series of cuts of the vector potential A in a $\mathrm{R}_m = 10^4$ solution, plotted along the coordinate line $y = L/2$, at equally spaced successive time intervals covering two turnover times. Clearly the inexorable winding of the fieldline leads to a general decrease of the length scale characterizing the evolving solution. In fact, each turnover time adds two new "layers" of alternating magnetic polarity to the spiraling sheet configuration, so that the average length scale ℓ decreases as t^{-1}:

$$\frac{\ell(t)}{L} \propto \frac{L}{u_0 t} \, , \qquad (7.45) \quad \{\texttt{E2.218}\}$$

which in turn implies that the *local* dissipation time is also decreasing as t^{-1}. On the other hand, examination of Fig. 7.9 soon reveals that the (decreasing) length scale characterizes the thickness of elongated magnetic structures that are themselves more or less *aligned* with the streamlines, so that the turnover time $\tau_{\rm c}$ remains the proper timescale measuring field induction. With $\tau_{\rm c}$ fixed and τ_η inexorably decreasing, the solution is bound to reach a point where $\tau_\eta \simeq \tau_{\rm c}$, *no matter how small dissipation actually is.* To reach that stage just takes longer in the higher R_m solutions, since more winding of the fieldlines is needed. Larger magnetic energy can build up in the transient phase, but the growth of the magnetic field is *always* arrested. Equating $\tau_{\rm c}$ ($\sim L/u_0$) to the *local* dissipation time ℓ^2/η, one readily finds that the length scale ℓ at which both process become comparable can be expressed in terms of the *global* R_m as

$$\frac{\ell}{L} = (\mathrm{R}_m)^{-1/2} \ , \qquad \mathrm{R}_m = \frac{u_0 L}{\eta} \ . \tag{7.46}$$ {E2.251}

That such a balance between induction and dissipation materializes means that a steady-state can be attained. Figure 7.11 shows four such steady states solutions for increasing values of the (global) magnetic Reynolds number R_m. The higher R_m solutions clearly show **flux expulsion** from the central regions of the domain. This is a general feature of steady, high-R_m magnetized flows with closed streamlines: magnetic flux is expelled from the regions of closed streamlines towards the edges of the flow cells, where it ends up concentrated in **boundary layers** which indeed have a thickness of order $\mathrm{R}_m^{-1/2}$, as suggested by eq. (7.46). It is important to understand how and why this happens.

To first get an intuitive feel for how flux expulsion operates, go back to Figure 7.9. As the flow wraps the fieldlines around one another, it does so in a manner that folds fieldlines of opposite polarity closer and closer to each other. When two such fieldlines are squeezed closer together than the dissipative length scale (eq. [7.46]), resistive decay takes over and destroys the field faster than it is being stretched. This is another instance of destructive folding, and can only be avoided along the boundaries, where the normal component of the field is held fixed. For flux expulsion to operate, flux-freezing must be effectively enforced on the spatial scale of the flow. Otherwise the field is largely insensitive to the flow, and fieldlines are hardly deformed with respect to their initial configuration (as on panel [A] of Fig. 7.11).

Consider now the implication for the total magnetic flux across the domain; flux conservation requires that the normal flux $B_0 L$ imposed at the right and left boundaries must somehow cross the interior, otherwise maxwell's equation $\nabla \cdot \mathbf{B} = 0$ would not be satisfied; because of flux expulsion, it can only do so in the thin layers along the bottom and top boundaries. Since

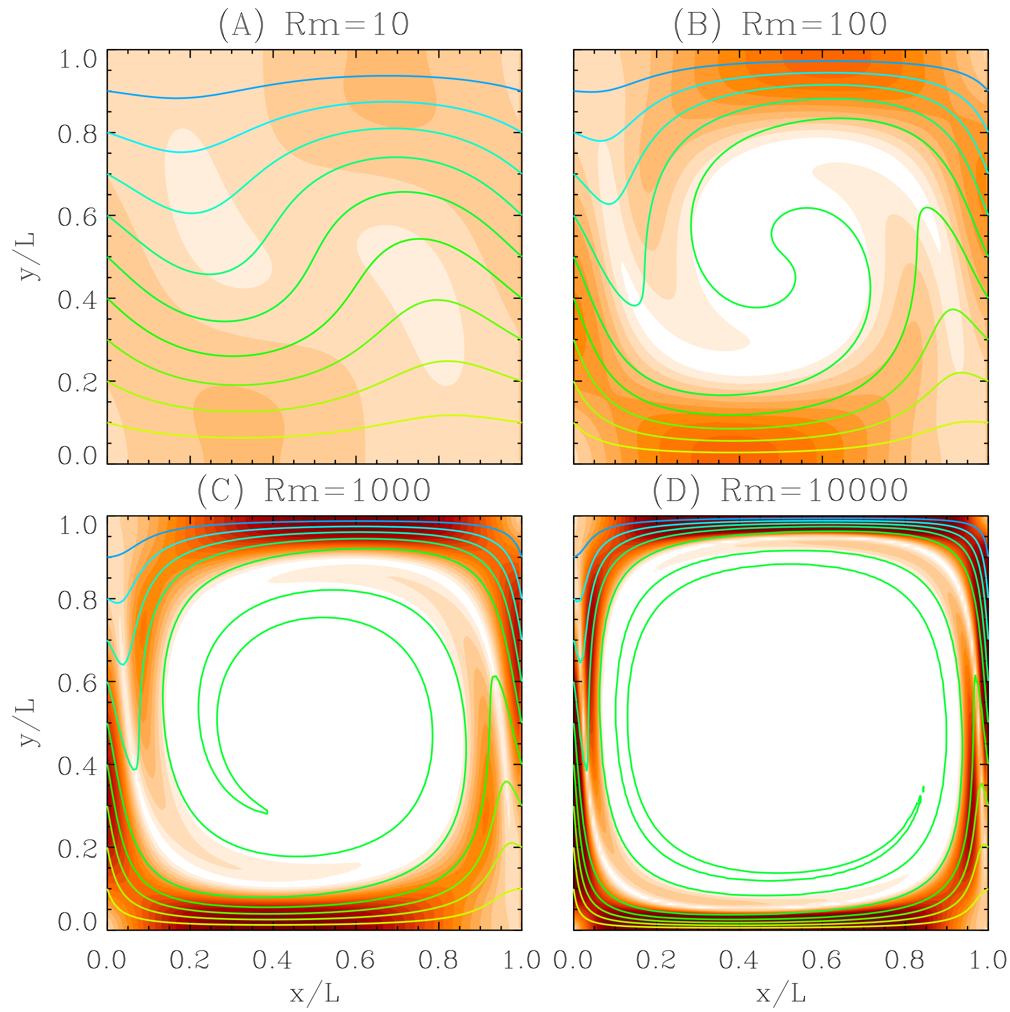

Figure 7.11: {F2.8} Steady-state solutions to the cellular flow problem, for increasing values of the magnetic Reynolds number R_m. The $R_m = 10^4$ solution is at the resolution limit of the $N_x \times N_y = 128 \times 128$ mesh used to obtain these solutions, as evidenced on part (D) by the presence of small scale irregularities where magnetic fieldlines are sharply bent. The color scale encodes the local magnitude of the magnetic field. Note how, in the higher R_m solutions, magnetic flux is expelled from the center of the flow cell. With $\mathcal{E}_B(0)$ denoting the energy of a purely horizontal field with same normal boundary flux distribution, the magnetic energy for these steady states is $\mathcal{E}_B/\mathcal{E}_B(0) = 1.37$, 2.80, 5.81 and 11.75, respectively, for panels (A) through (D).

the thickness of these layers scales as $\mathrm{R}_m^{-1/2}$, it follows that the field strength therein scales as $\sqrt{\mathrm{R}_m}$, which in turn implies that the total magnetic energy in the domain also scales as $\sqrt{\mathrm{R}_m}$ in the $t \gg \tau_c$ limit [13].

7.3.3 Digression: the electromagnetic skin depth {SSskdp}

You may recall that a sinusoidally oscillating magnetic field imposed at the boundary of a conductor will penetrate the conductor with an amplitude decreasing exponentionally iaway from the boundary and into the conductor, with a length scale called the **electromagnetic skin depth**:

$$\ell = \sqrt{\frac{2\eta}{\omega}} \qquad (7.47) \quad \{E2.260\}$$

Now, go back to the cellular flow and imagine that you are an observer located in the center of the flow cell, looking at the domain boundaries while rotating with angular velocity u_0/L; what you "see" in front of you is an "oscillating" magnetic field, in the sense that it flips sign with "angular frequency" u_0/L. The corresponding electromagnetic skin depth would then be

$$\frac{\ell}{L} = \sqrt{\frac{2\eta}{u_0 L}} =\equiv \sqrt{\frac{2}{\mathrm{R}_m}}\ . \qquad (7.48) \quad \{E2.261\}$$

which basically corresponds to the thickness of the boundary layer where significant magnetic field is present in the steady-states shown on Figure 7.11. How about that for a mind flip...

7.3.4 Timescales for field amplification and decay

Back to our cellular flow. Flux expulsion or not, it is clear from Figure 7.8 (solid lines) that some level of field amplification has occurred in the high R_m solutions, in the sense that $\mathcal{E}_B(t \to \infty) > \mathcal{E}_B(0)$. But is this a dynamo? The solutions of Fig. 7.11 have strong electric currents in the direction perpendicular to the plane of the paper, and these currents are subjected to resistive dissipation. Have we then reached the goal stated at the beginning of the chapter, namely, to amplify and maintain a weak, preexisting magnetic field against Ohmic dissipation?

In a narrow sense yes, but a bit of reflection will show that the boundary conditions are playing a crucial role. The only reason that the magnetic energy does not asymptotically go to zero is that the normal field component is held fixed at the boundaries, which, in the steady-state, implies a non-zero

[13] Hold it, $\mathcal{E}_\mathrm{B} \propto \mathbf{B}^2$ as per eq. (1.81); how can the magnetic field strength and magnetic energy *both* scale as $\sqrt{\mathrm{R}_m}$?

Poynting flux into the domain across the left and right vertical boundaries. The magnetic field is not avoiding resistive decay because of field induction within the domain, but rather because external energy (and magnetic flux) is being pumped in through the boundaries. This is precisely what is embodied in the second and third terms on the RHS of eq. (1.76).

What if this were not the case? One way to work around the boundary problem is to replace the fixed flux boundary conditions by periodic boundary conditions:

$$A(x,0) = A(x,L) \ , \qquad A(0,y) = A(L,y) \ . \tag{7.49}$$ {E2.262}

There is still a net flux across the vertical boundary at $t = 0$, but the boundary flux is now free to decay away along with the solution. You get to compute such a solution in Problem 2.3. It is time to reveal that the hitherto unexplained dotted lines on Fig. 7.8 correspond in fact to solutions computed with such boundary conditions, for the same cellular flow and initial condition as before.

Evidently the magnetic energy now decays to zero, confirming that the boundaries indeed played a crucial role in the sustenance of the magnetic field in our previous solutions. What is noteworthy is the rate at which it does so. In the absence of the flow and with freely decaying boundary flux, the initial field would diffuse away on a timescale $\tau_\eta \sim L^2/\eta$, which is equal to R_m if we retain the scaling of τ in terms of L/u_0. With the flow turned on, the decay proceeds at an accelerated rate because of the inexorable decrease of the typical length scale associated with the evolving solution, which we argued earlier varied as t^{-1}. What then is the typical timescale for this enhanced dissipation? The decay phase of the field (for $t \gg L/u_0$) is approximately described by

$$\frac{\partial A}{\partial t} = \eta \nabla^2 A \ . \tag{7.50}$$ {E2.263}

An estimate for the dissipation timescale can be obtained once again via dimensional analysis, by replacing ∇^2 by $1/\ell^2$, as in §7.1 but now with the important difference that ℓ is now a function of time:

$$\ell \to \ell(t) = \left(\frac{L}{t}\right)\left(\frac{L}{u_0}\right) \ , \tag{7.51}$$ {E2.264}

in view of our previous discussion (cf. Fig. 7.10 and accompanying text). This leads to

$$\frac{\partial A}{\partial t} \simeq -\frac{\eta u_0^2 t^2}{L^4} A \ , \tag{7.52}$$ {E2.265}

where the minus sign is introduced in view of the fact that $\nabla^2 A < 0$ in the decay phase. Equation (7.52) integrates to

$$\frac{A(t)}{A_0} = \exp\left[-\frac{\eta u_0^2}{3L^4}t^3\right] = \exp\left[-\frac{1}{3\mathrm{R}_m}\left(\frac{u_0^3 t^3}{L^3}\right)\right] . \qquad (7.53)$$ {E2.266}

This last expression indicates that with t measured in units of L/u_0, the decay time scales as $\mathrm{R}_m^{1/3}$. This is indeed a remarkable situation: in the low magnetic diffusivity regime (i.e., high R_m), the flow has in fact *accelerated* the decay of the magnetic field, even though large field intensification can occur in the early, transient phases of the evolution. This is not at all what a dynamo should be doing!

As it turns out, flux expulsion is even trickier than the foregoing discussion may have led you to believe! Flux expulsion destroys the mean magnetic field component directed *perpendicular* to the flow streamlines. It cannot do a thing to a mean component oriented *parallel* to streamlines. For completely general flow patterns and initial conditions, the dissipative phase with timescale $\propto \mathrm{R}_m^{1/3}$ actually characterizes the approach to a state where the advected trace quantity —here the vector potential A— becomes constant *along each streamline*, at a value $\bar{A}$ equal to the initial value of A averaged on each of those streamlines. For the cellular flow and initial conditions used above, this average turns out to be $\bar{A} = 0.5$ *for every streamline*, so that the $\mathrm{R}_m^{1/3}$ decay phase corresponds to the true decay of the magnetic field to zero amplitude. If $\bar{A}$ varies from one fieldline to the next, however, the $\mathrm{R}_m^{1/3}$ phase is followed by a third decay phase, which proceeds on a timescale $\sim \mathrm{R}_m$, since induction no longer operates ($\mathbf{u}\cdot\nabla A = 0$) and the typical length scale for A is once again L. You get to explore this phenomenon in problem 2.4 [14]. At any rate, even with a more favorable initial condition we have further delayed field dissipation, but we still don't have a dynamo since dissipation will proceed inexorably, on the "long" timescale $\mathrm{R}_m(L/u_0)$.

7.3.5 Global flux expulsion in spherical geometry: axisymmetrization {ssec:axisymm}

You may think that the flux expulsion problem considered in the preceding section has nothing to do with any astronomical objects you are likely to encounter in your future astrophysical carreers. Wroooong!

Consider the evolution of a magnetic field pervading a sphere of electrically conducting fluid, with the solar-like differential rotation profile already encountered previously (§7.2.3, Fig. 7.5 and eqs. (7.34)—(7.35)), and with

[14] In case you're too lazy to do the problem, you can view an animation of this solution on the Course Home Page. But please do the problem anyway.

the field having initially the form of an dipole whose axis is inclined by an angle Θ with respect to the rotation axis ($\theta = 0$). Such a magnetic field can be expressed in terms of a vector potential having components:

$$A_r(r, \theta, \phi) = 0 \qquad (7.54) \quad \{\texttt{eq:2.69a}\}$$

$$A_\theta(r, \theta, \phi) = (R/r)^2 \sin\Theta(\sin\beta\cos\phi - \cos\beta\sin\phi) \qquad (7.55) \quad \{\texttt{eq:2.69b}\}$$

$$A_\phi(r, \theta, \phi) = (R/r)^2[\cos\Theta\sin\theta - \sin\Theta\cos\theta(\cos\beta\cos\phi + \sin\beta\sin\phi)] \qquad (7.56) \quad \{\texttt{eq:2.69c}\}$$

where β is the angle between the $\phi = 0$ plane, and the plane defined by the dipole and coordinate axes.

Now, the vector potential for an inclined dipole can be written as the sum of two contributions, the first corresponding to an aligned dipole ($\Theta = 0$), the second to a perpendicular dipole ($\Theta = \pi/2$), their relative magnitude being equal to $\tan\Theta$[15]. Since the governing equation is linear, the solution for an inclined dipole can be broken into two independent solutions for the aligned and perpendicular dipoles. The former is precisely what we investigated already in §7.2.3, where we concluded there that the shearing of an aligned dipole by an axisymmetric differential rotation would lead to the buildup of a toroidal component, whose magnitude would grow linearly in time at a rate set by the magnitude of the shear.

The solution for a perpendicular dipole is in many way similar to the cellular flow problem of §7.3. You can see how this may be the case by imagining looking from above onto the equatorial plane of the sphere; the fieldlines contained in that plane will have a curvature and will be contained within a circular boundary, yet topologically the situation is similar to the cellular flow studied in the preceding section: the (sheared) flow in the equatorial plane is made of closed, circular streamlines contained within that plane, so that we can expect flux expulsion to occur. The equivalent of the turnover time here is the differential rotation timescale, namely the time for a point located on the equator to perform a full 2π revolution with respect the poles:

$$\tau_{\rm DR} = (\Omega_{\rm Equ} - \Omega_{\rm Pole})^{-1} = \Omega_\odot(a_2 + a_4) \ , \qquad (7.57) \quad \{???\}$$

where the second equality follows directly from eq. (7.35). For a freely decaying dipole, the perpendicular component of the initial dipole will then be subjected to flux expulsion, and dissipated away, at a rate far exceeding purely diffusive decay in the high R_m limit, as argued earlier.

But here is the amusing thing; for an observer looking at the magnetic field at the surface of the sphere, the enhanced decay of the perpendicular

[15]Can you work out the corresponding vector potential components ?

component of the dipole will translate into a gradual decrease in the inferred tilt axis of the dipole. Figure 7.12 shows this effect, for the differential rotation profile given by eq. (7.34) and a magnetic Reynolds number $R_m = 10^3$. The equivalent of the turbnoiver time for this problem is Contours of constant B_r are plotted on the surface $r/R = 1$, with the neutral line ($B_r = 0$) plotted as a thicker line. At $t = 0$ the field has the form of a pure dipole tilted by $\pi/3$ with respect to the coordinate axis, and the sphere is oriented so that the observer (you!) is initially looking straight down the magnetic axis of the dipole. Advection by the flow leads to a distorsion of the initial field, with the subsequent buildup of small spatial scales in the r- and $\theta-$directions (only the latter can be seen here)[16]. After two turnover times (last frame), the surface field looks highly axisymmetric.

So, in a differentially rotating fluid system with high R_m, flux expulsion leads to the **symmetrization** of any non-axisymmetric magnetic field component initially present —or contemporaneously generated. The efficiency of the symmetrization process should make us a little cautious in assuming that the large-scale magnetic field of the Sun, which one would deem roughly axisymmetric upon consideration of surface things like the sunspot butterfly diagram, is characterized by the same level of axisymmetry in the deep-seated generating layers, where the dynamo is presumed to operate. After all, standing in between is a thick, axisymmetrically differentially rotating convective envelope that must be reckoned with. In fact, observations of coronal density structures in the descending phase of the solar cycle can be interpreted in terms of a large-scale, tilted dipole component, with the tilt angle steadily decreasing over 3—4 years towards solar minimum. Interestingly, the differential rotation timescale for the Sun is ~ 6 months. Are we seeing the axisymmetrization process in operation ? Maybe. Axisymmetry is certainly a very convenient modeling assumption when working on the large scales of the solar magnetic field, but it may be totally wrong.

You may recall from §2.1.2 that the magnetic field of Saturn stands out among other solar system planets as having a symmetry axis aligned exactly with its rotation axis. Saturn also has the strongest large-scale surface differential rotation, with a broad equatorial "jet" peaking at XXX times the polar angular velocity. Structural models of Saturn also indicate that this differential rotation may well extend in the interior, with the angular velocity being constant along cylinders concentric with the rotation axis. Saturn's magnetic field is most likely generated by a dynamo mechanism operating in its metallic Hydrogen core, extending to a fractional radius of about 0.55. So imagine now that the dynamo-generated field is indeed inclined with respect to the rotation axis, like in most other planets. In between this field

[16]An animation of this solution, as well as a few others for different R_m and/or tilt angle, can be viewed on the course Web Page.

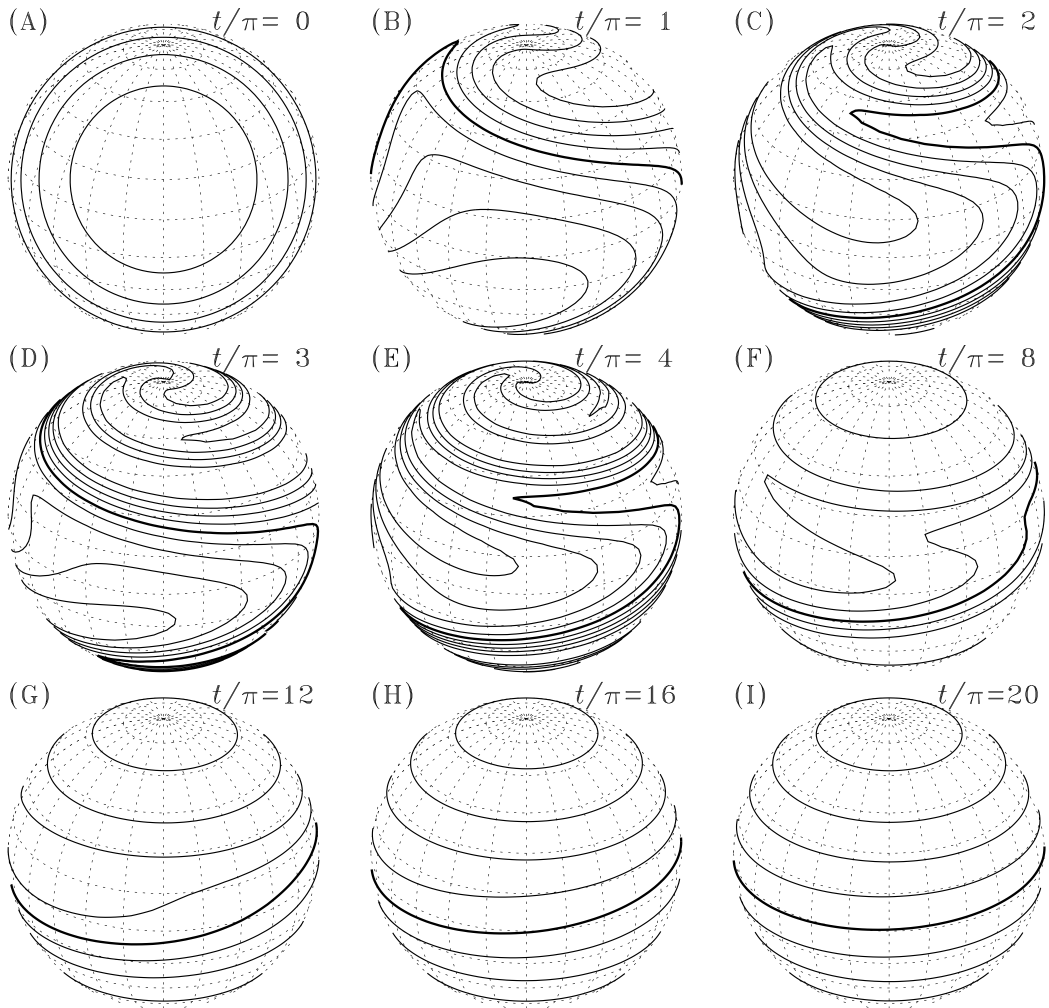

Figure 7.12: {`F2.13`} Symmetrization of an inclined dipole in a electrically conducting sphere in a state of solar-like axisymmetric differential rotation. Each panel shows contours of constant B_r at the surface of the sphere, and the solution is matched to a potential in the exterior ($r/R > 1$). The differential rotation is given by eq. (7.34). Time is given in units of $\tau_{\rm DR}$, in which the turnover period (or differential rotation period) is equal to 2π.

and the surface, where we make measurement, there stands a strongly differentially rotating partly conducting envelope, where axisymmetrization can take place. The key here is that the electrical conductivity in the molecular Hydrogen envelope must be sufficiently large for a coupling between the flow and field (in other words, we need $\mathrm{R}_m \gtrsim 1$, not $\mathrm{R}_m \ll 1$. See the references listed in the bibliography for more on this interesting Saturnian problem.

7.4 Two anti-dynamo theorems {Santidyn}

The cellular flow studied in §7.3, although it initially looked encouraging (cf. Fig. 7.8), proved not to be a dynamo after all. Is this peculiar to the flow defined by eqs. (7.41)–(7.42), or is this something more general? Exhaustively testing for dynamo action in all possible kinds of flow geometries is clearly impractical. However, it turns out that one can rule out *a priori* dynamo action in many classes of flows. These demonstrations are known as **anti-dynamo theorems**.

A powerful anti-dynamo theorem due to Ya. B. Zeldovich, has a lot to teach us about our cellular flow results. The theorem rules out dynamo action in steady planar flows in cartesian geometry, i.e., flows of the form

$$\mathbf{u}_2(x,y,z) = u_x(x,y,z)\hat{\mathbf{e}}_x + u_y(x,y,z)\hat{\mathbf{e}}_y \qquad (7.58)$$ {Eadt1}

in a bounded volume V at the boundaries (∂V) of which the magnetic field vanishes. Note that no other restrictions are placed on the magnetic field, which can depend on all three spatial coordinate as well as time. Nonetheless, in view of eq. (7.58) is will prove useful to consider separately the z-component of the magnetic field $B_z(x,y,z,t)$ from the (2D) field component in the $[x,y]$ plane (hereafter denoted $\mathbf{B}_2$). It is readily shown that the z-component of the induction equation then reduces to

$$\frac{\partial B_z}{\partial t} + \mathbf{u}\cdot\nabla B_z = \eta\nabla^2 B_z \qquad (7.59)$$ {Eadt2}

for spatially constant magnetic diffusivity. Now, the LHS is just a Lagrangian derivative, yielding the time variation of B_z as one moves along with the fluid. Multiplying this equation by B_z and integrating over V yields, after judicious use of a suitable vector identity and of the divergence theorem[17]:

$$\frac{1}{2}\int_V \frac{\mathrm{D}B_z^2}{\mathrm{D}t}dV = \int_{\partial V} B_z(\nabla B_z)\cdot\mathbf{n}\mathrm{d}S - \eta\int_V (\nabla B_z)^2\mathrm{d}V\ . \qquad (7.60)$$ {Eadt3}

Now, the first integral on the RHS vanishes since $\mathbf{B}=0$ on ∂V by assumption. The second integral is positive definite, *therefore B_z always decays on the diffusive timescale* (cf. §7.1).

[17]Try it!

Consider now the magnetic field $\mathbf{B}_2$ in $[x, y]$ planes. The most general such 2D field can be written as the sum of a solenoidal and potential component:

$$\mathbf{B}_2(x, y, z, t) = \nabla \times (A\hat{\mathbf{e}}_z) + \nabla\Phi \ , \qquad (7.61) \quad \{\texttt{Eadt4}\}$$

where the vector potential A and scalar potential Φ both depend on all three spatial coordinates and time. Evidently, the constraint $\nabla \cdot \mathbf{B} = 0$ implies

$$\nabla_2^2\Phi = -\frac{\partial B_z}{\partial z} \ , \qquad (7.62) \quad \{\texttt{Eadt5}\}$$

where $\nabla_2^2 \equiv \partial^2/\partial x^2 + \partial^2/\partial y^2$ is the 2D Laplacian operator in the $[x, y]$ plane. Clearly, once B_z has resistively dissipated, i.e., for times much larger than the global resistive decay time τ, Φ is simply a solution of the 2D Laplace equation $\nabla_2^2\Phi = 0$.

Here comes the sneaky part. We take the curl of the induction equation. Upon substituting eq. (7.61), the z-component of the resulting expression yields

$$\nabla \times \nabla \times \left[\frac{\partial A}{\partial t} + \mathbf{u}_2 \cdot \nabla A - \eta\nabla_2^2 A - \mathbf{u}_2 \times \nabla\Phi\right] = 0 \ , \qquad (7.63) \quad \{???\}$$

with $\nabla \cdot (A\hat{\mathbf{e}}_z) = 0$ as a choice of gauge. Note that only one term involving Φ survives, because $\nabla \times \nabla\Phi = 0$ identically. In general, the above expression is only satisfied if the quantity in square brackets itself vanishes, i.e.,

$$\frac{\mathrm{D}A}{\mathrm{D}t} = \eta\nabla_2^2 A + \mathbf{u}_2 \times \nabla\Phi \ . \qquad (7.64) \quad \{\texttt{Eadt6}\}$$

This expression is identical to that obtained above for B_z, except for the presence of the source term $\mathbf{u}_2 \times \nabla\Phi$. However, we just argued that for $t \gg \tau$, $\nabla_2^2\Phi = 0$. In addition, $\mathbf{B}$ vanishes on ∂V by assumption, so that the only possible asymptotic interior solutions are of the form $\Phi =$const, which means that the source term vanishes in the limit $t \gg \tau$. From this point on eq. (7.64) is indeed identical to eq. (7.59), for which we already demonstrated the inevitability of resistive decay. Therefore, dynamo action, i.e., maintenance of a magnetic field against resistive dissipation, is impossible in a planar flow for *any* 3D magnetic field.

Another powerful anti-dynamo theorem, predating in fact Zeldovich's, is due to T.G. Cowling. This anti-dynamo theorem is particularly important historically, since it rules out dynamo action for 3D but axisymmetric flows and magnetic fields, which happen to be the types of flows and fields one sees in the Sun, at least on the larger spatial scales. Rather than going over one of the many very mathematical proofs of Cowling's theorem found in the literature, we'll just follow the underlying logic of our proof of Zeldovich's theorem.

Assuming once again that there are no sources of magnetic field exterior to the domain boundaries, we consider the inductive action of a 3D, steady axisymmetric flow on a 3D axisymmetric magnetic field. Recall from §1.10.3 that under these circumstances the induction equation can be separated into the two components given by eqs. (1.94)–(1.95). The LHS of these expressions is again a Lagrangian derivative for the quantities in parentheses, and the first terms on each RHS are of course diffusion. The next term on the RHS of eq. (1.95) vanishes for incompressible flows, and remains negligible for very subsonic compressible flows. The last term on the RHS, however, is a source term, in that it can lead to the growth of B *as long as* A *does not decay away*. This is the very situation we have considered in §7.2.3, by holding A fixed as per eq. (7.31). However, there is no similar source-like term on the RHS of eq. (1.94), which governs the evolution of A.

This should now start to remind you of Zeldovich's theorem. In fact, eq. (1.94) is structurally identical to eq. (7.59), for which we demonstrated the inevitability of resistive decay in the absence of sources exterior to the domain. This means that A will inexorably decay, implying in turn that B will then also decay once A has vanished. Since axisymmetric flows cannot maintain A against Ohmic dissipation, *a 3D axisymmetric flow cannot act as a dynamo for a 3D axisymmetric magnetic field.* [18]. Cowling's theorem is not restricted to spherical geometry, and is readily generalized to any situation where both flow and field showing translational symmetry in one and the same spatial coordinate. Such physical systems are said to have an **ignorable coordinate**.

It is worth pausing and reflecting on what these two antidynamo theorems imply for the cellular flow of §7.3. It was indeed a planar flow ($u_z = 0$), and moreover the magnetic field had an ignorable coordinate ($\partial \mathbf{B}/\partial z \equiv 0$)! We thus fell under the purview of both Zeldovich's and Cowling's theorems, so in retrospect our failure to find dynamo action is now understood. Clearly, the way to evade both theorems is to consider flows and fields that are fully three-dimensional, and lack translational symmetry at least in the magnetic field. This is precisely what we do in the following chapter.

Problems:

1. Solve eq. (1.97) by the technique of separation of variables, and verify that eq. (1.99) is indeed the appropriate solution.

2. This problems gets you to further explore the diffusive decay problem of §7.1.4 as a numerical 1-D eigenvalue problem. Use the same magnetic diffusivity profile (with $\Delta\eta = 10^{-2}$), but to avoid having to deal

[18]A fact often unappreciated is that Cowling's theorem does not rule out the dynamo generation of a *non-axisymmetric* 3D magnetic field by a 3D axisymmetric flow.

with the matching of your solutions to a potential field in $r/R > 1$, focus instead on the decay of purely *toroidal* axisymmetric magnetic fields. You may use the computing language of your choice, but please do include listings of all your codes with your solutions. Some useful Fortran-77 routines for the solution of tridiagonal systems of linear algebraic equations, together with instructions for use, can be obtained from the course Web Page:

`http://www.astro.umontreal.ca/~paulchar/phy6795/phy6795.html`

From the top of the main page, click on *Problems: Software and hints*, locate the appropriate subsection, and follow the instructions given there.

(a) By assuming a spatial dependence of the form given by eq. (7.12), show that eq. (7.1) reduces to eq. (7.13).

(b) Use centered finite differences to discretize eq. (7.13), and solve the resulting system of algebraic equations using inverse iteration. As a test of your numerical implementation, do first a problem for constant η, and compare your numerical results to the analytic solutions found in §7.1.3.

(c) Using now the error function profile for the magnetic diffusivity, obtain solutions for the first three angular degrees l and radial harmonics degrees n (for a total of 9 modes). Labels your solutions in terms of (l, n) values, and rank them according to decay time.

(d) Compare the decay times of your toroidal eigenmodes to those of poloidal eigenmodes of corresponding angular and radial degrees, as shown on Fig. 7.1. Can you pick out a trend ? If so, try to come up with a sensible explanation for it.

3. Dynamical backreaction in shearing problem

4. Fill in the missing mathemarical steps leading to eq. (7.60).

5. Recompute the cellular flow solution of §7.3 using periodic boundary conditions. Use second order centered finite differences to discretize the RHS of eq. (7.43), and the leapfrog scheme for time stepping. Make use of the ghost cell formalism to enforce your periodic boundary conditions.

   (a) First compute a $\mathrm{R}_m = 10^3$ solution using the initial condition given by eq. (7.44) compute the time evolution of the total magnetic energy, and verify that it matches that plotted on Fig. 7.8.

(b) Now repeat your calculation, but use this time as an initial condition a vector potential $A(y)$ that is gaussian in y and peaks at $y/L = 0.5$:

$$A(x, y, 0) = 2B_0 \exp\left(-\frac{(y - L/2)^2}{(L/4)^2}\right).$$

Compare the resulting magnetic energy evolution to that you obtained earlier, and search your second solution for evidence of three more or less distinct amplification and decay phases:

(a) Growth of the field, on timescale $\sim L/u_0$;

(b) Enhanced resistive decay, on a timescale $\sim \mathrm{R}_m^{1/3} \equiv 10(L/u_0)$;

(c) A final decay phase, with timescale $\sim \mathrm{R}_m \equiv 10^3(L/u_0)$;

Bibliography:

The first detailed discussions of the diffusive decay of large-scale magnetic fields in astrophysical bodies are due to

Cowling, T.S. 1945, *Mon. Not. Roy. Astron. Soc.*, **105**, 166,
Wrubel, M.H. 1952, *Astrophys. J.*, **16**, 291.

On the numerical solution of algebraic eigenvalue problems in general, and on inverse iteration in particular, see

Golub, G.H., & Van Loan, C.F. 1989, *Matrix Computations (second edition)*, Baltimore: The Johns Hopkins University Press; §7.6.1 and 7.7.8.
Press, W.H., Teukolsky, S.A., Vetterling, W.T., & Flannery, B.P. 1992, *Numerical Recipes*, Second Ed., (Cambridge: Cambridge University Press), §11.7.

Chapter 11 of Press *et al.* also discusses numerical techniques that allow the computation of *all* eigenvalues of a matrix system, if you're interested in that. The book also contains a good introduction to the art of PDE discretization by finite differences. You will also find there a description of the leapfrog scheme for temporal dizcretisation. Among the numerous books discussing the calculation of macroscopic transport coefficients such as viscosity and electrical conductivity, starting from a microscopic point of view, our preference goes to an old classic:

Spitzer, L. Jr. 1962, *Physics of Fully Ionized Gases* (second ed.), New York: Wiley Interscience.

On the inference of large-scale magnetic fields on slowly rotating chemically peculiar stars, see

Deutsch, A.J. 1970, *Astrophys. J.*, **159**, 985,

Borra, E.F., Lanstreet, J.D., & Mestel, J. 1982, ARA&A, 20, 191,

Lanstreet, J.D. 2001, in *Magnetic field across the Hertzsprung-Russell diagram*, ASP Conf. Ser., vol. **248**, eds. G. Mathys, S.K. Solanki, and D.T. Wickramasinghe, 277,

and references therein. The flux rope dynamo of §7.2.2 is discussed and analyzed in

Vainshtein, S.I., & Zeldovich, Ya. B. 1972, *Soviet Physics Uspekhi*, **15**(2), 159.

Moffatt, H.K., & Proctor, M.R.E. 1985, *J. Fluid Mech.*, **154**, 493.

In addition, the mathematically-inclined will not want to miss the in-depth discussion of Stretch-Twist-Fold to be found in

Gilbert, A.D., 2003, "Dynamo theory", in *Handbook of Mathematical Fluid Dynamics*, vol. 2, eds. S. Frielander and D. Serre, Elsevier, 355-441,

from which Figure 7.4 was directly lifted. Magnetic flux expulsion from regions of closed streamlines is discussed in many textbooks dealing with magnetohydrodynamics. Analytical solutions for some specific cellular flows can be found for example in

Moffatt, H.K. 1978, *Magnetic field generation in electrically conducting fluids* (Cambridge: Cambridge University Press),

Parker, E.N. 1979, *Cosmical Magnetic Fields*, Oxford: Clarendon Press, chap. 16.

but for the first and last word on this topic, you should consult

Weiss, N.O. 1966, Proc. Roy. Soc. London A, 293, 310,

Rhines, P.B., & Young, W.R. 1983, J. Fluid Mech., 133, 133.

The Rhines & Young paper contains clean analytical examples of the two successive dissipative phases with characteristic timescales proportional to $\mathrm{R}_m^{1/3}$ and R_m, as discussed in §7.3.4. Flux expulsion is of course not restricted to 2-D flow; for a nice example in 3-D see

Galloway, D.J., & Proctor, M.R.E. 1983, *Geophys. Astrophys. Fluid Dyn.*, **24**, 109.

On the possible symmetrization of the Saturnian magnetic field by deep-seated differential rotation, see

Stevenson, D.J. 1982, *Geophys. Astrophys. Fluid Dyn.*, **21**, 113-127, Kirk, R.L., & Stevenson, D.J. 1987, *Astrophys. J.*, **316**, 836-846.

An insightful discussion of the symmetrization process in more general terms is that of

Rädler, K.-H. 1986, On the effect of differential rotation on axisymmetric and non-axisymmetric magnetic fields in cosmic bodies, in *Proceedings of the Joint Varenna-Abastumani International School and Workshop*, ESA Spec. Pub. SP-251, 569-574.

As for planetary magnetic field observations, a good recent overview is

Connerney, J.E.P. 1993, *J. Geophys. Res.*, **98**, 18659-18679.

On anti-dynamo theorems, see

Cowling, T.G. 1933, *Mon. Not. Roy. Astron. Soc.*, **94**, 39,
Bullard, E.C., & Gellman, H., *Phil. Trans. R. Soc. London A*, **247**, 213,
Zeldovich, Ya. B. 1956, J. Exp. and Theoretical Physics, 31, 154 [Russian];
Zeldovich, Ya. B., & Ruzmaikin, A.A. 1980, J. Exp. and Theoretical Physics, 78, 980 [Russian];

as well as pages 113-ff and 538–ff, respectively, of the books by Moffatt and Parker listed above. English translations of the last two papers are also reprinted in *Selected Works of Yakov Borisovich Zeldovich*, vol. 1 (ed. J.P. Ostriker, Princeton, 1992).

Chapter 8

Fast and slow dynamos

{chap:fastdyn}

It is nice to know that the computer understands the problem, but I would like to understand it too.

