Solar Physics and Magnetohydrodynamics

Axel Brandenburg
Nordita, AlbaNova University Center, Roslagstullsbacken 23
SE-10691 Stockholm, Sweden
http://www.nordita.org/~brandenb
brandenb@nordita.org

September 28, 2014

Contents

A very good text book on the overall problem of solar physics is:

Stix, M.: 1989, *The Sun: An introduction*. Springer-Verlag, Berlin Heidelberg

The present text tries to follow this book in many aspects.
It has in part been assembled using earlier notes; see
`http://www.nordita.org/~brandenb/teach/MAS218/notes.ps.gz`
`http://www.nordita.org/~brandenb/teach/MAS371/notes.ps.gz`
`http://www.nordita.org/~brandenb/tmp/phys_rep/`
`http://www.nordita.org/~brandenb/tmp/solint/`
In order to find scientific papers on the web, an important web site is ADS,
`http://adsabs.harvard.edu/abstract_service.html`
Here you find published papers in astronomy as well as preprints on the archive, `arXiv`; see `http://arxiv.org/`. (Looking at the `arXiv` every morning is a very

good idea!) Optionally, you can also click on the physics button, because sometimes relevant papers are only listed there. Another important address is the Web of Science, `http://scientific.thomson.com/products/wos/` but the access requires a license. I should also recommend looking every morning at the astronomy picture of the day `http://antwrp.gsfc.nasa.gov/apod/astropix.html`.

Table 1: Summary of some important physical constants

Newton's constant	G	6.67×10^{-11}	$\mathrm{m^3\,kg^{-1}\,s^{-1}}$
speed of light	c	3.00×10^{8}	$\mathrm{m\,s^{-1}}$
radiation constant	a	7.56×10^{-16}	$\mathrm{J\,m^{-3}\,K^{-4}}$
Stefan-Boltzmann constant	$\sigma_{\mathrm{SB}} = ac/4$	5.67×10^{-8}	$\mathrm{W\,m^{-2}\,K^{-4}}$

Table 2: Solar parameters

distance	$R_{\mathrm{E}\odot}$	1.5×10^{11}	m
radius	$M_\odot$	7×10^{8}	m
mass	$R_\odot$	2×10^{30}	kg
luminosity	$L_\odot$	4×10^{26}	W

Table 3: Derived solar parameters

average density	$\overline{\rho} = M_\odot/(\frac{4\pi}{3}R_\odot^3)$	1.4×10^{3}	$\mathrm{kg\,m^{-3}}$
surface gravity	$g = GM/R_\odot^2$	2.7×10^{2}	$\mathrm{m\,s^{-2}}$
solar constant	$S = L_\odot/(4\pi R_{\mathrm{E}\odot}^2)$	1.4×10^{3}	$\mathrm{W\,m^{-2}}$
effective temperature	$T_{\mathrm{eff}} = [L_\odot/(4\pi R_\odot^2 \sigma_{\mathrm{SB}})]^{1/4}$	5.8×10^{3}	K

Chapter 1

Introduction

The sun is our primary energy source. But it is not constant. Understanding its variability and especially its magnetic activity is important for many reasons. First, there is now mounting evidence for a clear connection between solar activity and variation of the earth climate; see, e.g., Fig. 1.1. The causal connection between sun and earth is hardly due to temperature changes of the sun (luminosity or irradiance[1] changes), but due to magnetic changes. According to a theory originally due to H. Svensmark this is caused by the solar magnetic field shielding the cosmic ray particles from the galaxy. Thus, an increase in the solar field strength increases the shielding, decreases the cosmic ray flux on earth, decreases the cloud cover and hence increases the temperature, as is illustrated symbolically as

$$B \uparrow \quad \mathrm{CR} \downarrow \quad \mathrm{clouds} \downarrow \quad T_{\mathrm{earth}} \uparrow . \tag{1.1}$$

This picture is rather simplified, and there can be drastic differences regarding clouds that are low or high, but the general idea is now gaining more and more acceptance, it seems.

Second, on a much shorter time scale (days and shorter) we should be concerned about extreme space weather[2] conditions which affect not only astronauts in space, but also the functionality and reliability of man-made satellites in orbit. This in turn affects the Global Positioning System (GPS) and hence also air traffic navigations, which now really becomes everybody's concern.

- Geomagnetic storms: strength G1–G5 (with G5 around 4 per cycle): voltage problems, blackouts, transformer damage.
- Radiation storms: S1–S5 (S5 fewer than 1/cycle): satellites rendered useless, memory impact, permanent damage to solar panels.

[1] Irradiance is what comes out of the sun sideways in the direction of the earth. This may not be quite the same as the flux integrated over 4π. Both are however bolometric fluxes, i.e. integrated over all frequencies.

[2] `http://www.sec.noaa.gov/NOAAscales/`

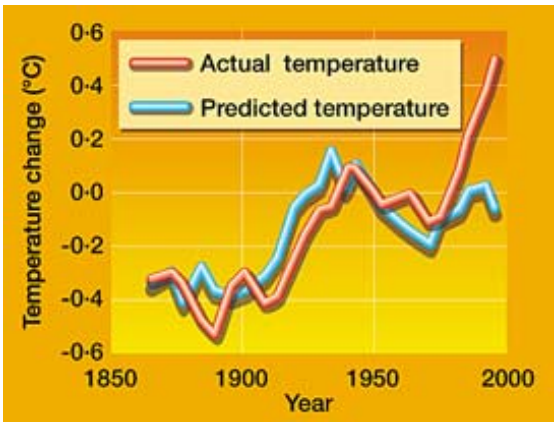


Figure 1.1: Observed temperature changes on earth and those predicted from the variations of solar activity.

- Radio blackouts: R1–R5 (R5 less than 1/cycle): HF blackout for hours.

These space weather conditions are being caused by various kinds of eruptions on the solar surface such as prominences and Coronal Mass Ejections (CMEs); see Fig. 1.2. Modeling all the physics between the actual eruption and the final effects on the ionospheric current system is one of the few Grand Challenge problems for which a lot of money has been set aside[3]. But in order to understand the origin of the eruptions, we have to understand where the magnetic field comes from. As will be explained later in this course, it has its origins in convective motions in the outer 30% of the sun. The drive a magnetic field that manifests itself in terms of an 11 year cycle of the sunspot number (Fig. 1.4) and of the magnetic field. This will be discussed in more detail at the end of this course where this field is explained as a result of a self-excited dynamo process.

Important conclusions about the nature of the energy source of the sun can be gained by looking at the energy budget. The luminosity of the Sun is $L_\odot = 4 \times 10^{26}\,\mathrm{W}$. The total interception with the Earth is

$$\frac{\pi R_E^2}{4\pi R_{E\odot}^2} = 4.4 \times 10^{-10}, \tag{1.2}$$

where R_E is the radius of the Earth (6400 km) and $R_{E\odot}$ is the distance between the Earth and the Sun ($= 1\,\mathrm{AU} = 1.5 \times 10^{11}\,\mathrm{m}$). So the total power reaching the Earth is $4.4 \times 10^{-10} \times 4 \times 10^{26}\,\mathrm{W} = 1.7 \times 10^{17}\,\mathrm{W}$.

For comparison we list here the fractions of energies that are extracted from the sun or that are consumed[4]:

[3] http://ct.gsfc.nasa.gov/grand.st3.html

[4] http://en.wikipedia.org/wiki/World_energy_resources_and_consumption

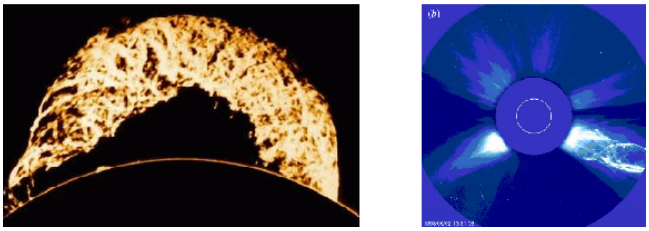


Figure 1.2: The famous "Grand daddy" prominence of 4 June 1946 (left) and a big coronal mass eruption of 2 June 1998 from the LASCO coronograph on board the SOHO satellite (right). Note the complexity of the ejected structures, being suggestive of helical nature. Courtesy of the SOHO consortium. SOHO is a project of international cooperation between ESA and NASA.

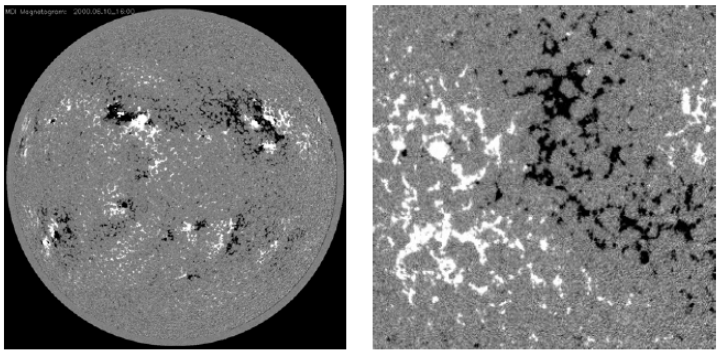


Figure 1.3: Distribution of magnetic flux in the visible solar surface layers (adapted from Solanki et al. 2006).

- 25% of it drives evaporation & cloud formation
- 1% drives winds
- 1% drives get turns into biological energy (photosynthesis)
- 0.01% of the world power consumption ($1.6 \times 10^{13}\,\mathrm{W} = 16\,\mathrm{TW}$)

The gravitational energy of the Sun is approximately

$$E_{\rm th} = \frac{G\overline{M}^2}{R_\odot} \approx \frac{GM_\odot^2}{2R_\odot} \approx 2 \times 10^{41}\,\mathrm{J} \tag{1.3}$$

The time it would take to use up all this energy to sustain the observed lumi-

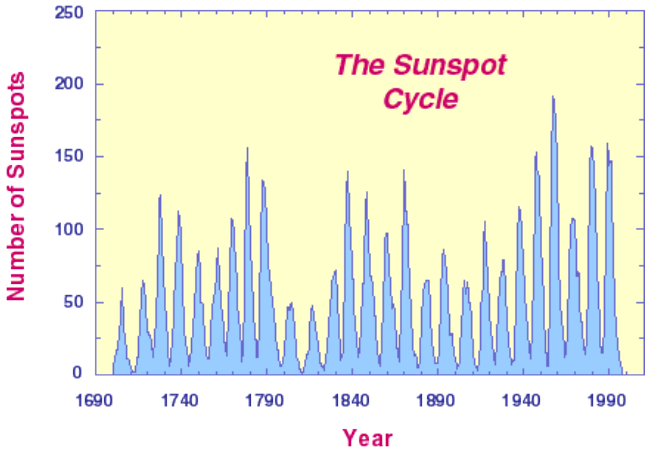


Figure 1.4: Sunspot number (yearly averages); `http://csep10.phys.utk.edu/astr162/lect/sun/sscycle.html`

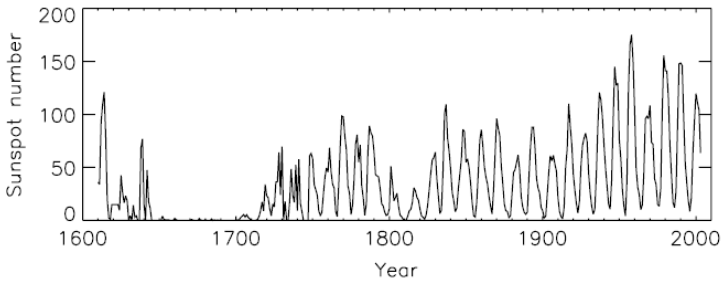


Figure 1.5: The record of yearly averaged group sunspot numbers dating back to 1600 (adapted from Solanki et al. 2006).

nosity is the Kelvin-Helmholtz time (note that $1\,\mathrm{yr} \approx 3 \times 10^7\,\mathrm{s}$)

$$\tau_{\mathrm{KH}} = E_{\mathrm{th}}/L_\odot = \frac{2 \times 10^{41}}{4 \times 10^{26}}\,\mathrm{s} \approx 10^7\,\mathrm{yr} \tag{1.4}$$

which is long compared to time scales we could observe directly, but short compared with the life time of the Sun and the solar system (more like 5×10^9 yr). This led to the discovery of the nuclear energy source of stars. The potential energy is also approximately the thermal energy (Virial theorem).

In fact, this similarity can be used to estimate the central temperature of a star by equating $GM/R = \mathcal{R}T_c/\mu$. Here, $\mathcal{R} = 8315\,\mathrm{m^2\,s^{-2}\,K^{-1}}$ (this is a script or curly R!) is the universal gas constant and μ is the mean molecular weight (in astrophysics this is defined dimensionless while in physics it is often defined

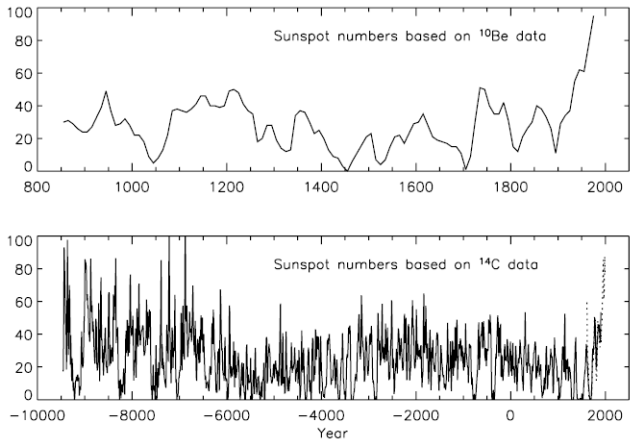


Figure 1.6: Reconstruction of the 10 year averaged sunspot number from the measured concentrations of cosmogenic isotopes (adapted from Solanki et al. 2006).

with units gram/mole), which is the atomic or molecular mass expressed in units of 1 amu (to a good approximation, $\mu_{\rm H} \approx 1$). The quantity $\mathcal{R}T/\mu$ has the dimensions of a velocity squared. For the Sun this yields the following estimate for the central temperature:

$$T_c = \frac{\mu}{\mathcal{R}} \frac{GM_\odot}{R_\odot} = \frac{0.6}{8300} \frac{7\times10^{-11}\,2\times10^{30}}{7\times10^{8}}\,{\rm K} = \frac{3}{4}\,10^{-4}\,10^{-11}\,2{\times}10^{22}\,{\rm K} = 1.5{\times}10^{7}\,{\rm K} \tag{1.5}$$

This estimate is actually spot on (but this is merely luck). It also tell us that the central temperature of the Sun is only determined by its mass and radius, and not, as one might have expected, by the effectiveness of the nuclear reactions taking place there.

The central pressure is given (to an order of magnitude) by

$$p_c = \frac{GM_\odot^2}{R_\odot^4} = \frac{7\times10^{-11}\,(2\times10^{30})^2}{(7\times10^{8})^4}\,{\rm kg\,m^{-1}\,s^{-2}} \approx 10^{15}\,{\rm kg\,m^{-1}\,s^{-2}}, \tag{1.6}$$

which turns out to be about an order of magnitude too small.

1.1 The main points of this chapter

The sun is not static, but shows an 11 year cycle dating back several thousand years (as indicated by measurements of radioactive decaying cosmogenic isotopes). This has tremendous effects on terrestrial climate changes (little ice age, etc).

Chapter 2

Radial structure

The sun is, to an excellent approximation, spherically symmetric. The equations governing the radial structure of the sun (or any other star) are plausible and easily derived. They can be written as a set of four ordinary differential equation, namely the

- equation for the sun's gravity (comes from the Poisson equation),
- hydrostatic equilibrium (comes from the momentum equation),
- thermal equilibrium (comes from the energy equation),
- radiative equilibrium (radiation transport equation, convection).

In the optically thick case, the radiative flux is in the direction of and proportional to the gradient of the radiative energy density, aT^4. (The connection between fluxes and concentration gradients is generally referred to as Fickian diffusion). Like in kinetic gas theory, the diffusion coefficient is 1/3 times the typical particle velocity (=speed of light c) and the mean free path ℓ, so

$$F = -\tfrac{1}{3}c\ell \,\frac{\mathrm{d}}{\mathrm{d}r}(aT^4) = -\tfrac{4}{3}ac\ell T^3 \,\frac{\mathrm{d}T}{\mathrm{d}r} = -K\,\frac{\mathrm{d}T}{\mathrm{d}r} \tag{2.1}$$

which defines the coefficient K.

2.1 Equations of stellar structure

We begin with the equation describing how much mass we have in each spherical shell. The volume of a shell of thickness $\mathrm{d}r$ is $4\pi r^2\mathrm{d}r$, so the mass in such a shell is $\mathrm{d}M_r = 4\pi r^2\rho\mathrm{d}r$. Knowing the mass M_r enclosed within radius r, we can calculate the change in pressure as $\mathrm{d}p = \rho g\mathrm{d}r$, where $g = GM_r/r^2$ is the modulus of the radial part of gravity. The change in luminosity, $\mathrm{d}L_r$, in each shell depends on the strength of the energy sources, ϵ, at that radius. The local luminosity gives the local radiative flux F, because $L_r = 4\pi r^2 F$, and the

flux is given by the negative radial temperature gradient times the radiative conductivity, K, i.e. $F = -K\nabla T$, so

$$\frac{\mathrm{d}M_r}{\mathrm{d}r} = 4\pi r^2 \rho \tag{2.2}$$

$$\frac{\mathrm{d}p}{\mathrm{d}r} = -\frac{GM_r}{r^2}\rho \tag{2.3}$$

$$\frac{\mathrm{d}L_r}{\mathrm{d}r} = 4\pi r^2 \epsilon \tag{2.4}$$

$$\frac{\mathrm{d}T}{\mathrm{d}r} = -L_r/(4\pi r^2 K) \tag{2.5}$$

At the simplest level of approximation, these equations can be integrated outwards subject to the boundary conditions $M_r = L_r = 0$ at $r = 0$, with reasonable estimated starting values for p and T.

In the following we highlight the connection between these equations and the more general equations of compressible, self-gravitating fluid dynamics.

2.2 Connection with full set of equations

The relevant equations are the Poisson equation for the gravity potential, ϕ, so gravity is $\boldsymbol{g} = -\boldsymbol{\nabla}\phi$, the momentum equation, the entropy equation in terms of entropy, and the equation of radiation equilibrium, i.e.

$$\nabla^2\phi = 4\pi G\rho, \tag{2.6}$$

$$\rho\frac{\mathrm{D}\boldsymbol{U}}{\mathrm{D}t} = -\boldsymbol{\nabla}p + \rho\boldsymbol{g}, \tag{2.7}$$

$$\rho T\frac{\mathrm{D}s}{\mathrm{D}t} = -\boldsymbol{\nabla}\cdot\boldsymbol{F} + \epsilon, \tag{2.8}$$

$$\hat{\boldsymbol{n}}\cdot\boldsymbol{\nabla}I = -\kappa\rho\,(I - S). \tag{2.9}$$

The latter equation is the radiation transfer equation for the intensity, $I = I(\boldsymbol{x}, t, \hat{\boldsymbol{n}}, \nu)$ along all possible ray directions $\hat{\boldsymbol{n}}$ for all possible frequencies ν, and S is the *source function* which we assume here to be the Planck function $S = aT^4$, but we note that things can be more complicated if there is scattering, for example. The quantity κ is the opacity (cross-section per unit mass) and $1/(\kappa\rho) \equiv \ell$ is the mean free path that we encountered in equation (2.1).

In § 2.3 we show that this equation couples to equation (2.8) via

$$-\boldsymbol{\nabla}\cdot\boldsymbol{F} = \int_0^\infty \int_{4\pi} \kappa\rho\,(I - S)\,\mathrm{d}\Omega\,\mathrm{d}\nu. \tag{2.10}$$

We won't bother with the ν dependence and pretend that we work with ν integrated quantities (this is the so-called gray approximation).

Let us begin with the Poisson equation, noting that from $\boldsymbol{g} = -\boldsymbol{\nabla}\phi$ we have $\boldsymbol{\nabla} \cdot \boldsymbol{g} = -\nabla^2\phi$. Since in spherical symmetry $\boldsymbol{g} = -GM_r/r^2$ we have

$$-\boldsymbol{\nabla} \cdot \boldsymbol{g} = -\frac{1}{r^2}\frac{\mathrm{d}}{\mathrm{d}r}\left(r^2 g_r\right) = \frac{1}{r^2}\frac{\mathrm{d}}{\mathrm{d}r}\left(GM_r\right) = 4\pi G\rho, \tag{2.11}$$

Cancelling G on both sides and moving r^2 to the right hand side, we have

$$\frac{\mathrm{d}M_r}{\mathrm{d}r} = 4\pi r^2\rho, \tag{2.12}$$

which is just the same as equation (2.2).

In the next two equations, (2.7)–(2.8), the left hand side vanishes in the hydrostatic state, so we recover equations (2.3) and (2.4) more ore less directly when noting that $\boldsymbol{g} = -GM_r/r^2$ and that

$$\boldsymbol{\nabla} \cdot \boldsymbol{F} = \frac{1}{r^2}\frac{\mathrm{d}}{\mathrm{d}r}\left[r^2\left(\frac{L_r}{4\pi r^2}\right)\right] = \frac{1}{4\pi r^2}\frac{\mathrm{d}L_r}{\mathrm{d}r}. \tag{2.13}$$

The last of the four main equations, equation (2.9), is a bit more complicated and deserves to be discussed in a separate section.

2.3 Solving radiative transfer using moments

Equation (2.9) can be solved by taking moments. We define 0th, 1st, and 2nd moments as follows,

$$J = \frac{1}{4\pi}\int_{4\pi} I\,\mathrm{d}\Omega, \tag{2.14}$$

$$\boldsymbol{F} = c\int_{4\pi} I\hat{\boldsymbol{n}}\,\mathrm{d}\Omega, \tag{2.15}$$

$$\mathsf{P} = \frac{1}{4\pi}\int_{4\pi} I\hat{\boldsymbol{n}}\hat{\boldsymbol{n}}\,\mathrm{d}\Omega. \tag{2.16}$$

Here, P is the radiation pressure tensor, $\boldsymbol{F}$ is the indeed the same flux as the one defined in equation (2.8), and c is the speed of light[1]. Next we take the 0th and 1st moments of transfer equation, so

$$\boldsymbol{\nabla} \cdot \boldsymbol{F} = -4\pi c\kappa\rho\,(J - S), \tag{2.17}$$

$$\boldsymbol{\nabla} \cdot \mathsf{P} = -\frac{\kappa\rho}{4\pi c}\boldsymbol{F}. \tag{2.18}$$

In principle this 'game' could go on forever by defining higher and higher moments, but in order to make progress, we have to make a "closure" assumption. A common *choice* is to assume isotropy of the radiation pressure tensor[2], i.e.

$$\mathsf{P}_{ij} = \tfrac{1}{3}\delta_{ij}J. \tag{2.19}$$

[1] In stars with strong radiation, this $\boldsymbol{F}$ also enters on the rhs of the momentum equation, so $\mathrm{D}\boldsymbol{U}/\mathrm{D}t = ... + (\kappa/c)\boldsymbol{F}$. For the sun this is not important, but for O and B stars it is, and certainly for discs around black holes.

[2] You should think of 'tensor' as just being a 3×3 matrix.

With this, equation (2.18) becomes

$$\tfrac{1}{3}\boldsymbol{\nabla} J = -\frac{\kappa\rho}{4\pi c}\boldsymbol{F}. \tag{2.20}$$

Next, in order to calculate $\boldsymbol{\nabla}\cdot\boldsymbol{F}$, we divide first by $\kappa\rho$ and then take the divergence, so

$$\boldsymbol{\nabla}\cdot\left(\frac{1}{3\kappa\rho}\boldsymbol{\nabla} J\right) = -\frac{1}{4\pi c}\boldsymbol{\nabla}\cdot\boldsymbol{F}. \tag{2.21}$$

Using equation (2.17) we obtain a closed equation for J.

$$\boldsymbol{\nabla}\cdot\left(\frac{1}{3\kappa\rho}\boldsymbol{\nabla} J\right) = \kappa\rho\,(J-S). \tag{2.22}$$

This equation is in its mathematical nature similar to the Poisson equation, so we talk about a 'Poisson-like' equation for the mean intensity J. This equation is also known under the name *Eddington approximation*.

Limb darkening

If you want to learn more about this topic, it is instructive to read about Eddington's solution to a stellar atmosphere which explains, at least qualitatively, the solar limb darkening[3]. However, comparison with observations indicates that quantitatively the solar limb darkening problem is still not well understood.

Optically thick limit

Let us define the optical depth along the vertical coordinate z as

$$\tau = \int_l^\infty \kappa\rho\,\mathrm{d}z'. \tag{2.23}$$

Below a certain height in the atmosphere the medium becomes optically thick (non-transparent). This is said to be the case where $\tau \gg 1$. Looking at equation (2.22) is it clear that when $\kappa\rho$ is large, $J - S$ must be small, so

$$J \approx S = aT^4 \quad \text{(optically thick)} \tag{2.24}$$

so the term under the divergence on the left hand side of equation (2.22) (and multiplied by $-c$, so it $\boldsymbol{F}$) can be written as

$$\boldsymbol{F} = -\frac{c}{3\kappa\rho}\boldsymbol{\nabla}\left(aT^4\right) = -\frac{ac}{3\kappa\rho}\boldsymbol{\nabla}T^4 = -\frac{4acT^3}{3\kappa\rho}\boldsymbol{\nabla}T. \tag{2.25}$$

We have therefore arrived at an equation like equation (2.1), i.e. $\boldsymbol{F} = -K\boldsymbol{\nabla}T$, where $K = \frac{4acT^3}{3\kappa\rho}$ is the radiative diffusivity. Since $ac/4 = \sigma_{\mathrm{SB}}$, we can write K also as

$$K = \frac{16\sigma_{\mathrm{SB}}T^3}{3\kappa\rho} \tag{2.26}$$

[3] `http://star-www.st-and.ac.uk/~kw25/teaching/stars/GRAY.pdf`

Absorption processes in astrophysical plasmas.

- *Electron scattering*: if an electromagnetic wave passes an electron the electric field makes the electron oscillate.
- *Free-free transitions*: if during its thermal motion a free electron passes an ion, the two charged particles form a system which can absorb and emit radiation.
- *Bound-free transitions*: a neutral hydrogen atom in its ground state is ionized by a photon.
- *Bound-bound transitions*: after absorption of a photon the electron jumps to a higher bound state, rather than leaving the atom altogether.
- *Negative Hydrogen ion*: a neutral hydrogen atom is polarized by a nearby charge and can then attract and bind another electron.

[Adapted from Kippenhahn and Weigert (1990).]

2.4 Polytrope solution

In the main parts of the Sun, energy is transported by photon diffusion: the optical mean-free path is short compared with other relevant length scales (e.g. pressure scale height), so we are in the optically thick limit and can make the diffusion assumption, i.e. $\boldsymbol{F} = -K\boldsymbol{\nabla}T$. In equation (2.26) the value of κ, and therefore of K, depends on the atomic physics involved in absorbing and scattering photons. It changes slowly, so to a first approximation, $K = \text{const}$. Since there are furthermore no local energy sources, the energy flux is constant ($\boldsymbol{F} = \text{const}$), so we also have $\boldsymbol{\nabla}T = \text{const}$. In a plane-parallel atmosphere $dT/dz = \text{const}$, i.e. T increases *linearly* with depth, which leads to the polytropic atmosphere considered in § 2.5.5.

A good approximation for the opacity κ is given by Kramer's formula

$$\kappa = \kappa_0 \varrho T^{-7/2}, \tag{2.27}$$

where $\kappa_0 = 6.6 \times 10^{18}\,\mathrm{m^5\,K^{7/2}\,kg^{-2}}$. This value may well be up to 30 times larger if the gas is "metal rich", i.e. a good electron supplier, so that bound-free processes become important as well. In practice, a good value is $\kappa_0 \approx 10^{20}\,\mathrm{m^5\,K^{7/2}\,kg^{-2}}$. With this coefficient and Kramer's formula, the conductivity is

$$K = \frac{16\sigma_{\mathrm{SB}} T^{13/2}}{3\kappa_0 \varrho^2}. \tag{2.28}$$

If the density is given by a power law of the temperature, $\varrho \sim T^m$, which is called a *polytrope solution* with polytropic index m, then

$$K \sim T^{13/2-2m}, \tag{2.29}$$

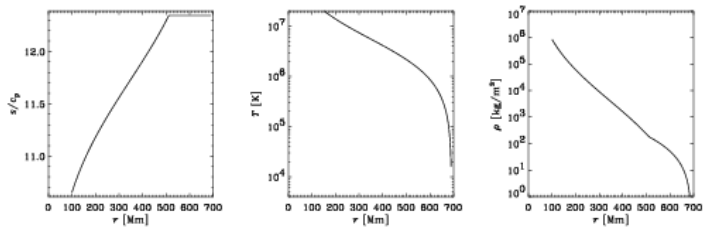


Figure 2.1: Simple polytropic stratification. The temperature gradient is assumed to be inversely proportional to the conductivity, $K(z)$, such that $K\boldsymbol{\nabla}T = F_{\rm tot} = L/(4\pi r^2)$. In the outer 30% the polytropic index is changed such that the entropy is constant. This is to model the convection zone.

which is constant for $m = 13/4 = 3.25$. This is gives a reasonable representation of the stratification of stars in convectively stable regions. At the bottom of the solar convection zone the density is about $200\,{\rm kg\,m^{-3}}$ and the temperature is about $2\times10^6\,{\rm K}$. This gives $K = (3...100)\times10^9\,{\rm kg\,m\,s^{-3}\,K^{-1}}$. In order to carry the solar flux the average temperature gradient then has to be around $0.01\,{\rm K/\,m}$.

However, in reality K does change slowly with height. Therefore the polytropic index effectively changes with height. In the outer layers the temperature decreases and there is ionization and recombination. In that layer there are many electrons allowing for the formation of negative hydrogen ions, H^-, from polarized neutral hydrogen atoms. This leads to low values of K. To transport the required energy flux, the temperature gradient has to go up. But this means the polytropic index goes down, see equation (2.66), and so the stratification will become unstable. This leads to convection in the outer parts of the Sun.

To a first approximation we can assume that convection leads to perfect mixing and therefore to a nearly uniform entropy distribution, corresponding to a state of marginal convective stability; see § 2.5.3. The star's stratification can then locally be described by a polytrope with $m = 3/2$ (for $\gamma = 5/3$). A plot of the resulting stratification is shown in Fig. 2.1. However, towards the surface layers cooling by radiation begins to play a role which causes the entropy to decrease gradually. A better approximation for the vertical stratification of density and temperature can be obtained by the mixing length theory, which will be discussed next. However, for more accurate and more detailed models one has to use numerical simulations, which are now beginning to become feasible.

2.5 Buoyancy and entropy

Gas motions in the Earth's and solar atmospheres are often driven by buoyancy. As an example we now calculate the buoyancy force from a hot air balloon. The buoyancy force results from a lower density inside the balloon (or any other

container) and the density outside it. So the buoyancy force is given by

$$F_{\text{buoy}} = -\Delta\varrho\, V g, \tag{2.30}$$

where g is the gravitational acceleration ($\approx 10\,\text{m}\,\text{s}^{-2}$), and $\Delta\varrho = \varrho_{\text{i}} - \varrho_{\text{e}}$ is the density difference between the interior and the exterior of the balloon. If the density within the balloon is smaller than outside, $\Delta\varrho < 0$ and F_{buoy} is positive. The density deficit depends on the temperature excess and, assuming that there is no pressure difference (which is justified for a hot air balloon), the two are proportional to each other, so

$$\frac{\Delta\varrho}{\varrho} = -\frac{\Delta T}{T}. \tag{2.31}$$

Let us assume that the temperature inside the balloon is $80^{\odot}$ C and the exterior temperature is 20° C. The temperature difference is then $60^{\odot}\,\text{C} = 60\,\text{K}$, and the *absolute* temperature is then is $T = 20\,\text{K} + 273\,\text{K} \approx 300\,\text{K}$ (Kelvin), so

$$\frac{\Delta\varrho}{\varrho} = -\frac{60}{300} \approx 0.2. \tag{2.32}$$

So, one cubic meter of hot air is about 20% lighter than cold air. Since the density of air is approximately $\varrho = 1\,\text{kg}\,\text{m}^{-3}$ we have $\Delta = 0.2\,\text{kg}\,\text{m}^{-3}$. The larger the balloon, the more hot air there is and the lighter is the balloon. Assuming a spherical shape with radius R the volume of the balloon is $V_{\text{balloon}} = (4\pi/3)R^3$, so the upward force of the entire balloon is

$$(-\Delta\varrho)Vg = \frac{4\pi}{3}R^3(-\Delta\varrho)g. \tag{2.33}$$

This has to be balanced against the weight of the balloon, which is mg, if m is the mass of the payload. Thus, we have

$$\frac{4\pi}{3}R^3\Delta\varrho = m. \tag{2.34}$$

If we want to know the size of the balloon necessary to carry, say, $m = 500\,\text{kg}$, we have

$$R = \left(\frac{3}{4\pi}\,\frac{m}{\Delta\varrho}\right)^{1/3} \approx \left(\frac{1}{4}\,\frac{500\,\text{kg}}{0.2\,\text{kg}\,\text{m}^{-3}}\right)^{1/3} = \sqrt[3]{625}\,\text{m} \approx 9\,\text{m}, \tag{2.35}$$

which seems quite plausible.

2.5.1 The perfect gas. Equation of state

We need an equation of state that relates the pressure p of a gas to its density ϱ and its temperature T. For a perfect gas this relation is

$$p = \frac{\mathcal{R}}{\mu}T\varrho. \tag{2.36}$$

The quantity $\mathcal{R}T/\mu$ has the dimensions of a velocity squared. As we will see later, this quantity equals the square of the sound speed in a situation where changes in the pressure and density are isothermal.[4]

The quantity

$$c_{\rm s}^{\rm (isoth)} = \left(\frac{\mathcal{R}T}{\mu}\right)^{1/2} \tag{2.37}$$

is therefore also referred to as the *isothermal* sound speed. For air the value of μ is 28.8, so

$$c_{\rm s}^{\rm (isoth)} \approx \left(\frac{8315 \times 300}{28.8}\right)^{1/2} \,{\rm m/\,s} \approx 300\,{\rm m/\,s} \tag{2.38}$$

For ionized hydrogen $\mu = 0.5$ (the atomic mass is 1 and the number of particles 2, because there are protons and electrons). However, in the Sun, as well as elsewhere in the cosmos, there is also helium and the value of μ is then around 0.6. On the other hand, the presence of neutral and molecular hydrogen increases the average value. Approximate values (to an order of magnitude) of $c_{\rm s}^{\rm (isoth)}$ and T are given in Table 2.1.

Table 2.1: Typical sound speeds in the Sun.

T	$c_{\rm s}^{\rm (isoth)}$
10^2 K	$1\,{\rm km/\,s}$
10^4 K	$10\,{\rm km/\,s}$
10^6 K	$100\,{\rm km/\,s}$

[4] When the changes are adiabatic, e.g. when thermal conduction is weak, the sound speed is slightly larger: $c_{\rm s} = \sqrt{\gamma} c_{\rm s}^{\rm (isoth)}$.

Specific heats: The energy content of a gas is measured by its specific heat, which is the energy needed to increase the temperature by one degree. This quantity can be measured by holding either the volume of the gas constant (specific heat at constant volume, c_v) or by keeping the pressure constant (specific heat at constant pressure, c_p). In general, the specific heat at constant volume is smaller than the specific heat at constant pressure, because when the pressure is constant the energy is not only used to increase the temperature, but also to increase the volume. The work associated with this is $p\,\Delta V = \mathcal{R}/\mu\Delta T$, and therefore

$$c_p - c_v = \frac{p\,\Delta V}{\Delta T} = \mathcal{R}/\mu \tag{2.39}$$

According to the kinetic theory of gases (i.e. the theory that describes the gas as noninteracting particles) the specific heat at constant volume is equal to $\mathcal{R}/(2\mu)$ times the number of degrees of freedom f of a single particle (atom or molecule), i.e. $c_v = f\mathcal{R}/(2\mu)$. Because of equation (2.39) we have $c_p = (f+2)\mathcal{R}/(2\mu)$. Therefore the ratio of the two specific heats, $\gamma \equiv c_p/c_v$, is equal to

$$\gamma = \frac{f+2}{f}. \tag{2.40}$$

For a mono-atomic gas $f = 3$, corresponding to the three directions of translation, so $\gamma = 5/3 = 1.67$. Further, for a bi-atomic (dumbbell-like) molecule there are two additional degrees of freedom corresponding to the rotation of the molecule about the axis connecting the two atoms and perpendicular to it, so $f = 3 + 2$ and $\gamma = 7/5 = 1.4$. The third rotation axis is only distinguished in molecules with more than two atoms. So, for example in CO_2 $f = 6$ and therefore $\gamma = 8/6 = 4/3 \approx 1.33$. Yet, values of γ closer to unity are possible when the molecules exhibit various kinds of oscillations that further increase the number of degrees of freedom.

2.5.2 The isothermal atmosphere

Things are changing as we rise. The exterior density decreases, decreasing therefore the buoyancy force. On the other hand, the pressure decreases, so the balloon (or gas parcel) expands and so the interior density also decreases. Which one decreases faster, depends on the temperature profile in the atmosphere. The simplest type of atmosphere is the isothermal atmosphere, i.e. one where the temperature is constant.

In any atmosphere in hydrostatic (or mechanical) equilibrium the weight (per unit area) of a thin layer of gas, $\varrho g dz$, increases the pressure by the amount $p_{\rm top} - p_{\rm bot} = -\Delta p = \varrho g dz$, so the condition of hydrostatic equilibrium is

$$\frac{dp}{dz} = -\varrho g, \tag{2.41}$$

where g is the gravitational acceleration ($\approx 10\,{\rm m/\,s^2}$ for the Earth and $\approx 300\,{\rm m/\,s^2}$ at the solar surface).

If the atmosphere is isothermal, then $p = c_{\rm s}^2 \varrho$, where $c_{\rm s}^2 = $ const. In that case we obtain

$$\frac{d}{dz}(c_{\rm s}^2 \varrho) = c_{\rm s}^2 \frac{d\varrho}{dz} = -\varrho g \tag{2.42}$$

or (after dividing by ϱ and $c_{\rm s}^2$)

$$\frac{1}{\varrho}\frac{d\varrho}{dz} = \frac{d\ln\varrho}{dz} = -g/c_{\rm s}^2 = \text{const}, \tag{2.43}$$

so

$$\ln\varrho = -gz/c_{\rm s}^2 + \ln\varrho_0, \tag{2.44}$$

where $\ln\varrho_0$ is an integration constant.

So, ϱ decreases exponentially with height, i.e.

$$\varrho = \varrho_0 e^{-z/H}, \tag{2.45}$$

where

$$H = c_{\rm s}^2/g \tag{2.46}$$

is also called the *scale height* of the atmosphere.

Furthermore, since $p = c_{\rm s}^2 \varrho$ we have

$$p = p_0 e^{-z/H}, \tag{2.47}$$

where $p_0 = c_{\rm s}^2 \varrho_0$.