Attributed to E.P. Wigner

In light of the anti-dynamo theorems considered in §7.4, our next move should be obvious: we need to consider three-dimensional flows and magnetic fields. In addition, another relevant class of flow not excluded by the theorems is that of time-dependent flows. In this chapter we focus on one example of each of these two potentially promising flow classes. These will in fact provide us with our first working dynamos.

The cell flow solution of the preceding chapter also illustrated the potentially dangerous role of boundary conditions in mimicking dynamo action. To bypass this difficulty, the flows (and magnetic fields) we consider in this chapter are chosen to be **spatially periodic**. Dynamo action, if and when it occurs, is then evidently a property of the flows themselves, rather than a boundary effect. Although this takes us somewhat farther away from the astrophysical context, much is to be learned about magnetic field amplification in electrically conducting fluids using such simplified models.

8.1 The Roberts cell dynamo

{SRCell}

8.1.1 The Roberts cell

The Roberts cell is a spatially periodic, incompressible flow defined over a 2D domain $(x, y) \in [0, 2\pi]$ in terms of a stream function

$$\Psi(x, y) = \cos x + \sin y. \qquad (8.1)$$ {E3.rc1b}

so that

$$\mathbf{u}(x,y) = \frac{\partial\Psi(x,y)}{\partial y}\hat{\mathbf{e}}_x - \frac{\partial\Psi(x,y)}{\partial x}\hat{\mathbf{e}}_y + \Psi(x,y)\hat{\mathbf{e}}_z \tag{8.2}$$ {E3.rc1a}

Note that the flow velocity is independent of the z-coordinate, even though the flow has a non-zero z-component. Equations (8.2)–(8.1) describes a periodic array of counterrotating flow cells in the $[x, y]$ plane, with a z-component that changes sign from one cell to the next; the total flow is then a series of helices, which have the same kinetic helicity $\mathbf{h} = \mathbf{u} \times \nabla \times \mathbf{u}$ in each cell. The Roberts cell flow represents one example of a **Beltrami flows**, i.e., it satisfies the relation $\nabla \times \mathbf{u} = \alpha\mathbf{u}$, where α is a numerical constant. Such flows are **maximally helical**, in the sense that their vorticity ($\boldsymbol{\omega} \equiv \nabla \times \mathbf{u}$) is everywhere parallel to the flow, which maximizes helicity for a given flow speed.

Figure 8.1 shows one periodic "unit" of the the Roberts cell flow pattern. Note the presence of two **stagnation points** in the periodic cell, where four flow cells meet at $(x, y) = (0, 3\pi/2)$ and $(\pi, \pi/2)$. Let's first pause and consider why one should expect the Roberts cell to evade Cowling's and Zeldovich's theorems. First, note that this is not a planar flow in the sense demanded by Zeldovich's theorem, since we do have three non-vanishing flow components. However, the z-coordinate is ignorable in the sense of Cowling's theorem, since all flow components are independent of z. *If this flow is to evade Cowling's theorem and act as a dynamo, it must act on a magnetic field that is dependent on all three spatial coordinates.*

Consequently, we consider the inductive effects of this flow acting on a fully three dimensional magnetic field $\mathbf{B}(x, y, z, t)$. Since the flow speed is independent of z, we can expect solutions of the linear induction equation to be separable in z, i.e.:

$$\mathbf{B}(x,y,z,t) = \mathbf{b}(x,y,t)e^{ikz} \tag{8.3}$$ {E3.rc2}

where k is a (specified) wavevector in the z-direction, and the 2D magnetic amplitude $\mathbf{b}$ is now a complex quantity. We are still dealing with a fully 3D magnetic field, but the problem has been effectively reduced to two spatial dimensions (x, y), which represents a great computational advantage.

8.1.2 Dynamo action at last

From the dynamo point of view, the idea is again to look for solutions of the induction equations where the magnetic energy does not fall to zero as $t \to \infty$. In practice this means specifying k, as well as some weak field as an initial condition, and solve the 2D linear initial value problem for $\mathbf{b}(x, y, t)$

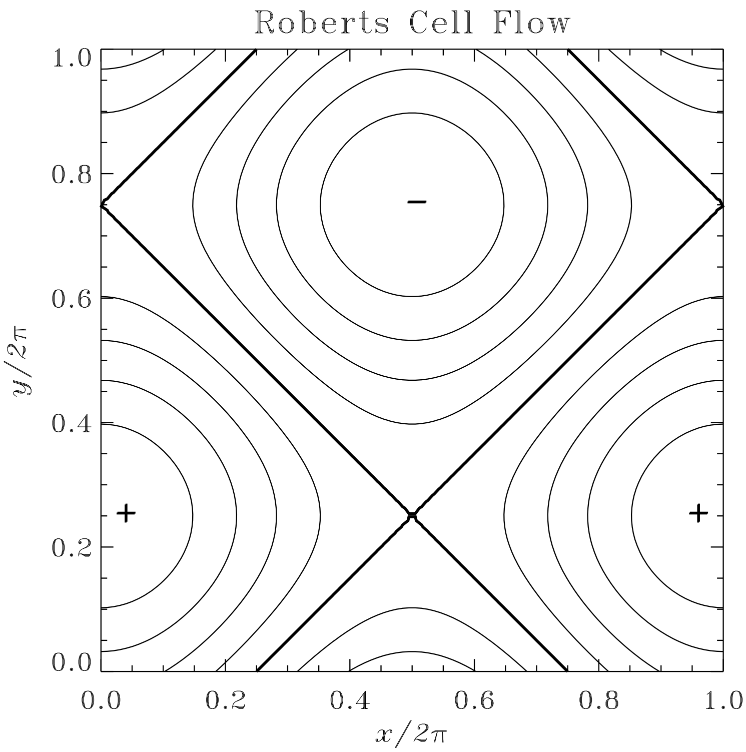

Figure 8.1: {F3.2} The Roberts cell flow. The flow is periodic in the $[x, y]$ plane, and independent of the z-coordinate (but $u_z \neq 0$!). Flow streamlines are shown projected in the $[x, y]$ plane, and the $+/-$ signs indicate the direction of the z-component of the flow. The thicker contour defines the network of **separatrix surfaces** in the flow, corresponding to cell boundaries. The u_z isocontours coincide with the projected streamlines.

resulting from the substitution of eq. (8.3) into the induction equation:

$$\frac{\partial \mathbf{b}}{\partial t} = (\mathbf{b} \cdot \nabla_{xy})\mathbf{u} - (\mathbf{u} \cdot \nabla_{xy})\mathbf{b} - iku_z\mathbf{b} + \mathrm{R}_m^{-1}(\nabla_{xy}^2\mathbf{b} - k^2\mathbf{b}) \qquad (8.4)$$ {E3.rc3}

subjected to periodic boundary conditions on $\mathbf{b}$. Here ∇_{xy} and ∇_{xy}^2 are the 2D gradient and Laplacian operators in the $[x, y]$ plane. As before we use as a time unit the turnover time τ_c, which is of order 2π here. All solutions described below were obtained numerically using second order finite difference in both space and time.

The time evolution of the can be divided into three more or less distinct phases, the first two being similar to the case of the 2D cellular flow considered in the preceding chapter: (1) quadratic growth of the magnetic energy for $t \lesssim \tau_c$; (2) flux expulsion for the subsequent few τ_c. However, and unlike the case considered in §7.3, for some values of k the third phase is one of *exponential growth* in the magnetic field (and energy).

Figure 8.2 shows a typical Roberts cell dynamo solution, here for $\mathrm{R}_m =$

10^2 and $k = 2$. What is plotted is the real part of the z-component of $\mathbf{b}(x, y, t)$, at time $t \gg \tau_c$. The thick dashed lines are again the separatrices of the flow. One immediately recognizes the workings of flux expulsion, in that very little magnetic flux is present near the center of the flow cells. Instead the field is concentrated in thin sheets parallel to the separatrix surfaces. Given our extensive discussion of flux expulsion in the preceding chapter, it should come as no surprise that the thickness of those sheets scales as $\mathrm{R}_m^{-1/2}$. For $t \gg \tau_c$, the field grows exponentially, but the shape of the "planform" remains fixed. In other words, even though we solved the induction equation as an initial value problem, the solution can be thought of as an eigensolution of the form $\mathbf{B}(x, y, z, t) = \mathbf{b}(x, y)e^{ikz+st}$, with $\mathrm{Re}(s) > 0$ and $\mathrm{Im}(s) = 0$.

In terms of the magnetic energy evolution, the growth rate s of $\mathbf{b}(x, y, t)$ is readily obtained by a linear least-squares fit to the $\log \mathcal{E}_\mathrm{B}$ vs t curves in the $t \gg \tau_c$ regime, or more formally defined as

$$s = \lim_{t\to\infty} \left[\frac{1}{2t} \log(\mathcal{E}_\mathrm{B}) \right] . \qquad (8.5)$$ {E3.rc4}

It turns out that the Roberts cell flows yields dynamo action (i.e., $s > 0$) over wide ranges of wavenumbers k and magnetic Reynolds number R_m. Figure 8.3 shows the variations in growth rates with k, for various values of R_m. The curves peak at a growth rate value k_max that gradually shifts to higher k as R_m increases. The largest growth rate is $k_\mathrm{max} \simeq 0.17$, and occurs at $\mathrm{R}_m \simeq 10$. It can be shown (see bibliography) that in the high R_m regimes the following scalings hold:

$$k_\mathrm{max} \propto \mathrm{R}_m^{1/2}, \qquad \mathrm{R}_m \gg 1 \ , \qquad (8.6)$$ {E3.rc5a}

$$s(k_\mathrm{max}) \propto \frac{\log(\log \mathrm{R}_m)}{\log \mathrm{R}_m}, \qquad \mathrm{R}_m \gg 1 \ . \qquad (8.7)$$ {E3.rc5b}

To understand the origin of these peculiar scaling relations, we need to take a closer look at the mechanism through which the magnetic field is amplified by the Roberts cell.

8.1.3 Exponential stretching and stagnation points {ssec:stagn}

Even cursory examination of Figure 8.2 suggests that magnetic field amplification in the Roberts cell is somehow associated with the network of separatrices and stagnation points. It will prove convenient in the foregoing analysis and discussion to first introduce new coordinates

$$x' = x - y \ , \qquad y' = x + y + \frac{3\pi}{2} \ , \qquad (8.8)$$ {E3.rc6a}

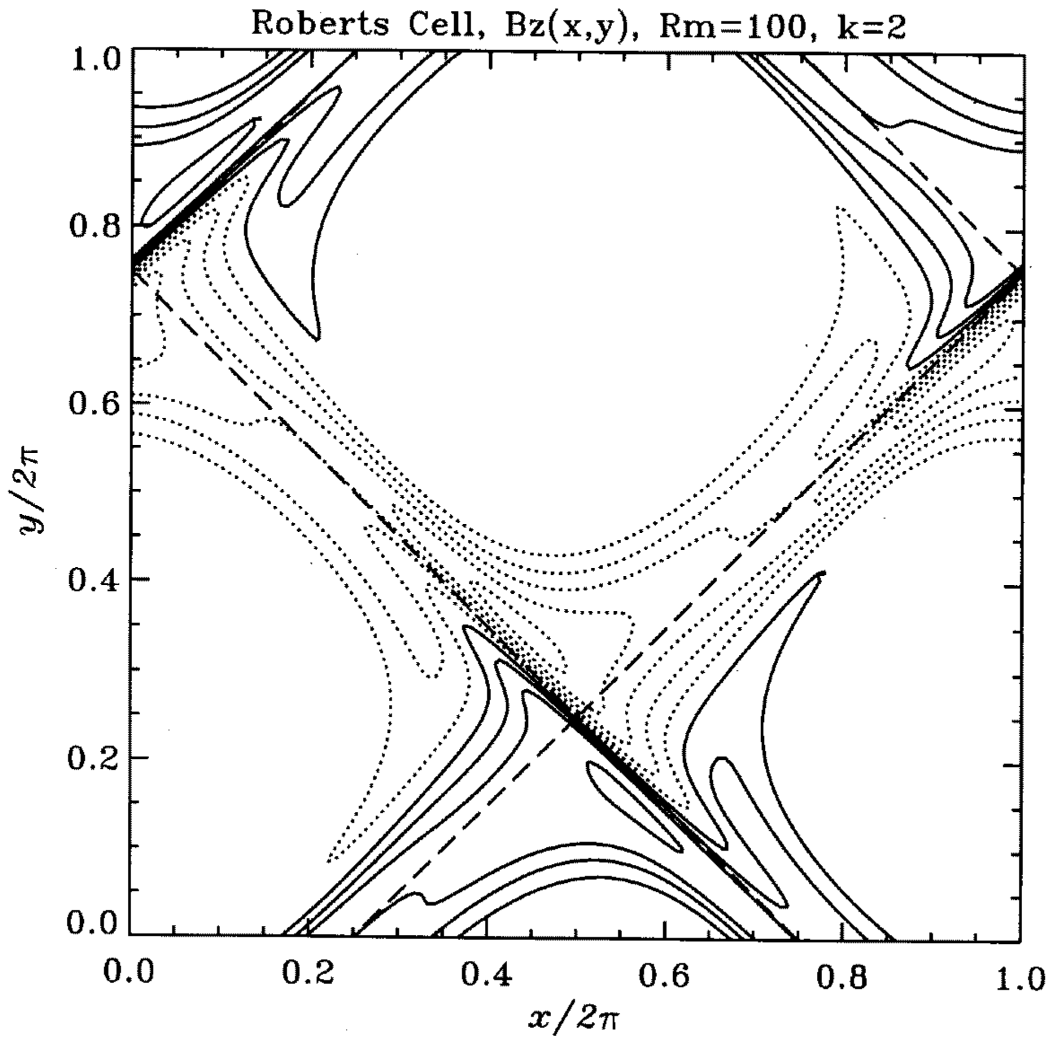

Figure 8.2: {F3.3} Isocontours for the z-component of the magnetic field in the $[x, y]$ plane, for a solutions with $\mathrm{R}_m = 100$ and $k = 2$, in the asymptotic regime $t \gg \tau_c$. The dashed straight lines indicate the separatrix surfaces of the underlying Roberts cell flow (see Fig. 8.1). Note the flux expulsion from the cell centers, and the concentration of the magnetic flux in thin sheets pressed against the separatrices. In the $t \gg \tau_c$ regime, the field grows exponentially but the shape of the planform is otherwise steady.

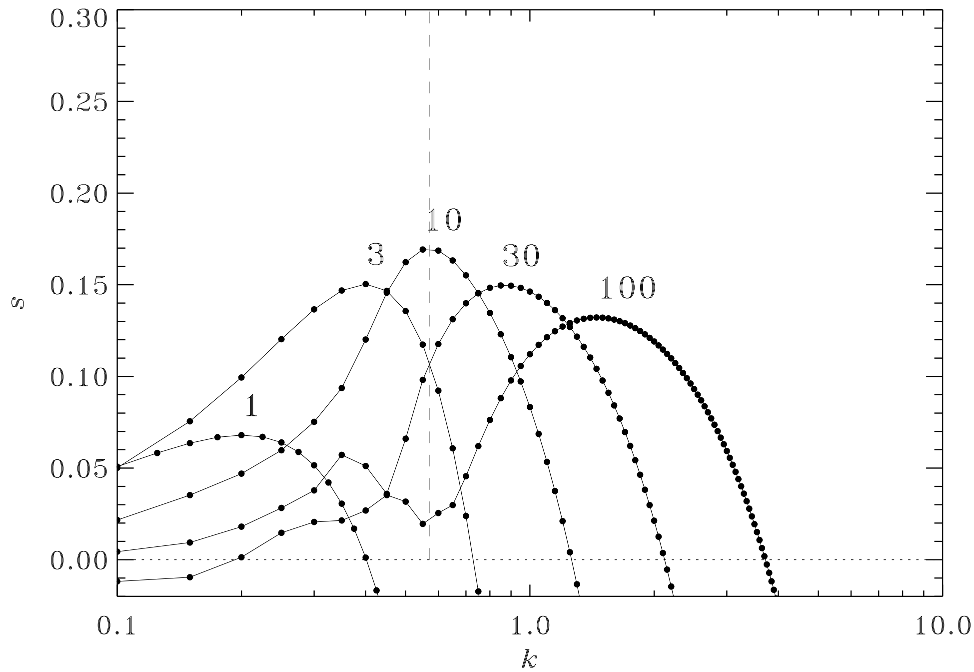

Figure 8.3: {F3.5} Growth rates of the magnetic energy in the Roberts cell, for sequences of solutions with increasing k and various values of R_m, as labeled near the maxima of the various curves. Growth typically occurs for a restricted range in k, and peaks at a value $k_{\max}$ that increases slowly with increasing R_m. Note however how the corresponding maximum growth rate decreases with increasing R_m. The small "dip" left of the main peaks for the high-R_m solutions is a real feature, although here it is not very well resolved in k.

corresponding to a $3\pi/2$ translation in the y-direction, followed by 45° rotation about the origin in the $[x, y]$ plane. The separatrices are now parallel to the coordinate lines $x' = n\pi$, $y' = n\pi$ $(n = 0, 1, ...)$, and the stream function has become

$$\Psi(x', y') = 2\sin(x')\sin(y') \ . \tag{8.9}$$ {E3.rc6b}

Close to the stagnation points, a good approximation to eq. (8.9) is

$$\Psi(x', y') \simeq 2x'y' \ , \qquad x', y' \ll 1 \tag{8.10}$$ {E3.rc6c}

which, if anything else, should now clarify why this is called a hyperbolic stagnation point... Consider now a fluid element flowing in the vicinity of this stagnation point. From a Lagrangian point of view its equations of motion are:

$$\frac{\partial x'}{\partial t} = u_{x'} = 2x' \ , \tag{8.11}$$ {E3.rc7a}

$$\frac{\partial y'}{\partial t} = u_{y'} = -2y' \ , \tag{8.12}$$ {E3.rc7b}

which immediately integrates to

$$x'(t) = x'_0 e^{2t}, \qquad y'(t) = y'_0 e^{-2t} \ , \tag{8.13}$$ {E3.rc7c}

where (x'_0, y'_0) is the location of the fluid element at $t = 0$. Evidently, the fluid element experiences **exponential stretching** in the x'-direction, and corresponding contraction in the y'-direction (since $\nabla \cdot \mathbf{u} = 0$!). Now, recall that in ideal MHD ($\mathrm{R}_m = \infty$) a magnetic fieldline obeys an equation identical to that of a line element, and that stretching leads to field amplification as per the mass conservation constraint (§7.2.1). Evidently stagnation point have quite a bit of potential, when it comes to amplifying exponentially a pre-existing magnetic field... providing that diffusion and destructive folding can be held at bay.

8.1.4 Mechanism of field amplification in the Roberts cell

We have shown that the Roberts cell can act as a dynamo, and that the field amplification mechanism is intimately tied to the presence of hyperbolic stagnation points at the cell corners. What we still need to do is figure out how the magnetic field generated by the Roberts cell manages to evade destructive folding.

We stick to the rotated Roberts cell used above, restrict ourselves to the $\mathrm{R}_m \gg 1$ regime, and pick up the field evolution after flux expulsion is completed and the magnetic field is concentrated in thin boundary layers (thickness $\propto \mathrm{R}_m^{-1/2}$) pressed against the separatrices (as on Fig. 8.2).

Consider a x'-directed magnetic fieldline crossing a vertical separatrix, as shown on Figure 8.4A (gray line labeled "a"). the y' component of the flow is positive on either side of the separatrix, and peaks on the separatrix. Consequently, the fieldline experiences stretching in the y'-direction ($a \to b \to c \to d$ on Fig. 8.4A). However, the induced y' component of the magnetic field changes sign across the separatrix, so that we seem to be heading towards our dreaded destructive folding. This is where the crucial role of the vertical (z) dimension becomes apparent. Figure 8.4B is a view of the same configuration in the $[x', z]$ plane, looking down onto the y' axis on part

A. At $t = 0$ the fieldlines have no component in the z-direction, but in view of the assumed e^{ikz} spatial dependency the x' component changes sign every half-wavelength k/π. Consider now the inductive action of the z-component of the velocity, which changes sign across the separatrix. After some time interval of order $k/(\pi u_z)$ the configuration of Fig. 8.4B will have evolved to that shown on part C. Observe what has happened: the fieldlines have been sheared in such a way that y'-components of the magnetic field of like signs have been brought in close proximity. Contrast this to the situation on part B, where magnetic footpoints in closest proximity have oppositely directed y'-components.

The end result of this process is that a y'-directed magnetic field is produced by shearing of the initial x'-directed field, with a phase shift in the z-direction such that destructive folding is avoided. Clearly, this requires *both* a z-component of velocity, *and* a z-dependency in the magnetic field. Either alone won't do the trick.

Now, the same reasoning evidently applies to a y'-directed magnetic fieldline crossing a horizontal separatrix: a x'-directed magnetic field will be induced. That magnetic field will be swept along the horizontal separatrix, get further amplified by exponential stretching as it zooms by the stagnation point, and continue along the vertical separatrix, where it can now serve as a seed field for the production of a y'-directed field. The dynamo "loop" is closed, at any time the rate of field production is proportional to the local field strength, and exponential growth of the field follows. The process works best if the half wavelength k/π is of order of the boundary layer thickness, which in fact is what leads to the scaling law given by eq. (8.6). The scaling for the growth rate (eq. (8.7)), in turn, is related to the time spent by a fluid element in the vicinity of the stagnation point.

8.2 Fast versus slow dynamos

One noteworthy aspect of the Roberts cell dynamo is the general decrease of the growth rates with increasing R_m (see Fig. 8.3). This is worrisome, because the $\mathrm{R}_m \to \infty$ limit is the one relevant to most astrophysically interesting circumstances. A dynamo exhibiting this property is called a **slow dynamo**, in contrast to a **fast dynamo**, which (by definition) retains a finite growth rate as $\mathrm{R}_m \to \infty$. In view of eq. (8.7), the Roberts cell is thus formally a slow dynamo. However the RHS of eq. (8.7) is such a slowly decreasing function of R_m that the Roberts cell is arguably the closest thing it could be to a fast dynamo... without formally being one. The distinction hinges on the profound differences between the strict mathematical case of $\mathrm{R}_m = \infty$ (ideal MHD), and the more physically relevant limit $\mathrm{R}_m \to \infty$.

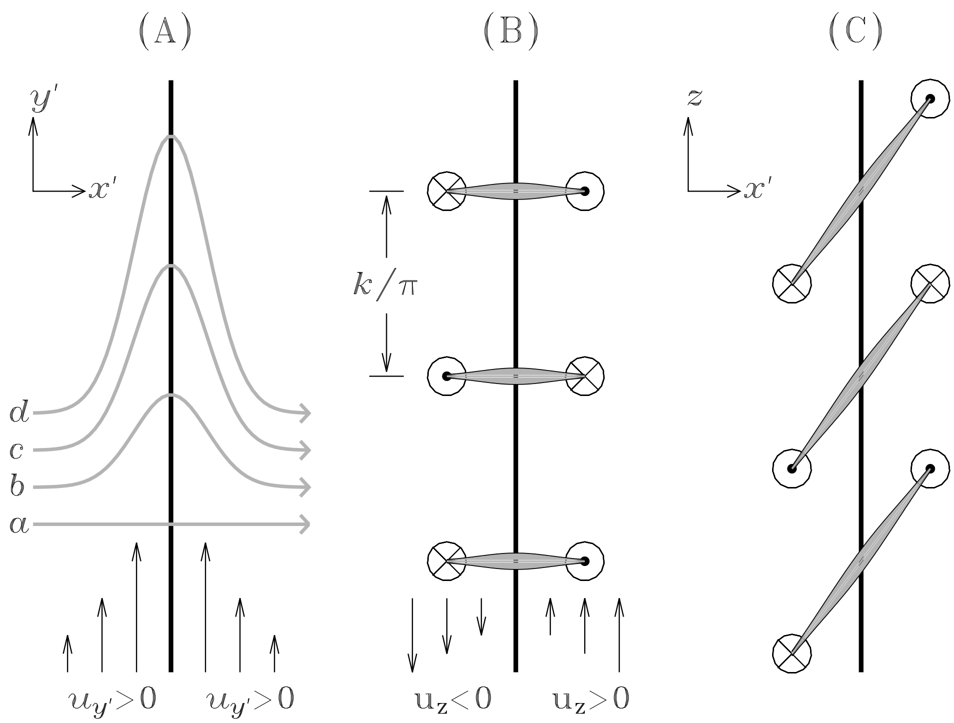

Figure 8.4: {F3.6} Mechanism of magnetic field amplification in the Roberts cell flow. The diagram is plotted in terms of the rotated $[x', y']$ Roberts cell. The thick vertical line is a separatrix surface, and the gray lines are magnetic fieldlines. Part (A) is a view in the horizontal plane $[x', y']$, and shows the production of a y'-directed magnetic component from an initially x'-directed magnetic field (line labeled "a"). Parts (B) and (C) are views in the $[x', z]$ plane looking down along the y' axis, and illustrate the phase shift in the z-direction of the y' magnetic component caused by the z-component of the velocity. The symbol $\odot$ ($\otimes$) indicates a magnetic field coming out (into) the plane of the page. Note on part (C) how footpoints of identical polarity are brought in close proximity, thus avoiding the destructive folding that would have otherwise characterized the situation depicted on part B in the $u_z = 0$ 2D case.

8.2.1 The singular limit $\mathrm{R}_m \to \infty$

From the physical point of view, the distinction between strict ideal MHD ($\eta = 0$) and the $\eta \to 0$ limit (or, equivalently, $\mathrm{R}_m \to \infty$) is a crucial one. One example will suffice. Recall that in the absence of dissipation **magnetic helicity** is a conserved quantity in any evolving magnetized fluid:

$$\frac{\mathrm{d}\mathcal{H}_\mathrm{B}}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\int_V \mathbf{A}\cdot\mathbf{B}\,\mathrm{d}V = 0, \tag{8.14}$$ {E3.fd1}

where $\mathbf{B} = \nabla \times \mathbf{A}$. Dynamo action, (in the sense of amplifying a weak initial field) is then clearly impossible except for the subset of initial fields having $\mathcal{H}_\mathrm{B} = 0$. This is a very stringent constraint on dynamo action! Go back now to the Roberts cell dynamo in the high-R_m regime. We saw that magnetic structures builp up on a horizontal length scale $\propto \mathrm{R}_m^{-1/2}$, and that the vertical wavelength of the fastest growing mode also decreases as $\mathrm{R}_m^{-1/2}$. *The inexorable shrinking of the length scales ensures that dissipation always continue to operate in the* $\mathrm{R}_m \to \infty$ *limit.*. This is why the Roberts cell dynamo can evade the constraint of helicity conservation. This is also why it is a slow dynamo. On the other hand, the Vainshtein & Zeldovich Stretch-Twist-Fold dynamo of §7.2, with its growth rate $\sigma = \ln 2$, is a fast dynamo since nothing prevents it from operating in the $\mathrm{R}_m \to \infty$ limit.

But is this really the case? In the flows we have considered up to now, the existence of dynamo action hinges on stretching winning over destructive folding; in the 2D cellular flow of §7.3, destructive folding won over stretching everywhere away from boundaries. In the Roberts cell, destructive folding is avoided only for vertical wavenumbers such that magnetic fields of like signs are brought together, minimizing dissipation. The STF dynamo actually combines stretching and constructive folding, such that folding *reinforces* stretching. The fact that destructive folding is avoided entirely is why the growth rate does not depend on R_m.

Well, upon further consideration it turns out that magnetic diffusivity must play a role in the STF rope dynamo after all. Diffusion comes in at two levels; the first and most obvious one is at the "knot" formed by the STF sequence. The second and less obvious arises from the fact that as one applies the STF operation n times, the resulting "flux rope" is in fact made up of n closely packed flux ropes, each of cross-section $\propto 2^{-n}$ times smaller than the original circular flux rope, so that the total cross-section looks more like a handful of spaghettis that it does a single monolithic flux rope of strength $\propto 2^n$. If one waits long enough, the magnetic length scale perpendicular to the loop axis shrinks to zero, so that even in the $\mathrm{R}_m \to \infty$ limit dissipation is bound to come into to play.

8.3 Fast dynamo action: the CP flow

{SCPF}

Knotty pasta notwithstanding, and despites its cartoon nature, the STF dynamo exemplifies the importance of constructive folding for fast dynamo action. However, it turns out to be exceedingly difficult (though possible, see bibliography) to find a smooth, continuous flow than achieves the requires stretch-twist-fold action. Fortunately, there exists wide classes of relatively simple (and analytically expressible) flows that, at least in the kinematic regime, achieves something essentially similar. In this section, we concentrate on one such flow, the so-called CP flow (for "Circularly Polarized"), as a prototypical flow yielding fast dynamo action.

8.3.1 The CP flow

The CP flow is nothing more than our familiar Roberts cell flow, with one important twist: an explicit time dependency is introduced in the flow:

$$u_x(x,y,t) = A\cos(y+\epsilon\sin\omega t)\ , \tag{8.15}$$ {E3.cp1a}

$$u_y(x,y,t) = C\sin(x+\epsilon\cos\omega t)\ , \tag{8.16}$$ {E3.cp1b}

$$u_z(x,y,t) = A\sin(y+\epsilon\sin\omega t) + C\cos(x+\epsilon\cos\omega t)\ . \tag{8.17}$$ {E3.cp1c}

What we have now is a periodic array of maximally helical counterotating flow cells, as on Fig. 8.1, with all cells "precessing" in unison in the $[x,y]$ plane along circular paths of radius ϵ, undergoing a full revolution in a time interval $2\pi/\omega$ [1]. In what follows we set $\omega = 1$, $\epsilon = 1$, $A = C = \sqrt{3/2}$, without any loss of generality.

The time-dependence of the flow turns out to have profound consequences for particle trajectories. Figure 8.5 shows the distances between two particles whose trajectories are being followed in the CP flow and in the the Roberts cell flow of §8.1, for the same starting positions and over the same time interval. The differences are striking. The short line element initially joining the two particles is stretched exponentially in the CP flow, but lengthens more or less linearly with time in the Roberts cell, as shown by the two fits on Figure 8.5.

Now, exponential streching, or, equivalently, exponential divergence of initially neighbouring trajectories, is the hallmark of **chaos**. Chaos has generated much hype (and occasional nonsense) in the literature, but the mathematical concept of chaos turns out to be extremely useful in analyzing flows for (potential) fast dynamo action.

[1] It is left as an (easy) exercise to verify that this is yet another Beltrami flow, and to figure out the form of the time-dependent stream function that describes it.

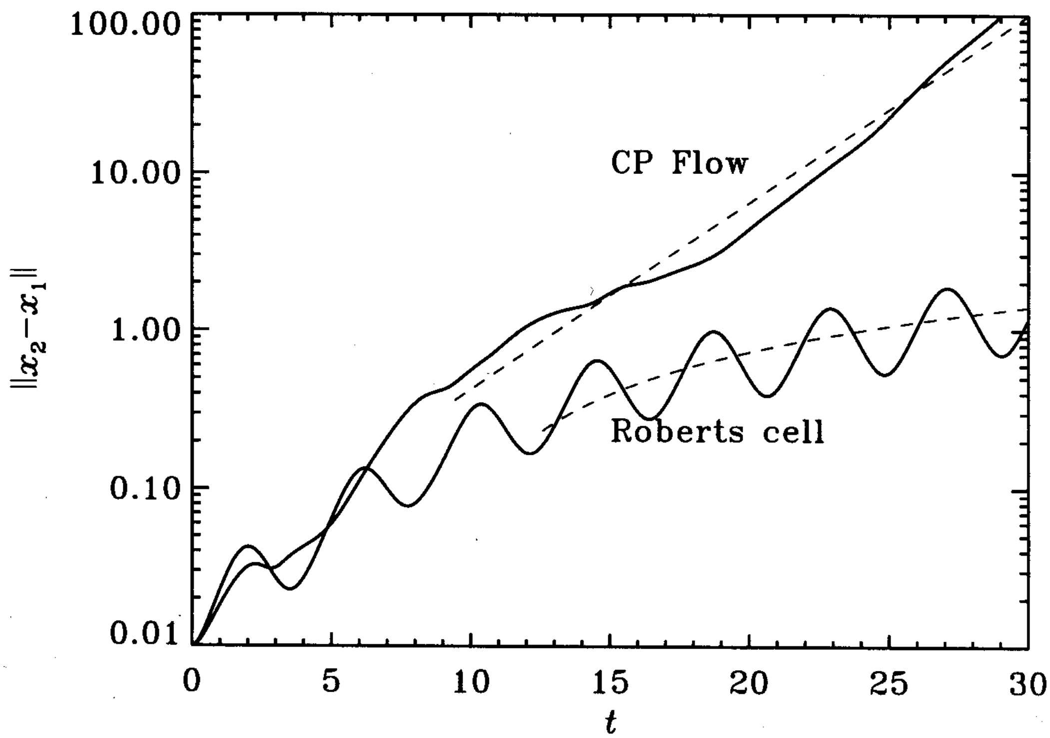

Figure 8.5: {F3.CP1} Stretching of a short line element initially located in the $z = 0$ plane, and "released" at $t = 0$ in the CP flow or Roberts cell. The two dashed lines are: (1) a linear least-squares fit of $\log ||x_2 - x_1||$ vs t to the CP flow curve, indicating *exponential* stretching; a linear least-squares of $||x_2 - x_1||$ vs t for the Roberts cell trajectory, indicating *linear* stretching.

8.3.2 Measures of chaos

The usefulness of chaos lies here with the fact that it can offer "measures" of fast dynamo action, without actually having to solve the induction equation! We now briefly consider two graphical measures of chaos: **Poincaré sections** and **Lyapunov exponents**.

A Poincaré section of the CP flow is shown on Figure 8.6. It is constructed by launching tracer particles at $z = 0$ (and $t = 0$), and following their trajectories as they are carried by the flow. At every 2π time interval, the position of the particle is plotted in the $[x, y]$ plane (modulo 2π in x and y, since most particles leave the original 2π-domain within which they were released). Some particles never venture too far away from their starting position in the $[x, y]$ plane. They end up tracing close curves (the so-called **KAM tori**, after Kolmogorov, Arnold, and Moser). Those curves, however distorted they may end up looking, identify regions of space where trajectories are integrable. Other particles, on the other hand, never return to

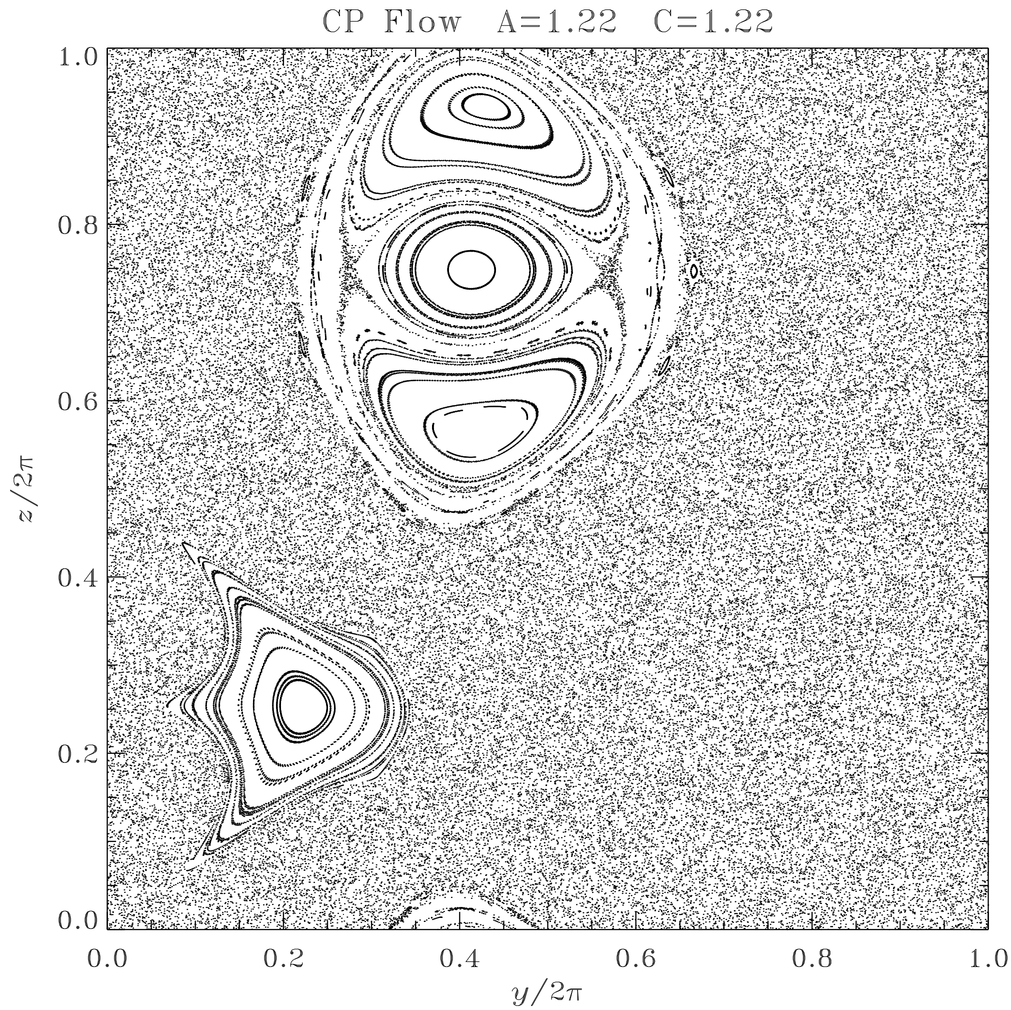

Figure 8.6: {F3.CP3} Poincaré section for the CP flow, for $\epsilon = 1$, $\omega = 1$, and $A = C = \sqrt{3/2}$. The plot is constructed by repeatedly "launching" particles at $z = 0$, $t = 0$, following their trajectories in time, and plotting their (projected) position (modulo 2π) in the $[x, y]$ plane at interval $\Delta t = 2\pi$. The flow is chaotic within the featureless "salt-and-pepper" regions, and integrable in regions threaded by close curves.

their starting position. If one waited long enough, one such particle would eventually come arbitrarily close to all points in the $[x, y]$ plane outside of the integrable regions. The corresponding particle trajectory is said to be **space filling**, and the associated particle motion *chaotic*. The region of the $[x, y]$ plane defined by the starting positions of all particles with space filling trajectories is called the *chaotic region* of the flow. [2].

Poincaré sections are useful to quickly eyeball the size of chaotic regions for a given set of flow parameters, but have little quantitative predictive values as to the potential efficiency of the flow as a dynamo. For this the **Lyapunov exponent** turns out to be a more useful quantity. The Lyapunov exponent is another fancy name for a rather simple concept; one, moreover that we encountered already on Fig. 8.5: the rate of exponential divergence of two neighbouring fluid element located at $\mathbf{x}_1$, $\mathbf{x}_2$ at $t = 0$ somewhere in the flow. The Lyapunov exponent λ_L can be (somewhat loosely) defined via

$$\ell(t) = \ell(0)\exp(\lambda_L t) \ , \tag{8.18}$$ {???}

where $\ell \equiv \|\mathbf{x}_2 - \mathbf{x}_1\|$ is the length of the *tangent vector* between the two fluid elements. Conceptually, Λ_L is nothing more than the slope of the dotted line on Fig. 8.5! Note however that, in general, Λ_L is likely to be a function of the position and relative orientations of $\mathbf{x}_1$ and $\mathbf{x}_2$. Strictly speaking Λ_L is mathematically defined in terms of stretching of an *infinitesimal* line element, located at a and oriented in direction e:

$$\lambda_L = \lim_{t\to\infty}\left(\frac{1}{2t}\log(\Lambda_{ij}e_i e_j)\right) \tag{8.19}$$ {???}

where

$$\Lambda_{ij} = \frac{\partial x_k}{\partial a_i}\frac{\partial x_k}{\partial a_j} \tag{8.20}$$ {???}

is the rate of strain tensor, so that $\Lambda_{ij}e_i e_j$ is the square of the stretching factor at time t. Because there are three independent possible directions in 3D space, one can compute *three* distinct Lyapunov exponents at any given point in the flow, and it can be shown that for an incompressible flow their sum is zero[3]. Now, the important thing about Lyapunov exponents is that $\lambda_L > 0$ somewhere in the flow indicates that this flow has chaotic regions.

Numerically, Lyapunov exponents are most often computed by repeatedly launching a set of particles defining a short line segment, advecting them over a finite time t, and measuring the rate of exponential stretching of that line

[2]Try sketching (or computing) a Poincaré section for the time-independent Roberts cell flow of §8.1. Does it differ much from Fig. 8.6?

[3]Demonstrate this result. Hint: start by thinking about what happens in the vicinity of a simple stagnation point, such as in §8.1.3

segment by summing the (exponentially increasing) distances between successive particles. The procedure is repeated for particle pairs with varying starting positions and relative orientations. Figure 8.7 shows a map $\Lambda_L(x, y)$ of the *largest* Lyapunov exponent for the CP flow as a function of position in the $[x, y]$ plane. The dark regions correspong to $\Lambda_L(x, y) \leq 0$, and the bright salt-and-pepper regions to $\Lambda_L(x, y) > 0$. The absolute largest Lyapunox exponent is $\Lambda_L^{\max} = 1.45$ here. Comparing Figure 8.7 to the Poincaré section on Fig. 8.6, one observes some definite similarities. For example, the integrable KAM regions on the Poincaré section correspond roughly to dark regions on the Lyapunov map. Yet the correspondence is far from perfect, illustrating the fact that trajectories and stretching of line elements are two related, but nonetheless distinct beasts.

8.3.3 Necessary conditions for fast dynamo action

Figures 8.6 and 8.7 might be aesthetically pleasing, but do they teach us anything *quantitative* about fast dynamo action? The Lyapunov exponent certainly does. There exists two important theorems stating that

1. A smooth flow cannot be a fast dynamo if $\lambda_L = 0$, so that $\lambda_L > 0$, or, equivalently, the existence of chaotic regions in the flow, is a necessary (although not sufficient) condition for fast dynamo action;

2. In the limit $\mathrm{R}_m \to \infty$, the largest Lyaponuv exponent of the flow is an upper bound on the dynamo growth rate.

Proofs of these theorems need not concern us here (but see bibliography). The theorems are indeed very useful information, in that they allows us to *rule out* fast dynamo action in many classes of flows. However, if one wants to *prove* fast dynamo action in a flow, at this writing there is no option but to integrate the induction equation. Time to return to the CP flow and do just that.

8.3.4 Fast dynamo action

Our search for dynamo action in the CP flow closely parallels what we did in the context of the Roberts cell. The time-dependency of the CP flow does not preclude the existence of solutions separable in z, so we again express the magnetic field via eq. (8.3), and solve the 2D induction equation (8.4) as an initial-boundary value problem, for specified vertical wavenumbers k. Periodic boundary conditions are again imposed on $\mathbf{b}(x, y, t)$. The time variation of the magnetic energy is again used as a test of dynamo action, and a

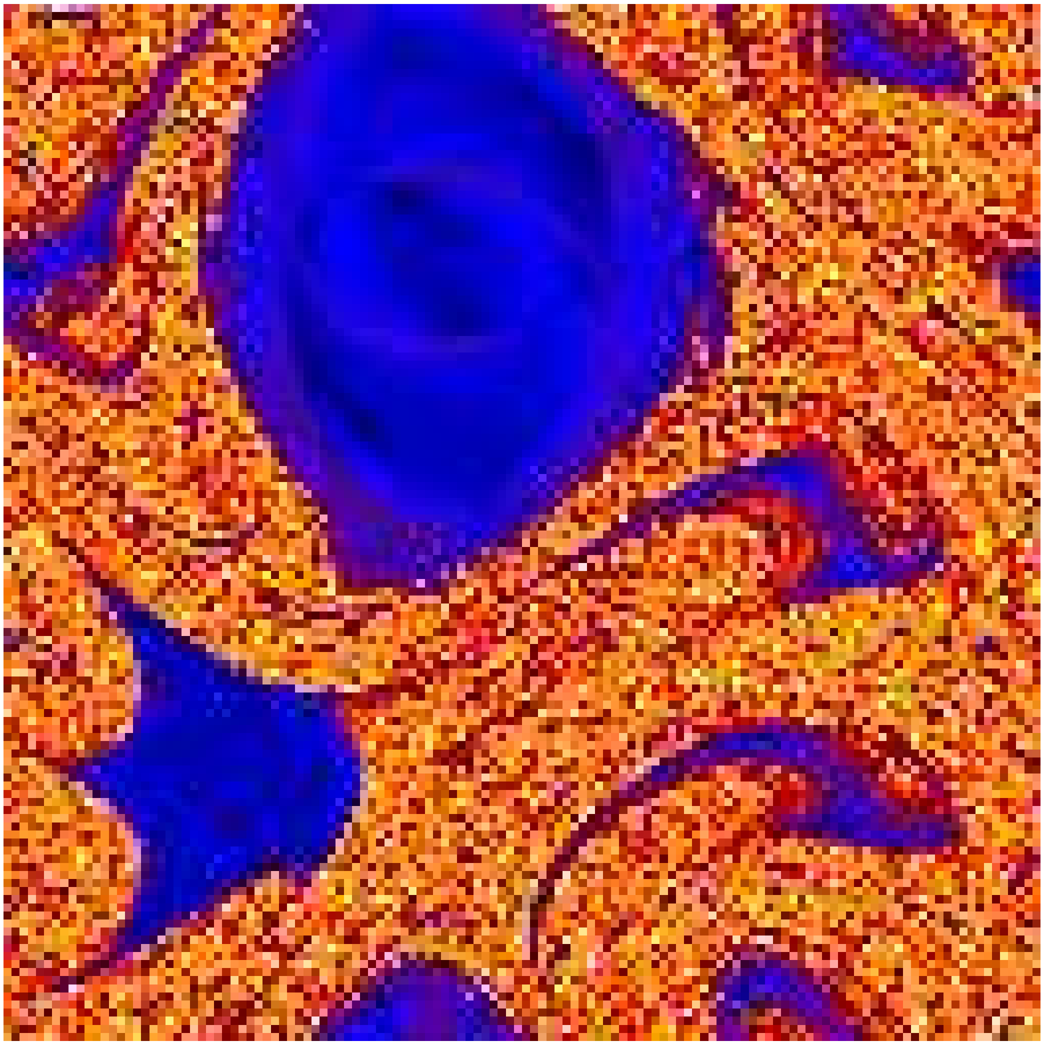

Figure 8.7: {F3.CP4} Finite time Lyaponuv exponent map for the CP flow with $\epsilon = 1$, $\omega = 1$, and $A = C = \sqrt{3/2}$. The dark part of the color scale correspond to negative Λ_L, and the brighter regions to $\Lambda_L > 0$. Compare this map to the Poincaré section of Figure 8.6.

growth rate is computed using eq. (8.5) for solutions exhibiting exponential growth in the $t \gg \tau_c$ regime.

As with the Roberts cell, dynamo action (i.e., positive growth rates $s(k, \mathrm{R}_m)$) occur in a finite range of vertical wavenumber k. Once again the phase of exponential growth sets in after a time of order of the turnover time. Figure 8.8 is similar in format to Fig. 8.2, and shows isocontours of the vertical magnetic field $b_z(x, y, t)$ in the phase of exponential growth, for a $\mathrm{R}_m = 2000$ solutions with $k = 0.57$. The solution is fully time-dependent, and its behavior is best appreciated by viewing it as an animation[4]. The solution is characterized by multiple sheets on intense magnetic field, of thickness once again $\propto \mathrm{R}_m^{-1/2}$.

The CP flow solution of Fig. 8.8 exhibits **spatial intermittency**. If one were to randomly choose a location somewhere in the $[x, y]$ plane, chances are good that only a weakish magnetic field would be found. In high-R_m solutions, strong fields are concentrated in small regions of the domain; in other words, their **filling factor** is small. This can be quantified by computing the **probability distribution function** (hereafter PDF) of the magnetic field strength, $f(|B_z|)$. This involves measuring B_z at every (x, y) mesh point in the solution domain, and simply counting how many mesh points have $|B_z|$ between values B and $B + \mathrm{d}B$. The result of such a procedure is shown in histogram form on Figure 8.9. The PDF shows a power-law tail at high field strengths,

$$f(|B_z|) \propto |B_z|^{-\gamma} , \qquad |B_z| \gtrsim 10^{-5} , \tag{8.21}$$ {E3.pltail}

spanning over four orders of magnitude in field strength, and with $\gamma \sim 1$ here. This indicates that strong field are still far more likely to be detected than if the magnetic field was simply a normally-distributed random variable (for example)[5]. The fact that the PDF's logarithmic slope is flatter than -2 indicates that the *largest* local field strength found in the domain will always dominate the computation of the spatially-averaged field strength[6].

The CP flow dynamo solutions also exhibit **temporal intermittency**; if one sits at one specific point (x, y) point in the domain and measures B_z at subsequent time steps, a weak B_z is measured most of the time, and only occasionally are large values detected. Once again the PDF shows a power-law tail with slope flatter than -2 indicating that a temporal average of B_z at one location will always be dominated by the largest B_z measured to date[7].

[4]which you can do, of course, on the course's Web Page, and for a few R_m values, moreover...

[5]What would be the shape of a Gaussian PDF on a log-log plot such as Fig. 8.9?

[6]Prove this; it begins with writing down an certain integral involving the PDF that yields the average value the variable of interest.

[7]Could you make an educated guess at the value of the logarithmic slope of this temporal PDF?

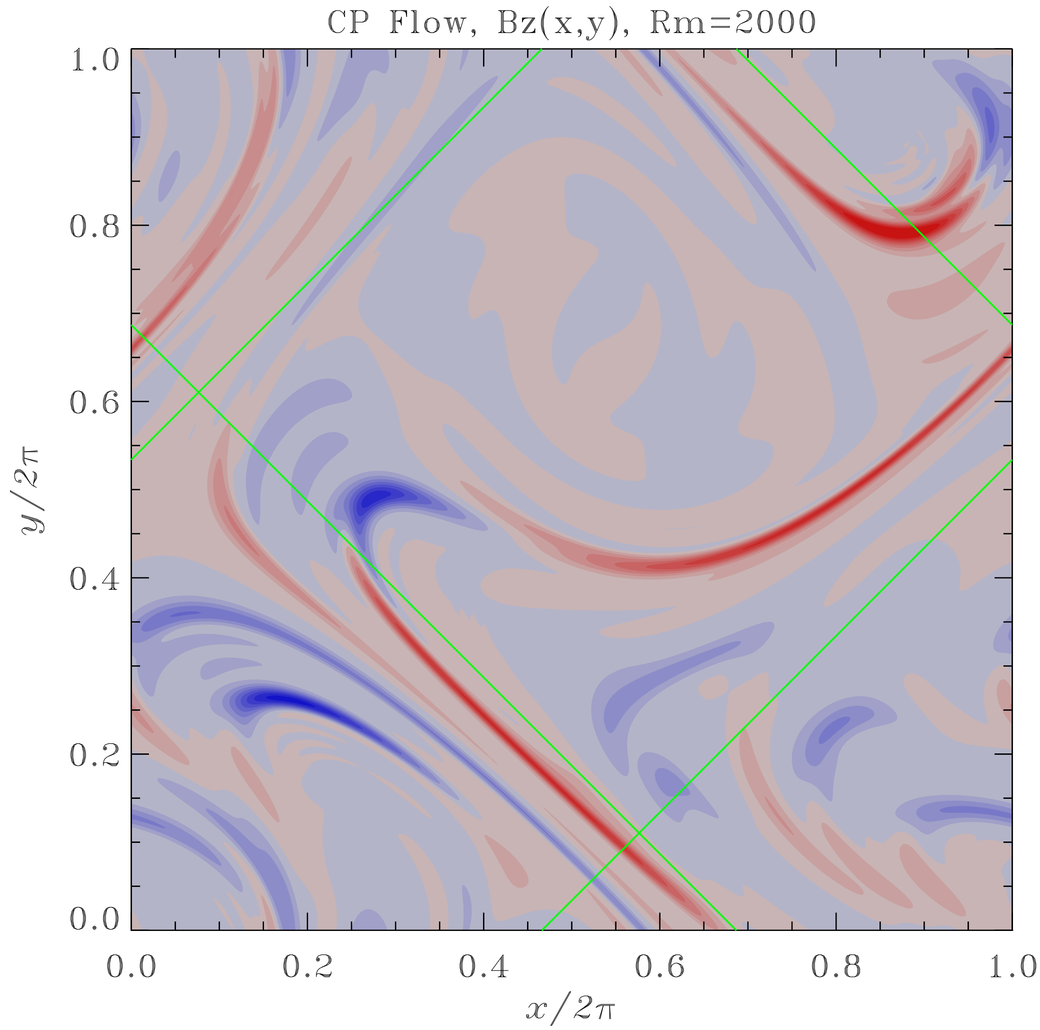

Figure 8.8: {F3.CPF5} Snapshot of the z-component of the magnetic field in the $[x, y]$ plane, for a CP Flow solution with $\mathrm{R}_m = 2000$ and $k = 0.57$, in the asymptotic regime $t \gg \tau_c$. The color scale codes the field strength (gray-to-blue is negative, gray-to-red positive). The green straight lines indicate the separatrix surfaces of the underlying Roberts cell flow (see Fig. 8.1). Unlike the Roberts cell solution of Fig. 8.2, this is a strongly time-dependent solution, although still exhibiting overall exponential growth of the magnetic field.

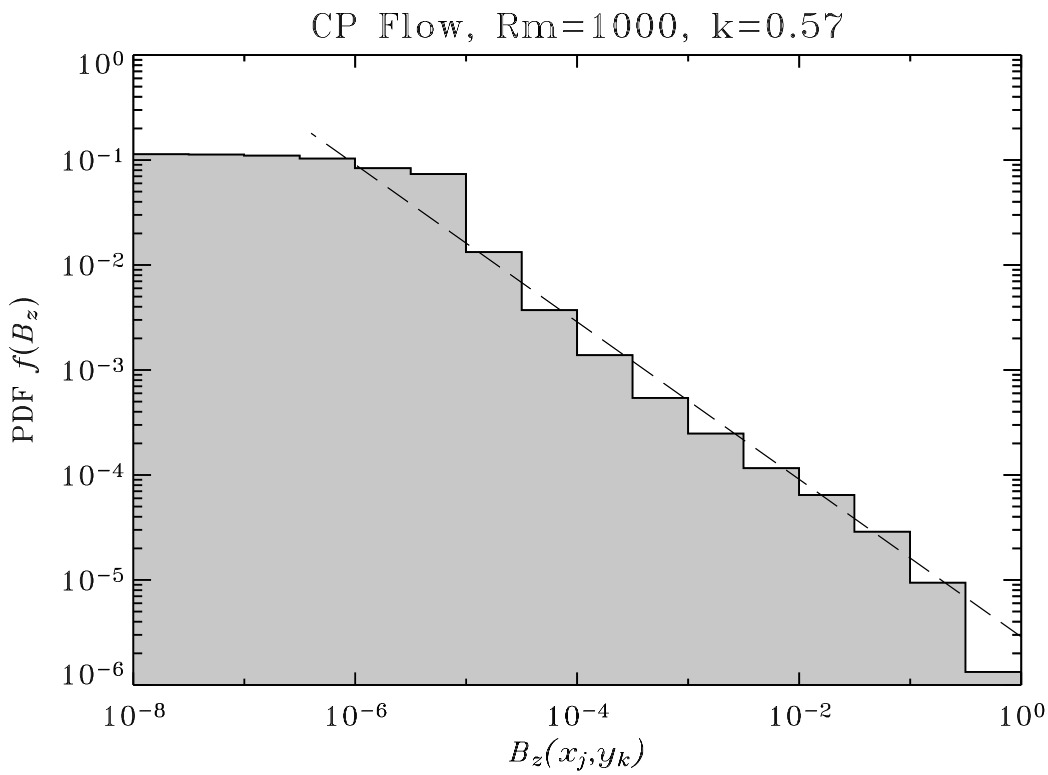

Figure 8.9: {`F3.histo`} Probability distribution function for the (unsigned) strength of the z-component of the magnetic field, for a $\mathrm{R}_m = 10^3$, $k = 0.57$ CP flow dynamo. The peak field strength has been normalized to a value of unity. Note the power-law tail at large field strength (straight line in this log-log plot, with slope ~ -0.75).

Unlike in the Roberts cell, the range of k yielding dynamo action does *not* shift significantly to higher k as R_m is increased, and in the high R_m regime the corresponding maximum growth rate $k_{\max}$ does *not* decrease with increasing R_m (see Fig. 8.3). In the CP flow considered here ($A = C = \sqrt{3/2}$, $\omega = 1$, $\epsilon = 1$), $k_{\max} \simeq 0.57$, with $s(k_{\max}) \simeq 0.3$ for $\mathrm{R}_m \gtrsim 10^2$, as shown on Figure 8.10 (solid line). Figure 8.10 suggests (but does not rigorously prove!) that the CP flow acts as a fast dynamo, since by all appearances

$$\lim_{\mathrm{R}_m \to \infty} s(k_{\max}) > 0 \ . \tag{8.22}$$ {`???`}

8.3.5 Magnetic flux versus magnetic energy {`S3.flux`}

With the CP flow, we definitely have a pretty good dynamo on our hands. But how are those dynamo solutions to be related to the Sun (or other

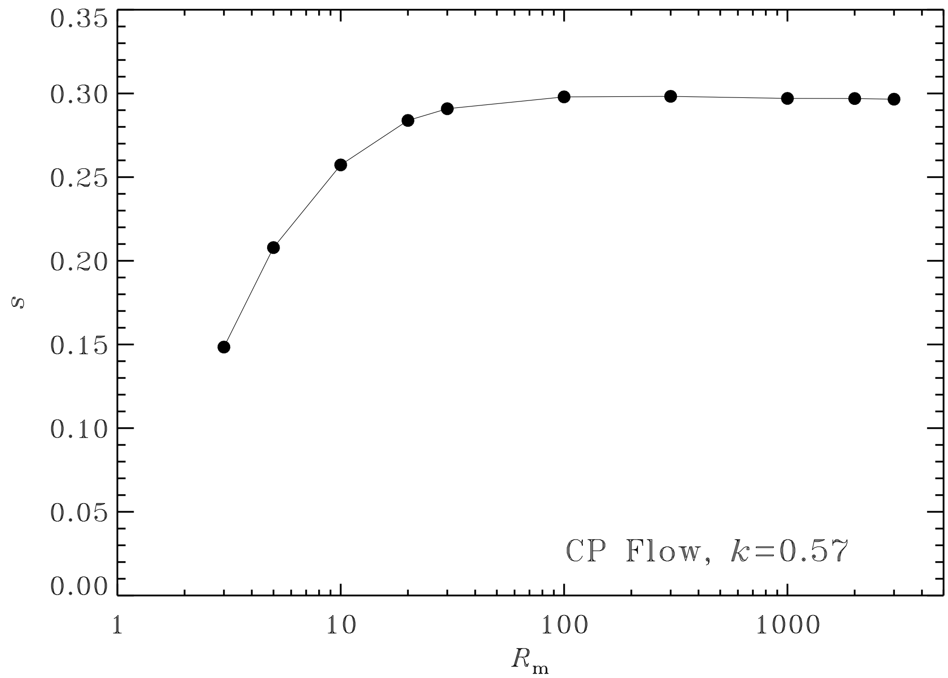

Figure 8.10: {F3.CPF6} Growth rate of $k = 0.57$ CP flow dynamo solutions, plotted as a function of the magnetic Reynolds number (solid line). The constancy of the growth rate in the high-R_m regime suggests (but does not strictly prove) that this dynamo is fast.

astrophysical bodies)? So far we have concentrated on the magnetic energy as a measure of dynamo action, but in the astrophysical context **magnetic flux** is also important. Consider the following two (related) measures of magnetic flux:

$$\Phi = \mid \langle \mathbf{B} \rangle \mid \ , \qquad F = \langle \mid \mathbf{B} \mid \rangle \ , \tag{8.23}$$ {E3.30}

where the angular brackets indicate some sort of suitable spatial average over the whole computational domain. The quantity Φ is nothing but the average magnetic flux, while F is the average **unsigned flux**. Under this notation the magnetic energy can evidently be written as $\mathcal{E}_{\mathrm{B}} = \langle \mid \mathbf{B} \mid^2 \rangle$. Consider now the scaling of the two following ratios as a function of the magnetic Reynolds number:

$$\mathcal{R}_1 = \frac{\mathcal{E}_{\mathrm{B}}}{\Phi^2} \propto \mathrm{R}_m^n \ , \tag{8.24}$$ {E3.31a}

$$\mathcal{R}_2 = \frac{F^2}{\Phi^2} \propto \mathrm{R}_m^\kappa \ . \tag{8.25}$$ {E3.31b}

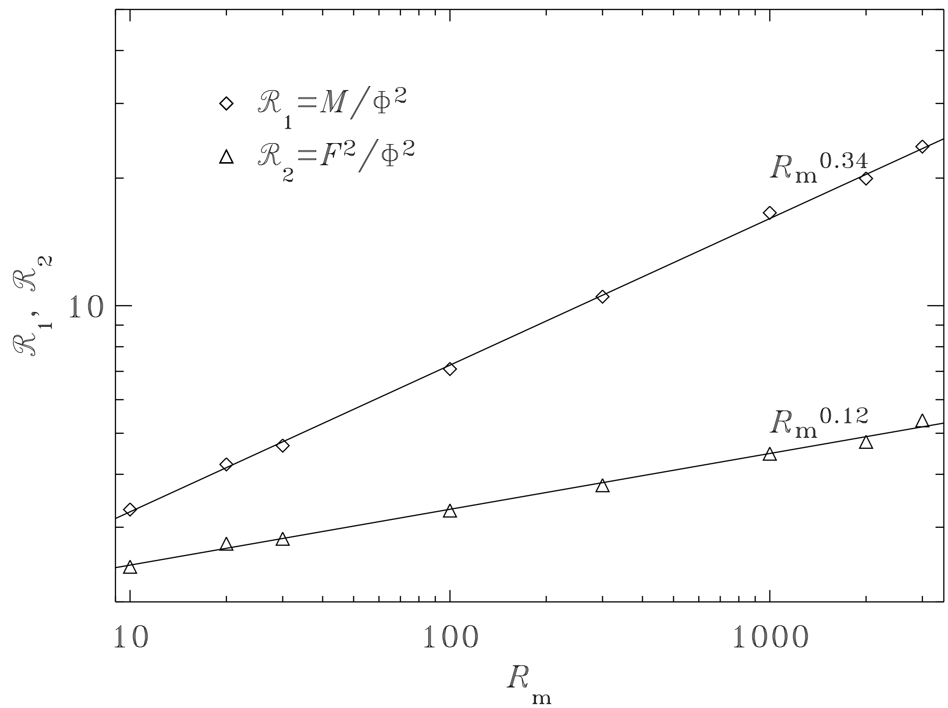

Figure 8.11: {F3.CPF7} Variations with R_m of the two ratios defined in eqs. (8.24)–(8.25). Least squares fits (solid lines) yield power law exponents $n = 0.35$ and $\kappa = 0.13$.

A little reflection will reveal that a large value of $\mathcal{R}_1$ indicates that the magnetic field is concentrated in a small total fractional area of the domain, i.e., the **filling factor** is much smaller than unity[8]. The ratio $\mathcal{R}_2$, on the other hand, is indicative of the dynamo's ability to generate a net signed flux. The exponent κ measures the level of **folding** in the solution; large values of κ indicate that while the dynamo may be vigorously producing magnetic flux on small spatial scales, it does so in a manner such that very little *net* flux is being generated on the spatial scale of the computational domain. Figure 8.11 shows the variations with R_m of the two ratios defined above. Least squares fits to the curves yields $n = 0.35$ and $\kappa = 0.13$. Positive values for the exponents κ and n indicate that the CP flow dynamo is relatively inefficient at producing magnetic flux in the high R_m regime, and even less efficient at producing net signed flux. While other flows yielding fast dynamo actions lead to different values for these exponents, in general they seem to always turn out positive, with $\kappa < n$, so that the (relative) inability to

[8]If you can't figure it out try this: take a magnetic field of strength B_1 crossing a surface area A_1; now consider a more intense magnetic field, of strength $B_2 = 4B_1$, concentrated in one quarter of the area A_1; calculate $\mathcal{E}_{\rm B}$, Φ, and $\mathcal{R}_1$... get it?

produce net signed flux seems to be a generic property of fast dynamos in the high-R_m regime.

8.3.6 Fast dynamo action in the nonlinear regime

We conclude this section by a brief discussion of fast dynamo action in the nonlinear regime. Evidently the exponential growth of the magnetic field will be arrested once the Lorentz force becomes large enough to alter the original CP flow. What might the nature of the backreaction on $\mathbf{u}$ look like?

Naively, one might think that the Lorentz force will simply reduce the amplitude of the flow components, leaving the overall geometry of the flow more or less unaffected, i.e., $\mathbf{u}_1 \simeq \mathbf{u}_0$. That this cannot be the case becomes obvious upon recalling that in the high R_m regimes the eigenfunction is characterized by magnetic structures of typical thickness $\propto R_m^{-1/2}$, while the flow has a typical length scale $\sim 2\pi$ in our dimensionless units. The extreme disparity between these two length scales in the high-R_m regime suggests that the saturation of the dynamo-generated magnetic field will involve alterations of the flow field on small spatial scales, so that a flow very much different from the original CP flow is likely to develop in the nonlinear regime.

That this is indeed what happens was was nicely demonstrated some years ago by F. Cattaneo and collaborators (see references in bibliography), who computed simplified nonlinear solutions of dynamo action in a suitably forced CP flow. They could show that

1. the r.m.s. flow velocity in nonlinear regime is comparable to that in the original CP flow;
2. magnetic dissipation actually *decreases* in the nonlinear regime;
3. dynamo action is suppressed by the disappearance of chaotic trajectories in the nonlinear flow.

8.4 The solar small-scale magnetic field {Ssolmag}

Of course, the problem with small-scale solar magnetic fields is precisely that—they *are* small-scale. And being small-scale makes them very difficult to resolve. Being unresolved, in truth there is not a lot one can discover about them, even with current state-of-the-art high precision spectropolarimetry.

All flows yielding dynamo actions that have been considered up to now are very artificial, and are arguably more akin to malfunctioning washing machines than any sensible astrophysical object. Nonetheless some of the things we have learned do carry over to more realistic circumstances. Most importantly, fast dynamos

1. produce flux concentrations on scales $\propto \mathrm{R}_m^{-1/2}$;

2. produce little or no mean-field, i.e., signed magnetic flux on a spatial scale comparable to the size of the system;

3. require chaotic flow trajectories to operate.

As a kind of proof of these sweeping statements, consider Figure 8.12 herein. It is a snapshot of a numerical simulation of dynamo action in a stratified, thermally-driven turbulent fluid being heated from below, and spatially periodic in the horizontal directions. This flow acts as a vigorous nonlinear fast dynamo, with a ratio of magnetic to kinetic energy of about 20%. The Figure shows a snapshot of the vertical magnetic field component $B_z(x, y)$ essentially at the top of the simulation box[9].

Thermally convecting flows in a stratified background have long been known to be characterized by cells of broad upwellings of warm fluid. These cells have a horizontal size set by, among other things, the density scale height within the box; On the other hand, the downwelling of cold fluid needed to satisfy mass conservation end up being concentrated in a network of narrow lanes at the boundaries between adjacent upwelling cells. This asymmetry is due to the vertical pressure and density gradient in the box: rising fluid expands laterally into the lower density layers above, and descending fluid is compressed laterally in the higher density layers below. Near the top of the simulation box, this leads to the concentration of magnetic structures in the downwelling lanes, as they are continuously being swept horizontally away from the centers of upwelling cells. This is the origin of the cellular pattern so striking on Fig. 8.12.

While this flow is far more complex (spatially and temporally) than the Roberts cell or CP flow, is exhibits some of the characteristics we have already encountered in the context of these simpler flows:

1. The magnetic field is highly intermittent, both spatially and temporally.

2. Magnetic flux concentrations are found on scales $\propto \mathrm{R}_m^{-1/2}$;

3. little or no **mean** magnetic field is produced on the scale of the computational box.

The fundamental physical link between this MHD simulation and the CP flow is the presence of chaotic trajectories in the flow, which in both cases is the culprit behind fast dynamo action.

[9]...and, as usual, you can view an animation of this simulation on the course Web Page.

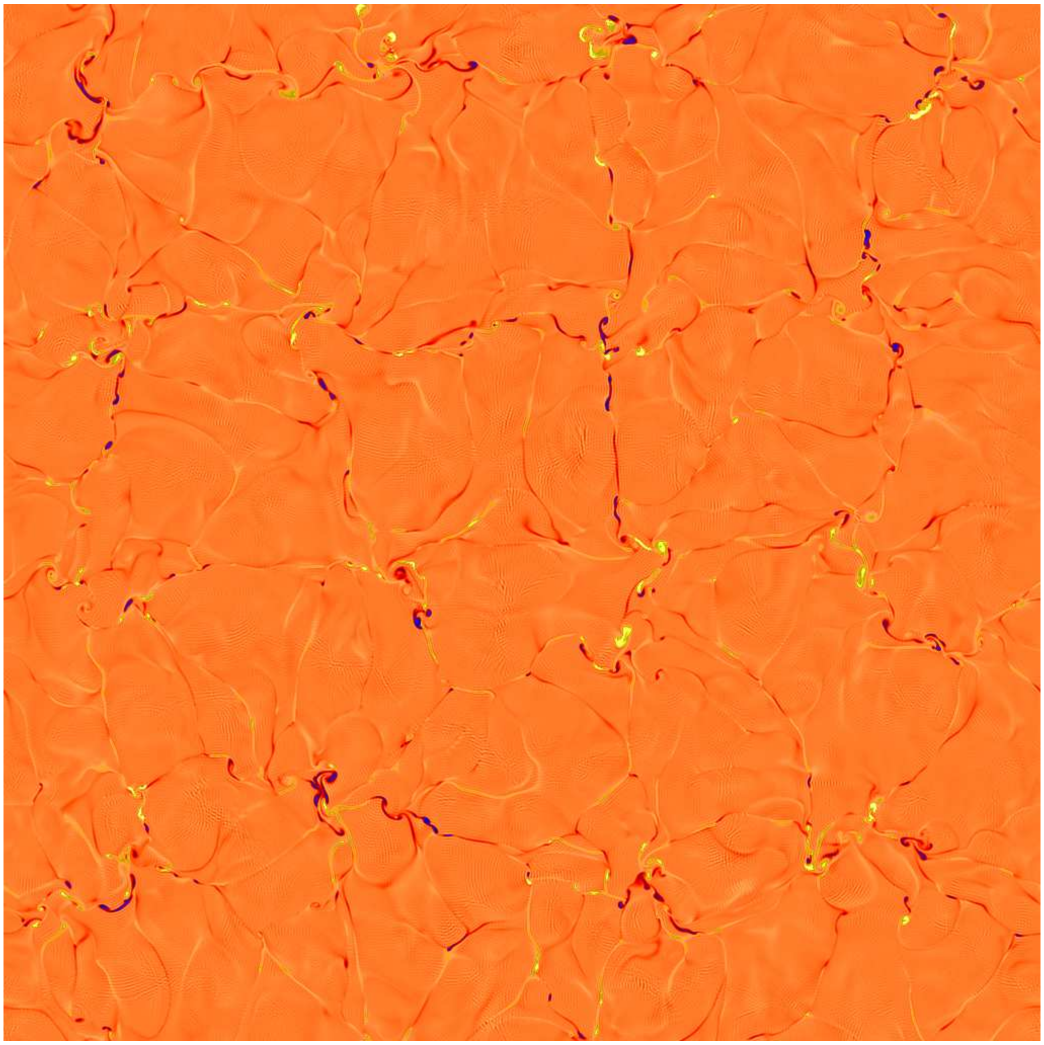

Figure 8.12: {F3.SMAG1} Snapshot of the top "horizontal" $[x, y]$ plane of a MHD numerical simulation of thermally-driven stratified turbulent convection in a box of aspect ratio $x : y : z = 10 : 10 : 1$, at a Viscous Reynolds number of 245 and $\mathrm{R}_m = 1225$. The simulation uses a pseudo-spectral spatial discretrization scheme, with 1024 collocation points in the x and y directions, and 97 in z. The color scale encodes the vertical (z) component of the magnetic field (orange-to-yellow is positive B_z, orange-to-blue negative). Numerical simulation results kindly provided by F. Cattaneo, University of Chicago.

Now consider figure 8.13 which shows a high-resolution magnetogram of a small piece of the solar photosphere, far away from sunspots or active regions. Note how the magnetic field is spatially very intermittent, and seems to have no marked preference for negative (black) or positive (white), except perhaps for the plage-like structure in the upper left corner. Here also the magnetic field is very intermittent, both spatially and temporally[10]. This is all *qualitatively* similar to the field distribution characterizing Fig. 8.12.

Fast dynamo action therefore offers an attractive explanation for the *small-scale* solar magnetic fields. Nice and fine, but the Sun also has a fairly well-defined large-scale component, for which something else than fast dynamo action must then be invoked. It turns out that the turbulent nature of the flow in the solar convective envelope can still do the trick, but to examine this we will need to adopt as statistical approach to turbulence and to the associated flow-field interactions. This is the focus of the following chapter.

Problems:

1. Calculate $\nabla \cdot \mathbf{u}$ and $\nabla \times \mathbf{u}$ for the Roberts Cell flow. Confirm that it is a Beltrami flow, in the sense discussed in §8.1.

2. This problem aims at getting you to investigate in more detail what can happen to a magnetic field in the vicinity of a stagnation point. Consider the 2D cartesian incompressible flow defined by the stream function

$$\Psi(x, y) = u_0 xy \tag{8.26}$$ {E3.11}

so that $u_x(x) = u_0 x$ and $u_y(y) = -u_0 y$ (note that the quantity u_0 has then units of s^{-1}!). We now want to consider the inductive action of this flow on a purely horizontal magnetic field, held fixed at values of $+B_0$ and $-B_0$, at $y = +L$ and $-L$ respectively. Evidently, this flow will tend to push the magnetic field towards the x-axis, where dissipation will occur since the field is oppositely directed on either side of the x-axis.

   (a) Show that the above flow has a stagnation point at the origin, and that its divergence is zero.

   (b) Show that in view of the imposed boundary condition, B_x can only be a function of the y coordinate everywhere in the domain $-L \leq y \leq +L$, $-\infty \leq x \leq \infty$.

[10] A magnetogram animation can be viewed on the course web page, and illustrates quite well the temporally intermittent nature of the solar small-scale magnetic field.

(c) Show that under these circumstances the x-component of the induction equation reduces to

$$\frac{1}{u_0}\frac{\partial B_x}{\partial t} = B_x + y\frac{\partial B_x}{\partial y} + \frac{\eta}{u_0 L}\frac{\partial^2 B_x}{\partial y^2},$$

where lengths are expressed in units of L.

(d) Show now that this equation accepts steady-state solutions of the form

$$B_x(y) = C\exp\left(-\alpha y^2\right) + D\exp\left(-\alpha y^2\right)\int_0^y \exp\left(\alpha (y')^2\right) \mathrm{d}y'$$

where the parameter $\alpha = u_0 L^2/2\eta \equiv \mathrm{R}_m/2$ controls the relative importance of magnetic dissipation, as measured by the usual magnetic Reynolds number $\mathrm{R}_m = u_0 L/\eta$, and C and D are integration constants.

(e) Show that the assumed boundary conditions imply that $C = 0$ here;

(f) Now show that the thickness of the current sheet forming in the vicinity of $y = 0$ scales as $1/\sqrt{\alpha}$;

(g) Evaluate numerically the integral on the above solution for $B_x(y)$ and plot the variation of B_x as a function of y, for values of $\alpha = 10, 100$ and 10^3.

(h) Finally, compute the magnitude of the electric current in the z-direction, and show that the rate of energy dissipation is independent of the assumed value of η. Explain this physically.

3. The so-called ABC flow is another long-time candidate for fast dynamo action. It is a steady periodic flow in cartesian geometry, defined as

$$\mathbf{u}(x, y, z) = (A\sin z + C\cos y, B\sin x + A\cos z, C\sin y + B\cos x)$$

(a) Verify whether or not this is a Beltrami flow;

(b) Find the position(s) of the stagnation point(s) in the flow, for the specific case $A = B = C = 1$.

(c) Calculate a Poincaré section for this flow, using now parameter values A = 1+(your birth month/12), B = 1+(your birth day/30), $C = 1$. This involves repeatedly launching a particle somewhere on the $z = 0$ plane, and plotting its position at every crossing of $2\pi n$ planes in the z-direction ($n = 1, 2, ...$). Is this flow chaotic?

4. The flow near the 3D stagnation points in the ABC flow can be approximated in cylindrical polar coordinates (r, θ, z) by

$$\mathbf{u} = (\alpha r/2, 0, -\alpha z) \ ,$$

with $\alpha = \pm\sqrt{2}$.

   (a) Calculate the the three Lyapunov exponents for $\alpha = +\sqrt{2}$ and $\alpha = -\sqrt{2}$, and show that in both cases their sum is zero.

   (b) Obtain a solution to the steady ($\partial/\partial t = 0$) form of the induction equation, with $\mathbf{u}$ given by the above expression.

   (c) On the basis of your solution, where would you expect to find magnetic fields in the flow?

   (d) Again on the basis of your solution, estimate a length scale characterizing the thickness of the magnetic structures present in the solutions. How does this characteristic length scale with the magnetic Reynolds number?

5. This problem gets you to compute and compare the PDFs associated with the CP flow solution discussed in detail in this chapter, and the numerical simulation of Cattaneo *et al.* discussed in §8.4. First go to the Course Web Page, and grab the two data files containing snapshots of $B_z(x, y)$ for a CP flow solution, and for the turbulent dynamo solution plotted on Fig. 8.12.

   (a) Compute the mean signed and unsigned fluxes for the two solutions; how do the corresponding ratii Φ/F compare?

   (b) Compute the PDFs of $|B_z|$ in both cases, and compare/contrast their shape. How similar are they? Is this surprising? Why?

Bibliography:

The mathematical aspects of fast dynamo theory are discussed at length in the book

Childress, S., & Gilbert, A.D. 1995, *Stretch, Twist, Fold: The Fast Dynamo*, (Berlin: Springer),

although the reader prefering a shorter introduction might want to first work through the two following review articles:

Childress, S. 1992, in *Topological Aspects of the Dynamics of Fluids and Plasmas*, eds. H.K. Moffatt *et al.*, Kluwer, 111-147.

Soward, A.M. 1994, in *Lectures on Solar and Planetary Dynamos*, eds. M.R.E. Proctor & A.D. Gilbert, Cambridge University Press, 181.

as well as the recent review paper by A.D. Gilbert cited in the bibliography of the preceeding chapter. On the Roberts cell dynamo, see chapter 5 of the Childress & Gilbert book cited above, as well as

Roberts, G.O. 1972, *Phil. Trans. R. Soc. London*, A**271**, 411,
Soward, A.M. 1983, *J. Fluid Mech.*, **180**, 267.

The scaling relations given by eqs. (8.6)–(8.7) are derived in the Soward paper.

On dynamo action in the CP flow see

Galloway, D.J., & Proctor, M.R.E. 1992, *Nature*, **356**, 691,
Ponty, Y., Pouquet, A., & Sulem, P.L. 1995, *Geophys. Astrophys. Fluid Dyn.*, **79**, 239,
Cattaneo, F., Kim, E.-J., Proctor, M.R.E., & Tao, L. 1995, *Phys. Rev. Lett.*, **75**, 1522.