2.5.3 Adiabatic changes. Entropy

If a fluid parcel preserves its heat content, i.e. if radiative losses or other heating mechanisms are unimportant on time scales of interest, pressure and density changes are said to be *adiabatic*. This is described by a quantity called the entropy which is then unchanged. For a perfect gas, we define the specific entropy (i.e. entropy per unit mass) as

$$s = c_v \ln p - c_p \ln\varrho. \tag{2.48}$$

(In principle there could be an additive constant s_0, but we can put it to zero, because only *changes* in s matter.) The entropy per unit mass is relevant, because we consider a bubble of a given mass, while its volume may change. The specific entropy, in units of c_p, is

$$s/c_p = \frac{1}{\gamma}\ln p - \ln\varrho \equiv \frac{1}{\gamma}\ln\left(p/\varrho^\gamma\right), \tag{2.49}$$

so if changes in p and ϱ are adiabatic, i.e. if $s = $ const, then

$$p = e^{\gamma s/c_p}\, \varrho^\gamma \tag{2.50}$$

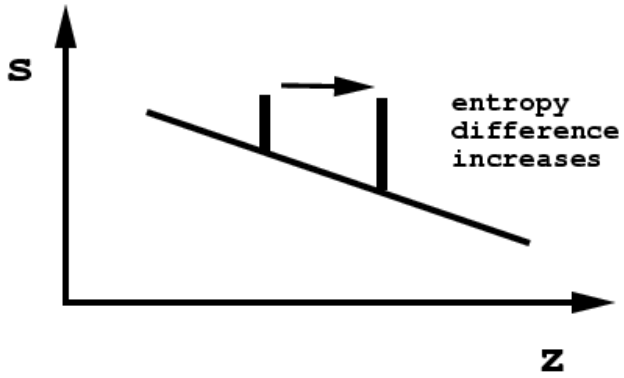


Figure 2.2: Entropy profile for an unstable atmosphere. The entropy difference between the bubble and the exterior increases constantly as the bubble ascends.

or

$$\frac{p}{p_0} = \left(\frac{\varrho}{\varrho_0}\right)^{\gamma} \tag{2.51}$$

In order to understand the evolution of a parcel in an atmosphere it is convenient to compute the vertical dependence of s. For an isothermal atmosphere the vertical gradient is

$$\frac{1}{c_p}\frac{ds}{dz} = -\frac{1}{\gamma H} + \frac{1}{H} = \frac{\gamma - 1}{\gamma}\frac{1}{H} > 0, \tag{2.52}$$

so s increases with height. This means that a rising fluid parcel, whose entropy is conserved, will end up in a location where the surrounding entropy is higher. At the same time, however, the pressure inside and outside the bubble will be the same, so $p_\mathrm{i} - p_\mathrm{e} \equiv \Delta p = 0$, where subscripts i and e refer to interior and exterior values. Thus, from equation (2.49) we have

$$\Delta s/c_p = -\Delta \ln \varrho. \tag{2.53}$$

So, since the rising fluid parcel ends up in a location of higher exterior entropy, we have $\Delta s < 0$ and therefore $\Delta \ln \varrho > 0$, so the fluid parcel becomes *heavier* and will be pulled back by the gravity. This provides a restoring force proportional to the vertical displacement, which leads to

2.5.4 Brunt-Väisälä oscillations

When the bubble rises over a distance z (from its original position, where $\Delta s = 0$), the relative change of density $\Delta\varrho/\varrho = \Delta \ln \varrho$ between interior and exterior is equal to

$$\Delta \ln \varrho = -\Delta s/c_p = +\frac{1}{c_p}\frac{ds}{dz}z, \tag{2.54}$$

so it is proportional to the displacement z. Now the buoyancy force acting on the fluid parcel per unit volume, is $-\Delta\varrho\, g$, so

$$\varrho\ddot{z} = -\Delta\varrho\, g \tag{2.55}$$

(dots indicate differentiation with respect to time), or [5]

$$\ddot{z} = -\frac{\Delta\varrho}{\varrho}\, g = -g\,\Delta\ln\varrho = g\,\frac{\Delta s}{c_p} = -\frac{g}{c_p}\frac{ds}{dz}\, z. \tag{2.56}$$

A solution of this differential equation is

$$z = z_0\cos(\omega_{\rm BV}t), \tag{2.57}$$

where z_0 is the initial displacement from the equilibrium state and $\omega_{\rm BV}$ is the Brunt-Väisälä (or buoyancy) frequency (sometimes also called $N_{\rm BV}$) with

$$\omega_{\rm BV}^2 = \frac{g}{c_p}\frac{ds}{dz}. \tag{2.58}$$

This expression only makes sense if the atmosphere is stably stratified, i.e. if $\boldsymbol{g}\cdot\boldsymbol{\nabla}s < 0$, and so we can express (2.58) in vector notation,

$$\omega_{\rm BV}^2 = -\boldsymbol{g}\cdot\boldsymbol{\nabla}s/c_p. \tag{2.59}$$

2.5.5 Polytropic atmospheres

Real atmospheres are not isothermal. A better approximation is to assume that T increases linearly with depth. *Warning*: depth increases downwards, whereas height increase upwards. To distinguish between the two we denote depth by $\tilde{z}$, with $\tilde{z} = z_{\rm max} - z$. Thus, $dz = -d\tilde{z}$, and therefore equation (2.41) becomes

$$\frac{1}{\varrho}\frac{dp}{d\tilde{z}} = g. \tag{2.60}$$

Now for a polytropic atmosphere we have

$$T/T_0 = \tilde{z}/H_T, \tag{2.61}$$

where H_T is the temperature scale height, the value of which will be determined below. The density is then assumed to be a power law of T, so

$$\varrho/\varrho_0 = (T/T_0)^m, \tag{2.62}$$

where m is the polytropic index. Because of equation (2.36) the pressure is then

$$p/p_0 = (T/T_0)^{m+1}. \tag{2.63}$$

[5] The change in sign in the last equation may be counter-intuitive, but one should keep in mind that

$$\Delta s = s_{\rm i} - s_{\rm e} = s(z{=}0) - s(z) = \frac{ds}{dz}\cdot(0-z) = -\frac{ds}{dz}z\,.$$

Plugging this into (2.60) we have

$$\varrho_0^{-1}(z/H_T)^{-m}\frac{d}{d\tilde{z}}\left[p_0(z/H_T)^{m+1}\right]=g \tag{2.64}$$

or

$$\frac{\mathcal{R}T_0}{\mu H_T}(m+1)=g. \tag{2.65}$$

The (constant) temperature gradient is then[6]

$$\beta\equiv \mathrm{d}T/\mathrm{d}\tilde{z}=T_0/H_T=\frac{g}{(m+1)\mathcal{R}/\mu}. \tag{2.66}$$

In order to see what happens to our fluid blob we determine again the entropy gradient, using equation (2.49),

$$\frac{1}{c_p}\frac{ds}{dz}=-\frac{1}{c_p}\frac{ds}{d\tilde{z}}=-\frac{m+1}{\gamma\tilde{z}}+\frac{m}{\tilde{z}}=\frac{1}{\tilde{z}}\frac{(\gamma-1)m-1}{\gamma}. \tag{2.67}$$

Here we have used the fact that $d\ln p/dz = d\ln(p/p_0)/d\tilde{z}$, because the second term in $\ln(p/p_0)=\ln p-\ln p_0$ is constant, and so

$$\frac{d\ln p}{d\tilde{z}}=\frac{d\ln(p/p_0)}{d\tilde{z}}=(m+1)\frac{d\ln(T/T_0)}{d\tilde{z}}=(m+1)\frac{d\ln\tilde{z}}{d\tilde{z}}=\frac{m+1}{\tilde{z}}. \tag{2.68}$$

Evidently, for

$$m<\frac{1}{\gamma-1} \tag{2.69}$$

the entropy gradient ds/dz turns negative, so a rising blob would find itself in an environment whose entropy is getting smaller and smaller as it rises further. This means its density relative to the exterior density is getting smaller and smaller, so the bubble becomes even more unstable. Since such bubbles may break lose all over the place the whole medium will start to bubble. This process is called convection. In this case the Brunt-Väisälä frequency becomes formally imaginary, corresponding to an exponentially growing solution with growth rate $\sigma_{\rm BV}=\mathrm{Im}(\omega_{\rm BV})$. This is because $\cos\omega_{\rm BV}t=\mathrm{Re}\left(e^{-i\omega_{\rm BV}t}\right)=e^{\sigma_{\rm BV}t}$. For $\gamma=5/3$ the criterion (2.67) for instability is $m<3/2$ (or $m>5/2$ for $\gamma=7/5$ and $m>3$ for $\gamma=4/3$).

We shall return to convective instability in connection with the solar convection zone; see § 4.2.

2.6 The main points of this chapter

The equations of stellar structure are just a subset of the equations of fluid mechanics (with some extensions). The equation of radiation transport determines

[6] In the adiabatic case, i.e. when $\mathrm{d}s/\mathrm{d}z=0$, we have $m=3/2$ (for $\gamma=5/3$) and then the temperature gradient is simply $\mathrm{d}T/\mathrm{d}\tilde{z}=g/c_p$. Remember that then $c_p=(5/2)\mathcal{R}/\mu$.

the temperature gradient. Except for the surface layers, radiation can be solved in the diffusion approximation. This means that gas and radiation have the same temperature ($J = S$). The smaller the radiative diffusivity, the steeper the temperature gradient. When it becomes too steep, such that the specific entropy gradient decreases in the outward direction, convection sets in that acts such as the minimize the entropy gradient, so the temperature gradient is then close to the adiabatic value, $\mathrm{d}T/\mathrm{d}z = -g/c_p$.

Chapter 3

Some fluid dynamics

3.1 Derivation of the energy equation

The first law of thermodynamics states that

$$T\mathrm{d}S = \mathrm{d}E + p\mathrm{d}V, \tag{3.1}$$

where T is the temperature, S the entropy, E the internal energy, p the pressure and V the volume. For later reference we mention here that TdS can also be expressed as

$$T\mathrm{d}S = \mathrm{d}H - V\mathrm{d}p, \tag{3.2}$$

where H is enthalpy. This expression can be important when rewriting the pressure gradient term in the momentum equation. In fluid dynamic we measure those quantities with respect to some unit mass. Thus, we have

$$T\mathrm{d}s = \mathrm{d}e + p\mathrm{d}v, \tag{3.3}$$

where s is the specific entropy per unit mass, e is the specific internal energy per unit mass, and v is the specific volume per unit mass (not to be confused with the velocity). For a perfect gas the temperature is related to e via $e = c_v T$, where c_v is the specific heat for constant volume. The specific volume per unit mass is just the ordinary fluid density, so $v = 1/\rho$. Since $\mathrm{d}v = \mathrm{d}(\rho^{-1}) = -\rho^{-2}\mathrm{d}\rho = -\rho^{-1}\mathrm{d}\ln\rho$ we have

$$T\mathrm{d}s = \mathrm{d}e - \frac{p}{\rho}\mathrm{d}\ln\rho. \tag{3.4}$$

In the following we shall always work with the density and never with the specific volume. Hence, from now on, v shall always refer to velocity. In fluid dynamics we are interested in the total time derivative, so we have

$$T\frac{\mathrm{D}s}{\mathrm{D}t} = \frac{\mathrm{D}e}{\mathrm{D}t} - \frac{p}{\rho}\frac{\mathrm{D}\ln\rho}{\mathrm{D}t}. \tag{3.5}$$

Now, from the continuity equation we have

$$0 = \frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot (\rho \boldsymbol{u}) = \frac{\partial \rho}{\partial t} + \boldsymbol{u} \cdot \boldsymbol{\nabla} \rho + \rho \boldsymbol{\nabla} \cdot \boldsymbol{u} = \frac{\mathrm{D}\rho}{\mathrm{D}t} + \rho \boldsymbol{\nabla} \cdot \boldsymbol{u}, \tag{3.6}$$

or, dividing by ρ,

$$\frac{\mathrm{D} \ln \rho}{\mathrm{D}t} = -\boldsymbol{\nabla} \cdot \boldsymbol{u}. \tag{3.7}$$

Using this in (3.5) yields

$$T \frac{\mathrm{D}s}{\mathrm{D}t} = \frac{\mathrm{D}e}{\mathrm{D}t} + \frac{p}{\rho} \boldsymbol{\nabla} \cdot \boldsymbol{u}. \tag{3.8}$$

In an adiabatic fluid the specific entropy is constant, i.e. $Ds/Dt = 0$. Hence, the evolution of e is given by

$$\frac{\mathrm{D}e}{\mathrm{D}t} = -\frac{p}{\rho} \boldsymbol{\nabla} \cdot \boldsymbol{u}. \tag{3.9}$$

This equation shows that if there is a compression, then $\boldsymbol{\nabla} \cdot \boldsymbol{u} < 0$ and so e increases. This effect is known as compressional heat. Conversely, if there is an expansion then $\boldsymbol{\nabla} \cdot \boldsymbol{u} > 0$ and e decreases. This effect is utilized in every fridge.

3.2 Sound waves

Shocks may occur in situations where the velocity is supersonic. This is typically when the Mach number, $\mathrm{Ma} = |\boldsymbol{u}|/c_s$, exceeds unity. Shocks can form when the amplitude of sound waves becomes comparable to the sound speed, so that the Mach number becomes of order unity. Since the sound speed is proportional to the temperature, this may occur when the gas cools, or when the wave reaches a region where the temperature is low.

In the previous chapter on winds we have seen, however, that the formation of a shock is not compulsory and that smooth transonic transitions are quite possible. Conversely, a shock may also occur under subsonic conditions if one just waits long enough for a shock wave to pile up.

In the linear regime, i.e. $\mathrm{Ma} \ll 1$, the nonlinear terms $\boldsymbol{u} \cdot \boldsymbol{\nabla}$ can be neglected and the governing equations are then, in 1 dimension,

$$\frac{\partial \ln \rho}{\partial t} = -\frac{\partial v_x}{\partial x} \tag{3.10}$$

$$\frac{\partial v_x}{\partial t} = -c_s^2 \frac{\partial \ln \rho}{\partial x} \tag{3.11}$$

Differentiating the second equation in time and substituting the first one for $\ln \rho$ yields the wave equation,

$$\frac{\partial^2 v_x}{\partial t^2} = c_s^2 \frac{\partial^2 v_x}{\partial x^2}, \tag{3.12}$$

which permits solutions of the form

$$\ln \rho = \mathrm{Ma}\, \cos k(x \mp c_s t), \tag{3.13}$$

$$v_x = \pm c_\mathrm{s} \mathrm{Ma}\, \cos k(x \mp c_s t), \tag{3.14}$$

where the upper and lower signs correspond to forward and backward propagating waves, respectively. (Here we have ignored the possibility of an arbitrary phase shift; for traveling waves this would just renormalize the zero point of time.)

3.3 Fluid equations in conservative form

In one dimension the ideal fluid equations can be written in conservative as form

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x}(\rho v_x) = 0, \tag{3.15}$$

$$\frac{\partial}{\partial t}(\rho v_x) + \frac{\partial}{\partial x}(\rho v_x^2 + p) = 0, \tag{3.16}$$

$$\frac{\partial}{\partial t}(\tfrac{1}{2}\rho v_x^2 + \rho e) + \frac{\partial}{\partial x}[v_x(\tfrac{1}{2}\rho v_x^2 + \rho e + p)] = 0. \tag{3.17}$$

where e is the internal energy density per unit mass, and the other variables have their usual meaning.

3.4 Shocks

In a frame of reference comoving with the shock the shock is stationary. Therefore $\partial/\partial t = 0$ and so the terms under the x-derivatives in the equations (3.15)-(3.17) are constant, i.e.

$$\rho v_x = \mathrm{const} = J, \tag{3.18}$$

$$\rho v_x^2 + p = \mathrm{const} = I, \tag{3.19}$$

$$(\rho v_x)\,(\tfrac{1}{2}v_x^2 + e + p/\rho) = \mathrm{const} = J\,E, \tag{3.20}$$

We assume a perfect gas with

$$p = (\gamma - 1)\rho e \tag{3.21}$$

and substitute

$$e = \frac{1}{\gamma - 1}\,\frac{p}{\rho}, \quad \text{and} \quad e + \frac{p}{\rho} = \frac{\gamma}{\gamma - 1}\,\frac{p}{\rho}. \tag{3.22}$$

This allows us rewrite the quantities that are constant across the shock in the form

$$\rho v_x = J, \tag{3.23}$$

$$\rho v_x^2 + p = I, \tag{3.24}$$

$$\tfrac{1}{2} v_x^2 + \frac{\gamma}{\gamma - 1} \frac{p}{\rho} = E. \tag{3.25}$$

We now eliminate ρ and p to derive a quadratic equation for the velocity. Dividing (3.24) by ρ gives

$$v_x^2 + \frac{p}{\rho} = \frac{I}{\rho}. \tag{3.26}$$

We now eliminate p/ρ using (3.25), so we have

$$v_x^2 + \frac{\gamma - 1}{\gamma} \left(E - \tfrac{1}{2} v_x^2 \right) = \frac{I}{\rho}. \tag{3.27}$$

Finally, we get rid of ρ using (3.23), which gives $\rho^{-1} = v_x/J$, so we have

$$v_x^2 + \frac{\gamma - 1}{\gamma} \left(E - \tfrac{1}{2} v_x^2 \right) - \frac{I}{J} v_x = 0, \tag{3.28}$$

which is a quadratic equation for v_x, which we write as

$$v_x^2 \left(1 - \frac{\gamma - 1}{2\gamma} \right) - \frac{I}{J} v_x + \frac{\gamma - 1}{\gamma} E = 0. \tag{3.29}$$

Since

$$1 - \frac{\gamma - 1}{2\gamma} = \frac{\gamma + 1}{2\gamma} \tag{3.30}$$

we have, after multiplying by $2\gamma/(\gamma + 1)$,

$$v_x^2 - \frac{2\gamma}{\gamma + 1} \frac{I}{J} v_x + 2 \frac{\gamma - 1}{\gamma + 1} E = 0. \tag{3.31}$$

Remembering that the quadratic equation can be written in the form $(v - v_1)(v - v_2) = v^2 - (v_1 + v_2) + v_1 v_2 = 0$, were v_1 and v_2 are the two solutions, we get an expression for $v_1 + v_2$,

$$v_1 + v_2 = \frac{2\gamma}{\gamma + 1} \frac{I}{J} \tag{3.32}$$

Note also that

$$\frac{I}{J v_x} = 1 + \frac{1}{\gamma \mathrm{Ma}^2}, \tag{3.33}$$

where $\mathrm{Ma} = v/c$ is the Mach number and $c^2 = \gamma p/\rho$ is the square of the adiabatic sound speed. We now assume that $\mathrm{Ma} \to \infty$ on the upstream side of the shock. This means that

$$\frac{I}{J v_1} = 1, \tag{3.34}$$

or $v_1 = I/J$. Substituting this for I/J in (3.32) we have

$$v_1 + v_2 = \frac{2\gamma}{\gamma+1} v_1, \tag{3.35}$$

or

$$v_2 = \left(\frac{2\gamma}{\gamma+1} - 1 \right) v_1 = \frac{2\gamma - (\gamma+1)}{\gamma+1} v_1 = -\frac{\gamma-1}{\gamma+1} v_1. \tag{3.36}$$

The mass ratio on the two sides of the shock is then by (3.23) given by

$$\frac{\rho_2}{\rho_1} = \frac{v_1}{v_2} = \frac{\gamma+1}{\gamma-1}. \tag{3.37}$$

For $\gamma = 5/3$ we have

$$\frac{\rho_2}{\rho_1} = \frac{5+3}{5-3} = \frac{8}{2} = 4. \tag{3.38}$$

For smaller values of γ, for example $\gamma = 7/5$, we have

$$\frac{\rho_2}{\rho_1} = \frac{7+5}{7-5} = \frac{12}{2} = 6. \tag{3.39}$$

Generalization to arbitrary shock strength

The assumption of arbitrarily strong shocks was not necessary. All we need to know is all the quantities on one side of the shock. We use equation (3.33) as it is (without assuming $\mathrm{Ma} \to \infty$) in equation (3.32) and obtain

$$v_1 + v_2 = \frac{2\gamma}{\gamma+1} \left(1 + \frac{1}{\gamma \mathrm{Ma}^2} \right) v_1, \tag{3.40}$$

so

$$\frac{v_2}{v_1} = \frac{2\gamma}{\gamma+1} \left(1 + \frac{p_1}{\rho_1 v_1^2} \right) - 1 \tag{3.41}$$

Example: suppose $v_1 = 5$, $\rho_1 = p_1 = 1$ and $\gamma = 5/3$, then

$$\frac{v_2}{v_1} = \frac{2 \times 5}{5+3} \left(1 + \frac{1}{25} \right) - 1 = \frac{5}{4} \frac{26}{25} - 1 = \frac{1}{2} \frac{13}{5} - 1 = \frac{13}{10} - 1 = 0.3, \tag{3.42}$$

so $v_2 = 5 \times 0.3 = 1.5 = 3/2$. From continuity we have $\rho_2 = \rho_1 v_1 / v_2 = 5/1.5 = 10/3 \approx 3.33$. Finally, for the pressure we

$$\rho_1 v_1^2 + p_1 = \rho_2 v_2^2 + p_2, \tag{3.43}$$

so

$$p_2 = p_1 + \rho_1 v_1^2 - \rho_2 v_2^2 = 1 + 25 - \frac{10}{3} \left(\frac{3}{2} \right)^2 = 26 - \frac{15}{2} = \frac{52-15}{2} = \frac{37}{2} = 18.5 \tag{3.44}$$

With this we can now also work out the entropies on the 2 sides of the shock. In the upstream (unshocked) part we have for the entropy in units of c_p,

$$s_1/c_p = \frac{1}{\gamma}\ln p_1 - \ln \rho_1 = 0, \tag{3.45}$$

because $p_1 = \rho_1 = 1$ and $\ln 1 = 0$. On the downstream side the shocked material has $s_2 = 0.55$.

The results are plotted in Fig. 3.1 for different values of the Mach number.

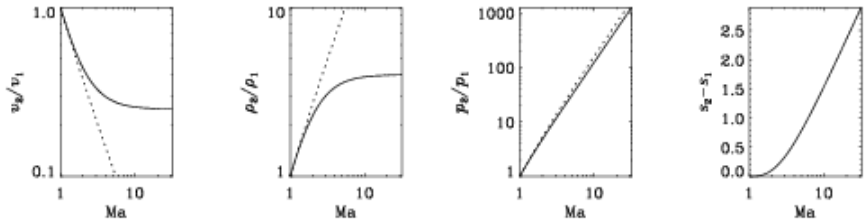


Figure 3.1: Relative velocity, density, pressure and entropy behind the shock. The dotted line refers to the polytropic case where the entropy cannot increase behind the shock.

For polytropic equations of state, with $p = K\rho^\Gamma$, the energy equation is no longer used, so there are only the following two conserved quantities,

$$J = \rho v, \quad I = \rho v^2 + K\rho^\gamma. \tag{3.46}$$

The dependence of velocity, density, pressure, and entropy jumps on the upstream Mach number is plotted in Fig. 3.1 for the case $\gamma = 5/3$ and compared with the polytropic case using $\Gamma = \gamma$. Note that in this case the density jump can be arbitrarily strong and is no longer limited by the ratio 4 found for $\gamma = 5/3$. This is however really just an artifact.

3.5 The main points of this chapter

Sound waves are solutions to the fluid dynamics equations. When their amplitude becomes large, i.e. $|u| \sim c_{\rm s}$, shocks can develop and kinetic energy is converted into heat.

Chapter 4

Helioseismology

The five-minute oscillations were discovered by Bob Leighton (1962) at Caltech from spectral line shifts (Na, Ca, Fe, etc). They were first thought to be the oscillatory response of the the atmosphere to granules pushing upwards. This idea turned later out to be quite wrong, and the oscillations are actually *global* oscillations permeating deep layers of the sun. In fact, they are just sound waves that are *trapped* in a cavity formed by reflection at the top[1] and refraction in deeper layers (see Fig. 4.1). The bending of sound waves may be illustrated by a troop of soldiers (walking of course in parallel lines) or a toy car moving at an oblique angle into troublesome terrain (Fig. 4.2). This effect is supposedly noticed by fishermen in the early mornings when there is a temperature inversion.

The decisive observation came in 1975 when Deubner[2] showed that these modes have large scale spatial coherence with wavenumbers corresponding to 20–60 Mm (1 Mm=1000 km); see Fig. 4.4. By now the field has growth to perfectionism; see Fig. 4.5.

4.1 Qualitative story

Since the beginning of the eighties, standing acoustic waves in the Sun have been used to gain information about the interior of the Sun. It was possible to measure directly (i.e. without the use of any solar model)

(i) the radial dependence of the sound speed, $c_{\rm s}(r)$ — and hence temperature. We recall that $c_{\rm s}^2 = \gamma p/\rho = \gamma \mathcal{R} T/\mu$.

(ii) the radial and latitudinal dependence of the internal angular velocity of the Sun.

[1] Sound waves can't penetrate if their wave length exceeds the scale on which density changes.

[2] Deubner, F.-L.: 1975, "Observations of low wavenumber nonradial eigenmodes of sun," *Astron. Astrophys.* **44**, 371–379

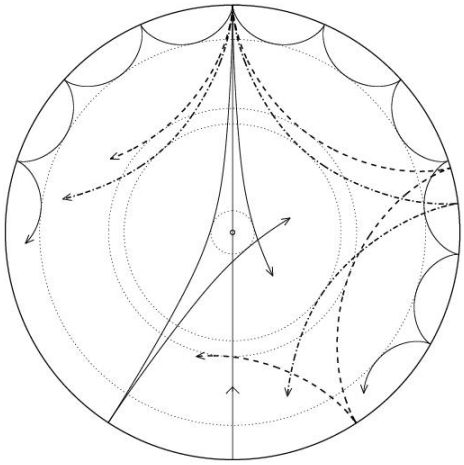


Figure 4.1: Refraction of sound waves in the sun. Courtesy J. Christensen-Dalsgaard.

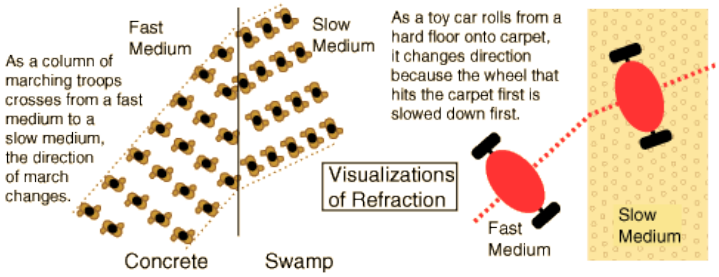


Figure 4.2: Refraction of a troop of soldiers (left) and toy cars (right).

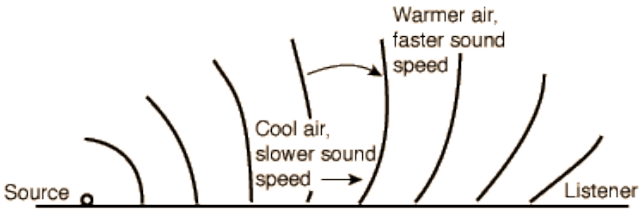


Figure 4.3: Refraction of sound waves for fisher men; see `http://hyperphysics.phy-astr.gsu.edu/hbase/sound/refrac.html`

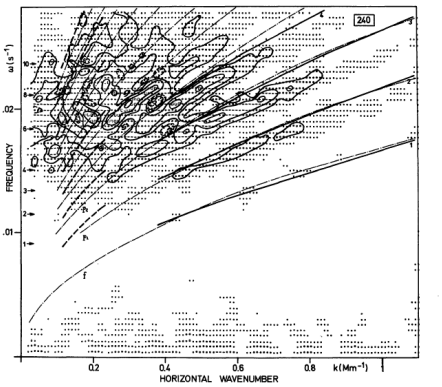


Figure 4.4: These are the original (and hence rather poor) result from Deubner's paper. However, this was the paper that first demonstrated and convinced the community that the five-minute oscillations are really global modes.

This technique is called *helioseismology*, because it is mathematically similar to the techniques used in seismology of the Earth. Qualitatively, the radial dependence of the sound speed can be measured, because standing sound waves of different horizontal wave number penetrate to different depths. Therefore, the frequencies of those different waves depend on how exactly the sound speed changes with depth. Since the Sun rotates, the waves that travel in the direction of rotation will be blue-shifted, and those that travel against the direction of rotation will be red-shifted. Therefore, the frequencies are split, depending on the amount of rotation in different layers. There are many reviews on the subject.[3]

Figure 4.6 shows the spatial pattern of a standing wave in three dimensions. It is the frequencies belonging to different latitudinal wave patterns that allow us to determine the latitudinal dependence of the angular velocity as well.

These acoustic waves are possible, because they are constantly being excited by the "noise" generated in the convection zone. The random fluctuations in the convection are turbulent and contain noise at all frequencies, similar to the noise generated by a glider going through the air. Now the Sun is a harmonic oscillator for sound waves and the different sound modes can be excited stochastically. This is similar to a bell in a sand storm starting to ring.

Figure 4.7 gives the result an inversion procedure that computes the radial dependence of the sound speed on depth, using the different frequency modes as input.

Helioseismology is now a big "industry", and more accurate data have now emerged due to the SOHO satellite and the GONG project (GONG = Global

[3] P. Demarque and D. B. Guenther: 1999, "Helioseismology: Probing the interior of a star," *Proc. Natl. Acad. Sci.***96**, 5356–5359

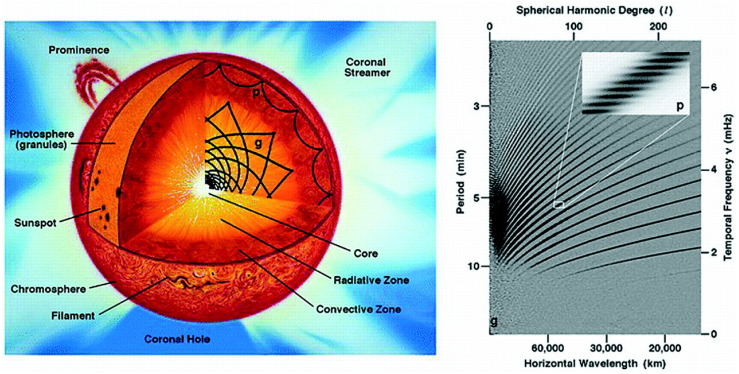


Figure 4.5: Ray paths in the sun (in addition to pressure or p-modes, also gravity or g-modes (not yet detected!) are sketched near the center. Courtesy J. Leibacher; see`http://www.pnas.org/cgi/content/full/96/10/5356`

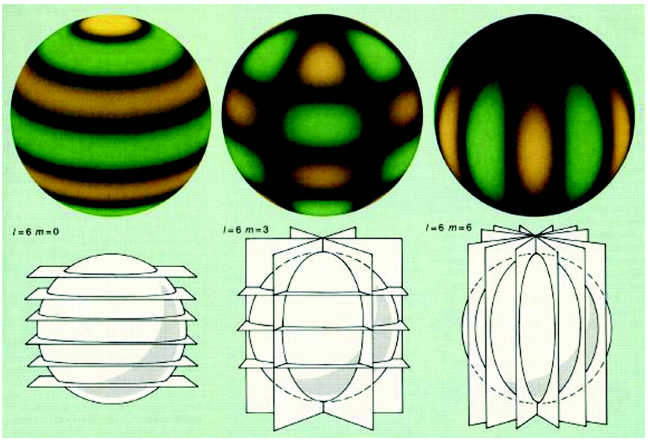


Figure 4.6: Three-dimensional wave pattern of a single wave mode. In reality, millions of different wave patterns are all superimposed. Courtesy J. Leibacher.

Oscillation Network Group), which has six stations around the globe to eliminate nightly gaps in the data.

4.1.1 Inverting the frequency spectrum

Like with a violin string, the acoustic frequency of the wave increases as the wavelength decreases, i.e.

$$\text{frequency} \propto 1/\text{wavelength}. \tag{4.1}$$

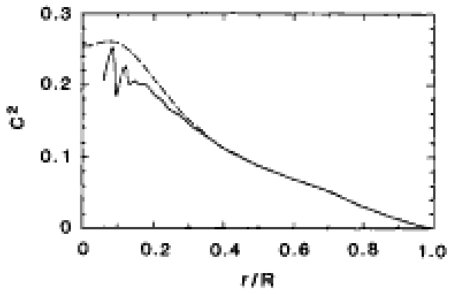


Figure 4.7: Radial dependence of the sound speed on radius in the Sun. Note the change in slope near $r = 0.7$ solar radii. The oscillations near the center are not physical. The theoretical model (dotted line) is in fair agreement with the direct measurements. The sound speed has its maximum not in the center, because the mean molecular weight μ increases towards the center, which causes $c_{\rm s}$ to decrease. [Again, we recall that $c_{\rm s}^2(r) = \gamma \mathcal{R} T/\mu$.]

More precisely, the frequency ω is given by

$$\omega = c_{\rm s} k, \tag{4.2}$$

where $c_{\rm s}$ is the sound speed and $k = 2\pi/\lambda$ is the wavenumber (λ is the wavelength). If sound waves travel an oblique path then we can express the wavenumber in terms of its horizontal and vertical wave numbers, $k_{\rm h}$ and $k_{\rm v}$, respectively. We do this because only the horizontal wavenumber can be observed. This corresponds to the horizontal pattern in Fig. 4.6. Thus, we have

$$k^2 = k_{\rm h}^2 + k_{\rm v}^2 \; . \tag{4.3}$$

The number of radial nodes of the wave is given by the number of waves that fit into the Sun, or at least the part of the Sun where the corresponding wave can travel. This part of the Sun will be referred to as cavity. The larger the cavity is, the more nodes there will be for a given wavelength. The number of modes n is then given by

$$n = 2\Delta r/\lambda = 2\Delta r \frac{k_{\rm v}}{2\pi} = \Delta r \, k_{\rm v}/\pi \; , \tag{4.4}$$

where Δr is the radial extent of the cavity. If the sound velocity and, hence, $k_{\rm v}$ depends on radius, this formula must be generalized to

$$n = \frac{1}{\pi} \int\limits_{r_{\rm min}}^{R_\odot} k_{\rm v} \, dr, \tag{4.5}$$

supposing the cavity to be the spherical shell $r_{\rm min} < r < R_\odot$.

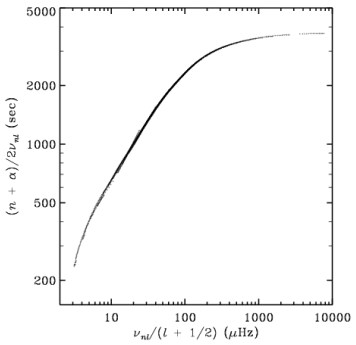


Figure 4.8: Duvall law. The vertical axis (ordinate) is the same as F in equation (4.12) and the horizontal axis (abscissa) is basically the same as u^{-1}. (He found this law well before it's significance was understood in terms one of the functions in Abel's integral transformation. Courtesy J. Christensen-Dalsgaard.

The horizontal pattern of the proper oscillation is described by spherical harmonics with indices l and m, hence the horizontal wave number is

$$k_{\rm h}^2 = \frac{\ell(\ell+1)}{r^2} \tag{4.6}$$

and we can write

$$k_{\rm v} = \sqrt{\frac{\omega^2}{c_{\rm s}^2} - \frac{\ell(\ell+1)}{r^2}} = \frac{\omega}{r}\sqrt{\frac{r^2}{c_{\rm s}^2} - \frac{\ell(\ell+1)}{\omega^2}} \tag{4.7}$$

Therefore, the number n of radial nodes is given by

$$\frac{\pi(n+\alpha)}{\omega} = \int_{r_{\rm min}}^{R_\odot} \sqrt{\frac{r^2}{c_{\rm s}^2} - \frac{\ell(\ell+1)}{\omega^2}}\,\frac{dr}{r}. \tag{4.8}$$

The phase shift $\alpha \approx 1.5$ accounts for the fact that the standing waves are confined by "soft" and extended, rather than fixed, boundaries.

The location of the inner radius is given by the point where the wave vector has turned horizontal. Using

$$\omega^2/c_{\rm s}^2 = k^2 = k_{\rm h}^2 + k_{\rm v}^2 \tag{4.9}$$

together with $k_{\rm v} = 0$ at $r = r_{\rm min}$ and $k_{\rm h}^2 = \ell(\ell+1)/r^2$, we have $(r_{\rm min}/c_{\rm s})^2 = \ell(\ell+1)/\omega^2$. This implies that

$$r_{\rm min} = \frac{c_{\rm s}}{\omega}\sqrt{\ell(\ell+1)}, \tag{4.10}$$

so only modes with low-ℓ values can be used to examine the sun's core. We now introduce new variables

$$\xi \equiv \frac{r^2}{c_{\rm s}^2}, \qquad u \equiv \frac{\ell(\ell+1)}{\omega^2}, \tag{4.11}$$

so the inner turning point of the modes corresponds to $\xi = u$. Furthermore, we denote the left hand side of equation (4.8) by $F(u)$, so we can write

$$F(u) = \int_u^{\xi_\odot} \sqrt{\xi - u}\, \frac{d\ln r}{d\xi}\, d\xi. \tag{4.12}$$

where the location of the inner refraction point corresponds to $u = \xi$. The function $F(u)$ was obtained from observations by Tom Duvall[4] on the grounds that this curious combination of data makes the different branches collapse onyo one (see Fig. 4.8). He discovered this well before its significance was understood by Douglas Gough[5] a few years later.

Since we know, at least in principle, $F(u)$ from observations and are interested in the connection between r and ξ (i. e. r and $c_{\rm s}$), we interpret (4.12) as an integral equation for the unknown function $r(\xi)$. Most integral equations cannot be solved in closed form, but this one can. It was Douglas Gough who realized that it can be cast in the form of Abel's integral equation[6]. Here is the pair of complementary equations (primes denote derivatives):

$$F(u) = \int_u^{\xi_\odot} \sqrt{\xi - u}\, G'(\xi)\, d\xi, \tag{4.13}$$

$$G(\xi) = \frac{2}{\pi} \int_\xi^{\xi_\odot} \frac{1}{\sqrt{\xi - u}}\, F'(u)\, du. \tag{4.14}$$

In the following we give an explicit derivation. Differentiate (4.12) with respect to u and, in the final result, rename ξ to ξ':

$$\frac{dF}{du} \equiv F'(u) = -\frac{1}{2} \int_u^{\xi_\odot} \frac{1}{\sqrt{\xi - u}} \frac{d\ln r}{d\xi}\, d\xi - 1 \cdot \sqrt{u - u}\, \frac{d\ln r}{d\xi}\bigg|_{\xi=u}, \tag{4.15}$$

i. e.

$$F'(u) = -\frac{1}{2} \int_u^{\xi_\odot} \frac{1}{\sqrt{\xi' - u}} \frac{d\ln r}{d\xi'}\, d\xi'. \tag{4.16}$$

[4]Duvall, T. L., Jr.: 1982, "A dispersion law for solar oscillations," *Nature* **300**, 242–243
[5]D. Gough: 1985, "Inverting helioseismic data," *Solar Phys.* **100**, 65–99
[6]Niels Abel (1802–1829); see `http://www.shu.edu/projects/reals/history/abel.html`

This can be solved analytically. Multiply by $1/\sqrt{u-\xi}$ and integrate from $u=\xi$ to $u=\xi_\odot$.

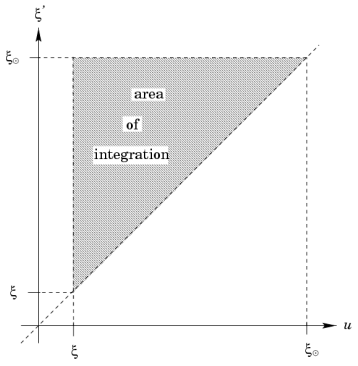
$$\begin{aligned}\int_\xi^{\xi_\odot} \frac{F'(u)}{\sqrt{u-\xi}}\,du &= -\frac{1}{2}\int_\xi^{\xi_\odot} du\,\frac{1}{\sqrt{u-\xi}}\int_u^{\xi_\odot} d\xi'\,\frac{1}{\sqrt{\xi'-u}}\,\frac{d\ln r}{d\xi'} \\ &= -\frac{1}{2}\int_\xi^{\xi_\odot} du\int_u^{\xi_\odot} d\xi'\,\frac{1}{\sqrt{(u-\xi)(\xi'-u)}}\,\frac{d\ln r}{d\xi'} \qquad (4.17)\end{aligned}$$

Due to the unknown function $\ln r(\xi')$, we cannot explicitly carry out the integration with respect to ξ', but integrating over u will be possible. In order to do this, we interchange the order of integration.

Interchanging the order of integration: This would be easy if integration were over a rectangular region. In our case, where the integration bounds for ξ' depend on u, the graphical representation on the right shows us what to do.
We can *either* let u run from ξ to $\xi_\odot$ (the outer integral) and for every given u integrate over ξ' from $\xi'=u$ to $\xi'=\xi_\odot$ (the inner integral).
Or we can choose the outer integration to be over ξ' from $\xi'=\xi$ to $\xi'=\xi_\odot$ and for every given value of ξ' integrate over u from $u=\xi$ to $u=\xi'$.
Hence,

$$\int_\xi^{\xi_\odot} du\int_u^{\xi_\odot} d\xi'(\ldots) = \int_\xi^{\xi_\odot} d\xi'\int_\xi^{\xi'} du(\ldots), \qquad (4.18)$$

where $(\ldots)$ stands for an arbitrary integrand.



With this, (4.17) becomes

$$\int_\xi^{\xi_\odot} \frac{F'(u)}{\sqrt{u-\xi}}\,du = -\frac{1}{2}\int_\xi^{\xi_\odot} d\xi'\,\frac{d\ln r}{d\xi'}\int_\xi^{\xi'} \frac{du}{\sqrt{(u-\xi)(\xi'-u)}} \qquad (4.19)$$

The integral over u can now be evaluated analytically. One way to do this involves a trigonometric substitution of variables. Alternatively, we can have a close look at the integrand, which is sketched in Fig. 4.9.

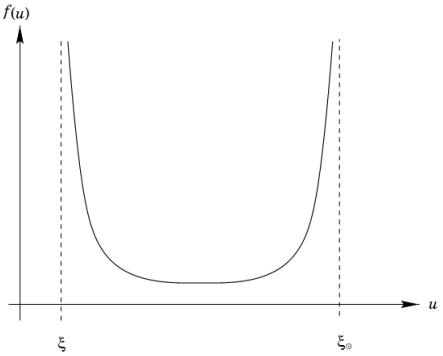


Figure 4.9: Sketch of the integrand $f(u) \equiv \dfrac{1}{\sqrt{(u-\xi)(\xi'-u)}}$ as a function of u.

The integrand $f(u)$ is symmetric with respect to $(\xi'+\xi)/2$, which motivates introduction of a new integration variable v in the following way:

$$u = \frac{\xi'+\xi}{2} + \frac{\xi'-\xi}{2} v \tag{4.20}$$

with

$$du = \frac{\xi'-\xi}{2}\, dv, \qquad u-\xi = \frac{\xi'-\xi}{2}\,(1+v), \qquad \xi'-u = \frac{\xi'-\xi}{2}\,(1-v), \tag{4.21}$$

Obviously, $u=\xi$ corresponds to $v=-1$, while $u=\xi'$ for $v=1$. Thus,

$$\begin{aligned} \int_{\xi}^{\xi'} \frac{du}{\sqrt{(u-\xi)(\xi'-u)}} &= \int_{-1}^{1} \frac{\frac{\xi'-\xi}{2}}{\sqrt{(\frac{\xi'-\xi}{2})^2(1-v^2)}}\, dv \\ &= \int_{-1}^{1} \frac{dv}{\sqrt{1-v^2}} \\ &= \arcsin v \Big|_{-1}^{1} \\ &= \frac{\pi}{2} - (-\frac{\pi}{2}) = \pi \end{aligned} \tag{4.22}$$

Inserting this into (4.17), we obtain

$$\int_{\xi}^{\xi_\odot} \frac{F'(u)}{\sqrt{u-\xi}}\, du = -\frac{\pi}{2} \int_{\xi}^{\xi_\odot} \frac{d\ln r}{d\xi'}\, d\xi' = -\frac{\pi}{2}\, \ln r \Big|_{\xi'=\xi}^{\xi_\odot} = \frac{\pi}{2} \ln \frac{r(\xi)}{R_\odot} \tag{4.23}$$

We can solve this equation for $r(\xi)$:

$$r(\xi) = R_\odot \exp\left(\frac{2}{\pi}\int_\xi^{\xi_\odot} \frac{F'(u)}{\sqrt{u-\xi}}\,du\right). \tag{4.24}$$

This is the final result of inverting the integral equation (4.12). It establishes the link we were looking for between the observable function $F(u)$ and the function $r(\xi)$, from which the radial profile of the sound velocity c_s can directly be obtained.

It should be noted, however, that this approach cannot be used in practice when the input data are noisy. One therefore usually adopts a minimization procedure where the resulting function is by construction smooth. This procedure falls under the general name of *inverse theory* and is used a lot in various branches of astrophysics. (Another interesting example is to so-called surface imaging of stars using only spectral line profile measurements of rotating stars.)

4.2 Current problems and issues

Opacities depend on metallicity abundance. The solar models calculated with the old tables agreed quite well ($Z/X = 0.017$). However, the metallicities were based on fits of observed spectra to synthetic line spectra calculated from model atmospheres. These models parameterize the three-dimensional convection rather crudely. New synthetic line spectra calculated from three-dimensional time-dependent hydrodynamical models of the solar atmosphere give a lower value of the solar oxygen abundance (Allende Prieto et al 2001, Asplund et al. 2004).[7] With the new values ($Z/X = 0.013$) it became difficult to reconcile the previously good agreement; see Fig. 4.10.

Another very important problem is to calculate the internal angular velocity of the sun (Fig. 4.11). This has already been possible for the past 20 years, but the accuracy has been ever improving. Several important features have emerged.

- The contours of constant angular velocity are not aligned with the axis of rotation, as would would have expected, and as many models still show (Fig. 4.12).

- The angular velocity in the radiative interior in nearly constant, so there is no rapidly rotating core, as one might have expected, because the sun started off as a rapid rotator, and it must have subsequently spun down. The fact that also the core has spun down means that there must be some

[7] Asplund, M., Grevesse, N., Sauval, A. J., Allende Prieto, C., & Kiselman, D.: 2004, "Line formation in solar granulation. IV. [O I], O I and OH lines and the photospheric O abundance," *Astron. Astrophys.* **417**, 751–768
Allende Prieto, C., Lambert, D. L., & Asplund, M.: 2001, "The forbidden abundance of oxygen in the Sun," *Astrophys. J. Lett.* **556**, L63–L66

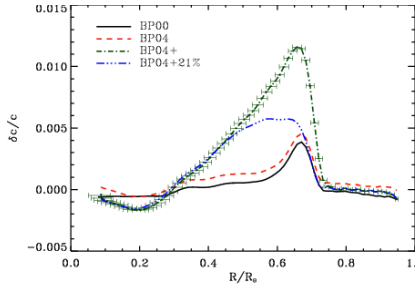


Figure 4.10: Relative sound-speed differences, $\delta c/c = (c_\odot - c_{\rm model})/c_{\rm model}$, between solar models and helioseismological results from MDI data. The vertical error bars show the 1σ error in the inversion due to statistical errors in the data. The horizontal error bars are a measure of the resolution of the inversions, defined as the distance between the first and third quartile points of the averaging kernels (approximately the half-width in radius of the measurement in regions of good resolution). Courtesy J. Bahcall et al. (2004).

efficient torques accomplishing the angular momentum transport. The most likely candidate is magnetic field: only a weak field suffices.

- The transition layer at the bottom of the convection zone, where the latitudinal differential rotation goes over into rigid rotation, is called *tachocline*. Below 30° the radial angular velocity gradient is here negative.

- Near the top layers (outer 5%) the angular velocity gradient is negative and quite sharp.

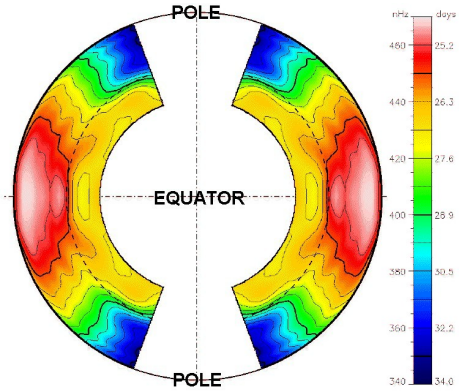


Figure 4.11: Internal angular velocity of the sun.

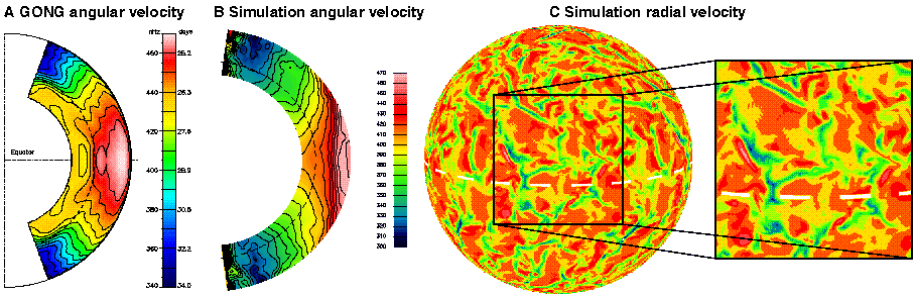


Figure 4.12: Models by Mark Miesch (High Altitude Observatory, Boulder, Colorado) and collaborators (right), compared with the helioseismologically inferred internal angular velocity of the sun (left). `http://lcd-www.colorado.edu/SPTP/sptp_global.html`

4.3 The main points of this chapter

Helioseismology is an advanced technique that allows the internal structure and distribution of angular velocity of the sun to be determined by measuring a large number of discrete acoustic frequencies ω_{nlm}. This technique is independent of any prior model assumption about the sun[8]

[8]This was important in the 1980ies when the neutrino emission from the sun was thought to be 1/3 too small, suggesting that perhaps the theoretical models were wrong. However, the models were right; the explanation were the neutrino oscillations that were not yet discovered back then.

Chapter 5

Atmospheric waves

Before we come to the general case, we first consider one-dimensional sound waves without entropy perturbations. This means that we ignore Brunt-Väisälä oscillations. The main thing that we shall learn here is that it is important to bring the equations into a form such that it has constant coefficients.

5.1 p- and g-modes

In addition to sound waves (also called p-modes, indicating that the restoring force is the pressure gradient) there are also gravity waves (also called g-modes). Gravity waves are not to be confused with gravitational waves that emerge from the linearized Einstein equations. Gravitational waves have not yet been detected, but it believed that they can be excited in connection with the collapse of a neutron star binary. Gravity waves, on the other hand, have also not yet been seen in the sun. This is because they can only propagate in the radiative interior, and their remains would be strongly throughout the convection zone.

Gravity waves are however observed in the earth atmosphere. Some impressive examples[1]; see are shown in Fig. 5.1.

5.2 Sound waves in a stratified atmosphere

In the presence of gravity there is density stratification which affects sounds waves if their wave length becomes comparable to the density scale height. The one-dimensional continuity and momentum equations for an isothermal atmosphere with constant speed of sound, c_s, no entropy perturbations, and uniform gravity, g, can be written in the form

$$\rho\frac{\partial v_z}{\partial t}+\rho v_z\frac{\partial v_z}{\partial z}+c_s^2\frac{\partial \rho}{\partial z}+\rho g=0, \tag{5.1}$$

[1]An excellent web site with pretty picture of g-modes is given in `http://taylor.math.ualberta.ca/~bruce/imagelinks/earth.html`

Figure 5.1: Gravity waves, as manifested in the could patter. These waves are often excited in the wave of wind flow over a mountain ridge. Adapted from `http://taylor.math.ualberta.ca/~bruce/imagelinks/earth.html`

$$\frac{\partial \rho}{\partial t} + v_z \frac{\partial \rho}{\partial z} + \rho \frac{\partial v_z}{\partial z} = 0, \tag{5.2}$$

where ρ is density and v_z vertical velocity. With the exception of the ρg term in equation (5.1), these equations are same as those used in § 3.2 for ordinary sound waves. The hydrostatic equilibrium solution of these equations is $v_z = \overline{v_z} = 0$ and $\rho = \overline{\rho}(z) = \rho_0\, e^{-z/H}$, where ρ_0 is a constant and $H = c_s^2/g$ is the vertical scale height. By writing $\rho = \overline{\rho} + \rho'$ and $v_z = v_z'$ one can linearize equations (5.1) and (5.2) with respect to $v_z = v_z'$ and ρ' about the solution $v_z = 0$ and $\rho = \overline{\rho}(z)$. This yields

$$\overline{\rho} \frac{\partial v_z'}{\partial t} + c_s^2 \frac{\partial \rho'}{\partial z} + \rho' g = 0, \tag{5.3}$$

$$\frac{\partial \rho'}{\partial t} + v_z' \frac{\mathrm{d}\overline{\rho}}{\mathrm{d}z} + \overline{\rho} \frac{\partial v_z'}{\partial z} = 0, \tag{5.4}$$

We may assume that $v_z = v_z'$ and ρ' take the form

$$v_z'(z,t) = \hat{v}_z e^{ikz - i\omega t + z/2H}. \tag{5.5}$$

$$\rho'(z,t) = \hat{\rho} e^{ikz - i\omega t - z/2H}, \tag{5.6}$$

The linearized equations can then be written as

$$\begin{pmatrix} -i\omega & [ik + (2H)^{-1}]c_s^2 \\ [ik - (2H)^{-1}] & -i\omega \end{pmatrix} \begin{pmatrix} \rho_0 \hat{v}_z \\ \hat{\rho} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{5.7}$$

The dispersion relation is then given by

$$\omega^2 = \omega_0^2 + k^2 c_s^2, \tag{5.8}$$

where we have introduced the acoustic cutoff frequency $\omega_0 = c_s/2H$. Figure 5.2 shows a plot of the dispersion relation.

As an example, we calculate the value of the period $T = 2\pi/\omega_0$ for the solar atmosphere, assuming $c_s = 6\,\mathrm{km/s}$ and $g = 270\,\mathrm{ms}^2$. First we have $H =$

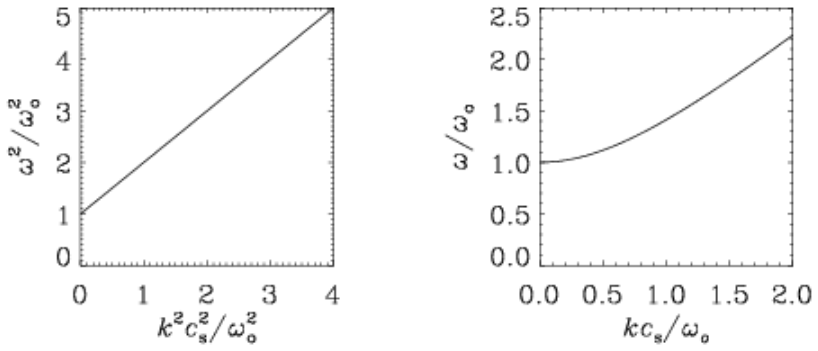


Figure 5.2: Dispersion relation for sound waves in a stratified atmosphere. On the left is shown $\omega^2(k^2)$ and on the right just $\omega(k)$.

$c_s^2/g = (6\times10^3)/270\,\mathrm{m} = 1.3\times10^5\,\mathrm{m} = 130\,\mathrm{km}$. Thus, $\omega_0 = c_s/2H = g/2c_s = 0.023\,\mathrm{s}^{-1}$, so $T = 2\pi/0.023\,\mathrm{s} = 280\,\mathrm{s} \approx 5\,\mathrm{min}$. This value is in fact close to the period of the 5 minutes oscillations of the sun.

5.3 What could have gone wrong

The linearized equations could have also been written in the form

$$\frac{\partial}{\partial t}\left(\frac{\rho'}{\overline{\rho}}\right) + v_z'\frac{\mathrm{d}\ln\overline{\rho}}{\mathrm{d}z} + \frac{\partial v_z'}{\partial z} = 0, \tag{5.9}$$

$$\frac{\partial v_z'}{\partial t} + c_s^2\frac{\partial}{\partial z}\left(\frac{\rho'}{\overline{\rho}}\right) = 0, \tag{5.10}$$

Here we have made use of the fact that

$$\frac{\partial}{\partial z}\left(\frac{\rho'}{\overline{\rho}}\right) = -\frac{\rho'}{\overline{\rho}^2}\frac{\mathrm{d}\overline{\rho}}{\mathrm{d}z} + \frac{1}{\overline{\rho}}\frac{\partial\rho'}{\partial z} = +\frac{1}{H}\frac{\rho'}{\overline{\rho}} + \frac{1}{\overline{\rho}}\frac{\partial\rho'}{\partial z}, \tag{5.11}$$

Equations (5.9) and (5.10) have constant coefficients, so we can make the ansatz

$$\frac{\rho}{\overline{\rho}} = \hat{\vartheta}e^{ikz-i\omega t}, \tag{5.12}$$

$$v_z'(z,t) = \hat{v}_z e^{ikz-i\omega t}. \tag{5.13}$$

which leads to the matrix equation

$$\begin{pmatrix} -i\omega & ikc_s^2 \\ ik - H^{-1} & -i\omega \end{pmatrix}\begin{pmatrix} \hat{v}_z \\ \hat{\vartheta} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{5.14}$$

For nontrivial solutions the determinant of the matrix has to vanish, which leads to the dispersion relation

$$\omega^2 = c_s^2 k^2 + ikc_s^2/H. \tag{5.15}$$

This result does not seem compatible with the previous one. In particular, instead of having stable waves, there now seems to be an instability: the frequency has an imaginary part with the growth rate $+kc_s^2/H$.

The reason for this apparent puzzle is that the eigenfunction is a wave whose amplitude increases with height. At any given point, however, the amplitude does not change. This is a classical case of what is called in Landau-Lifshitz *convective instability*[2] (this does have nothing to do with convection, however!).

In this case we can make ω real by allowing k to be complex. To see this we substitute

$$k = k' + \mathrm{i}k'', \tag{5.16}$$

where k' and k'' refer to real and imaginary parts, respectively. Note that $k^2 = k'^2 + 2\mathrm{i}k'k'' - k''^2$, so the dispersion relation becomes

$$\omega^2 = c_s^2(k'^2 - k''^2 - k''/H) + \mathrm{i}k'c_s^2(2k'' + 1/H). \tag{5.17}$$

We now require that the imaginary part vanishes, i.e. $k'' = -1/(2H)$, and obtain

$$\omega^2 = c_s^2\left(k'^2 + \frac{1}{4H^2}\right), \tag{5.18}$$

which is identical to equation (5.8).

5.4 The main points of this chapter

In the sun, because the density is not uniform, sound waves become non-propagating when the wavelength becomes comparable with the pressure scale height of the atmosphere. Effectively, this makes the sound waves reflect near the top.

[2] Physical Kinetics, Section VI

Chapter 6

The solar wind

Astrophysical flows are usually highly compressible. There are some specific effects resulting from that, one of which concerns the formation of winds and other types of outflows. The general principle of how a flow can attain supersonic speeds in a smooth manner (without going through a shock) can best be explained with the example of the Laval nozzle that is used in all rocket motors where the exhaust velocity can be 2000...3000 m/ s, which is up to ten times the speed of sound. The wind problem is actually analogous to the problem Bondi accretion, which is spherical accretion onto a central object. The wind solution was first proposed by Gene Parker[1] in the context of the sun.

We begin this section by first discussing the general issue of how a flow can be made supersonic. We then apply this to winds and outflows from stars and discs and mention other circumstances where this formalism can be used. A rather important application is jets, i.e. outflows from accretion discs that become highly collimated. The collimation is probably due to the magnetic field, but this is very much an open research topic at the moment. The highly relativistic version of the wind problem is relevant for understanding outflows from active galactic nuclei (AGNs) and the gamma-ray burst (GRB) phenomenon.

6.1 The analogy with the Laval nozzle

We are used to think that the flow through a pipe is becoming faster when the cross-section becomes smaller (garden hose experiment!). However, this is only true when the fluid is incompressible. That's often not a good approximation in astrophysics. The following problem illustrates this.

In a steady state the mass flux through a pipe is conserved, i.e.

$$\rho v_x S = \text{const.} \tag{6.1}$$

[1] Parker, E. N.: 1966, "The dynamical state of the interstellar gas and field," *Astrophys. J.* **145**, 811–833

Here, $S = S(x)$ is the cross-section, $v_x(x)$ the streamwise velocity and $\rho(x)$ the density. This equation follows from integrating the equation $\boldsymbol{\nabla} \cdot (\rho \boldsymbol{u}) = 0$ over the surface of a pipe that becomes narrower, and by applying Gauss' divergence theorem; see Fig. 6.1.

Indeed, if ρ is constant, (6.1) shows that a larger cross-sectional area S implies a smaller speed v_x. But what if ρ actually decreases? In the following we will see an example where supersonic speeds can be achieved by decreasing density and increasing the cross-sectional area in a suitable manner.

The velocity is obtained by solving the steady, one-dimensional, isothermal momentum equation without any extra forces,

$$v_x \frac{\mathrm{d}v_x}{\mathrm{d}x} = -c_s^2 \frac{\mathrm{d}\ln\rho}{\mathrm{d}x}. \tag{6.2}$$

We now use equation (6.1), differentiate logarithmically, i.e.

$$0 = \frac{1}{\rho v_x S}\frac{\mathrm{d}v_x}{\mathrm{d}x}(\rho v_x S) = \frac{d}{dx}\ln(\rho v_x S) = \frac{\mathrm{d}\ln\rho}{\mathrm{d}x} + \frac{\mathrm{d}\ln|v_x|}{\mathrm{d}x} + \frac{\mathrm{d}\ln S}{\mathrm{d}x}, \tag{6.3}$$

so we have

$$\frac{\mathrm{d}\ln\rho}{\mathrm{d}x} = -\frac{\mathrm{d}\ln|v_x|}{\mathrm{d}x} - \frac{\mathrm{d}\ln S}{\mathrm{d}x}. \tag{6.4}$$

Using that in equation (6.2) yields

$$v_x \frac{\mathrm{d}v_x}{\mathrm{d}x} = -c_s^2 \left(-\frac{\mathrm{d}\ln|v_x|}{\mathrm{d}x} - \frac{\mathrm{d}\ln S}{\mathrm{d}x} \right). \tag{6.5}$$

We move the first term on the rhs to the left and on the left we write

$$v_x \frac{\mathrm{d}v_x}{\mathrm{d}x} = v_x^2 \frac{\mathrm{d}\ln|v_x|}{\mathrm{d}x}. \tag{6.6}$$

This yields the equation

$$(v_x^2 - c_s^2)\frac{\mathrm{d}\ln|v_x|}{\mathrm{d}x} = c_s^2 \frac{\mathrm{d}\ln S}{\mathrm{d}x}, \tag{6.7}$$

where $S(x)$ is the cross-sectional area of the nozzle, which is a known function of x. To obtain $v_x(x)$ we can integrate

$$\frac{\mathrm{d}\ln|v_x|}{\mathrm{d}x} = c_s^2 \frac{d\ln S/\mathrm{d}x}{v_x^2 - c_s^2}, \tag{6.8}$$

but in order for the solution to be regular when $|v_x| = c_s$ we have to require that the numerator vanishes at the same point. Thus, the critical point is where $S(x)$ has a minimum, because then $d\ln S/dx = 0$, which must be where $v_x = c_s$. Define $\rho v_x S = \dot{M}$ and integrate to obtain an implicit equation for $v(x)$:

$$\tfrac{1}{2}v_x^2 - c_s^2 \ln|v_x| - c_s^2 \ln S = \mathcal{E} - c_s^2 \ln \dot{M} \equiv \mathcal{E}'. \tag{6.9}$$

This is just the usual Bernoulli equation that we have encountered elsewhere. We can now determine $\mathcal{E}'_{\rm crit}$ by applying the known values of r and v_r at the critical. For $c_s = \dot{M} = 1$ we find $\mathcal{E}'_{\rm crit} = \frac{1}{2}c_s^2 - c_s^2 \ln S_{\rm min}$.

6.2 The isothermal wind problem

We now consider the steady, isothermal wind problem. We adopt spherical polar coordinates, (r, θ, ϕ), but assume spherical symmetry in this case, so $\partial/\partial\theta = \partial/\partial\phi = 0$. The continuity and Euler equations are then

$$\frac{1}{r^2}\frac{\mathrm{d}}{\mathrm{d}r}(r^2 \rho v_r) = 0, \tag{6.10}$$

$$v_r \frac{\mathrm{d}v_r}{\mathrm{d}r} = c_s^2 \frac{d\ln\rho}{dr} - \frac{GM}{r^2}. \tag{6.11}$$

These equations can be brought into the form

$$(v_r^2 - c_s^2)\frac{\mathrm{d}\ln|v_r|}{\mathrm{d}r} = \frac{2c_s^2}{r} - \frac{GM}{r^2}. \tag{6.12}$$

There is a critical point r_* where the flow becomes transonic. At that point the left hand side and the right hand side of the equation must vanish simultaneously, i.e.

$$\frac{2c_s^2}{r} = \frac{GM}{r^2}. \tag{6.13}$$

This yields for the value of the critical point

$$r_* = \frac{GM}{2c_s^2} \tag{6.14}$$

Assuming $c_s = 100\,\mathrm{km/\,s}$, and using $G \approx 7{\times}10^{-11}\,\mathrm{m^3\,s^{-2}\,kg^{-1}}$, $M = 2{\times}10^{30}\,\mathrm{kg}$, $1\,\mathrm{AU} = 1.5 \times 10^{11}\,\mathrm{m}$, we have $r_* \approx 0.05\,\mathrm{AU}$.

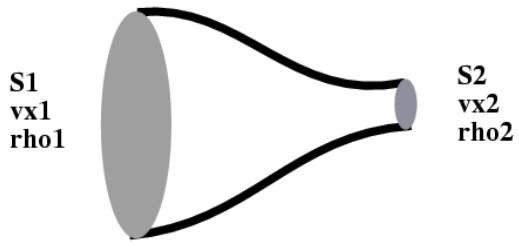


Figure 6.1: In the steady case we have $\boldsymbol{\nabla}\cdot\rho\boldsymbol{u} = 0$. Using Gauss' divergence theorem it follows that $\oint \rho\boldsymbol{u}\cdot d\boldsymbol{S} = 0$. The two surfaces, S_1 and S_2, are the only places where the $\boldsymbol{u}\cdot d\boldsymbol{S} \neq 0$.

The steady wind problem is characterized by two integrals of motions:

$$\tfrac{1}{2}v_r^2 + c_s^2\ln\rho - \frac{GM}{r} = \mathcal{E}, \tag{6.15}$$

$$\dot{M} = 4\pi r^2 \rho v_r, \tag{6.16}$$

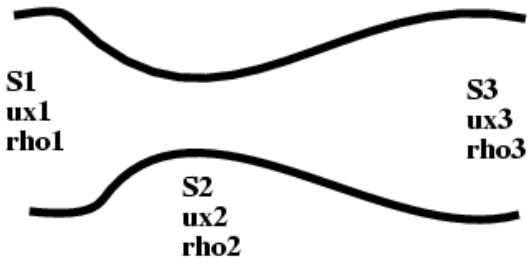


Figure 6.2: Cross-sectional area of a laval nozzle. If the flow is able to reach the sound speed at the point of minimal cross-section, it must go supersonic behind that point.

see Eq. (9) in Shore's book. From this we get

$$\tfrac{1}{2}v_r^2 - c_s^2 \ln v_r - 2c_s^2 \ln r - \frac{GM}{r} = \mathcal{E}' \tag{6.17}$$

where $\mathcal{E}' = \mathcal{E} - c_s^2 \ln(\dot{M}/4\pi)$. (Note that the $\mathcal{C}$ in Shore's book changes its meaning all the time).

To find the energy for the critical solution we assume a critical radius r_*, so $c_s = \sqrt{GM/2r_*}$. At $r = r_*$ we have $v = c_s$. This yields $\mathcal{E}$ for the critical solution that goes through the critical point:

$$\mathcal{E}'_{\rm crit} = \tfrac{1}{2}c_s^2 - c_s^2 \ln c_s - 2c_s^2 \ln r_* - \frac{GM}{r_*} \tag{6.18}$$

We may find the solution through iteration:

$$v_{\rm new} = \sqrt{2\left(\mathcal{E}_r + c_s^2 \ln v\right)} \tag{6.19}$$

where $\mathcal{E}_r \equiv \mathcal{E}' + 2c_s^2 \ln r_* + GM/r_*$ is introduced for convenience. As usual, we can get another solution from the inverse iteration formula

$$v_{\rm new} = \exp\left(\tfrac{1}{2}v^2 - \mathcal{E}_r\right) \tag{6.20}$$

For $\mathcal{E}' < \mathcal{E}'_{\rm crit}$ the upper solutions are obtained from equation (6.19) whilst the lower solutions are obtained by iterating equation (6.20). The critical solution, as well as two more solutions are shown in Fig. 6.3.

Another possibility of obtaining a graphical representation of the possible wind solutions is to do a contour plot of $\mathcal{E}$ as a function of r and v_r; see Fig. 6.4.

Note that the iteration process always yields that solution that has a minimum or a maximum at $r = r_*$. To get the physically sensible solution one needs to reconnect the right branches by hand, This is done in Fig. 6.5, where we have plotted the critical solution together with the density profile.

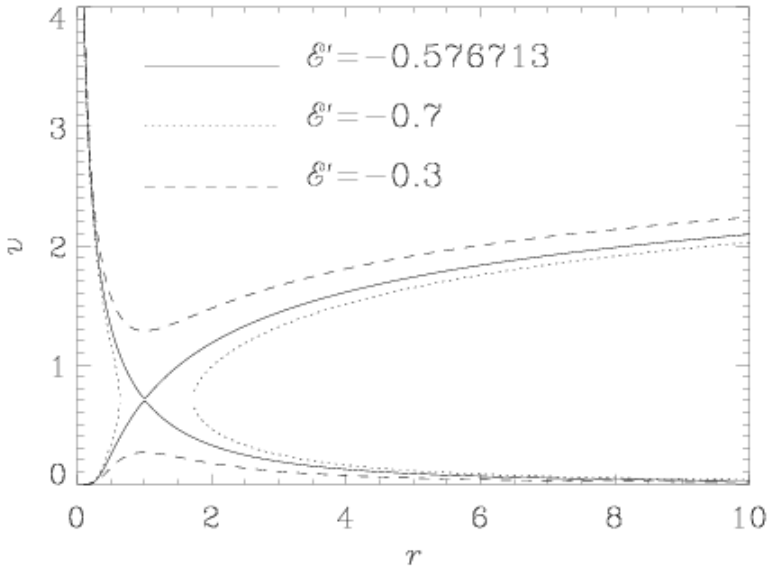


Figure 6.3: The critical wind solution, together with two other solutions for different values of $\mathcal{E}'$. Here we use $r_* = 1$, $GM = 1$, and so $c_s^2 = 0.5$.

6.3 The time-dependent wind problem

It is illuminating to consider the time-dependent problem. For example, one may wonder what happens when one were to increase the density or the velocity at the bottom. The governing equations are

$$\frac{\partial \ln \rho}{\partial t} = -v_r \frac{\partial \ln \rho}{\partial r} - 2\frac{v_r}{r} - \frac{\partial v_r}{\partial r} \tag{6.21}$$

$$\frac{\partial v_r}{\partial t} = -v_r \frac{\partial v_r}{\partial r} - c_s^2 \frac{\partial \ln \rho}{\partial r} - \frac{GM}{r^2}. \tag{6.22}$$

Since the flow comes from the inner boundary at $r = r_0$ we have to specify boundary conditions at $r = r_0$. Video animations of numerical experiments show that specifying the value of v_r at the lower boundary is inconsistent. The solution has a strong desire to come back to the steady solution discussed above (Fig. 6.6). On the other hand, the density on the lower boundary may well be specified arbitrarily. The solution has then just another value of $\dot{M}$.

The solution can actually be reversed and then we have the problem of spherically symmetric accretion (Bondi accretion). This is the reason why one should always write $\ln |v_r|$.

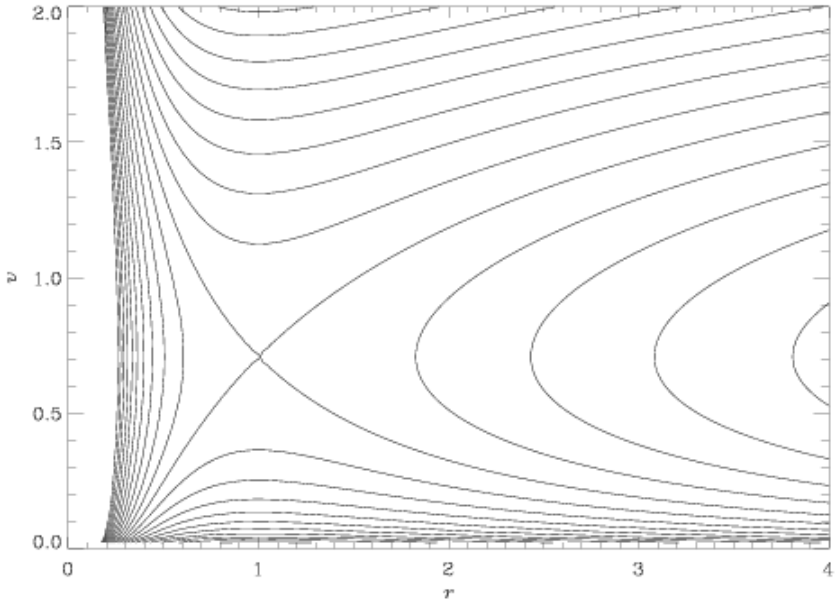


Figure 6.4: Contour plot of $\mathcal{E}'$. The contour levels are equidistant in steps of 0.2, symmetrically about the critical value $\mathcal{E}' = -0.576713$. Again, $r_* = 1$, $GM = 1$, and so $c_s^2 = 0.5$.

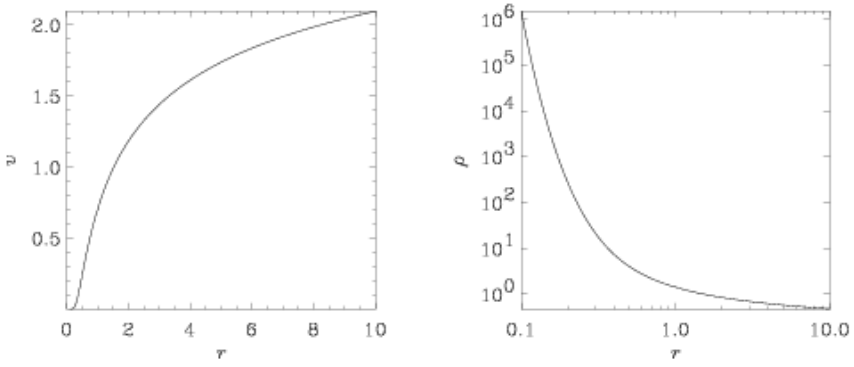


Figure 6.5: Critical solution for v and ρ (assuming $\dot{M} = 4\pi$).

Summary to solving the wind problem. The trick is to write the equations in the form $(v_r^2 - c_s^2) d \ln |v_r|/dx =$ something. This something then tells us *where* the critical point is. Once we know that, we plug that back into the integrated momentum equations to obtain an implicit equation for v_r that goes through the critical point.

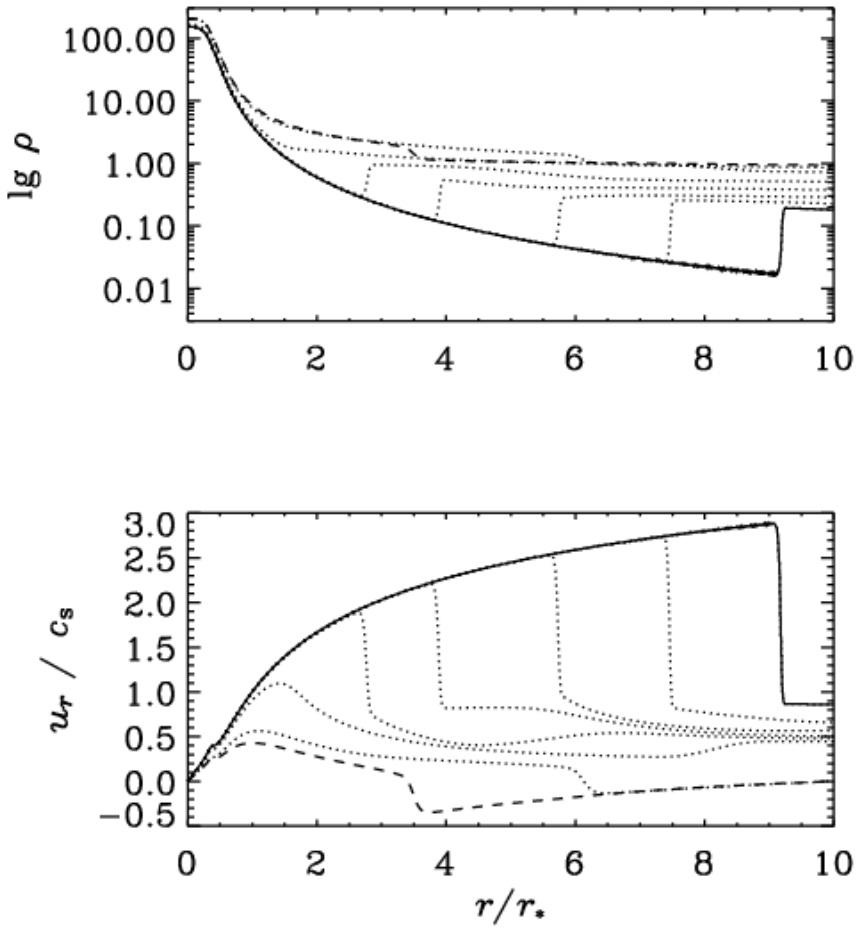


Figure 6.6: Evolution of ρ and u_r for the isothermal wind problem. The solid line shows the last time ($t = 30$) and the dashed line the first time ($t = 4$). The dotted lines show intermediate times ($t = 7$, 13, 18, 21, 24, and 27).

6.4 What drives the wind? The need for a non-static corona

The following consideration illustrates that there cannot be a static solar corona. In the solar corona heat is transported via conduction (as opposed to radiation or convection). In the corona (Spitzer-type) heat conduction is important, so

the conductivity K satisfies

$$K = K_0 \left(\frac{T}{T_0} \right)^{5/2}. \tag{6.23}$$

The heat flux is then given by

$$\boldsymbol{F} = -K \boldsymbol{\nabla} T. \tag{6.24}$$

In order to have a steady state (vanishing divergence of the heat flux, i.e. $\boldsymbol{\nabla} \cdot \boldsymbol{F} = 0$) we have to require that the temperature satisfies

$$T = T_0 \left(\frac{r}{r_0} \right)^{-2/7}, \tag{6.25}$$

where T_0 is the temperature at the base of the corona at $r = r_0$. This is significantly shallower than the temperature for a polytropic stratification, which would imply a r^{-1} behavior with

$$T_{\rm poly} = \left(1 - \frac{1}{\gamma} \right) \frac{\mu}{\mathcal{R}} \frac{GM}{r}. \tag{6.26}$$

Assuming a perfect gas, i.e. $p = \mathcal{R} T \rho$ one can show that

$$p = p_0 \exp \left\{ \frac{7 r_0}{5 H_0} \left[\left(\frac{r}{r_0} \right)^{-5/7} - 1 \right] \right\}. \tag{6.27}$$

Here p_0 is the pressure at the base of the corona and H_0 is the scale height of the pressure, i.e. the distance over which the pressure changes by a factor of e. The problem with this expression is that at infinity, $r \to \infty$, the pressure remains finite,

$$p(\infty) = p_0 \exp \left[-\frac{7 r_0}{5 H_0} \right]. \tag{6.28}$$

In fact, for the sun H_0 is comparable with the solar radius, r_0, and so $p(\infty)$ is just somewhat smaller than p_0, which is definitely quite unrealistic. For a correct solution the pressure must become comparable with the interstellar pressure, which is orders of magnitude smaller than the pressure near the sun. This inconsistency leads then to the idea of a nonstatic corona with an outflow.

6.5 The main points of this chapter

Under certain conditions, an infinitely extended atmosphere around a gravitating body can become unstable and develop a wind, i.e. an outward increasing velocity field that becomes supersonic at some critical radius. In the steady state, this transition to supersonic flows can be completely smooth, i.e. without a shock.

Chapter 7

Coronal mass ejections

7.1 Phenomenology

Coronal mass ejections (CMEs) are spontaneous erruptions on the solar surface. The power associated with such events in on the order of 10^{12} W. CME events often lead to strong auroal activity and can be responsible for power outages and can cause damage to satellites.

Most CMEs originate from active regions. These are groupings of sunspots associated with frequent flares. These regions have closed magnetic field lines, where the magnetic field strength is large enough to allow the containment of the plasma. The CME must open these field lines at least partially to escape from the Sun. However, CMEs can also be initiated in quiet sun regions (although in many cases the quiet region was recently active). During solar minimum, CMEs form primarily in the coronal streamer belt near the solar magnetic equator. During solar maximum, CMEs originate from active regions whose latitudinal distribution is more homogeneous.

Table 7.1: Physical properties of CMEs

power	10^{12} W
energy	10^{25} J
velocity	20...2700 km/s
mass	10^{12}...10^{13} kg
frequency	0.5...6 d^{-1}

7.2 CME models

At first, it was thought that CMEs might be driven by the heat of an explosive flare. However, it soon became apparent that many CMEs were not associated with flares, and that even those that were often began before the flare did.

Because CMEs are initiated in the solar corona (which is dominated by magnetic energy), their energy source must be magnetic. Only flares could provide enough heat energy to drive the CME, and flares get their energy from the magnetic field anyway.

Because the energy of CMEs is so high, it is unlikely that their energy could be directly driven by emerging magnetic fields in the photosphere (although this is still a possibility). Therefore, most models of CMEs assume that the energy is stored up in the coronal magnetic field over a long period of time and then suddenly released by some instability or a loss of equilibrium in the field. There is still no consensus on which of these release mechanisms is correct, and observations are not currently able to constrain these models very well.

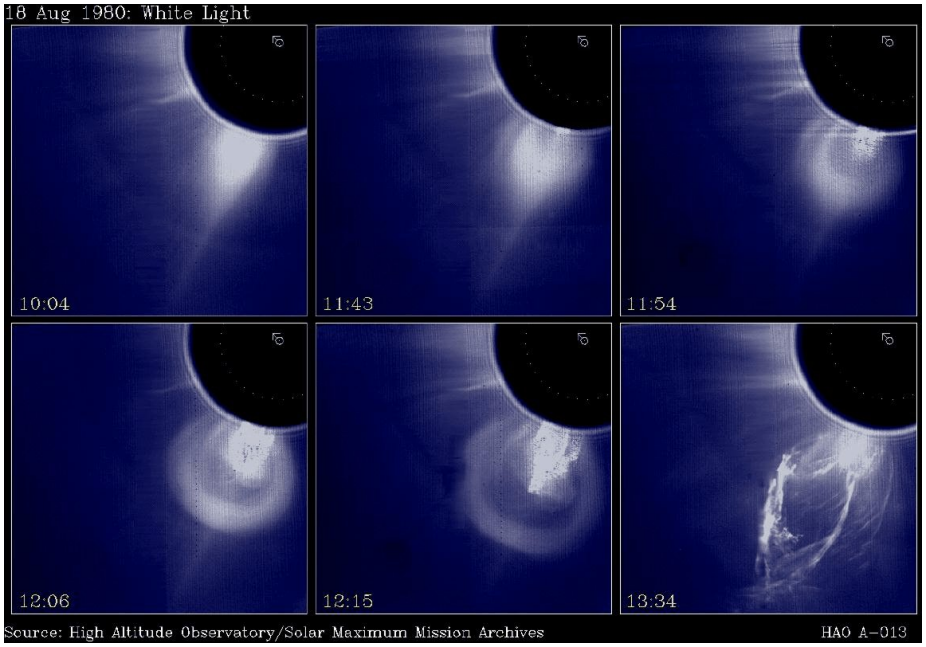


Figure 7.1: A coronal mass ejection and prominence eruption observed in white light from the SMM (Solar Maximum Mission) spacecraft. The time of each panel increases from left to right. The dashed inner circle in each panel is the solar radius, the occulting radius is at 1.6 solar radii. Image courtesy of the High Altitude Observatory.