For all you would ever want to know about PDFs with power law tails, their statistical properties, and the mechanisms producing them, see the first few chapters in

Sornette, D. 2000, *Critical phenomena in natural sciences*, Springer.

The discussion in §8.3.5 follows closely that in the Cattaneo *et al.* (1995) paper cited above, with the results shown on Fig. 8.11 taken directly from that paper. On dynamo action in the ABC flow (problem 3.3), see

Galloway, D.J., & Frisch, U. 1984, *Geophys. Astrophys. Fluid Dyn.*, **29**, 13,
Galloway, D.J., & Frisch, U. 1986, *Geophys. Astrophys. Fluid Dyn.*, **36**, 53.

The mathematically inclined reader wishing to delve deeper into the theorems for fast dynamo action mentioned in §8.3 will get a solid and character building workout out of

Vishik, M.M. 1989, *Geophys. Astrophys. Fluid Dyn.*, **48**, 151,
Klapper, I., & Young, L.S. 1995, *Comm. Math. Phys.*, **173**, 623.

Our discussion of nonlinear effects in the CP flow is taken directly from

Cattaneo, F., Hughes, D.W., & Kim, E.-J. 1996, *Phys. Rev. Lett.*, **76**, 2057.

On dynamo action in three-dimensional thermally-driven convective turbulence, see

Cattaneo, F. 1999, *Astrophys. J.*, **515**, L39,

Cattaneo, F., Emonet, T., & Weiss, N.O. 2003, *Astrophys. J.*, **588**, 1183-1198.

The numerical simulation results shown on Fig. 8.12 are taken from the second of these papers.

The detection and statistics of small-scale magnetic flux concentrations has garnered much attention over the years. As the both resolution and sophistication of the detection methods improve, the picture continues to evolve and ideas must be discarded or recycled. The following chronological list of papers illustrates this process,

Spruit, H.C., & Zwaan, C. 1981, *Solar Phys.*, **70**, 207,

Zwaan, C. 1987, *Ann. Rev. Astron. Ap.*, **25**, 83,

Topka, K.P., et al. 1992, *Astrophys. J.*, **396**, 351,

Keller, C.U. 1992, Nature, **359**, 307,

Berger, T.E., et al. 1995, *Astrophys. J.*, **454**, 531,

Lin, H. 1995, *Astrophys. J.*, **446**, 421,

Berger, T.E., & Title, A.M. 1996, *Astrophys. J.*, **463**, 365,

Grossman-Doerth, U., et al. 1996, *Astron. Ap.*, **315**, 610,

Schrijver, C.J., et al. 1997, *Astrophys. J.*, **487**, 424.

Simon, G.W., Title, A.M., & Weiss, N.O. 2001, *Astrophys. J.*, **561**, 427.

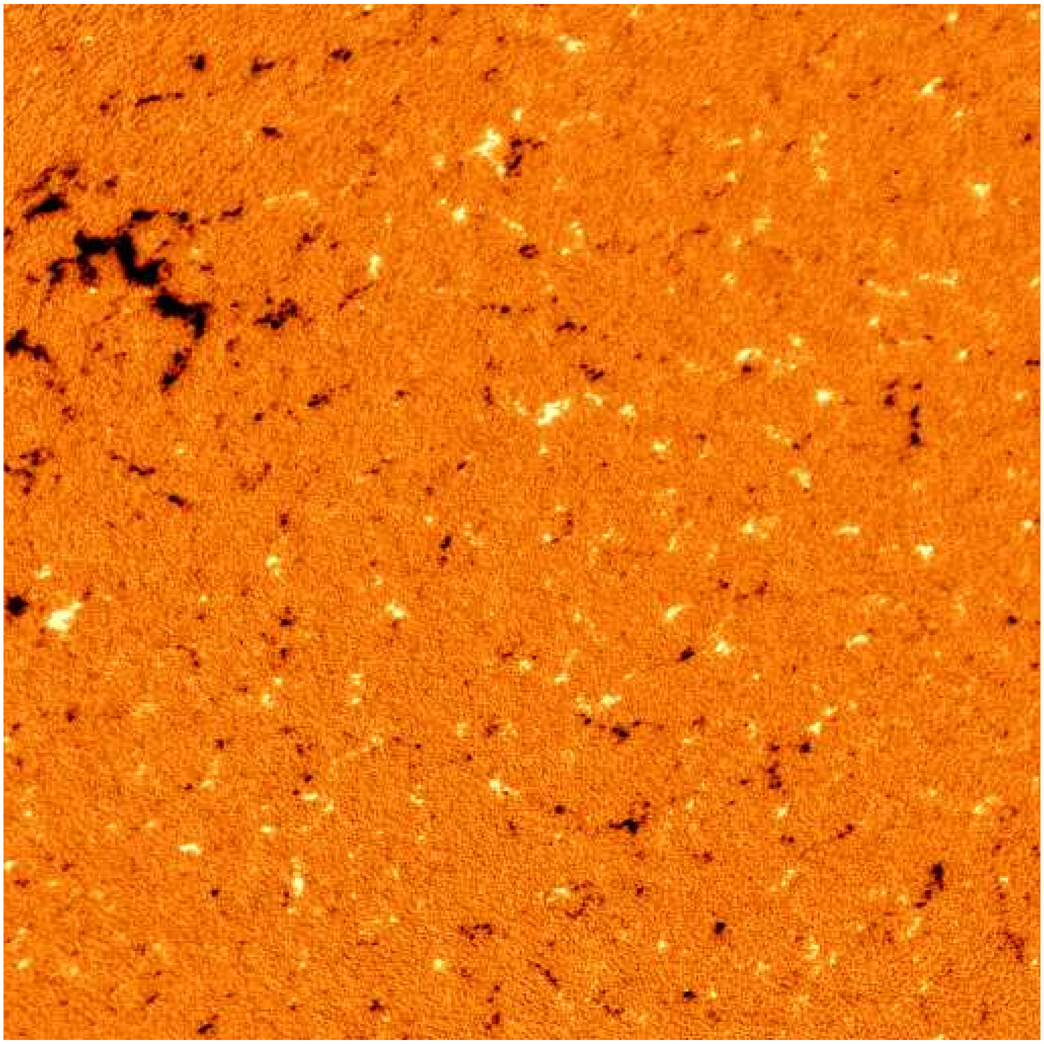

Figure 8.13: {F3.SMAG2} High resolution magnetogram (0.6 arcsec/pixel) of a small piece of "quiet sun", obtained my the MDI instrument onboard SOHO.

Chapter 9

Mean-field theory

{chap:MFE}

In my opinion nothing is contrary to nature save the impossible, and that never happens.

Galileo Galilei
Discourses on Two New Sciences (1638; trans. S. Drake)

This chapter is concerned with the topic of mean-field electrodynamics, which encompasses the "classical" underpinning of dynamo theory from the period before the advent of of large super-computers and parallel-processing, when a megaflop was an over-budget Hollywood film that died on arrival at the box office. A number of themes which we have run across at earlier junctures in these notes reappear in this chapter in slightly different guises and with somewhat altered agendas. The principal achievement of these deliberations is some crucial physical insights—provided by the analytic mathematics upon which mean-field theory is based—on the operation of the α–effect, which is the cornerstone of nearly all astrophysical dynamos.[1]

9.1 Scale separation and statistical averages

The fundamental idea on which mean field theory rests is the **two scale approach**, which consists of a decomposition of the field variables into mean and fluctuating parts. This process naturally implies that an averaging procedure can meaningfully be defined. The derivation of mean field theory can proceed equally from the choice of space averages, time averages or ensemble averages. Space averages are somewhat easier to understand physically,

[1]The material presented in this chapter is written by Thomas J. Bogdan, and is an abridged and slightly modified variant of lecture notes prepared for the APAS-7500 course by Dr. Fausto Cattaneo (Department of Astronomy, University of Chicago) during the fall semester of 1994. The Cattaneo notes, in turn, were strongly inspired by the wonderful 1978 book *Magnetic field generation in electrically conducting fluids*, by H. K. Moffatt.

and that is what we shall *implicitly* adopt here. Ensemble averages are more convenient from a purely mathematical perspective. It is the **ergodic hypothesis** which provides the physical and mathematical justification for our penchant of weaving back and forth between these various definitions of "$\langle\ \rangle$",

$$\langle A \rangle = \frac{1}{\lambda^3} \int_V A \, d\mathbf{x}, \qquad \text{or} \qquad \langle A \rangle = \frac{1}{\tau} \int A \, dt, \tag{9.1}$$ {E4.1}

or the ensemble average.

We assume that the velocity and magnetic field can be decomposed into a mean and fluctuating part so that

$$\mathbf{U} = \langle \mathbf{U} \rangle + \mathbf{u}, \qquad \text{and} \qquad \mathbf{B} = \langle \mathbf{B} \rangle + \mathbf{b}. \tag{9.2}$$ {E4.2}

The decomposition (9.2) makes sense provided $\langle \mathbf{u} \rangle = \langle \mathbf{b} \rangle = 0$. The physical interpretation of (9.2) is as follows. The velocity and magnetic fields are characterized by a slowly varying component, $\langle \mathbf{U} \rangle$ and $\langle \mathbf{B} \rangle$, which vary on the characteristic large scale L, plus rapidly fluctuating parts, $\mathbf{u}$ and $\mathbf{b}$, which vary on the much smaller scale ℓ. The volume averages are computed over some intermediate scale λ such that

$$\ell \ll \lambda \ll L. \tag{9.3}$$ {E4.3}

Whenever (9.3) is satisfied we say that we have a "good" scale separation.[2]

The objective of mean field theory is to produce a closed set of equations for the mean quantities. Substituting (9.2) into the induction equation (2.1), and averaging, we obtain equations for the mean and fluctuating quantities, namely

$$\frac{\partial \langle \mathbf{B} \rangle}{\partial t} = \nabla \times (\langle \mathbf{U} \rangle \times \langle \mathbf{B} \rangle) + \nabla \times \boldsymbol{\mathcal{E}} + \eta \nabla^2 \langle \mathbf{B} \rangle, \tag{9.4}$$ {E4.4}

and

$$\frac{\partial \mathbf{b}}{\partial t} = \nabla \times (\langle \mathbf{U} \rangle \times \mathbf{b}) + \nabla \times (\mathbf{u} \times \langle \mathbf{B} \rangle) + \nabla \times \mathbf{G} + \eta \nabla^2 \mathbf{b}, \tag{9.5}$$ {E4.5}

where

$$\boldsymbol{\mathcal{E}} = \langle \mathbf{u} \times \mathbf{b} \rangle, \qquad \text{and} \qquad \mathbf{G} = \mathbf{u} \times \mathbf{b} - \langle \mathbf{u} \times \mathbf{b} \rangle. \tag{9.6}$$ {E4.6}

The important thing is that (9.4) now contains a source term associated with the average of products of fluctuations. The term $\boldsymbol{\mathcal{E}}$, which is called

[2] In chapter 2 we used ℓ to denote the typical length over which $\mathbf{B}$ varies appreciably. Consistent with our scale separation hypothesis of this chapter, $\mathbf{B}$ is endowed with *two* characteristic length scales for the mean (L) and fluctuating (ℓ) constituents. The intermediate averaging length scale λ is related to the integration volume V by the obvious relation $\lambda \equiv V^{1/3}$.

the average electromotive force, or emf for short, plays a central role in this theory. It is clear that to solve (9.4), $\boldsymbol{\mathcal{E}}$ must be expressed in terms of $\langle \mathbf{U} \rangle$ and $\langle \mathbf{B} \rangle$.

In order to obtain the the desired expression, we note that (9.5) is a *linear* equation for $\mathbf{b}$ with the term $\nabla \times (\mathbf{u} \times \langle \mathbf{B} \rangle)$ acting as a source. There must therefore exist a *linear* relationship between $\mathbf{B}$ and $\mathbf{b}$, and hence, one between $\mathbf{B}$ and $\langle \mathbf{u} \times \mathbf{b} \rangle$. The latter relationship can be expressed formally by the following series

$$\mathcal{E}_i = \alpha_{ij} \langle B \rangle_j + \beta_{ijk} \partial_k \langle B \rangle_j + \gamma_{ijkl} \partial_j \partial_k \langle B \rangle_l + \cdots, \tag{9.7}$$ {E4.7}

where the tensorial coefficients, α, β, γ, and so forth must depend on $\langle \mathbf{U} \rangle$, what we might loosely term the *statistics* of the turbulent velocity fluctuations, $\mathbf{u}$, and on the diffusivity η—but *not* on $\langle \mathbf{B} \rangle$. In this sense, equations (9.4) and (9.7), constitute a closed set of equations for the evolution of $\langle \mathbf{B} \rangle$. The convergence of the series representation provided by equation (9.7) can be anticipated in those cases where the good separation of scales applies. For in these cases each successive derivative in equation (9.7) is smaller than the previous one by approximately a factor of $\ell/L \ll 1$. With any luck, we may expect equation (9.7) to be dominated by the first few terms.

9.2 The α–effect and turbulent diffusivities

We have already remarked that $\boldsymbol{\mathcal{E}}$ in (9.4) acts as a source term for the mean field. It is instructive to examine the contributions to $\boldsymbol{\mathcal{E}}$ deriving from the individual terms in the expansion (9.7). The first contribution is associated with the second-rank tensor, α_{ij}, thus

$$\mathcal{E}_i^{(1)} = \alpha_{ij} \langle B \rangle_j. \tag{9.8}$$ {E4.8}

The first thing to note is that α_{ij} must be a pseudo–tensor since it establishes a linear relationship between a polar vector–the mean emf, and an axial vector–the mean magnetic field.[3] We can divide α_{ij} into its symmetric and antisymmetric parts, thus[4]

$$\alpha_{ij} = \alpha_{ij}^s - \epsilon_{ijk} a_k, \tag{9.10}$$ {E4.9}

[3]The difference between polar and axial vectors derives from their behavior under parity transformations. Let P be the parity transformation associated with reflection through the origin. Under the action of P a vector field $\mathbf{F}$ transforms according to

$$\mathbf{F}(P\mathbf{x}) = \lambda \mathbf{F}(\mathbf{x})$$

where $\lambda = -1$ for a polar vector and $+1$ for an axial one.

[4]Here, ϵ_{ijk} is the Levi-Civita tensor density, also known as the unit alternating tensor, and has the values $\epsilon_{ijk} = 0$ when i, j, k are not all different, $\epsilon_{ijk} = +1$ or -1 when i, j, k

where $2a_k = -\epsilon_{ijk}\alpha_{ij}$. From (4.8) we have

$$\mathcal{E}_i^{(1)} = \alpha_{ij}^s \langle B \rangle_j + (\mathbf{a} \times \langle \mathbf{B} \rangle)_i. \tag{9.11}$$ {E4.10}

The effect of the antisymmetric part is to provide an additional advective velocity (not in general solenoidal) so that the effective mean velocity becomes $\langle \mathbf{U} \rangle + \mathbf{a}$. The nature of the symmetric part is most easily illustrated in the case when $\mathbf{u}$ is an **isotropic** random field.[5] Then $\mathbf{a}$ is zero, α_{ij} must be an isotropic tensor of the form $\alpha_{ij} = \alpha\delta_{ij}$, and (9.11) reduces to

$$\boldsymbol{\mathcal{E}}^{(1)} = \alpha \langle \mathbf{B} \rangle. \tag{9.12}$$ {E4.11}

Using Ohm's law, this component of the emf is found to generate a contribution to the mean current of the form

$$\mathbf{j}^{(1)} = \alpha\sigma_e \langle \mathbf{B} \rangle, \tag{9.13}$$ {E4.12}

where σ_e is the electrical conductivity. For nonzero α, equation (9.13) implies the appearance of a mean current everywhere *parallel* to the mean magnetic field—the so-called α–effect. This is in sharp contrast to the more conventional case where the induced current $\sigma_e(\mathbf{U} \times \mathbf{B})$ is *perpendicular* to the magnetic field. The importance of the α–effect is immediately apparent. We recall from our deliberations in §2.2.1 that a toroidal field could be generated from a poloidal one by differential rotation (velocity shear). The α–effect makes it possible to drive a mean toroidal current parallel to the mean toroidal field, which, in turn will regenerate a poloidal field thereby closing the dynamo cycle. This idea of inducing a toroidal current by the α–effect is at the heart of almost all models of astrophysical dynamos.

To appreciate the physical nature of the α–effect we pause to examine the original model of E.N. Parker (1955). We define a cyclonic event to be the rising of a fluid element associated with a definite circulation, say

are all different and in cyclic, or acyclic, order respectively. A particularly useful formula is (Einstein summation over repeated indices in force):

$$\epsilon_{ijk}\epsilon_{klm} = \delta_{il}\delta_{jm} - \delta_{im}\delta_{jl} \tag{9.9}$$ {???}

where δ_{ij} is the Kronecker-delta, and has the value $\delta_{ij} = 0$ if i, j are different, and $\delta_{ij} = 1$ when $i = j$.

[5] Throughout the rest of this chapter, we will have cause to repeatedly refer to the statistical properties of the turbulent velocity field. In order to avoid confusion we state the following definitions: a (random) field is **stationary** if its probability distribution function (pdf) is time independent, it is **homogeneous** if its pdf is independent of position, it is **isotropic** if its pdf is independent of orientation (or equivalently, invariant under rotations), and it is **reflectionally symmetric** if its pdf is invariant under parity reversal. We should note that isotropy and reflectional symmetry are taken here to be distinct properties, although this protocol is not universally accepted.

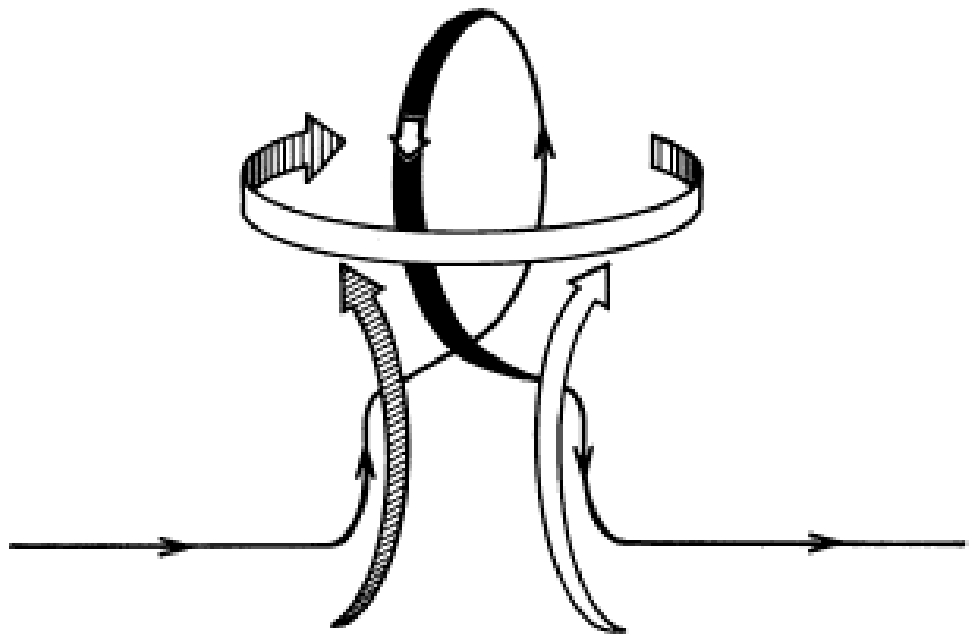

Figure 9.1: {`F4.1a`} A sketch of magnetic line of force entrained by a cyclonic, rising fluid element in the frozen-in limit. Note that the resulting cyclonic loop can be viewed as resulting from an element of electric current flowing parallel to the original, uniform magnetic field. [*from:* Parker 1970, *The Astrophysical Journal*, vol. **162**, Figure 1].

anticlockwise when seen from below (see Figure 9.1). In spherical geometry, we consider the effect of many such events on an initially purely toroidal field line (cf. Figure 9.2). Each cyclonic event creates an elemental loop of field with an associated current distribution that will have a component parallel to the initial field if the angle of rotation is less than π and antiparallel if it is greater. By assuming that the individual events are short lived we can rule out rotations of more than 2π. It is clear that the combined effect of many such events is to give rise to a net current with a component along $\langle \mathbf{B} \rangle$.

An important property of α is its pseudoscalar nature, i.e. α changes sign under parity transformations. This implies that α can be nonzero only if the statistics of $\mathbf{u}$ lacks reflectional symmetry. In other words the velocity field must have a definite handedness (also called chirality). In the example above there is a definite relationship between vertical displacements and sense of circulation.[6] In general the lack of reflectional symmetry of the fluid velocity manifests itself through a nonzero value of the fluid *helicity*, $\langle \mathbf{u} \cdot (\nabla \times \mathbf{u}) \rangle$, itself a pseudo scalar. As we shall presently see there is an important relation

[6] Why?

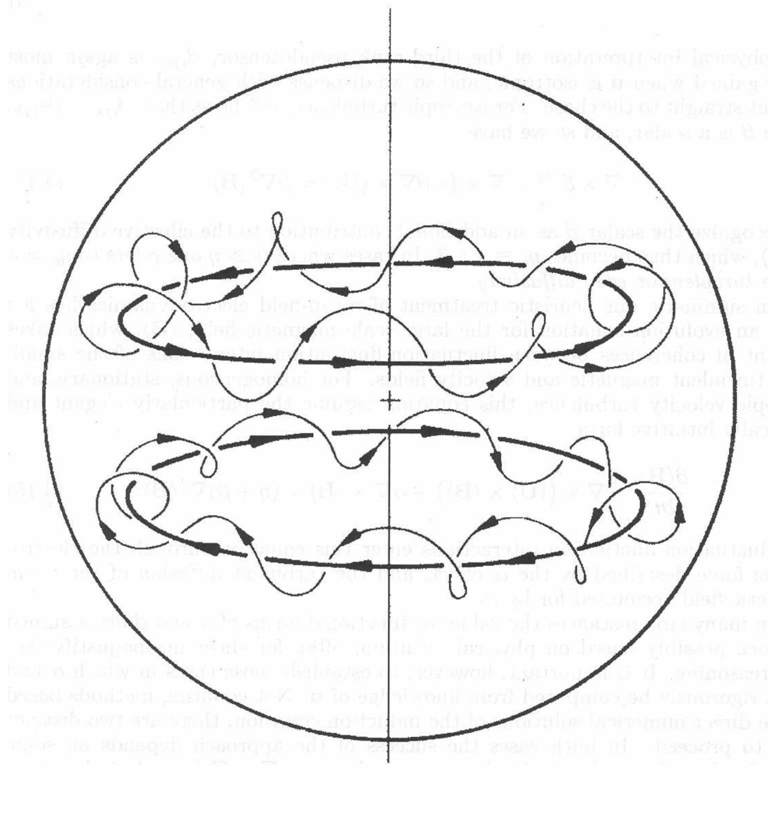

Figure 9.2: {F4.1} A sketch of the azimuthal (toroidal) magnetic lines of force (heavy lines) in the northern and southern hemisphere, carried into spirals by local cyclonic convection cells (thin lines). The collective effect of these events is a mean electric current flowing in the azimuthal direction, which can sustain a poloidal magnetic component. [*from:* Parker 1979, *Cosmical Magnetic Fields*, (Oxford: Clarendon Press), p. 548.]

between fluid helicity and the α–effect.

We now turn to the next term in the expansion (9.7), namely

$$\mathcal{E}_i^{(2)} = \beta_{ijk}\partial_k\langle B\rangle_j. \quad (9.14) \quad \{E4.13\}$$

The physical interpretation of the third-rank pseudotensor, β_{ijk}, is again most easily gained when $\mathbf{u}$ is isotropic, and so we dispense with general considerations and cut straight to the chase. For isotropic turbulence, it follows that, $\beta_{ijk} = \beta\epsilon_{ijk}$, where β is a scalar, and so we have

$$\nabla \times \mathcal{E}^{(2)} = \nabla \times (-\beta\nabla \times \langle \mathbf{B}\rangle) = \beta\nabla^2\langle \mathbf{B}\rangle. \quad (9.15) \quad \{E4.14\}$$

We recognize the scalar β as an additional contribution to the effective diffusivity of $\langle \mathbf{B}\rangle$, which thus becomes $\eta_e \equiv \eta + \beta$. In cases where $\beta \gg \eta$ one refers to $\eta_e \approx \beta$ as the *turbulent* or *eddy diffusivity*.

In summary, our heuristic treatment of mean-field electrodynamics has led us to an evolution equation for the large-scale magnetic field, $\langle \mathbf{B} \rangle$, which takes account of coherences between fluctuation-fluctuation interactions of the small-scale turbulent magnetic and velocity fields. For homogeneous, stationary, and isotropic velocity turbulence, this equation assume the particularly elegant and physically intuitive form

$$\frac{\partial \langle \mathbf{B} \rangle}{\partial t} = \nabla \times (\langle \mathbf{U} \rangle \times \langle \mathbf{B} \rangle) + \alpha \nabla \times \langle \mathbf{B} \rangle + (\eta + \beta) \nabla^2 \langle \mathbf{B} \rangle. \qquad (9.16) \quad \{\texttt{E4.15}\}$$

The fluctuation-fluctuation interactions enter this equation through the electromotive force described by the α–effect, and the turbulent diffusion of the mean magnetic field accounted for by β.

In many circumstances the values or functional forms of α and β are assumed *a priori*, possibly based on physical intuition, often for sheer means-justify-the-ends reasoning. It is important, however, to establish those cases in which α and β can rigorously be computed from knowledge of $\mathbf{u}$. Not counting methods based on the direct numerical solutions of the induction equation, there are two distinct ways to proceed. In both cases the success of the approach depends on some simplification of equation (9.5). In one case the term $\nabla \times \mathbf{G}$ is neglected leading to the so-called first order smoothing approximation (FOS). In the other, the term $\eta \nabla^2 \mathbf{b}$ is neglected, leading to the Lagrangian approximation. The two approaches are complementary in the sense that the former is applicable (for most physically relevant circumstances) when the diffusivity is large and the latter when it is small.

9.2.1 First order smoothing

We begin with the case where $\nabla \times \mathbf{G}$ may be neglected. Assuming that $\langle \mathbf{U} \rangle = 0$, (9.5) becomes[7]

$$\frac{\partial \mathbf{b}}{\partial t} = \nabla \times (\mathbf{u} \times \langle \mathbf{B} \rangle) + \nabla \times \mathbf{G} + \eta \nabla^2 \mathbf{b}, \qquad (9.17) \quad \{\texttt{E4.16}\}$$

$$\mathcal{O}(b_o/\tau) \quad \mathcal{O}(B_o u_o/\ell) \quad \mathcal{O}(u_o b_o/\ell) \quad \mathcal{O}(\eta b_o/\ell^2) \qquad (9.18) \quad \{\texttt{???}\}$$

where the magnitudes of the terms in (9.17) are as indicated. Here ℓ and τ are the characteristic length and time scales associated with $\mathbf{u}$, and u_o, b_o and B_o are the rms values of $\mathbf{u}$, $\mathbf{b}$, and $\langle \mathbf{B} \rangle$. Two distinct situations are of physical interest:

$$\tau \approx \ell / u_o \;, \qquad (9.19) \quad \{\texttt{E4.17a}\}$$

[7]This is permissible since we can always effect a Galilean transformation into the co-moving frame of the mean flow.

$$\tau \ll \ell/u_o \ . \tag{9.20}$$ {E4.17b}

The first case corresponds to conventional fluid turbulence where the characteristic time, or the correlation time, is comparable with the eddy-turnover time. In the second case, the correlation time is much less than the turnover time. This corresponds, for example, to an ensemble of random waves. This latter case is sometimes also referred to as the Markovian approximation.

If (9.20) is satisfied, then $|\nabla \times \mathbf{G}| \ll |\partial_t \mathbf{b}|$, and then to a good approximation,

$$\frac{\partial \mathbf{b}}{\partial t} = \nabla \times (\mathbf{u} \times \langle \mathbf{B} \rangle) + \eta \nabla^2 \mathbf{b}, \tag{9.21}$$ {E4.18}

is valid. If, on the other hand, it is equation (9.19) that is satisfied, then $|\nabla \times \mathbf{G}|$ and $|\partial_t \mathbf{b}|$ are necessarily of the same order. Our basic goal is to find a way to discard the $\nabla \times \mathbf{G}$ term since it leads to a very complicated equation for $\mathbf{b}$. We notice that both $|\nabla \times \mathbf{G}|$ and $|\partial_t \mathbf{b}|$ are negligible compared to $\eta \nabla^2 \mathbf{b}$ if we can assume that

$$r_m = \frac{u_o \ell}{\eta} \ll 1, \tag{9.22}$$ {E4.19}

where r_m is the magnetic Reynolds number that pertains to the *small-scale* magnetic fluctuations. While we have repeatedly stressed that the magnetic Reynolds number for the large-scale magnetic field is necessarily a very large number in most astrophysical applications, owing to the large values for L, it is not quite so obvious that r_m should also be much in excess of unity. If we accept for the moment that ℓ *may* be sufficiently small that equation (9.22) is valid, then equation (9.17) reduces to

$$0 = \nabla \times (\mathbf{u} \times \langle \mathbf{B} \rangle) + \eta \nabla^2 \mathbf{b} \ . \tag{9.23}$$ {E4.20}

For all intents and purposes, both of these limiting arguments lead to equation (9.21), since equation (9.23) is basically contained within equation (9.21) as a further special case. In either example, therefore, fluctuations in $\mathbf{b}$ are generated solely by the interaction of the random velocity $\mathbf{u}$ with the mean field $\langle \mathbf{B} \rangle$, and fluctuation–fluctuation interactions, described by the $\nabla \times \mathbf{G}$ term can safely be neglected. A little thought reveals that the success of the present approach hinges on the existence of a short memory time. In case (9.20) the correlation time of the turbulence is short, so that the effects of past history are small. In case (9.19) the further requirement that $r_m \ll 1$ ensures that diffusion acts quickly enough to remove any effects of past history, even though the turbulence per se now has a rather long memory. As we shall see serious difficulties can arise when the memory time is not small.

With these remarks being said, our next task is to solve equation (9.21) within the volume $V = \lambda^3$, for a specified (turbulent) velocity field, $\mathbf{u}$, and a prescribed (effectively) constant mean magnetic field, $\langle \mathbf{B} \rangle$. Of course, it is not $\mathbf{b}$ per se that is of interest, but rather the mean emf $\boldsymbol{\mathcal{E}}$ generated within the volume V. Hence it will prove necessary to specify the statistical properties of $\mathbf{u}$, so that α and β can be related to them. Within FOS only second-order moments of $\mathbf{u}$ are required, which can be specified entirely in terms of a beast called the velocity spectrum tensor. The mathematics gets rather intricate, and those having never seen an octuple integral are encourage to consult §X.Y of the monograph my Moffatt listed in the bibliography at the end of this chapter.

Rather than work out general expressions for the α and β tensors, to better appreciate some of the problems to be encountered within FOS in the limit of small η we examine a particularly simple example. Consider the following velocity field consisting of a single helical wave:

$$\mathbf{u}(\mathbf{x}, t) = u_o(\sin(kz - \omega t), \cos(kz - \omega t), 0) = \mathrm{Re}\left\{\mathbf{u}_o \mathrm{e}^{\mathrm{i}(\mathbf{k}\cdot\mathbf{x} - \omega t)}\right\}, \quad (9.24) \quad \{\mathtt{E4.43}\}$$

where

$$\mathbf{u}_o = u_o(-\mathrm{i}, 1, 0), \qquad \mathbf{k} = (0, 0, k). \quad (9.25) \quad \{\mathtt{???}\}$$

For this velocity field

$$\nabla \times \mathbf{u} = k\mathbf{u}, \quad \mathbf{u}\cdot(\nabla \times \mathbf{u}) = k u_o^2, \quad \text{and} \quad \mathrm{i}\mathbf{u}_o \times \mathbf{u}_o^* = 2u_o^2(0, 0, 1) \;. \quad (9.26) \quad \{\mathtt{E4.44}\}$$

The corresponding periodic solution of (9.21) has the form

$$\mathbf{b}(\mathbf{x}, t) = \mathrm{Re}\left\{\mathbf{b}_o \mathrm{e}^{\mathrm{i}(\mathbf{k}\cdot\mathbf{x} - \omega t)}\right\}, \quad \text{with} \;\; \mathbf{b}_o = \frac{\mathrm{i}\langle \mathbf{B}\rangle \cdot \mathbf{k}}{-\mathrm{i}\omega + \eta k^2}\mathbf{u}_o. \quad (9.27) \quad \{\mathtt{E4.45}\}$$

Hence we can obtain

$$\boldsymbol{\mathcal{E}} = \langle \mathbf{u} \times \mathbf{b} \rangle = -\frac{\eta u_o^2(\langle \mathbf{B}\rangle \cdot \mathbf{k})k^2}{\omega^2 + \eta^2 k^4}(0, 0, 1) \;, \quad (9.28) \quad \{\mathtt{E4.46}\}$$

which gives

$$\alpha_{ij} = \alpha^{(3)}\delta_{i3}\delta_{j3}, \quad \alpha^{(3)} = -\frac{\eta u_o^2 k^3}{\omega^2 + \eta^2 k^4} \;. \quad (9.29) \quad \{\mathtt{E4.47}\}$$

In the example above, we should note that $\mathbf{u} \times \mathbf{b}$ is uniform, therefore $\mathbf{G}$ is zero, and the FOS approximation is *exact.* Expression (9.29) then states that $\alpha \to 0$ as $\eta \to 0$, and that some diffusion is *necessary* for the α–effect to work. In order to appreciate some additional subtle effects associated with η, we note that the above solution does not satisfy the nominal initial condition,

$\mathbf{b}(\mathbf{x}, 0) = 0$. If we insist that this condition be satisfied, then we must add to the particular solution (9.27), a transient term of the form

$$\mathbf{b}_1 = -\mathrm{Re}\left\{\mathbf{b}_0 \mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{x}} \mathrm{e}^{-\eta k^2 t}\right\}, \tag{9.30}$$ {E4.48}

which is simply a magnetic diffusion mode of the homogeneous equation. This additional term also contributes to $\boldsymbol{\mathcal{E}}$, and therefore to α. This transient contribution will decay to zero in a time $\mathcal{O}(\eta k^2)^{-1}$, and, clearly, the memory of the initial conditions will then be forgotten after a time $t \geq (\eta k^2)^{-1}$. However, the limit $\eta \to 0$ poses some interesting problems. If we fix η to some small positive value and let $t \to \infty$ then the transient disappears and we recover (9.29). If, on the other hand, we *first* let $\eta \to 0$, and *then* try to ascertain the long-time behavior, we have

$$\boldsymbol{\mathcal{E}} = \langle \mathbf{u} \times \mathbf{b} \rangle = -\frac{1}{\omega} u_o^2 k \sin \omega t \ (0, 0, 1) \ . \tag{9.31}$$ {E4.49}

The mean emf $\boldsymbol{\mathcal{E}}$, and therefore α, never settles down to any definite value as $t \to \infty$. For this latter case the initial conditions are never forgotten.

9.2.2 The Lagrangian approximation

We saw that in the limit of small diffusivity the FOS approximation cannot consistently be used for standard turbulence and for the case of random waves it may run into difficulties if zero frequency waves are present. It is therefore desirable to derive another approximation that does not require the neglect of the $\nabla \times \mathbf{G}$ term in equation (9.17). This is the basis of the Lagrangian approximation which retains the $\nabla \times \mathbf{G}$ term but neglects instead the diffusive term $\eta \nabla^2 \mathbf{b}$. Clearly the Lagrangian approximation may most likely be justified in the limit of vanishing η.

The Lagrangian approximation leads to expressions for α and β in terms of second order statistics of the *Lagrangian* velocity field. Since these are less commonly used in turbulence work than their Eulerian counterparts, it is instructive to begin with a simpler case and examine the diffusion of a passive scalar, as was first considered by G.I. Taylor (1921). Let θ be a scalar quantity advected by the random velocity field $\mathbf{u}$. Then the evolution of θ is governed by

$$\frac{\partial \theta}{\partial t} + \mathbf{u} \cdot \nabla \theta = 0 \ , \tag{9.32}$$ {E4.50}

Again we assume that the velocity correlation length is ℓ and define averages over some scale $\lambda \gg \ell$. We anticipate that the evolution of $\langle \theta \rangle$ will be governed by a diffusion equation of the type

$$\frac{\partial \langle \theta \rangle}{\partial t} = \kappa_e \nabla^2 \langle \theta \rangle \ , \tag{9.33}$$ {E4.51}

where κ_e is the effective, or turbulent diffusivity. For homogeneous, isotropic turbulence, we expect further that κ_e will be a scalar satisfying $\kappa_e = \mathcal{O}(u_o \ell)$. The physical basis for this expectation is that the effects of turbulent motions are to convect the quantity θ over a distance ℓ at a typical velocity u_o. We notice a similarity between this argument and the procedure used in kinetic theory of gases to compute the *collisional* diffusivity in terms of the mean free path and velocity distribution function.