An important observational constraint on theoretical models is that the stressing of the coronal magnetic field is slow compared to characteristic timescales in the corona. Photospheric driving velocities are typically of order 1 km/s, whereas characteristic coronal speeds are of order several hundred km/s for sound waves and several thousand km/s for Alfvenic motions. Therefore, to a good approximation the energy for a CME is pumped into the corona quasi-

statically.

apparent contradiction between the Aly-Sturrock energy limit and CME observations. Hence, the magnetic system consists of a single coronal arcade (see, e.g., Roumeliotis et al. 1994 ; Mikic & Linker 1994).

The important point is that, for a single arcade system, the observational requirement that the innermost flux near the neutral line open up requires that all the flux in the system open. However, this is forbidden by the Aly-Sturrock limit, at least for a purely magnetically driven eruption, because no sheared field state can have enough energy to open up all the flux.

Chapter 8

Convection, mixing length theory

8.1 Mixing length theory and convection simulations

- Buoyancy is balanced against advection, so $F_{\rm buoy} = F_{\rm D}^{\rm (turb)}$ and therefore $\Delta\varrho\, V g = C_{\rm D}\varrho v^2 S$, where V is the volume of the bubble and S is cross-sectional area. Denoting by $\ell \equiv V/(C_{\rm D}S)$ the mixing length, we have

$$v^2 = \frac{\Delta T}{T}\, g\ell. \tag{8.1}$$

- The only natural length scale in the problem is the scale height, so we assume that the mixing length is some fraction α of the local vertical pressure scale height, i.e.

$$\ell = \alpha H_p. \tag{8.2}$$

- The definition of the scale height is

$$H_p = \frac{\mathcal{R}T}{\mu g} = \frac{c_{\rm s}^2}{g}. \tag{8.3}$$

- The definition of the convective flux is

$$F_{\rm conv} = \varrho v c_p \Delta T. \tag{8.4}$$

  This expression shows that there is a positive convective flux if both the velocity is positive (upwards) and the temperature fluctuation is positive, i.e. if the upward moving fluid parcel is indeed warmer than its surroundings.

- There is actually a second expression for the convective flux, which is however more an approximation than a definition. We know convection can only occur when there is a downward entropy gradient, i.e. if the entropy decreases upwards. The entropy transport is stronger if the downward entropy gradient is stronger. To a first approximation the two are proportional to each other, i.e.

$$\boldsymbol{F}_{\rm conv} = -\chi_{\rm t}\varrho T\boldsymbol{\nabla}s \quad \text{if} \quad \boldsymbol{g}\cdot\boldsymbol{\nabla}s > 0, \tag{8.5}$$

  where $\chi_{\rm t}$ is a (turbulent) diffusion coefficient, and the ϱ and T factors have to be there on dimensional grounds.

- Like with all other types of diffusion coefficients, they are proportional to the speed of the fluid parcels accomplishing the diffusion, as well as the length over which such parcels stay coherent. Thus, we have

$$\chi_{\rm t} = c_\chi v\ell, \tag{8.6}$$

  where c_χ is another free parameter of order unity. (The other free parameter was α.)

This is the set of equations that we need to calculate the stratification in a convection zone. Some of the expressions above, especially (8.1), are severe approximations, and so one usually allows for extra non-dimensional factors in some of those expressions. This is why in other text books some of the expressions may involve somewhat different coefficients in places.

8.1.1 The entropy gradient

To calculate the stratification, the simplest approach was to assume that there was no entropy gradient at all within the convection zone (perfect mixing). With the equations derived above we can do better than that. The main thing we see from those equations is that v is proportional to $(\Delta T/T)^{1/2}$, see equation (8.1), and that therefore $F_{\rm conv}$ is proportional to $(\Delta T/T)^{3/2}$, see equation (8.4). Therefore,

$$\Delta T/T \sim F_{\rm conv}^{2/3} \quad \text{and} \quad v/c_{\rm s} \sim F_{\rm conv}^{1/3}. \tag{8.7}$$

Using equation (8.5) and the fact that $\chi_{\rm t} \sim v \sim F_{\rm conv}^{1/3}$ we have

$$F_{\rm conv} \sim F_{\rm conv}^{1/3}|ds/dz|, \tag{8.8}$$

or

$$|ds/dz| \sim F_{\rm conv}^{2/3}. \tag{8.9}$$

A proper calculation using the equations above shows that

$$\boldsymbol{g}\cdot\boldsymbol{\nabla}s/c_p = k\left(\frac{g}{c_{\rm s}}\right)^2\left(\frac{F_{\rm conv}}{\varrho c_{\rm s}^3}\right)^{2/3} \tag{8.10}$$

where

$$k=\left(\frac{\gamma-1}{\gamma}\right)^{2/3}c_\chi^{-1}\alpha^{-4/3}, \tag{8.11}$$

which is around unity for $c_\chi=1/3$ and $\alpha=1.5$.

8.1.2 Calculating the stratification

If we assume that *all* the flux is carried by convection, i.e. if $F_{\rm tot}=F_{\rm conv}$, then we just have the following system of equations governing the stratification:

$$\frac{ds/c_p}{dz}=-\frac{g}{c_{\rm s}^2}\,k\left(\frac{F_{\rm tot}}{\varrho c_{\rm s}^3}\right)^{2/3}, \tag{8.12}$$

$$\frac{dp}{dz}=-\frac{g}{c_{\rm s}^2}, \tag{8.13}$$

together with $\ln\varrho=\frac{1}{\gamma}\ln p-s/c_p$, and $c_{\rm s}^2=p/\varrho$. A solution of those equations is given in Fig. 8.1. Note the almost perfectly flat s-gradient within the convection zone.

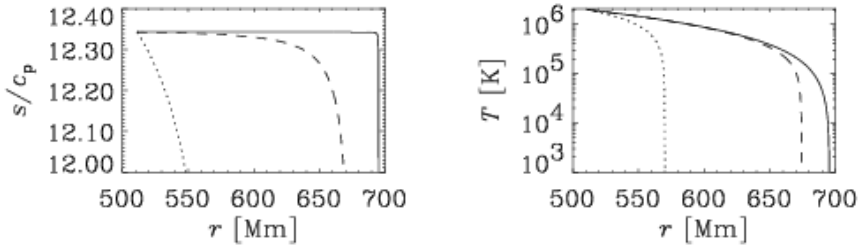


Figure 8.1: Mixing length stratification. The solid line is for solar luminosity, whilst the dashed and dotted lines are respectively for 10^6 and 10^9 times the solar value.

8.1.3 Including the radiative flux consistently

If the radiative flux is to be included consistently we have

$$F_{\rm rad}+F_{\rm conv}=F_{\rm tot}. \tag{8.14}$$

The radiative flux is

$$F_{\rm rad}=-K\frac{dT}{dz}=-KT\frac{d\ln T}{dz}=-KT\frac{d\ln p}{dz}\,\frac{d\ln T}{d\ln p}. \tag{8.15}$$

Here, $d\ln p/dz = -g/c_{\rm s}^2$, and $T = c_{\rm s}^2/(\mathcal{R}/\mu)$, so $c_{\rm s}^2$ cancels out and so

$$F_{\rm rad} = \frac{Kg}{\mathcal{R}/\mu}\,\frac{d\ln T}{d\ln p} \tag{8.16}$$

This can be expressed in terms of the entropy gradient using the differentiated form of equation (2.49),

$$ds/c_p = \frac{1}{\gamma} d\ln p - d\ln\varrho \tag{8.17}$$

which leads to

$$F_{\rm rad} = \frac{Kg}{\mathcal{R}/\mu}\,\left[\left(1-\frac{1}{\gamma}\right) + \frac{ds/c_p}{d\ln p}\right]. \tag{8.18}$$

Thus, the radiative flux has two contributions, one from the adiabatic temperature gradient, and one from the super-adiabatic temperature gradient, so

$$F_{\rm rad} = F_{\rm rad}^{\rm (ad)} + K_r\,\frac{ds/c_p}{d\ln p}, \tag{8.19}$$

where

$$K_r = \frac{Kg}{\mathcal{R}/\mu} \quad\text{and}\quad F_{\rm rad}^{\rm (ad)} = K_r\left(\frac{\gamma-1}{\gamma}\right). \tag{8.20}$$

Inserting this into equation (8.14) we have

$$F_{\rm rad}^{\rm (ad)} + K_r\,\frac{ds/c_p}{d\ln p} + K_c\left(\frac{ds/c_p}{d\ln p}\right)^{3/2} = F_{\rm tot}, \tag{8.21}$$

where $K_c = \varrho c_{\rm s}^3/k^{3/2}$. We now introduce the additional abbreviations

$$\mathcal{G} \equiv \frac{ds/c_p}{d\ln p}, \quad \mathcal{F} \equiv \frac{F_{\rm tot}-F_{\rm rad}^{\rm (ad)}}{K_c}, \quad q = K_r/K_c, \tag{8.22}$$

so we can write

$$\mathcal{G}^{3/2} = q\mathcal{G} + \mathcal{F}, \tag{8.23}$$

which leads to a cubic equation for the entropy gradient, $\mathcal{G}$,

$$\mathcal{G}^3 - q^2\mathcal{G}^2 - 2q\mathcal{F}\mathcal{G} - \mathcal{F}^2 = 0. \tag{8.24}$$

A table of some values is given below, where we compare with the gradient, $\mathcal{G}_0 = (F_{\rm tot}/K_c)^{2/3}$, which is obtained when the radiative flux is neglected. An excellent approximation is to neglect q, in which case $\mathcal{G} = \mathcal{F}^{2/3}$, which is also given in the table.

A solution of the full system of equations, which include more realistic physics than what has been described here, has been given by Spruit (1974); see Table 8.2.

Table 8.1: Solutions of the cubic equation for a solar model

r [Mm]	$T[10^6\,\mathrm{K}]$	$\mathcal{G}/10^{-6}$	$\mathcal{G}_0/10^{-6}$	$\mathcal{F}/10^{-9}$	$\mathcal{F}^{2/3}/10^{-6}$	$q/10^{-9}$
530	1.72	0.20	0.36	0.09	0.20	0.31
550	1.45	0.36	0.47	0.22	0.36	0.26
570	1.21	0.58	0.65	0.44	0.58	0.22
590	0.98	0.90	0.95	0.85	0.90	0.19
610	0.76	1.45	1.49	1.75	1.45	0.15
630	0.56	2.60	2.62	4.19	2.60	0.12
650	0.38	5.65	5.66	13.44	5.65	0.10
670	0.20	19.43	19.44	85.67	19.42	0.07
690	0.03	720.24	720.21	19329.29	719.98	0.02

8.2 Convective overshoot

Both numerical calculations of solar mixing length models as well as helioseismology (described below) indicate that the depth of the convection zone is about 200 Mm. However, near and beyond the boundaries of the convection zones the approximation (8.1) becomes very bad, because it ignores the fact that convective elements have inertia and can therefore overshoot a significant distance into the stably stratified regions. In those layers where the entropy gradient has reversed, a downward moving fluid parcel becomes hotter than its surroundings. Thus, in those layers the convection carries convective flux downwards, so its sign is reversed. This is shown in Fig. 8.4 where profiles of entropy and convective flux, (8.4), are shown from a three-dimensional convection simulation.

Because of strong stratification, convection will be highly inhomogeneous, with narrow downdrafts and broad upwellings. This leads to a characteristic (but irregular) pattern of convection; see Fig. 8.2. A sketch showing how stratification causes upwellings to broaden and downdrafts to converge is given in Fig. 8.3.

Table 8.2: The solar mixing length model of Spruit (1974).

z [Mm]	T [K]	$\rho\,[\mathrm{g\,cm^{-3}}]$	H_p [Mm]	$u_{\rm rms}$ [m/s]	τ [d]	$\nu_{\rm t}\,[\mathrm{cm^2/s}]$	$\Omega_0\tau$
24	1.8×10^5	0.004	8	70	1.3	1.5×10^{12}	0.6
39	3.0×10^5	0.010	13	56	2.8	2.0×10^{12}	1.3
155	1.6×10^6	0.12	48	25	22	3.2×10^{12}	10
198	2.2×10^6	0.20	56	4	157	0.6×10^{12}	70

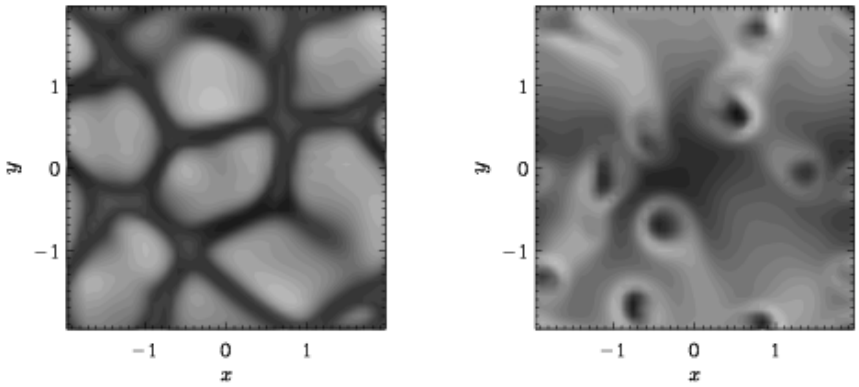


Figure 8.2: Images of temperature in surface and deeper layers of a convection simulation. Dark means lower temperatures and light higher temperatures.

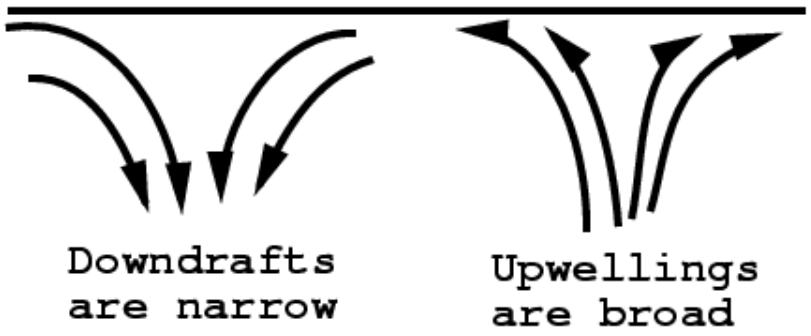


Figure 8.3: Downdrafts contract as they dive into deeper layers where the density is higher. On the other hand, upwellings expand as they ascend into the upper layers where the density is low.

8.3 The main points of this chapter

By using simple scaling arguments between convection velocities and buoyancy force, it is possible to a simple theory between the convective flux and the entropy gradient (which is only approximately zero by comparison with the radiative interior). This yields a reasonable theory for the radial stratification

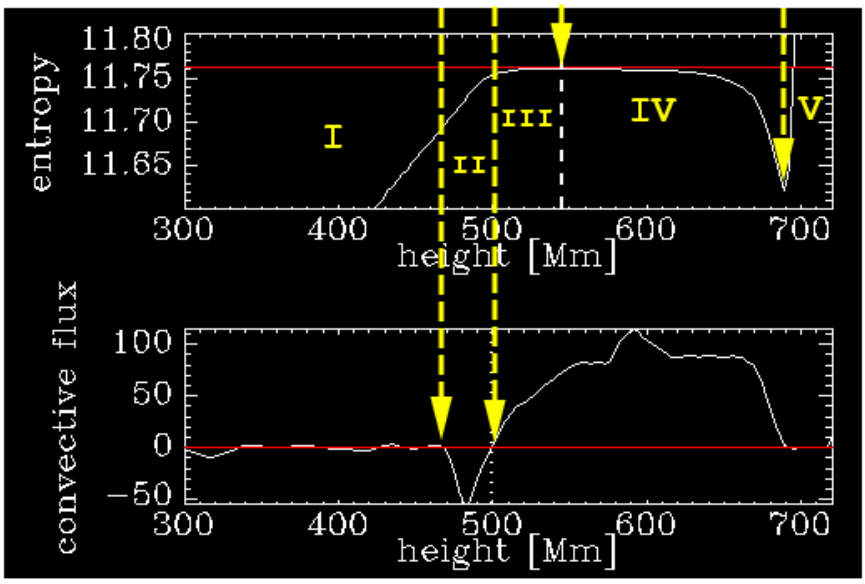


Figure 8.4: Profiles of entropy and convective flux. Region I is the radiative interior, II the overshoot layer, III the radiative heating layer, IV the bulk of the convection zone, and V the surface layer.

of solar/stellar convections zones. As a good approximation, one has $F_{\rm conv} \approx \rho u_{\rm rms}^3$.

Chapter 9

Dimensional Analysis

Dimensional analysis can sometimes be a useful means for understanding the underlying physics. The main difficulties with this approach are twofold. Firstly, it is not always easy to identify the few quantities that govern some physical process. Secondly, there could be some *nondimensional* parameter that may not be close to unity. In that case one gets the right scaling, but quantitatively the result may be off by a few orders of magnitude. In the following we consider a few examples where dimensional analysis works alright: blast waves and various types of turbulence.

9.1 Three-dimensional turbulence

In three-dimensional turbulence the flux of energy, ϵ, is constant and independent of wavenumber k. Using dimensional arguments one can find the energy spectrum $E(k)$ by noting that the energy spectrum is normalized such that $\int_0^\infty E(k)\,dk = \frac{1}{2}\langle \mathbf{u}^2 \rangle$ and that $E(k)$ depends only ϵ and k. Thus, we write

$$E(k) = C\,\epsilon^a k^b, \tag{9.1}$$

where C is a dimensionless constant, and a and b are exponents that are determined by matching the dimensions on both sides of the equation. The dimensions (indicated by square brackets) of $E(k)$, ϵ, and k are

$$[E(k)] = \mathrm{m}^3\,\mathrm{s}^{-2} \tag{9.2}$$

$$[\epsilon] = \mathrm{m}^2\,\mathrm{s}^{-3} \tag{9.3}$$

$$[k] = \mathrm{m}^{-1} \tag{9.4}$$

Thus, the dimensions on both sides of equation (9.1) are

$$\mathrm{m}^3\,\mathrm{s}^{-2} = \mathrm{m}^{2a}\,\mathrm{s}^{-3a}\,\mathrm{m}^{-b}. \tag{9.5}$$

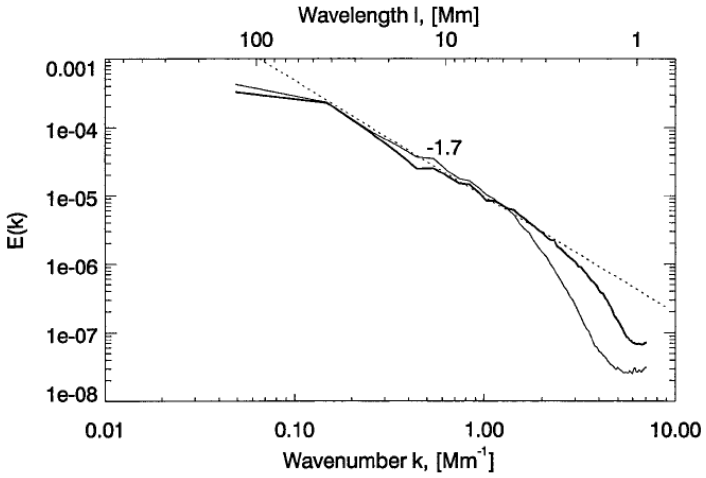


Figure 9.1: Magnetic support of a flare.

To match the dimensions on both sides of the equation we have

$$\text{for the dimension of m:} \quad 3 = 2a - b, \tag{9.6}$$

$$\text{for the dimension of s:} \quad -2 = -3a. \tag{9.7}$$

Thus, $a = 2/3$ and $b = 2a - 3 = 4/3 - 3 = -5/3$, so the spectrum is given by

$$E(k) = C\,\epsilon^{2/3} k^{-5/3}, \tag{9.8}$$

Such a spectra is often seen in turbulent flows. It is also seen in the solar granulation and even in active regions; see Fig. 9.1.

9.2 Two-dimensional turbulence

In two-dimensional turbulence the rate of enstrophy injection, $\beta = \frac{d}{dt}\langle \boldsymbol{\omega}^2 \rangle$, is constant and independent of wavenumber k. Using dimensional arguments one can find the energy spectrum $E(k)$ by noting that the energy spectrum is normalized such that $\int_0^\infty E(k)\,dk = \frac{1}{2}\langle \mathbf{u}^2 \rangle$ and that $E(k)$ depends only β and k. Thus, we write

$$E(k) = C\,\beta^a k^b, \tag{9.9}$$

where C is a dimensionless constant, and a and b are exponents that are determined by matching the dimensions on both sides of the equation. The dimensions (indicated by square brackets) of $E(k)$, β, and k are

$$[E(k)] = \mathrm{m}^3\,\mathrm{s}^{-2} \tag{9.10}$$

$$[\beta] = \mathrm{s}^{-3} \tag{9.11}$$

$$[k] = \mathrm{m}^{-1} \tag{9.12}$$

Thus, the dimensions on both sides of equation (9.9) are

$$\mathrm{m}^3\,\mathrm{s}^{-2} = \mathrm{s}^{-3a}\,\mathrm{m}^{-b}. \tag{9.13}$$

To match the dimensions on both sides of the equation we have

$$\text{for the dimension of m:} \quad 3 = -b, \tag{9.14}$$

$$\text{for the dimension of s:} \quad -2 = -3a. \tag{9.15}$$

Thus, $b = -3$ and $a = 2/3$, so the spectrum is given by

$$E(k) = C\,\beta^{2/3} k^{-3}, \tag{9.16}$$

9.3 Hydromagnetic turbulence (Iroshnikov–Kraichnan theory)

Using dimensional arguments one can find the energy spectrum $E(k)$ for hydromagnetic turbulence, by assuming that the energy transfer is governed by the geometric mean of the Alfvén speed and the energy transfer rate ϵ. Thus, we assume the spectrum to be of the form

$$E(k) = C\,(v_A \epsilon)^a k^b, \tag{9.17}$$

where C is a dimensionless constant, v_A is the Alfvén speed, ϵ (with dimension $\mathrm{m}^2\,\mathrm{s}^{-3}$) the energy injection rate, and k the wavenumber. The dimensions (indicated by square brackets) of $E(k)$, v_A, ϵ, and k are

$$[E(k)] = \mathrm{m}^3\,\mathrm{s}^{-2} \tag{9.18}$$

$$[v_A] = \mathrm{m}\,\mathrm{s}^{-1} \tag{9.19}$$

$$[\epsilon] = \mathrm{m}^2\,\mathrm{s}^{-3} \tag{9.20}$$

$$[k] = \mathrm{m}^{-1} \tag{9.21}$$

Thus, the dimensions on both sides of equation (9.17) are

$$\mathrm{m}^3\,\mathrm{s}^{-2} = \mathrm{m}^a\,\mathrm{s}^{-a}\,\mathrm{m}^{2a}\,\mathrm{s}^{-3a}\,\mathrm{m}^{-b}. \tag{9.22}$$

To match the dimensions on both sides of the equation we have

$$\text{for the dimension of m:} \quad 3 = 3a - b, \tag{9.23}$$

$$\text{for the dimension of s:} \quad -2 = -4a. \tag{9.24}$$

Thus, $a = 1/2$ and so $b = 3a - 3 = 3/2 - 3 = -3/2$, so the spectrum is given by

$$E(k) = C\,(v_A \epsilon)^{1/2} k^{-3/2}, \tag{9.25}$$

A final comment is here in order. Although the $k^{-3/2}$ scaling has been verified in two-dimensional numerical simulations, there is now mounting evidence in favor of the more conventional $k^{-5/3}$ scaling even in the magnetic case.[1]

[1] See recent paper by Biskamp & Müller 1999, Phys. Plasmas **7**, 4889

Chapter 10

Magnetic fields

This chapter contains more material than will be presented in the course. The extra material is included mainly for interest.

10.1 The Lorentz force

Astrophysical bodies are almost always electrically conducting and can thus interact with magnetic fields. The first example concerns the support of prominences against gravity by a magnetic field. There, we considered only one particle and the force exerted on a single particle was given by $q\boldsymbol{u}\times\boldsymbol{B}$. Now, if there are many particles with given number density n (= number of particles per unit volume), then the force on the gas per unit volume will be $nq\boldsymbol{u}\times\boldsymbol{B}$. The expression $nq\boldsymbol{u}$ is the current density, so

$$\boldsymbol{J}\equiv nq\boldsymbol{u}. \tag{10.1}$$

Thus the Lorentz force per unit volume is $\boldsymbol{J}\times\boldsymbol{B}$ and, if there are no other terms, this will accelerate the gas. The acceleration per unit volume is $\varrho d\boldsymbol{u}/dt$, so the equation of motion takes the form

$$\varrho\frac{d\boldsymbol{u}}{dt}=\boldsymbol{J}\times\boldsymbol{B}+\text{possibly further terms.} \tag{10.2}$$

In order to calculate this force, we have to express $\boldsymbol{J}$ in terms of other known quantities. At this point we may express $\boldsymbol{J}$ using one of Maxwell's equations

$$\mu_0\boldsymbol{J}=\boldsymbol{\nabla}\times\boldsymbol{B}, \tag{10.3}$$

which is also called Ampère's law, and where the Faraday displacement current $c^{-2}\partial\boldsymbol{E}/\partial t$ is neglected. Thus the Lorentz force takes the form $\boldsymbol{F}_L\equiv\boldsymbol{J}\times\boldsymbol{B}=\frac{1}{\mu_0}(\boldsymbol{\nabla}\times\boldsymbol{B})\times\boldsymbol{B}$, which can also be written in the form

$$\boldsymbol{F}_L=-\boldsymbol{\nabla}(\boldsymbol{B}^2/2\mu_0)+(\boldsymbol{B}\cdot\boldsymbol{\nabla})\boldsymbol{B}/\mu_0, \tag{10.4}$$

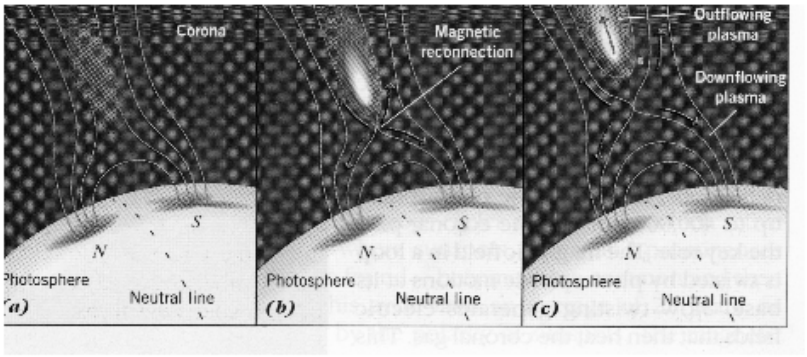


Figure 10.1: Magnetic support of a flare.

which shows that the Lorentz force has a gradient term (magnetic pressure gradient) and a derivative of the field strength along the direction of the field. The latter tends to contract field lines (magnetic tension).

10.2 Magnetic support of prominences

The magnetic field is able to support fluid against gravity. An example is the quasi-steady support of prominences in the solar corona. Figure 10.1 shows a simple cartoon picture of a V-shaped magnetic field line where the Lorentz force points upwards, trying to move fluid with the field lines such as to shorten the field lines.

We take a simple parabola-shaped field line. In order to automatically satisfy the condition $\boldsymbol{\nabla} \cdot \boldsymbol{B} = 0$ we write $\boldsymbol{B} = \boldsymbol{\nabla} \times (A\hat{\boldsymbol{y}})$, where $\hat{\boldsymbol{y}}$ is the unit vector in the direction out of the paper. The lines $A = \text{const}$ are parallel to field lines (at least in two dimensions). This can be verified by showing that the gradient of A is perpendicular to $\boldsymbol{B}$:

$$\boldsymbol{B} \cdot (\boldsymbol{\nabla} A) = \begin{pmatrix} -\partial_z A \\ 0 \\ \partial_x A \end{pmatrix} \cdot \begin{pmatrix} \partial_x A \\ 0 \\ \partial_z A \end{pmatrix} = 0. \tag{10.5}$$

Now, we write our parabola-shaped field lines as

$$A = z - x^2, \tag{10.6}$$

see Fig. 10.2. Let us now calculate the resulting field, current and Lorentz force:

$$\boldsymbol{B} = \begin{pmatrix} \partial_x \\ 0 \\ \partial_z \end{pmatrix} \times \begin{pmatrix} 0 \\ z - x^2 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ -2x \end{pmatrix} \tag{10.7}$$

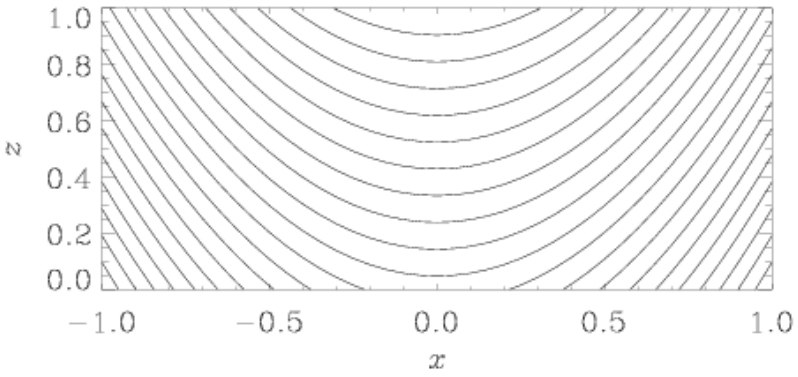


Figure 10.2: Contours of $A = z - x^2$, as a model for magnetic field lines supporting prominences against gravity.

Thus, the current is then

$$\boldsymbol{J} = \frac{1}{\mu_0}\begin{pmatrix}\partial_x\\0\\\partial_z\end{pmatrix}\times\begin{pmatrix}-1\\0\\-2x\end{pmatrix} = \frac{1}{\mu_0}\begin{pmatrix}0\\2\\0\end{pmatrix} \tag{10.8}$$

and so the Lorentz force is

$$\boldsymbol{J}\times\boldsymbol{B} = \frac{1}{\mu_0}\begin{pmatrix}0\\2\\0\end{pmatrix}\times\begin{pmatrix}-1\\0\\-2x\end{pmatrix} = \frac{1}{\mu_0}\begin{pmatrix}-4x\\0\\2\end{pmatrix} \tag{10.9}$$

i.e. the Lorentz force has a vertical component upwards and points towards the z-axis.

10.3 Magnetic field evolution

The evolution of the magnetic field is governed by the Faraday equation

$$\frac{\partial\boldsymbol{B}}{\partial t} = -\boldsymbol{\nabla}\times\boldsymbol{E}, \tag{10.10}$$

Ampère's equation

$$\boldsymbol{J} = \boldsymbol{\nabla}\times\boldsymbol{B}/\mu_0, \tag{10.11}$$

and Ohm's law,

$$\boldsymbol{J} = \sigma(\boldsymbol{E} + \boldsymbol{u}\times\boldsymbol{B}). \tag{10.12}$$

Eliminating $\boldsymbol{E}$ using (10.10) and (10.12) we obtain the induction equation,

$$\frac{\partial \boldsymbol{B}}{\partial t} = \boldsymbol{\nabla} \times (\boldsymbol{u} \times \boldsymbol{B} - \boldsymbol{J}/\sigma). \tag{10.13}$$

The quantity σ is here the conductivity (not to be confused with the Stefan-Boltzmann constant!!). The magnetic diffusivity is $\eta = (\mu_0 \sigma)^{-1}$ and has dimensions m^2/s. If η is constant, then the induction equation can also be written in the form

$$\frac{\partial \boldsymbol{B}}{\partial t} = \boldsymbol{\nabla} \times (\boldsymbol{u} \times \boldsymbol{B}) + \eta \nabla^2 \boldsymbol{B}. \tag{10.14}$$

In SI units Maxwell's equations can be written in the form

$$\frac{\partial \boldsymbol{B}}{\partial t} = -\boldsymbol{\nabla} \times \boldsymbol{E}, \quad \text{Faraday's law} \tag{10.15}$$

$$\mu_0 \boldsymbol{J} + \frac{1}{c^2} \frac{\partial \boldsymbol{E}}{\partial t} = \boldsymbol{\nabla} \times \boldsymbol{B}, \quad \text{Ampère's law} \tag{10.16}$$

together with

$$\boldsymbol{\nabla} \cdot \boldsymbol{B} = 0, \quad \text{and} \quad \boldsymbol{\nabla} \cdot \boldsymbol{E} = \varrho_e \tag{10.17}$$

where ϱ_e is the charge density. The $\partial \boldsymbol{E}/\partial t$ term is also called the Faraday displacement current. It is usually small compared with the other two terms in that equation. There are two exceptions where it can become important: (i) if there is a vacuum, i.e. if the ordinary current $\boldsymbol{J}$ vanishes, see equation (10.12), or if there are rapid variations over large length scales so that the Faraday displacement current becomes comparable to $\boldsymbol{\nabla} \times \boldsymbol{B}$, i.e. the typical velocity becomes comparable with the speed of light. The former occurs in the atmosphere, where it is responsible for radio waves, whilst the latter may become important near a black hole, where all velocities become comparable with the speed of light. In all other cases the Faraday displacement current may safely be neglected. The resulting equations are also called the pre-Maxwell equations.

10.4 Frozen-in magnetic fields

If magnetic diffusion vanishes, i.e. $\eta \to 0$ and $\sigma \to \infty$ (high conductivity limit), we may neglect the diffusion term and the induction equation then takes the form

$$\frac{\partial \boldsymbol{B}}{\partial t} = \boldsymbol{\nabla} \times (\boldsymbol{u} \times \boldsymbol{B}) = -(\boldsymbol{u} \cdot \boldsymbol{\nabla})\boldsymbol{B} + (\boldsymbol{B} \cdot \boldsymbol{\nabla})\boldsymbol{u} - \boldsymbol{B}(\boldsymbol{\nabla} \cdot \boldsymbol{u}). \tag{10.18}$$

Compare this now with the evolution equation of a material line element, $\delta \boldsymbol{l}$. Let $\boldsymbol{u}$ be the velocity on one end of the line element, then the velocity at the other end of the line element $\delta \boldsymbol{l}$ is $\boldsymbol{u} + (\delta \boldsymbol{l} \cdot \boldsymbol{\nabla})\boldsymbol{u}$. Thus, within the time dt the change of $\delta \boldsymbol{l}$ is equal to $dt \delta \boldsymbol{l}$. Therefore the evolution of $\delta \boldsymbol{l}$ satisfies the equation

$$\frac{d}{dt} \delta \boldsymbol{l} = (\delta \boldsymbol{l} \cdot \boldsymbol{\nabla})\boldsymbol{u} \tag{10.19}$$

This equation is equivalent to the evolution equation of the magnetic field if $\eta = 0$ and if $\boldsymbol{\nabla} \cdot \boldsymbol{u} = 0$ is assumed. The latter assumption is unessential: if $\boldsymbol{\nabla} \cdot \boldsymbol{u} \neq 0$ then we have to invoke the continuity equation

The derivative d/dt is here taken in a frame co-moving with the fluid at speed $\boldsymbol{u}$. This derivative is also called lagrangian derivative; to emphasize that one deals with a lagrangian derivative one often uses a capital D for the differential. Expressing it in terms of the normal non-moving frame of reference this becomes

$$D/Dt \equiv \partial/\partial t + (\boldsymbol{u} \cdot \boldsymbol{\nabla}). \tag{10.20}$$

If the velocity is divergence-free (no sources or sinks, i.e. the flow is incompressible) then the magnetic field evolves according to

$$\frac{D\boldsymbol{B}}{Dt} = (\boldsymbol{B} \cdot \boldsymbol{\nabla})\boldsymbol{u}, \tag{10.21}$$

which is exactly the same equation as that for $\delta\boldsymbol{l}$. Thus,we conclude that the magnetic field vectors evolve in the same way as material line elements do. If the $\boldsymbol{B}$-vectors had finite length, the two ends of the vector would coincide with the locations of particles in the flow. Furthermore, if the flow diverges locally it will stretch the magnetic field lines, which will leads to their enhancement. This stretching is an important ingredient of all dynamos.

10.5 The magnetic vector potential

It is sometimes convenient to consider the evolution of the vector potential $\boldsymbol{A}$, because then the magnetic field $\boldsymbol{B} = \boldsymbol{\nabla} \times \boldsymbol{A}$ is guaranteed to be divergence-free. The induction equation (10.13) can be "uncurled", i.e. the curl can be removed on both sides of the equation. However, this leads to an uncertainly, because a gradient term could always be added to the uncurled equation without changing $\boldsymbol{B}$. Thus, we have

$$\frac{\partial \boldsymbol{A}}{\partial t} = \boldsymbol{u} \times \boldsymbol{B} - \boldsymbol{J}/\sigma - \boldsymbol{\nabla}\phi, \tag{10.22}$$

where ϕ is called the gauge potential, which is really like an integration constant. We are free to choose any gauge that is convenient. Note that

$$\mu_0 \boldsymbol{J} = \boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \boldsymbol{A} \equiv -\nabla^2 \boldsymbol{A} + \boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot \boldsymbol{A}). \tag{10.23}$$

A convenient gauge is the so-called Lorentz gauge,

$$\phi = \boldsymbol{\nabla} \cdot \boldsymbol{A}, \tag{10.24}$$

in which the evolution equation for $\boldsymbol{A}$ becomes

$$\frac{\partial \boldsymbol{A}}{\partial t} = \boldsymbol{u} \times \boldsymbol{B} + \eta \nabla^2 \boldsymbol{A}. \tag{10.25}$$

However, this works only if the magnetic diffusivity $\eta = (\mu_0 \sigma)^{-1}$ is constant.

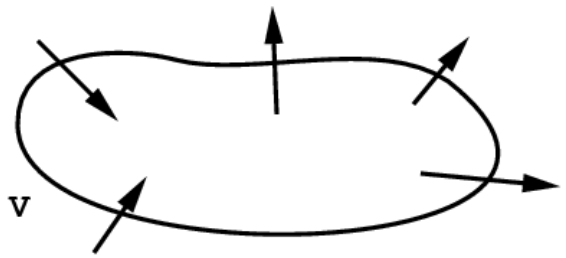


Figure 10.3: Flux conservation. As many field lines enter the volume as field lines leave the volume.

10.6 Flux conservation

The condition $\boldsymbol{\nabla} \cdot \boldsymbol{B} = 0$ means that there are no magnetic monopoles, from which magnetic field lines would originate. This becomes obvious when taking the volume integral of $\boldsymbol{\nabla} \cdot \boldsymbol{B} = 0$, which can be turned into a surface integral by Gauss' integral theorem; so

$$0 = \int_V \boldsymbol{B}\, dV = \oint_S \boldsymbol{B} \cdot d\boldsymbol{S}, \tag{10.26}$$

see Fig. 10.3. Here, $S = \partial V$ is the closed surface bounding the volume V. If field lines go out of the volume, then there must be an equal amount of field lines going into the volume, such that $\oint_S \boldsymbol{B} \cdot d\boldsymbol{S} = 0$.

In Fig. 10.4 we consider the magnetic field in a tube (flux tube). Only at the two ends does the field stick out of the tube. Integrating over the tube we have

$$0 = \oint_S \boldsymbol{B} \cdot d\boldsymbol{S} = \oint_{S_1} \boldsymbol{B} \cdot d\boldsymbol{S} + \oint_{S_2} \boldsymbol{B} \cdot d\boldsymbol{S}. \tag{10.27}$$

Note that in Fig. 10.4 the normal of the surface element points outwards. We now define the flux though a surface S as

$$\Phi = \int_S \boldsymbol{B} \cdot d\boldsymbol{S}, \tag{10.28}$$

where now $d\boldsymbol{S}$ points always in the same direction. Then we see that

$$\Phi = \text{constant along the tube.} \tag{10.29}$$

This property is referred to as flux conservation.

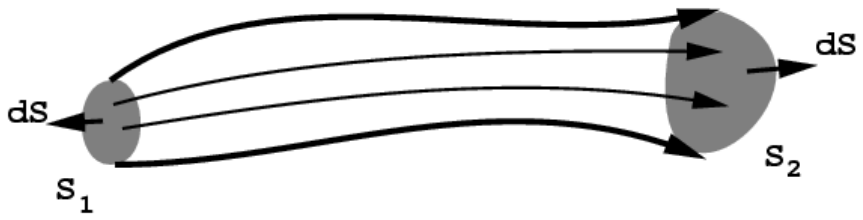


Figure 10.4: Flux conservation. The total surface integral gives zero: nothing comes from the wall of the tube and the contributions from the two ends must be equal in magnitude, but of opposite sign.

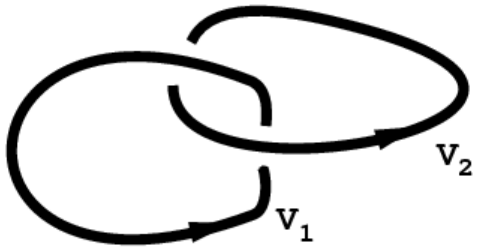


Figure 10.5: Two interlinked flux loops.

10.7 Connection with topology

The dot product $\boldsymbol{A} \cdot \boldsymbol{B}$ is of some importance in that it can be related to the topology of magnetic flux ropes. Let us define the quantity

$$H = \int_V \boldsymbol{A} \cdot \boldsymbol{B}\, dV, \tag{10.30}$$

where the integration volume is chosen such that the normal component of $\boldsymbol{\mathcal{B}}$ vanishes on the boundary, i.e. $\hat{\boldsymbol{n}} \cdot \boldsymbol{B} = 0$. Let us consider two interlinked flux loops (Fig. 10.5).

For the volume V_1 of the first loop we have

$$H_1 = \int_{V_1} \boldsymbol{A} \cdot \boldsymbol{B}\, dV = \int_{V_1} \boldsymbol{A} \cdot \boldsymbol{B}\, (d\boldsymbol{l} \cdot d\boldsymbol{S}), \tag{10.31}$$

where $d\boldsymbol{S}$ is the surface element across the tube and $d\boldsymbol{l}$ is the line element along the tube. Note that $\boldsymbol{B} \parallel d\boldsymbol{l} \parallel d\boldsymbol{S}$, so the integral can also be written as

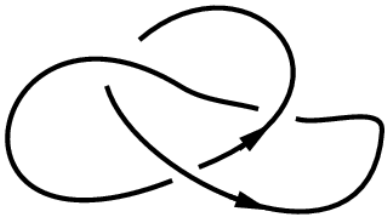


Figure 10.6: The trefoil knot.

$\int(\boldsymbol{A}\cdot d\boldsymbol{l})(\boldsymbol{B}\cdot d\boldsymbol{S})$. In $\boldsymbol{A}\cdot\boldsymbol{B}$ only the component of $\boldsymbol{A}$ that is parallel to $\boldsymbol{B}$ (and $d\boldsymbol{l}$) matters. This component is only affected by the field of the other tube, but since the field of the other tube vanishes inside V_1, the parallel component of $\boldsymbol{A}$ must be *constant* in V_1. Therefore we can split the integral into two separate integrals, so

$$H_1 = \left(\int_{S_1} \boldsymbol{B}\cdot d\boldsymbol{S}\right)\left(\oint_{C_1} \boldsymbol{A}\cdot d\boldsymbol{l}\right), \tag{10.32}$$

where C_1 is a closed line along the tube. (It doesn't matter where across the tube this line goes because the parallel component of $\boldsymbol{A}$, and therefore also $\boldsymbol{A}\cdot d\boldsymbol{l}$, is constant). Now, the first integral is just the flux Φ_1 of the first tube. The second integral is actually the flux of the second tube, Φ_2, because, according to Stokes' integral theorem,

$$\Phi_2 = \int_{S_2} \boldsymbol{B}\cdot d\boldsymbol{S} = \int_{S(C_1)} (\boldsymbol{\nabla}\times\boldsymbol{B})\cdot d\boldsymbol{S} = \oint_{C_1} \boldsymbol{A}\cdot d\boldsymbol{l}. \tag{10.33}$$

Here $S(C_1)$ is the surface enclosed by the curve C_1. We were able to take the integral over this bigger cross-section, because the field outside S_2 and inside $S(C_1)$ vanishes. Therefore,

$$H_1 = \Phi_1\Phi_2. \tag{10.34}$$

By the same arguments we find the same result for H_2 when considering the other tube, so $H_2 = \Phi_1\Phi_2$. Therefore the integral over all space is

$$H = \int_V \boldsymbol{A}\cdot\boldsymbol{B}\, dV = H_1 + H_2 = 2\Phi_1\Phi_2. \tag{10.35}$$

For loops that are linked in more complicated ways one gets the product of the fluxes multiplied by the winding number. In particular, for the trefoil knot shown in Fig. 10.6, one gets the result $H = 2\Phi^2$, even through this knot consists only of a single (knotted) flux tube.

10.8 Conservation of magnetic helicity

We can easily derive an equation for the evolution of the magnetic helicity,

$$\frac{dH}{dt} = \frac{d}{dt}\int_V \boldsymbol{A}\cdot\boldsymbol{B}\,dV = \int_V \frac{\partial}{\partial t}(\boldsymbol{A}\cdot\boldsymbol{B})\,dV = \int_V \left(\frac{\partial \boldsymbol{A}}{\partial t}\cdot\boldsymbol{B} + \boldsymbol{A}\cdot\frac{\partial \boldsymbol{B}}{\partial t}\right) dV \tag{10.36}$$

using equation (10.10) and its uncurled form,

$$\frac{\partial \boldsymbol{A}}{\partial t} = -\boldsymbol{E} + \boldsymbol{\nabla}\phi. \tag{10.37}$$

This leads to

$$\frac{dH}{dt} = \int_V \left(-\boldsymbol{E}\cdot\boldsymbol{B} + \boldsymbol{B}\cdot\boldsymbol{\nabla}\phi - \boldsymbol{A}\cdot\boldsymbol{\nabla}\times\boldsymbol{E}\right)\,dV. \tag{10.38}$$

Using integration by parts using

$$\boldsymbol{A}\cdot\boldsymbol{\nabla}\times\boldsymbol{E} = -\boldsymbol{E}\cdot\boldsymbol{B} - \boldsymbol{\nabla}\cdot(\boldsymbol{E}\times\boldsymbol{A}) \tag{10.39}$$

and, because $\boldsymbol{\nabla}\cdot\boldsymbol{B} = 0$,

$$\boldsymbol{B}\cdot\boldsymbol{\nabla}\phi = \boldsymbol{\nabla}\cdot(\boldsymbol{B}\phi) + \phi\boldsymbol{\nabla}\cdot\boldsymbol{B} = \boldsymbol{\nabla}\cdot(\boldsymbol{B}\phi), \tag{10.40}$$

we have

$$\frac{dH}{dt} = -2\int_V (\boldsymbol{E}\cdot\boldsymbol{B})\,dV - \oint (\boldsymbol{E}\times\boldsymbol{A} + \boldsymbol{B}\phi)\cdot d\boldsymbol{S}. \tag{10.41}$$

If the integration is over all space we can ignore surface integrals. The surface integral gives also no contribution if the boundary conditions are periodic or if the boundaries are perfectly conduction (no horizontal electric field; i.e. a short circuit), i.e.

$$\frac{dH}{dt} = -2\int_V (\boldsymbol{E}\cdot\boldsymbol{B})\,dV. \tag{10.42}$$

In all other cases, however, the surface integral can give a contribution.

Now, looking at Ohm's law, equation (10.12), we have

$$\boldsymbol{E} = -\boldsymbol{u}\times\boldsymbol{B} + \eta\mu_0\boldsymbol{J}, \tag{10.43}$$

and so the helicity equation takes the form

$$\frac{dH}{dt} = -2\eta\mu_0\int_V (\boldsymbol{J}\cdot\boldsymbol{B})\,dV. \tag{10.44}$$

In the ideal case, $\eta = 0$, and magnetic helicity is conserved.

10.9 The main points of this chapter

Magnetic fields have some similarities to velocity fields (diffusion and advection properties), but there are also some new effects (restoring forcing allowing MHD waves, and the importance of helicity).

Chapter 11

Alfvén and magnetosonic waves

In this chapter we demonstrate the technique of linearizing the MHD equations, which can then be used to study phenomena of small amplitude. This applies mainly to waves (provided they are of small amplitude), but it also applies to weak perturbations that can in some systems grow exponentially and would eventually no longer be small. In that case such an analysis can establish the possibility of an instability.

11.1 Linearizing the MHD equations

We now consider the isothermal MHD equations in the form

$$\rho\frac{\partial \boldsymbol{v}}{\partial t}+\rho\boldsymbol{v}\cdot\boldsymbol{\nabla}\boldsymbol{v}=-c_s^2\boldsymbol{\nabla}\rho-\boldsymbol{\nabla}\left(\frac{\boldsymbol{B}^2}{2\mu_0}\right)+\frac{1}{\mu_0}(\boldsymbol{B}\cdot\boldsymbol{\nabla})\boldsymbol{B}. \tag{11.1}$$

Whenever we linearize, we have to linearize about something, which must be a solution to the equations. Those solutions may well be complicated, but then we won't be able to do the analysis so easily on the blackboard. Therefore we now go for simple solutions. One solution is where the magnetic field is constant and uniform, i.e. $\boldsymbol{B}=\boldsymbol{B}_0$, and where the density is also constant, i.e. $\rho=\rho_0$, but the velocity is zero, i.e. $\boldsymbol{v}=\boldsymbol{v}_0=0$. In that case the terms on the lhs of equation (11.1) vanish. The terms on the rhs of equation (11.1) also vanish because because ρ_0 and $\boldsymbol{B}_0$ are uniform. So we do have a solution now.

We now proceed by linearizing term by term, i.e. we write

$$\rho=\rho_0+\rho', \tag{11.2}$$

$$\boldsymbol{v}=\boldsymbol{v}_0+\boldsymbol{v}'=\boldsymbol{v}', \tag{11.3}$$

$$\boldsymbol{B}=\boldsymbol{B}_0+\boldsymbol{B}'. \tag{11.4}$$

The first term in (11.1) then becomes

$$\rho\frac{\partial \boldsymbol{v}}{\partial t} = (\rho_0 + \rho')\frac{\partial \boldsymbol{v}'}{\partial t} = \rho_0\frac{\partial \boldsymbol{v}'}{\partial t} + \rho'\frac{\partial \boldsymbol{v}'}{\partial t}. \tag{11.5}$$

Here the second term is quadratic in the perturbations. Since those perturbations are small, then something small squared will be even smaller and will hence be neglected. The advection term is already nonlinear in the perturbations, because there is no zero-order term, so it vanishes altogether. Next, the pressure gradient term is linear already, so we have

$$c_s^2\boldsymbol{\nabla}\rho = c_s^2\boldsymbol{\nabla}\rho', \tag{11.6}$$

because ρ_0 is constant. For the magnetic pressure gradient we have

$$\boldsymbol{\nabla}\left(\frac{\boldsymbol{B}^2}{2\mu_0}\right) = \boldsymbol{\nabla}\left(\frac{\boldsymbol{B}_0\cdot\boldsymbol{B}'}{\mu_0}\right) + \text{quadratic terms}, \tag{11.7}$$

because

$$\boldsymbol{B}^2 = (\boldsymbol{B}_0 + \boldsymbol{B}')\cdot(\boldsymbol{B}_0 + \boldsymbol{B}') = \boldsymbol{B}_0^2 + 2\boldsymbol{B}_0\cdot\boldsymbol{B}' + \boldsymbol{B}'^2, \tag{11.8}$$

of which only the second term survives. The first one is constant and gives no contribution under the gradient, and the last term is quadratic in the perturbations. Finally, the magnetic stretching term gives simply

$$\frac{1}{\mu_0}(\boldsymbol{B}\cdot\boldsymbol{\nabla})\boldsymbol{B} = \frac{1}{\mu_0}(\boldsymbol{B}_0\cdot\boldsymbol{\nabla})\boldsymbol{B}' + \text{quadratic terms}. \tag{11.9}$$

Thus, our linearized momentum equation takes the form

$$\rho_0\frac{\partial \boldsymbol{v}'}{\partial t} = -c_s^2\boldsymbol{\nabla}\rho' - \boldsymbol{\nabla}\left(\frac{\boldsymbol{B}_0\cdot\boldsymbol{B}'}{\mu_0}\right) + \frac{1}{\mu_0}(\boldsymbol{B}_0\cdot\boldsymbol{\nabla})\boldsymbol{B}'. \tag{11.10}$$

You can imagine that with a little bit of experience one can write down those equations straight away. We do this now with the remaining equations, the induction and continuity equations, looking at equations (**??**) and (**??**),

$$\frac{\partial \boldsymbol{B}'}{\partial t} = \boldsymbol{B}_0\cdot\boldsymbol{\nabla}\boldsymbol{v}' - \boldsymbol{B}_0\boldsymbol{\nabla}\cdot\boldsymbol{v}'. \tag{11.11}$$

$$\frac{\partial \rho'}{\partial t} + \rho_0\boldsymbol{\nabla}\cdot\boldsymbol{v}' = 0, \tag{11.12}$$

Next, we go through some examples of various degrees of interest.

11.2 Alfvén waves in the presence of a vertical magnetic field

We assume that all variables can be decomposed into plane waves in the form

$$v'_x = \hat{v}_x e^{i\boldsymbol{k}\cdot\boldsymbol{x}-i\omega t}, \tag{11.13}$$

$$v'_y = \hat{v}_y e^{i\boldsymbol{k}\cdot\boldsymbol{x}-i\omega t}, \tag{11.14}$$

$$v'_z = \hat{v}_z e^{i\boldsymbol{k}\cdot\boldsymbol{x}-i\omega t}, \tag{11.15}$$

$$B'_x = \hat{B}_x e^{i\boldsymbol{k}\cdot\boldsymbol{x}-i\omega t}, \tag{11.16}$$

$$B'_y = \hat{B}_y e^{i\boldsymbol{k}\cdot\boldsymbol{x}-i\omega t}, \tag{11.17}$$

$$B'_z = \hat{B}_z e^{i\boldsymbol{k}\cdot\boldsymbol{x}-i\omega t}, \tag{11.18}$$

$$\rho' = \hat{\rho} e^{i\boldsymbol{k}\cdot\boldsymbol{x}-i\omega t}, \tag{11.19}$$

which allows us to derive a set of algebraic equations. For the velocity components we have

$$-i\omega\hat{v}_x = -ik_x\hat{\rho}\frac{c_s^2}{\rho_0} - ik_x\hat{B}_z\frac{B_0}{\mu_0\rho_0} + ik_z\hat{B}_x\frac{B_0}{\mu_0\rho_0} \tag{11.20}$$

$$-i\omega\hat{v}_y = -ik_y\hat{\rho}\frac{c_s^2}{\rho_0} - ik_y\hat{B}_z\frac{B_0}{\mu_0\rho_0} + ik_z\hat{B}_y\frac{B_0}{\mu_0\rho_0} \tag{11.21}$$

$$-i\omega\hat{v}_z = -ik_z\hat{\rho}\frac{c_s^2}{\rho_0} \tag{11.22}$$

where the last two magnetic terms have canceled. The continuity equation is

$$-i\omega\hat{\rho} = -ik_x\rho_0\hat{u}_x - ik_y\rho_0\hat{u}_y - ik_z\rho_0\hat{u}_z \tag{11.23}$$

The magnetic equations are now for the form

$$-i\omega\hat{B}_x = +ik_zB_0\hat{u}_x \tag{11.24}$$

$$-i\omega\hat{B}_y = +ik_zB_0\hat{u}_y \tag{11.25}$$

$$-i\omega\hat{B}_z = -ik_xB_0\hat{u}_x - ik_yB_0\hat{u}_y \tag{11.26}$$

where, again, the ik_z terms have canceled in the last equation.

It is now convenient to write those equations in matrix form

$$\begin{pmatrix} -i\omega & 0 & 0 & ik_x\frac{c_s^2}{\rho_0} & -ik_z\frac{B_0}{\mu_0\rho_0} & 0 & +ik_x\frac{B_0}{\mu_0\rho_0} \\ 0 & -i\omega & 0 & ik_y\frac{c_s^2}{\rho_0} & 0 & -ik_z\frac{B_0}{\mu_0\rho_0} & +ik_y\frac{B_0}{\mu_0\rho_0} \\ 0 & 0 & -i\omega & ik_z\frac{c_s^2}{\rho_0} & 0 & 0 & 0 \\ ik_x\rho_0 & ik_y\rho_0 & ik_z\rho_0 & -i\omega & 0 & 0 & 0 \\ -ik_zB_0 & 0 & 0 & 0 & -i\omega & 0 & 0 \\ 0 & -ik_zB_0 & 0 & 0 & 0 & -i\omega & 0 \\ ik_xB_0 & ik_yB_0 & 0 & 0 & 0 & 0 & -i\omega \end{pmatrix} \begin{pmatrix} \hat{v}_x \\ \hat{v}_y \\ \hat{v}_z \\ \hat{\rho} \\ \hat{B}_x \\ \hat{B}_y \\ \hat{B}_z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}. \tag{11.27}$$

We won't consider this rather large system of equations at this point. Instead we want to observe what happens when we assume that the system is one-dimensional. Hence, we ignore the x and y components of the wave vector, so our matrix becomes

$$\begin{pmatrix} -i\omega & 0 & 0 & 0 & -ik_z\frac{B_0}{\mu_0\rho_0} & 0 & 0 \\ 0 & -i\omega & 0 & 0 & 0 & -ik_z\frac{B_0}{\mu_0\rho_0} & 0 \\ 0 & 0 & -i\omega & ik_z\frac{c_s^2}{\rho_0} & 0 & 0 & 0 \\ 0 & 0 & ik_z\rho_0 & -i\omega & 0 & 0 & 0 \\ -ik_zB_0 & 0 & 0 & 0 & -i\omega & 0 & 0 \\ 0 & -ik_zB_0 & 0 & 0 & 0 & -i\omega & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -i\omega \end{pmatrix} \begin{pmatrix} \hat{v}_x \\ \hat{v}_y \\ \hat{v}_z \\ \hat{\rho} \\ \hat{B}_x \\ \hat{B}_y \\ \hat{B}_z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}. \tag{11.28}$$

We notice that the last equation disappears altogether. We notice further that the third and fourth equations decouple from the remaining four equations. In order to have non-trivial solutions (where the hatted variables themselves don't vanish), we have to require that the determinant of the matrix vanishes. Because the matrix decouples into two matrices, M_1 and M_2, we have to require that either $\det M_1 = 0$ or $\det M_2 = 0$, where

$$M_1 = \begin{pmatrix} -i\omega & 0 & -ik_z\frac{B_0}{\mu_0\rho_0} & 0 \\ 0 & -i\omega & 0 & -ik_z\frac{B_0}{\mu_0\rho_0} \\ -ik_zB_0 & 0 & -i\omega & 0 \\ 0 & -ik_zB_0 & 0 & -i\omega \end{pmatrix}, \quad M_2 = \begin{pmatrix} -i\omega & ik_z\frac{c_s^2}{\rho_0} \\ ik_z\rho_0 & -i\omega \end{pmatrix} \tag{11.29}$$

There are two separate dispersion relations. From $\det M_2 = 0$ we have

$$\omega^2 = c_s^2 k_z^2. \tag{11.30}$$

The resulting dispersion relation from $\det M_2 = 0$ is

$$\omega^4 - 2v_A^2 k^2 \omega^2 + v_A^4 k^4 = 0. \tag{11.31}$$

This is a biquadratic equation, or a quadratic equation in ω^2. However, in this case it can also be written as

$$(\omega^2 - v_A^2 k^2)^2 = 0, \tag{11.32}$$

which simply means that there are Alfvén waves were either the x or the y components of the field are involved.

11.3 One-dimensional Alfvén waves revisited

Given that in one dimension the Alfvén waves simplify significantly we begin all over again, focusing attention immediately on the essential points. The

linearized, pressureless, ideal MHD equations in one dimension, in the presence of a vertical magnetic field, are

$$\frac{\partial v_x'}{\partial t} = \frac{B_0}{\mu_0 \rho_0} \frac{\partial B_x'}{\partial z}, \tag{11.33}$$

$$\frac{\partial B_x'}{\partial t} = B_0 \frac{\partial v_x'}{\partial z}, \tag{11.34}$$

Differentiating the first equation in time and inserting the second equation for $\partial B_x'/\partial t$ yields a wave equation, similar to the equation describing sound waves. In the more complicated situations described below it turns out to be easier however to take the solution to be in the form

$$v_x' = \hat{v}_x e^{ikz - i\omega t}, \tag{11.35}$$

$$B_x' = \hat{B}_x e^{ikz - i\omega t}, \tag{11.36}$$

which allows us to derive a set of two algebraic equations,

$$-i\omega \hat{v}_x - 2\Omega \hat{v}_y = \frac{B_0}{\rho_0 \mu_0} ik \hat{B}_x, \tag{11.37}$$

$$-i\omega \hat{B}_x = B_0 ik \hat{v}_x. \tag{11.38}$$

It is convenient to write those equations in matrix form

$$\begin{pmatrix} -i\omega & -ik\frac{B_0}{\mu_0\rho_0} \\ -ikB_0 & -i\omega \end{pmatrix} \begin{pmatrix} \hat{v}_x \\ \hat{B}_x \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}. \tag{11.39}$$

For a nontrivial solution the determinant of the governing matrix has to vanish

$$\det M \equiv \det \begin{pmatrix} -i\omega & -ik\frac{B_0}{\mu_0\rho_0} \\ -ikB_0 & -i\omega \end{pmatrix} = -\omega^2 + \frac{B_0^2}{\mu_0 \rho_0} k^2 = 0, \tag{11.40}$$

This leads to the dispersion relation $\omega = \omega(k)$ with

$$\omega = \pm v_A k, \tag{11.41}$$

where v_A is the Alfvén speed with $v_A^2 = B_0^2/(\rho_0 \mu_0)$.

Unlike sound waves, where the restoring force is the pressure gradient, for Alfvén waves the restoring force is the magnetic field. Another important difference is the fact that now the field has only components perpendicular to the direction of propagation. Here the direction of propagation is z, but $\boldsymbol{B}$ has only a nonvanishing component in the x direction. For sound waves, on the other hand, the velocity has only a component in the direction of propagation.

11.3.1 The effects of rotation

We now assume that we are in a rotating system of reference, where the rotation axis is the z axis, which is also the direction of the applied magnetic field, so $\boldsymbol{\Omega} = (0, 0, \Omega)$ and so the Coriolis force (per unit mass) is

$$\boldsymbol{F}_{\rm Cor} = -2\boldsymbol{\Omega} \times \boldsymbol{v} = \begin{pmatrix} 0 \\ 0 \\ -2\Omega \end{pmatrix} \times \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix} = \begin{pmatrix} +2\Omega v_y \\ -2\Omega v_x \\ 0 \end{pmatrix}. \tag{11.42}$$

Adding this term to the equations causes immediately some coupling to the y components of both the velocity and the magnetic field. The linearized, pressureless, ideal MHD equations in one dimension, in the presence of rotation, are then

$$\frac{\partial v'_x}{\partial t} - 2\Omega v'_y = \frac{B_0}{\mu_0 \rho_0} \frac{\partial B'_x}{\partial z}, \tag{11.43}$$

$$\frac{\partial v'_y}{\partial t} + 2\Omega v'_x = \frac{B_0}{\mu_0 \rho_0} \frac{\partial B'_y}{\partial z}, \tag{11.44}$$

$$\frac{\partial B'_x}{\partial t} = B_0 \frac{\partial v'_x}{\partial z}, \tag{11.45}$$

$$\frac{\partial B'_y}{\partial t} = B_0 \frac{\partial v'_y}{\partial z}. \tag{11.46}$$

As usual, we assume the solution to be of the form

$$v'_x = \hat{v}_x e^{ikz - i\omega t}, \tag{11.47}$$

$$v'_y = \hat{v}_y e^{ikz - i\omega t}, \tag{11.48}$$

$$B'_x = \hat{B}_x e^{ikz - i\omega t}, \tag{11.49}$$

$$B'_y = \hat{B}_y e^{ikz - i\omega t}, \tag{11.50}$$

which allows us to derive a set of algebraic equations.

$$-i\omega \hat{v}_x - 2\Omega \hat{v}_y = \frac{B_0}{\rho_0 \mu_0} ik \hat{B}_x, \tag{11.51}$$

$$-i\omega \hat{v}_y + 2\Omega \hat{v}_x = \frac{B_0}{\rho_0 \mu_0} ik \hat{B}_y, \tag{11.52}$$

$$-i\omega \hat{B}_x = B_0 ik \hat{v}_x, \tag{11.53}$$

$$-i\omega \hat{B}_y = B_0 ik \hat{v}_y. \tag{11.54}$$

It is convenient to write those equations in matrix form

$$\begin{pmatrix} -i\omega & -2\Omega & -ik\frac{B_0}{\mu_0\rho_0} & 0 \\ 2\Omega & -i\omega & 0 & -ik\frac{B_0}{\mu_0\rho_0} \\ -ikB_0 & 0 & -i\omega & 0 \\ 0 & -ikB_0 & 0 & -i\omega \end{pmatrix} \begin{pmatrix} \hat{v}_x \\ \hat{v}_y \\ \hat{B}_x \\ \hat{B}_y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}. \tag{11.55}$$

For a nontrivial solution the determinant of the governing matrix has to vanish

$$\det M \equiv \det \begin{pmatrix} -i\omega & -2\Omega & -ik\frac{B_0}{\mu_0\rho_0} & 0 \\ 2\Omega & -i\omega & 0 & -ik\frac{B_0}{\mu_0\rho_0} \\ -ikB_0 & 0 & -i\omega & 0 \\ 0 & -ikB_0 & 0 & -i\omega \end{pmatrix}. \tag{11.56}$$

To calculate the determinant we split it into sub-determinants:

$$\begin{aligned} \det M = -i\omega \det \begin{pmatrix} -i\omega & 0 & -ik\frac{B_0}{\mu_0\rho_0} \\ 0 & -i\omega & 0 \\ -ikB_0 & 0 & -i\omega \end{pmatrix} + 2\Omega \det \begin{pmatrix} 2\Omega & 0 & -ik\frac{B_0}{\mu_0\rho_0} \\ -ikB_0 & -i\omega & 0 \\ 0 & 0 & -i\omega \end{pmatrix} \\ -ik\frac{B_0}{\mu_0\rho_0} \det \begin{pmatrix} 2\Omega & -i\omega & -ik\frac{B_0}{\mu_0\rho_0} \\ -ikB_0 & 0 & 0 \\ 0 & -ikB_0 & -i\omega \end{pmatrix} \end{aligned} \tag{11.57}$$

Determinants of 3×3 matrices are easy to calculate, so

$$\begin{aligned} \det M = -i\omega \left[(-i\omega)^3 - (-i\omega)\left(-k^2\frac{B_0^2}{\mu_0\rho_0}\right)\right] + 2\Omega\left[2\Omega(-i\omega)^2\right] \\ -ik\frac{B_0}{\mu_0\rho_0}\left[-k^2\frac{B_0^2}{\mu_0\rho_0}(-ikB_0) - (-i\omega)^2(-ikB_0)\right]. \end{aligned} \tag{11.58}$$

Thus,

$$\det M = -\omega^2\left[-\omega^2 + k^2\frac{B_0^2}{\mu_0\rho_0}\right] - 4\Omega^2\omega^2 - k^2\frac{B_0^2}{\mu_0\rho_0}\left[-k^2\frac{B_0^2}{\mu_0\rho_0} + \omega^2\right], \tag{11.59}$$

or

$$\det M = \omega^4 - \omega^2 k^2 v_A^2 - 4\Omega^2\omega^2 + (k^2 v_A^2)^2 - k^2 v_A^2\omega^2, \tag{11.60}$$

where $v_A^2 = B_0^2/(\mu_0\rho_0)$ is the square of the Alfvén speed. Setting the determinant to zero leads to the dispersion relation

$$\omega^4 - \omega^2(2v_A^2k^2 + 4\Omega^2) + v_A^4k^4 = 0. \tag{11.61}$$

This is a biquadratic equation, or a quadratic equation in ω^2.

11.3.2 The effects of rotation and shear

The linearized, pressureless, ideal MHD equations in one dimension, in the presence of rotation and shear with a vertical magnetic field, are

$$\frac{\partial v_x'}{\partial t} - 2\Omega v_y' = \frac{B_0}{\mu_0\rho_0}\frac{\partial B_x'}{\partial z}, \tag{11.62}$$

$$\frac{\partial v_y'}{\partial t} + \tfrac{1}{2}\Omega v_x' = \frac{B_0}{\mu_0\rho_0}\frac{\partial B_y'}{\partial z}, \tag{11.63}$$

$$\frac{\partial B'_x}{\partial t} = B_0 \frac{\partial v'_x}{\partial z}, \tag{11.64}$$

$$\frac{\partial B'_y}{\partial t} = B_0 \frac{\partial v'_y}{\partial z} - \tfrac{3}{2}\Omega B'_x, \tag{11.65}$$

We assume the solution to be of the form

$$v'_x = \hat{v}_x e^{ikz - i\omega t}, \tag{11.66}$$

$$v'_y = \hat{v}_y e^{ikz - i\omega t}, \tag{11.67}$$

$$B'_x = \hat{B}_x e^{ikz - i\omega t}, \tag{11.68}$$

$$B'_y = \hat{B}_y e^{ikz - i\omega t}, \tag{11.69}$$

which allows us to derive a set of algebraic equations.

$$-i\omega\hat{v}_x - 2\Omega\hat{v}_y = \frac{B_0}{\rho_0\mu_0} ik\hat{B}_x, \tag{11.70}$$

$$-i\omega\hat{v}_y + \tfrac{1}{2}\Omega\hat{v}_x = \frac{B_0}{\rho_0\mu_0} ik\hat{B}_y, \tag{11.71}$$

$$-i\omega\hat{B}_x = B_0 ik\hat{v}_x, \tag{11.72}$$

$$-i\omega\hat{B}_y = B_0 ik\hat{v}_y - \tfrac{3}{2}\Omega\hat{B}_x. \tag{11.73}$$

It is convenient to write those equations in matrix form

$$\begin{pmatrix} -i\omega & -2\Omega & -ik\frac{B_0}{\mu_0\rho_0} & 0 \\ \frac{1}{2}\Omega & -i\omega & 0 & -ik\frac{B_0}{\mu_0\rho_0} \\ -ikB_0 & 0 & -i\omega & 0 \\ 0 & -ikB_0 & \frac{3}{2}\Omega & -i\omega \end{pmatrix} \begin{pmatrix} \hat{v}_x \\ \hat{v}_y \\ \hat{B}_x \\ \hat{B}_y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}. \tag{11.74}$$

For a nontrivial solution the determinant of the governing matrix has to vanish

$$\det M \equiv \det \begin{pmatrix} -i\omega & -2\Omega & -ik\frac{B_0}{\mu_0\rho_0} & 0 \\ \frac{1}{2}\Omega & -i\omega & 0 & -ik\frac{B_0}{\mu_0\rho_0} \\ -ikB_0 & 0 & -i\omega & 0 \\ 0 & -ikB_0 & \frac{3}{2}\Omega & -i\omega \end{pmatrix}. \tag{11.75}$$

To calculate the determinant we split it into sub-determinants:

$$\begin{aligned} \det M = -i\omega \det \begin{pmatrix} -i\omega & 0 & -ik\frac{B_0}{\mu_0\rho_0} \\ 0 & -i\omega & 0 \\ -ikB_0 & \frac{3}{2}\Omega & -i\omega \end{pmatrix} + 2\Omega \det \begin{pmatrix} \frac{1}{2}\Omega & 0 & -ik\frac{B_0}{\mu_0\rho_0} \\ -ikB_0 & -i\omega & 0 \\ 0 & \frac{3}{2}\Omega & -i\omega \end{pmatrix} \\ -ik\frac{B_0}{\mu_0\rho_0} \det \begin{pmatrix} \frac{1}{2}\Omega & -i\omega & -ik\frac{B_0}{\mu_0\rho_0} \\ -ikB_0 & 0 & 0 \\ 0 & -ikB_0 & -i\omega \end{pmatrix} \end{aligned} \tag{11.76}$$

Determinants of 3×3 matrices are easy to calculate, so

$$\det M = -i\omega \left[(-i\omega)^3 - (-i\omega)\left(-k^2\frac{B_0^2}{\mu_0\rho_0}\right)\right] + 2\Omega\left[\tfrac{1}{2}\Omega(-i\omega)^2 - k^2\frac{B_0^2}{\mu_0\rho_0}\tfrac{3}{2}\Omega\right] - ik\frac{B_0}{\mu_0\rho_0}\left[-k^2\frac{B_0^2}{\mu_0\rho_0}(-ikB_0) - (-i\omega)^2(-ikB_0)\right] \quad (11.77)$$

Thus,

$$\det M = -\omega^2\left[-\omega^2 + k^2\frac{B_0^2}{\mu_0\rho_0}\right] + 2\Omega\left[-\tfrac{1}{2}\Omega\omega^2 - k^2\frac{B_0^2}{\mu_0\rho_0}\tfrac{3}{2}\Omega\right] - k^2\frac{B_0^2}{\mu_0\rho_0}\left[-k^2\frac{B_0^2}{\mu_0\rho_0} + \omega^2\right], \quad (11.78)$$

or

$$\det M = \omega^4 - \omega^2k^2v_A^2 - \Omega^2\omega^2 - 3\Omega^2k^2v_A^2 + (k^2v_A^2)^2 - k^2v_A^2\omega^2, \quad (11.79)$$

where $v_A^2 = B_0^2/(\mu_0\rho_0)$ is the square of the Alfvén speed. Setting the determinant to zero leads to the dispersion relation

$$\omega^4 - \omega^2(2v_A^2k^2 + \Omega^2) + v_A^2k^2(v_A^2k^2 - 3\Omega^2) = 0. \quad (11.80)$$

This is a biquadratic equation, or a quadratic equation in ω^2. There are two solutions for ω^2, ω_1^2 and ω_2^2, say. We can therefore write the dispersion relation in the form

$$(\omega^2 - \omega_1^2)(\omega^2 - \omega_2^2) = 0, \quad (11.81)$$

or

$$\omega^4 - \omega^2(\omega_1^2 + \omega_2^2) + \omega_1^2\omega_2^2 = 0. \quad (11.82)$$

Comparing with equation (11.105) we see that one of the two solutions is negative, i.e. $\omega_1^2\omega_2^2 < 0$ when

$$k^2v_A^2 < 3\Omega^2. \quad (11.83)$$

In that case ω is plus or minus an imaginary number. Since the solution is proportional to $e^{-i\omega t}$ that means that the solution behaves like

$$e^{-i\omega t} = e^{\pm|\mathrm{Im}(\omega)|t}. \quad (11.84)$$

The solution for which the upper sign applies grows fastest. We thus say that the solution is *unstable*. Eventually the magnitude of the velocity and magnetic field perturbations will be so large that the linearized equations are no longer valid.

Note that we were able to draw conclusions about stability and instability even though we did not actually solve the dispersion relation. This happens quite often, especially with higher order polynomials. But in the present case an explicitly solution can easily be written down. It is

$$\omega_{1,2}^2 = v_A^2k^2 + \tfrac{1}{2}\Omega^2 \pm \sqrt{4v_A^2k^2\Omega^2 + \tfrac{1}{4}\Omega^4} \quad (11.85)$$

We see that the presence of shear has a destabilizing effect on the slow magnetosonic waves. This instability is also known as Balbus-Hawley instability and it has great relevance for causing turbulence in accretion discs. The maximum growth rate is reached when $v_A^2 k^2 = \frac{15}{16}\Omega^2$, and the corresponding value of ω^2 is then $-\omega^2 = \frac{9}{16}\Omega^2 \approx 0.56\Omega^2$, or $\mathrm{Im}\,\omega = \frac{3}{4}\Omega$.

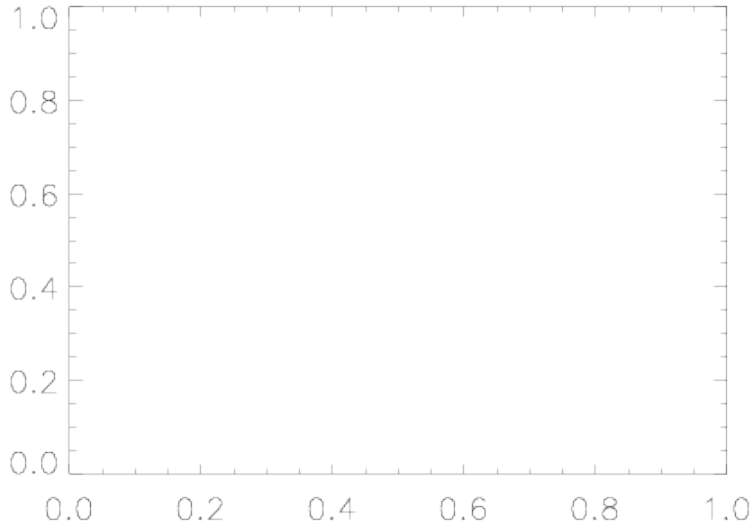


Figure 11.1: Dispersion relation for slow magnetosonic and Alfven waves. Solid lines denote Alfven waves. For them ω^2 is always positive. Dashed lines refer to slow magnetosonic waves. For not too large values of k they can become unstable ($\omega^2 < 0$, so ω becomes imaginary!).