The solution of equation (9.32) can be developed through a great many nefarious means. Of all of these possibilities, by far the most efficient is to recognize that equation (9.32) is identical to the continuity equation for a solenoidal flow field, e.g., equation (I.1.3). In our discussion of the Lagrangian formulation of wave propagation, in §III.2.3, we devised a means to integrate the continuity equation that was even valid under the more general circumstance in which $\nabla \cdot \mathbf{u} \neq 0$. This method hinged upon viewing the dynamics as a mapping

$$\mathbf{x}(\mathbf{a}, t) = \mathbf{a} + \boldsymbol{\xi}(\mathbf{a}, t) \ , \qquad (9.34)$$ {E4.52}

which takes an element of fluid situated at the point $\mathbf{a}$ at time $t = 0$, to the point $\mathbf{x}$ at any subsequent time $t \geq 0$.[8]Two points are worth mentioning. First, we use $\mathbf{a}$ here instead of $\mathbf{x}^\star$ to represent the initial location. Notation, notation, notation! Second, I have finally outsmarted TEX and figured out how to boldface greek letters! And then equations (III.2.54)–(III.2.58) give the so-called *Cauchy solution* to the problem in terms of the Jacobian of the mapping

$$\mathcal{J}_{ij}(\mathbf{a}, t) = \frac{\partial x_i}{\partial a_j} = \delta_{ij} + \frac{\partial \xi_i}{\partial a_j} \ . \qquad (9.35)$$ {E4.53}

Since we have specialized our discussion to strictly solenoidal flows, it follow that $\mathcal{J} \equiv \det(\mathcal{J}_{ij}) = 1$, and so from equation (III.2.54) we find,

$$\theta(\mathbf{x}(\mathbf{a}, t), t) = \theta(\mathbf{a}, 0) \ , \qquad (9.36)$$ {E4.54}

where $\mathbf{a}$ is the initial position of the fluid trajectory that passes through $\mathbf{x}$ at time t. Recall from our extensive discussion presented in §III.2.3 that equation (4.54) is a Lagrangian statement. The analogous Eulerian statement requires and inversion of the dynamic mapping. For short times, or equivalently, small Lagrangian displacements $|\boldsymbol{\xi}| = |\mathbf{x} - \mathbf{a}| \leq \ell$, we can Taylor-expand equation (9.36) to obtain the (approximate) corresponding Eulerian statement,

$$\theta(\mathbf{x}, t) = \theta(\mathbf{x}, 0) - \xi_i \partial_i \theta(\mathbf{x}, 0) + \frac{1}{2} \xi_i \xi_j \partial_i \partial_j \theta(\mathbf{x}, 0) + \cdots \qquad (9.37)$$ {E4.55}

814

Ensemble-averaging (9.37) and assuming that there are no initial correlations between θ and $\mathbf{u}$ we obtain

$$\langle\theta(t)\rangle = \langle\theta(0)\rangle + \frac{1}{2}\langle\xi_i\xi_j\rangle\partial_i\partial_j\langle\theta(0)\rangle + \cdots . \qquad (9.38) \quad \{\mathtt{E4.56}\}$$

For isotropic flow we may further simplify (9.38) to get

$$\langle\theta(t)\rangle = \langle\theta(0)\rangle + \frac{1}{6}\langle\xi^2\rangle\nabla^2\langle\theta(0)\rangle + \cdots . \qquad (9.39) \quad \{\mathtt{E4.57}\}$$

After a correlation time, deviations from the initial configuration will become substantial and the square displacement field will behave like a random walk, i.e

$$\langle\xi^2\rangle \sim t. \qquad (9.40) \quad \{\mathtt{E4.58}\}$$

In this regime, (9.39) can be regarded as a solution of the diffusion equation (9.33) (in perturbation theory) with

$$\kappa_e = \frac{1}{6}\frac{d}{dt}\langle\xi^2\rangle \ . \qquad (9.41) \quad \{\mathtt{E4.59}\}$$

It is also useful to express the diffusivity in terms of velocity correlations. This can easily be achieved by noting that

$$\xi_i = \int_0^t v_i(\mathbf{a},t')dt', \qquad (9.42) \quad \{\mathtt{E4.60}\}$$

where $v_i(\mathbf{a},t)$ is the Lagrangian velocity. Then

$$\langle\xi^2\rangle = \int_0^t\int_0^t\langle\mathbf{v}(\mathbf{a},t_1)\cdot\mathbf{v}(\mathbf{a},t_2)\rangle dt_1dt_2 = 2\int_0^t[tR_L(s) - sR_L(s)]ds$$

$$\approx t\int_{-\infty}^{+\infty} R_L(s)ds \ , \qquad (9.43) \quad \{\mathtt{E4.61}\}$$

$$R_L(s) \equiv \langle\mathbf{v}(\mathbf{a},t)\cdot\mathbf{v}(\mathbf{a},t+s)\rangle. \qquad (9.44) \quad \{\mathtt{E4.62}\}$$

In order to derive (9.43) we have assumed that for stationary turbulence the correlation function depends on the time difference $|t_1 - t_2|$ but not on t_1 or t_2 separately. Furthermore we also assumed that most of the contributions to the last integral come from $s \sim 0$. Both assumptions are believed to be justified for turbulent flows. The last integral in (9.43) is equal to the zero frequency component of the Lagrangian energy spectrum, and so we obtain another useful expression for the diffusivity, namely

$$\kappa_e = \frac{1}{6}\Phi_L(0) \ . \qquad (9.45) \quad \{\mathtt{E4.63}\}$$

Where the energy spectrum is defined in analogy with equation (**??**)

$$\Phi_L(\omega) = \int_{-\infty}^{\infty} R_L(t) e^{-\mathrm{i}\omega t} \;\; dt. \tag{9.46}$$ {E4.64}

Having practiced on the scalar case we are ready to tackle the more complicated case of the magnetic field. The Cauchy solution for the magnetic field reads [cf. equation (III.2.56)]

$$B_i(\mathbf{x}, t) = \frac{\partial x_i}{\partial a_j} B_j(\mathbf{a}, 0), \tag{9.47}$$ {E4.65}

which is the vector equivalent of (9.36). It shows that the magnetic field is both advected and stretched by the velocity field. From equation (9.47) we can immediately calculate the emf, namely

$$\mathcal{E}_i = \langle \mathbf{u} \times \mathbf{b} \rangle_i = \langle \mathbf{u} \times \mathbf{B} \rangle_i = \epsilon_{ijk} \Big\langle v_j(\mathbf{a}, t) B_l(\mathbf{a}, 0) \frac{\partial x_k}{\partial a_l} \Big\rangle. \tag{9.48}$$ {E4.66}

The calculation of α follows from (9.48) most simply if we assume that $\langle \mathbf{B} \rangle$ is uniform (and therefore constant), and that $\mathbf{b}(\mathbf{x}, 0) = 0$, so that $\mathbf{B}(\mathbf{a}, 0) = \langle \mathbf{B} \rangle$. Then

$$\alpha_{il}(t) = \epsilon_{ijk} \Big\langle v_j(\mathbf{a}, t) \frac{\partial x_k(\mathbf{a}, t)}{\partial a_l} \Big\rangle, \tag{9.49}$$ {E4.67}

where now α_{il} is explicitly a function of time. As before, we use (9.42) to express (9.49) in terms of velocities. We get

$$\alpha_{il}(t) = \epsilon_{ijk} \int_0^t \Big\langle v_j(\mathbf{a}, t) \frac{\partial v_k(\mathbf{a}, s)}{\partial a_l} \Big\rangle ds. \tag{9.50}$$ {E4.68}

The time dependence derives from the requirement that $\mathbf{b}(\mathbf{x}, 0) = 0$ which trivially implies that $\alpha(0) = 0$. For times longer than the correlation time we again expect that the imprint of the initial conditions should be forgotten and that α should rapidly approach its asymptotic value. In other words we expect that as in (9.43) we may carry the integration to infinity and write

$$\alpha_{il} \approx \epsilon_{ijk} \int_0^\infty \Big\langle v_j(\mathbf{a}, t) \frac{\partial v_k(\mathbf{a}, s)}{\partial a_l} \Big\rangle ds. \tag{9.51}$$ {E4.69}

There are however some important differences between the integrand of (9.43) and that of (9.51) that may severely undermine the convergence of the integral in (9.51). The problem is associated with the long time behavior of the derivative in the correlation term in (9.51), namely

$$\frac{\partial v_k}{\partial a_l} = \Big(\frac{\partial v_k}{\partial x_m} \Big) \Big(\frac{\partial x_m}{\partial a_l} \Big) \; . \tag{9.52}$$ {E4.70}

The first term on the LHS of (9.52) is in general stationary for stationary turbulence, however the second is not, since two initially adjacent particles tend to drift apart so that $|\delta\mathbf{x}|/|\delta\mathbf{a}| \sim t^{1/2}$ as $t \to \infty$. It follows that the integrand of (9.51) is both a function of t and s and not of s alone as in (9.43). Expression (9.52) was obtained for zero diffusivity, the convergence of (9.51) in the limit of $\eta \to 0$ is still largely an open question.

For isotropic turbulence (9.51) simplifies to

$$\alpha(t) = -\frac{1}{3}\int_0^\infty \langle \mathbf{v}(\mathbf{a},t)\cdot(\nabla^{(\mathbf{a})}\times\mathbf{v}(\mathbf{a},s))\rangle ds \ , \tag{9.53}$$ {E4.71}

where the differentiation is with respect to $\mathbf{a}$. We may interpret the integrand as a Lagrangian helicity correlation.

Assuming that the initial mean field has a uniform *gradient* and that $\mathbf{b}(\mathbf{x},0) = 0$ leads to a rather similar calculation for the diffusivity β. In the isotropic case we have

$$\beta(t) = \frac{1}{3}\int_0^t \langle \mathbf{v}(t)\cdot\mathbf{v}(s)\rangle ds + \int_0^t \alpha(t)\alpha(s) ds$$

$$+\frac{1}{6}\int_0^t\int_0^t \langle \mathbf{v}(t)\cdot\mathbf{v}(s_2)\nabla^{(\mathbf{a})}\cdot\mathbf{v}(s_1) - (\mathbf{v}(t)\cdot\nabla^{(\mathbf{a})}\mathbf{v}(s_1))\cdot\mathbf{v}(s_2)\rangle ds_1 ds_2. \tag{9.54}$$ {E4.72}

The first term in (9.54) is identical to the expression for a passive scalar, the second and third terms are associated with the vector character of the field $\mathbf{B}$. In particular the term involving products of α at different times suggests that helicity fluctuations may play an important role.[9] The convergence of the term involving triple Lagrangian correlations is open to the same doubts as (9.51). It is important to note that (9.54) implies that β may have a *negative* value. That being the case, and further if $\eta + \beta < 0$ then the effects of the diffusion term are to *amplify* rather than suppress high frequency components. This behavior is probably incompatible with the two scale approach used to derive (9.54).

9.3 Dynamo waves {sec:dynwave}

Having derived the mean field dynamo equations and having established that, at least in some regimes, the α and β coefficients are well behaved, it is instructive to study some elementary solutions. We distinguish different types of solution in terms of the dominant regenerative processes. Although the distinction applies in general, it is most easily illustrated in a simplified Cartesian geometry.

[9] For a discussion of the last point see Kraichnan (1976)

To this end, we begin by recalling our quintessential mean-field equation in the presence of homogeneous, stationary and isotropic turbulence,

$$\frac{\partial\langle\mathbf{B}\rangle}{\partial t} = \nabla\times(\langle\mathbf{U}\rangle\times\langle\mathbf{B}\rangle) + \alpha\nabla\times\langle\mathbf{B}\rangle + (\eta+\beta)\nabla^2\langle\mathbf{B}\rangle. \qquad (9.55) \quad \{\texttt{E4.15b}\}$$

The simplest Cartesian problem which comes equipped with all the standard features of the fancy astrophysical dynamos we shall presently contemplate in chapter 5, arises from the basic shear flow

$$\langle\mathbf{U}\rangle = \Omega z\ \hat{\mathbf{e}}_y\ , \qquad (9.56) \quad \{\texttt{E4.73}\}$$

where Ω is a constant [units: s^{-1}]. We shall further assume that the mean-field coefficients α [units: cm s^{-1}] and $\eta_e = \beta + \eta$ [units: $\mathrm{cm}^2\ \mathrm{s}^{-1}$] are constant.

We begin by uncurling equation (9.16) to obtain[10]

$$\Big\{\frac{\partial}{\partial t} + \Omega z\frac{\partial}{\partial y} - \eta_e\nabla^2\Big\}\langle\mathbf{A}\rangle = \alpha\langle\mathbf{B}\rangle - \Omega(\hat{\mathbf{e}}_y\cdot\langle\mathbf{A}\rangle)\hat{\mathbf{e}}_z. \qquad (9.57) \quad \{\texttt{E4.74}\}$$

The two terms on the RHS of this equation parameterize the α–effect and the Ω–effect. Recall that the Ω–effect describes generation of a toroidal magnetic field by the shearing out of a poloidal field. The (mean-field) α–effect accounts for the regeneration of *both* poloidal and toroidal magnetic fields due to the chirality, or handedness, of the turbulent flow field. These two terms offer the possibility of dynamo action overcoming the magnetic diffusion term which resides on the LHS of this equation. We shall soon see that dynamo action is possible in the absence of shear ($\Omega = 0$), leading to what is called an α^2–dynamo. When both α and Ω are nonzero we have an $\alpha\Omega$–dynamo. And when only Ω is nonzero we have—well, no dynamo at all![11]

Equation (9.57) is very nearly another example of a PDE with constant coefficients. The offending term is the advective derivative $\langle\mathbf{U}\rangle\cdot\nabla$. One means to circumvent the phase-mixing and related chicanery this term has waiting in the wings for us (cf. Figure 7.9) is to focus our attention of two-dimensional dynamo waves which are invariant under translation in the streamwise direction (i.e., $\partial/\partial y \equiv 0$). With the advective term summarily dealt with, we are now free to look for elementary plane-wave solutions of the form

$$\langle\mathbf{A}\rangle = \mathbf{a}_0\exp[\lambda t + \mathrm{i}k(z\cos\vartheta + x\sin\vartheta)]\ . \qquad (9.58) \quad \{\texttt{E4.75}\}$$

[10]Notice that in uncurling equation (9.16) we have made astute use of our ability to add the net divergence of any scalar to the RHS of equation (9.56. What scalar did we pick? Is this at all related to our freedom to to make a gauge transformation in choosing a representation for the vector potential?

[11]Why?

We may assume that $k \geq 0$ and $0 \leq \vartheta \leq 2\pi$ are prescribed (real) parameters. If equation (9.58) is substituted into equation (9.57), the requirement that there be nontrivial $\mathbf{a}_0$ eigenvectors leads to the dispersion relation,[12]

$$(\lambda + \eta_e k^2)^2 = \alpha k(\alpha k + \mathrm{i}\Omega \sin \vartheta) \; . \qquad (9.59) \quad \{\mathrm{E4.76}\}$$

Equation (9.59) provides us with a quadratic equation for λ, with the two solutions,[13]

$$\lambda_{\pm} = \; -\eta_e k^2 \pm \sqrt{\frac{|\alpha| k}{2}} \; \left\{ \left(\sqrt{\Omega^2 \sin^2 \vartheta + \alpha^2 k^2} + |\alpha| k \right)^{\frac{1}{2}} \right.$$

$$\left. + \mathrm{i} \; \mathrm{sign}(\Omega \alpha \sin \vartheta) \; \left(\sqrt{\Omega^2 \sin^2 \vartheta + \alpha^2 k^2} - |\alpha| k \right)^{\frac{1}{2}} \right\} . \qquad (9.60) \quad \{\mathrm{E4.77}\}$$

The λ_- solution can only produce a disturbance which decays with the passage of time, and so the possibility of an exponentially growing mean-field rests on the properties of the λ_+ root. Dynamo action occurs when $\mathrm{Re}(\lambda_+) > 0$. Examination of equation (9.60) indicates that an exponentially growing *dynamo wave* obtains when $0 < k < k_\star$, where the critical wavenumber $k_\star$ is one of the (six) roots of the equation,

$$k_\star^6 - \frac{\alpha^2}{\eta_e^2} k_\star^4 - \frac{\alpha^2 \Omega^2}{4\eta_e^4} \sin^2 \vartheta = 0 \; , \qquad (9.61) \quad \{\mathrm{E4.78}\}$$

If $k_\star \rightarrow 0$ then the "window" for dynamo action disappears. This occurs when $\alpha \rightarrow 0$, which confirms that there is no such beast as an Ω^2–dynamo. From a physical perspective it makes a good deal of sense that the dynamo window inhabits the small-wavenumber, large-wavelength, end of the range of possible parameters. Clearly dynamo waves with rapid spatial fluctuations are susceptible to severe damping due to the enhanced diffusivity $\eta_e \approx \beta$. On the other hand, if the spatial variations of $\langle \mathbf{A} \rangle$ are too large, then there is

[12] Derive this result. Hint: Try to find a pair of PDE's for the y-components of $\langle \mathbf{A} \rangle$ and $\langle \mathbf{B} \rangle$.

[13] Verify that the second term on the RHS of equation (9.60) is indeed the square-root of the RHS of equation (9.59). This complex square-root formula is particularly useful result which is needed quite often in dealing with contour integration and complex variables. Notice that this definition of the complex square-root guarantees that the real part of the square-root is positive-definite, while the imaginary component can change its sign depending upon the sign of the product $\alpha\Omega \sin \vartheta$. It is also permitted to multiply the real part of the square-root by the factor $\mathrm{sign}(\alpha\Omega \sin \vartheta)$, instead of the imaginary part. This would guarantee that the imaginary part of the square-root is positive definite. In the parlance of complex analysis, choosing between either of these options is called picking a particular *Riemann sheet*.

very little $\langle \mathbf{B} \rangle$ for the α–effect to work on, and so the dynamo process again stalls as $k \to 0$.

To solve equation (9.61) for the critical dynamo wavenumber, it is helpful to view equation (9.61) as a *cubic* equation for $\zeta \equiv k_\star^2$. Unlike the sixth-order polynomial equation, the cubic is exactly solvable. Once we find the three (generally complex) values for ζ by standard means, we can take the square-root of each (see footnote # 19) to obtain the six choices for $k_\star$. Based solely on the coefficients of equation (9.61 it is possible to show that there is one real positive root, and a pair of complex-conjugate roots for the cubic ζ-equation. The lone positive root is the ticket, since (one) of its square-roots will also be positive and will provide us with the critical dynamo wavenumber that we seek. Rather than write out the result in all its detail, we will just remark that the critical dynamo wavenumber is readily estimated from equation (9.61 by inspection in the limiting cases:

$$k_\star \approx \begin{cases} \left[\frac{|\alpha\Omega\sin\vartheta|}{2\eta_e^2}\right]^{\frac{1}{3}} & \text{if } |\alpha| \ll \sqrt{\eta_e|\Omega\sin\vartheta|} \\ \frac{|\alpha|}{\eta_e} & \text{if } |\alpha| \gg \sqrt{\eta_e|\Omega\sin\vartheta|} \end{cases} \qquad (9.62) \quad \{\texttt{E4.79}\}$$

The upper line is generally thought to be most applicable to astrophysical situations, and the growing dynamo waves it predicts are called $\alpha\Omega$–dynamos. The lower line is associated with the α^2–dynamo wave.

We use the word "wave" to describe these exponentially growing solutions of the mean field equations because it is clear from equation (9.60) that $\mathrm{Im}(\lambda_+) \neq 0$. The direction of propagation clearly depends upon the sign of the product of α and Ω, and the magnitude of the oscillation period is comparable to the growth rate for the $\alpha\Omega$–dynamo, but it is very much longer than this characteristic growth time for the α^2–dynamo wave. If we think about applying this simple Cartesian example to "explain" the solar cycle and the Maunder butterfly diagram, then our best bet is to hope that the $\alpha\Omega$–dynamo is in operation.

To conclude this section, let's see how well the $\alpha\Omega$–dynamo λ_+-solution that we found above will do in accounting for Figure 6.7. Before we plug in the numbers, we'll first get the geometry straight. The shear flow, you will recall, points in the $\hat{\mathbf{e}}_y$ direction, which we should associate locally with the $\hat{\mathbf{e}}_\phi$ direction in the spherical coordinate system. The $\alpha\Omega$–dynamo works best when the propagation direction of the dynamo wave is perpendicular *both* to the flow direction ($\hat{\mathbf{e}}_y$) and to the direction of shear ($\hat{\mathbf{e}}_z$). Therefore, to optimize our effort we should take $\vartheta = \pi/2$, so the dynamo wave propagates in the $\pm\hat{\mathbf{e}}_x$ direction in the Cartesian coordinate system, or equivalently the $\pm\hat{\mathbf{e}}_\theta$ on the Sun. So far so good. Using the right-hand-rule, this leaves $\hat{\mathbf{e}}_z$ corresponding to $\hat{\mathbf{e}}_r$. Hence, we have a radial shear of the mean zonal (azimuthal) flow (a.k.a. the differential rotation!), which in the presence of a non-zero α–effect, will lead to $\alpha\Omega$–dynamo waves propagating in the latitudinal direction. Excellent!

Now let's go back to the expression we have for λ_+ and put in the numbers. If τ is the assumed dynamo wave period, then, our requirement that we have a good working dynamo solution is,

$$|\alpha| = \frac{8\pi^2}{|\Omega|k\tau^2} \geq \frac{2\eta_e^2 k^3}{|\Omega|} \ . \qquad (9.63)$$ {E4.80}

The inequality guarantees that we have a growing dynamo wave solution, and the equality pegs its period to the observed value of τ. Following earlier discussions, we should place this dynamo wave on the tachocline between the solar envelope and the rigidly rotating solar radiative interior. This has the advantage of gaining us quite a hefty value for Ω, which in turn reduces the required efficiency of the α–effect. From the references provided at the end of chapter 5 of part II, we may deduce that reasonable ballpark values for the parameters appearing on the RHS of equation (9.63) are: $k \approx 4/(0.7R_\odot)$, $\eta_e \approx \beta \approx 10^{10}$ cm^2 s^{-1}, $\Omega \approx -130$ nHz, and $\tau \approx 22$ yr. If you do the arithmetic, you find that we require $\alpha \approx +15$ cm s^{-1}—*positive* in order to get the dynamo wave to propagate from the pole *toward* the equator—and that we safely satisfy the required inequality by something like 5 orders of magnitude.

9.3.1 Numerical simulations

9.4 The mean-field dynamo equations

{S5prob}

9.4.1 Axisymmetric formulation

We close this admittedly very mathematical chapter by getting back to the solar/stellar dynamo problem. Obviously, serious simplifications of the mean-field machinery is needed to yield as tractable problem. The stated goal, remember, is to produce models for the spatiotemporal evolution of the large-scale component of the magnetic field, while subsuming the inductive action of the small scale turbulent flow into the α- and β-effect terms of mean-field theory, as developed above. It is worth repeating that these are the two terms retained from a (severely) truncated series expansion of the mean electromotive force $\mathcal{E} = \langle \mathbf{u} \times \mathbf{b} \rangle$ associated with the small-scale, fluctuating components of the velocity and magnetic field. You should also recall that the physical conditions under which this truncation can be expected to be meaningful may well not be satisfied under solar interior conditions, and that the rotationally-induced break of axisymmetry which allows to circumvent Cowling's theorem is completely contained in the α-effect.

We now proceed to reformulate the mean-field induction equation (**??**) into a form suitable for axisymmetric large-scale magnetic fields. We proceed as we did way back in §1.10.3), which is to express the poloidal field as the curl of a toroidal vector potential, and restrict the large-scale flow to the axisymmetric forms given by eq. (1.92). Henceforth dropping the averaging brackets for notational simplicity, the poloidal/toroidal separation procedure now leads to

$$\frac{\partial A}{\partial t} = \eta\left(\nabla^2 - \frac{1}{\varpi^2}\right)A - \frac{1}{\varpi}\mathbf{u}_p \cdot \nabla(\varpi A) + \alpha B\ , \tag{9.64}$$ {E5.14a}

$$\frac{\partial B}{\partial t} = \quad \eta\left(\nabla^2 - \tfrac{1}{\varpi^2}\right)B - (\nabla\eta)\times(\nabla\times\mathbf{B})$$

$$-\varpi\nabla\cdot\left(\tfrac{B}{\varpi}\mathbf{u}_p\right) + \varpi(\nabla\times A)\cdot(\nabla\Omega) + \nabla\times[\alpha\nabla\times(A\hat{\mathbf{e}}_\phi)]\ , \tag{9.65}$$ {E5.14b}

[Add CM terms] which, structurally, only differs from eqs. (1.94)—(1.95) by the presence of two new terms on the RHS associated with the α-effect. The appearance of this term in eq. (9.64) is crucial, since this is allows us to evade Cowling's theorem.

Equations (9.64)–(9.65) will hereafter be refered to as the *dynamo equations*. For simplicity of notation, we continue to use η for the net magnetic diffusivity, with the understanding that this now includes the (presumably dominant) contribution from the β-term of mean-field theory.

In general, solutions are sought in a meridional plane of a sphere of radius R, and as with the diffusive problem of §7.1 are matched to a potential field in the exterior ($r/R > 1$). Regularity requires that the following boundary conditions be imposed on the symmetry axis:

$$A(r,0) = A(r,\pi) = 0, \qquad B(r,0) = B(r,\pi) = 0\ . \tag{9.66}$$ {E5.15}

In practice it is often useful to solve explicitly for mode having odd and even symmetry with respect to the equatorial plane. To do so, one simply solves the dynamo equations in a meridional *quadrant*, and imposes the following boundary conditions along the equatorial plane. For a dipole-like antisymmetric mode,

$$\frac{\partial A(r,\pi/2)}{\partial\theta} = 0, \qquad B(r,\pi/2) = 0\ , \qquad \text{[Antisymmetric]} \tag{9.67}$$ {E5.22b}

while for symmetric (quadrupole-like) modes one sets instead

$$A(r,\pi/2) = 0, \qquad \frac{\partial B(r,\pi/2)}{\partial\theta} = 0\ , \qquad \text{[Symmetric]}\ . \tag{9.68}$$ {E5.22c}

9.4.2 Scalings and dynamo numbers

Our next step is to put the dynamo equations into nondimensional form. This can actually be carried out in a number of ways. We begin by scaling all lengths in terms of R, and time in terms of the diffusion time $\tau = R^2/\eta$. Equations (9.64)–(9.65) become

$$\frac{\partial A}{\partial t} = \left(\nabla^2 - \frac{1}{\varpi^2}\right) A - \frac{\mathrm{R}_m}{\varpi}\mathbf{u}_p \cdot \nabla(\varpi A) + C_\alpha \alpha B \ , \qquad (9.69) \quad \{\texttt{E5.16a}\}$$

$$\frac{\partial B}{\partial t} = \left(\nabla^2 - \tfrac{1}{\varpi^2}\right) B - (\nabla\eta) \times (\nabla \times \mathbf{B}) - \mathrm{R}_m \varpi \nabla \cdot \left(\tfrac{B}{\varpi}\mathbf{u}_p\right)$$

$$+ C_\Omega \varpi (\nabla \times A) \cdot (\nabla\Omega) + C_\alpha \nabla \times [\alpha \nabla \times (A\hat{\mathbf{e}}_\phi)] \ , \qquad (9.70) \quad \{\texttt{E5.16b}\}$$

where the following three nondimensional numbers have materialized:

$$C_\alpha = \frac{\alpha_0 R}{\eta_0} \ , \qquad (9.71) \quad \{\texttt{E5.17a}\}$$

$$C_\Omega = \frac{\Omega_0 R^2}{\eta_0} \ , \qquad (9.72) \quad \{\texttt{E5.17b}\}$$

$$\mathrm{R}_m = \frac{u_0 R}{\eta_0} \ , \qquad (9.73) \quad \{\texttt{E5.17c}\}$$

with α_0 (dimension cm s^{-1}), η_0 (dimension cm s^{-1}), u_0 (dimension cm s^{-1} and Ω_0 (dimension s^{-1}) as reference values for the α-effect, diffusivity, meridional flow and shear, respectively. Remember that the functionals α, η, $\mathbf{u}_p$ and Ω are hereafter dimensionless quantities. The quantities C_α and C_Ω are *dynamo numbers*, measuring the importance of inductive versus diffusive effects on the RHS of eqs. (9.69)–(9.70). The third dimensionless number, R_m, is none other than our old friend the magnetic Reynolds number, which here measures the relative importance of advection (by meridional circulation) versus diffusion (by Ohmic dissipation) in the transport of A and B in meridional planes.

9.4.3 The little zoo of mean-field dynamo models {ssec:mfzoo}

We now have a *two* source terms on the RHS of (9.70). As we will get to explore in subsequent chapters, whether or not one dominates over the other can lead to distinct modes of dynamo action.

Note first that dynamo action is now possible in the absence of a large-scale shear, i.e., with $\nabla\Omega = 0$ in eq. (9.70). Such dynamos are known as

α^2 dynamos, and regenerate their magnetic field entirely via the inductive action of small-scale turbulence. Traditionally, dynamo action in planetary cores has been assumed to belong to this variety (at least from the point of view of mean-field theory).

Another possibility is that the shearing terms entirely dominates over the α-effect term, in which case the latter is altogether dropped out of eq. (9.70). This leads top the $\alpha\Omega$ dynamo model, which is believed to be most appropriate to the Sun and solar-type stars.

Finally, retaining both source terms in eq. (9.70) defines, you guessed it I hope, the $\alpha^2\Omega$ dynamo model. This has received comparatively little attention in the context of solar/stellar dynamos, since (simple) a priori estimates of the dynamo numbers C_α and C_Ω usually yield $C_\alpha/C_\Omega \ll 1$; caution is however warranted if dynamo action takes place in a thin shell...

Problems:

1. Carry out the averaging and separation procedure on the MHD induction equation, as described in §4.1, and show that it does lead to eqs. (9.4) and (9.5) for the mean and fluctuating parts of the magnetic field.
2. In the context of the plane-wave solutions discussed in §4.3, complete all missing mathematical steps leading to the dispersion relation given by eq. (9.59).

Bibliography:

The basic presentation of this chapter follows the excellent monograph,

Moffatt, H.K. 1978, *Magnetic Field Generation in Electrically Conducting Fluids*, (Cambridge: Cambridge Univ. Press).

Mean-field electrodynamics grew out of the original pioneering efforts of,

Parker, E.N. 1955, ApJ, 122, 293,
Braginskii, S.I. 1964, Sov. Phys. JETP, 20, 726; 1462,
Steenbeck, M., et al. 1966, Z. Naturforsch, 21a, 369; 1285.

Moffatt steers a middle course between the overtly mathematical and physical approaches to the problem. If your inclinations tend more toward the latter try,

Parker, E.N. 1979, *Cosmical Magnetic Fields*, (Oxford: Clarendon Press), ch. 18,

while those with an insatiable appetite for contracting tensors over various sets of indices will derive much satisfaction from perusing,

Krause, F., & Rädler, K.-H. 1980, *Mean-Field Magnetohydrodynamics and Dynamo Theory*, (Oxford: Pergamon Press).

A very reasonable introduction to correlation functions and attendant statistical concepts for random fields can be obtained from,

Batchelor, G.K. 1953, *The Theory of Homogeneous Turbulence*, (Cambridge: Cambridge Univ. Press).

It should me mentioned that to Batchelor, the appellation "homogeneous" also implies "mirror symmetric", which can lead to no small amount of confusion if you begin comparing some of our formulas with those in Batchelor's tome.

The turbulent mixing of a passive scalar is treated by,

Taylor, G.I. 1921, Proc. London Math. Soc., A20, 196,

while the potential for *negative* turbulent diffusion implied by equation (9.54) is discussed by,

Kraichnan, R.H. 1976, J. Fluid Mech., 75, 657; 77, 753,

Parker, E.N. 1979, *Cosmical Magnetic Fields*, (Oxford: Clarendon Press), pp. 584-592.

It is of course possible to run a numerical simulation of MHD turbulence and then look for the α-effect by directly averaging the turbulent flow field generated in the simulation. Here are a few good entry points into this literature:

Pouquet, A., Frisch, U., and Lorat, J., *J. Fluid Mech.*, **77**, 321,

Nordlund, Å, Brandenburg, A., Jennings, R.L., Rieutord, M., Ruokalaien, J., Stein, R.F., & Tuominen, I. 1992, *Astrophys. J.*, **392**, 647,

Ossendrijver, M., Stix, M., and Brandenburg, A. 2001, *Astron. Ap.*, **376**, 731.

Chapter 10

Dynamo models of the solar cycle

{chap:dynmod}

The time has now come to put everything (well... almost) we have learned so far to construct dynamo models for solar and stellar magnetic fields. In this chapter we concentrate on the Sun, for which the amount of observational data available constrains dynamo models to a degree much greater than for other stars, to the extent that the latter will be considered in a separate, subsequent chapter.

We concentrate here on axisymmetric mean-field-like models, in the sense that we will be setting and solving partial differential equations for poloidal and toroidal large-scale magnetic components, and subsume the effects of small-scale fluid motions and magnetic fields into coefficients of these PDEs:

$$\frac{\partial A}{\partial t} = \underbrace{\eta\left(\nabla^2 - \frac{1}{\varpi^2}\right)A}_{\text{resistive decay}} - \underbrace{\frac{\mathbf{u}_p}{\varpi}\cdot\nabla(\varpi A)}_{\text{advection}} \qquad +[+\text{Source}]\ , \tag{10.1}$$ {eq:cowa}

$$\frac{\partial B}{\partial t} = \underbrace{\eta\left(\nabla^2 - \frac{1}{\varpi^2}\right)B}_{\text{resistive decay}} + \underbrace{\frac{1}{\varpi}\frac{\partial(\varpi B)}{\partial r}\frac{\partial \eta}{\partial r}}_{\text{diamagnetic transport}} - \underbrace{\varpi\mathbf{u}_p\cdot\nabla\left(\frac{B}{\varpi}\right)}_{\text{advection}}$$

$$- \underbrace{B\nabla\cdot\mathbf{u}_p}_{\text{compression}} + \underbrace{\varpi(\nabla\times(A\hat{\mathbf{e}}_\phi))\cdot\nabla\Omega}_{\text{shearing}}\ . \tag{10.2}$$ {eq:cowb}

As you will hopefully recall (cf. §1.10.3), these two PDEs result from the separation of the MHD induction equation upon substitution of axisymmetric flow and magnetic fields having the general form:

$$\mathbf{u}(r,\theta) = \mathbf{u}_p(r,\theta) + \varpi\Omega(r,\theta)\hat{\mathbf{e}}_\phi\ , \tag{10.3}$$ {???}

$$\mathbf{B}(r,\theta,t) = \nabla \times (A(r,\theta,t)\hat{\mathbf{e}}_\phi) + B(r,\theta,t)\hat{\mathbf{e}}_\phi \ . \tag{10.4}$$ {eq:axiB}

You will also recall that the presence of a "[Source]" term in eq. (10.1), usually taken to depend on the toroidal field B, is essential for sustained dynamo action, in order to bypass Cowling's theorem (cf. §7.4). With the poloidal source a function of B we recover a nice reciprocal symmetry between eqs. (10.1) and (10.1); the toroidal field production is proportional to the poloidal field strength via the differential rotation. The poloidal field production, in turn, is proportional to the toroidal field strength via the as-yet unspecified poloidal source term; schematically,

$$\nabla\Omega \otimes A \to B \ , \tag{10.5}$$ {E5.18c}

$$[\text{Source}] \otimes B \to A \ , \tag{10.6}$$ {E5.18d}

where the symbol "$\otimes$" and "$\to$" stand for "acting on" and "produces". Evidently we have here —at least conceptually— the ingredients needed for self-regeneration (and exponential growth) of both A and B[1]. It will often prove useful to envision dynamo action as the two-step process as outlined above; even though both mechanisms operate simultaneously and concurrently, it is quite possible that they in fact do so in spatially distinct regions of the solar interior, in which case a suitable transport mechanism must exist to link the two source regions.

Moreover, you will certainly also recall (if not `goto` Fig. 6.11 and `return`) that the sun's poloidal magnetic component, as measured on photospheric magnetograms, flips polarity near sunspot cycle maximum, which –presumably– corresponds to the epoch of peak internal toroidal field strength. The poloidal component (P), in turn, peaks at time of sunspot minimum. The cyclic regeneration of the sun's full large-scale field can thus be thought of as a temporal sequence of the form

$$A(+) \to B(-) \to A(-) \to B(+) \to A(+) \to \dots \ , \tag{10.7}$$ {eq:reversal}

where the (+) and (−) refer to the signs of the poloidal and toroidal components, as established observationally. The dynamo problem can thus be broken into two sub-problems: generating a toroidal field from a pre-existing poloidal component, and a poloidal field from a pre-existing toroidal component.

With shearing by differential rotation taking care of the $A \to B$ step, the whole game will hinge on the specification of the poloidal source term in eq. (10.1). The mean-field electrodynamics approach of the preceding

[1] Can you see the similarity here with the mode of operation of the Roberts Cell dynamo, discussed two chapter ago?

chapter is one, mathematically formal way to calculate possible forms (the "α-effect", where [Source]$\equiv \alpha B$), but there exist also some more empirical approaches that we will look into in due time.

Indeed, the different types of dynamo models we will consider in what follows differ primarily in the choice they make regarding the physical origin and mathematical form of this poloidal source term. They all share the shearing of a poloidal field by differential rotation (§7.2.3) as a source of toroidal field, and all invoke some sort of enhanced, "turbulent" magnetic diffusivity in the solar convective envelope (the "β-effect" of the preceding chapter).

For the sake of convenience, we first (§10.1) collect and review these various common model ingredients. We then consider (§10.2) solar cycle models based on simple forms the α-effect of mean-field electrodynamics. We then look into what currently stands as their main "competitors", namely solar cycle models based on poloidal field regeneration by the surface decay of active regions, more succinctly known as Babcock-Leighton models (§**??**). We then consider (§10.4) cycle models relying on various MHD instabilities to provide a poloidal source term. We then look into the nonlinear behavior and response to stochastic forcing of some of these models (§10.5), with an eye on understanding some of the observed pattern of solar cycle fluctuations reviewed in chap. 6. We close with a brief survey of the current state of model-based solar cycle prediction schemes (§10.6).

10.1 Basic model design {sec:design}

10.1.1 The differential rotation

For the differential rotation $\Omega(r,\theta)$ we retain our now familiar solar-like parametrization (see also Figure 7.12), scaled in terms of the surface equatorial rotation rate:

$$\Omega(r,\theta) = \Omega_C + \frac{\Omega_S(\theta) - \Omega_C}{2}\left[1 + \mathrm{erf}\left(\frac{r - r_C}{w}\right)\right] , \qquad (10.8)$$ {E5.68a}

where

$$\Omega_S(\theta) = (1 - a_2\cos^2\theta - a_4\cos^4\theta) \qquad (10.9)$$ {E5.68b}

with parameter values $\Omega_C = 0.939$, $a_2 = 0.1264$, $a_4 = 0.1591$, $r_c/R = 0.7$, and $w/R = 0.05$, as inferred helioseismologically. Figure 10.1 below shows the corresponding isocontours of angular velocity, together with radial cuts at the pole, equator and mid-latitudes.

It should be noted once again that such a solar-like differential rotation profile is quite complex from the point of view of dynamo modelling, in that it

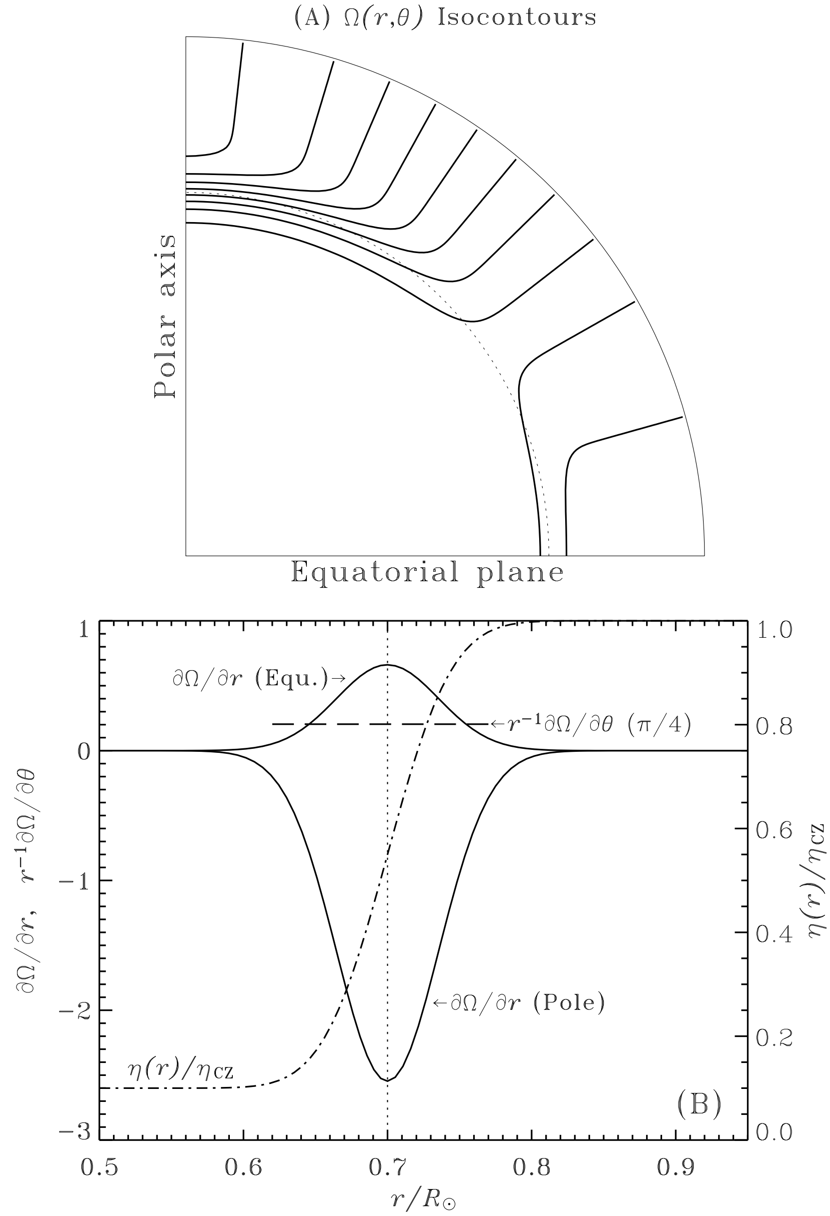

Figure 10.1: Isocontours of angular velocity generated by eqs. (10.8)—(10.9), with parameter values $w/R = 0.05$, $\Omega_C = 0.8752$, $a_2 = 0.1264$, $a_4 = 0.1591$ (panel A). The radial shear changes sign at colatitude $\theta = 55°$. Panel B shows the corresponding angular velocity gradients, together with the total magnetic diffusivity profile defined by eq. (10.10) (dash-dotted line, here with $\Delta\eta = 0.1$ for illustrative purposes). The core-envelope interface is located at $r/R_\odot = 0.7$ (dotted lines). {`fig:dr`}

is characterized by *three* partially overlapping shear regions: a strong positive radial shear in the equatorial regions of the tachocline, an even stronger negative radial shear in its the polar regions, and a significant latitudinal shear throughout the convective envelope and extending partway into the tachocline. As shown on panel B of Fig. 10.1, for a tachocline of half-thickness $w/R_\odot = 0.05$, the mid-latitude latitudinal shear at $r/R_\odot = 0.7$ is comparable in magnitude to the equatorial radial shear; its potential contribution to dynamo action should not be casually dismissed.