11.3.3 Eigenfunction for the Balbus-Hawley instability

A simple way to get the eigenfunction of the Balbus-Hawley instability, of in fact of any linear eigenvalue problem, is to use the original equations starting with the most simple one. The simplest one is equation (11.72), because it involves just 2 terms. We may choose the coefficient in front of $\hat{v}_x$ to one, so

$$\hat{v}_x = 1 \tag{11.86}$$

Using equation (11.72) it follows that

$$\hat{B}_x = -\frac{ik}{i\omega} B_0 \tag{11.87}$$

Next we use equation (11.70) to find $\hat{v}_y$

$$\hat{v}_y = -\frac{i\omega + \frac{B_0}{\rho_0\mu_0} ik \left(-\frac{ik}{i\omega} B_0\right)}{2\Omega} \tag{11.88}$$

or

$$\hat{v}_y = \frac{\omega^2 - v_A^2 k^2}{2\Omega i\omega} \tag{11.89}$$

Finally, we need to calculate $\hat{B}_y$, but we have two equations still unused, (11.71) and (11.73). We use first equation (11.71) to obtain

$$\hat{B}_y = \frac{-i\omega\hat{v}_y + \frac{1}{2}\Omega\hat{v}_x}{ikv_A^2} B_0 \tag{11.90}$$

or, using the above results,

$$\hat{B}_y = \frac{-(\frac{\omega^2 - v_A^2 k^2}{2\Omega}) + \frac{1}{2}\Omega}{ikv_A^2} B_0 \tag{11.91}$$

so

$$\hat{B}_y = ikB_0 \frac{\omega^2 - v_A^2 k^2 - \Omega^2}{2\Omega k^2 v_A^2} \tag{11.92}$$

Finally, we can use the last equation, (11.73) to check everything, so we plug in the results obtained so far. We have

$$-i\omega\hat{B}_y = B_0 ik\hat{v}_y - \tfrac{3}{2}\Omega\hat{B}_x. \tag{11.93}$$

so

$$-i\omega(ikB_0)\frac{\omega^2 - v_A^2 k^2 - \Omega^2}{2\Omega k^2 v_A^2} = B_0 ik \frac{\omega^2 - v_A^2 k^2}{2\Omega i\omega} - \tfrac{3}{2}\Omega\left(-\frac{ik}{i\omega}B_0\right). \tag{11.94}$$

We cancel ikB_0 on both sides, and multiply by $i\omega$, so

$$\omega^2 \frac{\omega^2 - v_A^2 k^2 - \Omega^2}{2\Omega k^2 v_A^2} = \frac{\omega^2 - v_A^2 k^2}{2\Omega} + \tfrac{3}{2}\Omega. \tag{11.95}$$

Multiplying by $2\Omega k^2 v_A^2$ yields

$$\omega^2(\omega^2 - v_A^2 k^2 - \Omega^2) = (\omega^2 - v_A^2 k^2)k^2 v_A^2 + 3\Omega^2 k^2 v_A^2. \tag{11.96}$$

which gives the old biquadratic equation,

$$\omega^4 - \omega^2(2v_A^2 k^2 - \Omega^2) + (k^2 v_A^2 - 3\Omega^2)k^2 v_A^2 = 0, \tag{11.97}$$

so everything seems to have gone alright. Anyway, the eigenfunction is therefore

$$\begin{pmatrix} \hat{v}_x \\ \hat{v}_y \\ \hat{B}_x \\ \hat{B}_y \end{pmatrix} = \begin{pmatrix} 1 \\ \frac{\omega^2 - v_A^2 k^2}{2\Omega i\omega} \\ -\frac{ik}{i\omega}B_0 \\ ikB_0 \frac{\omega^2 - v_A^2 k^2 - \Omega^2}{2\Omega k^2 v_A^2} \end{pmatrix} \tag{11.98}$$

11.3.4 The effect of magnetic and ambipolar diffusion

For a nontrivial solution the determinant of the governing matrix has to vanish

$$\det M \equiv \det \begin{pmatrix} -i\omega & -2\Omega & -ik\frac{B_0}{\mu_0\rho_0} & 0 \\ \frac{1}{2}\Omega & -i\omega & 0 & -ik\frac{B_0}{\mu_0\rho_0} \\ -ikB_0 & 0 & -i\tilde{\omega} & 0 \\ 0 & -ikB_0 & \frac{3}{2}\Omega & -i\tilde{\omega} \end{pmatrix}. \tag{11.99}$$

where $-i\tilde{\omega} = -i\omega - \eta k^2$. To calculate the determinant we split it into sub-determinants:

$$\det M = -i\omega \det \begin{pmatrix} -i\omega & 0 & -ik\frac{B_0}{\mu_0\rho_0} \\ 0 & -i\tilde{\omega} & 0 \\ -ikB_0 & \frac{3}{2}\Omega & -i\tilde{\omega} \end{pmatrix} + 2\Omega \det \begin{pmatrix} \frac{1}{2}\Omega & 0 & -ik\frac{B_0}{\mu_0\rho_0} \\ -ikB_0 & -i\tilde{\omega} & 0 \\ 0 & \frac{3}{2}\Omega & -i\tilde{\omega} \end{pmatrix}$$
$$-ik\frac{B_0}{\mu_0\rho_0} \det \begin{pmatrix} \frac{1}{2}\Omega & -i\omega & -ik\frac{B_0}{\mu_0\rho_0} \\ -ikB_0 & 0 & 0 \\ 0 & -ikB_0 & -i\tilde{\omega} \end{pmatrix} \tag{11.100}$$

Determinants of 3×3 matrices are easy to calculate, so

$$\det M = -i\omega \left[(-i\omega)(-i\tilde{\omega})^2 - (-i\tilde{\omega})\left(-k^2\frac{B_0^2}{\mu_0\rho_0}\right)\right] + 2\Omega\left[\frac{1}{2}\Omega(-i\tilde{\omega})^2 - k^2\frac{B_0^2}{\mu_0\rho_0}\frac{3}{2}\Omega\right]$$
$$-ik\frac{B_0}{\mu_0\rho_0}\left[-k^2\frac{B_0^2}{\mu_0\rho_0}(-ikB_0) - (-i\omega)(-i\tilde{\omega})(-ikB_0)\right] \tag{11.101}$$

$$\det M = -i\omega \left[(-i\omega)(-i\tilde{\omega})^2 - (-i\tilde{\omega})\left(-k^2\frac{B_0^2}{\mu_0\rho_0}\right)\right] + 2\Omega\left[\frac{1}{2}\Omega(-i\tilde{\omega})^2 - k^2\frac{B_0^2}{\mu_0\rho_0}\frac{3}{2}\Omega\right]$$
$$-ik\frac{B_0}{\mu_0\rho_0}\left[-k^2\frac{B_0^2}{\mu_0\rho_0}(-ikB_0) - (-i\omega)(-i\tilde{\omega})(-ikB_0)\right] \tag{11.102}$$

Thus,

$$\det M = -\omega^2\left[-\tilde{\omega}^2 + \sigma k^2\frac{B_0^2}{\mu_0\rho_0}\right] + 2\Omega\left[-\frac{1}{2}\Omega\tilde{\omega}^2 - k^2\frac{B_0^2}{\mu_0\rho_0}\frac{3}{2}\Omega\right] - k^2\frac{B_0^2}{\mu_0\rho_0}\left[-k^2\frac{B_0^2}{\mu_0\rho_0} + \sigma\omega^2\right], \tag{11.103}$$

where $\sigma = \tilde{\omega}/\omega$ has been introduced as a short hand. This can be simplified to give

$$\det M = \omega^4 - \sigma\omega^2 k^2 v_A^2 - \Omega^2\omega^2 - 3\Omega^2 k^2 v_A^2 + (k^2 v_A^2)^2 - k^2 v_A^2 \sigma\omega^2, \tag{11.104}$$

where $v_A^2 = B_0^2/(\mu_0\rho_0)$ is the square of the Alfvén speed. Setting the determinant to zero leads to the dispersion relation

$$\omega^4 - \omega^2(2\sigma v_A^2 k^2 + \Omega^2) + v_A^2 k^2(v_A^2 k^2 - 3\Omega^2) = 0. \tag{11.105}$$

11.4 Nonaxisymmetric Balbus-Hawley instability

Due to shear, one expects a strong toroidal magnetic field in accretion discs. In that case ... $\boldsymbol{B}_0 = (0, B_0, 0)$, so

$$\boldsymbol{j} \times \boldsymbol{B}_0 = \boldsymbol{B}_0 \cdot \boldsymbol{\nabla} \boldsymbol{b} - \boldsymbol{\nabla}(\boldsymbol{B}_0 \cdot \boldsymbol{b}) = B_0(\partial_y \boldsymbol{b} - \boldsymbol{\nabla} b_y) = B_0 \begin{pmatrix} \partial_y & -\partial_x & 0 \\ 0 & 0 & 0 \\ 0 & -\partial_z & \partial_y \end{pmatrix} \begin{pmatrix} b_x \\ b_y \\ b_z \end{pmatrix} \tag{11.106}$$

$$\boldsymbol{\nabla} \times (\boldsymbol{u} \times \boldsymbol{B}_0) = \boldsymbol{B}_0 \cdot \boldsymbol{\nabla} \boldsymbol{u} - \boldsymbol{B}_0 \boldsymbol{\nabla} \cdot \boldsymbol{u} = B_0 \begin{pmatrix} \partial_y & 0 & 0 \\ -\partial_x & 0 & -\partial_z \\ 0 & 0 & \partial_y \end{pmatrix} \begin{pmatrix} u_x \\ u_y \\ u_z \end{pmatrix} \tag{11.107}$$

$$\boldsymbol{\nabla} \times (\boldsymbol{U}_0 \times \boldsymbol{b}) = -\boldsymbol{U}_0 \cdot \boldsymbol{\nabla} \boldsymbol{b} + \boldsymbol{b} \cdot \boldsymbol{\nabla} \boldsymbol{U}_0 = -Sx\partial_y \boldsymbol{b} + S \begin{pmatrix} 0 \\ b_x \\ 0 \end{pmatrix} \tag{11.108}$$

In the general nonaxisymmetric case we can solve the problem by letting k_x be time-dependent, ie $k_x = k_x(t)$, with

$$k_x(t) = k_x(0) + q\Omega t. \tag{11.109}$$

In that case the solutions are of the form $\boldsymbol{q} = \hat{\boldsymbol{q}}(t) \exp[\mathrm{i}\boldsymbol{k}(t) \cdot \boldsymbol{x}]$, and we are left with a system of *ordinary* differential equations,

$$\frac{\mathrm{d}\hat{\boldsymbol{q}}}{\mathrm{d}t} = -\mathrm{i}\hat{\mathsf{L}}\hat{\boldsymbol{q}} \tag{11.110}$$

with the matrix

$$\hat{\mathsf{A}} \equiv -\mathrm{i}\hat{\mathsf{L}}(t) = \begin{pmatrix} 0 & 2\Omega & 0 & \mathrm{i}k_y & -\mathrm{i}k_x(t) & 0 \\ -(2-q)\Omega & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\mathrm{i}k_z & \mathrm{i}k_y \\ \mathrm{i}k_y & 0 & 0 & 0 & 0 & 0 \\ -\mathrm{i}k_x(t) & 0 & -\mathrm{i}k_z & -q\Omega & 0 & 0 \\ 0 & 0 & \mathrm{i}k_y & 0 & 0 & 0 \end{pmatrix} \tag{11.111}$$

being explicitly time-dependent. Note that in the absence of shear, $q = 0$, the matrix A is *antihermitian*, ie $\mathsf{A}^*_{ij} = -\mathsf{A}_{ji}$. This means that

$$\sum_{i=1}^{6} q_i^* \mathsf{A}_{ij} q_j \quad \text{is purely imaginary} \tag{11.112}$$

If either $q = 0$ or $k_x = 0$, we have the dispersion relation $-\omega^2 D(\omega, \boldsymbol{k}) = 0$ with

$$D = \omega^4 - \omega^2(\boldsymbol{k}^2 + k_y^2 + \kappa^2) + k_y^2 \boldsymbol{k}^2 + \kappa^2(k_y^2 + k_z^2). \tag{11.113}$$

In the general case, however, if one ignores the time dependence of $k_x(t)$, one has

$$\mathrm{i}\omega D + (\omega^2 - k_y^2) q\Omega k_x k_y = 0. \tag{11.114}$$

If pressure forces are included, we have

$$\hat{\mathbf{A}} \equiv -\mathrm{i}\hat{\mathbf{L}}(t) = \begin{pmatrix} 0 & 2\Omega & 0 & \mathrm{i}k_y v_\mathrm{A} & -\mathrm{i}k_x(t) v_\mathrm{A} & 0 & -\mathrm{i}k_x c_\mathrm{s} \\ -(2-q)\Omega & 0 & 0 & 0 & 0 & 0 & -\mathrm{i}k_y c_\mathrm{s} \\ 0 & 0 & 0 & 0 & -\mathrm{i}k_z v_\mathrm{A} & \mathrm{i}k_y v_\mathrm{A} & -\mathrm{i}k_z c_\mathrm{s} \\ \mathrm{i}k_y v_\mathrm{A} & 0 & 0 & 0 & 0 & 0 & 0 \\ -\mathrm{i}k_x(t) v_\mathrm{A} & 0 & -\mathrm{i}k_z v_\mathrm{A} & -q\Omega & 0 & 0 & 0 \\ 0 & 0 & \mathrm{i}k_y v_\mathrm{A} & 0 & 0 & 0 & 0 \\ -\mathrm{i}k_x c_\mathrm{s} & -\mathrm{i}k_y c_\mathrm{s} & -\mathrm{i}k_z c_\mathrm{s} & 0 & 0 & 0 & 0 \end{pmatrix} \tag{11.115}$$

11.4.1 Alternative formulation

Ignore magnetic pressure gradient (assume that it is being balanced by the gas pressure gradient), so we have

$$\boldsymbol{B}_0 \cdot \boldsymbol{\nabla} \boldsymbol{b} = B_0 \partial_y \boldsymbol{b} = B_0 \begin{pmatrix} \partial_y & 0 & 0 \\ 0 & \partial_y & 0 \\ 0 & 0 & \partial_y \end{pmatrix} \begin{pmatrix} b_x \\ b_y \\ b_z \end{pmatrix} \tag{11.116}$$

$$\boldsymbol{B}_0 \cdot \boldsymbol{\nabla} \boldsymbol{u} = B_0 \begin{pmatrix} \partial_y & 0 & 0 \\ 0 & \partial_y & 0 \\ 0 & 0 & \partial_y \end{pmatrix} \begin{pmatrix} u_x \\ u_y \\ u_z \end{pmatrix} \tag{11.117}$$

so the full matrix is

$$\hat{\mathbf{A}} \equiv -\mathrm{i}\hat{\mathbf{L}}(t) = \begin{pmatrix} 0 & 2\Omega & 0 & \mathrm{i}k_y & 0 & 0 \\ -(2-q)\Omega & 0 & 0 & 0 & \mathrm{i}k_y & 0 \\ 0 & 0 & 0 & 0 & 0 & \mathrm{i}k_y \\ \mathrm{i}k_y & 0 & 0 & 0 & 0 & 0 \\ 0 & \mathrm{i}k_y & 0 & -q\Omega & 0 & 0 \\ 0 & 0 & \mathrm{i}k_y & 0 & 0 & 0 \end{pmatrix} \tag{11.118}$$

Corresponds to the dispersion relation

$$-\omega^6 + \omega^4 \left[3v_\mathrm{A}^2 k^2 + 2(2-q)\Omega^2\right] - \omega^2 v_\mathrm{A}^2 k^2 \left[3v_\mathrm{A}^2 k^2 + 4(1-q)\Omega^2\right] + v_\mathrm{A}^4 k^4 (v_\mathrm{A}^2 k^2 - 2q\Omega^2) = 0. \tag{11.119}$$

Has unstable eigenvalues for $3\Omega^2 > v_\mathrm{A}^2 k^2$, just like in the axisymmetric case.

11.5 Magnetosonic waves

We now consider the linearized MHD equations in vector notation,

$$-i\omega\hat{\boldsymbol{v}} = -i\boldsymbol{k}(c_s^2/\rho_0)\hat{\rho} - i\boldsymbol{k}\frac{\boldsymbol{B}_0 \cdot \hat{\boldsymbol{B}}}{\mu_0\rho_0} + \frac{i\boldsymbol{k}\cdot\boldsymbol{B}_0}{\mu_0\rho_0}\hat{\boldsymbol{B}}, \tag{11.120}$$

$$-i\omega\hat{\rho} = -i\boldsymbol{k}\cdot\hat{\boldsymbol{v}}, \tag{11.121}$$

$$-i\omega\hat{\boldsymbol{B}} = (i\boldsymbol{k}\cdot\boldsymbol{B}_0)\hat{\boldsymbol{v}} - (i\boldsymbol{k}\cdot\hat{\boldsymbol{v}})\boldsymbol{B}_0. \tag{11.122}$$

We multiply equation (11.120) by $-i\omega$ and substitute for $-i\omega\hat{\rho}$ and $-i\omega\hat{\boldsymbol{B}}$ using Eqs. (11.121) and (11.122), so

$$-\omega^2\hat{\boldsymbol{v}} = -i\boldsymbol{k}(c_s^2/\rho_0)(-i\boldsymbol{k}\cdot\hat{\boldsymbol{v}}) - \frac{i\boldsymbol{k}}{\mu_0\rho_0}\left[(i\boldsymbol{k}\cdot\boldsymbol{B}_0)(\hat{\boldsymbol{v}}\cdot\boldsymbol{B}_0) - (i\boldsymbol{k}\cdot\hat{\boldsymbol{v}})\boldsymbol{B}_0^2\right] + \frac{i\boldsymbol{k}\cdot\boldsymbol{B}_0}{\mu_0\rho_0}\left[(i\boldsymbol{k}\cdot\boldsymbol{B}_0)\hat{\boldsymbol{v}} - (i\boldsymbol{k}\cdot\hat{\boldsymbol{v}})\boldsymbol{B}_0\right], \tag{11.123}$$

or

$$(k^2v_A^2 - \omega^2)\hat{\boldsymbol{v}} = +\boldsymbol{k}\left[-(\boldsymbol{k}\cdot\hat{\boldsymbol{v}})(c_s^2 + v_A^2) + (\boldsymbol{k}\cdot\boldsymbol{B}_0)(\hat{\boldsymbol{v}}\cdot\boldsymbol{B}_0)v_A^2\right] + \hat{\boldsymbol{B}}_0(\boldsymbol{k}\cdot\boldsymbol{B}_0)(\boldsymbol{k}\cdot\hat{\boldsymbol{v}})v_A^2. \tag{11.124}$$

We now assume that the wave vector points in the x direction and that the magnetic field vector lies in the $x-y$ plane, i.e.

$$\boldsymbol{k} = \begin{pmatrix} k \\ 0 \\ 0 \end{pmatrix}, \quad \boldsymbol{B}_0 = \begin{pmatrix} \cos\psi \\ \sin\psi \\ 0 \end{pmatrix} B_0. \tag{11.125}$$

Therefore,

$$\hat{\boldsymbol{v}}\cdot\hat{\boldsymbol{B}}_0 = B_0\left(\hat{v}_x\cos\psi + \hat{v}_y\sin\psi\right), \quad \boldsymbol{k}\cdot\boldsymbol{B}_0 = kB_0\cos\psi, \quad \boldsymbol{k}\cdot\hat{\boldsymbol{v}} = k\hat{v}_x. \tag{11.126}$$

With this the z component of the right hand side vanishes, but because $\hat{v}_z$ does not generally vanish we have $k^2v_A^2 - \omega^2 = 0$, which gives already two possible solutions:

$$\omega = \pm kv_A. \tag{11.127}$$

For the remaining two components we have to satisfy the x and y components separately, i.e.

$$(k^2v_A^2 - \omega^2)\hat{v}_x = k^2\left[-\hat{v}_x(c_s^2 + v_A^2) + \cos\psi(\hat{v}_x\cos\psi + \hat{v}_y\sin\psi)\,v_A^2\right] + \cos\psi(\hat{v}_x\cos\psi)\hat{v}_x k^2 v_A^2. \tag{11.128}$$

$$(k^2v_A^2 - \omega^2)\hat{v}_y = \sin\psi\cos\psi\; v_A^2k^2\hat{v}_x \tag{11.129}$$

This is a system of two algebraic equations. Written in matrix form we have,

$$\begin{pmatrix} (k^2v_A^2 - \omega^2) + (c_s^2 + v_A^2)k^2 - 2\cos^2\psi\, v_A^2k^2 & \cos\psi\sin\psi\, v_A^2k^2 \\ -\cos\psi\sin\psi\; v_A^2k^2 & (k^2v_A^2 - \omega^2) \end{pmatrix} \begin{pmatrix} \hat{v}_x \\ \hat{v}_y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{11.130}$$

which leads to the dispersion relation

$$\left[(k^2v_A^2 - \omega^2) + (c_s^2 + v_A^2)k^2 - 2\cos^2\psi\, v_A^2k^2\right](k^2v_A^2 - \omega^2) + \cos^2\psi\sin^2\psi\, v_A^4k^4 = 0 \tag{11.131}$$

We note that there are altogether three wave types, slow and fast magnetosonic waves and Alfvén waves. The frequency depends on the values of k, the angle between the wave vector $\boldsymbol{k}$ and the magnetic field $\boldsymbol{B}_0$, and of course the values of the sound speed and the Alfén speed. The dependence of the frequency as a function of the angle between $\boldsymbol{k}$ and $\boldsymbol{B}_0$ is shown in Fig. 11.2 for different values of c_s and v_A. In the special case $\psi = 0$ we have

$$(k^2c_s^2 - \omega^2)(k^2v_A^2 - \omega^2) = 0. \tag{11.132}$$

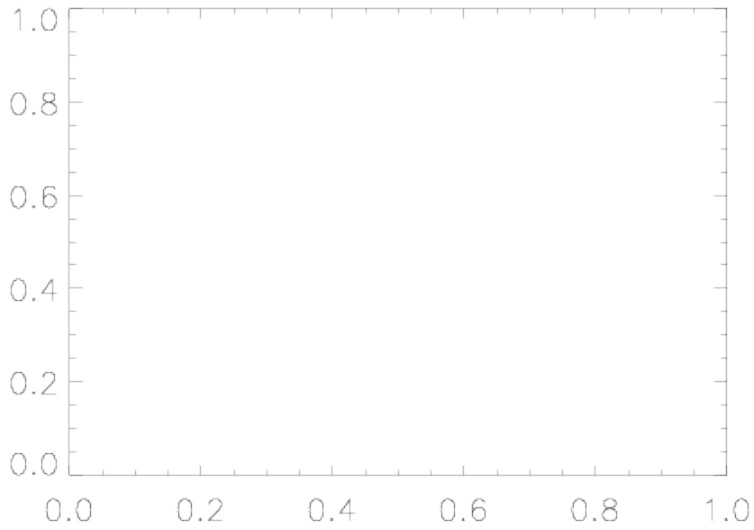


Figure 11.2: Dispersion relation for Alfven waves. Solid lines denote the fast magnetosonic waves, dashed lines Alfvén waves, and dotted lines slow magnetosonic waves.

11.5.1 Eigenfunctions

Here we only bother to give one particular eigenfunction for the case where $\boldsymbol{B}_0$ and $\boldsymbol{k}$ are aligned. In that case, Alfvén waves and fast magnetosonic waves are degenerate (see Fig. 11.2), so we are left with two modes.

$$v_x = \epsilon_{\rm s} c_{\rm s} \sin k(x - c_{\rm s} t), \quad \ln \rho = \epsilon_{\rm s} \sin k(x - c_{\rm s} t), \tag{11.133}$$

$$v_y = \epsilon_{\rm A} v_{\rm A} \sin k(x - v_{\rm A} t), \quad b_y = \epsilon_{\rm A} B_0 \sin k(x - v_{\rm A} t). \tag{11.134}$$

We note that the vector potential is $\boldsymbol{A} = (0, 0, A_z)$ with $A_z = \epsilon_{\rm A} k^{-1} B_0 \sin k(x - v_{\rm A} t)$.

Chapter 12

Dynamos

The dynamo mechanism provides a means of converting kinetic energy into magnetic energy. We shall focus on the astrophysically relevant case of a turbulent dynamo, as opposed to a laminar one. Laminar dynamos are easier to understand – and we shall discuss some simple examples – while turbulent dynamos have to be tackled via direct numerical simulations or by stochastic methods. Both will be discussed below.

12.1 Energetics

Important insight can be gained by considering the magnetic energy equation. By taking the dot product of equation (10.13) with $\boldsymbol{B}/(2\mu_0)$ and integrating over the volume V we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_V \frac{\boldsymbol{B}^2}{2\mu_0}\,\mathrm{d}V = -\int_V \boldsymbol{U}\cdot(\boldsymbol{J}\times\boldsymbol{B})\,\mathrm{d}V - \int_V \frac{\boldsymbol{J}^2}{\sigma}\,\mathrm{d}V - \oint_{\partial V}\frac{\boldsymbol{E}\times\boldsymbol{B}}{\mu_0}\,\mathrm{d}\boldsymbol{S}. \quad (12.1)$$

This equation shows that the magnetic energy can be increased by doing work against the Lorentz force, provided this terms exceeds resistive losses (second term) or losses through the surface (Poynting flux). Likewise, by taking the dot product of equation (2.7) with $\rho\boldsymbol{U}$ and integrating one arrives at the kinetic energy equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_V \tfrac{1}{2}\rho\boldsymbol{U}^2\,\mathrm{d}V = +\int_V p\boldsymbol{\nabla}\cdot\boldsymbol{U}\,\mathrm{d}V + \int_V \boldsymbol{U}\cdot(\boldsymbol{J}\times\boldsymbol{B})\,\mathrm{d}V + \int_V \rho\boldsymbol{U}\cdot\boldsymbol{g}\,\mathrm{d}V - \int_V 2\nu\rho\mathsf{S}^2\,\mathrm{d}V, \quad (12.2)$$

where $\mathsf{S}_{ij} = \frac{1}{2}(u_{i,j}+u_{j,i}) - \frac{1}{3}\delta_{ij}u_{k,k}$ is the rate of strain tensor, and commas denote derivatives. In deriving equation (12.2) we have assumed stress-free boundary conditions, so there are no surface terms and no kinetic energy is lost through the boundaries. Equations (12.1) and (12.2) show that the generation

Figure 12.1: Energy budget in a local accretion disc simulation where the turbulence is maintained by the Balbus-Hawley instability. The numbers on the arrows indicate the approximate energy conversion rates in units of $\Omega E_{\rm M}$, where Ω is the angular velocity and $E_{\rm M}$ is the steady state value of the magnetic energy.

of magnetic energy goes at the expense of kinetic energy, without loss of net energy.

In many astrophysical settings one can distinguish four different energy reservoirs that are involved in the dynamo process: magnetic, kinetic, thermal, and potential energy. In accretion discs the magnetic energy comes ultimately from potential energy which is first converted into kinetic energy. This is only possible by getting rid of angular momentum via Reynolds and/or Maxwell stresses. Half of the potential energy goes into orbital kinetic energy and the other half goes into turbulent kinetic energy which is then dissipated into heat and radiation. This requires turbulence to produce small enough length scales so that enough kinetic energy can indeed be dissipated on a dynamical time scale. This turbulence is most likely driven by the Balbus-Hawley (or magneto-rotational) instability. In Fig. 12.1 we show the energy diagram from a local simulation of the Balbus-Hawley instability. Here, the magnetic field necessary for the instability is maintained by a dynamo process. Most of the turbulent energy is dissipated by Joule heating. The magnetic energy typically exceeds the kinetic energy by a factor of about 3 or more, but is below the thermal energy by a factor of about 10–20.

In the case of solar convection the energy for the dynamo comes ultimately from the nuclear reactions in the center of the star. This acts as a source of thermal energy which gets converted into kinetic energy via the convection instability. The corresponding energy diagram for this case is shown in Fig. 12.2. Potential energy does not contribute directly: it only contributes through rearranging the mean density stratification.

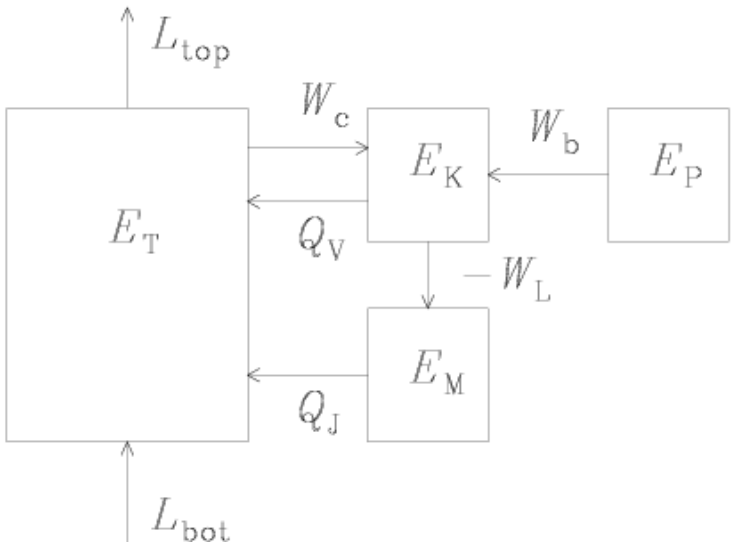


Figure 12.2: Energy budget in a local convection simulation. The dynamo is convectively driven by the luminosity entering from below, giving rise to convection via work done by adiabatic compression, $W_{\rm c} \equiv \int p\boldsymbol{\nabla}\cdot\boldsymbol{u}\,{\rm d}V$, and through work done against the Lorentz force, $W_{\rm L} = \int \boldsymbol{u}\cdot(\boldsymbol{J}\times\boldsymbol{B})\,{\rm d}V$. Energy is being fed back from magnetic and kinetic energy to thermal energy via Joule and viscous heating, $Q_{\rm J}$ and $Q_{\rm v}$. Some of the kinetic energy is constantly being exchanged with potential energy $E_{\rm P}$ via $W_{\rm b} = \int \rho\boldsymbol{g}\cdot\boldsymbol{u}\,{\rm d}V$

12.2 Kinematic dynamos

The onset of dynamo action can be studied in the linear approximation, i.e. the velocity field is assumed to be given (kinematic problem). There is in general a critical value of the magnetic Reynolds number above which the magnetic field grows exponentially. A lot of work has been devoted to the question of whether the growth rate can remain finite in the limit $R_{\rm m} \to \infty$ (the so-called fast dynamo). Fast dynamos are physically meaningful only until nonlinear effects begin to modify the flow to limit further growth of the field.

In the following we consider two simple examples of a dynamo. Both are slow dynamos, i.e. magnetic diffusion is crucial for the operation of the dynamo. We also discuss the stretch-twist-fold dynamo as a qualitative example of what is possibly a fast dynamo.

12.2.1 The Herzenberg dynamo

In the wake of Cowling's antidynamo theorem[1] the Herzenberg dynamo played an important role as an early example of a dynamo where the existence of excited solutions could be proven rigorously. The Herzenberg dynamo does not attempt to model an astrophysical dynamo. Instead, it was complementary to some of the less mathematical and more phenomenological models at the time, such as Parker's migratory dynamo as well as the observational model of Babcock, and the semi-observational model of Leighton, all of which were specifically designed to describe the solar cycle.

The Herzenberg dynamo is based on the mutual interaction of the magnetic fields produced by two spinning spheres in a conducting medium. In its simplest variant, the axes of the two spheres lie in planes that are parallel to each other, but the axes have an angle φ to each other; see Fig. 12.3, which shows the field vectors from a numerical simulation of the Herzenberg dynamo.

Dynamo action is possible unless the angle φ is exactly 0°, 90°, or 180°. For $90^\circ < \varphi < 180^\circ$ nonoscillatory dynamo action is possible. In the limit where the radius of the spheres, a, is small compared with their separation d, one can expand the field locally in terms of multipoles to lowest order. Defining a magnetic Reynolds number as $R_{\rm m} = \omega d^2/\eta$, where ω is the spin frequency of each of the spheres, the critical magnetic Reynolds number for dynamo action, $R_{\rm crit}$, is found to be

$$R_{\rm crit}^{-2} = -\frac{1}{4800}\left(\frac{a}{d}\right)^6 \sin^2\varphi\cos\varphi \quad (\text{for } 90^\circ < \varphi < 180^\circ), \tag{12.3}$$

which shows that the smallest value of $R_{\rm crit}$ is reached for $\varphi \approx 125^\circ$. Critical magnetic Reynolds numbers are several hundreds, but they would be only around ten if we were to redefine the magnetic Reynolds number based on some typical wavenumber. The dynamo works on the principle that each sphere winds up its ambient field, creates thereby a strong toroidal field around itself, but because there is an angle between the two spheres the toroidal field of one sphere acts as a poloidal field for the other sphere. For the toroidal field of each sphere to propagate to the other sphere, a non-zero diffusion is necessary, hence making this dynamo a slow dynamo.

Already back in the sixties, the idea of the Herzenberg dynamo has been verified experimentally using two conducting cylinders embedded in a solid block

[1]Larmor prosed in 1919 that the solar field might be generated by a self-excited dynamo. However, in 1933 Cowling published his anti-dynamo theorem, which states that two-dimensional (axisymmetric) magnetic fields cannot be sustained by dynamo action. Larmor (1934) responds to Cowling (1933) with the words "The view that I advanced briefly and tentatively long ago, which has come to be referred to as, perhaps too precisely, the self-exciting dynamo analogy, is still, so far as I know, the only foundation on which a gaseous body such as the Sun could possess a magnetic field: so that if it is demolished there could be no explanation of the Sun's magnetism even remotely in sight."

Cowling, T. G.: 1933, "The magnetic field of sunspots," *Mon. Not. Roy. Astron. Soc.* **94**, 39–48

Larmor, J.: 1934, "The magnetic field of sunspots," *Mon. Not. Roy. Astron. Soc.* **94**, 469–471

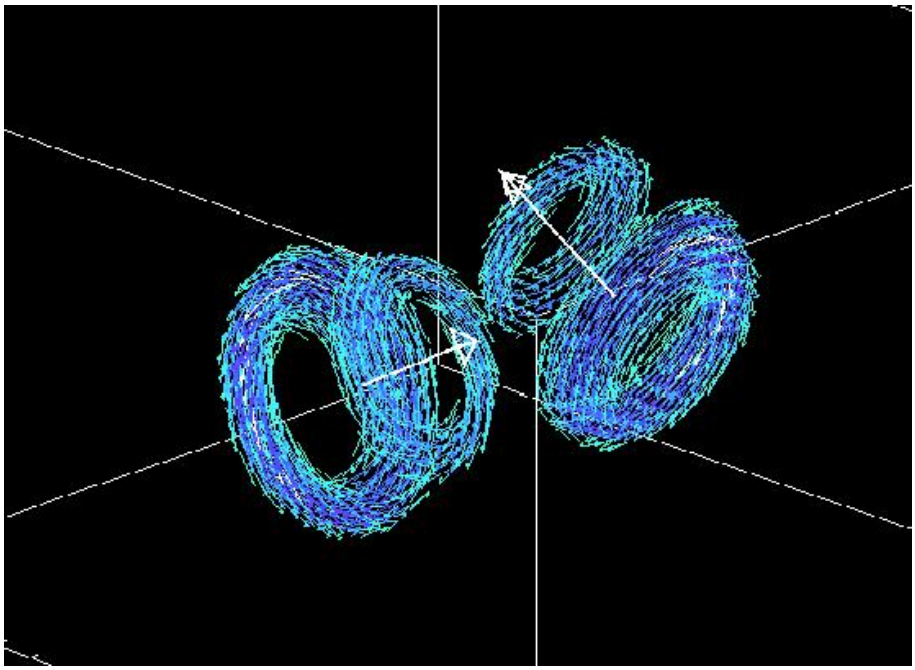


Figure 12.3: Three-dimensional visualization of the magnetic field geometry of the Herzenberg dynamo. $\boldsymbol{B}$-vectors are shown when their length exceeds about 25% of the maximum value

of the same material. The cylinders were in electric contact with the block through a thin lubricating film of mercury.

The asymptotic theory of Herzenberg assumed that $a/d \ll 1$; for excellent reviews of the Herzenberg dynamo see. Using numerical simulations it has been shown that equation (12.3) remains reasonably accurate even when $a/d \approx 1$. These simulations also show that in the range $0^\circ < \varphi < 90^\circ$ dynamo action is still possible, but the solutions are no longer steady but oscillatory. In the early papers, only steady solutions were sought, which is the reason why no solutions were originally found for $0^\circ < \varphi < 90^\circ$.

12.2.2 The Roberts flow dynamo

In the early years of dynamo theory most examples were constructed and motivated based on what seems physically possible and plausible. An important element of astrophysical dynamos is that the flow is bounded in space and that the magnetic field extends to infinity. Later, and especially in recent years, these restrictions were relaxed in may approaches. One of the first examples is the G. O. Roberts dynamo. The flow depends on only two coordinates, $\boldsymbol{U} = \boldsymbol{U}(x, y)$,

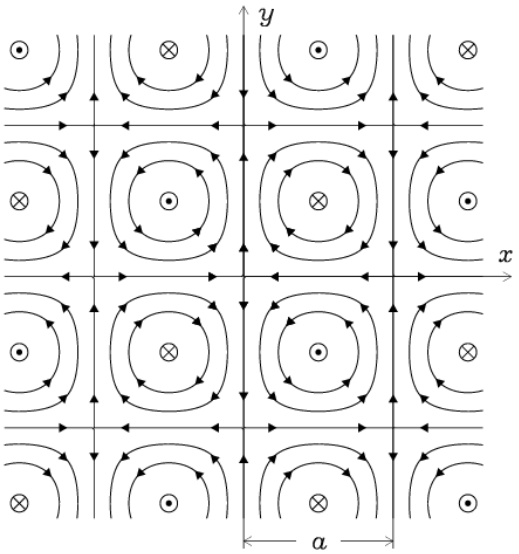


Figure 12.4: Roberts flow pattern with periodicity $2a$, corresponding to equation (12.4).

and can be written in the form

$$\boldsymbol{U}(x,y) = \sqrt{2}\,\boldsymbol{\nabla}\times(\varphi\boldsymbol{z}) + \boldsymbol{\nabla}\times\boldsymbol{\nabla}\times(\varphi\boldsymbol{z}), \tag{12.4}$$

with the stream function $\varphi = (U_0/k_\mathrm{f})\sin k_x x \sin k_y y$, where $k_x = k_y = \pi/a$; see Fig. 12.4. This flow is fully helical with $\boldsymbol{W} = k_\mathrm{f}\boldsymbol{U}$, were $k_\mathrm{f}^2 = k_x^2 + k_y^2$ and $\boldsymbol{W} = \boldsymbol{\nabla}\times\boldsymbol{U}$. While the flow is only two-dimensional (in the sense that $\boldsymbol{U}$ is a function only of x and y), the magnetic field must be three-dimensional for all growing solutions (dynamo effect). The field must therefore also depend on z.

The governing equations are homogeneous with coefficients that are independent of z and t. The solutions of the kinetic problem can therefore be written in the form

$$\boldsymbol{B}(x,y,z,t) = \mathrm{Re}\left[\hat{\boldsymbol{B}}_{k_z}(x,y)\exp(\mathrm{i}k_z z + \lambda t)\right], \tag{12.5}$$

where $\hat{\boldsymbol{B}}_{k_z}$ is the eigenfunction, which is obtained by solving the eigenvalue problem

$$\lambda\hat{\boldsymbol{A}}_{k_z} = \boldsymbol{U}\times\hat{\boldsymbol{B}}_{k_z} + \eta\left(\nabla^2 - k_z^2\right)\hat{\boldsymbol{A}}_{k_z}, \tag{12.6}$$

where $\hat{\boldsymbol{B}}_{k_z} = \boldsymbol{\nabla}\times\hat{\boldsymbol{A}}_{k_z} + \mathrm{i}\boldsymbol{k}_z\times\hat{\boldsymbol{A}}_{k_z}$ is expressed in terms of $\hat{\boldsymbol{A}}_{k_z}$, which is a mixed representation of the vector potential; in real space the vector potential would be $\mathrm{Re}\left[\hat{\boldsymbol{A}}_{k_z}(x,y)\exp(\mathrm{i}k_z z + \lambda t)\right]$. For $k_z = k_x = k_y \approx 0.71 k_\mathrm{f}$, the marginal state ($\lambda = 0$) is reached when $R_\mathrm{crit} \equiv U_0/(\eta k_\mathrm{f})_\mathrm{crit} \approx 3.90$. The larger the domain in

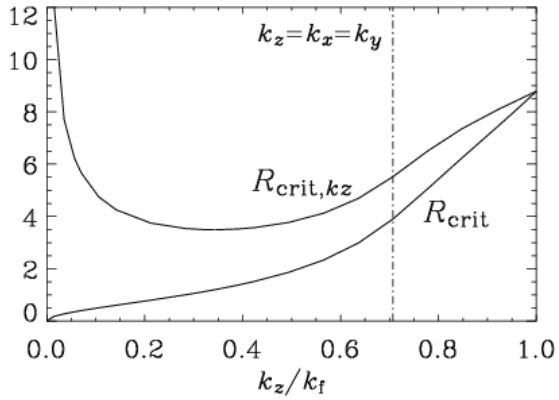


Figure 12.5: Critical magnetic Reynolds number, $R_{\rm crit} = U_0/(\eta k_{\rm f})_{\rm crit}$, for the Roberts flow as a function of $k_z/k_{\rm f}$, where $k_{\rm f}^2 = k_x^2 + k_y^2$. The critical magnetic Reynolds number based on k_z, $R_{{\rm crit},kz} = U_0/(\eta k_z)_{\rm crit}$, has a minimum at $k_z \approx 0.34 k_{\rm f} \approx 0.48 k_x$. The case of a squared domain with $k_x = k_y = k_z$, i.e. $k_z/k_{\rm f} = 1/\sqrt{2}$, is indicated by the vertical dash-dotted line.

the z-direction, the lower is the critical magnetic Reynolds number. However, the critical magnetic Reynolds number based on k_z, $R_{{\rm crit},kz} = U_0/(\eta k_z)_{\rm crit}$, has a minimum at $k_z \approx 0.34 k_{\rm f} \approx 0.48 k_x$ with $R_{{\rm crit},kz} \approx 3.49$; cf. Fig. 12.5.

The horizontally averaged eigenfunction is $\overline{\hat{\boldsymbol{B}}}_{k_z} = ({\rm i}, 1, 0)$, corresponding to a Beltrami wave, which has maximum magnetic helicity with a sign that is opposite to that of the flow. In the present case, the kinetic helicity of the flow is positive, so the magnetic and current helicities of the mean field are negative.

The significance of this solution is two-fold. On the one hand, this dynamo is the prototype of any fully helical dynamo capable of generating a large scale field. On the other hand, it is a simple model of the Karlsruhe dynamo experiment where a similar flow of liquid sodium is generated by an arrangement of pipes with internal 'spin generators' making the flow helical. It is also an example of a flow where the generation of the magnetic field can be described in terms of mean field electrodynamics.

Unlike the original Roberts flow dynamo, the flow in the Karlsruhe dynamo experiment is bounded and embedded in free space. Within the dynamo domain, the mean field, $\overline{\hat{\boldsymbol{B}}}_{k_z} = ({\rm i}, 1, 0)$, has only (x, y)-components. The field lines must close outside the dynamo domain, giving therefore rise to a dipole lying in the (x, y)-plane. Similar fields have long been predicted for rapidly rotating stars.

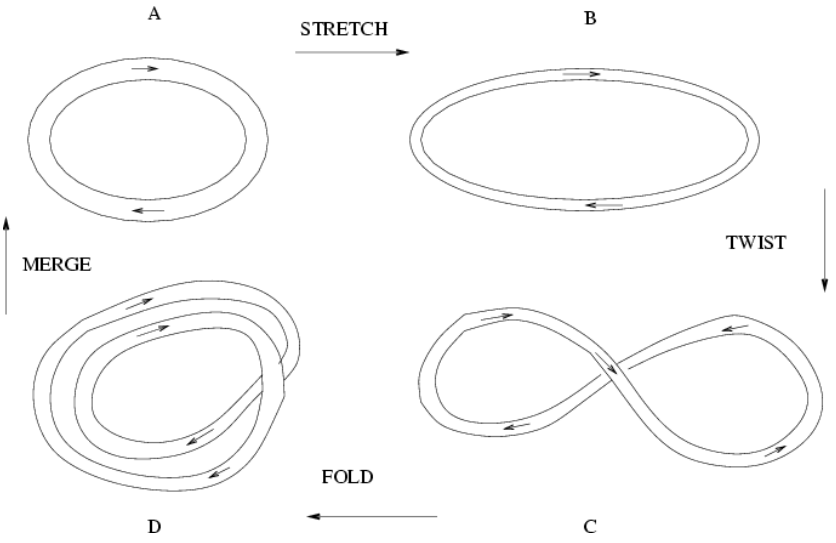


Figure 12.6: A schematic illustration of the stretch-twist-fold-merge dynamo.

12.3 Fast dynamos: the stretch-twist-fold picture

An elegant heuristic dynamo model illustrating the possibility of fast dynamos is what is often referred to as the Zeldovich 'stretch-twist-fold' (STF) dynamo (see Fig. 12.6). We briefly outline it here, as it illustrates nicely several features of more realistic dynamos.