10.1.2 The total magnetic diffusivity

For the total magnetic diffusivity $\eta(r)$ we use the same error-function radial profile as before, normalized to the turbulent diffusivity in the convective envelope:

$$\frac{\eta(r)}{\eta_e} = \Delta\eta + \frac{1-\Delta\eta}{2}\left[1+\mathrm{erf}\left(\frac{r-r_c}{w}\right)\right] \ . \qquad (10.10)$$ {eq:eta}

The corresponding profile is plotted on Fig. 10.1 as a dash-dotted line. In practice, the core-to-envelope diffusivity ratio $\Delta\eta \equiv \eta_c/\eta_e$ is treated as a model parameter, with of course $\Delta\eta \ll 1$, since we associate η_c with the microscopic magnetic diffusivity, and η_e with the presumably much larger mean-field turbulent diffusivity β [2]. With the microscopic diffusivity $\eta_c \sim 10^4\,\mathrm{cm}^2\,\mathrm{s}^{-1}$ below the core-envelope interface, and taking the mean-field estimates of β at face value, one obtains $\Delta\eta \sim 10^{-9}$—10^{-6}. The solutions discussed below have $\Delta\eta = 10^{-3}$—10^{-1}, which is much larger, but still small enough to nicely illustrate some important consequence of radial gradients in diffusivity.

10.1.3 The meridional circulation

Meridional circulation is unavoidable in turbulent, compressible rotating convective shells. It basically results from an imbalance between Reynolds stresses and buoyancy forces. The $\sim 15\,\mathrm{m\ s}^{-1}$ poleward flow observed at the surface has been detected helioseismically, down to $r/R_\odot \simeq 0.85$ without significant departure from the poleward direction (except locally and very close to the surface, in the vicinity of active region belts). Mass conservation evidendly requires an equatorward flow deeper down.

For all models discussed below including a meridional circulation $\mathbf{u}_p(r,\theta)$,

[2]We should perhaps repeat that this assumption is a somewhat dubious one, that moreover has been called into question by direct numerical simulation.

we use the following convenient parametric form:

$$u_r(r,\theta) = 2u_0 \left(\frac{R}{r}\right)^2 \left[-\frac{1}{m+1} + \frac{c_1}{2m+1}\xi^m - \frac{c_2}{2m+p+1}\xi^{m+p} \right]$$

$$\times \xi[(q+2)\cos^2\theta - \sin^2\theta]\sin^q\theta \ , \qquad (10.11) \quad \{\texttt{eq:cm1}\}$$

$$u_\theta(r,\theta) = 2u_0 \left(\frac{R}{r}\right)^3 \left[-1 + c_1\xi^m - c_2\xi^{m+p} \right] \sin^{q+1}\theta \cos\theta \ , \qquad (10.12) \quad \{\texttt{eq:cm2}\}$$

with

$$c_1 = \frac{(2m+1)(m+p)}{(m+1)p}\xi_b^{-m} \ , \qquad (10.13) \quad \{\texttt{eq:cm3}\}$$

$$c_2 = \frac{(2m+p+1)m}{(m+1)p}\xi_b^{-(m+p)} \ , \qquad (10.14) \quad \{\texttt{eq:cm4}\}$$

$$\xi = \frac{R}{r} - 1 \ , \qquad (10.15) \quad \{\texttt{eq:cm5}\}$$

$$\xi_b = \frac{R}{r_b} - 1 \ . \qquad (10.16) \quad \{\texttt{eq:cm6}\}$$

This meridional flow satisfies mass conservation ($\nabla \cdot (\rho \mathbf{u}_p) = 0$) for a polytropic density profile of the form:

$$\frac{\rho(r)}{\rho_b} = \left(\frac{R}{r} - 1\right)^m \ . \qquad (10.17) \quad \{\texttt{eq:cm2}\}$$

Setting $m = 0.5$, $p = 0.25$ and $q = 0$, this defines a steady quadrupolar circulation pattern, with a single flow cell per quadrant extending from the surface down to a depth r_b. Circulation streamlines are shown on Fig. 10.2, together with radial cuts of the latitudinal component at mid-latitudes ($\theta = \pi/4$). The flow is poleward in the outer convection zone, with an equatorward return flow peaking slightly above the core-envelope interface, and rapidly vanishing below.

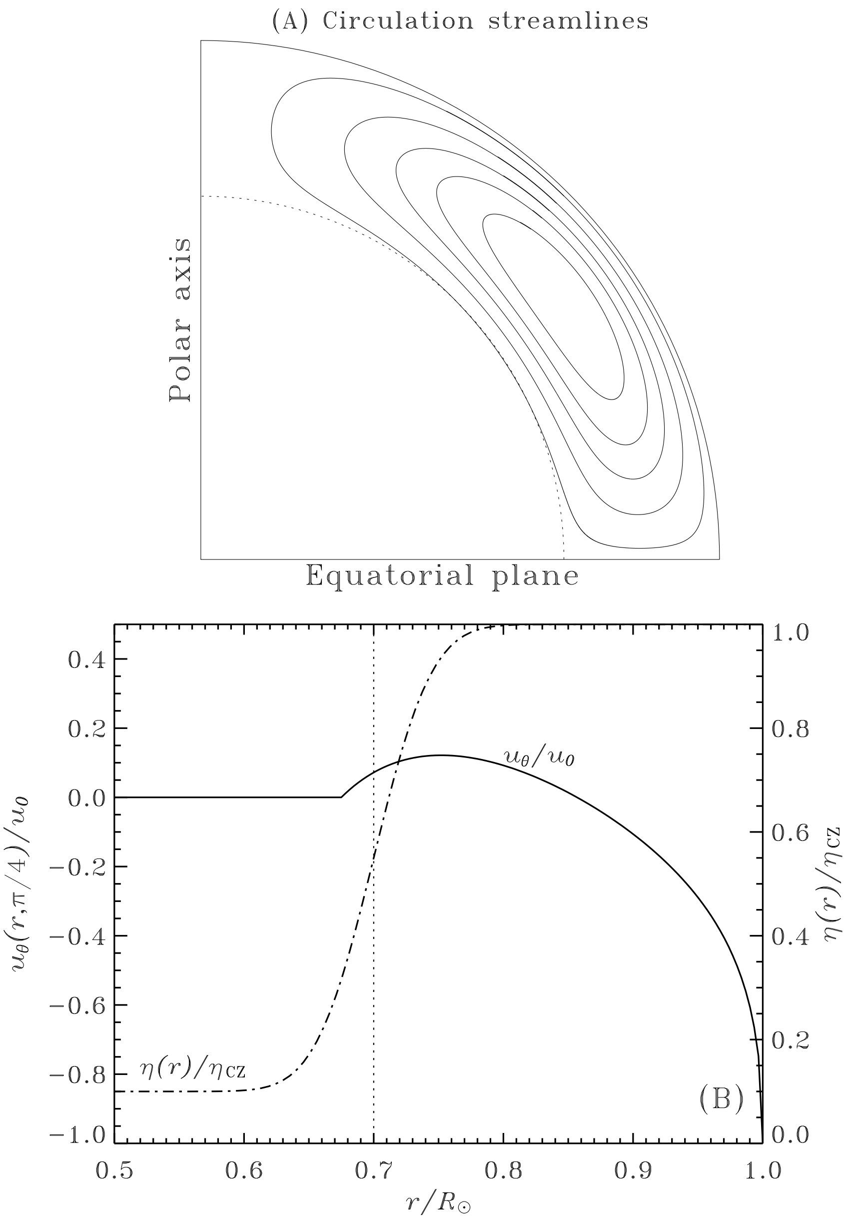

Figure 10.2: Streamlines of meridional circulation (panel A), together with the total magnetic diffusivity profile defined by eq. (10.10) (dash-dotted line, again with $\Delta\eta = 0.1$) and a mid-latitude radial cut of u_θ (bottom panel). the dotted line is the core-envelope interface. This is the analytic flow of van Ballegooijen and Choudhuri (see bibliography), with parameter values $m = 0.5$, $p = 0.25$, $q = 0$ and $r_b = 0.675$. {fig:cm}

10.2 Mean-field models {sec:mfsoldyn}

In this section we consider a series of dynamo models where the poloidal source is the α-effect of mean-field electrodynamics: For the time being we also restrict the models to the kinematic regime, i.e., all flow fields posed *priori* and deemed steady ($\partial/\partial t = 0$), as described by the functional forms given in §10.1. Unless specifically stated otherwise, we assume the parameter values:

$$\eta_T = 5 \times 10^{11}\,\mathrm{cm}^2\,\mathrm{s}^{-1}\ , \qquad \Delta\eta = 0.1\ , \tag{10.18}$$ {???}

$$\Omega_{\mathrm{eq}} = 2.6 \times 10^{-6}\,\mathrm{rad}^2\,\mathrm{s}^{-1}\ , \tag{10.19}$$ {???}

which leads to

$$C_\Omega = 2.5 \times 10^4\ , \tag{10.20}$$ {???}

$$\tau = \frac{R^2}{\eta_T} = 10^{10}\,\mathrm{s} = XXX\,\mathrm{yr}\ . \tag{10.21}$$ {???}

10.2.1 The $\alpha\Omega$ dynamo equations

In constructing mean-field dynamos for the sun, it has been a common procedure to neglect meridional circulation, on the grounds that it is a very weak flow (but more on this further below), and to adopt the $\alpha\Omega$ model formulation, on the grounds that with $R \simeq 7 \times 10^{10}\,\mathrm{cm}$, $\Omega_0 \sim 10^{-6}\,\mathrm{rad\ s}^{-1}$, and $\alpha_0 \sim 100\,\mathrm{cm\ s}^{-1}$, one finds $C_\alpha/C_\Omega \sim 10^3$, independently of the assumed (and poorly constrained) value for η_T. Using the non-dimensional scalings already introduced in §**??**, equations (10.1)—(10.2) then reduce to the so-called $\alpha\Omega$ dynamo equations:

$$\frac{\partial A}{\partial t} = \left(\nabla^2 - \frac{1}{\varpi^2}\right) A + C_\alpha B\ , \tag{10.22}$$ {E5.18a}

$$\frac{\partial B}{\partial t} = \left(\nabla^2 - \frac{1}{\varpi^2}\right) B + C_\Omega \varpi (\nabla \times A) \cdot (\nabla \Omega)\ . \tag{10.23}$$ {E5.18b}

In the spirit of producing a model that is solar-like we use a fixed value $C_\Omega = 2.5 \times 10^4$, obtained assuming $\Omega_0 = \Omega_{Eq}\Omega_S(0) \sim 10^{-6}\,\mathrm{rad\ s}^{-1}$ and $\eta_0 = 5 \times 10^{11}\,\mathrm{cm}^2\,\mathrm{s}^{-1}$.

In the parameter regime characterizing the strongly turbulent solar convection zone, the strength (or even sign) of the α-effect cannot be computed in any reliable manner from first principles, so this will remain the major unknown of the model. In accordance with the $\alpha\Omega$ approximation of the dynamo equations, we restrict ourselves to cases where $|C_\alpha| \ll C_\Omega$. For the dimensionless functional $\alpha(r, \theta)$ we use an expression of the form

$$\alpha(r, \theta) = f(r) g(\theta) \ , \tag{10.24}$$ {E5.26a}

where

$$f(r) = \frac{1}{4}\left[1 + \mathrm{erf}\left(\frac{r - r_c}{w}\right)\right]\left[1 - \mathrm{erf}\left(\frac{r - 0.8}{w}\right)\right] \ . \tag{10.25}$$ {E5.26b}

This combination of error functions concentrates the α-effect in the bottom half of the envelope, and let it vanish smoothly below, just as the net magnetic diffusivity does (i.e., we again set $r_c/R = 0.7$ and $w/R = 0.05$). Various lines of argument point to an α-effect peaking at the bottom of the convective envelope, since there the convective turnover time is commensurate with the solar rotation period, a most favorable setup for the type of toroidal field twisting at the root of the α-effect. Likewise, the hemispheric dependence of the Coriolis force suggests that the α-effect should be positive in the Northern hemisphere, and change sign across the equator ($\theta = \pi/2$). The "minimal" latitudinal dependency is thus

$$g(\theta) = \cos\theta \ . \tag{10.26}$$ {E5.26c}

The C_α dimensionless number, measuring the strength of the α-effect, is treated as a free parameter of the model. You may be shocked by the fact that we are, in a very very cavalier manner, effectively treating the α-effect as a (almost) free-function; this sorry situation is unfortunately the rule rather than the exception in mean-field dynamo modelling (references to some of the more noteworthy exceptions are provided in ther bibliography at thenend of this chapter).

10.2.2 Linear dynamo solutions {ssec:lmf}

With α, β and the large-scale flow given, The $\alpha\Omega$ dynamo equations (10.22)—(10.23) become linear in the mean-field $\mathbf{B}$. With none of the PDE coefficients depending explicitly on time, one can seek eigensolutions of the form

$$\begin{bmatrix} A(r, \theta, t) \\ B(r, \theta, t) \end{bmatrix} = \begin{bmatrix} a(r, \theta) \\ b(r, \theta) \end{bmatrix} e^{\lambda t} \ , \tag{10.27}$$ {E5.19}

where the amplitudes a and b are in general complex quantities. Substituting eqs. (10.27) into the dynamo equations yields a classical linear eigenvalue problem. The problem being linear, such eigensolutions leave the absolute scale of the magnetic field strength undetermined. It will prove convenient to write the eigenvalue explicitly as

$$\lambda = \sigma + i\omega \ , \tag{10.28}$$ {E5.20}

so that σ is the growth rate and ω the cyclic frequency, both expressed in terms of the inverse diffusion time[3] $\tau^{-1} = \eta/R^2$. In a model for the (oscillatory) solar dynamo, we are looking for solutions where $\sigma > 0$ and $\omega \neq 0$. You may think of a dynamo as a peculiar form of MHD instability!

Armed (and dangerous) with the above model, we plow ahead and solve the $\alpha\Omega$ as en eigenvalue problem, using inverse iteration (see appendix XX) We first produce a sequence of solutions for increasing values of $|C_\alpha|$, holding C_Ω fixed at a its "solar" value 2.5×10^4, and without meridional circulation ($\mathrm{R}_m = 0$)[4]. Figure 10.3 shows the variation of the growth rate σ and frequency ω as a function of C_α. Four sequences are shown, corresponding to modes that are either antisymmetric or symmetric with respect to the equatorial plane ("A" and "S" respectively), computed with either positive or negative C_α. For $|C_\alpha|$ smaller than some threshold value, the induction terms make too small a contribution to the RHS of eq. (10.22), leaving the dissipation terms dominant, so that solutions all have $\sigma < 0$, as per Cowling's theorem. As $|C_\alpha|$ increases, the growth rate eventually reaches $\sigma = 0$. At this point we also have $\omega \neq 0$, so that the corresponding solution oscillates with neither growth of decay of its amplitude. Further increases of $|C_\alpha|$ lead to $\sigma > 0$. We are now finally in the dynamo regime, where a weak initial field is amplified exponentially in time.

Ref to numerical appendix

Computing similar sequences for the same same model but different values of C_Ω soon reveals than the onset of dynamo activity ($\sigma > 0$) is controlled by the *product* of C_α and C_Ω:

$$D \equiv C_\alpha \times C_\Omega = \frac{\alpha_0 \Omega_0 R^3}{\eta_0^2}. \tag{10.29}$$ {E5.13}

The value of D for which $\sigma = 0$ is called the **critical dynamo number** (denoted $D_{\rm crit}$)[5]. This, at least, is similar to what we found for the analytical

[3] In view of our discussion in chapter 3, this then implies that all mean-field dynamo models produced by solution of eq. (10.27) are *by definition* slow dynamos. Can you figure that one out?

[4] Obtaining such sequences by inverse iteration is easy if one uses the eigenvalue obtained for a given value of C_α as a guess for the eigenvalue of the next solution incremented in C_α. The first eigenvalue of the sequence must be hunted down by trial and error, or estimated using a different numerical technique (see Appendix XX).

[5] Can you find a way of scaling the $\alpha\Omega$ dynamo equations so that the only nondimensional number appearing in the scaled version of the equation is the dynamo number D defined above?

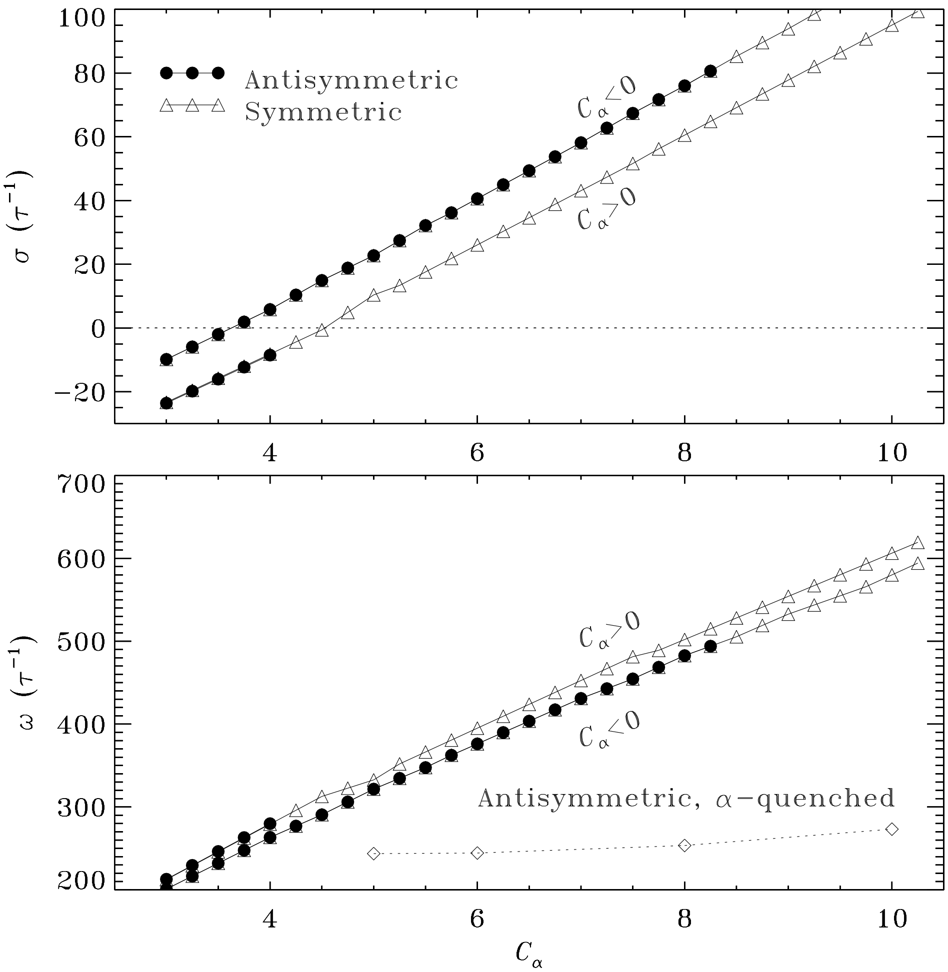

Figure 10.3: {**fig:grates**} Variations of the dynamo growth rate (A) and frequency (B) as a function of increasing $|C_\alpha$— in the minimal $\alpha\Omega$ model. Sequences are shown for either positive or negative dynamo number (as labeled), and symmetric (triangles) or antisymmetric (dots) parity. Modes having $\sigma < 0$ are decaying, and modes with $\sigma > 0$ exponentially growing. Here modes with A or S parity have very nearly identical eigenvalues. In this model the first mode to reach criticality is the negative C_α mode, for which $D_{\rm crit} = -0.9 \times 10^5$. The positive C_α mode reaches criticality at $D_{\rm crit} = 1.1 \times 10^5$. The diamonds on panel (B) correspond to the dynamo frequency measured in a nonlinear version of the same minimal $\alpha\Omega$ model, including algebraic α-quenching as discussed in §10.2.4.

solution of §4.XX[6] Modes having $\sigma < 0$ are called **subcritical**, and those having $\sigma > 0$ **supercritical**. Note on Fig. 10.3 how little the growth rate and dynamo frequency depend on the assumed solution parity.

Here the first mode to become supercritical is the negative C_α mode, for which $D_{\rm crit} = -0.9 \times 10^5$, followed shortly by the positive C_α mode ($D_{\rm crit} = -1.1 \times 10^5$). The dynamo frequency for these critical modes is $\omega \simeq 300$, which corresponds to a full cycle period of $\sim 6\,$yr. This is within a factor of three of the observed full solar cycle period. Once again we should not be too impressed by this, since we have quite a bit of margin of manoeuver in specifying numerical values for η_0 and C_α, and there is no reason to believe that the Sun should be exactly exactly at the critical threshold for dynamo action.

Figure 10.4 shows a half a cycle of the dynamo solution, in the form snapshot of the toroidal (color scale) and poloidal eigenfunctions (fieldlines) in a meridional plane, with the rotation/symmetry axis oriented vertically. The four frames are separated by a phase interval $\varphi = \pi/3$, so that panel (D) is identical to (A) except for reversed magnetic polarities in both magnetic components.

The toroidal field peaks in the vicinity of the core-envelope interface, which is not surprising since in view of eqs. (10.8)—(10.9) the radial shear is maximal there and the magnetic diffusivity and α-effect are undergoing their fastest variation with depth. But why is the amplitude of the dynamo mode vanishing so rapidly below the core-envelope interface? After all, the poloidal and toroidal diffusive eigenmodes investigated in §7.1 were truly global, and the diamagnetic effect should favor stronger fields in the lower diffusivity core. The crucial difference lies with the oscillatory nature of the solution: because the magnetic field produced in the vicinity of the core-envelope interface is oscillating with alternating polarities, its penetration depth in the core is limited by the electromagnetic skin depth $\ell = \sqrt{2\eta_c/\omega}$ (§7.3), with η_c the core diffusivity. Having assumed $\eta_T = 5 \times 10^{11}\,{\rm cm^2\,s^{-1}}$, we have $\eta_c = \eta_T \Delta\eta = 5 \times 10^8\,{\rm cm^2\,s^{-1}}$. A dimensionless dynamo frequency $\omega \simeq 300$ corresponds to $3 \times 10^{-8}\,{\rm s}^{-1}$, so that $\ell/R \simeq 0.026$, quite small indeed.

Careful examination of 10.4A→D also reveals that the toroidal/poloidal flux systems polarity present in the shear layer first show up at high-latitutes, and then *migrate equatorward* to finally disappear at mid-latitudes in the course of the half-cycle[7]. If you haven't already guessed it, what we are seeing on Figure 10.4 is the spherical equivalent of the dynamo waves investigated in §9.3 for the cartesian case with uniform α-effect and shear, if we identify r with z and x with θ. In more general terms, the dynamo wave travel in a

[6]...but does not hold for $\alpha^2\Omega$ dynamo solutions!

[7]An animation of this solution, as well as the one discussed next, can be viewed on the course Web Page.

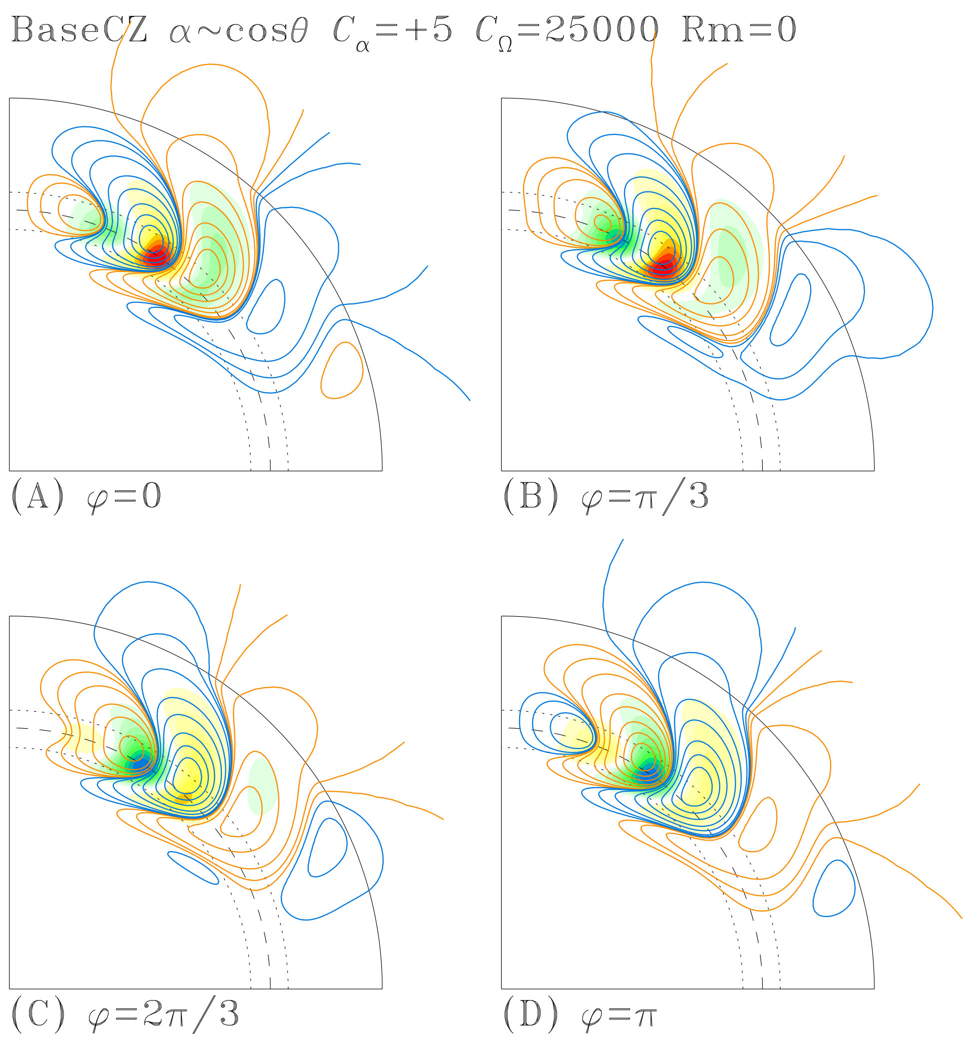

Figure 10.4: {F5.3} Four snapshots in meridional planes of our minimal linear $\alpha\Omega$ dynamo solution with defining parameters $C_\Omega = 25000$, $\eta_T/\eta_c = 10$, $\eta_T = 5 \times 10^{11}\,\mathrm{cm}^2\,\mathrm{s}^{-1}$. With $C_\alpha = +5$, this is a mildly supercritical solution (cf. Fig. 10.3). The toroidal field is plotted as filled contours (green to blue for negative B, yellow to red for positive B, normalized to the peak strength and with increments $\Delta B = 0.2$), on which poloidal fieldlines are superimposed (blue for clockwise-oriented fieldlines, orange for counter-clockwise orientation). The dashed line is the core-envelope interface at $r_c/R = 0.7$ The four snapshots shown here cover a half magnetic cycle, i.e., panel (D) is identical to (A) except for reversed magnetic polarities.

direction $\mathbf{s}$ given by

$$\mathbf{s} = \alpha \nabla \Omega \times \hat{\mathbf{e}}_\phi \ , \tag{10.30}$$ {eq:wave}

i.e., along isocontours of angular velocity. This result is known as the "Parker-Yoshimura sign rule". DISCUSS UPWARD MOTION

10.2.3 Nonlinearities and α-quenching

{ssec:alpqu}

Obviously the exponential growth characterizing supercritical ($\sigma > 0$) linear solutions must stop once the Lorentz force associated with the growing magnetic field becomes dynamically significant for the inductive flow. This magnetic backreaction can show up here in two distinc ways:

1. Reduction of the differential rotation,
2. Reduction of turbulent velocities, and therefore of the α-effect (and perhaps also of the total diffusivity).

Because the solar surface and internal differential rotation shows very little dependence on the phase of the solar cycle, it has been costumary that magnetic backreaction occurs at the level of the α-effect. In the mean-field spirit of *not* solving dynamical equations for the small-scales, it has been standard practice to *assume* a dependence of α on B that "does the right thing", namely reducing the α-effect once the magnetic field becomes "strong enough", the latter usually taken to mean when the growing dynamo-generated mean magnetic field reaches a magnitude such that its energy per unit volume is comparable to the kinetic energy of the underlying turbulent fluid motions. Denoting this **equipartition field strength** by $B_{\rm eq}$, one often introduces an *ad hoc* nonlinear dependency of α (and sometimes η_T as well) directly on the mean-toroidal field B by writing:

$$\alpha \to \alpha(B) = \frac{\alpha_0}{1 + (B/B_{\rm eq})^2} \ . \tag{10.31}$$ {eq:alpquench}

Needless to say, this remains an *extreme* oversimplification of the complex interaction between flow and field that is known to characterize MHD turbulence, but its wide usage in solar dynamo modeling makes it a nonlinearity of choice for the illustrative purpose of this section.

10.2.4 Kinematic $\alpha\Omega$ models with α-quenching

{ssec:nlmf}

With algebraic α-effect included in the poloidal source term, the mean-field $\alpha\Omega$ equations are now nonlinear, and are best solved as an initial-boundary-value problem. The initial condition is an arbitrary seed field of very low

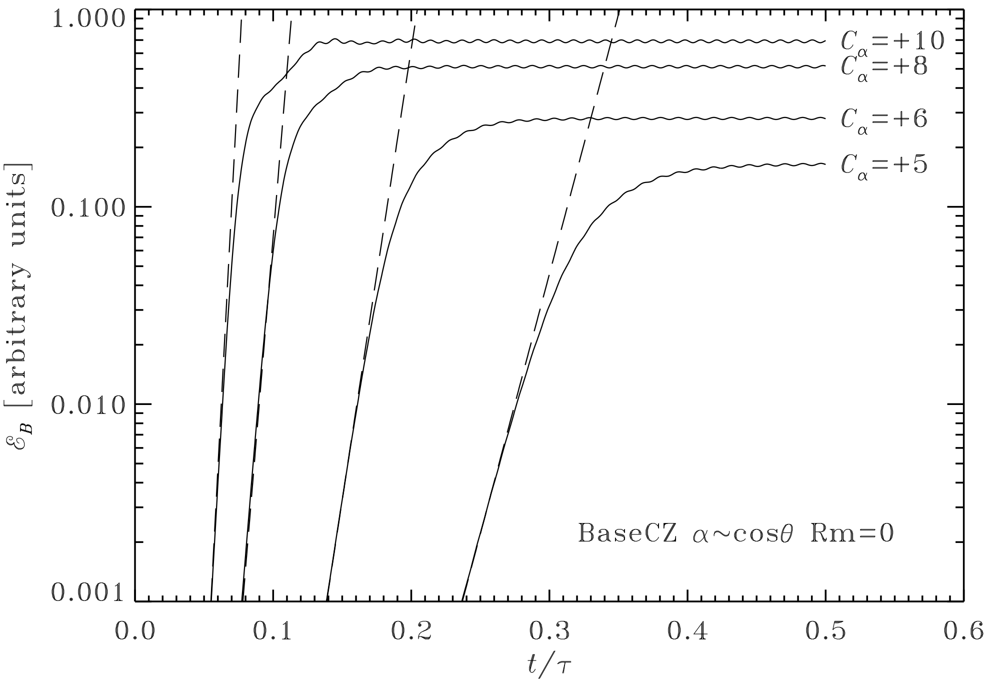

Figure 10.5: {`fig:2tsao`} Time series of magnetic energy for a set of $\alpha\Omega$ dynamo solutions using our minimal $\alpha\Omega$ model including algebraic α-quenching, and different values for C_α, as labeled. Magnetic energy is expressed in arbitrary units, and The dashed line indicates the exponential growth phase characterizing the linear regime.

amplitude, in the sense that $B \ll B_{\rm eq}$ everywhere in the domain. Boundary conditions remain the same as for the linear analysis of the preceding section.

Consider again the minimal $\alpha\Omega$ model of §10.2.2, where the α-effect assumes its simplest possible latitudinal dependency, $\propto \cos\theta$. We use again $C_\Omega = 2.5 \times 10^4$, so that with $C_\alpha = +10$ this places the corresponding linear solution in the supercritical regime (see Figure 10.3). With a very weak **B** as initial condition, early on the model is essentially linear and exponential growth is expected. This is indeed what is observed, as can be seen on Fig. 10.5A, showing time series of the total magnetic energy in the simulation domain for increasing values of C_α, all above criticality. Eventually however, B starts to become comparable to $B_{\rm eq}$ in the region where the α-effect operates, leading to a break in exponential growth, and eventual saturation at some constant value of magnetic energy. Evidently, α-quenching is doing what it was designed to do! Note how the saturation energy level increases with increasing C_α, an intuitively satisfying behavior since solutions with larger C_α have a more powerful poloidal source term. The cycle frequency for these solutions is plotted as diamonds on Fig. 10.3B and, unlike in the linear solutions, now shows very little increase with increasing C_α. Moreover, the dynamo frequency of these α-quenched solutions are found to be slightly *smaller* that the frequency of the linear critical mode (here by some 10—15%), a behavior that is typical of mean-field dynamo models. Yet the overall form of the dynamo solutions closely resembles that of the linear eigenfunctions plotted on Fig. 10.4. Indeed, the full cycle period is here $P/\tau \simeq 0.027$, which translates into 9 yr for our adopted $\eta_T = 5 \times 10^{11}\,\mathrm{cm}^2$ s^{-1}, i.e., a little over a factor of two shorter than the real thing. Not bad!

As a solar cycle model, these dynamo solutions does suffer from one obvious problem: magnetic activity is concentrated at too high latitudes (see Fig. 10.4). This is a direct consequence of the assumed $\cos\theta$ dependency for the α-effect. One obvious way to push the dynamo mode towards the equator is to (artificially) concentrate the α-effect at low latitude. By choosing in this manner an α-effect that "does the right thing", we are throwing away a significant chunk of whatever predictive capability our model might have had. The sad truth is that *ad hoc* specification of the α-effect is a long accepted practice in mean-field dynamo modeling (which of course does not make it any less *ad hoc*!). We therefore proceed nonetheless, using now a latitudinal dependency in $\propto \sin^2\cos\theta$ for the α-effect.

Figure 10.6 shows a selection of three $\alpha\Omega$ dynamo solutions, in the form of time-latitude diagrams of the toroidal field extracted at the core-envelope interface, here $r_c/R_\odot = 0.7$. If sunspot-producing toroidal flux ropes form in regions of peak toroidal field strength, and if those ropes rise radially to the surface, then such diagrams are directly comparable to the sunspot butterfly diagram of Fig 6.7. As before all models have $C_\Omega = 25000$, $|C_\alpha| = 10$, $\Delta\eta = 0.1$, and $\eta_T = 5\times10^{11}\,\mathrm{cm}^2\ \mathrm{s}^{-1}$. To facilitate comparison between solutions, antisymmetric parity was *imposed* via the boundary condition at the equator[8]. On such diagrams, the latitudinal propagation of dynamo waves shows up as a "tilt" of the flux contours away from the vertical direction.

The first solution, on Figure 10.6A, is once again our basic solution of Fig. 10.4, with an α-effect varying in $\cos\theta$. The other two use an α-effect varying in $\sin^2\cos\theta$, and so manage to produce dynamo action that materializes in two more or less distinct branches, one associated with the negative radial shear in the high latitude part of the tachocline, the other with the positive shear in the low-latitude tachocline. These two branches propagate in opposite directions, again in agreement with the Parker-Yoshimura sign rule, since the α-effect here does not change sign within an hemisphere.

It is noteworthy that co-existing dynamo branches, as on Fig. 10.6B and

[8]Animations of the evolving solutions in meridional quadrant are available for the time being at `http://www.astro.umontreal.ca/~paulchar/lrsp/lrsp.html`.

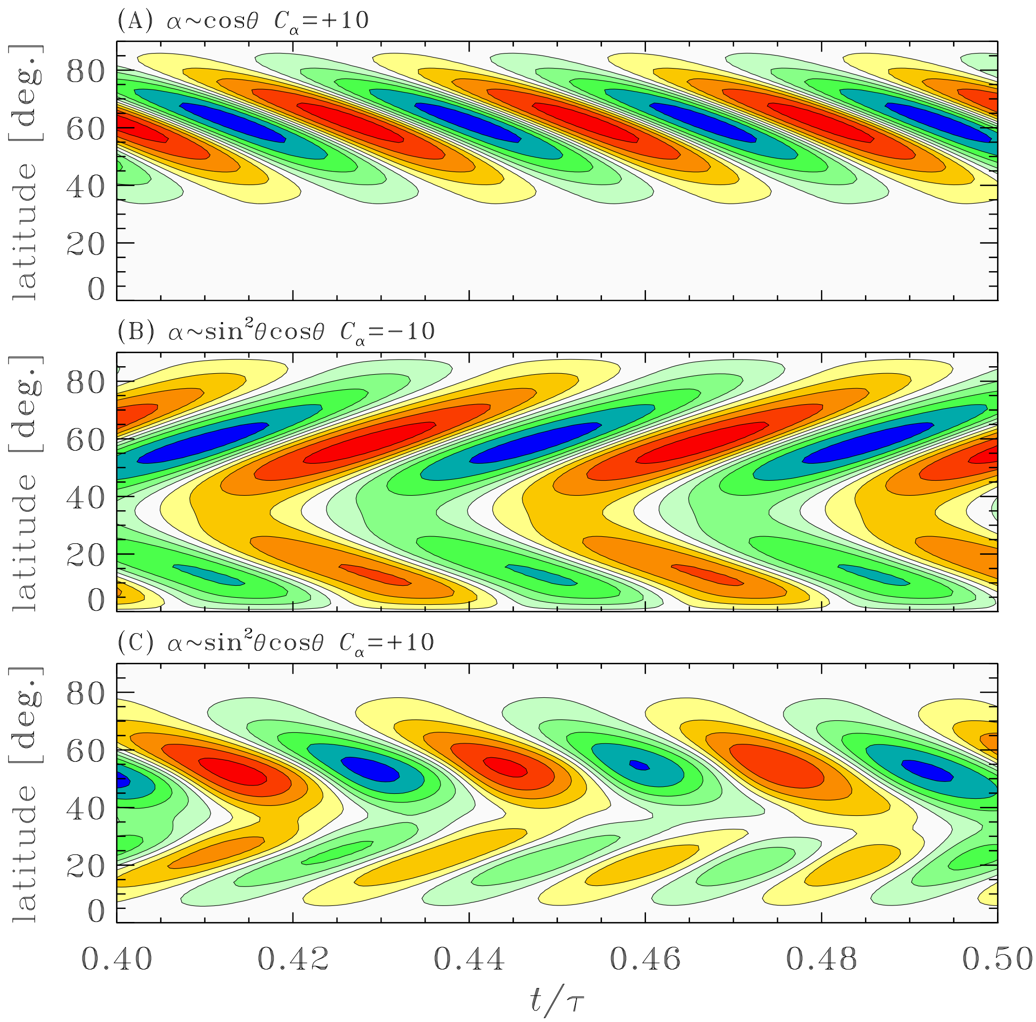

Figure 10.6: {fig:aosolns} Northern hemisphere time-latitude ("butterfly") diagrams for a selection of nonlinear $\alpha\Omega$ dynamo solutions including α-quenching, constructed at the depth $r/R_\odot = 0.7$ corresponding to the core-envelope interface. Isocontours of toroidal field are normalized to their peak amplitudes, and plotted for increments $\Delta B/\max(B) = 0.2$, with yellow-to-red (green-to-blue) contours corresponding to $B > 0$ (< 0). The assumed latitudinal dependence of the α-effect is on given each panel. Other model ingredients as on Fig. 10.1. Note the co-existence of two distinct cycles in the solution shown on panel C, with periods differing by about 25%, which translates in a modulation of the magnetic energy timeseries.

C, can have distinct dynamo periods, which in nonlinearly saturated solutions leads to long-term amplitude modulation. Such modulations are typically not expected in dynamo models where the only nonlinearity present is a simple algebraic quenching formula such as eq. (10.31). Note that this does not occur for the $C_\alpha < 0$ solution, where both branches propagate away from each other, but share a common latitude of origin and so are phased-locked at the onset (cf. Fig. 10.6B). We are seeing here a first example of potentially distinct dynamo modes interfering with one another, a direct consequence of the complex profile of solar internal differential rotation.

The solution of Fig. 10.6B is characterized by a low-latitude equatorially propagating branch, and a full cycle period of 16 yr, which is getting pretty close to the "target" 22yr. But again the strong high-latitude, poleward-propagating branch has no counterpart in the sunspot butterfly diagram. Well, no-problemo, we just concentrate the α-effect even more towards the equator, why not like $\propto \sin^4\theta\cos\theta$, say? It works, but I hope you are starting to find this general approach to the problem as silly as I do... let's try something else instead.

10.2.5 $\alpha\Omega$ models with meridional circulation {ssec:aocm}

Meridional circulation can bodily transport the dynamo-generated magnetic field (terms labeled "advective transport" in eqs. (10.1)–(10.2)), and therefore, for a (presumably) solar-like equatorward return flow that is vigorous enough —in the sense of R_m being large enough— can presumably overpower the Parker-Yoshimura propagation rule embodied in eq. (10.30) and produce equatorward propagation no matter what the sign of the α-effect is. This is readily demonstrated in simple $\alpha\Omega$ models using a purely radial shear at the core-envelope interface (see references in blbiiography), but with a solar-like differential rotation profile the situation turns out to be far more complex.

Starting from our three $\alpha\Omega$ dynamo solutions of Fig. 10.6, new solutions are now recomputed, this time including meridional circulation. Results are shown on Fig. 10.7, for three increasing values of the circulation flow speed, as measured by R_m. At $\mathrm{R}_m = 50$, little difference is seen with the circulation-free solutions, except for the $C_\alpha = +10$ solution with $\alpha \propto \sin^2\theta\cos\theta$, (Fig. 10.7C), where the equatorial branch is now dominant and the polar branch has shifted to mid-latitudes and is cyclic with twice the frequency of the equatorial branch. At $\mathrm{R}_m = 200$, correponding here to a solar-like circulation speed, drastic changes have materialized in all solutions. The negative C_α solution has now transited to a steady dynamo mode, that in fact persists to higher R_m values (panels E and H). The $C_\alpha = +10$ solution with $\alpha \propto \cos\theta$ is decaying at $\mathrm{R}_m = 200$, while the solution with equatorially-concentrated α-effect starts to show a hint of equatorward propagation at mid-latitudes

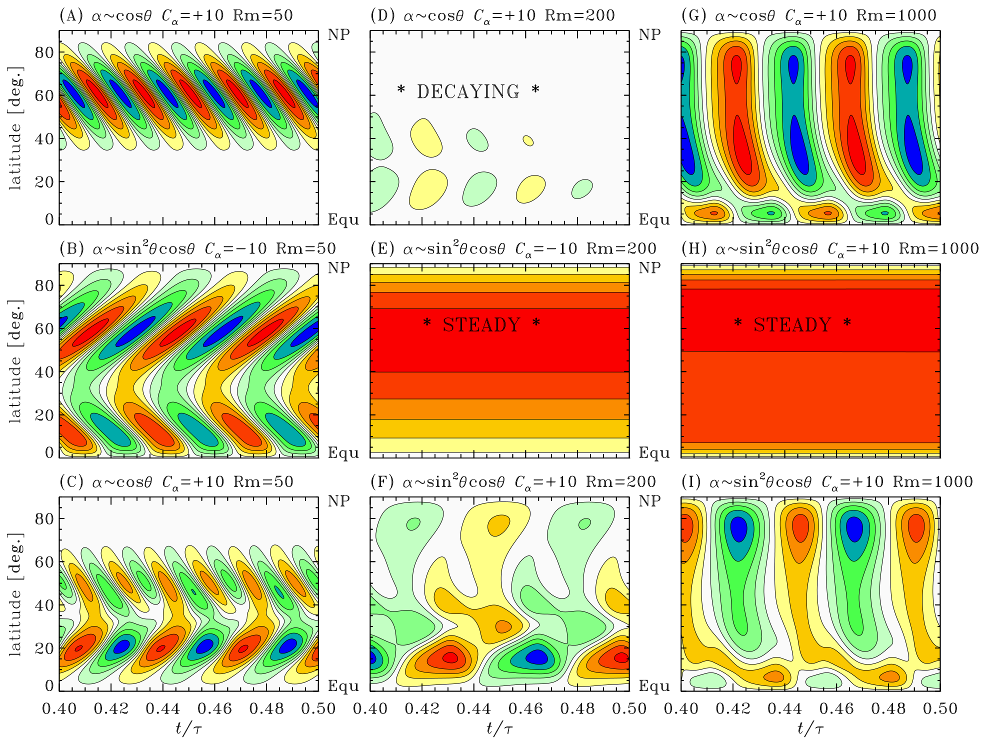

Figure 10.7: Time-latitude diagrams for three of the $\alpha\Omega$ solutions depicted on Fig. 10.6, with meridional circulation now included; the solutions have $\mathrm{R}_m = 50$ (left column), $\mathrm{R}_m = 200$ (middle column), and $\mathrm{R}_m = 10^3$ (left column). For the turbulent diffusivity value adopted here, $\eta_T = 5 \times 10^{11}$ cm^2 s^{-1}, $\mathrm{R}_m = 200$ corresponds to a solar-like circulation speed. {`fig:aocmsolns`}

(panel F). At $\mathrm{R}_m = 10^3$, the circulation has overwhelmed the dynamo wave, and both positive C_α solutions show equatorially-propagating toroidal fields (panels G and I).

Evidently, meridional circulation can have a profound influence on the overall character of the solutions. The behavioral turnover from dynamo wave-like solutions to circulation-dominated magnetic field transport sets in when the circulation speed becomes comparable to the propagation speed of the dynamo wave. In the circulation-dominated regime, the cycle period loses sensitivity to the assumed turbulent diffusivity value, and becomes determined primarily by the circulation's turnover time. This can be seen on Fig. 10.7: at $\mathrm{R}_m = 50$ the solutions on panels (A) and (C) have markedly distinct (primary) cycle periods, while at $\mathrm{R}_m = 10^3$ (panels G and I) the cycle periods are nearly identical. Note however that significant effects require a large R_m ($\gtrsim 10^3$ for the circulation profile used here), which, u_0 being fixed

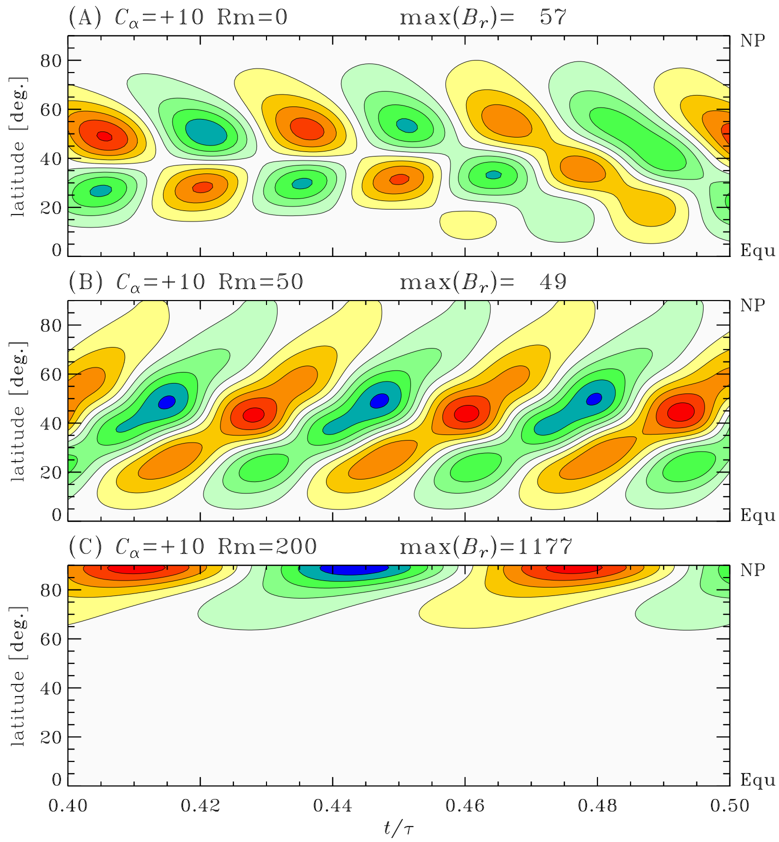

Figure 10.8: Time-latitude diagrams of the surface radial magnetic field, for increasing values of the circulation speed, as measured by the Reynolds number R_m. This is an $\alpha\Omega$ solution with the α-effect concentrated at low-latitude (see §**??** and Fig. 10.6B). Recall that the $\mathrm{R}_m = 0$ solution on panel A exhibits amplitude modulation (cf. Figs. 10.6C). {`fig:surfbrcm`}

by surface observations, translates into a magnetic diffusivity $\eta_T \lesssim 10^{11}$; by most orders-of-magnitude estimates constructed in the framework of mean-field electrodynamics, this is rather low.

Meridional circulation can also dominate the spatiotemporal evolution of the radial surface magnetic field, as shown on Figure 10.8 for a sequence of solutions with $\mathrm{R}_m = 0$, 50, and 200 (corresponding toroidal butterfly diagram at the core-envelope interface are plotted on Figs. 10.6C and 10.7C, F). In the circulation-free solution ($\mathrm{R}_m = 0$), the equatorward drift of the surface radial field is a direct reflection of the equatorward drift of the deep-seated toroidal field (see Fig. 10.6B). With circulation turned on, however, the surface magnetic field is swept instead towards the pole (Fig. 10.8B), becoming strongly concentrated and amplified there for solar-like circulation speeds ($R_m = 200$, Fig. 10.8C) as a consequence of magnetic flux conservation in a converging flow.

Discuss link to strong high-lat B in butterfly diagram

10.2.6 Other classes of mean-field solar cycle models {ssec:mfothers}

10.3 Babcock-Leighton models {sec:BLsoldyn}

Solar cycle models based on what is now called the Babcock-Leighton mechanism were first proposed by Babcock61 and further elaborated by Leighton69, yet they were all but eclipsed by the rise of mean-field electrodynamics in the mid- to late 1960's. Their revival was motivated not only by the mounting difficulties with mean-field models alluded to earlier, but also by the fact that synoptic magnetographic monitoring over solar cycles 21 and 22 has offered strong evidence that the surface polar field reversals are indeed triggered by the decay of active regions (see Fig. 6.11). The crucial question is whether this is a mere side-effect of dynamo action taking place independently somewhere in the solar interior, or a dominant contribution to the dynamo process itself.

Figure **??** illustrates the basic idea of the Babcock-Leighton mechanism. Consider the two bipolar magnetic regions (BMR) sketched on part (A). Recall that each of these is the photospheric manifestation of a toroidal flux rope emerging as an Ω-loop. The *leading* (*trailing*) component of each BMR is that located ahead (behind) in the direction of the Sun's rotation (from E to W). Joy's Law (§X.Y) states that, on average, the leading component is located at lower latitude than the trailing component, so that a line joining each component of the pair makes an angle with respect to the E-W line. Hale's polarity law also inform us that the leading/trailing magnetic polarity pattern is opposite in each hemisphere, a reflection of the equatorial antisymmetry of the underlying toroidal flux system.

Babcock demonstrated empirically from his observation of the sun's surface solar magnetic field that as the BMRs decay (presumably under the influence of turbulent convection), the trailing components drift to higher latitudes, leaving the leading components at lower latitudes, as sketched on Fig. **??**B. Babcock also argued that the trailing polarity poloidal flux released to high latitude by the cumulative effects of the emergence and subsequent decay of many BMRs was responsible for the reversal of the sun's large-scale dipolar field. More germane from the dynamo point of view, the Babcock-Leighton mechanism taps into the (formerly) toroidal flux in the BMR to produce a poloidal magnetic component. To the degree that a positive dipole moment is being produced from a toroidal field that is positive in the N-hemisphere, this is a bit like a positive α-effect in mean-field theory.

In both cases the Coriolis force is the agent imparting a twist on a magnetic field; with the α-effect this process occurs on the small spatial scales and operates on individual magnetic fieldlines. In contrast, the Babcock-Leighton mechanism operates on the large scales, the twist being imparted via the the Coriolis force acting on the flow generated along the axis of a buoyantly rising magnetic flux tube.

10.3.1 Sunspot decay and the Babcock-Leighton mechanism {ssec:BLmech}

Evidently this mechanism can operate as sketched on Figure **??** provided the magnetic flux in the leading and trailing components of each (decaying) BMR are separated in latitude faster than they can diffusively cancel with one another. Moreover, the leading components must end up at low enough latitudes for diffusive cancellation to take place acros the equator. This is not trivial to achieve, and we now take a more quantitative looks at the Babcock-Leighton mechanism, first with a simple 2D numerical model.

The starting point of the model is the grand sweeping assumption that, once the sunspots making up the bipolar active region lose their cohesiveness, their subsequent evolution can be approximated by the passive advection and resistive decay of the radial magnetic field component. This drastic simplification does away with any dynamical effect associated with magnetic tension and pressure within the spots, as well as any anchoring with the underlying toroidal flux system. The model is further simplified by treating the evolution of B_r as a two-dimensional transport problem on a spherical surface corresponding to the solar photosphere. Consequently, no subduction of the radial field can take place.

Even under these simplifying assumptions, the evolution is still governed by the MHD induction equation, specifically its r-component. The imposed flow is made of an axisymmetric "meridional circulation" and differential rotation:

$$\mathbf{u}(\theta) = 2u_0 \sin\theta\cos\theta\hat{\mathbf{e}}_\theta + \Omega_S(\theta)R\sin\theta\hat{\mathbf{e}}_\phi \ , \qquad (10.32)$$ {E6.10a}

where Ω_S is the surface differential rotation profile used in the preceeding chapter (see eq. (**??**)). Note that $\nabla\cdot\mathbf{u} \neq 0$, a direct consequence of working on a spherical surface without possibility of subduction. Introducing a new latitudinal variable $\mu = \cos\theta$ and neglecting all radial derivatives, the r-component of the induction equation (evaluated at $r = R$) becomes:

$$\frac{\partial B_r}{\partial t} = \frac{2u_0}{R}(1-\mu^2)\left[B_r + \mu\frac{\partial B_r}{\partial \mu}\right] - \Omega_S(1-\mu^2)^{1/2}\frac{\partial B_r}{\partial \phi}$$

$$+\frac{\partial}{\partial\mu}\left[\frac{\eta}{R^2}\frac{\partial B_r}{\partial\mu}\right]+\frac{\partial}{\partial\phi}\left[\frac{\eta}{R^2(1-\mu^2)}\frac{\partial B_r}{\partial\phi}\right] , \qquad (10.33) \quad \{E6.11\}$$

with η being the magnetic diffusivity. As usual, we work with the nondimensional form of eq. (10.33), obtained by expressing time in units of $\tau_c = R/u_0$, i.e., the advection time associated with the meridional flow. This leads to the appearance of the following two nondimensional numbers in the scaled version of eq. (10.33):

$$\mathrm{R}_m = \frac{u_0 R}{\eta} , \qquad \mathrm{R}_u = \frac{u_0}{\Omega_0 R} . \qquad (10.34) \quad \{E6.12\}$$

Using $\Omega_0 = 3\times10^{-6}\,\mathrm{rad\ s^{-1}}$, $u_0 = 1500\,\mathrm{cm\,s^{-1}}$, and $\eta = 6\times10^{12}\,\mathrm{cm^2\,s^{-1}}$ yields $\tau_c \simeq 1.5\,\mathrm{yr}$, $\mathrm{R}_m \simeq 20$ and $\mathrm{R}_u \simeq 10^{-2}$. The former is really a measure of the (turbulent) magnetic diffusivity, and is the only free parameter of the model, as R_u is well constrained by surface Doppler measurements. The corresponding magnetic diffusion time is $\tau_\eta = R^2/\eta \simeq 26\,\mathrm{yr}$, so that $\tau_c/\tau_\eta \ll 1$.

Figure 10.9 shows a representative solution. The initial condition (panel A, $t = 0$) mimics a series of eight BMRs, four per hemisphere, equally spaced 90^o apart at latitudes $\pm45°$. Each BMR consists of two Gaussian profiles of opposite sign and adding up to zero net flux, with angular separation $d = 10°$ and with a line joining the center of the two Gaussians tilted with respect to the E-W direction[9] by an angle γ, itself related to the latitude θ_0 of the BMR's midpoint according to the Joy's Law-like relation

$$\sin\gamma = 0.5\cos\theta_0 \qquad (10.35) \quad \{E6.13\}$$

The symmetry of the initial condition means that the problem can be solved in a single hemisphere with $B_r = 0$ enforced in the equatorial plane, in a 90° wide longitudinal wedge with periodic boundary conditions in ϕ.

The combined effect of circulation, diffusion and differential rotation is to concentrate the magnetic polarity of the trailing "spot" to high latitude, while the polarity of the leading spot remains near the original location of the active region. This is readily seen upon calculating the longitudinally averaged latitudinal profiles of B_r, as shown on Fig. 10.9F for the five successive epochs shown on (A)—(E). This is essentially equivalent to Babcock's original cartoon (cf. **??**). The time required to achieve this here is $t/\tau_c \sim 1$, and scales[10] as $(\mathrm{R}_m/\mathrm{R}_u)^{1/3}$.

[9]Remember that this is meant to represent the result of a toroidal flux rope erupting through the surface, so that in this case the underlying toroidal field is positive, which is the polarity the polarity of the trailing "spot", as measured with respect to the direction of rotation, from left to right here.

[10]Can you figure that one out?

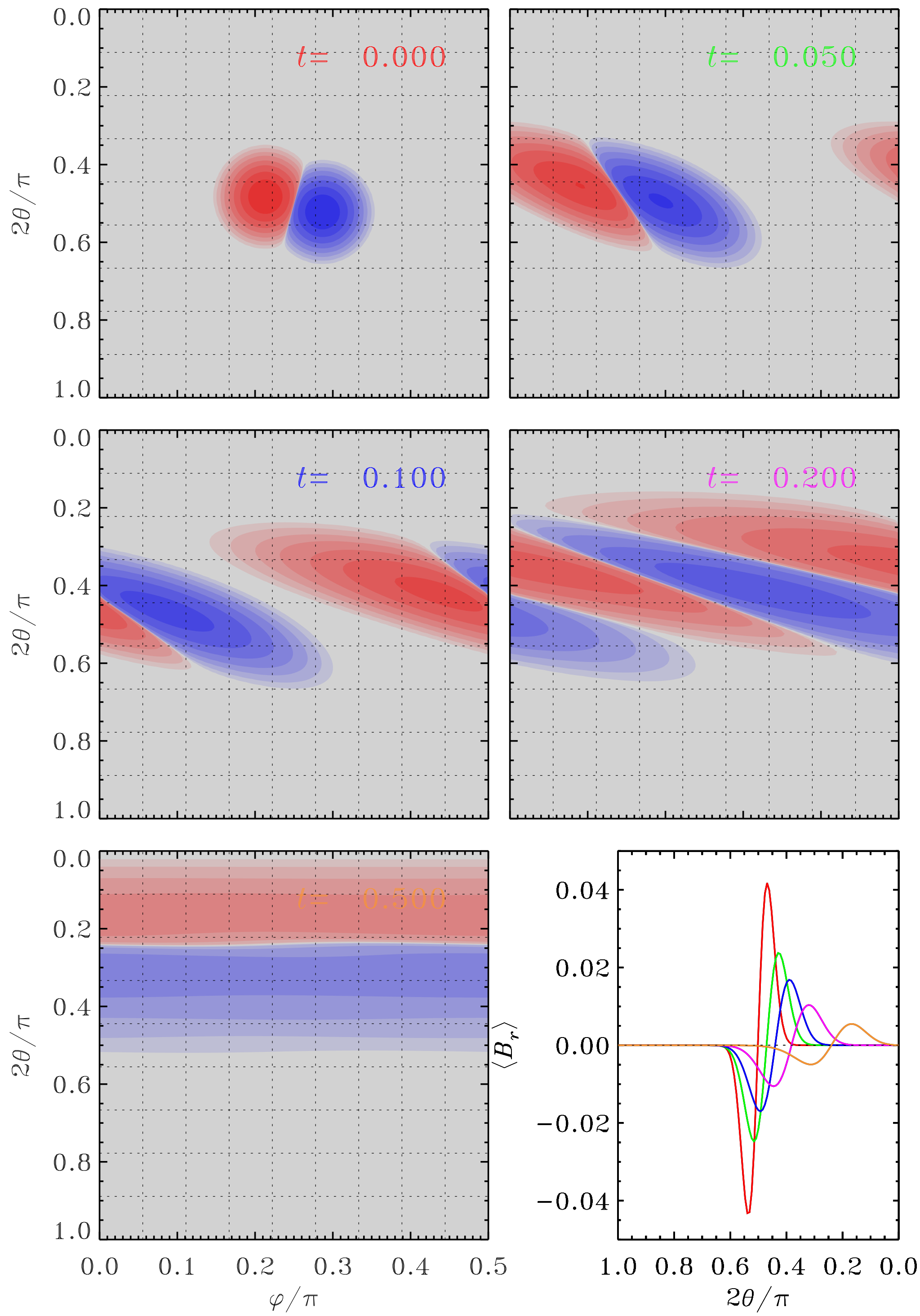

Figure 10.9: {F6.2} Evolution of the surface radial magnetic field component, as described by the 2D advection-diffusion equation (10.33). Parameter values are $\mathrm{R}_u = 10^{-2}$ and $\mathrm{R}_m = 50$, with time given in units of $\tau_c = R/u_0$. The bottom right panel shows the evolution of the longitudinally averaged radial magnetic field.

We can use these simulation results to estimate the "efficiency" of the Babcock-Leighton mechanism. First we define the mean signed and unsigned magnetic flux:

$$\Phi = \mid \langle B_r \rangle \mid \ , \qquad F = \langle \mid B_r \mid \rangle \ , \tag{10.36}$$ {E6.13a}

where the averaging operator on the spherical surface is simply

$$\langle B_r \rangle = -\!\!\!\!\!\!\int_0^{2\pi} \int_{-1}^{+1} B_r \mathrm{d}\mu \mathrm{d}\phi \ . \tag{10.37}$$ {E6.13b}

Figure 10.10A shows the time-evolution of the signed and unsigned flux F, for the solution of Fig. 10.9. The unsigned flux decreases rapidly at first, then settles into a slower decay phase[11]. Meanwhile a small but significant hemispheric signed flux is building up. This is a direct consequence of (negative) flux cancellation across the equator, mediated by diffusion, and is the Babcock-Leighton mechanism in action. Note the dual, conflicting role of diffusion here; it is needed to for cross-hemispheric flux cancellation, yet must be small enough to allow the survival of a significant trailing polarity flux on timescales of order τ_c.

The efficiency (Ξ) of the Babcock-Leighton mechanism, i.e., converting toroidal to poloidal field, can defined as the ratio of the signed flux at $t = \tau_\mathrm{c}$ to the BMR's initial unsigned flux[12]:

$$\Xi = 2\frac{\Phi(t = \tau_\mathrm{c})}{F(t = 0)} \ . \tag{10.38}$$ {E6.13c}

Note that Ξ is independent of the assumed initial field strength of the BMRs since eq. (10.33) is linear in B_r. The efficiency does depend on the tilt and separation of the initial bipolar region, and on the adopted values for R_u and The efficiency is highest for BMRs initially located at high latitudes, even though the tilt γ is smaller. This is because proximity to the equator favors transequatorial diffusive fluxcancellation of the leading component, while having $\mathrm{d}u_\theta/\mathrm{d}\theta < 0$ favors the separation of the two BMR components, thus minimizing diffusive flux cancellation between the leading and trailing components. The efficiency does depend non-trivially on many of the model's parameters,... something you get to explore further in Problem 6.1!

10.3.2 Axisymmetrization revisited

Take another look at Fig. 10.9; at $t = 0$ (panel A) the surface magnetic field distribution is highly non-axisymmetric. By $t/\tau_\mathrm{c} = 0.5$ (panel E), however,

[11]This should remind you of something encountered a few chapter ago...

[12]Can you figure out why a factor of two was inserted on the RHS of eq. (10.38?

the field distribution shows a far less pronounced ϕ-dependency, especially at high latitudes where in fact B_r is nearly axisymmetric. This should remind you of something we encountered earlier: axisymmetrization of a non-axisymmetric magnetic field by an axisymmetric differential rotation (§7.3.5), the spherical analog of flux expulsion. In fact a closer look at the behavior of the unsigned flux on Fig. 10.10A (dashed line) shows the two-timescale behavior we have come to expect of axisymmetrization: the rapid destruction of the non-axisymmetric flux component and slower ($\sim \tau_\eta$) diffusive decay of the remaining axisymmetric flux distribution.

Since the spherical harmonics represent a complete and nicely orthonormal functional basis on the sphere, it follows that the initial condition for the simulation of Fig. 10.9 can be written as

$$B_r^0(\theta,\phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{+l} b_{lm} Y_{lm}(\theta,\phi) \ , \qquad (10.39) \quad \{\texttt{E6.15}\}$$

where the Y_{lm}'s are the spherical harmonics:[13]

$$Y_{lm}(\theta,\phi) = \sqrt{\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}} P_l^m(cos\theta) e^{im\phi} \ , \qquad (10.40) \quad \{\texttt{E6.15b}\}$$

and with the coefficients b_{lm} given by

$$b_{lm} = \int_0^{2\pi} \int_0^{\pi} B_r^0(r,\theta) Y_{lm}^*(\theta,\phi) \qquad (10.41) \quad \{\texttt{E6.16}\}$$

where the "$*$" indicates complex conjugation[14]. Now, axisymmetrization will wipe all $m \neq 0$ modes, leaving only the $m=0$ modes to decay away on the slower diffusive timescale[15]. Therefore, at the end of the axisymmetrization process, the radial field distribution now has the form:

$$B_r(\theta) = \sum_{l=0}^{\infty} \sqrt{\frac{2l+1}{4\pi}} b_{l0} P_l(\cos\theta) \ , \qquad t/\tau_{\rm c} \gg {\rm R}_u \ . \qquad (10.42) \quad \{\texttt{E6.17}\}$$

FIGURE: b_{lm} gray scale plot, with resulting asymptotic $B_r(\theta)$, and corresponding $\langle B_r \rangle$ from above simulation. [16]

---

[13]Better rewrite those factorials differently when trying a numerical implementation...

[14]What are the non-zero b_{lm} for the inclined dipole treated in §7.3.5?

[15]With $\mathbf{u}=0$, the decay rate of those remaining modes are given by the eigenvalues of the 2D pure resistive decay problem, much like in §2.XX [THIS COULD BE A PROBLEM IN CHAP. IV.2].

[16]How would you go about computing the toroidal-to-poloidal efficiency factor Ξ within this modeling framework?

10.3.3 Dynamo models based on the Babcock-Leighton mechanism

So now we understand how the Babcock-Leighton mechanism can provide a poloidal source term in eq. (10.1). Now we need to construct a solar cycle model. One big difference with the $\alpha\Omega$ models considered in §10.2 is that the two source regions are now spatially segregated: production of the toroidal field takes place in the tachocline, as before, but now production of the poloidal field takes place in the surface layers.

The mode of operation of a generic solar cycle model based on the Babcock-Leighton mechanism is illustrated in cartoon form on Figure 10.11. Let P_n represent the amplitude of the high-latitude, surface ("A") poloidal magnetic field in the late phases of cycle n, i.e., after the polar field has reversed. The poloidal field P_n is advected downward by meridional circulation (A→B), where it then starts to be sheared by the differential rotation while being also advected equatorward (B→C). This leads to the growth of a new low-latitude (C) toroidal flux system, T_{n+1}, which becomes buoyantly unstable (C→D) and starts producing sunspots (D), which subsequently decay and release the poloidal flux P_{n+1} associated with the new cycle $n+1$. Poleward advection and accumulation of this new flux at high latitudes (D→A) then obliterates the old poloidal flux P_n, and the above sequence of steps begins anew. Meridional circulation clearly plays a key role in this "conveyor belt" model of the solar cycle, by providing the needed link between the two spatially segregated source regions.

10.3.4 The Babcock-Leighton poloidal source term

The definition of the Babcock-Leighton source term S in eq. (**??**) is evidently the crux of the model. Consider the following:

$$S(r,\theta,B(t)) = s_0 f(r)\sin\theta\cos\theta\left[1-\left(\frac{B(r_c,\theta,t)}{B_0}\right)^2\right]^{-1} B(r_c,\theta,t)\ ,\qquad (10.43)\quad \{\texttt{E6.21}\}$$

with

$$f(r)=\frac{1}{2}\left[1+\mathrm{erf}\left(\frac{r-r_2}{d_2}\right)\right]\left[1-\mathrm{erf}\left(\frac{r-r_3}{d_3}\right)\right]\ ,\qquad (10.44)\quad \{\texttt{E6.21b}\}$$

where s_0 is a numerical coefficient setting the strength of the source term (corresponding dynamo number being $C_S = s_0 R/\eta_0$), and with the various remaining numerical coefficient taking the values $r_2/R = 0.95$, $r_3/R = 1$, $d_2 = d_3 = 10^{-2}R$, and $B_0 = 10^5\,\mathrm{G}$. Note that the dependency on B is *non-local*, i.e., it involves the toroidal field evaluated at the core-envelope interface r_c, (but at the same colatitude θ). The combination of error functions concentrate the source term immediately beneath the surface, which is fine. The nonlocality in B represents the fact that the strength of the source term is proportional to the field strength in the bipolar active region, itself presumably reflecting the strength of the diffuse toroidal field near the core-envelope interface, where the magnetic flux ropes eventually giving rise to the bipolar active region originate. The nonlocal quenching nonlinearity reflects the fact that as the strength of the flux rope reaches about 10^5 G, the flux rope emerges without the tilt essential to the Babcock-Leighton mechanism. The $\cos\theta$ dependency is a first order description of Joy's Law, i.e., the tilt of active regions increases with latitude. Notably missing in eq. (10.43) is some sort of lower threshold on S, to mimic the fact that flux ropes with field strengths lower than a few tens of kG either fail to be destabilized in a short enough timescale, rise to the surface at high latitudes and without systematic tilt patterns, and/or fail altogether to survive their rise through the convective envelope.

The nonlocality of S notwithstanding, at this point the model equations are definitely mean-field like. Yet no averaging on small scales is involved. What is implicit in eq. (10.43) is some sort of averaging process at least in longitude and time.

10.3.5 A sample solution

Figure 10.12 shows N-hemisphere time-latitude diagrams for the toroidal magnetic field at the core-envelope interface (panel A), and the surface radial field (panel B), for a representative Babcock-Leighton dynamo solutions computed following the model implementation described above. The equatorward advection of the toroidal field by meridional circulation is here clearly apparent, as well as the concentration of the surface radial field near the pole. Note how the polar radial field changes from negative (blue) to positive (red) at just about the time of peak positive toroidal field at the core-envelope interface; this is the phase relationship inferred from synoptic magnetograms (e.g., Fig. 6.11 herein) as well as observations of polar faculae

Although it exhibits the desired equatorward propagation, the toroidal field butterfly diagram on Fig. 10.12A peaks at much higher latitude ($\sim 45°$) than the sunspot butterfly diagram ($\sim 15°$–$20°$, cf. Fig. 6.7). This occurs because this is a solution with high magnetic diffusivity contrast, where meridional circulation closes at the core-envelope interface, so that the *latitudinal* component of differential rotation dominates the production of the toroidal field. This difficulty can be alleviated by letting the meridional circulation penetrate below the core-envelope interface, but this often leads to the production of a strong polar branches, again a consequence of both the strong radial shear present in the high-latitude portion of the tachocline, and of the concentration of the poloidal field taking place in the high latitude-surface layer prior to this field being advected down into the tachocline by meridional circulation (viz. Figs. 10.11 and 10.12)

A noteworthy property of this class of model is the dependency of the cycle period on model parameters; over a wide portion of parameter space, the meridional flow speed is found to be the primary determinant of the cycle period (P). This behavior arises because, in these models, the two source regions are spatially segregated, and the time required for circulation to carry the poloidal field generated at the surface down to the tachocline is what effectively sets the cycle period. The corresponding time delay introduced in the dynamo process has rich dynamical consequences, to be discussed in §10.5 below. On the other hand, P is found to depend very weakly on the assumed values of the source term amplitude s_0, and turbulent diffusivity η_T; the latter is is very much unlike the behavior typically found in mean-field models, where P scales nearly as η_T^{-1} in α-quenched $\alpha\Omega$ mean-field models[17].

10.4 Models based on MHD instabilities {sec:instsoldyn}

10.5 Nonlinearities, fluctuations and intermittency {sec:fluct}

10.6 Predicting future cycles {sec:predict}

Bibliography:

Mean-field models for solar (and planetary) dynamos were first discussed by

Steenbeck, M., & Krause, F. 1969, *Astr. Nach.*, **291**, 49.

A lot of the material in §§10.2, 10.3 10.4 and 10.5 is adapted from the following recent review paper:

Charbonneau, P., 2005, *Liv. Rev. Sol. Phys.*, **2**,

which is available online at

`http://solarphysics.livingreviews.org/Articles/lrsp-2005-2/`

See also these other recent review papers:

[17]OK hold it, how do you reconcile this statement with the near independence of the cycle period on C_α for the periods of α-quenched models plotted in Fig. 10.3B (dimonds)?

Ossendrijver

The literature on mean-field solar cycle models is immense. The following is a short list of "classics":

Lerche, I., & Parker, E.N. 1972, *Astrophys. J.*, **176**, 213,
Yoshimura, Y. 1975, *Astrophys. J.*, **201**, 740,
Ivanova, T.S., & Ruzmaikin, A.A. 1976, *Sov. Astron.*, **20**, 227,
Stix, M. 1976, *Astron. Ap.*, **47**, 243,

On the impact of meridional circulation on dynamo waves, see

Roberts & Stix 1972
Choudhuri, A.R., Schüssler, M., & Dikpati, M. 1995, A&A309L29.
Schulz & Rüdiger, G.

The meridional circulation profile described in §10.1 is the creation of

van Ballegooijen, A.A., & Choudhuri, A.R. 1988, *Astrophys. J.*, **333**, 965.

What is now refered to as Babcock-Leighton solar-cycle models go back to the following three seminal papers by H. Babcock and R. Leighton:

Babcock, H.W. 1961, *Astrophys. J.*, **133**, 572,
Leighton, R.B. 1964, *Astrophys. J.*, **140**, 1547,
Leighton, R.B. 1969, *Astrophys. J.*, **156**, 1.

Although some details of the model are different, the 2D surface simulations described in §10.3.1 basically follow

Wang, Y.-M., Nash, A.G., & Sheeley, N.R. Jr 1991, *Science*, **245**, 712,
Wang, Y.-M., & Sheeley, N.R. Jr 1991, *Astrophys. J.*, **375**, 761.

The formulation of the Babcock-Leighton solar cycle model of §**??** is identical to

Dikpati, M., & Charbonneau, P. 1999, *Astrophys. J.*, **518**, 508.

For different modeling approaches, see

Wang, Y.-M., & Sheeley, N.R. Jr, & Nash, A.G. 1991, *Astrophys. J.*, **383**, 431,
Durney, B.R. 1995, *Solar Phys.*, **160**, 213,
Durney, B.R. 1997, *Astrophys. J.*, **486**, 1065.
Choudhuri, A.R., & Nandy, D. 2001, ApJin press?.

The latter paper, in particular, explores the correspondence between Durney's discrete eruption approach, and the mean-field-like formulation used in §10.3.

Figure 10.10: {F6.3} Evolution of the signed (solid line) and unsigned (dashed line) magnetic flux for the advection-diffusion solution of Fig. 10.9. Note the rapid initial decay of the unsigned flux, followed by a slower decay phase.