The dynamo algorithm starts with first stretching a closed flux rope to twice its length preserving its volume, as in an incompressible flow (A→B in Fig. 12.6). The rope cross-section then decreases by factor two, and because of flux freezing the magnetic field doubles. In the next step, the rope is twisted into a figure eight (B→C in Fig. 12.6) and then folded (C→D in Fig. 12.6) so that now there are two loops, whose fields now point in the same direction and together occupy a similar volume as the original flux loop. The flux through this volume has now doubled. The last important step consists of merging the two loops into one (D→A in Fig. 12.6), through small diffusive effects. This is important in order that the new arrangement cannot easily undo itself and the whole process becomes irreversible. The newly merged loops now become topologically the same as the original single loop, but now with the field strength scaled up by factor 2.

Repeating the algorithm n times, leads to the field in the flux loop growing by factor 2^n, or at a growth rate $\sim \ln 2/T$ where T is the time for the STF steps. This makes the dynamo potentially a fast dynamo, whose growth rate does not decrease with decreasing resistivity. Also note that the flux through

a fixed 'Eulerian surface' grows exponentially, although the flux through any Lagrangian surface is nearly frozen; as it should be for small diffusivities.

The STF picture illustrates several other features: first we see that shear is needed to amplify the field at step A→B. However, without the twist part of the cycle, the field in the folded loop would cancel rather than add coherently. To twist the loop the motions need to leave the plane and go into the third dimension; this also means that field components perpendicular to the loop are generated, albeit being strong only temporarily during the twist part of the cycle. The source for the magnetic energy is the kinetic energy involved in the STF motions.

Most discussions of the STF dynamo assume implicitly that the last step of merging the twisted loops can be done at any time, and that the dynamo growth rate is not limited by this last step. This may well be true when the fields in the flux rope are not strong enough to affect the motions, that is, in the kinematic regime. However as the field becomes stronger, and if the merging process is slow, the Lorentz forces due to the small scale kinks and twists will gain in importance compared with the external forces associated with the driving of the loop as a whole. This may then limit the efficiency of the dynamo.

In this context one more feature deserves mentioning: if in the STF cycle one twists clockwise and folds, or twists counter-clockwise and folds one will still increase the field in the flux rope coherently. However, one would introduce opposite sense of writhe in these two cases, and so opposite internal twists. So, although the twist part of the cycle is important for the mechanism discussed here, the sense of twist can be random and does not require net helicity. This is analogous to a case when there is really only a small scale dynamo, but one that requires finite kinetic helicity density locally. We should point out, however, that numerical simulations have shown that dynamos work and are potentially independent of magnetic Reynolds number even if the flow has zero kinetic helicity density everywhere.

If the twisted loops can be made to merge efficiently, the saturation of the STF dynamo would probably proceed differently. For example, the field in the loop may become too strong to be stretched and twisted, due to magnetic curvature forces. Another interesting way of saturation is that the incompressibility assumed for the motions may break down; as one stretches the flux loop the field pressure resists the decrease in the loop cross-section, and so the fluid density in the loop tends to decrease as one attempts to make the loop longer. (Note that it is B/ρ which has to increase during stretching.) The STF picture has inspired considerable work on various mathematical features of fast dynamos and some of this work can be found in the book by Childress and Gilbert which in fact has STF in its title!

12.4 Fast ABC-flow dynamos

ABC flows are solenoidal and fully helical with a velocity field given by

$$\boldsymbol{U} = \begin{pmatrix} C\sin kz + B\cos ky \\ A\sin kx + C\cos kz \\ B\sin ky + A\cos kx \end{pmatrix}. \tag{12.7}$$

When A, B, and C are all different from zero, the flow is no longer integrable and has chaotic streamlines. There is numerical evidence that such flows act as fast dynamos. The magnetic field has very small net magnetic helicity. This is a general property of any dynamo in the kinematic regime and follows from magnetic helicity conservation. Even in a nonlinear formulation of the ABC flow dynamo problem, where the flow is driven by a forcing function similar to equation (12.7) the net magnetic helicity remains unimportant. This is however not surprising, because the development of net magnetic helicity requires sufficient scale separation, i.e. the wavenumber of the flow must be large compared with the smallest wavenumber in the box ($k = k_1$). If this is not the case, helical MHD turbulence behaves similarly to nonhelical turbulence. A significant scale separation also weakens the symmetries associated with the flow and the field, and leads to a larger kinematic growth rate, more compatible with the turnover time scale.

Most of the recent work on nonlinear ABC flow dynamos has focused on the case with small scale separation and, in particular, on the initial growth and possible saturation mechanisms. In the kinematic regime, these authors find a near balance between Lorentz work and Joule dissipation. The balance originates primarily from small volumes where the strong magnetic flux structures are concentrated. The net growth of the magnetic energy comes about through stretching and folding of relatively weak field which occupies most of the volume. The mechanism for saturation could involve achieving a local pressure balance in these strong field regions.

Chapter 13

Solar cycle

In this section we discuss properties of magnetic fields observed in various astrophysical settings. We focus specifically on aspects that are believed to be important for nonlinear dynamo theory and its connection with magnetic helicity. We begin with a discussion of the solar magnetic field, which consists of small scale and large scale components. The typical length scale associated with the large scale field is the width of the toroidal flux belts with the same polarity which is around 30° in latitude, corresponding to about 300 Mm (1 Mm = 1000 km). The pressure scale height at the bottom of the convection zone is about 50 Mm, and all scales shorter than that may be associated with the small scale field.

The theory of the large scale component has been most puzzling, while the small scale field could always be explained by turbulence and convection shredding and concentrating the field into isolated flux bundles. The simultaneous involvement of a so-called small scale dynamo may provide another source for the small scale field, which needs to be addressed. We begin by outlining the observational evidence for large scale fields in the sun and in stars, and discuss then the evidence for magnetic fields in accretion discs and galaxies, as well as galaxy clusters.

13.0.1 Solar and stellar magnetic fields

The sun has a magnetic field that manifests itself in sunspots through Zeeman splitting of spectral lines. It has long been known that the sunspot number varies cyclically with a period between 7 and 17 years. The longitudinally averaged component of the radial magnetic field of the sun shows a markedly regular spatio-temporal pattern where the radial magnetic field alternates in time over the 11 year cycle and also changes sign across the equator (Fig. 13.1). One can also see indications of a migration of the field from mid latitudes toward the equator and the poles. This migration is also well seen in a sunspot diagram, which is also called a butterfly diagram, because the pattern formed by the positions of sunspots in time and latitude looks like a sequence of butterflies

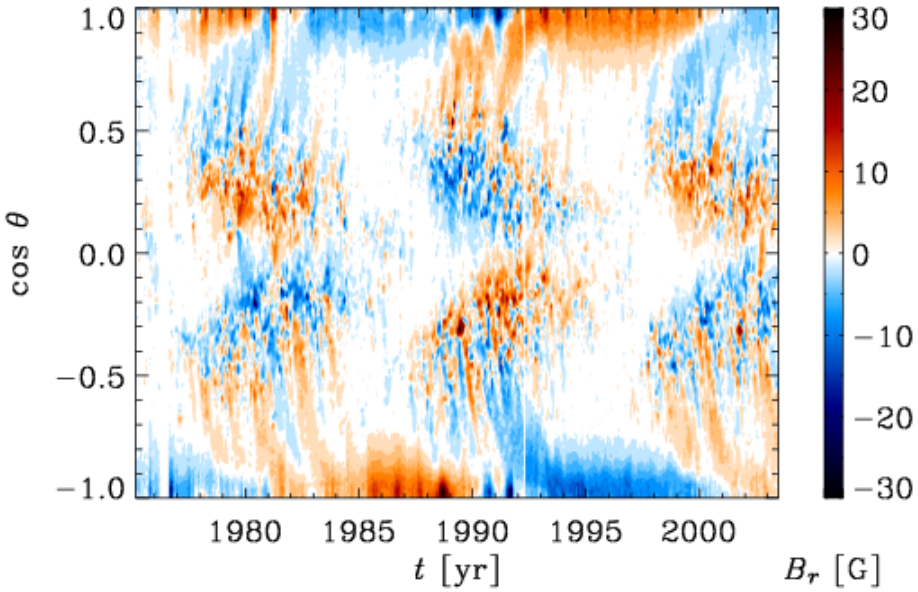


Figure 13.1: Longitudinally averaged radial component of the observed solar magnetic field as a function of cos(colatitude) and time. Dark (blue) shades denote negative values and light (yellow) shades denote positive values. Note the sign changes both in time and across the equator (courtesy of R. Knaack).

lined up along the equator (Fig. 13.2).

At the solar surface the azimuthally averaged radial field is only a few gauss ($1\,\mathrm{G} = 10^{-4}\,\mathrm{Tesla}$). This is rather weak compared with the peak magnetic field in sunspots of about 2 kG. In the bulk of the convection zone, because of differential rotation, the magnetic field is believed to point mostly in the azimuthal direction, and it is probably much larger near the bottom of the convection zone due to an effect known as downward pumping.

Estimates of the field strength in the deeper convection zone

In the bulk of the solar convection zone the thermal energy transport is reasonably well described by mixing length theory. This theory yields a rough estimate for the turbulent rms velocity which is around $u_{\rm rms} = 20\,\mathrm{m\,s^{-1}}$ near the bottom of the solar convection zone. With a density of about $\rho = 0.2\,\mathrm{g\,cm^{-3}}$ this corresponds to an equipartition field strength of about 3 kG. (The equipartition field strength is here defined as $B_{\rm eq} = \sqrt{\mu_0\rho}\,u_{\rm rms}$, where μ_0 is the magnetic permeability.)

A similar estimate is obtained by considering the total (unsigned) magnetic flux that emerges at the surface during one cycle. This argument is dubious,

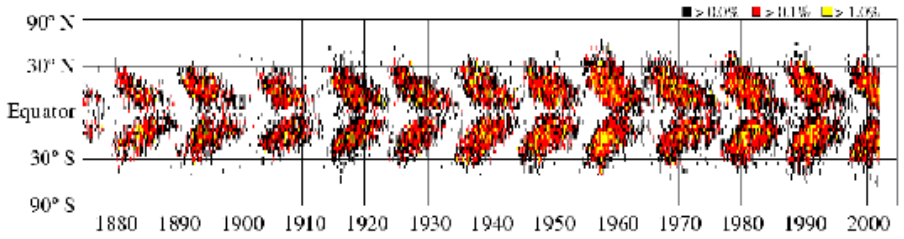


Figure 13.2: Solar butterfly diagram showing the sunspot number in a space-time diagram. Note the migration of sunspot activity from mid-latitudes toward the equator (courtesy of D. N. Hathaway).

because one has to make an assumption about how many times the flux tubes in the sun have emerged at the solar surface. Nevertheless, the notion of magnetic flux (and especially unsigned flux) is rather popular in solar physics, because this quantity is readily accessible from solar magnetograms. The total unsigned magnetic flux is roughly estimated to be 10^{24} Mx. Distributed over a meridional cross-section of about 500 Mm in the latitudinal direction and about 50 Mm in radius (i.e. the lower quarter of the convection zone) yields a mean field of about 4 kG, which is in fair agreement with the equipartition estimate above.

Another type of estimate concerns *not* the mean field but rather the peak magnetic field in the strong flux tubes. Such tubes are believed to be 'stored' either just below or at the bottom of the convection zone. By storage one means that the field survives reasonably undisturbed for a good fraction of the solar cycle and evolves mostly under the amplifying action of differential rotation. Once such a flux tube becomes buoyant in one section of the tube it rises, expands and becomes tilted relative to the azimuthal direction owing to the Coriolis force. Calculations based on the thin flux tube approximation predict field strengths of about 100 kG that are needed in order to produce the observed tilt angle of bipolar sunspots near the surface.

The systematic variation of the global field of the sun is important to understand both for practical reasons, e.g. for space weather forecasts, and for theoretical reasons because the solar field is a prime example of what we call large scale dynamo action.

The 11 year cycle of the sun is commonly explained in terms of $\alpha\Omega$ dynamo theory (§ 14.5), but this theory faces a number of problems that will be discussed later. Much of the resolution of these problems focuses around magnetic helicity. This has become a very active research field in its own right. Here we discuss the observational evidence.

Magnetic helicity of the solar field

Magnetic helicity studies have become an important observational tool to quantify the complexity of the sun's magnetic field. Examples of complex magnetic

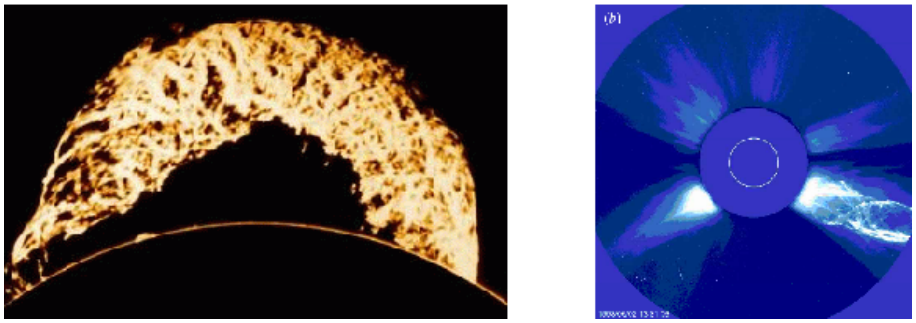


Figure 13.3: The famous "Grand daddy" prominence of 4 June 1946 (left) and a big coronal mass eruption of 2 June 1998 from the LASCO coronograph on board the SOHO satellite (right). Note the complexity of the ejected structures, being suggestive of helical nature. Courtesy of the SOHO consortium. SOHO is a project of international cooperation between ESA and NASA.

structures being ejected from the solar surface are shown in Fig. 13.3. The significance of magnetic helicity for understanding the nonlinear dynamo has only recently been appreciated. Here we briefly review some of the relevant observational findings.

The only information about the magnetic helicity of the sun available to date is from surface magnetic fields, and these data are necessarily incomplete. Nevertheless, some systematic trends can be identified.

Vector magnetograms of active regions show negative (positive) current helicity in the northern (southern) hemisphere. From local measurements one can only obtain the current helicity density, so nothing can be concluded about magnetic helicity, which is a volume integral. Under the assumption of isotropy, the spectra of magnetic and current helicity are however simply related by a wavenumber squared factor. This implies that the signs of current and magnetic helicities agree if they are determined in a sufficiently narrow range of length scales.

Berger and Ruzmaikin have estimated the flux of magnetic helicity from the solar surface using magnetograms. They discussed the α effect and differential rotation as the main agents facilitating the loss of magnetic helicity. Their results indicate that the flux of magnetic helicity due to differential rotation and the observed radial magnetic field component is negative (positive) in the northern (southern) hemisphere, and of the order of about $10^{46}\,\mathrm{Mx}^2$ integrated over the 11 year cycle; see Fig. 13.4.

Chae estimated the magnetic helicity flux based on counting the crossings of pairs of flux tubes. Combined with the assumption that two nearly aligned flux tubes are nearly parallel, rather than anti-parallel, his results again suggest that the magnetic helicity is negative (positive) in the northern (southern) hemisphere. The same sign distribution was also found by DeVore who con-

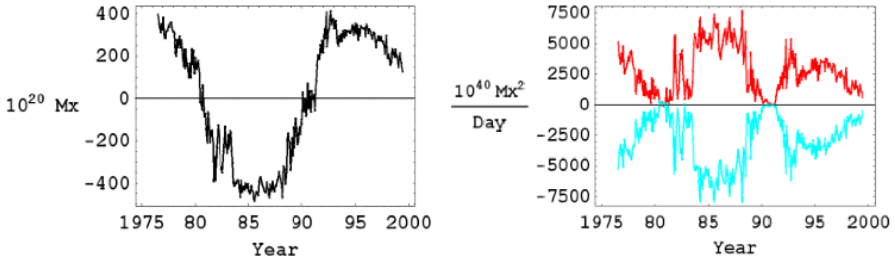


Figure 13.4: Net magnetic flux through the solar surface at the northern hemisphere (left hand panel) and magnetic helicity flux for northern and southern hemispheres (right hand panel, lower and upper curves, respectively). Adapted from Berger and Ruzmaikin.

Figure 13.5: X-ray image at 195 Å showing an N-shaped sigmoid (right-handed writhe) of the active region NOAA AR 8668 at the northern hemisphere (1999 August 21 at 18:51 UT).

sidered magnetic helicity generation by differential rotation. He finds that the magnetic helicity flux integrated over an 11 year cycle is about $10^{46}\,\mathrm{Mx}^2$ both from active regions and from coronal mass ejections. Thus, the sign agrees with that of the current helicity obtained using vector magnetograms. This idea of a bi-helical field is supported further by studies of sigmoids: an example is

Fig. 13.5, which shows a TRACE image of an N-shaped sigmoid (right-handed writhe) with left-handed twisted filaments of the active region NOAA AR 8668, which is typical of the northern hemisphere. This observation is quite central to our new understanding of nonlinear dynamo theory.

Chapter 14

Mean field dynamo theory

14.1 Solar and stellar magnetic fields

The sun has a magnetic field that manifests itself in sunspots through Zeeman splitting of spectral lines. It has long been known that the sunspot number varies cyclically with a period between 7 and 17 years. The longitudinally averaged component of the radial magnetic field of the sun shows a markedly regular spatio-temporal pattern where the radial magnetic field alternates in time over the 11 year cycle and also changes sign across the equator (Fig. 13.1). One can also see indications of a migration of the field from mid latitudes toward the equator and the poles. This migration is also well seen in a sunspot diagram, which is also called a butterfly diagram, because the pattern formed by the positions of sunspots in time and latitude looks like a sequence of butterflies lined up along the equator (Fig. 13.2).

At the solar surface the azimuthally averaged radial field is only a few gauss ($1\,\mathrm{G} = 10^{-4}\,\mathrm{Tesla}$). This is rather weak compared with the peak magnetic field in sunspots of about 2 kG; see Fig. 13.1. In the bulk of the convection zone, because of differential rotation, the magnetic field is believed to point mostly in the azimuthal direction, and it is probably much larger near the bottom of the convection zone due to an effect known as downward pumping.

14.2 Phenomenological considerations

Important insights into the operation of the solar dynamo have come from close inspection of magnetic fields on the solar surface. One important ingredient is differential rotation. At the equator the sun is rotating about 30% faster than at the poles. This means that any poloidal field will be sheared out and toroidal field aligned with the direction of the shear will be generated. Mathematically, this is described by the stretching term in the induction equation, i.e.

$$\frac{\mathrm{d}\boldsymbol{B}_{\mathrm{tor}}}{\mathrm{d}t} = \boldsymbol{B}_{\mathrm{pol}} \cdot \boldsymbol{\nabla} \boldsymbol{U}_{\mathrm{tor}} + \ldots . \tag{14.1}$$

This term describes the generation of magnetic field $\boldsymbol{B}_{\rm tor}$ in the direction of the flow $\boldsymbol{U}_{\rm tor}$ from a cross-stream poloidal magnetic field $\boldsymbol{B}_{\rm pol}$. To an order of magnitude, the amount of toroidal field generation from a 100 G poloidal field in a time interval $\Delta t = 10^8\,{\rm s} = 3\,{\rm yr}$ is

$$\Delta \boldsymbol{B}_{\rm tor} = \boldsymbol{B}_{\rm pol}\,\Delta\Omega_\odot \Delta t \approx 100\,{\rm G} \times 10^{-6} \times 10^8 = 10^4\,{\rm G}, \tag{14.2}$$

where we have used $\Omega_\odot = 3 \times 10^{-6}\,{\rm s}^{-1}$ for the solar angular velocity, and $\Delta\Omega_\odot/\Omega_\odot = 0.3$ for the relative latitudinal differential rotation. So, a 10 kG toroidal field can be regenerated completely from a 100 G poloidal field in about 3 years. However, in the bulk and the upper parts of the solar convection zone the poloidal fields are weaker (3-10 G), which would yield toroidal fields on the order of 300-1000 G. This would be far too weak a field if it was to rise coherently all the way from the bottom of the convection zone, which is still the standard picture. However, if the field of bipolar regions is produced locally in the upper parts of the convection zone, as recently supposed a 300 G field might well be sufficient. The 2 kG fields in sunspots could then be the result of local compression by an ambient flow.

Further, magnetic flux frequently emerges at the solar surface as bipolar regions. The magnetic field in sunspots is also often of bipolar nature. It was long recognized that such bipolar regions are tilted. This is now generally referred to as Joy's law. The sense of average tilt is clockwise in the northern hemisphere and counter-clockwise in the southern. This tilt is consistent with the interpretation that a toroidal flux tube rises from deeper layers of the sun to upper layers where the density is less, so the tube evolves in an expanding flow field which, due to the Coriolis force, attains a clockwise swirl in the northern hemisphere and counter-clockwise swirl in the southern hemisphere; see Fig. 14.1.

Observations suggest that once a tilted bipolar region has emerged at the solar surface, the field polarities nearer to the poles drift rapidly toward the poles, producing thereby new poloidal field. Underneath the surface, the field continues as before, but there it is also slightly tilted, although necessarily in the opposite sense (see Fig. 14.2). Because of differential rotation, the points nearest to the equator move faster, helping so to line up similarly oriented fields. As is evident from Fig. 14.2, a toroidal field pointing east in the northern hemisphere and west in the southern will develop into a global northward pointing field above the surface.

14.3 Mean-field electrodynamics

In 1955 Parker first proposed the idea that the generation of a poloidal field, arising from the systematic effects of the Coriolis force (Fig. 14.3), could be described by a corresponding term in the induction equation,

$$\frac{\partial \overline{\boldsymbol{B}}_{\rm pol}}{\partial t} = \boldsymbol{\nabla} \times \left(\alpha \overline{\boldsymbol{B}}_{\rm tor} + ...\right). \tag{14.3}$$

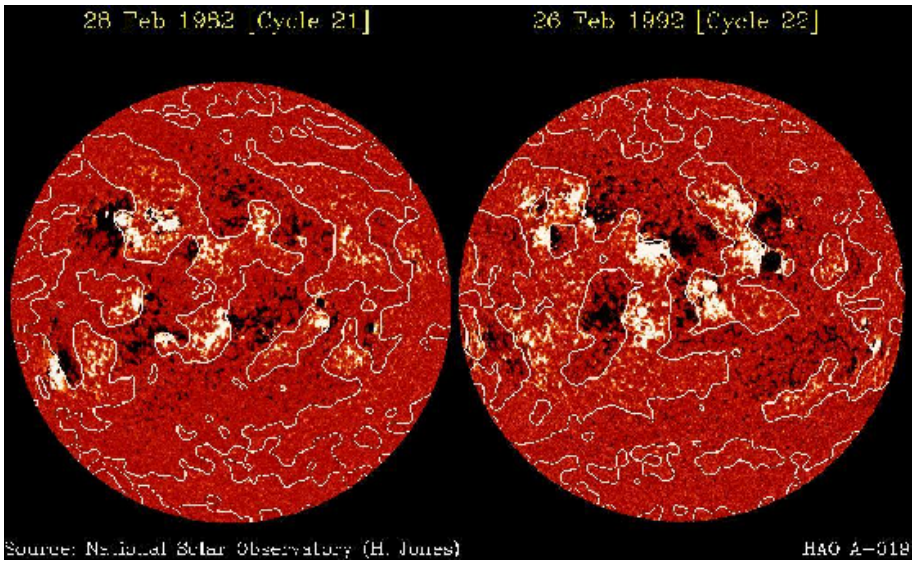


Figure 14.1: Solar magnetogram showing bipolar regions, their opposite orientation north and south of the equator, and the clockwise tilt in the northern hemisphere and the counter-clockwise tilt in the southern hemisphere. Note that the field orientation has reversed orientation at the next cycle (here after 10 years). Courtesy of the High Altitude Observatory.

It is clear that such an equation can only be valid for averaged fields (denoted by overbars), because for the actual fields, the induced electromotive force (EMF) $\boldsymbol{U}\times\boldsymbol{B}$, would never have a component in the direction of $\boldsymbol{B}$. While being physically plausible, this approach only received general recognition and acceptance after Roberts and Stix (1972) translated the work of Steenbeck, Krause, Rädler (1966) into English. In those papers the theory for the α effect, as they called it, was developed and put on a mathematically rigorous basis. Furthermore, the α effect was also applied to spherical models of the solar cycle (with radial and latitudinal shear) and the geodynamo (with uniform rotation).

In mean field theory one solves the Reynolds averaged equations, using either ensemble averages, toroidal averages or, in cases in cartesian geometry with periodic boundary conditions, two-dimensional (e.g. horizontal) averages. We thus consider the decomposition

$$\boldsymbol{U}=\overline{\boldsymbol{U}}+\boldsymbol{u},\quad \boldsymbol{B}=\overline{\boldsymbol{B}}+\boldsymbol{b}. \tag{14.4}$$

Here $\overline{\boldsymbol{U}}$ and $\overline{\boldsymbol{B}}$ are the mean velocity and magnetic fields, while $\boldsymbol{u}$ and $\boldsymbol{b}$ are their fluctuating parts. These averages satisfy the Reynolds rules,

$$\overline{\boldsymbol{U}_1+\boldsymbol{U}_2}=\overline{\boldsymbol{U}}_1+\overline{\boldsymbol{U}}_2,\quad \overline{\overline{\boldsymbol{U}}}=\overline{\boldsymbol{U}},\quad \overline{\overline{\boldsymbol{U}}\boldsymbol{u}}=0,\quad \overline{\overline{\boldsymbol{U}}\,\overline{\boldsymbol{U}}}=\overline{\boldsymbol{U}}\,\overline{\boldsymbol{U}}, \tag{14.5}$$

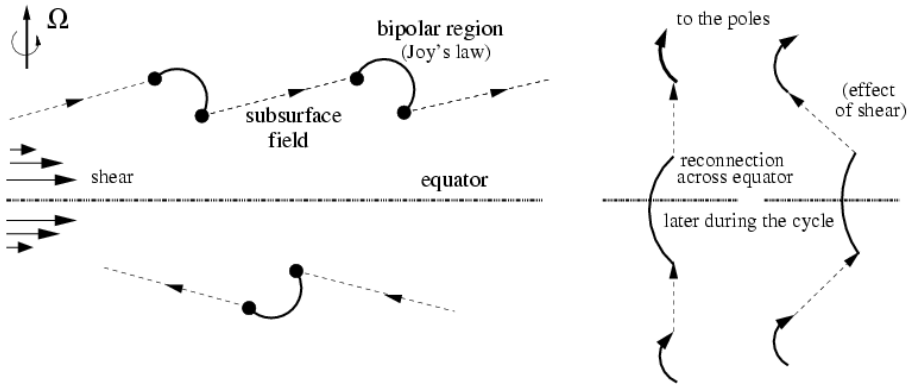


Figure 14.2: Sketch of the Babcock-Leighton dynamo mechanism. As bipolar regions emerge near the surface, they get tilted in the clockwise sense in the northern hemisphere and counter-clockwise in the southern hemisphere. Beneath the surface this process leaves behind a poloidal field component that points here toward the north pole on either side of the equator. Once the remaining subsurface field gets sheared by the surface differential rotation, it points in the opposite direction as before, and the whole process starts again.

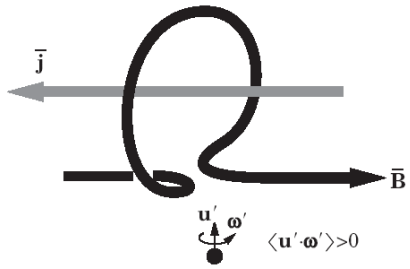


Figure 14.3: Production of positive writhe helicity by an uprising and expanding blob tilted in the clockwise direction by the Coriolis force in the southern hemisphere, producing a field-aligned current $\overline{\boldsymbol{J}}$ in the opposite direction to $\overline{\boldsymbol{B}}$.

$$\overline{\partial \boldsymbol{U}/\partial t} = \partial \overline{\boldsymbol{U}}/\partial t, \quad \overline{\partial \boldsymbol{U}/\partial x_i} = \partial \overline{\boldsymbol{U}}/\partial x_i. \tag{14.6}$$

Some of these properties are not shared by several other averages; for gaussian filtering $\overline{\overline{\boldsymbol{U}}} \neq \overline{\boldsymbol{U}}$, and for spectral filtering $\overline{\overline{\boldsymbol{U}}\,\overline{\boldsymbol{U}}} \neq \overline{\boldsymbol{U}}\,\overline{\boldsymbol{U}}$, for example. Note that $\overline{\overline{\boldsymbol{U}}} = \overline{\boldsymbol{U}}$ implies that $\overline{\boldsymbol{u}} = 0$.

In the remainder we assume that the Reynolds rules do apply. Averaging equation (10.13) yields then the mean field induction equation,

$$\frac{\partial \overline{\boldsymbol{B}}}{\partial t} = \boldsymbol{\nabla} \times \left(\overline{\boldsymbol{U}} \times \overline{\boldsymbol{B}} + \overline{\mathcal{E}} - \eta \overline{\boldsymbol{J}} \right), \tag{14.7}$$

where

$$\overline{\mathcal{E}} = \overline{\boldsymbol{u} \times \boldsymbol{b}} \tag{14.8}$$

is the mean EMF. Finding an expression for the correlator $\overline{\mathcal{E}}$ in terms of the mean fields is a standard closure problem which is at the heart of mean field theory. In the two-scale approach one assumes that $\overline{\mathcal{E}}$ can be expanded in powers of the gradients of the mean magnetic field. This suggests the rather general expression

$$\mathcal{E}_i = \alpha_{ij}(\boldsymbol{g}, \hat{\boldsymbol{\Omega}}, \overline{\boldsymbol{B}}, ...)\overline{B}_j + \eta_{ijk}(\boldsymbol{g}, \hat{\boldsymbol{\Omega}}, \overline{\boldsymbol{B}}, ...)\partial \overline{B}_j / \partial x_k, \tag{14.9}$$

where the tensor components α_{ij} and η_{ijk} are referred to as turbulent transport coefficient. They depend on the stratification, angular velocity, and mean magnetic field strength. The dots indicate that the transport coefficients may also depend on correlators involving the small scale magnetic field, for example the current helicity of the small scale field. We have also kept only the lowest large scale derivative of the mean field; higher derivative terms are expected to be smaller.

The general form of the expression for $\overline{\mathcal{E}}$ can be determined by rather general considerations. For example, $\overline{\mathcal{E}}$ is a polar vector and $\overline{\boldsymbol{B}}$ is an axial vector, so α_{ij} must be a pseudo-tensor. The simplest pseudo-tensor of rank two that can be constructed using the unit vectors $\boldsymbol{g}$ (symbolic for radial density or turbulent velocity gradients) and $\hat{\boldsymbol{\Omega}}$ (angular velocity) is

$$\alpha_{ij} = \alpha_1 \delta_{ij}\, \boldsymbol{g} \cdot \hat{\boldsymbol{\Omega}} + \alpha_2 \hat{g}_i \hat{\Omega}_j + \alpha_3 \hat{g}_j \hat{\Omega}_i. \tag{14.10}$$

Note that the term $\boldsymbol{g} \cdot \hat{\boldsymbol{\Omega}} = \cos\theta$ leads to the co-sinusoidal dependence of α on latitude, θ, and a change of sign at the equator. Additional terms that are nonlinear in $\boldsymbol{g}$ or $\hat{\boldsymbol{\Omega}}$ enter if the stratification is strong or if the body is rotating rapidly. Likewise, terms involving $\overline{\boldsymbol{U}}$, $\overline{\boldsymbol{B}}$ and $\boldsymbol{b}$ may appear if the turbulence becomes affected by strong flows or magnetic fields. In the following section we discuss various approaches to determining the turbulent transport coefficients.

One of the most important outcomes of this theory is a quantitative formula for the coefficient α_1 in equation (14.10) by Krause (1967)

$$\alpha_1\, \boldsymbol{g} \cdot \hat{\boldsymbol{\Omega}} = -\tfrac{16}{15} \tau_{\rm cor}^2 u_{\rm rms}^2 \boldsymbol{\Omega} \cdot \boldsymbol{\nabla} \ln(\rho u_{\rm rms}), \tag{14.11}$$

where $\tau_{\rm cor}$ is the correlation time, $u_{\rm rms}$ the root mean square velocity of the turbulence, and $\boldsymbol{\Omega}$ the angular velocity vector. The other coefficients are given by $\alpha_2 = \alpha_3 = -\alpha_1/4$. Throughout most of the solar convection zone, the product $\rho u_{\rm rms}$ decreases outward.[1] Therefore, $\alpha > 0$ throughout most of the

[1] This can be explained as follows: in the bulk of the solar convection zone the convective

northern hemisphere. In the southern hemisphere we have $\alpha < 0$, and α varies with colatitude θ like $\cos\theta$. However, this formula also predicts that α reverses sign very near the bottom of the convection zone where $u_{\rm rms} \to 0$. This is caused by the relatively sharp drop of $u_{\rm rms}$.

14.4 Turbulent transport coefficients

Various techniques have been proposed for determining turbulent transport coefficients. Even in the kinematic regime, where the changes in the velocity field due to Lorentz forces are ignored, these techniques have some severe uncertainties. Nevertheless, the various techniques produce similar terms, although the so-called minimal τ approximation (MTA) does actually predict an extra time derivative of the electromotive force. We only mention here that in MTA the triple correlations are not neglected, as they are in FOSA; see § 14.4.1. Instead, the triple correlations are approximated by quadratic terms. This is similar in spirit to the usual τ approximation used in the Eddy Damped Quasi-Normal Markovian (EDQNM) closure approximation. where the irreducible part of quartic correlations are approximated by a relaxation term proportional to the triple correlations. Other approaches include direct simulations, calculations based on random waves or individual blobs, or calculations based on the assumption of delta-correlated velocity fields.

14.4.1 First order smoothing approximation

The first order smoothing approximation (FOSA) or, synonymously, the quasi-linear approximation, or the second order correlation approximation is the simplest way of calculating turbulent transport coefficients. The approximation consists of linearizing the equations for the fluctuating quantities and ignoring quadratic terms that would lead to triple correlations in the expressions for the quadratic correlations. This technique has traditionally been applied to calculating the turbulent diffusion coefficient for a passive scalar or the turbulent viscosity (eddy viscosity).

Suppose we consider the induction equation. The equation for the fluctuating field can be obtained by subtracting equation (14.7) from equation (10.13), so

$$\frac{\partial \boldsymbol{b}}{\partial t} = \boldsymbol{\nabla} \times \left(\overline{\boldsymbol{U}} \times \boldsymbol{b} + \boldsymbol{u} \times \overline{\boldsymbol{B}} + \boldsymbol{u} \times \boldsymbol{b} - \overline{\mathcal{E}} - \eta \boldsymbol{j} \right), \tag{14.12}$$

where $\boldsymbol{j} = \boldsymbol{\nabla} \times \boldsymbol{b} \equiv \boldsymbol{J} - \overline{\boldsymbol{J}}$ is the fluctuating current density. The first order smoothing approximation consists of *neglecting* the term $\boldsymbol{u} \times \boldsymbol{b}$ on the RHS of equation (14.12), because it is nonlinear in the fluctuations. This can only be done if the fluctuations are small, which is a good approximation only under

flux is approximately constant, and mixing length predicts that it is approximately $\rho u_{\rm rms}^3$. This in turn follows from $F_{\rm conv} \sim \rho u_{\rm rms} c_p \delta T$ and $u_{\rm rms}^2 / H_p \sim g \delta T / T$ together with the expression for the pressure scale height $H_p = (1 - \frac{1}{\gamma}) c_p T / g$. Thus, since $\rho u_{\rm rms}^3 \approx$ const, we have $u_{\rm rms} \sim \rho^{-1/3}$ and $\rho u_{\rm rms}^3 \sim \rho^{2/3}$.

rather restrictive circumstances, for example if $R_{\rm m}$ is small. The term $\overline{\mathcal{E}}$ is also nonlinear in the fluctuations, but it is not a fluctuating quantity and gives therefore no contribution, and the $\overline{\boldsymbol{U}}\times\boldsymbol{b}$ is often neglected because of simplicity. The neglect of the $\overline{\boldsymbol{U}}$ term may not be justified for systems with strong shear (e.g. for accretion discs) where the inclusion of $\overline{\boldsymbol{U}}$ itself could lead to a new dynamo effect, namely the shear–current effect. In the case of small $R_{\rm m}$, one can neglect both the $\boldsymbol{G}$ term and the time derivative of $\boldsymbol{b}$, resulting in a linear equation

$$\eta\boldsymbol{\nabla}^2\boldsymbol{b}=-\boldsymbol{\nabla}\times\left(\boldsymbol{u}\times\overline{\boldsymbol{B}}\right). \tag{14.13}$$

This can be solved for $\boldsymbol{b}$, if $\boldsymbol{u}$ is given. $\overline{\mathcal{E}}$ can then be computed relatively easily.

However, in most astrophysical applications, $R_{\rm m}\gg1$. In such a situation, FOSA is thought to still be applicable if the correlation time $\tau_{\rm cor}$ of the turbulence is small, such that $\tau_{\rm cor}u_{\rm rms}k_{\rm f}\ll1$, where $u_{\rm rms}$ and $k_{\rm f}$ are typical velocity and correlation wavenumber, associated with the random velocity field $\boldsymbol{u}$. Under this condition, the ratio of the nonlinear term to the time derivative of $\boldsymbol{b}$ is argued to be $\sim(u_{\rm rms}k_{\rm f}b)/(b/\tau_{\rm cor})=\tau_{\rm cor}u_{\rm rms}k_{\rm f}\ll1$, and so $\boldsymbol{G}$ can be neglected (but see below). We then get

$$\frac{\partial\boldsymbol{b}}{\partial t}=\boldsymbol{\nabla}\times\left(\boldsymbol{u}\times\overline{\boldsymbol{B}}\right). \tag{14.14}$$

To calculate $\overline{\mathcal{E}}$, we integrate $\partial\boldsymbol{b}/\partial t$ to get $\boldsymbol{b}$, take the cross product with $\boldsymbol{u}$, and average, i.e.

$$\overline{\mathcal{E}}=\overline{\boldsymbol{u}(t)\times\int_0^t\boldsymbol{\nabla}\times\left[\boldsymbol{u}(t')\times\overline{\boldsymbol{B}}(t')\right]\,{\rm d}t'}. \tag{14.15}$$

For clarity, we have suppressed the common $\boldsymbol{x}$ dependence of all variables. Using index notation, we have[2]

$$\overline{\mathcal{E}}_i(t)=\int_0^t\left[\hat{\alpha}_{ip}(t,t')\overline{B}_p(t')+\hat{\eta}_{ilp}(t,t')\overline{B}_{p,l}(t')\right]{\rm d}t', \tag{14.16}$$

with $\hat{\alpha}_{ip}(t,t')=\epsilon_{ijk}\overline{u_j(t)u_{k,p}(t')}$ and $\hat{\eta}_{ilp}(t,t')=-\epsilon_{ijp}\overline{u_j(t)u_l(t')}$, where we have used $\overline{B}_{l,l}=0=u_{l,l}$, and an additional term $\epsilon_{ijk}\overline{u_j(t)u_k(t')}\delta_{lp}$ in $\hat{\eta}_{ilp}(t,t')$ has been omitted, because it will soon drop out. In the statistically steady state, we can assume that $\hat{\alpha}_{ip}$ and $\hat{\eta}_{ilp}$ depend only on the time difference, $t-t'$. Assuming isotropy (again only for simplicity), these tensors must be proportional to the isotropic tensors δ_{ip} and ϵ_{ilp}, respectively, so we have

$$\overline{\mathcal{E}}(t)=\int_0^t\left[\hat{\alpha}(t-t')\overline{\boldsymbol{B}}(t')-\hat{\eta}_{\rm t}(t-t')\overline{\boldsymbol{J}}(t')\right]{\rm d}t', \tag{14.17}$$

[2] Note that $[\boldsymbol{u}\times\boldsymbol{\nabla}\times(\boldsymbol{u}\times\overline{\boldsymbol{B}})]_i=\epsilon_{ijk}\epsilon_{klm}\epsilon_{mnp}u_j\partial_l(u_n\overline{B}_p)=\hat{\alpha}_{ip}\overline{B}_p+\hat{\eta}_{ilp}\overline{B}_{p,l}$, where commas denote partial differentiation and time arguments in $\hat{\alpha}_{ip}=\hat{\alpha}_{ip}(t,t')$, $\hat{\eta}_{ilp}=\hat{\eta}_{ilp}(t,t')$, and $\overline{B}_p=\overline{B}_p(t')$ has been omitted. Contracting $\epsilon_{klm}\epsilon_{mnp}=\delta_{kn}\delta_{lp}-\delta_{kp}\delta_{ln}$ gives $\hat{\alpha}_{ip}=\epsilon_{ijk}(\overline{u_ju_{k,p}}-\overline{u_ju_{l,l}\delta_{kp}}$, and $\hat{\eta}_{ilp}=\epsilon_{ijk}(\overline{u_ju_k}\delta_{lp}-\overline{u_ju_l}\delta_{kp})$.

where $\hat{\alpha}(t-t') = -\frac{1}{3}\overline{\boldsymbol{u}(t)\cdot\boldsymbol{\omega}(t')}$ and $\hat{\eta}_{\rm t}(t-t') = \frac{1}{3}\overline{\boldsymbol{u}(t)\cdot\boldsymbol{u}(t')}$ are integral kernels, and $\boldsymbol{\omega} = \boldsymbol{\nabla}\times\boldsymbol{u}$ is the vorticity of the velocity fluctuation (see the footnote[3] for details).

If we assume the integral kernels to be proportional to the delta function, $\delta(t-t')$, or, equivalently, if $\overline{\boldsymbol{B}}$ can be considered a slowly varying function of time, one arrives at

$$\overline{\mathcal{E}} = \alpha\overline{\boldsymbol{B}} - \eta_{\rm t}\overline{\boldsymbol{J}} \tag{14.18}$$

with

$$\alpha = -\tfrac{1}{3}\int_0^t \overline{\boldsymbol{u}(t)\cdot\boldsymbol{\omega}(t')}\,{\rm d}t' \approx -\tfrac{1}{3}\tau_{\rm cor}\overline{\boldsymbol{u}\cdot\boldsymbol{\omega}}, \tag{14.19}$$

$$\eta_{\rm t} = \tfrac{1}{3}\int_0^t \overline{\boldsymbol{u}(t)\cdot\boldsymbol{u}(t')}\,{\rm d}t' \approx \tfrac{1}{3}\tau_{\rm cor}\overline{\boldsymbol{u}^2}. \tag{14.20}$$

When t becomes large, the main contribution to these two expressions comes only from late times, $t-t' \ll t$, because then the contributions from early times are no longer strongly correlated with $\boldsymbol{u}(t)$. By using FOSA we have thus solved the problem of expressing $\overline{\mathcal{E}}$ in terms of the mean field. The turbulent transport coefficients α and $\eta_{\rm t}$ depend, respectively, on the helicity and the energy density of the turbulence.

One must however point out the following caveat to the applicability of FOSA in case of large $R_{\rm m}$. Firstly note that even if $\tau_{\rm cor}u_{\rm rms}k_{\rm f} \ll 1$, one can have $R_{\rm m} = (\tau_{\rm cor}u_{\rm rms}k_{\rm f})/(\eta\tau_{\rm cor}k_{\rm f}^2) \gg 1$, because the diffusion time $(\eta\tau_{\rm cor}k_{\rm f}^2)^{-1}$ can be much larger than the correlation time of the turbulence. When $R_{\rm m} > R_{\rm crit} \sim 30$, small scale dynamo action will take place to produce exponentially growing fluctuating fields, independent of the mean field. So the basic assumption of FOSA of small $\boldsymbol{b}$ relative to $\overline{\boldsymbol{B}}$ will be rapidly violated and the $\boldsymbol{u}\times\boldsymbol{b}$ term in equation (14.12) cannot be neglected. Nevertheless, the functional form of the expressions for the turbulent transport coefficients obtained using FOSA seem to be not too different from that found in simulations. For example, it is likely that strong fluctuations produced by small scale dynamo action do not correlate well with $\boldsymbol{u}$ in $\overline{\boldsymbol{u}\times\boldsymbol{b}}$, so they would not contribute to $\overline{\mathcal{E}}$. This interpretation will be developed further in § 14.4.2 on the τ approximation, which works specifically only with those parts that do correlate.

14.4.2 MTA – the 'minimal' τ approximation

The 'minimal' τ approximation is a simplified version of the τ approximation as it has been introduced by Orszag (1970) and used by Pouquet, Frisch and Léorat (1976) in the context of the Eddy Damped Quasi Normal Markovian (EDQNM) approximation. In that case a damping term is introduced in order to express fourth order moments in terms of third order moments. In the τ approximation,

[3] Note that under isotropy we have $\hat{\alpha}_{ip} = \hat{\alpha}\delta_{ip}$ and $\hat{\eta}_{ilp} = \hat{\eta}_{\rm t}\epsilon_{ilp}$. Multiplying these equations by δ_{ip} and ϵ_{ilp}, respectively, and noting that $\delta_{ip}\delta_{ip} = 3$ and $\epsilon_{ilp}\epsilon_{ilp} = 6$ we have $\hat{\alpha} = \frac{1}{3}\hat{\alpha}_{ip}\delta_{ip} = \frac{1}{3}\overline{u_j\epsilon_{jki}u_{k,i}} = -\frac{1}{3}\boldsymbol{u}\cdot\boldsymbol{\omega}$, and $\hat{\eta}_{\rm t} = \frac{1}{6}\hat{\eta}_{ilp}\epsilon_{ilp} = -\frac{1}{6}\epsilon_{ijp}\overline{u_ju_l}\epsilon_{ilp} = -\frac{1}{3}\overline{\boldsymbol{u}^2}$, where we have used $\epsilon_{ijp}\epsilon_{ilp} = 2\delta_{jl}$, and the t and t' arguments have been omitted.

as introduced by Vainshtein and Kitchatinov and Kleeorin and Rogachevskii one approximates triple moments in terms of quadratic moments via a wavenumber-dependent relaxation time $\tau(k)$. The 'minimal' τ approximation (MTA), as it is introduced by Blackman and Field is applied in real space in the two-scale approximation. We will refer to both the above types of closures (where triple moments are approximated in terms of quadratic moments and a relaxation time τ) as the 'minimal' τ approximation or MTA.

There are some technical similarities between FOSA and the minimal τ approximation. The main advantage of the τ approximation is that the fluctuations do *not* need to be small and so the triple correlations are no longer neglected. Instead, it is assumed (and this can be and has been tested using simulations) that the one-point triple correlations are proportional to the quadratic correlations, and that the proportionality coefficient is an inverse relaxation time that can in principle be scale (or wavenumber) dependent.

In this approach, one begins by considering the time derivative of $\overline{\mathcal{E}}$,

$$\frac{\partial \overline{\mathcal{E}}}{\partial t} = \overline{\boldsymbol{u} \times \dot{\boldsymbol{b}}} + \overline{\dot{\boldsymbol{u}} \times \boldsymbol{b}}, \tag{14.21}$$

where a dot denotes a time derivative. For $\dot{\boldsymbol{b}}$, we substitute equation (14.12) and for $\dot{\boldsymbol{u}}$, we use the Euler equation for the fluctuating velocity field,

$$\frac{\partial \boldsymbol{u}}{\partial t} = -\frac{1}{\rho_0}\boldsymbol{\nabla} p + \boldsymbol{f} + \boldsymbol{F}_{\rm vis} + \boldsymbol{H}, \tag{14.22}$$

where $\boldsymbol{H} = -\boldsymbol{u}\cdot\boldsymbol{\nabla}\boldsymbol{u} + \overline{\boldsymbol{u}\cdot\boldsymbol{\nabla}\boldsymbol{u}}$ is the nonlinear term, $\boldsymbol{f}$ is a stochastic forcing term (with zero divergence), and $\boldsymbol{F}_{\rm vis}$ is the viscous force. We have also assumed for the present that there is no mean flow ($\overline{\boldsymbol{U}} = 0$), and have considered the kinematic regime where the Lorentz force is set to zero. All these restrictions can in principle be lifted (see below). For an incompressible flow, the pressure term can be eliminated in the standard fashion in terms of the projection operator. Now since $\boldsymbol{f}$ does not correlate with $\boldsymbol{b}$, the only contribution to $\overline{\dot{\boldsymbol{u}} \times \boldsymbol{b}}$, is the small viscous term and the triple correlation involving $\boldsymbol{b}$ and $\boldsymbol{H}$. The $\overline{\boldsymbol{u} \times \dot{\boldsymbol{b}}}$ term however has non-trivial contributions. We get

$$\frac{\partial \overline{\mathcal{E}}}{\partial t} = \tilde{\alpha}\,\overline{\boldsymbol{B}} - \tilde{\eta}_{\rm t}\,\overline{\boldsymbol{J}} - \frac{\overline{\mathcal{E}}}{\tau}, \tag{14.23}$$

where the last term subsumes the effects of all triple correlations, and

$$\tilde{\alpha} = -\tfrac{1}{3}\overline{\boldsymbol{u}\cdot\boldsymbol{\omega}} \quad \text{and} \quad \tilde{\eta}_{\rm t} = \tfrac{1}{3}\overline{\boldsymbol{u}^2} \quad \text{(kinematic theory)} \tag{14.24}$$

are coefficients that are closely related to the usual α and $\eta_{\rm t}$ coefficients in equation (14.18). We recall that in this *kinematic* calculation the Lorentz force, and in fact the entire $\dot{\boldsymbol{u}}$ equation in equation (14.21) has been ignored. Its inclusion turns out to be extremely important: it leads to the emergence of a small scale magnetic correction term in the expression for $\tilde{\alpha}$.

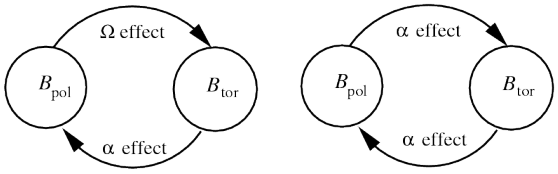


Figure 14.4: Mutual regeneration of poloidal and toroidal fields in the case of the $\alpha\Omega$ dynamo (left) and the α^2 dynamo (right).

One normally neglects the explicit time derivative of $\overline{\mathcal{E}}$, and arrives then at almost the same expression as equation (14.18), except that now one deals directly with one-point correlation functions and not only via an approximation. Furthermore, the explicit time derivative can in principle be kept, although it becomes unimportant on time scales long compared with τ. In comparison with equation (14.17), we note that if one assumes $\hat{\alpha}(t-t')$ and $\hat{\eta}_{\rm t}(t-t')$ to be proportional to $\exp[-(t-t')/\tau]$ for $t > t'$ (and zero otherwise), one recovers equation (14.23) with the relaxation time τ playing now the role of a correlation time.

14.5 α^2 and $\alpha\Omega$ dynamos: simple solutions

For astrophysical purposes one is usually interested in solutions in spherical or oblate (disc-like) geometries. However, in order to make contact with turbulence simulations in a periodic box, solutions in simpler cartesian geometry can be useful. Cartesian geometry is also useful for illustrative purposes. In this subsection we review some simple cases.

Mean field dynamos are traditionally divided into two groups; $\alpha\Omega$ and α^2 dynamos. The Ω effect refers to the amplification of the toroidal field by shear (i.e. *differential* rotation) and its importance for the sun was recognized very early on. Such shear also naturally occurs in disk galaxies, since they are differentially rotating systems. However, it is still necessary to regenerate the poloidal field. In both stars and galaxies the α effect is the prime candidate. This explains the name $\alpha\Omega$ dynamo; see the left hand panel of Fig. 14.4. However, large scale magnetic fields can also be generated by the α effect alone, so now also the toroidal field has to be generated by the α effect, in which case one talks about an α^2 dynamo; see the right hand panel of Fig. 14.4. (The term $\alpha^2\Omega$ model is discussed at the end of § 14.5.2.)

14.5.1 α^2 dynamo in a periodic box

We assume that there is no mean flow, i.e. $\overline{\boldsymbol{U}} = 0$, and that the turbulence is homogeneous, so that α and $\eta_{\rm t}$ are constant. The mean field induction equation

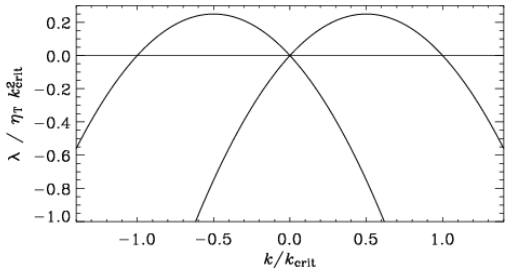


Figure 14.5: Dispersion relation for α^2 dynamo, where $k_{\rm crit}=\alpha/\eta_{\rm T}$.

then reads

$$\frac{\partial\overline{\boldsymbol{B}}}{\partial t}=\alpha\boldsymbol{\nabla}\times\overline{\boldsymbol{B}}+\eta_{\rm T}\nabla^2\overline{\boldsymbol{B}},\quad\boldsymbol{\nabla}\cdot\overline{\boldsymbol{B}}=0, \tag{14.25}$$

where $\eta_{\rm T}=\eta+\eta_{\rm t}$ is the sum of microscopic and turbulent magnetic diffusivity. We can seek solutions of the form

$$\overline{\boldsymbol{B}}(\boldsymbol{x})=\mathrm{Re}\left[\hat{\boldsymbol{B}}(\boldsymbol{k})\,\exp(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}+\lambda t)\right]. \tag{14.26}$$

This leads to the eigenvalue problem $\lambda\hat{\boldsymbol{B}}=\alpha\mathrm{i}\boldsymbol{k}\times\hat{\boldsymbol{B}}-\eta_{\rm T}k^2\hat{\boldsymbol{B}}$, which can be written in matrix form as

$$\lambda\hat{\boldsymbol{B}}=\begin{pmatrix}-\eta_{\rm T}k^2 & -\mathrm{i}\alpha k_z & \mathrm{i}\alpha k_y\\ \mathrm{i}\alpha k_z & -\eta_{\rm T}k^2 & -\mathrm{i}\alpha k_x\\ -\mathrm{i}\alpha k_y & \mathrm{i}\alpha k_x & -\eta_{\rm T}k^2\end{pmatrix}\hat{\boldsymbol{B}}. \tag{14.27}$$

This leads to the dispersion relation, $\lambda=\lambda(\boldsymbol{k})$, given by

$$(\lambda+\eta_{\rm T}k^2)\left[(\lambda+\eta_{\rm T}k^2)^2-\alpha^2k^2\right]=0, \tag{14.28}$$

with the three solutions

$$\lambda_0=-\eta_{\rm T}k^2,\quad\lambda_\pm=-\eta_{\rm T}k^2\pm|\alpha k|. \tag{14.29}$$

The eigenfunction corresponding to the eigenvalue $\lambda_0=-\eta_{\rm T}k^2$ is proportional to $\boldsymbol{k}$, but this solution is incompatible with solenoidality and has to be dropped. The two remaining branches are shown in Fig. 14.5.

Unstable solutions ($\lambda>0$) are possible for $0<\alpha k<\eta_{\rm T}k^2$. For $\alpha>0$ this corresponds to the range

$$0<k<\alpha/\eta_{\rm T}\equiv k_{\rm crit}. \tag{14.30}$$

For $\alpha<0$, unstable solutions are obtained for $k_{\rm crit}<k<0$. The maximum growth rate is at $k=\frac{1}{2}k_{\rm crit}$. Such solutions are of some interest, because they have been seen as an additional hump in the magnetic energy spectra from fully three-dimensional turbulence simulations.

14.5.2 $\alpha\Omega$ dynamo in a periodic box

Next we consider the case with linear shear, and assume $\overline{\boldsymbol{U}} = (0, Sx, 0)$, where $S = \text{const}$. This model can be applied as a local model to both accretion discs (x is radius, y is longitude, and z is the height above the midplane) and to stars (x is latitude, y is longitude, and z is radius). For keplerian discs, the shear is $S = -\frac{3}{2}\Omega$, while for the sun (taking here only radial differential rotation into account) $S = r\partial\Omega/\partial r \approx +0.1\Omega_\odot$ near the equator.

For simplicity we consider axisymmetric solutions, i.e. $k_y = 0$. The eigenvalue problem takes then the form

$$\lambda\hat{\boldsymbol{B}} = \begin{pmatrix} -\eta_{\rm T}k^2 & -{\rm i}\alpha k_z & 0 \\ {\rm i}\alpha k_z + S & -\eta_{\rm T}k^2 & -{\rm i}\alpha k_x \\ 0 & {\rm i}\alpha k_x & -\eta_{\rm T}k^2 \end{pmatrix} \hat{\boldsymbol{B}}, \tag{14.31}$$

where $\eta_{\rm T} = \eta + \eta_{\rm t}$ and $\boldsymbol{k}^2 = k_x^2 + k_z^2$. The dispersion relation is now

$$(\lambda + \eta_{\rm T}k^2)\left[(\lambda + \eta_{\rm T}k^2)^2 + {\rm i}\alpha S k_z - \alpha^2 k^2\right] = 0, \tag{14.32}$$

with the solutions

$$\lambda_\pm = -\eta_{\rm T}k^2 \pm (\alpha^2 k^2 - {\rm i}\alpha S k_z)^{1/2}. \tag{14.33}$$

Again, the eigenfunction corresponding to the eigenvalue $\lambda_0 = -\eta_{\rm T}k^2$ is not compatible with solenoidality and has to be dropped. The two remaining branches are shown in Fig. 14.6, together with the *approximate* solutions (valid for $\alpha k_z/S \ll 1$)

$${\rm Re}\lambda_\pm \approx -\eta_{\rm T}k^2 \pm |\tfrac{1}{2}\alpha S k_z|^{1/2}, \tag{14.34}$$

$${\rm Im}\lambda_\pm \equiv -\omega_{\rm cyc} \approx \pm|\tfrac{1}{2}\alpha S k_z|^{1/2}, \tag{14.35}$$

where we have made use of the fact that ${\rm i}^{1/2} = (1+{\rm i})/\sqrt{2}$.

Sometimes the term $\alpha^2\Omega$ dynamo is used to emphasize that the α effect is not neglected in the generation of the toroidal field. This approximation, which is sometimes also referred to as the $\alpha\Omega$ approximation, is generally quite good provided $\alpha k_z/S \ll 1$. However, it is important to realize that this approximation can only be applied to axisymmetric solutions.

14.5.3 Eigenfunctions, wave speed, and phase relations

We now make the $\alpha\Omega$ approximation and consider the marginally excited solution (${\rm Re}\lambda = 0$), which can be written as

$$\overline{B}_x = B_0 \sin k_z(z - ct), \quad \overline{B}_y = \sqrt{2}\, B_0 \left|\frac{c}{\alpha}\right| \sin[k_z(z-ct) + \varphi], \tag{14.36}$$

where B_0 is the amplitude (undetermined in linear theory), and $c = \omega_{\rm cyc}/k_z$ is the phase speed of the dynamo wave, which is given by

$$c = \frac{\alpha S}{|2\alpha S k_z|^{1/2}} = \pm\eta_{\rm T} k_z, \tag{14.37}$$

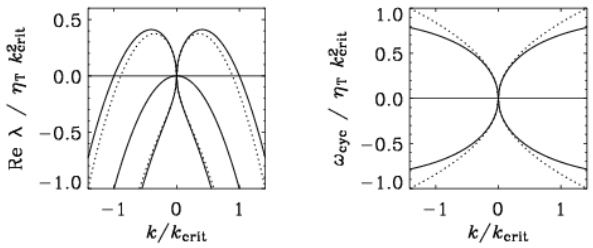


Figure 14.6: Dispersion relation for $\alpha^2\Omega$ dynamo with $\alpha k_{\rm crit}/S = 0.35$. The dotted line gives the result for the $\alpha\Omega$ approximation equations (14.34) and (14.35). The axes are normalized using $k_{\rm crit}$ for the full $\alpha^2\Omega$ dynamo equations.

Table 14.1: Summary of propagation directions and phase relation for $\alpha\Omega$ dynamos.

object	α	S	φ	c	wave propagation	$\lvert\omega_{\rm cyc}\rvert\,\Delta t$
disc	$-$	$-$	$-3\pi/4$	$+$	away from midplane	$-3\pi/4$
disc/star?	$+$	$-$	$+3\pi/4$	$-$	equatorward	$-3\pi/4$
star?	$-$	$+$	$-\pi/4$	$-$	equatorward	$+\pi/4$
star	$+$	$+$	$+\pi/4$	$+$	poleward	$+\pi/4$

where the upper (lower) sign applies when αS is positive (negative). The sign of c gives the direction of propagation of the dynamo wave; see Table 14.1 for a summary of the propagation directions in different settings.

An important property of the $\alpha\Omega$ dynamo solutions that can be read off from the plane wave solutions is the phase shift of $\pm\frac{3}{4}\pi$ (for $S<0$) and $\pm\pi/4$ (for $S>0$) between the poloidal and toroidal fields. It is customary to quote instead the normalized time delay $\lvert\omega_{\rm cyc}\rvert\,\Delta t = \varphi\,{\rm sgn}(c)$, by which the toroidal field lags behind the radial field. These values are given in the last column of Table 14.1. Note that the temporal phase shift only depends on the sign of the shear S and not on α.

Chapter 15

Differential rotation

It became clear from the discussion in § 14 that differential rotation plays an important role in producing a large scale magnetic field in the Sun. It may also be important for the dynamo in disposing of its excess small scale current helicity, as discussed in the previous section. In this section we discuss the theoretical basis for explaining the origin and properties of solar and stellar differential rotation.

15.1 Mean field theory of differential rotation

The origin of differential rotation has long been understood to be a consequence of the anisotropy of convection. It has long been clear that the vertical exchange of momentum by convection should lead to a tendency toward constant angular momentum in the radial direction, i.e. $\overline{\Omega}\varpi^2 = \text{const}$, and hence the mean angular velocity scales with radius like $\overline{\Omega}(r) \sim r^{-2}$. Here, $\varpi = r\sin\theta$ denotes the cylindrical radius (i.e. the distance from the rotation axis).

The $r\phi$ component of the viscous stress tensor contributes to the angular momentum equation,

$$\frac{\partial}{\partial t}\left(\rho\varpi^2\overline{\Omega}\right) + \boldsymbol{\nabla}\cdot\left[\rho\varpi\left(\overline{\boldsymbol{U}}\,\overline{U}_\phi + \overline{\boldsymbol{u}u_\phi}\right)\right] = 0, \tag{15.1}$$

where $\overline{u_i u_j} = Q_{ij}$ are the components of the Reynolds tensor. In spherical coordinates the full mean velocity vector is written as $\overline{\boldsymbol{U}} = (\overline{U}_\varpi, \varpi\overline{\Omega}, \overline{U}_z)$.

The early treatment in terms of an anisotropic viscosity tensor was purely phenomenological. A rigorous calculation of the Reynolds stresses shows that the mean Reynolds stress tensor is described not only by diffusive components that are proportional to the components of the rate of strain tensor of the mean flow, but that there are also non-diffusive components that are directly proportional to the local angular velocity. In particular the $r\phi$ and $\theta\phi$ components of the Reynolds tensor are of interest for driving r and θ gradients of $\overline{U}_\phi \equiv \varpi\overline{\Omega}$. Thus, for ordinary isotropic turbulent viscosity one has, using Cartesian index

notation,

$$Q_{ij} = -\nu_{\rm t}\left(\overline{U}_{i,j} + \overline{U}_{j,i}\right) - \zeta_{\rm t}\delta_{ij}\overline{U}_{k,k}, \tag{15.2}$$

where $\zeta_{\rm t}$ is a turbulent bulk viscosity, and commas denote partial differentiation. This expression implies in particular that

$$Q_{\theta\phi} = -\nu_{\rm t}\sin\theta\frac{\partial\overline{\Omega}}{\partial\theta}. \tag{15.3}$$

Note that for the Sun, where $\partial\overline{\Omega}/\partial\theta > 0$ in the northern hemisphere, this formula would predict that $Q_{\theta\phi}$ is negative in the northern hemisphere. However, it was noted long ago from correlation measurements of sunspot proper motions that $Q_{\theta\phi}$ is in fact *positive* in the northern hemisphere. The observed profile of $Q_{\theta\phi}$ is also known as the Ward profile. The observed positive sign was used to motivate that there must be an additional term in the expression for Q_{ij}. Using a closure approach, such as the first order smoothing approximation that is often used to calculate the α effect in dynamo theory, one can find the coefficients in the expansion

$$Q_{ij} = \Lambda_{ijk}\overline{\Omega}_k - \mathcal{N}_{ijkl}\overline{U}_{k,l}, \tag{15.4}$$

where Λ_{ijk} describes the so-called Λ effect and $\mathcal{N}_{ijkl}$ is the turbulent viscosity tensor. The viscosity tensor $\mathcal{N}_{ijkl}$ must in general be anisotropic. When anisotropies are included, $\mathcal{N}_{ijkl}$ gets modified (but it retains its overall diffusive properties), and Λ_{ijk} takes the form

$$\Lambda_{ijk}\overline{\Omega}_k = \begin{pmatrix} 0 & 0 & V\sin\theta \\ 0 & 0 & H\cos\theta \\ V\sin\theta & H\cos\theta & 0 \end{pmatrix}\overline{\Omega}, \tag{15.5}$$

where V and H are still functions of radius, latitude, and time; V is thought to be responsible for driving vertical differential rotation ($\partial\overline{\Omega}/\partial r \neq 0$) while H is responsible for latitudinal differential rotation ($\partial\overline{\Omega}/\partial\theta \neq 0$).

The first order smoothing approximation predicts the following useful approximations for V and H:

$$V \approx \tau\left(\overline{u_\phi^2} - \overline{u_r^2}\right), \tag{15.6}$$

$$H \approx \tau\left(\overline{u_\phi^2} - \overline{u_\theta^2}\right). \tag{15.7}$$

These expressions show that when the rms velocity in the radial direction is larger than in the azimuthal direction we must expect $V < 0$ and hence $\partial\overline{\Omega}/\partial r < 0$. In the Sun, this effect is responsible for the negative radial shear near the surface where strong downdrafts may be responsible for a comparatively large value of $\overline{u_r^2}$. Likewise, when the rms velocity in the latitudinal direction is larger than in the azimuthal direction we expect $H < 0$ and hence $\partial\overline{\Omega}/\partial\theta < 0$, so the equator would spin slower than the poles. This does not apply to the Sun, but it may be the case in some stars, especially when the flows are dominated by large scale meridional circulation.

15.2 The Λ effect from turbulence simulations

Several of the relationships described above have been tested using convection simulations, both in local Cartesian boxes located at different latitudes as well as in global spherical shells. Generally, the various simulations agree in that the sign of the horizontal Reynolds stress is positive in the northern hemisphere and negative in the southern, reproducing thus the Ward profile. The simulations also show that the off-diagonal components of the turbulent heat transport tensor are mostly positive in the northern hemisphere, and negative in the southern hemisphere. This agrees with the sign required if the baroclinic term is to produce a tendency toward spoke-like angular velocity contours. Simulations also reproduce the sudden drop of angular velocity at the top of the convection zone. This agrees with a predominantly negative sign of the vertical Reynolds stress at a similar depth. Furthermore, some of the more recent simulations show an unexpectedly sharp increase of the horizontal Reynolds stress just near the equator (at around $\pm5^\circ$ latitude), before changing sign right at the equator. The significance of this result for the solar differential rotation pattern is still unclear.

15.3 Meridional flow and the baroclinic term

According to the formalism described in the previous section, a finite differential rotation can be obtained by ignoring meridional flows and solving equation (15.1) in isolation. However, this would only be a poor approximation that becomes quickly invalid when the angular velocity becomes large compared with the turbulent viscous decay rate. This is quantified by the Taylor number

$$\mathrm{Ta} = \left(2\overline{\Omega}_0 R^2/\nu_\mathrm{t}\right)^2. \tag{15.8}$$

Using the first order smoothing expression from Rüdiger (1989), $\nu_\mathrm{t} = (2/15)\,\tau u_\mathrm{rms}^2$, we have for values typical for the Sun (see Table 8.2), i.e. $\nu_\mathrm{t} \approx 10^{12}\mathrm{cm}^2/\mathrm{s}$, $\mathrm{Ta} \approx 10^9$. This value of Ta is rather large so that nonlinearities produce strong deviations from linear theory.

As the value of Ta is increased, the Coriolis force increases, which then drives a meridional flow. This meridional flow first increases with increasing values of Ta, but then it reaches a maximum at $\mathrm{Ta} \approx 3\times10^5$, and later declines with increasing values of Ta. (The solar value is $\mathrm{Ta} \approx 3\times10^7$.) This decline is because eventually the Coriolis force can no longer be balanced against advection or diffusion terms. This can best be seen by considering the curl of the momentum equation,

$$\frac{\partial\overline{W}_\phi}{\partial t} + \varpi\overline{\boldsymbol{U}}\cdot\boldsymbol{\nabla}\left(\frac{\overline{W}_\phi}{\varpi}\right) - \nu_\mathrm{t}\mathrm{D}^2\overline{W}_\phi = \varpi\frac{\partial\overline{\Omega}^2}{\partial z} + \hat{\boldsymbol{\phi}}\cdot\overline{\boldsymbol{\nabla}T\times\boldsymbol{\nabla}S}. \tag{15.9}$$

We recall that we consider here a nonrotating frame of reference, so there is no Coriolis force. Nevertheless, part of the inertial term takes a form that is quite

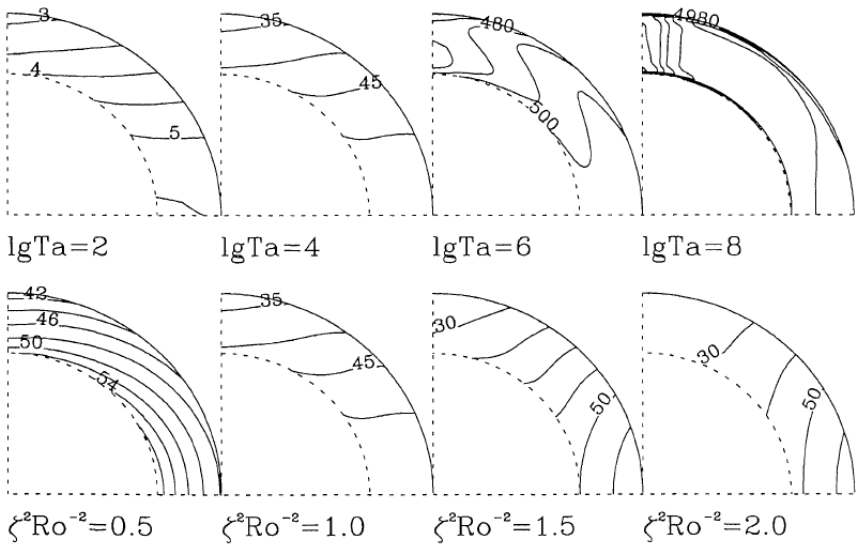


Figure 15.1: Contours of constant $\overline{\Omega}$ for different values of Taylor number (upper panel) and different values of the inverse Rossby number, affecting the relative importance of H over V (lower panel). [Adapted from Brandenburg et al. (1990).]

similar to the Coriolis term, but here $\overline{\Omega}$ is a function of position, while in the Coriolis term the angular velocity would normally be a constant.

In the barotropic case one has $\boldsymbol{\nabla} T \parallel \boldsymbol{\nabla} S$ so there is no baroclinic term, i.e. $\hat{\boldsymbol{\phi}} \cdot (\boldsymbol{\nabla} T \times \boldsymbol{\nabla} S) = 0$. So, if viscous and inertial terms are small, which is indeed the case for rapid rotation, then $\partial \overline{\Omega}^2 / \partial z$ has to vanish, so $\overline{\Omega}$ would be constant along cylinders; see Fig. 15.1. It is generally believed that the main reason for $\overline{\Omega}$ not having cylindrical contours in the Sun is connected with the presence of the baroclinic term. The presence of magnetic forces may also play a role, but unlike the baroclinic term, magnetic forces tend to produce a rather variable $\overline{\Omega}$ patterns, often connected with rapid motions near the poles where the inertia is lower.

Currently the highest resolution simulations of global convection in spherical shells are those by Miesch et al. (2000). These simulations show a great amount of detail and reproduce some basic features of the Sun's differential rotation such as the more rapidly spinning equator. However, in low latitudes they show strongly cylindrical $\overline{\Omega}$ contours that deviate markedly from the more spoke-like contours inferred for the Sun using helioseismology. These simulations also do not show the near-surface shear layer where the rotation rate drops by over 20 nHz over the last 30 Mm below the surface.

Mean field simulations using the Λ effect show surprisingly good agreement

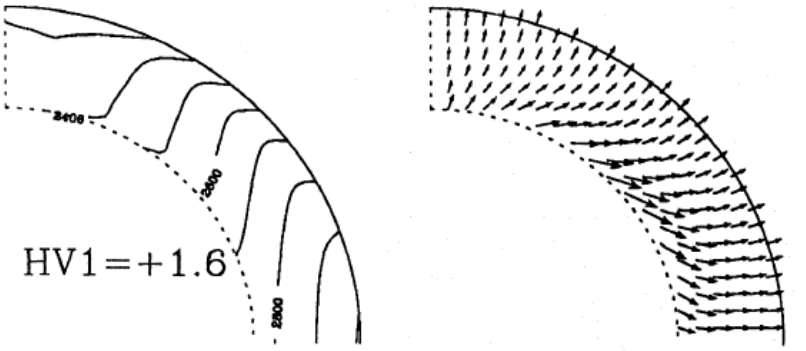


Figure 15.2: Contours of angular velocity (left) and turbulent convective energy flux (right) for a model with anisotropic heat transfer tensor. [Adapted from Brandenburg et al. (1992).]

with the helioseismologically inferred $\overline{\Omega}$ pattern, and they are also beginning to address the problem of the near-surface shear layer. In these simulations it is indeed the baroclinic term that is responsible for causing the departure from cylindrical contours. This, in turn, is caused by an anisotropy of the turbulent heat conductivity which causes a slight enhancement in temperature and entropy at the poles. In the bulk of the convection zone the entropy is nearly constant, so the radial entropy variation is smallest compared with the radial temperature variation. It is therefore primarily the latitudinal entropy variation that determines the baroclinic term, with

$$\varpi\frac{\partial\overline{\Omega}^2}{\partial z}\approx-\hat{\boldsymbol{\phi}}\cdot\overline{\boldsymbol{\nabla}T\times\boldsymbol{\nabla}S}\approx-\frac{1}{r}\frac{\partial\overline{T}}{\partial r}\frac{\partial\overline{S}}{\partial\theta}<0. \tag{15.10}$$

The inequality shows that negative values of $\partial\overline{\Omega}^2/\partial z$ require that the pole is slightly warmer than the equator ($\partial\overline{S}/\partial\theta<0$). However, this effect is so weak that it cannot at present be observed. Allowing for these conditions in a simulation may require particular care in the treatment of the outer boundary conditions. In Fig. 15.2 we show the plots of angular velocity contours and convective energy transport in a model with anisotropic turbulent conductivity tensor, χ_{ij}. Given that the flux, $\boldsymbol{F}$, is proportional to $-\chi_{ij}\nabla_j\overline{S}$, a negative $\partial\overline{S}/\partial\theta$ can be produced from a positive F_r with a positive value of $\chi_{r\theta}$.

In the discussion above we ignored in the last step a possible correlation between entropy and temperature fluctuations, i.e. a contribution from the term $\overline{\boldsymbol{\nabla}T'\times\boldsymbol{\nabla}S'}$ where primes denote fluctuations. Such correlations, if of suitable sign, might provide yet a further explanation for a non-zero value of $\partial\overline{\Omega}^2/\partial z$.

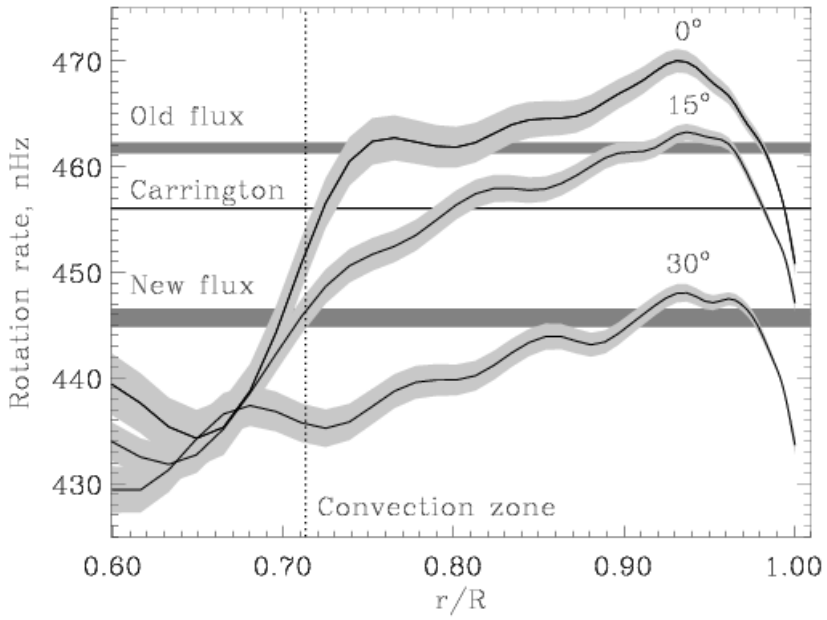


Figure 15.3: Radial profiles of the internal solar rotation rate, as inferred from helioseismology (sidereal, i.e. in a fixed frame). The rotation rate of active zones at the beginning of the cycle (at $\approx 30°$ latitude) and near the end (at $\approx 4°$) is indicated by horizontal bars, which intersect the profiles of rotation rate at $r/R_\odot \approx 0.97$. For orientation, the conventionally defined Carrington rotation period of 27.3 days (synodic value, corresponding to 424 nHz) has been translated to the sidereal value of 456 nHz. Courtesy of Benevolenskaya et al. (1999).

15.4 Near-surface shear layer

The first results of helioseismology indicated significantly higher angular velocities in the sub-surface than what is seen at the surface using Doppler measurements. This apparent conflict is now resolved in that helioseismological inversions of the data from the SOHO spacecraft show a sharp negative gradient, connecting the observed surface values smoothly with the local maximum of the angular velocity at about 35 Mm depth; see Fig. 15.3.

The theory of this negative near-surface shear layer is still a matter of ongoing research, but it is clear that negative shear would generally be the result of predominantly vertical turbulent velocities such as strong downdrafts near the radiating surface. However, such a layer that is dominated by strong downdrafts was only thought to be several megameters deep, and not several tens

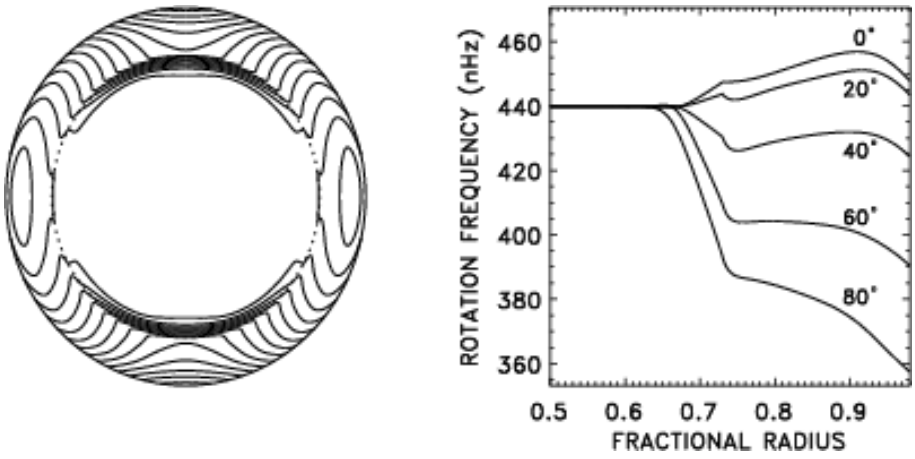


Figure 15.4: Rotation law obtained by Kitchatinov & Rüdiger (2005) taking the anisotropy of the turbulence near the surface into account. [Courtesy Kitchatinov & Rüdiger (2005).]

of megameters. With an improved theory for the anisotropy of the turbulence especially near the surface layers, one obtains a clear radial decline of the local angular velocity near the surface, although still not quite as much as is observed; see Fig. 15.4. In any case, these results do at least reproduce the near-surface shear layer qualitatively correctly. A proper understanding of this layer is now quite timely in view of the fact that near-surface shear is likely to contribute to the production of strong toroidal fields.

15.5 Magnetic effects

Using photospheric Doppler measurements and helioseismology so-called torsional oscillations were discovered in the 1980ies. They are a cyclic modulation of the latitudinal profile of the angular velocity at the surface of the sun. Model calculations suggest that these oscillations can well be modeled by restoring the Lorentz force by adding a term $-\varpi\overline{\boldsymbol{B}}\overline{B}_\phi$ under the divergence in equation (15.1). Unfortunately, given that there is no definitive solar dynamo model, models for the Sun's torsional oscillations are equally preliminary and still a matter ongoing research.

Chapter 16

Final words

In the past few decades there have been significant developments in understanding the physics of the Sun. Even regarding the radial structure of the Sun, which was thought to be qualitatively well understood, major revisions have emerged just recently with the refinement of three-dimensional simulations of solar granulation. Such simulations have led to new spectral line fits that imply a drastically reduced abundance of the heavier elements. This has consequences for the opacities that affect the deep parts of the Sun's interior.

There are many aspects of solar physics where a detailed understanding of the three-dimensional flow pattern of the Sun is crucial. It is not surprising that effects involving details of the turbulent flow field in the solar convection zone, such as the theory of differential rotation and magnetic field generation, provide other examples where the three-dimensional dynamics is important. Fully three-dimensional simulations of solar convection with magnetic fields produce flow and magnetic field structures in great detail, but at present they deviate in some important aspects from the Sun (e.g. the fraction of small scale to large scale field is rather large; and the angular velocity contours are still too strongly aligned with the rotation axis). Some tentative explanations are available (magnetic Prandtl number not small enough in the simulations to reduce or even suppress small scale dynamo action, and surface conditions not realistic enough to allow for sufficiently large a baroclinic term). Future advances in computer technology will bring a steady increase in numerical resolution. However, increase of spatial resolution by a factor of two will always be very difficult when close to the machine capacity. Substantial progress may rather hinge on new insights that may emerge from a closer interrelation between local simulations where turbulence is well resolved and mean field calculations that benefit from input and calibration of detailed simulations.