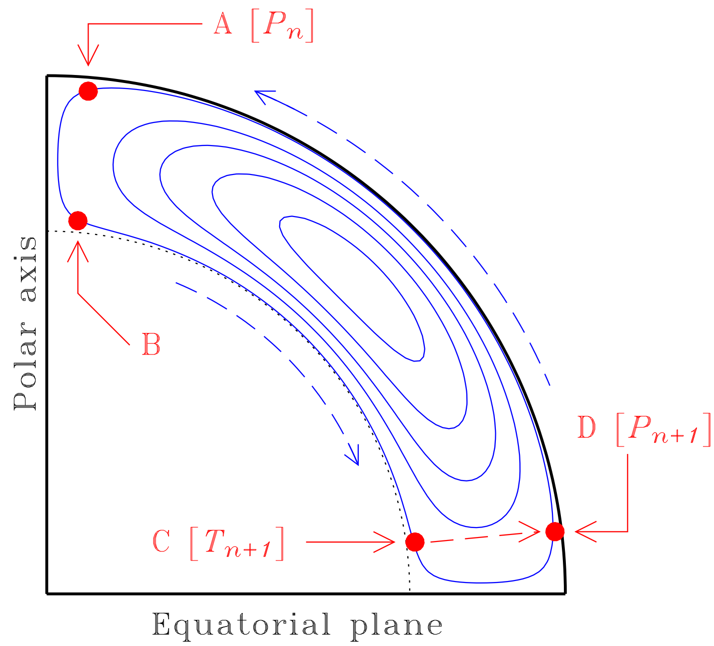

Figure 10.11: Operation of a solar cycle model based on the Babcock-Leighton mechanism. The diagram is drawn in a meridional quadrant of the sun, with streamlines of meridional circulation plotted in blue. Poloidal field having accumulated in the surface polar regions ("A") at cycle n must first be advected down to the core-envelope interface (dotted line) before production of the toroidal field for cycle $n+1$ can take place (B→C). Buoyant rise of flux rope to the surface (C→D) is a process taking place on a much shorter timescale. {`fig:BLcbelt`}

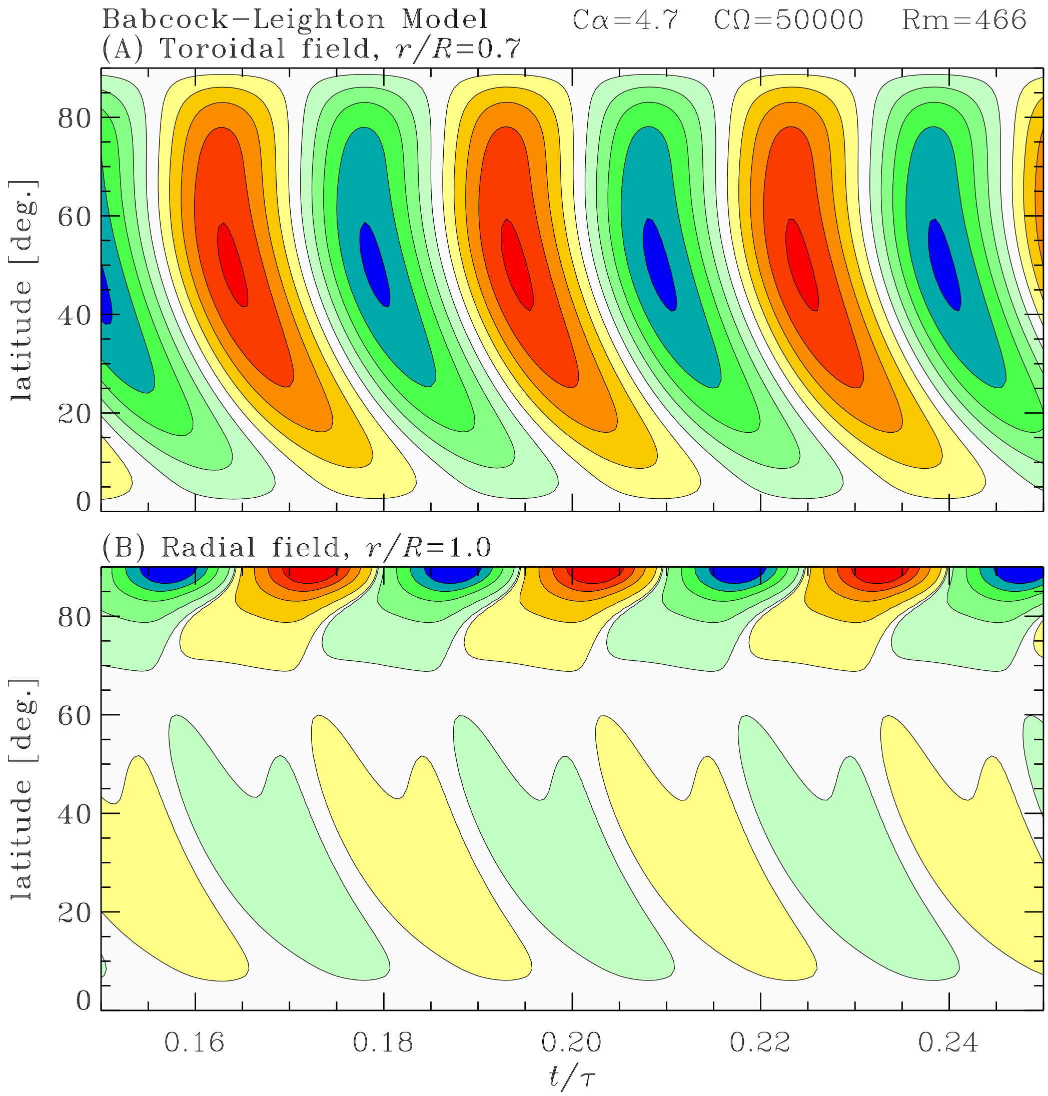

Figure 10.12: Time-latitude diagrams of the surface toroidal field at the core-envelope interface (panel A), and radial component of the surface magnetic field (panel B) in a Babcock-Leighton model of the solar cycle. This solution is computed for solar-like differential rotation and meridional circulation, the latter here closing at the core-envelope interface. The core-to-envelope contrast in magnetic diffusivity is $\Delta\eta = 1/300$, the envelope diffusivity $\eta_T = 2.5 \times 10^{11}\,\mathrm{cm}^2\ \mathrm{s}^{-1}$, and the (poleward) mid-latitude surface meridional flow speed is $u_0 = 16\,\mathrm{m\ s}^{-1}$. {`fig:BLsoln`}

Chapter 11

Stellar dynamos

{chap:dynstar}

The problem —and the beauty— with the Sun is that it overwhelms us with data. Many of the intricacies we have busied ourselves in the preceding chapter were directly motivated by the detailed observations and magnetic measurements made possible by the sun's astronomical proximity. The sun remains for sure an exemplar, but with other stars observational contraints are much more sparse, and theoretical considerations take on an enlarged role.

So, it's back to basics. What have we learned in the preceding three chapters about dynamo action in electrically conducting fluids? At the most fundamental level, a top-three list could run as follows:

- We learned in chapter 7 that rotation, and especially differential rotation, is one very powerful mechanism allowing to build a large-scale magnetic field;
- We learned in chapter 8 that flows with chaotic trajectories, such as arising from strongly turbulent convection, can act as dynamos;
- We learned in chapter 9 that in turbulent flows, the presence of rotation and stratification can break rotational symmetry and produce a self-amplifying large-scale magnetic field.

So, offhand we are not in too bad a shape with regards to stellar dynamos. Stars certainly are stratified, and certainly rotate. Thermally-driven convection is also present across large-part of the HR diagram, but here we start to encounter complications that restrict the use of the "solar exemplar". Figure 11.1 illustrates, in schematic form, the internal structure of main-sequence stars, more specifically the presence or absence of convection zones. A G-star like the Sun has a thick outer convection zone, spanning the outer 30% in radius in the solar case. As one moves down to less massive stars, the relative thickness of the convective envelope increases until, somewhere in the M spectral range, stars become fully convective. Exactly at what mass the radiative core disappears depends on metallicity, opacities, and so on. Moving instead from the Sun to higher masses, the convective envelope becomes ever thinner, until somewhere around spectral-type A0 it essentially vanishes. However, at around the same spectral type Hydrogen burning switches from the *p-p* chain to the CNO cycle, for which nuclear reaction rates are much more sensitively dependent on temperature. Core energy release becomes strongly depth-dependent, leading to a steep —and convectively unstable— temperature gradient. This produces a small convective core, which grows in size as one moves up to larger masses. In a "typical" B-star of solar metalicity, the convective core spans the inner 25% or so in radius of the star.

From these simple considerations, A-stars immediately stand out as the least likely to support dynamo action, because they lack a convective region of substantial size. This squares well with various lines of observations; in particular, main-sequence A-stars are amongst the most "magnetically quiet" stars in the HR diagram, as far as things like X-Ray emission and flaring is concerned. Indeed, the chemically peculiar Ap stars discussed in 2 do show strong magnetic fields, but even those show no sign of anything even mildly analogous to solar activity. This is why to this day the fossil field hypothesis remains the favored explanatory model for the magnetic field of Ap stars.

Until strong evidence to the contrary is brought to the fore, we are allowed to assume that late-type stars with a thick convective envelopes overlying a radiative core host a solar-type dynamo. This is buttressed by the observation of solar-like cyclic activity in many such stars (as briefly discussed already many, many pages ago in §2.1.4). We will therefore begin (§**??**) by looking into the way(s) the various types of solar-cycle models considered in the preceding chapter can be "scaled" to other solar-type stars, of varying masses, rotation rates, etc.

With fully convective stars, we encounter potential deviations from a solar-type dynamo mechanism; without a tachocline and radiative core to store and amplify toroidal flux ropes, the Babcock-Leighton mechanism becomes problematic. Mean-field models based on the turbulent α-effect remain viable, but as we shall see in §11.1 below the dynamo behavior becomes dependent on the presence and strength of differential rotation.

Finally, at the other end of the main-sequence mass range, i.e. O and B stars, the presence of a turbulent convective core combined with high rotation (viz. §5.3) makes dynamo action more than likely. As we shall see in §11.2 below, the challenge is actually to bring the magnetic field produced in the core to the surface.

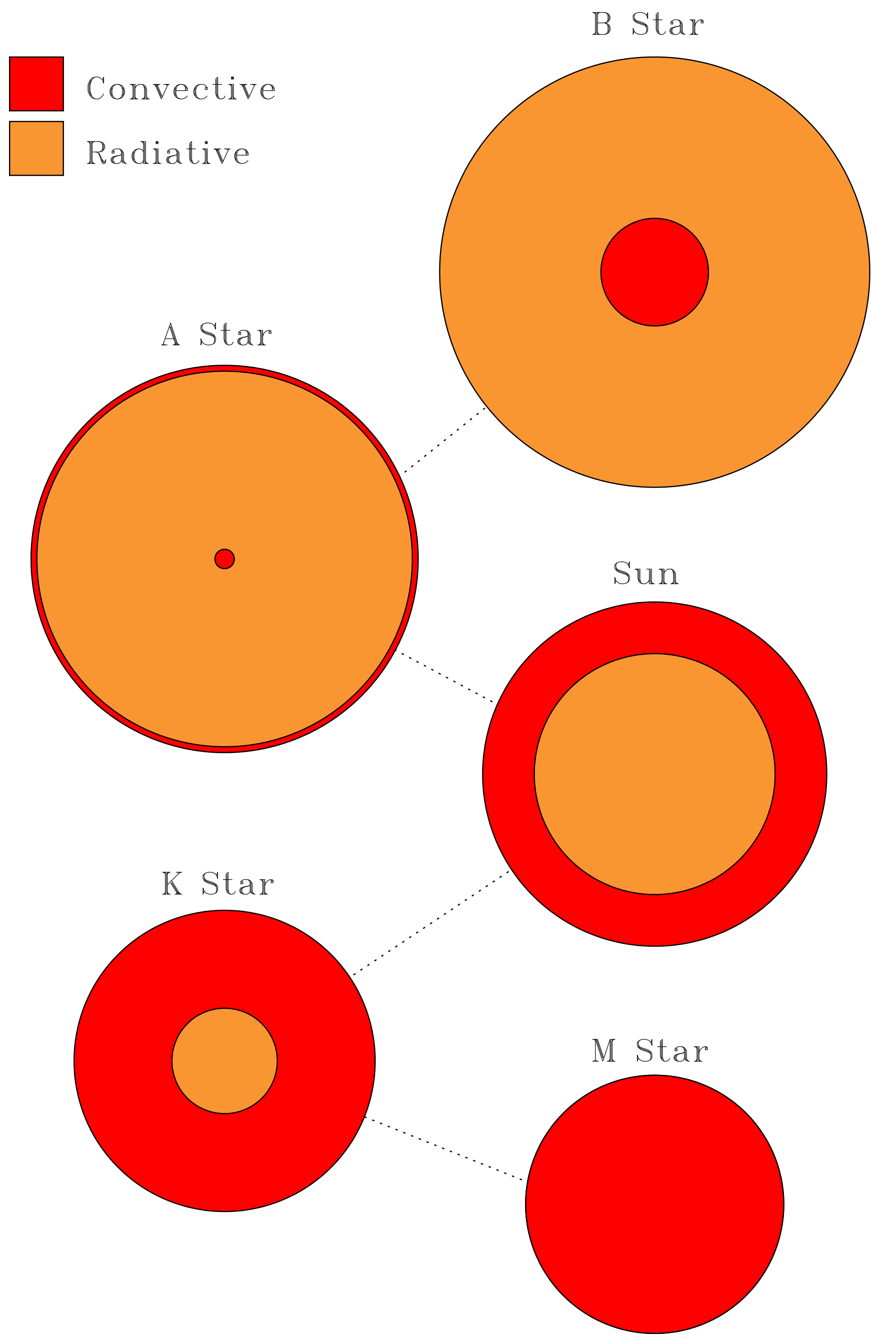

Figure 11.1: {fig:msstruct} Schematic representation of the radiative/convective internal structure of main-sequence stars. The thickness of the outer convection zone for the A-star is here greatly exaggerated; drawn to scale it would be thinner than the black circle delineating the stellar surface on this drawing. Relative stellar sizes are also not to scale.

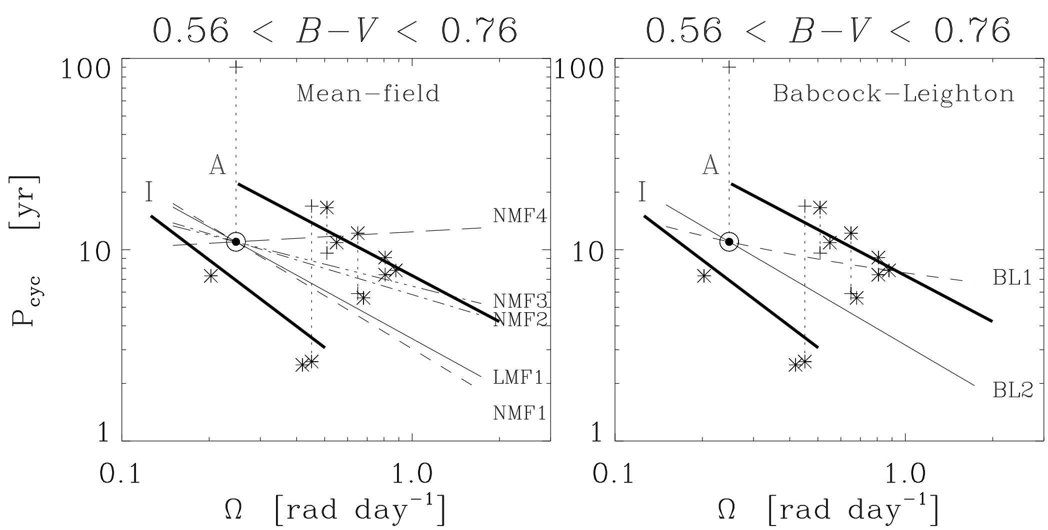

Figure 11.2: {fig:ls1} Cycle Period-Rotation rate relationship in solar-type stars of the Mt Wilson sample for which good cycle period determinations are available.

11.1 Late-type stars other than the Sun {sec:fullconv}

Making dynamo models for these stars is your end-or-year class project! Details provided in early November...

11.2 Early-type stars {sec:mstardyn}

In this section[1] we consider a set of representative mean-field dynamo calculations pertaining to the convective core of a $9M_\odot$ ZAMS stellar model, with luminosity $L = 3767L_\odot$, effective temperature $T_{\rm eff} = 23,600\,$K, and radius $R = 3.678R_\odot$ (spectral type B2). The radius of the convective core (r_c) in this model is at $r_c = 0.232R$. Within the core, thermally-driven turbulent fluid motions are assumed to give rise to an α-effect and turbulent diffusivity, which both vanish for $r \gtrsim r_c$ (under the assumption that the radiative envelope is turbulence free). In the spirit of the other dynamo models discussed in this chapter, we consider kinematic dynamos with parametric profiles for α and η:

$$\alpha(r,\theta) = \frac{1}{2}\left[1+\mathrm{erf}\left(\frac{r-r_c}{w}\right)\right]\mathrm{erf}\left(\frac{2r}{w}\right)\cos(\theta)\ , \qquad (11.1) \quad \{E1.15a\}$$

$$\eta(r) = \eta_e + \frac{\eta_c-\eta_e}{2}\left[1-\mathrm{erf}\left(\frac{r-r_c}{w}\right)\right]\ , \qquad (11.2) \quad \{E1.15b\}$$

[1]The set of dynamo solutions presented here are all taken directly from the Charbonneau & MacGregor 2001 paper cited in the bibliography.

where $\mathrm{erf}(x)$ is once again the error function. Equations (11.1) represent "minimal" assumptions on the spatial dependency of the α-effect: it changes sign across the equator ($\theta = \pi/2$), vanishes at $r = 0$, rises to a maximum value within the convective core, and falls again to zero for $r \gtrsim r_c$, the transition occurring across a spherical layer of thickness $\sim 2w$. we consider models with both positive and negative α-effect. The $\cos\theta$ dependency in eq. (11.1) is the latitudinal variation of the Coriolis force felt by a radially rising/sinking fluid element, and in the context of mean-field theory represents the "minimal" θ-dependency for the α-effect.

Various lines of argument related to the rotational evolution of early-type stars suggest that significant differential rotation may exist between the convective core and overlying radiative envelope. In what follows we restrict ourselves to the (simple) case of a convective core and radiative envelope both rotating rigidly but at different rates Ω_c, Ω_e, joined smoothly across a thin spherical shear layer coinciding with the core-envelope interface at $r = r_c$:

$$\Omega(r,\theta) = \Omega_c + \frac{\Omega_e - \Omega_c}{2}\left[1 + \mathrm{erf}\left(\frac{r - r_c}{w}\right)\right] \ . \tag{11.3}$$ {E3.1}

The rotation increases inward, i.e., $\Omega_c > \Omega_e$, leading to a negative radial shear in the vicinity of the core-envelope interface[2]. The parameter w used to specify the thickness of the shear layer is the same as that used to specify the width of the transition region for the turbulent diffusivity and α-effect. We are now solving the dynamo equations in their $\alpha^2\Omega$ (9.69)–(9.70) with $\mathrm{R}_m = 0$ but with all other terms present. All dynamo solutions discussed below are obtained as eigenvalue problems, as in §**??**. Remember that such linear solutions leave the absolute scale of the magnetic field unspecified.

An interesting physical quantity accessible from linear models is the ratio of the surface field field strength to the field strength in the dynamo region, here the convective core. In what follows we use towards this purpose the ratio (Σ) of the r.m.s. surface poloidal field to the r.m.s. poloidal field at the core-envelope interface r_c:

$$\Sigma = \left(\frac{R^2\int|\nabla\times A|^2_{r=R}\sin\theta\mathrm{d}\theta}{r_c^2\int|\nabla\times A|^2_{r=r_c}\sin\theta\mathrm{d}\theta}\right)^{1/2} \ . \tag{11.4}$$ {E1.16b}

In practice, the finite numerical accuracy at which the eigenfunctions are computed leads to a lower bound on meaningful values of Σ, here at about 10^{-8}.

[2]Negative radial shear profiles are the only ones considered here, since steep positive radial shears are in all likelihood hydrodynamically unstable.

11.2.1 α^2 dynamos

{Sa2}

We first consider solutions where magnetic field generation occurs exclusively through the agency of the α-effect, i.e., α^2 dynamo models, in the terminology introduced in §9.4.3. Figure **??**fig:ms1 shows a series of typical linear α^2 solution with increasing diffusivity contrasts between the core and envelope. The value of C_α for the solutions on panels B, C and D were adjusted to yield solutions with growth rates similar to that of the contant-η solution in A, so that the four eigenfunctions are in some sense comparable.

The constant-η solution transits from decaying ($\sigma < 0$) to growing ($\sigma > 0$) at $C_\alpha \simeq -32.8$, and the growth rate keeps increasing as $|C_\alpha|$ is further increased. The solution plotted on Fig. 11.32A is computed for $C_\alpha = -34.5$, and is supercritical ($\sigma = 10.82\tau^{-1}$). A solution with $C_\alpha = +34.5$ has an identical growth rate and eigenfunction, but shows an opposite relative polarity between the poloidal and toroidal components. For $\eta_e/\eta_c \lesssim 0.1$, the symmetric modes now have slightly larger growth rate ($\sigma = 11.0$, 11.8, and $12.5\,\tau^{-1}$ for $\eta_e/\eta_c = 0.1$, 0.01, and 0.001, respectively). Nonetheless, to facilitate comparison with the constant diffusivity solution of part A, the antisymmetric modes are plotted on Figure 11.3B–D. Defining parameters for all solutions plotted on Fig. 11.3 are listed in the top part of Table 1 below.

Linear mean-field dynamo of the α^2 type with a time-independent scalar functional $\alpha(r)$ always produce *steady* magnetic fields, i.e., the solution eigenvalue is purely real ($\omega = 0$ in eq. (10.28)). The solution plotted on Figure 11.3A is dipole-like (i.e., antisymmetric), and is the fastest growing solution for our model with constant η, at the adopted value for C_α. The next fastest growing mode is symmetric with respect to the equatorial plane, and has a growth rate only slightly smaller, $\sigma = 10.79\,\tau^{-1}$. This situation is typical of α^2 dynamo solutions using a scalar α-effect[3]. Note that $\sigma = 10$ in dimensionless units amounts to an e-folding time of about 20 yr in dimensional units, leaving no doubt that ample time is available to amplify a weak seed magnetic field in the core of a massive star.

Table 1

Parameters and eigenvalues for various α^2 and $\alpha^2\Omega$ solutions

[3]The α^2 form of the mean-field dynamo equations also admits growing solutions than are non-axisymmetric even though the α-effect profile exhibits axisymmetry with respect to the rotation axis. Growth rates for non-axisymmetric modes are often comparable to those of their axisymmetric counterparts For simplicity, we restrict ourselves here to axisymmetric modes. We note nonetheless that, motivated largely by the challenge posed by planetary magnetic fields, α^2 models can and have been constructed where non-axisymmetric modes are the fastest growing, and dominate in the moderately supercritical nonlinear regime (see, e.g., Rädler *et al.* 1990).

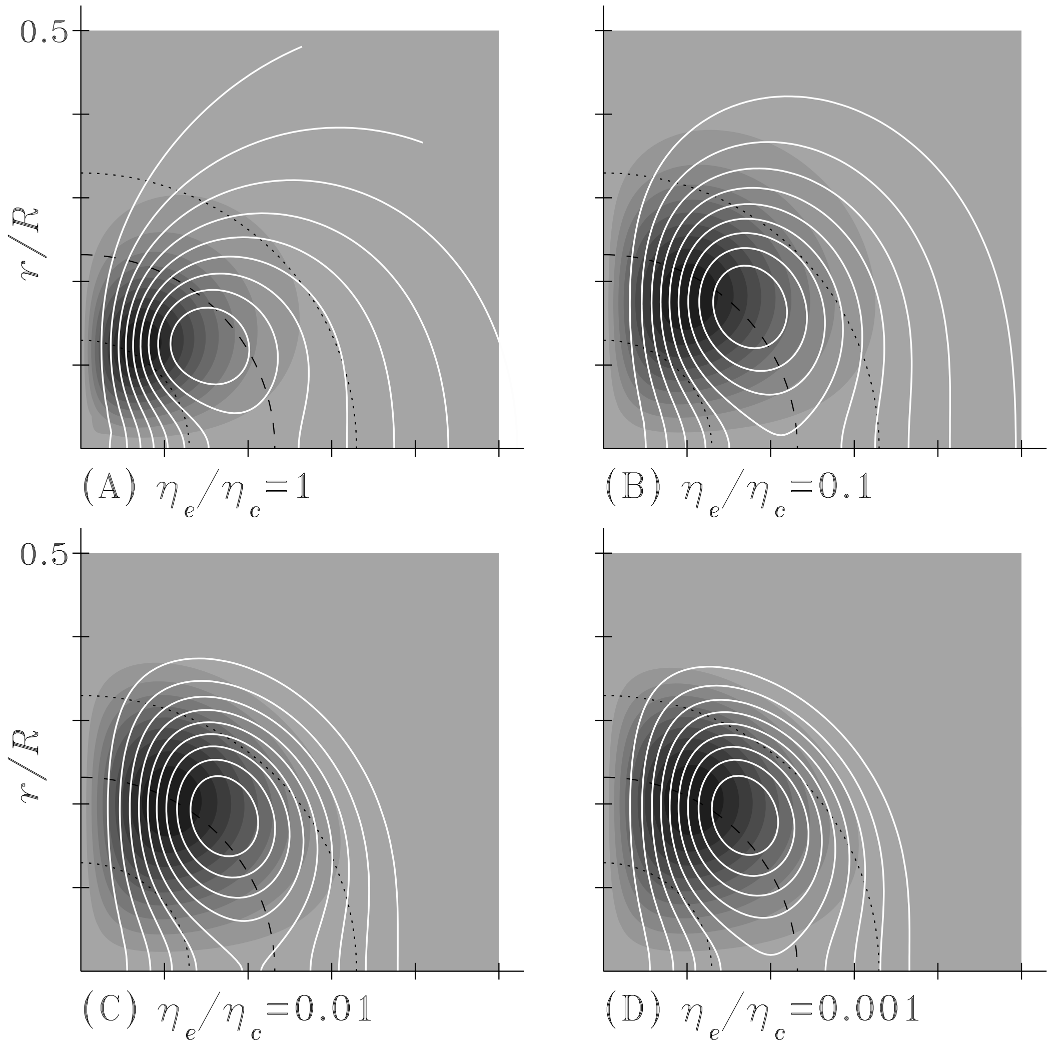

Figure 11.3: Four antisymmetric steady α^2 dynamo solutions, computed using varying magnetic diffusivity ratios between the core and envelope. The solutions are plotted in a meridional quadrant, with the symmetry axis coinciding with the left quadrant boundary. Poloidal fieldlines are plotted superimposed on a gray scale representation for the toroidal field (light to dark is weaker to stronger field). The dashed line marks the core-envelope interface depth r_c, and the two dotted lines indicates the depths $r_c \pm w$ corresponding to the width of the transition layer between core and envelope. Note how the solutions with $\eta_e/\eta_c \lesssim 10^{-2}$ have their toroidal field peaking across the core-envelope interface. This behavior is generic and materializes for smaller values of w and r_c, and for symmetric (i.e., quadrupolar-like) solutions. Parameters for these solutions are listed in Table 1. {`fig:ms1`}

Type	Parity	C_α	C_Ω	η_e/η_c	w/R	σ	ω	Σ^b
α^2	A	-34.5	0	1	0.1	10.8	0	1.2×10^{-2}
α^2	A	-23.0	0	0.1	0.1	8.96	0	2.7×10^{-4}
α^2	A	-21.0	0	0.01	0.1	9.91	0	$< 10^{-8}$
α^2	A	-21.0	0	0.001	0.1	10.58	0	$< 10^{-8}$
$\alpha^2\Omega$	A	-21.0	2000	0.01	0.1	14.6	175	$< 10^{-8}$
$\alpha^2\Omega$	S	-21.0	2000	0.01	0.1	21.8	186	$< 10^{-8}$
$\alpha^2\Omega$	A	$+21.0$	2000	0.01	0.1	21.2	184	$< 10^{-8}$
$\alpha^2\Omega$	S	$+21.0$	2000	0.01	0.1	14.0	172	$< 10^{-8}$
$\alpha^2\Omega$	S	-24.0	2000	0.01	0.05	19.9	287	$< 10^{-8}$
$\alpha^2\Omega$	S	-35.0	2000	0.01	0.025	17.7	494	$< 10^{-8}$

{tab:mstardyn}

The most significant consequence of a η_e/η_c being smaller than one is perhaps the "trapping" of the magnetic field in the lower part of the radiative envelope, a direct consequence of the difficulty experienced by an external magnetic field to diffusively penetrate a good electrical conductor. This is clearly evident from Table 1, in the rapid decrease of the surface-to-core field ratio Σ (see eq. (11.4)) with decreasing diffusivity ratio η_e/η_c. This is long-recognized property of stellar core dynamos (e.g., Schüssler & Pähler 1978), and represents a rather formidable obstacle to be bypassed if the magnetic fields generated by dynamo action in the convective core are to become observable at the stellar surface. As discussed in Schüssler & Pähler, the situation is even worse than Table 1 may suggest. In a time-dependent situation, the time needed for the magnetic field to resistively diffuse to the surface can become larger than the star's main-sequence lifetime, for masses in excess of about $5\,M_\odot$.

Less striking but equally important in what follows is the fact that in solutions with $\eta_e/\eta_c < 1$, the locus of peak dynamo action —as measured by the peak in toroidal field strength— moves out to the core-envelope boundary. Note on Fig. 11.3 how, for $\eta_e/\eta_c \lesssim 0.01$, toroidal fields are present out to $r \simeq r_c + w$. This is a direct consequence of the α/η ratio remaining equal to unity over a significant radial distance outside of the core, as per eqs. (11.1)—(11.2). As η_e/η_c decreases, the magnetic field is increasingly trapped in the interior, yet is increasingly concentrated near the core-envelope interface. This behavior is robust, in that it also materializes in solutions computed using different parameter values.

Mean-field dynamo models of the α^2 variety typically generate magnetic fields that have poloidal and toroidal components of comparable strengths. Indeed we find here that the toroidal-to-poloidal ratio defined in eq. (6.9) are of order unity and vary very slightly with η_e/η_c (see Table 1).

11.2.2 $\alpha^2\Omega$ and $\alpha\Omega$ dynamos

{Sdr}

Perhaps the most significant difference between $\alpha^2\Omega$ solutions and the α^2 solutions considered previously is the fact that while the latter are spatially steady (in the sense that $\omega = 0$), the former usually yield oscillatory solutions, with solution eigenvalues occurring in complex conjugate pairs $\sigma \pm i\omega$.

Figure 11.4 illustrates a half-cycle of a representative $\alpha^2\Omega$ solution. This symmetric solution has $C_\alpha = -21$, $C_\Omega = 2000$, $w/R = 0.1$, $\eta_e/\eta_c = 10^{-2}$, and is characterized by a growth rate $\sigma = 21.8\,\tau^{-1}$ and frequency $\omega = 186\,\tau^{-1}$. For $\eta_c = 10^{13}\,\mathrm{cm}^2\,\mathrm{s}^{-1}$, this corresponds to a dynamo period of about 7 yr, quite short compared to any other relevant timescales. The magnetic field distribution is shown at five distinct phases, at constant intervals of $\Delta\varphi = \pi/4$, in a format identical to that of Fig. 11.3 for each panel (note in particular that the eigenmodes are again plotted only in the inner half of the star). At a given phase the solutions bear some resemblance to the α^2 solutions of Fig. 11.3C, in that the magnetic field is again trapped in the interior. As before, the toroidal field is concentrated near the core-envelope interface, and in fact here peaks slightly outside $r = r_c$ (dashed circular arc).

As with the α^2 solutions considered previously, the growth rate of the $\alpha^2\Omega$ solution increases with increasing values of either or both the dynamo numbers C_α and C_Ω. The dynamo frequency ω also increases with C_α and C_Ω. In the $\alpha\Omega$ limit, where the α-effect makes a vanishing contribution to the RHS of eq. (9.70), the eigenvalue is completely determined by the value of the product $C_\alpha C_\Omega$, but this property does not hold in general for $\alpha^2\Omega$ models.

Examination of Figure 11.4 soon reveals that the magnetic field distribution migrates steadily poleward in the course of the half-cycle shown on Figure 3, with the solutions at $\varphi/\pi = 1$ being a mirror image of that at $\varphi/\pi = 0$, i.e., the magnetic polarity has undergone a polarity reversal after half an oscillation cycle. This is the "dynamo wave" we already encountered previously, and indeed the poleward propagation observed here is what one would expect from a negative radial shear acting in conjunction with a negative α-effect (cf. §9.3). [4]. Note that the toroidal field gains in strength as the dynamo wave proceeds from low to mid-latitudes, peaking at about 60° and falling thereafter as the wave experiences enhanced dissipation upon converging toward the symmetry axis.

A solution with $C_\alpha = +21$ but otherwise identical to that shown on Fig. 11.4 has growth rate and frequency that are comparable to, but not

[4] Parker's original dynamo wave solutions were obtained in Cartesian geometry, and in the so-called $\alpha\Omega$ limit, in which the α-effect is omitted on the RHS of the toroidal component of the dynamo equation. Similar dynamo wave solutions are also readily found in the more general $\alpha^2\Omega$ case; see for example Choudhuri (1990).

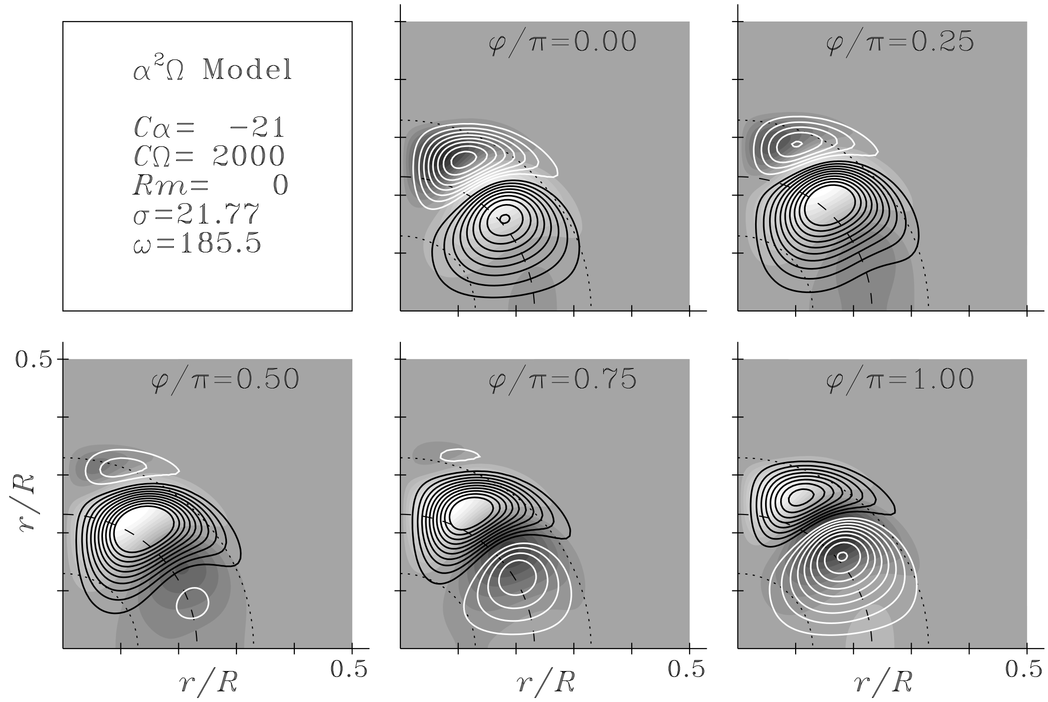

Figure 11.4: A representative $\alpha^2\Omega$ solution. As this is an oscillatory solution, the eigenfunction is plotted at five equally spaced phase intervals ($\Delta\varphi = \pi/4$), covering half an oscillation cycle. The format in each panel is similar to Fig. 11.3. White (black) lines indicate fieldlines oriented in a clockwise (counterclockwise) direction. Note the wave-like propagation of the magnetic field from low to high latitudes. Parameter values are listed in Table 1. {Fig3}

identical to the $C_\alpha = -21$ solution (see Table 1). The difference is due to spherical geometry; a poleward-propagating dynamo wave suffers greater diffusive decay as it converges towards the symmetry axis, than an equatorward propagating wave does converging towards the equatorial plane, where the symmetry imposed via the boundary condition also affects the dissipation. Solutions with thinner transition layers require a larger value of $|C_\alpha|$ to maintain comparable growth rates, and are thus characterized by higher oscillation frequencies. Table 1 lists solution parameters and characteristics for a few representative such solutions.

Not surprisingly, in $\alpha^2\Omega$ models the availability of an additional energy source in the toroidal component of the dynamo equations leads to solutions where the toroidal field strength in general exceeds that of the poloidal field. For the solution plotted on Fig. 11.4, the toroidal-to-poloidal field ratio (see eq. (6.9)) reaches a value $\Theta \simeq 3$. Further increases of C_Ω lead to increasing Θ (e.g., $\Theta \simeq 3.4$ and 4.3 at $C_\Omega = 5000$ and 10^4), until in the $\alpha\Omega$ limit Θ

scales roughly as C_Ω/C_α. For a given diffusivity ratio η_e/η_c, oscillatory $\alpha^2\Omega$ solutions have a smaller surface-to-core field strength ratio Σ than α^2 models, a direct consequence of the oscillatory nature of the field, which restricts the radial extent of the eigenfunction above the core-envelope interface to a distance comparable to the electromagnetic skin depth, which is very much smaller than the stellar radius for $\eta_e/\eta_c \ll 1$.

The markedly different spatial distributions and temporal behavior of α^2 and $\alpha^2\Omega$ eigenmodes naturally leads one to suspect that both dynamo modes should have some difficulty operating simultaneously. That this is indeed the case can be seen in Figure 11.5, showing isocontours of the linear growth rate σ in the $[C_\Omega, C_\alpha]$ plane, for antisymmetric negative-C_α solutions. Dynamo solutions ($\sigma > 0$) are located below the thick contour, and the thick dashed line delineates the regions where steady ($\omega = 0$, α^2-like) and oscillatory ($\omega \neq 0$, $\alpha^2\Omega$-like) solutions are found. At a fixed value of C_α, introducing differential rotation first leads to a *decrease* of the growth rate, reflecting the perturbative influence of differential rotation on the basic α^2 mode. Once C_Ω exceeds a certain (C_α-dependent) threshold at about $C_\Omega \simeq 300$, the dynamo becomes $\alpha^2\Omega$-like ($\omega \neq 0$). However, growth rates comparable to that of the pure α^2 mode ($C_\Omega = 0$) materialize only for much larger values of C_Ω. Much the same behavior is seen in symmetric solutions, and/or for positive-C_α solutions. Nonetheless, the transition from the α^2 to the $\alpha^2\Omega$ dynamo regime occurs smoothly as differential rotation is increased.

11.3 Getting the magnetic field to the surface {sec:totop}

For our adopted value $\eta_c = 10^{13}\,\mathrm{cm}^2\,\mathrm{s}^{-1}$, $C_\Omega = 300$ amounts to $\Delta\Omega_0/\Omega_* \simeq 10^{-3}$, i.e., very weak differential rotation. The extant observations and inferences of magnetic fields in upper main-sequence stars reviewed in §**??** currently have little to say about the steady/oscillatory character of the underlying field. Even if it were oscillating with a regular period of the order of a few years, as do the $\alpha^2\Omega$ solutions discussed here, it is not at all clear that the mechanism(s) responsible for bringing the field to the surface may not introduce additional temporal variabilities that would mask the underlying cycle period. If on the other hand the magnetic fields are shown to be strictly steady, one would then be forced to conclude that the same magnetic fields have obliterated any angular velocity difference between the core and envelope, something which they can in fact achieve quite efficiently in the absence of internal or external forcing.

Problems:

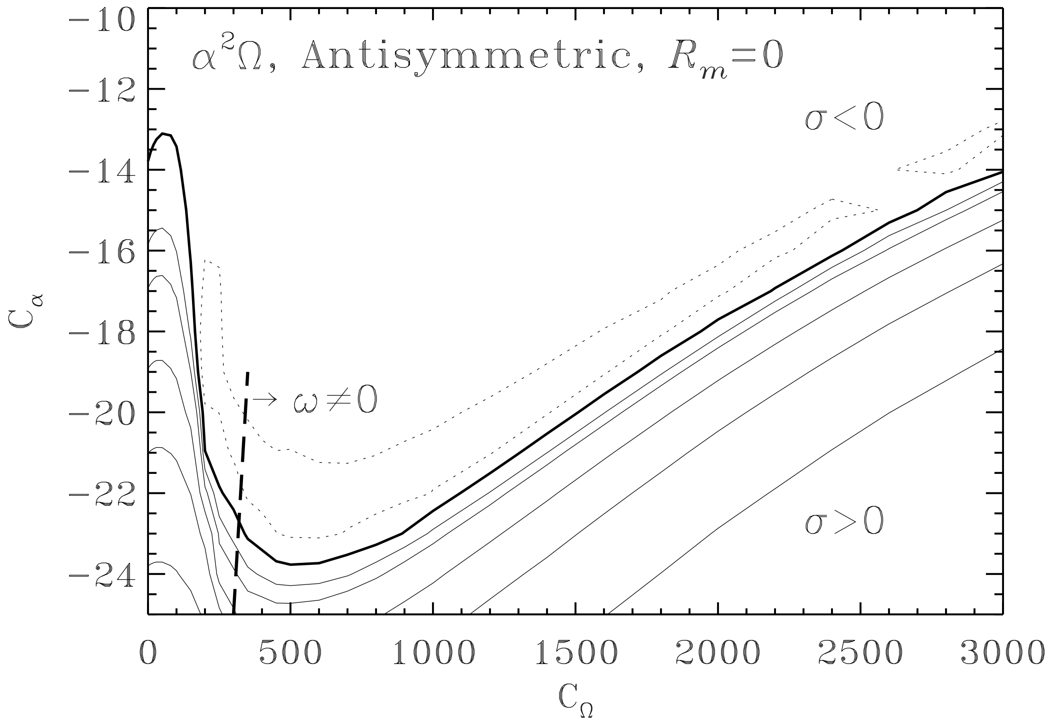

Figure 11.5: Isocontours of the $\alpha^2\Omega$ linear growth rate in the $[C_\Omega, C_\alpha]$ plane. The thicker contours corresponds to $\sigma = 0$, and solid contours to $\sigma > 0$. All solutions are of antisymmetric parity and have $w/R = 0.1$, $\eta_e/\eta_c = 0.01$, and $\mathrm{R}_m = 0$. Solutions left of the thick-dashed line are steady ($\omega = 0$, α^2-like), and oscillatory to its right. Qualitatively similar diagrams are obtained for symmetric modes, other values of w/R, and/or solutions with positive C_α. {Fig4}

Bibliography: