AN INTRODUCTION TO HYDRODYNAMICS

Jonathan Braithwaite

March 2011

Preface

These notes accompany a course in hydrodynamics taught in Bonn to Master's students. The course has been designed for students who have completed an undergraduate degree in physics and are familiar with the basics of thermodynamics and vector calculus. As the vast majority of undergraduate courses do not include fluid mechanics, no prior knowledge of the subject is assumed. Also not assumed is any knowledge of astrophysics, so that this course is accessible to students studying for a Master's degree in physics. However there is a definite bias towards applications in astrophysics and geophysics, as opposed to applications in engineering – in practice this means that topics given much attention elsewhere, such as boundary layers, pipe flow, and aerodynamics are mentioned only relatively briefly in this course. Many of the principles are illustrated using examples from everyday experience, as in this way the student can develop an intuitive understanding which can then be applied in other contexts. For instance, reference is made to the hydraulic shock formed as water from a tap spreads out across the surface of a wash basin as a connection to the phenomenon of astrophysical shocks. Phenomena in atmospheric physics are also used as a bridge between terrestrial intuition and the astrophysical context. In fact, atmospheric fluid mechanics has a longer history than astrophysical fluids and can be considered more 'advanced'; astrophysicists are well advised to learn from this neighbouring field to avoid reinventing the wheel.

At the moment these notes contain little or no material on shocks, convection or turbulence. At present, these topics are taught as a part of this course in two lectures by M. Cantiello.

The last quarter of the course (the last two chapters of these notes) concerns magnetohydrodynamics (MHD), essentially an extension of hydrodynamics to electrically conducting fluids; here we cannot draw on terrestrial intuition and must rely purely on theory. These chapters are 'stand alone' in that they can be read without the rest of the notes.

Jon Braithwaite
Bonn, October 2010

Suggested reading:

- L.D. Landau & E.M. Lifshitz: Fluid mechanics (Pergamon Press, 2nd edition, 1987)
- S.N. Shore: Astrophysical hydrodynamics: an introduction (Wiley-VCH, 2nd edition, 2007)
- A.R. Choudhuri: The physics of fluids and plasmas (Cambridge University Press, 1998)
- J. Pedlosky: Geophysical fluid dynamics (Springer, 2nd edition, 1990)

- P.H. Roberts: An introduction to magnetohydrodynamics (American Elsevier, 1967)
- H.C. Spruit: Essential magnetohydrodynamics for astrophysics (http://arxiv.org/abs/1301.5572)

Contents

Chapter 1

Introduction

We first examine what is meant by a fluid before deriving the equations of motion.

1.1 The fluid approximation

The ancient Greeks amongst others debated over two thousand years ago whether matter is made from discrete particles or is a continuum, divisible ad infinitum. This question was not properly resolved until well into nineteenth century (Brownian motion, etc.), by which time useful theories of thermodynamics had already been developed, driven largely by the need to build more efficient steam engines. Therefore, it is not necessary to think about particles in order to understand thermodynamics; it is just necessary to accept a small number of experimentally-supported axioms (the laws of thermodynamics) and the rest of classical thermodynamics follows. Later, when statistical mechanics was developed, it become possible to understand where the laws of thermodynamics come from, in terms of more fundamental physics. However, for practical purposes this is unnecessary and complicates matters.

The same is true of hydrodynamics, the study of fluid flow, which was also developed prior to the conclusion of the atom vs. continuum debate. In many situations it is sufficient to treat a fluid as a continuous substance. Now that we know that fluids are made of particles, we can explain some fluid phenomena in terms of more fundamental physics, for instance we can predict the viscosity of a gas (a macroscopic quantity) by consideration of particles, mean-free paths and so on. However, in this course, we shall cover *classical* hydrodynamics, meaning without consideration for the particle nature of matter, making only occasional reference to particles. Before deriving the equations of hydrodynamics, it is useful to look at the *fluid approximation* and its limitations so that we know not to try to use hydrodynamics where it does not apply.

In the fluid approximation, we treat the ensemble of particles as a single fluid. To describe an ensemble of particles precisely we need to know the position and velocity of each particle (ignoring quantum mechanics); if the number of particles is large enough to perform statistics then it makes sense to describe the ensemble with a distribution function n:

$$\delta N(t) = N(\mathbf{r}, \mathbf{u}, t)\, \delta\mathbf{r}^3\, \delta\mathbf{u}^3 \tag{1.1}$$

where δN is the number of particles in a small volume in position/velocity space at time t; $\mathbf{r}$ is the space coordinate (a vector with as many components as the space has dimensions) and $\mathbf{u}$ is the velocity. A small volume in physical space (i.e. $\delta\mathbf{r}^3 = \delta x\,\delta y\,\delta z$) can contain particles with completely different velocities. In contrast to this, the fluid approximation describes the system in the following way. First we integrate N over all velocity space to obtain a space density $n(\mathbf{r}, t)$, and then we introduce a mean velocity $\bar{\mathbf{u}} = \bar{\mathbf{u}}(\mathbf{r}, t)$, the mean velocity of the particles[1] at position $\mathbf{r}$. This is arrived at by integrating $N(\mathbf{r}, \mathbf{u}, t)\,\mathbf{u}$ over velocity space and dividing by $n(\mathbf{r}, t)$. [Hereafter the bar on the mean velocity is dropped.] Obviously in doing this we have lost all information about the spread of particle velocities about the mean. However, we can make up for this by noting that if we have local thermodynamic equilibrium (LTE), there is only one degree of freedom in the spread of velocities which we characterise with temperature $T = T(\mathbf{r}, t)$.

In assuming LTE, we are assuming in effect that the particles in some small volume are able come into equilibrium with each other via collisions, rather than wandering larger distances before this has been achieved; only in this way can temperature be defined locally. For this condition to hold, it is necessary that the mean free path of the particles is significantly less than any other length scales of interest to us. For instance, in the Earth's atmosphere the mean free path is of order 10^{-5}cm while the smallest length scales of interest to us in weather forecasting are perhaps 100m, so that we may safely treat air as a fluid. In some contexts the fluid approximation is not applicable, for instance in the solar wind where the mean free path of protons is 10^{15}cm $\approx$ 20AU (an astronomical unit is the distance between the Sun and the Earth). This example brings us onto another point: in the fluid approximation we are assuming that all particle species making up a fluid are in LTE amongst themselves and with each other. In other words, all species have the same velocity $\mathbf{u}$ and temperature T at any point in space and time. In the solar wind, the electrons have a significantly shorter mean free path and may come into thermal equilibrium with each other while the protons can still be considered collisionless. A proper study of these phenomena is outside the scope of this course; the interested student should consider taking a course in plasma physics.

Finally it is worth noting that the equations of hydrodynamics which we derive using the fluid approximation can sometimes predict situations which violate the applicability of the approximation. A good example of this is shocks (section 6) – the fluid equations predict in some circumstances the appearance of discontinuities in the fluid quantities such as $\mathbf{u}$ and T. The relevant length scale in the fluid has gone to zero, which is clearly less than the mean free path and violates the fluid approximation! Fortunately there is a way out of this apparently unpleasant predicament without completely abandoning the fluid picture. [In reality, the discontinuity has a thickness roughly equal to the mean free path.]

1.2 The hydrodynamic equations

In this section the equations of hydrodynamics are derived.

We know from thermodynamics that the state of a fluid can be described in terms of a number of 'functions of state', which in a simple fluid is two, for instance pressure and temperature; all other variables, for instance density or entropy, can be found from the equation of state. In the following we use a simple fluid, but the equations can easily be generalised to include more complex fluids such as a fluid in which the mean molecular weight is not fixed, which one encounters sometimes in astrophysics, or the salinity in an ocean or water vapour concentration in the atmosphere, for example. Note that these quantities

[1] If the particles do not have uniform mass, we take a mass-weighted mean. This ensures that the resulting equations respect conservation of momentum.

are called *intensive variables* as they can be defined and measured at any particular point in space, as opposed to *extensive variables* such as volume or mass which are properties of a whole system. In addition to these functions of state, in a fluid flow we also need the velocity $\mathbf{u}$ for a complete description. The velocity and the thermodynamic variables are functions of position $\mathbf{r}$ and time t.

There are three equations of hydrodynamics, which come from the conservation of momentum, mass and energy. They are partial differential equations containing the time derivatives of the velocity and the two thermodynamics variables. First of all, the application of Newton's second law to a fluid element of volume δV gives us:

$$\rho\,\delta V \frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t} = \delta\mathbf{F} \tag{1.2}$$

where ρ is the density of the fluid and $\delta\mathbf{F}$ is the force on the fluid element. Dividing by δV and splitting the right-hand side up into different types of force we have

$$\begin{aligned}\rho\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t} &= \mathbf{F}_{\mathrm{body}} + \mathbf{F}_{\mathrm{surface}} \\ &= \rho\mathbf{g} - \nabla P + \mathbf{F}_{\mathrm{visc}} \end{aligned} \tag{1.3}$$

where the terms on the right hand side now represent various forces per unit volume. These forces fall into two classes. First there are *body forces* such as gravity ($\mathbf{g}$ is the local gravitational force per unit mass). In section 7 we look at effects of the Coriolis force present in any fluid in a rotating frame of reference. In ionised gases there is also generally an electromagnetic body force. Secondly the *surface forces*, where the force on a fluid element comes from its immediate neighbours: the pressure gradient force, present in all fluids, and the viscous force. One can consider that the pressure is defined (apart from some additive constant) by this equation. Alternatively pressure is defined in a non-viscous fluid as the force per unit area exerted by a fluid element on its neighbours; the net force per unit volume appearing above is found by equating $\oint P\,\mathrm{d}\mathbf{S} = \int \nabla P\,\mathrm{d}V$. In a viscous fluid the force exerted by an element on its neighbours is generally not the same in all directions and the average is not necessarily equal to P; the definition of pressure in this case is less straightforward – see section 5. Finally, note that the derivative on the left-hand side of (1.3), $\mathrm{d}/\mathrm{d}t$, is the Lagrangian (co-moving) derivative, which is related to the Eulerian (stationary) derivative $\partial/\partial t$ in the following way. Remembering that an infinitesimal change δf in a function $f(x,t)$ can be expressed as

$$\delta f = \left(\frac{\partial f}{\partial t}\right)_x \delta t + \left(\frac{\partial f}{\partial x}\right)_t \delta x, \tag{1.4}$$

we can express the rate of change of any quantity $q(\mathbf{r},t)$ in a fluid element moving with velocity $\mathbf{u}$ as

$$\frac{\mathrm{d}q}{\mathrm{d}t} \equiv \frac{\partial q}{\partial t} + \mathbf{u}\cdot\nabla q; \tag{1.5}$$

in other words, the co-moving rate of change of a quantity in a particular fluid element momentarily located at $\mathbf{r}$ is equal to the rate of change fixed at that location $\mathbf{r}$ plus the spatial derivative in the direction of the fluid velocity multiplied by the magnitude of the fluid velocity. Note that often a capital D is used for the Lagrangian derivative instead of d.

Now we use conservation of mass to derive the second equation. Imagining a volume V with boundary S, the rate of change of mass in the volume is equal to the mass flux $\rho\mathbf{u}$ into the volume through the boundaries, giving

$$\frac{\partial}{\partial t}\int \rho \; \mathrm{d}V \;\; = \;\; -\oint \rho \mathbf{u}\cdot \mathrm{d}\mathbf{S}, \tag{1.6}$$

$$\int \frac{\partial \rho}{\partial t} \; \mathrm{d}V \;\; = \;\; -\int \boldsymbol{\nabla}\cdot(\rho\mathbf{u}) \; \mathrm{d}V, \tag{1.7}$$

where the second line follows from Gauss' theorem. Taking the time derivative inside the integrand, and noting that this relation is valid for any volume V, gives us the usual form of the continuity equation

$$\frac{\partial \rho}{\partial t} = -\boldsymbol{\nabla}\cdot(\rho\mathbf{u}). \tag{1.8}$$

Often, we wish to have this equation in a form containing the Lagrangian derivative. Using (1.5), we have

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = -\rho\boldsymbol{\nabla}\cdot\mathbf{u}. \tag{1.9}$$

So far, we have two equations (1.3) and (1.8) and three unknowns $\mathbf{u}$, ρ and P. To close this set, one option is to find some way of directly relating ρ to P without involving any new variables. This is known as a 'barotropic' equation of state where $\rho = \rho(P)$. A special case is to assume a constant density: ρ = const, so that ρ can be replaced by a constant ρ_0 in (1.3) and then (1.8) reduces to $\boldsymbol{\nabla}\cdot\mathbf{u} = 0$. However, we often have two or more independent thermodynamic variables and the equation of state of the fluid is expressible as $\rho = \rho(P, X_1, X_2...)$ where X_1, X_2 are some other thermodynamic variables. For instance, the ideal gas equation of state, often applicable in astrophysics, is $P = \rho RT/\mu$ where R is the gas constant, T is temperature and μ is mean molecular weight (in atomic mass units, approximately equal to the hydrogen atom mass). Having introduced two more thermodynamic variables, we need two extra equations to close the set – these equations describe the evolution of thermal energy and of chemical composition. Often (and everywhere in this course) the latter is simply μ = const; the former is now derived.

We know that $\mathrm{d}U = \mathrm{d}Q - P\mathrm{d}V$ as a standard result of thermodynamics[2] so that if we consider a unit mass of the fluid, which has a volume $1/\rho$ and internal energy ϵ,

$$\mathrm{d}\epsilon = \mathrm{d}Q - P\mathrm{d}(1/\rho), \tag{1.10}$$

where $\mathrm{d}Q$ is heat energy is deposited into the fluid from an as yet unspecified source. Making these into time derivatives we have

$$\frac{\mathrm{d}\epsilon}{\mathrm{d}t} = \frac{Q}{\rho} - \frac{P}{\rho}\boldsymbol{\nabla}\cdot\mathbf{u} \tag{1.11}$$

where (1.9) has been used. Q has units energy per unit volume per unit time. Of course a new variable ϵ has been introduced but the fluid has just two independent thermodynamic variables and every other variable can be expressed as a function of those two, including $\epsilon = \epsilon(P, \rho)$; in the case of an ideal gas $\epsilon = P/\rho(\gamma - 1)$ where $\gamma = c_p/c_v$ is the ratio of specific heats, equal to 5/3 in a monatomic gas. In this case, with some rearrangement the energy equation can be written as

$$\frac{\mathrm{d}P}{\mathrm{d}t} = (\gamma - 1)Q - \gamma P\,\boldsymbol{\nabla}\cdot\mathbf{u}. \tag{1.12}$$

[2]The first law of thermodynamics is $\mathrm{d}U = \mathrm{d}Q + \mathrm{d}W$. In this context we can equate the work done on the gas to $-P\mathrm{d}V$ only because the change in volume is reversible, which means that a given fluid element is pushing against neighbouring elements with the same pressure at which the neighbouring elements are pushing back, rather than the fluid element being allowed to expand into a neighbouring vacuum, for instance, where $-P\mathrm{d}V$ is non-zero but work $\mathrm{d}W = 0$. In section 5 we consider viscous, irreversible fluid flow.

Heat rate Q could contain contributions from thermal conduction, viscous heating, dissipation of electric currents, nuclear energy generation, release of latent heat, radiative cooling, and so on, and it should be possible to express Q as a function of the other variables; for instance nuclear energy generation can be expressed as a function of density, pressure and the abundance of the relevant chemical species.

Note that there are various ways of writing down the energy equation. Since specific entropy s can also be expressed in terms of the two thermodynamic variables we have already introduced: $s = s(P, \rho)$, as can temperature $T = T(P, \rho)$, we can also express the energy equation as

$$\frac{\mathrm{d}s}{\mathrm{d}t} = \frac{Q}{\rho T} \tag{1.13}$$

to complete the set.

So, we now have three partial differential equations containing three unknowns $\mathbf{u}$, P and ρ. To begin with, rather than trying to solve the general equations we shall make some simplifications. In the next section we look at 'ideal fluids', which have no diffusion of momentum or heat, and no additional heating from any source. This means that $Q = 0$ and $\mathbf{F}_{\mathrm{visc}} = \mathbf{0}$.

Exercises

1.1 A different form of the energy equation

Assuming the ideal gas equation of state, express the energy equation with $\mathrm{d}T/\mathrm{d}t$ on the left-hand side and P, ρ, $\mathbf{u}$ and Q on the right.

Chapter 2

Ideal fluids: basic concepts

The equations governing the motion of an ideal fluid are:

$$\begin{aligned}
\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \boldsymbol{\nabla})\mathbf{u} &= -\frac{1}{\rho}\boldsymbol{\nabla} P + \mathbf{g}; && (2.1)\\
\frac{\partial \rho}{\partial t} &= -\boldsymbol{\nabla} \cdot (\rho \mathbf{u}); && (2.2)\\
\frac{\mathrm{d}s}{\mathrm{d}t} &= 0. && (2.3)
\end{aligned}$$

These are often called the Euler equations. In the third equation, s is the specific entropy; this form may be obtained by setting $Q = 0$ in (1.12).

2.1 Hydrostatics

The condition that the fluid remains stationary is:

$$\rho \mathbf{g} - \boldsymbol{\nabla} P = 0. \qquad (2.4)$$

This relation can be used, for instance, to determine the structure of the atmosphere. Taking gravity to be directed downwards, we have the well-known equation of hydrostatic equilibrium

$$\frac{\mathrm{d}P}{\mathrm{d}z} = -\rho g, \qquad (2.5)$$

where the z axis points upwards. In an isothermal gas atmosphere this equation is easily integrated, using the equation of state $P = \rho R_\mu T$ where $R_\mu \equiv R/\mu$, which is often used where mean molecular weight μ is constant. The solution is

$$P = P_0 \exp\left(-\frac{z}{H_P}\right) \qquad \text{where} \qquad H_P \equiv \left(\frac{\mathrm{d}\ln P}{\mathrm{d}z}\right)^{-1} = \frac{R_\mu T}{g}. \qquad (2.6)$$

The definition of the pressure scale height H_P is valid also for non-isothermal atmospheres. In the Earth's atmosphere, its value is around 8 km, approximately equal to the height of Mount Everest.

2.2 Bernoulli's equation

Let us consider a steady flow, i.e. a flow where $\partial/\partial t = 0$ (but the co-moving derivatives are in general non-zero). In a steady flow the *streamlines*, which are defined as those lines which are everywhere along their length tangential to the velocity, are also the paths of individual fluid elements (called *path lines*). Consider a volume bounded at the sides by streamlines at each end and by a surface perpendicular to the flow – the fluid is flowing into the volume at one end and exiting at the other end. Denoting quantities at the inflow and outflow end of the volume with 0 and 1, the rate of change of energy in this volume is given by the difference between the energy entering and exiting and the difference between the $P\,\mathrm{d}V$ work done by the fluid at the ends on the fluid ahead of it, and must vanish in a steady flow:

$$A_0 u_0 \varepsilon_0 \rho_0 - A_1 u_1 \varepsilon_1 \rho_1 + A_0 u_0 P_0 - A_1 u_1 P_1 = 0 \tag{2.7}$$

where ε is total energy per unit mass, the sum of internal, kinetic and gravitational potential:

$$\varepsilon \equiv \epsilon + \frac{1}{2}u^2 + \Phi \tag{2.8}$$

and A and u are the cross-sectional area of, and the velocity at, the ends. The mass in the volume must also be constant in time, so that $A_0 u_0 \rho_0 = A_1 u_1 \rho_1$. Substituted back into (2.7) this gives:

$$\varepsilon_0 + \frac{P_0}{\rho_0} = \varepsilon_1 + \frac{P_1}{\rho_1}, \tag{2.9}$$

which is valid for any volume in the flow, so that we can say more generally that

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\varepsilon + \frac{P}{\rho}\right) = 0. \tag{2.10}$$

Physically this represents just conservation of energy. This condition holds only in an ideal fluid, where no energy can be transferred between neighbouring fluid elements *across* the streamlines since the component of the pressure gradient perpendicular to the streamlines is of course perpendicular to the velocity. [Note that unlike energy and mass, momentum *can* be transferred across streamlines.] In a viscous fluid this is no longer true: energy can be transferred by viscous stress; also heat conduction can transfer energy perpendicular to the flow.

The same result can be derived in a different way, providing us with a more intuitive understanding. The momentum equation (2.1) equates the acceleration of the fluid to the force per unit mass, and taking the dot product with $\mathbf{u}$ equates the rate of change of kinetic energy to the rate at which work is done by the various forces:

$$\begin{aligned} \mathbf{u}\cdot\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t} &= -\mathbf{u}\cdot\left(\frac{1}{\rho}\boldsymbol{\nabla}P\right) - \mathbf{u}\cdot\boldsymbol{\nabla}\Phi, \\ \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{2}u^2\right) &= -\mathbf{u}\cdot\boldsymbol{\nabla}\left(\frac{P}{\rho}\right) + P\mathbf{u}\cdot\boldsymbol{\nabla}\left(\frac{1}{\rho}\right) - \frac{\mathrm{d}\Phi}{\mathrm{d}t} \end{aligned} \tag{2.11}$$

where the pressure gradient term has been broken up into two parts, the first of which is simply (minus) the Lagrangian time derivative of P/ρ; the second part is pressure times the Lagrangian derivative of specific volume, which in an ideal fluid (where there is no viscous or other heating and no conduction of heat) can be equated to the rate of PdV work done on a unit mass – see (1.10). The gravity term is simply the Lagrangian derivative of the specific potential energy. Collecting terms we have therefore

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{2}u^2 + \epsilon + \Phi + \frac{P}{\rho}\right) = 0. \tag{2.12}$$

This demonstrates that a change in velocity along a streamline is produced by a pressure gradient along that streamline.

We shall see in due course that Bernoulli's equation is an incredibly useful form of the principle of energy conservation. Sometimes this equation appears with enthalpy $h \equiv \epsilon + P/\rho$, reducing the number of terms by one. Note that in an ideal gas it can be seen from the equation of state that ϵ and P/ρ are related simply by $P/\rho = (\gamma - 1)\epsilon$ where γ is the ratio of specific heats, so that $h = \gamma\epsilon$.

One device which can be easily understood with the help of this equation is the Venturi meter which measures the flow of air through a pipe: see fig. 2.1. Another is the calculation of the flow of water out of a hole in a barrel.

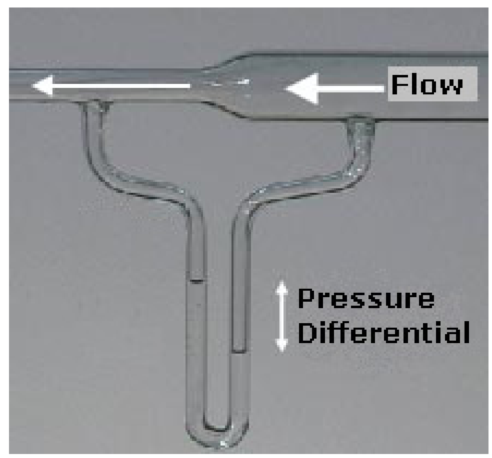


Figure 2.1: A Venturi meter, which measures the flow of air through a pipe. As the flow is constricted the velocity increases, meaning that P/ρ must decrease. Since this is an adiabatic process we know that P/ρ^γ is constant; given that γ is always greater than unity we see that pressure must drop through the constriction. Note that the pressure and velocity difference change sign if the gas is moving supersonically through the pipe (see section 3.1), i.e. the velocity goes *down* as the fluid enters the constriction.

The interpretation of various phenomena is however not always as straightforward as it seems and it is often easier to go back to the momentum equation, considering the acceleration of, and the forces on, a fluid element. A good example of this is the aeroplane wing: it is often said that the pressure above the wing is lower than that below the wing because looking at the streamlines it is obvious that the air above the wing has further to travel and must therefore be moving faster, implying a lower pressure. This is misleading because there is no reason that the air flowing above the wing must meet up again with its former neighbour so that it does not necessarily have to travel faster. To see the flawed argument, consider the lift generated by a thin curved aerofoil which is tilted with respect to the flow so that the air is deflected downwards: here the length of the streamlines above and below are the same, and yet the aerofoil still generates lift. A good example of this kind of aerofoil is the sail of a boat in the situation when the boat is travelling upwind or perpendicular to the wind. How is the lift generated? The easiest explanation comes from consideration of the acceleration perpendicular to the streamlines: below the aerofoil the airflow must curve downwards. The only thing which can produce this acceleration is a pressure gradient perpendicular to the streamlines, namely such that the pressure near the aerofoil is higher than that further away. Above the aerofoil the fluid must also accelerate downwards so that the pressure just above the aerofoil must be lower than that further away. Since the pressure further away tends towards the ambient pressure, the pressure just above the wing must be lower than that just below it, accounting for the lift. Alternatively, one can think of the pressure changes as arising from the inertia of the oncoming fluid. Either way, Bernoulli's equation does not provide us with any quick explanation. See fig. 2.2 for an illustration of the two types of aerofoil. We see that the lift comes essentially from the deflection of air downwards – indeed how is the aerofoil to gain momentum upwards if the air is not given downwards momentum?

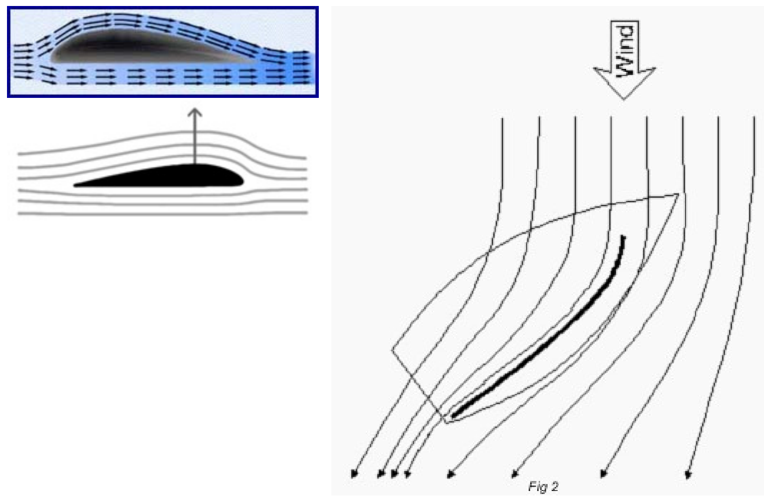


Figure 2.2: *Left:* two examples of a commonly seen but misleading diagram of the Bernoulli effect producing lift on an aerofoil (from the NASA website and wikimedia.org). The streamlines are longer over than under the aerofoil, apparently resulting in a velocity and therefore pressure difference. However, it is in fact not at all obvious why the flow should be faster over the wing than under it. *Right:* the flow of air around the sail of a boat. This time, the streamlines immediately either side of the sail are of equal length, and yet we still have a velocity and pressure difference on either side.

2.3 Sound waves

The most basic kind of wave in a fluid is the sound wave, which can propagate in any compressible fluid. Consider a system where variables vary only in the x direction, so that any wave must also propagate in that direction (we call this a plane wave), and the velocity in the x direction is u; in addition there is no gravity. The momentum and mass conservation equations can be written:

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = -\frac{1}{\rho}\frac{\partial P}{\partial x}, \tag{2.13}$$

$$\frac{\partial \rho}{\partial t} = -\frac{\partial}{\partial x}(\rho u). \tag{2.14}$$

In order to look at the properties of low-amplitude, linear waves, it is necessary to linearise these equations. To do this, we first define some background equilibrium state with pressure P_0, density ρ_0 and zero velocity, and write pressure $P = P_0 + \delta P$, $\rho = \rho_0 + \delta\rho$ where $\delta P \ll P_0$ and $\delta\rho \ll \rho_0$. It follows that the velocity u is also small. In addition we need some relation between the pressure and density perturbations: we define $c_s^2 \equiv \partial P/\partial\rho$. The linearised equations are

$$\frac{\partial u}{\partial t} = -\frac{1}{\rho_0}\frac{\partial(\delta P)}{\partial x}, \tag{2.15}$$

$$\frac{\partial(\delta\rho)}{\partial t} = -\rho_0 \frac{\partial u}{\partial x}, \tag{2.16}$$

$$\delta P = c_s^2 \delta\rho, \tag{2.17}$$

where only first-order terms have been retained. At this juncture it is worthwhile examining the condition under which this linearisation is valid. Looking at the ratio of the two terms on the left-hand side of (2.13), we see that we can ignore the second provided that

$$\frac{u}{\tau} \gg \frac{u^2}{L} \qquad \Rightarrow \qquad L \gg a \tag{2.18}$$

where τ is the timescale, i.e. period of the waves, and $a \sim u\tau$ is the amplitude. This condition is easily shown to be equivalent to the conditions $\delta P \ll P_0$ and $\delta\rho \ll \rho_0$ introduced above.

Substituting for $\delta\rho$ in (2.16), differentiating w.r.t. t and combining with (2.15) gives the wave equation

$$\frac{\partial^2(\delta P)}{\partial t^2} = c_s^2 \frac{\partial^2(\delta P)}{\partial x^2}. \tag{2.19}$$

The solution to this is

$$\delta P = A(x - c_s t) + B(x + c_s t). \tag{2.20}$$

The two terms represent waves travelling in the positive and negative x-directions, respectively. If we consider only waves travelling in the positive direction (i.e. $B = 0$), then whatever the form of δP at time $t = 0$ is preserved in shape but is shifted a distance $c_s t$ at a later time t, i.e. it moves with speed c_s. In an ideal gas, the sound speed $c_s = \sqrt{\gamma P/\rho} = \sqrt{\gamma RT/\mu}$, which follows from the assumption that the motion is adiabatic and the fluid elements have constant entropy. In the Earth's atmosphere it is equal to around 300 m s^{-1}; in the interstellar medium at 10^4 K, $c_s \approx 10$ km s^{-1}.

Alternatively we could have dealt with (2.19) by assuming a solution of the form $\delta P = Ae^{i(kx-\omega t)}$, giving a dispersion relation of

$$\omega = \omega(k) = kc_s. \tag{2.21}$$

The phase speed ω/k and group speed $\partial\omega/\partial k$ are both equal to c_s, which does not depend on the frequency. This non-dependence of speed on frequency we call *non-dispersiveness*. We return to this topic in the following sections.

2.4 Compressibility

In many situations the variations in density in a flow are small and this enables us to make approximations which simplify the equations. We can estimate the expected fractional density variation in a flow in the following way. The system has characteristic length and time scales L and T and typical velocity U. The sound speed is $c_s^2 = (\partial P/\partial\rho)_\mathrm{s}$ (see section 2.3) and the Mach number is defined as $M \equiv U/c_s$. Ignoring the gravity term for the time being and comparing the size of the terms in the momentum equation (2.1), we find the following sizes:

$$\begin{array}{ccccc} \dfrac{\partial \mathbf{u}}{\partial t} & + & (\mathbf{u}\cdot\boldsymbol{\nabla})\mathbf{u} & = & -\dfrac{1}{\rho}\boldsymbol{\nabla} P \\ \dfrac{U}{T} & & \dfrac{U^2}{L} & & \dfrac{\delta P}{\rho L} \\ M^2\dfrac{L}{UT} & & M^2 & & \dfrac{\delta\rho}{\rho}, \end{array} \tag{2.22}$$

where δP is the typical departure of the pressure from the mean or equilibrium pressure. The third line is given by multiplying the second by L/c_s^2. Generally $UT \sim L$ and so the first two terms will be of comparable size. However, we see in section 2.3 that in a stationary fluid hosting a sound wave it is appropriate to equate L to the wavelength and T to the period, so $L/T \approx c_s \gg U$, provided that the wave is not of large amplitude, so that the second term can be neglected and the equation 'linearised'. The third term must be comparable to the terms on the left, since density differences within the fluid can only arise through the inertia of the fluid. If the motion is subsonic, i.e. if $M < 0$, then the fractional density differences are small. In the literature it is customary therefore to say that the flow is incompressible provided that $M < 0.3$ so that pressure differences are less than a tenth. This allows us to replace the density ρ appearing in the momentum equation next to the pressure gradient with some constant density ρ_0:

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \boldsymbol{\nabla})\mathbf{u} = -\frac{1}{\rho_0}\boldsymbol{\nabla} P \tag{2.23}$$

and the continuity equation (2.2) with

$$\boldsymbol{\nabla} \cdot \mathbf{u} = 0. \tag{2.24}$$

In principle this completes the incompressible set of equations because we have the same number of equations as variables. Normally though we solve for P by taking the divergence of the momentum equation and using (2.24) to give the Laplace equation

$$\nabla^2 P = 0, \tag{2.25}$$

which is a boundary value problem.

2.5 Steady flow of a compressible fluid: subsonic and supersonic flow

In this section we look at properties of a steady compressible flow, finding quite different behaviour according to whether the flow is subsonic or supersonic. Gravity is assumed to be absent, so Bernoulli's equation is

$$h + \frac{1}{2}u^2 = h_0 \tag{2.26}$$

where $h \equiv \epsilon + P/\rho$ is the specific enthalpy. The quantity h_0 is a constant along each streamline and is equal to the enthalpy the fluid has where the flow speed is zero. We are ignoring viscous processes and so the entropy is constant along streamlines:

$$s = s_0. \tag{2.27}$$

Now, changes in enthalpy are given in (A.17) which in the case of *specific* quantities, i.e. per unit mass, becomes $\mathrm{d}h = T\mathrm{d}s + \mathrm{d}P/\rho$. In an ideal fluid the first term is zero so we see that changes in h and P always have the same sign; this means that where the flow accelerates and the enthalpy drops – as is clear from (2.26) – the pressure must also drop. Physically this comes from the fact that the gas is being made to accelerate by the pressure gradient. While the maximum value of h along a streamline is h_0, it is not immediately obvious what the maximum value of u should be since we do not know the minimum value of h, which is a consequence of the fact that its absolute value has not yet been defined – as for internal energy ϵ only changes $\mathrm{d}h$ have been defined. The way out of this problem is to recognise that the maximum velocity is where $P = 0$ since the flow cannot be accelerated any more if there is no pressure. We can then define the zero point of h to be where temperature $T = 0$ and therefore $P = 0$. The maximum possible value of u along a streamline is then given by $u_{\max} = \sqrt{2h_0}$. This value is reached where the pressure goes to zero. A good example of this happening in nature is the solar wind: gas moves from a place with finite pressure to a place with essentially zero pressure and ignoring gravity the velocity is determined simply by the initial enthalpy. We have perfect conversion of thermal to kinetic energy, which if we think in terms of heat engines is only possible because the cold reservoir is at absolute zero. Microscopically we can think of a collection of particles with random thermal velocities being released into a vacuum where after some time a particle's position will depend only on its initial velocity and therefore its velocity is a function only of its position; there is no spread in velocities of particles in the same location and therefore the thermal energy has vanished.

The flow of fluid in terms of mass per unit area per unit time is $\rho\mathbf{u}$. This mass flux increases in the direction of the flow where streamlines converge and drops where they diverge. To calculate how it changes along a streamline we look at the component of the momentum equation (2.1) along a streamline, which can be written $u\mathrm{d}u = -\mathrm{d}P/\rho$. We also see that since $s = s(P,\rho) = s_0$, we must have during adiabatic changes $\mathrm{d}P = c^2\mathrm{d}\rho$ where $c = c(P,\rho)$. We shall see later in section 2.3 that c is the sound speed. We know from experience that c is real, since we want $\mathrm{d}P$ and $\mathrm{d}\rho$ to have the same sign in an adiabatic expansion or compression. Therefore $u\mathrm{d}u = -c^2\mathrm{d}\rho/\rho$ and so

$$\mathrm{d}(\rho u) = \rho\,\mathrm{d}u + u\,\mathrm{d}\rho = \rho\,\mathrm{d}u\left(1 - \frac{u^2}{c^2}\right) \tag{2.28}$$

meaning that in a subsonic flow, converging streamlines accompany acceleration, whereas acceleration in a supersonic flow is found where the streamlines are diverging. The latter is outside of our everyday experience and therefore somewhat counterintuitive.

The maximum possible mass flux ρu along a given streamline with stationary enthalpy h_0 must occur where $u = c$.

2.6 Vorticity

Let us define the circulation Γ around a closed loop as

$$\Gamma \equiv \oint \mathbf{u}\cdot\delta\mathbf{s} \tag{2.29}$$

where $\delta\mathbf{s}$ is the infinitesimal displacement along the loop. As the loop moves with the flow, the rate of change of this circulation is

$$\begin{aligned}
\frac{\mathrm{d}\Gamma}{\mathrm{d}t} &= \oint \frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t}\cdot\delta\mathbf{s} + \oint \mathbf{u}\cdot\frac{\mathrm{d}}{\mathrm{d}t}(\delta\mathbf{s}) \\
&= \int \boldsymbol{\nabla}\times\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t}\cdot\delta\mathbf{S} + \oint \mathbf{u}\cdot\delta\left(\frac{\mathrm{d}\mathbf{s}}{\mathrm{d}t}\right) \\
&= \int \boldsymbol{\nabla}\times\left(-\frac{1}{\rho}\boldsymbol{\nabla}P + \mathbf{g}\right)\cdot\delta\mathbf{S} + \oint \mathbf{u}\cdot\delta\mathbf{u} \\
&= \int \left(\frac{1}{\rho^2}\boldsymbol{\nabla}P\times\boldsymbol{\nabla}\rho\right)\cdot\delta\mathbf{S} + \oint \frac{1}{2}\delta\left(u^2\right),
\end{aligned}$$

where some vector calculus identities have been used. Obviously the second term on the right is zero; the first term vanishes only in a *barotropic flow* where the gradients of pressure and density are parallel. In summary, in an inviscid barotropic flow[1] where the body forces such as gravity are conservative, i.e. curl-free, we have

$$\frac{\mathrm{d}\Gamma}{\mathrm{d}t} = 0. \tag{2.30}$$

This is known as *Kelvin's circulation theorem*.

Finally, we define a useful quantity: the *vorticity* is the curl of velocity $\boldsymbol{\omega} \equiv \boldsymbol{\nabla} \times \mathbf{u}$. Using Stokes' theorem we see that

$$\Gamma = \oint \mathbf{u} \cdot \mathrm{d}\mathbf{s} = \int \boldsymbol{\omega} \cdot \mathrm{d}\mathbf{S} \tag{2.31}$$

where $\mathrm{d}\mathbf{s}$ and $\mathrm{d}\mathbf{S}$ are line and surface elements respectively of an infinitesimally small loop. This leads us to the following statement: if at some point in time the vorticity vanishes at every location, then it also vanishes at all other times. Such a flow is called *irrotational*.

We look at vorticity and rotation in more detail in chapter 7.

2.7 Potential flow

The conclusion of the last section is useful because we can simplify the equations by expressing a curl-free velocity field as the gradient of a scalar: $\mathbf{u} = \boldsymbol{\nabla}\phi$. This is called *potential flow* or *irrotational flow*, a subset of flow of an ideal fluid. First of all we make use of the vector identity $(\mathbf{u}.\boldsymbol{\nabla})\mathbf{u} = (1/2)\boldsymbol{\nabla}(u^2) - \mathbf{u} \times (\boldsymbol{\nabla} \times \mathbf{u})$, losing the last term because the flow is irrotational, to write the momentum equation (2.1) in the following form:

$$\frac{\partial}{\partial t}(\boldsymbol{\nabla}\phi) + \boldsymbol{\nabla}\left(\frac{1}{2}u^2\right) = -\boldsymbol{\nabla}h - \boldsymbol{\nabla}\Phi. \tag{2.32}$$

where the pressure gradient term has been reorganised with the help of a new function $h = h(P)$ where $\boldsymbol{\nabla}h \equiv (1/\rho)\boldsymbol{\nabla}P$, making use of the barotropic condition that $\rho = \rho(P)$, i.e. that there is only one independent thermodynamic variable. In the case of an isentropic flow, $h = \epsilon + P/\rho$ is the enthalpy we used above, otherwise it is some other function of pressure. Collecting terms we have

$$\frac{\partial\phi}{\partial t} + \frac{1}{2}u^2 + h + \Phi = f(t). \tag{2.33}$$

In an unsteady irrotational inviscid barotropic flow the sum of the four terms on the left is constant in space but not in time. In a steady flow on the other hand, we lose the first term on the left and the right hand side becomes a constant. This is a stronger statement than the form of Bernoulli's equation for a rotational inviscid flow (2.12), which stated that a fluid element retains the same value of a certain quantity as it moves; in other words that the quantity is constant along streamlines. Here, it is constant in all space, i.e. along all streamlines. The difference comes from the extra assumption here of barotropicity.

[1] Note that in a barotropic flow the fluid does not necessarily need to have a barotropic equation of state, such as that of the cold degenerate material that white dwarf stars are made of with $P = K\rho^\gamma$. A fluid with a non-barotropic equation of state can participate in a barotropic flow: for instance an ideal gas with $P = \rho RT$ could be isothermal so that T is constant everywhere in the volume of interest or it could be isentropic with $P\rho^{-\gamma}$ constant everywhere. If however we have non-parallel pressure and density gradients we say the flow is *baroclinic*, and circulation can be generated and destroyed without recourse to viscosity or non-conservative body forces.

Exercises

2.1 Flow of water through a hole in a barrel

Consider a barrel containing water with a hole through which water is exiting. The hole has cross-sectional area A, which is small compared to the size of the barrel. Use Bernoulli's equation to calculate the time taken for a barrel containing volume of water V and height h to empty.

Chapter 3

Some problems in one-dimensional flow

In this section we look at various phenomena and contexts in which the fluid flow can be considered one-dimensional, in that only one space dimension appears in the equations. Flow along a pipe of varying cross-section is an obvious example. In astrophysics, the most common context is a spherical geometry with spherical symmetry, only the radius r appears in the equations.[1]

3.1 Flow through a nozzle

Imagine we have a steady flow of compressible fluid through a tube of varying cross section A between two large volumes at pressures P_0 and P_1 where $P_0 > P_1$. We make the assumption that changes in the cross-section are gradual (i.e. that the diameter of the tube changes over length scales much larger than the diameter) and that the flow can be considered uniform across the cross-section of the tube. The fluid begins from rest in the first reservoir with enthalpy h_0 and entropy s_0, whose values in the tube are given by (2.26) and (2.27). Since the flow through the tube is steady the mass flux must be constant along the tube:

$$A\rho u = \text{const} \tag{3.1}$$

The tube is connected smoothly to the first volume in such a way that the cross section A is large where it joins the first reservoir and becomes smaller further away. The flow starts from rest and so is initially subsonic; we saw in section 2.5 that as the streamlines converge – in other words as ρu increases – the fluid accelerates. If the pressure difference between the two reservoirs is small, then the flow does not reach the sound speed and the pressure in the tube drops from P_0 at one end to P_1 at the other. It is important to note that the gas has already reached pressure P_1 as it exits the tube, giving lateral pressure balance between the emerging jet and the surroundings. Ignoring friction, the flow speed and therefore total flow in terms of mass per unit time can be calculated from (2.26) and (2.27). The fractional pressure difference required to reach the sound speed can be calculated from the properties of the fluid. For instance, an ideal gas has a ratio of specific heats γ, sound speed $c^2 = \gamma P/\rho$, enthalpy $h = c^2/(\gamma - 1)$ and the pressure and density during adiabatic expansion and compression obey $P/\rho^\gamma = P_0/\rho_0^\gamma$. The

[1] It has been said that astrophysics is the study of spheres and discs. This is a consequence of gravity and, in the latter, angular momentum.

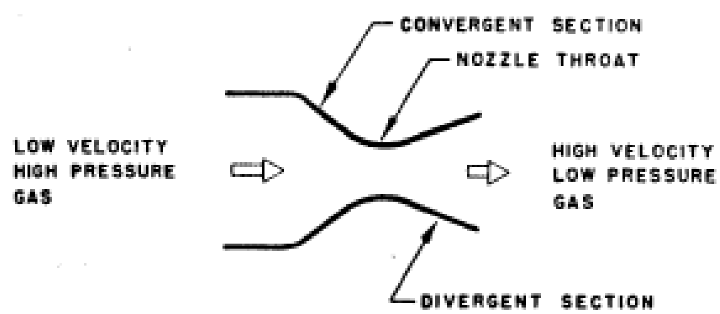
Mach number $M \equiv u/c$ during the acceleration is

$$
\begin{aligned}
M^2 &= \frac{2}{c^2}(h_0 - h) \\
&= \frac{2}{\gamma - 1}\left(\frac{c_0^2}{c^2} - 1\right) \\
&= \frac{2}{\gamma - 1}\left(\left(\frac{P_0}{P}\right)^{\frac{\gamma-1}{\gamma}} - 1\right) \qquad \Rightarrow \qquad P_{M=1} = P_0\left(\frac{2}{\gamma - 1}\right)^{\frac{\gamma-1}{\gamma}}.
\end{aligned}
\tag{3.2}
$$

The numerical factor is 0.49 for $\gamma = 5/3$ and 0.53 for $\gamma = 7/5$. Therefore if the pressure in the second reservoir is equal to this $P_{M=1}$ then the sound speed is reached exactly at the exit of the tube.

If the pressure of the second reservoir P_1 is decreased below $P_{M=1}$ then the behaviour depends on the form of the tube. We can see from (2.28) and (3.1) that continued acceleration along the tube beyond the sound speed can only happen if A increases. In other words, the tube must become narrower further away from the first reservoir only until the sound speed is reached, and then it must be flared. Such a shaped tube is called a *de Laval nozzle* after the Swedish engineer. If it is not flared, then the fluid cannot accelerate beyond the sound speed and the pressure in the tube cannot drop below $P_{M=1}$, meaning that the remaining pressure drop from $P_{M=1}$ and P_1 must take place *after* the fluid has exited the tube, while the tube fluid is mixing into the ambient fluid. In a flared tube, however, the fluid can continue to accelerate past the throat, or sonic point, driven by the remaining drop in pressure from $P_{M=1}$ downwards. The pressure after the sonic point can be calculated simply from the cross-section A; if the cross-section at the exit is such that the pressure in the tube is greater than P_1, then the remaining pressure drop takes place outside of the tube as it does with an unflared tube when $P_1 < P_{M=1}$. If the pressure calculated at the exit is *lower* than P_1 then the flow has a tendency to break away from the boundaries of the tube and a stationary shock wave enters the tube, the details of which are very much an engineering problem and beyond the scope of this course.

Figure 3.1: The de Laval nozzle.



3.2 Stellar winds and accretion

In this section we build on the analysis of nozzles to look at two very important astrophysical settings – stellar winds and accretion. Fortunately it turns out that both can be analysed with the same set of equations.

Let us imagine a steady, spherically-symmetric wind with velocity u coming from a star. Negative u signifies accretion. All quantities depend only on r, the distance from the origin. In nozzle parlance, the cross-section area of the flow is $4\pi r^2$ and so conservation of mass gives $4\pi r^2 \rho u$ =const, which we shall use later in the form:

$$-\frac{1}{\rho}\frac{\mathrm{d}\rho}{\mathrm{d}r} = \frac{1}{u}\frac{\mathrm{d}u}{\mathrm{d}r} + \frac{1}{r^2}\frac{\mathrm{d}r^2}{\mathrm{d}r} = \frac{1}{2u^2}\frac{\mathrm{d}u^2}{\mathrm{d}r} + \frac{2}{r}. \tag{3.3}$$

The momentum equation (in the radial direction) is

$$\begin{aligned} u\frac{\mathrm{d}u}{\mathrm{d}r} &= -\frac{1}{\rho}\frac{\mathrm{d}P}{\mathrm{d}r} - \frac{GM}{r^2} \\ \frac{1}{2}\frac{\mathrm{d}u^2}{\mathrm{d}r} &= -\frac{c^2}{\rho}\frac{\mathrm{d}\rho}{\mathrm{d}r} - \frac{1}{2}\frac{v_{\mathrm{esc}}^2}{r} \end{aligned} \tag{3.4}$$

where $v_{\mathrm{esc}}^2 = 2GM/r$ is the escape velocity, a function of radius. This equation becomes, on substituting from (3.3),

$$\begin{aligned} \frac{1}{2}\frac{\mathrm{d}u^2}{\mathrm{d}r} &= c^2\left(\frac{1}{2u^2}\frac{\mathrm{d}u^2}{\mathrm{d}r} + \frac{2}{r}\right) - \frac{1}{2}\frac{v_{\mathrm{esc}}^2}{r} \\ \frac{1}{2}\frac{\mathrm{d}u^2}{\mathrm{d}r}\left(1 - \frac{c^2}{u^2}\right) &= \frac{1}{2r}(4c^2 - v_{\mathrm{esc}}^2) \end{aligned} \tag{3.5}$$

which is perhaps more elegantly written

$$\frac{r}{u^2}\frac{\mathrm{d}u^2}{\mathrm{d}r} = \frac{4c^2 - v_{\mathrm{esc}}^2}{u^2 - c^2}, \tag{3.6}$$

which we can also write in the form

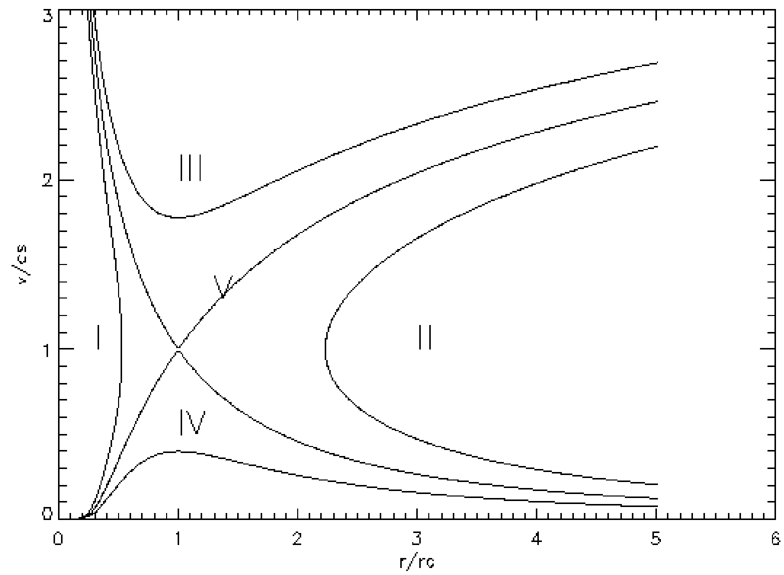
$$\frac{r^2}{4r_s}\frac{1}{u^2}\frac{\mathrm{d}u^2}{\mathrm{d}r} = \frac{1 - r/r_s}{1 - u^2/c^2} \qquad \text{where} \qquad r_s \equiv \frac{GM}{2c^2}. \tag{3.7}$$

Solutions to this equation are sketched in fig. 3.2. They fall into six categories; two which pass through the sonic point at $r = r_s$ with $u = c$, and four in the quadrants separated by them. The main two solutions of interest are the two which pass through the sonic point; the solution which is supersonic at $r < r_s$ and subsonic at $r > r_s$ is accretion and the other is the stellar wind solution.

Although r_s is a function of c which is a function of r, it is possible to make a rough estimate of the rate of spherical accretion onto a star from an interstellar medium (ISM) of given temperature and density. We simply assume that the sound speed at the sonic point is not too different from the sound speed at infinity c_0, an assumption which is justified by the relatively small numerical factors found in (3.2). Taking the temperature of the ISM to be 10^4K, the sound speed is $c = \sqrt{\gamma RT/\mu} \approx 10$ km s^{-1} if the gas is neutral hydrogen/helium. The accretion radius of a star of one solar mass is therefore $r_s = 7 \cdot 10^{13}$ cm which is about the radius of Jupiter's orbit. As a star passes through the ISM the gas outside of this radius is affected little by the gravitational pull of the star, but inside this radius the gas is falling supersonically. Using these numbers the accretion rate is

$$\dot{M} = 4\pi r_s^2 \rho c = \frac{\pi G^2 M^2 \rho}{c^3} \tag{3.8}$$

Figure 3.2: Solutions to (3.6). Solutions in regions I and II are double valued, i.e. velocity has two values at the same radius. Solutions of type III are supersonic everywhere, and type IV are subsonic everywhere. Type V has a sonic point.



which has the value 10^{11}g s^{-1} or around $10^{-15} M_\odot$ yr^{-1} if we take the ISM density to be 1 cm^{-3}, i.e. $\rho \sim 10^{-24}$ g cm^{-3}. Clearly this is not sufficient to form a star! This is why stars can only form in cold, dense environments; this state is achieved by radiative cooling. In binary systems, stars can accrete material much faster since the density is higher.

This analysis explains how gas at rest accretes onto a star, but it does not explain the origin of stellar winds; for this it is necessary to look in more detail at the nature of a stellar atmosphere. We know that the Sun is surrounded by a hot tenuous medium called the *corona*, which is the Latin for ‘crown’. We can attempt to find the structure of this atmosphere by assuming that the solar corona is static, and that the hydrostatic equation is satisfied:

$$\nabla P = \rho \mathbf{g} \qquad \Rightarrow \qquad \frac{\mathrm{d}P}{\mathrm{d}r} = -\rho g_r \tag{3.9}$$

where the form on the right hand side is given in spherical coordinates, where spherical symmetry is assumed. In the static problem the energy equation reduces to

$$\nabla \cdot (K \nabla T) = 0 \qquad \Rightarrow \qquad \frac{\mathrm{d}}{\mathrm{d}r}\left(r^2 K \frac{\mathrm{d}T}{\mathrm{d}r}\right) = 0. \tag{3.10}$$

which comes from the theory of heat diffusion (see section 5.7), expressing the condition that the net heat flux into a fluid element is zero in a static equilibrium. It turns out that the thermal conductivity K is proportional to $T^{5/2}$, which comes from kinetic theory of gases. Note the implicit assumption here that there are no heat sources or sinks, such as radiative losses. Imposing the boundary conditions $T = T_0$ at $r = r_0$, i.e. at the surface of the Sun, and $T = 0$ at infinity, the solution of the equation $r^2 T^{5/2} \mathrm{d}T/\mathrm{d}r = \mathrm{const}$ is:

$$\frac{T}{T_0} = \left(\frac{r}{r_0}\right)^{-2/7}. \tag{3.11}$$

Substituting this back into (3.9) gives:

$$\frac{\mathrm{d}P}{\mathrm{d}r} = -\frac{r_0^2}{r^2}\frac{g_0 P}{R_\mu T} = -\left(\frac{r_0}{r}\right)^{12/7}\frac{g_0 P}{R_\mu T_0} = -\left(\frac{r_0}{r}\right)^{12/7}\frac{P}{H_0} \tag{3.12}$$

where the equation of state $P = \rho R_\mu T$ and the inverse-square law $g = g_0 r_0^2/r^2$ have been used where g_0 is the gravitational acceleration at the solar surface. The pressure scale height at the surface $H_0 = R_\mu T_0/g_0$ has been defined. The solution is

$$\ln\left(\frac{P}{P_0}\right) = \frac{7r_0}{5H_0}\left[\left(\frac{r}{r_0}\right)^{5/7} - 1\right]. \tag{3.13}$$

The crux of the matter is that the pressure does not drop to zero as r goes to infinity, in fact it drops only by around three orders of magnitude if, as is the case in reality, that $H_0 \approx r_0/10$. This asymptotic pressure is much greater than the actual gas pressure in the interplanetary space. Somewhere we have made an incorrect assumption! It turns out (not surprisingly) that the incorrect assumption is that the corona is in hydrostatic equilibrium: in fact, the material is moving outwards, accelerating as it does so and being observed as the *solar wind* as it passes by the Earth, where is has a velocity of around 500 km s^{-1}. Now, looking at (3.13) we can see that the extent to which the static solution is wrong, so to speak, depends on the ratio H_0/r_0. If the atmosphere of a star is very cold and this ratio is consequently very small, only a small correction to the static solution is required, i.e. the mass loss rate of the star is very small. This can be thought of as a situation where the sound velocity (comparable to the thermal velocity of the particles) is very much less than the escape velocity from the surface. To achieve significant mass loss, a star must have a sufficiently hot atmosphere so that the sound speed is not much less than the escape velocity. In this way the mass-loss rate of the Sun is around $10^{-14}\, M_\odot\ \mathrm{yr}^{-1}$, since the corona is heated to around a million kelvin by some process involving magnetic fields. However, if the corona were so hot that its sound speed approached and exceeded the escape velocity, the result would be explosive mass loss; this corresponds to the everywhere-supersonic solution to the wind equation.

This type of wind is called a *thermal wind*. It is the dominant mechanism of mass loss in cool main-sequence stars with hot coronae such as the Sun. Higher-mass main-sequence stars lack hot coronae (a consequence of the lack of convection in their envelope) and so this mechanism does not work; in these stars winds can still be driven by radiative mechanisms, especially as a star's luminosity approaches the Eddington limit. These are fundamentally different from thermal winds in that it is the outwards momentum imparted by photons on individual particles which drives the wind, rather than thermal energy deposited into the corona as a whole. Mathematically this takes the form of an extra term in the momentum equation (3.4). This question of whether mass loss is driven by momentum or energy from stars and supernovae is also often encountered in discussions of mass loss from star clusters.

Exercises

3.1 Matching exit pressure to external pressure

We wish to construct a rocket engine which converts as much of the combustion thermal energy as possible into kinetic energy, in order to maximise propulsion. This means taking account of the extarnal pressure. Comment on practical difficulties in building the optimum rocket engine to work in space.

3.2 Stellar winds and mass loss

An important but poorly understood process in stellar physics is the mass loss. A very simple model of a stellar wind is the isothermal model. Assuming that the temperature is constant, calculate the mass-loss rate from a star as a function of its mass M, radius R, the temperature in the wind T and the pressure at the base of the wind (i.e. at $r = R$) P. Entering realistic numbers, estimate the mass-loss rate of the Sun, and comment.

3.3 Accretion into galaxy clusters

Galaxy clusters are the largest gravitationally-bound structures in the universe. They grow by accreting matter from their surroundings. Estimate the accretion rate, given realistic parameters. Comment on the limits of assuming spherical symmetry.

Chapter 4

Waves and instabilities

In this section we look at various types of waves and instabilities which are present in fluids. They can be categorised according to the nature of the restoring force. Sound waves, where the restoring force is pressure and which exist in any compressible fluid, we looked at in section 2.3. First, we look at various types of gravity wave, where the restoring force is gravity (obviously), and related instabilities which are driven by gravity. At the end of this chapter we add self-gravity to the equation, giving rise to the so-called Jeans instability. In chapter 7 we shall look at waves which require rotation, such as inertial waves and Rossby waves, and in chapter 8.1 we look at magnetic waves.

4.1 Surface gravity waves

The first type of gravity wave to look at is the kind which propagates in a body of water. The water has uniform depth h and the z coordinate points upwards so the the bottom and (equilibrium) surface of the water are at $z = -h$ and $z = 0$ respectively. We need now to look at the equations of motion of a incompressible fluid in a gravitational field; since the fluid is incompressible, the continuity equation and the momentum equation are

$$\nabla \cdot \mathbf{u} = 0; \tag{4.1}$$

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla)\mathbf{u} = -\frac{1}{\rho}\nabla P - g\hat{\mathbf{z}}. \tag{4.2}$$

No sound waves are permitted in an incompressible medium. Formally the incompressibility condition is equivalent to letting $(\partial P/\partial \rho)_s$ go to infinity, which causes the sound speed also to go to infinity. Any disturbance to the pressure field immediately spreads to the rest of the fluid. In the real ocean, the sound speed is not quite infinite ($c_s \approx 1.5$ km s^{-1}) but is significantly faster than other relevant speeds so that incompressibility is a good approximation. Note that the system of equations (4.1) and (4.2) provides the correct number of equations to determine $\mathbf{u}$ and P, but it is not obvious how to find P: the divergence of (4.2) is taken, so that the first and last terms vanish, and the resulting (Laplace) equation can be solved for P.

Sound waves aside, the only other possible restoring force for a wave is gravity. Linearising the momentum equation we have

$$\frac{\partial \mathbf{u}}{\partial t} = -\frac{1}{\rho}\boldsymbol{\nabla} P' - g\hat{\mathbf{z}} \tag{4.3}$$

in addition to

$$\nabla^2 P' = 0, \tag{4.4}$$

where again the linearisation is valid as long as the amplitude of the wave is much less than both the wavelength λ and the depth of the liquid h. At rest, the surface of the liquid is at height $z = 0$ and the lower boundary is at height $z = -h$. The perturbation to the height of the surface is ζ. We shall consider only two dimensions, the horizontal dimension being x; the horizontal and vertical components of velocity are u and w. The boundary conditions of the system are

$$w = 0 \quad \text{at} \quad z = -h; \qquad w = \frac{\mathrm{d}\zeta}{\mathrm{d}t} = \frac{\partial \zeta}{\partial t} + u\frac{\partial \zeta}{\partial x} \quad \text{at} \quad z = \zeta \qquad \text{and} \qquad P = 0 \quad \text{at} \quad z = \zeta, \tag{4.5}$$

where the first two are known as *kinematic boundary conditions* and the third as a *dynamic boundary condition*. Now, since the motion is irrotational we may express the velocity field as the gradient of a scalar potential ψ with the incompressibility condition as Laplace's equation $\nabla^2\psi = 0$. The kinetic boundary conditions can be expressed as

$$\frac{\partial \psi}{\partial z} = 0 \quad \text{at} \quad z = -h \qquad \text{and} \qquad \frac{\partial \psi}{\partial z} = \frac{\mathrm{d}\zeta}{\mathrm{d}t} \approx \frac{\partial \zeta}{\partial t} \quad \text{at} \quad z = 0, \tag{4.6}$$

where some linearisation has been performed on the second condition: the last term is dropped because it is second order in small quantities and the conditions can be approximated to apply at $z = 0$ rather than $z = \zeta$. We now make use of the form of Bernoulli's equation applicable in unsteady, irrotational flows (2.33):

$$\frac{\partial \psi}{\partial t} + \frac{1}{2}(u^2 + w^2) + \frac{P}{\rho} + gz = f(t). \tag{4.7}$$

The second term is second order and can be dropped, and the term on the right-hand side can be absorbed into $\partial\psi/\partial t$. Substituting from here for P into the dynamic boundary condition in (4.5) gives

$$\frac{\partial \psi}{\partial t} + g\zeta = 0 \qquad \text{at} \qquad z = 0 \tag{4.8}$$

where the same replacement of $z = \zeta$ by $z = 0$ has been made as before.

We now consider solutions of the form

$$\zeta = \hat{\zeta}\, e^{i(kx-\omega t)} \qquad \text{and} \qquad \psi = \hat{\psi}\, Z(z)\, e^{i(kx-\omega t)}. \tag{4.9}$$

Substituting this ψ into the incompressibility condition $\nabla^2\psi = 0$ gives

$$-k^2 + \frac{1}{Z}\frac{\mathrm{d}^2 Z}{\mathrm{d}z^2} = 0 \qquad \Rightarrow \qquad Z(z) = e^{kz} + ae^{-kz}, \tag{4.10}$$

where only one constant a is needed because we can absorb the rest into $\hat{\psi}$. The value of a we find from the first kinetic boundary condition of (4.6):

$$ke^{k(-h)} - ake^{-k(-h)} = 0 \qquad \Rightarrow \qquad a = e^{-2kh} \qquad \Rightarrow \qquad Z(z) = e^{kz} + e^{-kz-2kh}. \tag{4.11}$$

Now the second kinetic boundary condition of (4.6) and the dynamic boundary condition (4.8) become

$$k\hat{\psi}\left(1 - e^{-2kh}\right) + i\omega\hat{\zeta} \;=\; 0, \tag{4.12}$$

$$-i\omega\hat{\psi}\left(1 + e^{-2kh}\right) + g\hat{\zeta} \;=\; 0. \tag{4.13}$$

Eliminating $\hat{\psi}$ and $\hat{\zeta}$ from these two simultaneous equations we finally arrive at the dispersion relation, i.e. the relation between ω and k:

$$\omega^2 = gk \tanh kh. \tag{4.14}$$

We see immediately from this relation that there are two regimes – one long wavelength and one short wavelength with dispersion relations

$$\omega^2 \approx gk^2h \qquad \text{where} \quad kh \ll 1 \qquad\qquad \text{and} \qquad\qquad \omega^2 \approx gk \qquad \text{where} \quad kh \gg 1. \tag{4.15}$$

The phase speed ω/k and group speed $\mathrm{d}\omega/\mathrm{d}k$ are readily calculated in the two cases. The phase speeds are:

$$\frac{\omega}{k} \approx \sqrt{gh} \qquad \text{where} \quad kh \ll 1 \qquad\qquad \text{and} \qquad\qquad \frac{\omega}{k} \approx \sqrt{\frac{g}{k}} \qquad \text{where} \quad kh \gg 1. \tag{4.16}$$

In the long wavelength case, the group speed is the same as the phase speed and neither depend on the wavelength. We say that the waves are *non-dispersive*, meaning that a wave train consisting of various wavelengths will stay intact as it propagates. This is also true of sound waves. In the short wavelength case, on the other hand, the group velocity is a factor of two smaller than the phase velocity and more importantly they do depend on the wavelength. If we create some waves by making a localised non-sinusoidal disturbance for a finite period of time, the initially superimposed waves of different wavelength will propagate outwards at different speeds and so the shape of the waves will not be preserved; in fact the waveforms will approach sinusoidality. We call this kind of wave *dispersive*.

We can now examine in more detail the structure of these waves in the two limits. In the long wavelength limit these waves are normally called shallow water waves; tidal waves and tsunamis are good examples of these at different frequencies. The function $Z(z)$ depends only weakly on z and $\mathrm{d}Z/\mathrm{d}z \approx 2k^2(h + z)$. This derivative appears in the vertical velocity w which varies therefore from a maximum on the surface linearly down to zero at $z = -h$. The horizontal velocity on the other hand is roughly constant with depth. In addition to this we see from (4.9) that u and w are 90° out of phase with each other, meaning that the particle paths are ellipses.

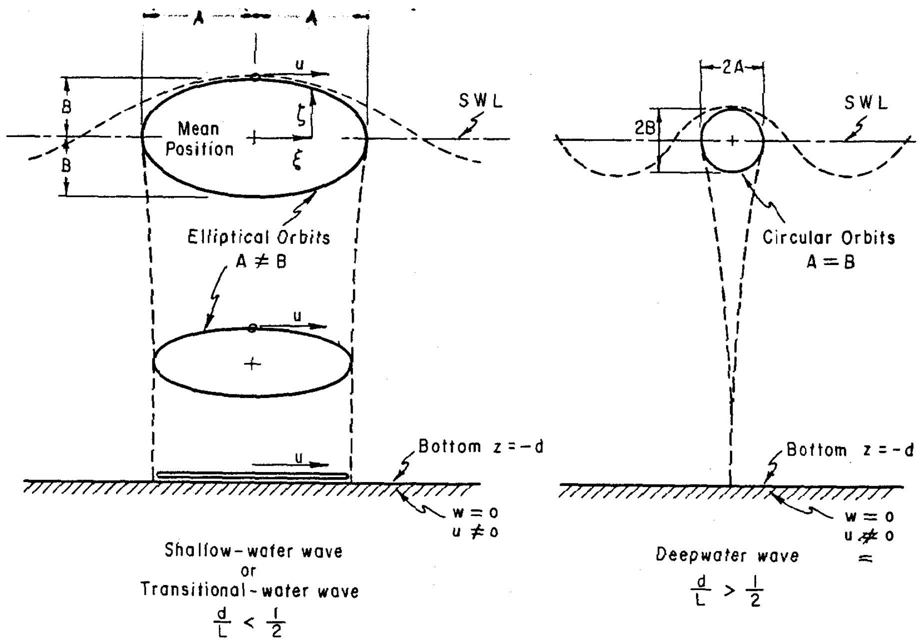
Deep water waves are excited on the sea by the wind (see section 4.3). Here, $Z(z)$ and therefore also the flow velocity decay exponentially away from the surface are negligible at $z = -h$. The particle paths are circular, as can be seen by comparing the vertical and horizontal derivatives of ψ.

Aside: the shallow-water equations

We saw above that in the case where the characteristic length scale of motion in the horizontal direction is much greater than the depth of the water, the horizontal velocity of the fluid is independent of z. In this situation we can simplify the equations of motion at the beginning of the analysis; since $\partial u/\partial x$ is independent of height, we see from (4.1) that $\partial w/\partial z$ is also independent of height. This means that the continuity equation can be easily integrated over the depth of the water and thus be expressed as

$$\frac{\partial \xi}{\partial t} = -h\frac{\partial u}{\partial x} \tag{4.17}$$

Figure 4.1: Particle paths in shallow and deep water waves.



where we have made the implicit assumption that $\xi \ll h$. In addition the horizontal component of the momentum equation can be similarly integrated to give

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = -g\frac{\partial \xi}{\partial x}, \tag{4.18}$$

which is a complete set of equations since the number of variables is now just two: u and ξ. Linearising to drop the second term on the left-hand side of (4.18) gives the same equations as we met with sound waves in section 2.3. The speed of the waves is $\sqrt{gh}$, which corresponds as expected to the long-wavelength limit in (4.16). We use the shallow water equations again in section 7 where the y dimension is added (intellectually trivial) and a Coriolis force is added to the momentum equation (intellectually more interesting).

4.2 Rayleigh-Taylor instability

We now turn our attention to an instability which occurs whenever a dense fluid lies in a gravitational field on top of a less dense fluid. The setup is similar to that in the previous section except that we now have two fluids with a boundary between them at $z = \zeta$. The fluid above has density ρ_1 and that below has ρ_2. Furthermore, to keep things simple we shall restrict ourselves to the case where both fluids are "deep" compared to the length scale of the disturbance at the interface. As one might intuitively expect, the situation is unstable if $\rho_1 > \rho_2$. The kinetic boundary conditions (4.5) become

$$\begin{gathered} u_1 = w_1 = 0 \quad \text{as} \quad z \to \infty; \qquad u_2 = w_2 = 0 \quad \text{as} \quad z \to -\infty \\ w_1 = \frac{\mathrm{d}\zeta}{\mathrm{d}t} = \frac{\partial \zeta}{\partial t} + u_1\frac{\partial \zeta}{\partial x} \quad \text{at} \quad z = \zeta; \qquad w_2 = \frac{\mathrm{d}\zeta}{\mathrm{d}t} = \frac{\partial \zeta}{\partial t} + u_2\frac{\partial \zeta}{\partial x} \quad \text{at} \quad z = \zeta, \end{gathered} \tag{4.19}$$

Figure 4.2: Deep-water waves resulting from a disturbance (localised in both time and space) on the surface of a pond. Note that the longer wavelength modes have travelled further than the shorter wavelength modes.

and the pressure boundary condition (4.5) becomes $P_1 = P_2$ at $z = \zeta$. The boundary conditions become, using the velocity potentials ψ_1 and ψ_2 in the two fluids, and linearising (compare to (4.6))

$$\nabla\psi_1 = 0 \quad \text{as} \quad z \to \infty; \quad \nabla\psi_2 \quad \text{as} \quad z \to -\infty; \qquad \frac{\partial\psi_1}{\partial z} = \frac{\partial\psi_2}{\partial z} = \frac{\partial\zeta}{\partial t} \quad \text{at} \quad z = 0. \tag{4.20}$$

The pressure condition similarly becomes

$$\rho_1 \left[\frac{\partial\psi_1}{\partial t} + g\zeta\right] = \rho_2 \left[\frac{\partial\psi_2}{\partial t} + g\zeta\right]. \tag{4.21}$$

Assuming as before solutions of the form $\psi_1 = \hat{\psi}_1 Z(z) \exp\left[i(kx - \omega t)\right]$, this gives us the dispersion relation

$$\omega^2 = gk \frac{\rho_2 - \rho_1}{\rho_1 + \rho_2}. \tag{4.22}$$

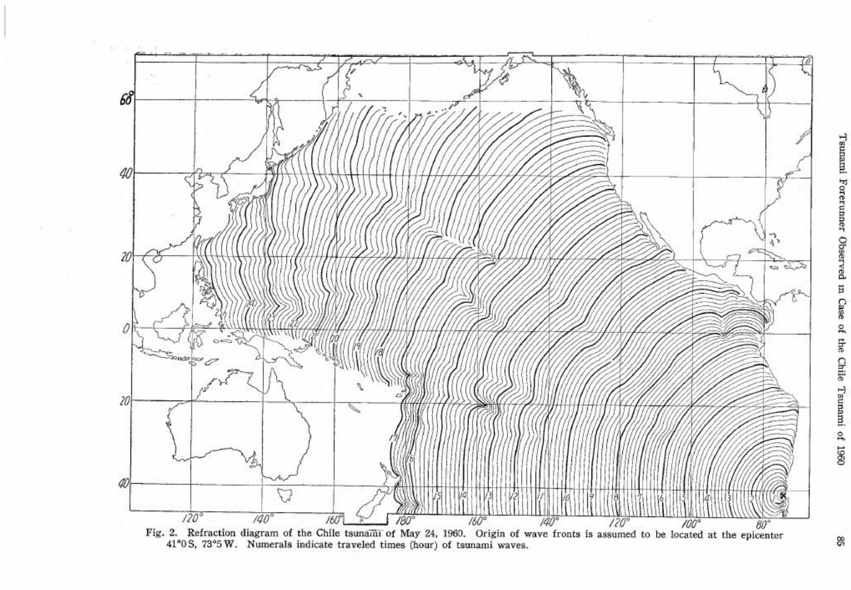
By setting $\rho_1 = 0$ we arrive back at the deep-water limit of the dispersion relation (4.15) we derived in the previous section. Clearly, ω can be either real or imaginary, depending on which of the two fluids is more dense.

In the unstable case, the growth timescale is given by

$$\tau_{\mathrm{R-T}} = \left(\frac{\rho_1 + \rho_2}{\rho_1 - \rho_2} \cdot \frac{\lambda/2\pi}{g}\right)^{1/2}. \tag{4.23}$$

This expression contains some ratio of the densities – in this case the difference between the densities is what drives the growth, so it is not surprising to see the fractional density difference in the expression for the timescale. The remaining part of this expression is simply the freefall timescale over a distance comparable to the wavelength; the shorter wavelengths grow more quickly. However, it is also important to note that the short wavelengths become nonlinear at smaller amplitude ζ than the larger wavelengths, and that the larger wavelengths will ultimately be crucial for the eventual turnover and/or mixing of the fluids.

Figure 4.3: The propagation of the tsunami of 1960 across the Pacific. Note how the speed changes in response to the changing depth and how this affects the direction of propagation.



4.3 Shear instability between two fluids: Kelvin-Helmholtz case

Before describing this instability, it is informative to have a look at the phenomenon of discontinuities in a fluid. Imagine a fluid contains a surface and that velocity, temperature and density have different values on either side of the surface. Such discontinuities may arise, for example, as a result of gravitational separation of two fluids (e.g. atmosphere and ocean) or as a result of supersonic flow.

Let us transfer to a frame of reference in which the discontinuity is at rest; the components of velocity normal and parallel to the discontinuity are $u_\perp$ and $u_\parallel$ respectively. We have now two possibilities: $u_\perp = 0$ on both sides of the discontinuity, or $u_\perp \neq 0$ on both sides. In the latter case various properties of the gas – enthalpy, specific momentum parallel to the surface, mean molecular weight and so on – must be conserved as it passes through the discontinuity. This case will be looked at in section 6. In the case where the fluid is not passing through the surface of discontinuity, there are different relations between properties of the fluid on either side; the fluids may have any velocity parallel to the surface and may actually be quite different in nature, but the pressure must be the same on both sides. It turns out that the discontinuity in parallel velocity gives rise to an instability in which an initially planar surface of discontinuity becomes rippled. The most obvious example is that of the wind blowing over the surface of the sea, which excites the waves discussed in section 4.1, but many other examples occur in nature such as the boundary between an astrophysical jet and its surroundings. This instability at surface of velocity discontinuity is called the *Kelvin-Helmholtz instability*; note that the astrophysics literature contains many references to Kelvin-Helmholtz in the case of a *continuous* shear flow such as those found in stars and discs, which is strictly speaking incorrect. The existence and nature of instability in such continuous shear flows is a topic of debate and will not be covered in this course.

Imagine two fluids separated by a horizontal planar discontinuity. The coordinate in the vertical direction is z and that in the horizontal direction is x. The densities of the upper and lower fluids are ρ_1 and ρ_2, and the undisturbed velocities in the x direction are U_1 and U_2. Note that there is no loss of generality, since it is always possible to change to a frame of reference in which both velocities are parallel to the

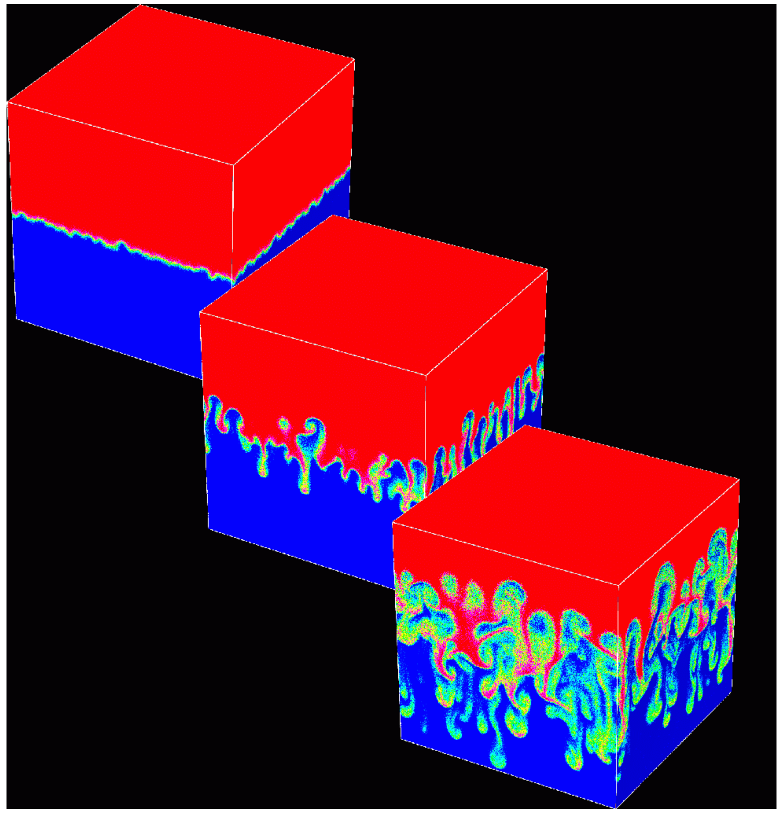


Figure 4.4: Simulations of the Rayleigh-Taylor instability. *K. Kadau, University of Duisburg.*

x-axis. Making the further assumption of incompressibility and inviscidity we can express the velocity as the gradient of a scalar, since from Kelvin's circulation theorem we see that the vorticity must be zero everywhere at all times (see section 2.7). We can express the flow as the sum of the undisturbed flow and a (small) perturbation:

$$\phi_1 = U_1 x + \psi_1, \tag{4.24}$$

$$\phi_2 = U_2 x + \psi_2, \tag{4.25}$$

where ϕ and ψ are the total and perturbation velocity potentials. The vertical perturbation to the position of the boundary is ζ, and since the fluid does not pass through the boundary we have at the boundary

$$w_1 = (U_1 + u_1)\frac{\partial \zeta}{\partial x} + \frac{\partial \zeta}{\partial t} \tag{4.26}$$

where u_1 and w_1 are the perturbation velocity components in the x and z directions. Linearising this equation, we can drop u_1 and take this condition to hold at $z = 0$ rather than $z = \xi$. We have therefore at $z = 0$:

$$\frac{\partial \psi_1}{\partial z} = U_1 \frac{\partial \zeta}{\partial x} + \frac{\partial \zeta}{\partial t}, \tag{4.27}$$

$$\frac{\partial \psi_2}{\partial z} = U_2 \frac{\partial \zeta}{\partial x} + \frac{\partial \zeta}{\partial t}. \tag{4.28}$$

We also know that the pressure must be continuous across the boundary. We can find the pressure from

Figure 4.5: Simulations of the Rayleigh-Taylor instability in a stellar wind context. *From Woitke 2006*

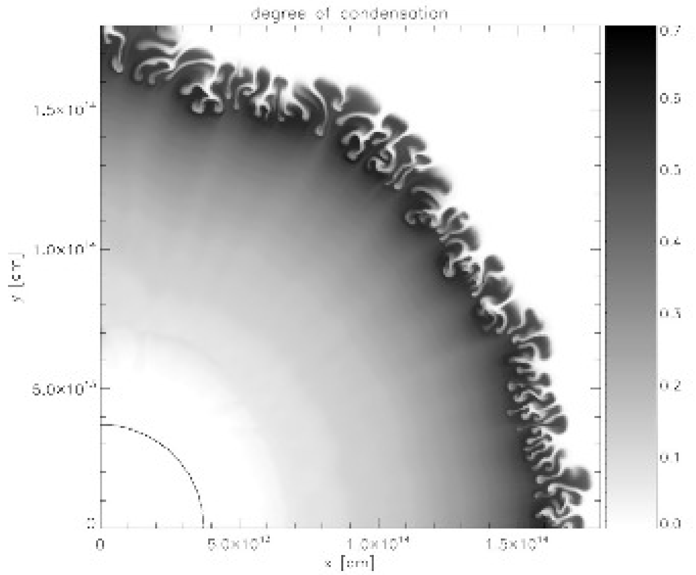


Fig. 7. Rayleigh-Taylor instabilities in an expanding dust shell shortly after the start ($t = 3.2\,$yr) of a new axisymmetric (2D) simulation. Stellar parameters $M_\star = 1\,M_\odot$, $L_\star = 10^4\,L_\odot$, $T_{\rm eff} = 2500\,$K and C/O = 1.9. The additional contour line for $T_{\rm rad} = 2500\,$K indicates the size of the star.

the form of Bernoulli's equation applicable to unsteady irrotational flow (2.33):

$$\frac{\partial \phi}{\partial t} + \frac{1}{2}\left((U+u)^2 + w^2\right) + \frac{p}{\rho} + gz = f(t) \tag{4.29}$$

which applies on both sides of the discontinuity. Equating pressure on both sides gives

$$\rho_1\left[f_1(t) - \frac{\partial \phi_1}{\partial t} - \frac{1}{2}\left((U_1+u_1)^2 + w_1^2\right) - g\zeta\right] = \rho_2\left[f_2(t) - \frac{\partial \phi_2}{\partial t} - \frac{1}{2}\left((U_2+u_2)^2 + w_2^2\right) - g\zeta\right]. \tag{4.30}$$

The unperturbed state must of course also satisfy this equation, so subtracting the unperturbed state and performing some reorganisation we find that at $z = 0$

$$\rho_1\left[\frac{\partial \psi_1}{\partial t} + U_1\frac{\partial \psi_1}{\partial x} + g\zeta\right] = \rho_2\left[\frac{\partial \psi_2}{\partial t} + U_2\frac{\partial \psi_2}{\partial x} + g\zeta\right], \tag{4.31}$$

where non-linear terms have been dropped. Since the motion at $z \to \pm\infty$ must vanish we see that the perturbations to f_1 and f_2 must also vanish since these quantities are constant in space; they can therefore be dropped from the equation above.

Let us assume a solution of the form $\psi_1 = \hat{\psi}_1\, Z(z)\, \exp\left[i(kx - \omega t)\right]$ and similar form for ψ_2. For ζ there will be no dependence on z. Since we have assumed that the fluid is incompressible we have $\nabla^2\psi = 0$, so that

$$\frac{Z''}{Z} - k^2 = 0 \qquad \Rightarrow \qquad Z = e^{\pm kz}, \tag{4.32}$$

Figure 4.6: Kelvin-Helmholtz instability in the laboratory. Two immiscible fluids of similar densities are placed in a long narrow tank and made to flow past one another by tipping the tank and letting gravity take its course.

where the sign in the exponent must be chosen so that solutions which diverge at infinity are dropped. Substituting these solutions into (4.27) and (4.27) gives

$$-k\hat{\psi}_1 = (-i\omega + U_1 ik)\hat{\zeta}, \tag{4.33}$$

$$k\hat{\psi}_2 = (-i\omega + U_2 ik)\hat{\zeta}. \tag{4.34}$$

We can also substitute the solutions into (4.31)

$$\rho_1(-i\omega + U_1 ik)\hat{\psi}_1 + \rho_1 g\hat{\zeta} = \rho_2(-i\omega + U_2 ik)\hat{\psi}_2 + \rho_2 g\hat{\zeta} \tag{4.35}$$

and using (4.33) and (4.34) to substitute for $\hat{\psi}_1$ and $\hat{\psi}_2$, this becomes

$$\rho_1(U_1 k - \omega)^2 + \rho_2(U_2 k - \omega)^2 = gk(\rho_2 - \rho_1). \tag{4.36}$$

The solutions are

$$\frac{\omega}{k} = \frac{\rho_1 U_1 + \rho_2 U_2 \pm \sqrt{(g/k)(\rho_2 - \rho_1)(\rho_2 + \rho_1) - \rho_1\rho_2(U_1 - U_2)^2}}{\rho_1 + \rho_2}. \tag{4.37}$$

It is a good idea to check at this stage that we recover (4.22) when we set $U_1 = U_2 = 0$. In the more general case, clearly there are no real roots to this quadratic equation (except for the trivial case $U_1 = U_2 = \omega/k$) without both non-zero gravity and $\rho_2 > \rho_1$; to be more precise the condition that the roots are real is

$$\frac{g}{k} > \frac{\rho_1\rho_2}{\rho_2^2 - \rho_1^2}(U_1 - U_2)^2. \tag{4.38}$$

If this condition is fulfilled, the waves are stable, i.e. their amplitude is constant. If not, the two solutions correspond to exponential growth and decay. Physically, one can think of stabilisation occurring if the energy released by the instability does not exceed the work which needs to be done against gravity to move the fluid vertically. An unstable mode will grow until non-linear effects become important; in the case of ocean waves excited by the wind, this happens at a particular ratio of amplitude to wavelength,

Figure 4.7: Schematic representation of the linear and nonlinear development of the Kelvin-Helmholtz instability.

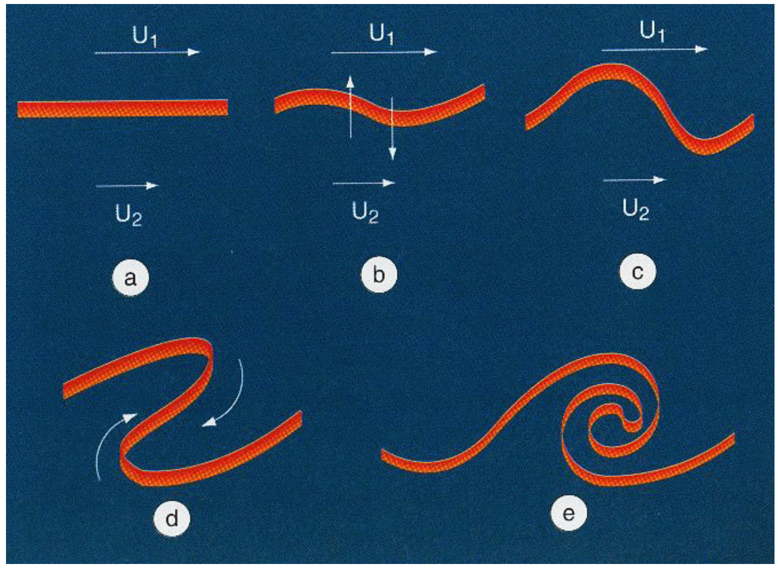
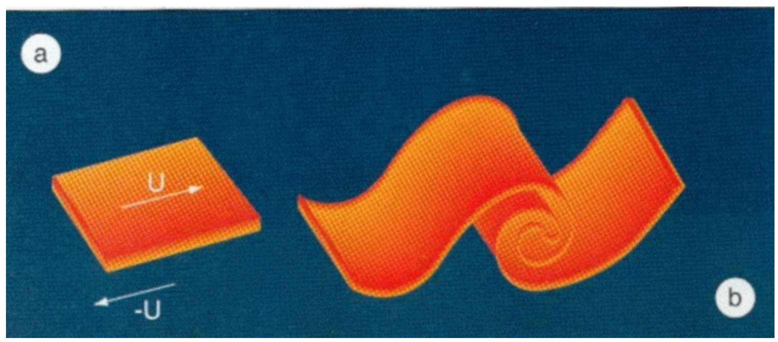


Figure 4.8: Break up of a vortex sheet into vortices.



and is visible as wave-breaking – at the onset of breaking, white froth appears. Note that the wavelength threshold increases with increasing velocity shear, explaining why the wavelength (and therefore height) of the largest waves on the sea is limited by the wind speed. Finally, a matter of terminology: if ω is real, the system is said to be stable; if it is imaginary, such as in the Rayleigh-Tayler case in the previous section where the displacement ζ increases exponentially, the system is unstable. If ω is complex and the solution represents oscillations of exponentially increasing amplitude, we say the system is *overstable.*

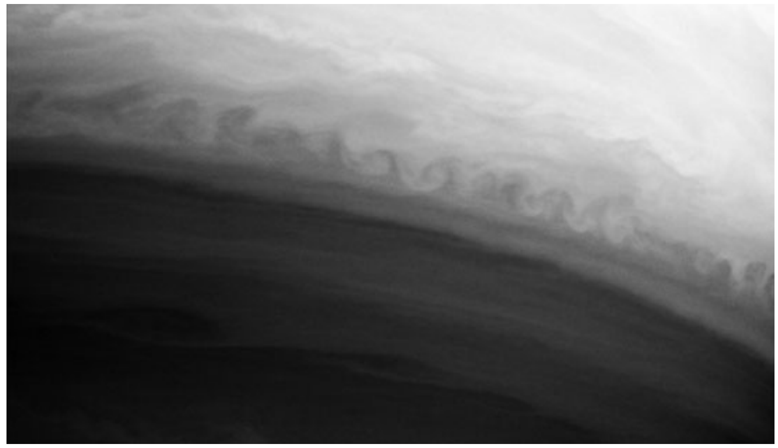
In many astrophysical contexts, gravity can be neglected, in which case a shear flow is unstable at all wavelengths. The timescale of the exponential growth is given by

$$\tau_{\text{K-H}} \equiv \frac{1}{\text{Im}(\omega)} = \frac{\rho_1 + \rho_2}{(\rho_1\rho_2)^{1/2}} \frac{\lambda/2\pi}{\Delta U} \tag{4.39}$$

where $\Delta U \equiv |U_1 - U_2|$ and $\lambda = 2\pi/k$. If the two densities are comparable the the growth time then ignoring factors of order unity we have $\tau_{\text{K-H}} \sim \lambda/\Delta U$, or in other words, it is equal to the time taken for one fluid to travel past the other a distance comparable to the wavelength. At shorter wavelength the instability grows faster, indeed the growth rate diverges at small wavelength and problems are prevented in reality by viscous damping of the shortest wavelengths.

We have now derived the dispersion relation and growth rate, but that has given us no insight into the physical mechanism which drives this instability. There are various ways to understand the mechanism;

Figure 4.9: Photograph taken by the Cassini spacecraft of the Kelvin-Helmholtz instability in the atmosphere of Saturn.



the simplest to perhaps to think in terms of how we apply Bernoulli's equation to the flow around a solid body. Where a fluid is forced to flow over an aeroplane wing, its speed must increase and its pressure therefore decreases. Here, consider that one fluid is forced to flow over undulations in the discontinuity separating the two fluids; its pressure falls, encouraging the other fluid to flow further into the space occupied by the first fluid, thereby increasing the displacement ξ.

The above describes the linear development of the instability. As the amplitude grows the instability becomes nonlinear and eventually produces mixing of the two fluids. This provides an intuitive way of understanding where the instability comes from: if we mix the two fluids together until they move with a uniform velocity, whilst conserving linear momentum, the kinetic energy must be lower: imagine an observer in an inertial frame of reference such that $\rho_1 U_1 + \rho_2 U_2 = 0$, i.e. where the combined momentum of the two fluids is zero (or rather, where the combined momentum of layers of thickness $1/k$ on either side of the discontinuity is zero). Incidentally, note that the phase speed of the waves as given by the real part of ω/k in (4.37) vanishes in this frame. Now, initially the kinetic energy is obviously non-zero; however as the two fluids become mixed the kinetic energy does tend to zero. Every instability works off some kind of free energy; in this instance the free energy is the kinetic energy of the shear. This kinetic energy, which is directed in the x direction, is converted first into kinetic energy in the z direction and eventually also into thermal energy as viscosity damps motion on short length scales. In some sense the energy originally present wants to convert into other forms and ultimately into thermal, as that represents the greatest entropy.

If the two fluids are in fact the same, with $\rho_1 = \rho_2$, then the equations simplify somewhat. Gravity, if present, has no effect, and the shear discontinuity is unstable to all wavelengths. Consequently, if such 'vortex sheets' develop in any kind of flow, we should expect them to be unstable and break up.

Both the Kelvin-Helmholtz and the Rayleigh-Taylor instabilities have divergent growth rate as the wavelength goes to zero. In reality of course a fluid is viscous and without working through the equations properly, we can estimate the shortest unstable wavelength if we know the kinematic viscosity ν. The timescale on which any disturbance of wavenumber k is damped is given by

$$\tau_{\rm visc} \approx \frac{1}{k^2 \nu} \tag{4.40}$$

and stabilisation occurs if this timescale is shorter than the growth timescale of the instability. Comparing with the instability growth times given by (4.39) and (4.23) it is easily verified that an upper limit on

Figure 4.10: Shear flow in the atmosphere. The properties of water vapour near saturation enable us to 'see' the flow of air.

unstable wavenumber k appears.

4.4 Internal gravity waves

In the previous section we looked at gravity waves in a liquid with a free surface and between two fluids with a discontinuity in density; the restoring force comes from the tendency of gravity to flatten out this free surface. In this section we look at fluids in which gravity waves can propagate without the need for any surface. For this reason, they are called *internal* gravity waves; density variations within the fluid give rise to the restoring force as gravity pulls more strongly on the more dense fluid elements. Both kinds of wave are important in geophysics, while in astrophysics the internal waves are more important and appear in a greater variety of objects than surface waves. Note that in the astrophysical literature, these waves (or oscillations, if the waves are 'standing') are often referred to as 'g-modes', where the letter g refers to the nature of the restoring force. In addition, if the restoring force is pressure, one speaks of 'p-modes', and of 'r-modes' if the restoring force comes from the rotation of the fluid body, via the Coriolis force (see chapter 7).

Consider a fluid in hydrostatic equilibrium, i.e. where $\frac{\partial P_0}{\partial z} = -\rho_0 g$ where the subscript 0 denotes the equilibrium quantities and gravity is directed downwards along the z axis. The equilibrium quantities are functions only of z. We look at small deviations from this equilibrium where the pressure field, for instance, becomes $P = P_0 + P'$. We want to look at motions which are subsonic, and where characteristic length scales in the vertical are much less than the length scale over which the density varies due to pressure differences (scale height H_ρ). Since the sound speed $c^2 = (\partial P/\partial\rho)_s$, we see from the hydrostatic equilibrium equation that $H_\rho \equiv (\partial \ln\rho/\partial z)^{-1} \sim c^2/g$. In a gas therefore it is comparable to the pressure scale height H_P but in a liquid it is much greater, for instance in the ocean it is $\approx$ 200km. So we ignore variations in density caused by pressure and keep just the variation due to entropy variation, since it is this variation which gives rise to the phenomena of interest. The is called the *Boussinesq approximation*. We have for the density

$$\rho' = -\alpha s' \qquad \text{where} \qquad \alpha \equiv -\left(\frac{\partial\rho_0}{\partial s_0}\right)_P \tag{4.41}$$

Figure 4.11: Rising cigarette smoke. This behaviour is generic to jets of fluid passing through surrounding fluid at rest.

where α has been defined so that it generally positive. The subsonic motion allows a simplification of the continuity equation to

$$\nabla \cdot \mathbf{u} = 0 \tag{4.42}$$

and the momentum equation can be written

$$\begin{aligned}\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla)\mathbf{u} &= -\frac{1}{\rho}\nabla P - g\hat{\mathbf{z}} \\ &= -\frac{1}{\rho_0}\nabla P' - g\frac{\rho'}{\rho_0}\hat{\mathbf{z}},\end{aligned} \tag{4.43}$$

where in reaching the second line the assumption that $\rho' \ll \rho_0$ has been used together with the hydrostatic balance equation for the equilibrium state. In our analysis of waves we can drop the second term on the left since it is second order in $\mathbf{u}$, a 'small' quantity. All we need now to complete the set of equations is the evolution of entropy, which in an ideal fluid is

$$\frac{\partial s'}{\partial t} + (\mathbf{u} \cdot \nabla)s = 0 \tag{4.44}$$

which can be linearised in the wave context to give

$$\frac{\partial s'}{\partial t} + w\frac{\mathrm{d}s_0}{\mathrm{d}z} = 0. \tag{4.45}$$

We now assume solutions of the form $\exp[i(\mathbf{k} \cdot \mathbf{r} - \omega t)]$. The continuity equation (4.42) becomes

$$\mathbf{k} \cdot \mathbf{u} = 0 \tag{4.46}$$

which means that the motion is perpendicular to the wavevector $\mathbf{k}$. Equations (4.43) and (4.45) are

$$-i\omega\mathbf{u} = -\frac{1}{\rho_0}i\mathbf{k}P' + g\frac{\alpha s'}{\rho_0}\hat{\mathbf{z}}, \tag{4.47}$$

$$-i\omega s' + w\frac{\mathrm{d}s_0}{\mathrm{d}z} = 0. \tag{4.48}$$

These three equations (4.46), (4.47) and (4.48) can be combined to eliminate $\mathbf{u}$, P' and s' to find the dispersion relation

$$\omega^2 = N^2 \sin^2\theta \tag{4.49}$$

where θ is the angle between the wavevector $\mathbf{k}$ and the vertical and N is the *buoyancy frequency* or *Brunt-Väisälä frequency*, which has the value

$$N^2 = \frac{\alpha g}{\rho} \cdot \frac{\mathrm{d}s}{\mathrm{d}z}. \tag{4.50}$$

The first thing we see is that, since oscillations only occur if ω is real, we need a positive entropy gradient if α is positive; if this is not the case then instead of oscillations we get convective turnover.[1] The second thing we see is that this dispersion relation is clearly quite different from those of surface gravity waves in that the frequency depends only on the angle of the wavevector and not on its magnitude. Physically this can be understood in the following way. A system hosting deep water surface gravity waves has no particular length scale; the only constant appearing in the equations is g. Therefore the waves can be made as small or as large as desired and correspondingly the frequency can have any value. In a continuously stratified fluid, however, the change in density takes place over a finite distance H_ρ. It is impossible to make the oscillations happen faster by reducing the length scale of the disturbance, because that also reduces the fractional difference in density and therefore the restoring force. In fact it is impossible to make the oscillations happen faster than the buoyancy frequency N.

Note also that the frequency goes to zero when the wavevector is vertical. In this case, the motion is entirely horizontal and it is clear that the restoring force vanishes. This is the reason that flows in stratified fluids tend to reside in horizontal surfaces; for instance the motion in the atmosphere is almost perfectly horizontal, except in those places where the entropy gradient becomes negative and convection appears. The same is true of flows around stars; this manifests itself for instance in that chemical elements are mixed efficiently on surfaces of constant radius but that the mixing in radius is very slow in radiative (non-convective) zones.

4.5 Sounds waves and the Jeans instability

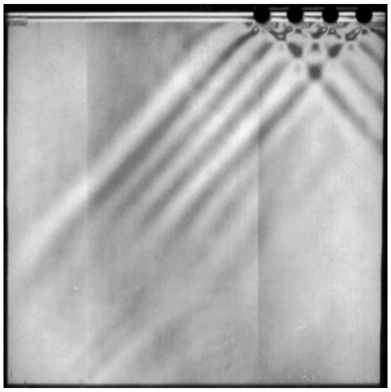
We saw in section 2.3 that sound waves can exist in any kind of fluid, since the only ingredient required is pressure, which is present in every fluid. Here, we have a gravity term in the momentum equation (2.1) $\mathbf{g} = -\nabla\Phi$ where Φ is the gravitational potential. This term can give rise to an instability if the gravitational potential is a result of the self-gravity of the fluid itself, rather than some distant fixed mass as in the previous sections. Recalling that for a fluid of density ρ, the gravitational potential Φ is given by the Poisson equation[2]

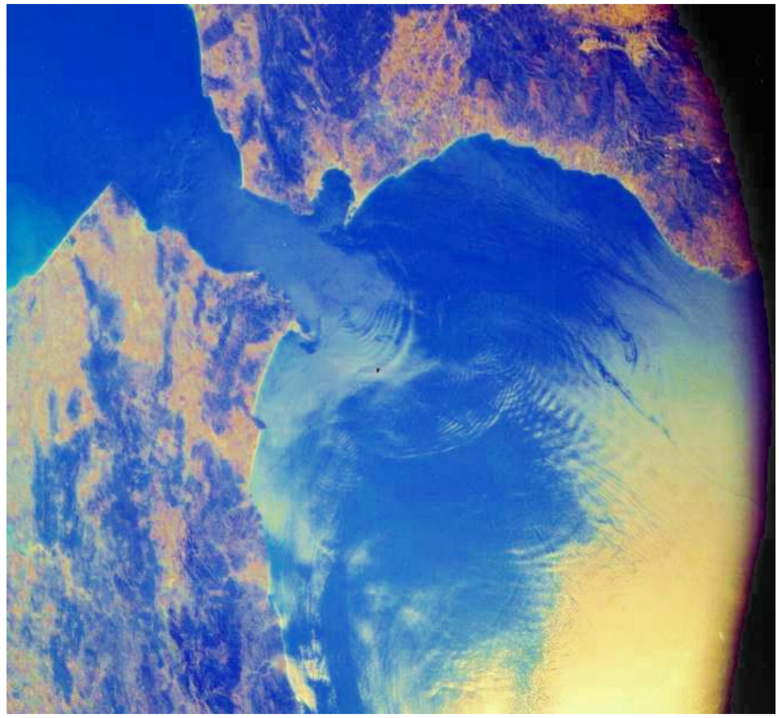
$$\nabla^2\Phi = 4\pi G\rho, \tag{4.51}$$

[1]Exceptions to $\alpha > 0$ include water between 0 and 4°C. This allows ice to form on the surface of a lake in the winter without all the water right down to the bottom having to cool to zero first.

[2]The intermediate step here is $4\pi G\rho = -\nabla \cdot \mathbf{g}$. Note the similarity with Maxwell's equation $4\pi\rho_e = \nabla \cdot \mathbf{E}$. No constant is required in the electromagnetic case because it is built into the definition of the unit of charge (in c.g.s., but not SI units); the other difference is the minus sign.

Figure 4.12: *Below:* internal gravity waves in the laboratory, propagating diagonally downards. *Right:* waves excited by the passage of air through the Straights of Gibraltar.





we can make the simplification (as in section 2.3) that all motion is in the x direction and that all quantities have vanishing gradients in the y and z directions, making the momentum equation, continuity equation, and self-gravity equation:

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = -\frac{1}{\rho}\frac{\partial P}{\partial x} - \frac{\partial \Phi}{\partial x}, \tag{4.52}$$

$$\frac{\partial \rho}{\partial t} = -\frac{\partial}{\partial x}(\rho u), \tag{4.53}$$

$$\frac{\partial^2 \Phi}{\partial x^2} = 4\pi G\rho. \tag{4.54}$$

Just as in section 2.3, we linearise these equations. We first define some background equilibrium state with pressure P_0, density ρ_0 and zero velocity, writing pressure $P = P_0 + \delta P$, $\rho = \rho_0 + \delta\rho$, $\Phi = \Phi_0 + \delta\Phi$ where $\delta P \ll P_0$ and $\delta\rho \ll \rho_0$. It follows that the velocity u is also small. As before, we define $c_s^2 \equiv \partial P/\partial\rho$. The linearised equations are

$$\frac{\partial u}{\partial t} = -\frac{1}{\rho_0}\frac{\partial(\delta P)}{\partial x} - \frac{\partial(\delta\Phi)}{\partial x}, \tag{4.55}$$

$$\frac{\partial(\delta\rho)}{\partial t} = -\rho_0\frac{\partial u}{\partial x}, \tag{4.56}$$

$$\frac{\partial^2(\delta\Phi)}{\partial x^2} = 4\pi G\,\delta\rho, \tag{4.57}$$

$$\delta P = c_s^2\delta\rho. \tag{4.58}$$

Substituting from (4.58) for $\delta\rho$ into (4.56) and (4.57) gives

$$\frac{1}{c_s^2}\frac{\partial(\delta P)}{\partial t} = -\rho_0\frac{\partial u}{\partial x}, \tag{4.59}$$

$$\frac{\partial^2(\delta\Phi)}{\partial x^2} = \frac{4\pi G}{c_s^2}\,\delta P. \tag{4.60}$$

and then differentiating (4.55) w.r.t. x, substituting from (4.59) for u and from (4.60) for Φ, and tidying, gives

$$\frac{\partial^2(\delta P)}{\partial t^2} = c_s^2\frac{\partial^2(\delta P)}{\partial x^2} + 4\pi G\rho_0\,\delta P. \tag{4.61}$$

We now assume solutions of the form $\exp[i(kx - \omega t)]$, which upon substitution into (4.61) gives

$$\omega^2 = c_s^2k^2 - 4\pi G\rho_0 \tag{4.62}$$

which clearly gives oscillations (i.e. real ω) only if the right-hand side is positive. The stability criterion is often expressed as an equality in terms of the wavelength $\lambda = 2\pi/k$ as

$$\lambda < c_s\sqrt{\frac{\pi}{G\rho}} \tag{4.63}$$

where the subscript on the density has been omitted. The critical wavelength of marginal stability is known as the *Jeans length*, named after the English astronomer James Jeans. Above this length scale, gas tends to collapse under its own gravity.

Exercises

4.1 Energy density and flux

(a) Show that the average potential and kinetic energies of a wave are equal. By integrating the work done per unit time and area Pu, derive the energy flux of the wave, and show that the group velocity is simply the ratio of energy flux to energy density.

(b) Explore what happens as a wave moves from deep water to ever shallower water, assuming that the energy flux is constant.

(c) By considering refraction, explain qualitatively what happens to the direction of propagation of a wave as it approaches a beach.

4.2 Waves at interface between two liquids

(a) Imagine two liquids of different densities lying on top of each other, the more dense fluid below. Derive the dispersion relation for gravity waves at the interface, assuming that the wavelength is much smaller than the depth of either liquid. To do this, the boundary conditions required are the same as those used in this section except that they must be applied to both fluids, also the dynamic condition is that the pressure on both sides of the discontinuity is the same, but non-zero.

(b) Construct a rudimentary experiment in a coffee cup to demonstrate the existence of these waves. *Hint: milk is denser than water.*

The excitation of waves between two layers of liquid is the explanation for a phenomenon called ‘dead water’, where boats entering Norwegian fjords experience increased drag. The fjords contain fresh water lying on top of salty seawater.

4.3 Stokes’ drift

By keeping second order terms, show that in the deep water case the particle paths in a wave of finite amplitude are not quite circular but that averaged over a cycle a particle moves slightly in the direction of wave propagation.

4.4 Alternative calculation of buoyancy frequency

Consider a fluid element in a stratified fluid. Initially the fluid element has the same pressure and density as its surroundings at that particular height z, and it is displaced vertically from this equilibrium position a distance δz. Calculate the restoring force on the fluid element as a function of this displacement and use this to calculate the frequency at which the element oscillates about its equilibrium position.

4.5 Gravitational collapse

Building on the analysis in section 4.5, show that the Jeans length can also be estimated by equating the time taken for a sound wave to travel a certain distance to the freefall time over that distance. Furthermore, by considering a spherical cloud (of constant density, to simplify matters), show that the Jeans length is simply the size of the cloud in which the thermal and gravitational energies are comparable. Comment on the meaning of this, in terms of the virial theorem. Finally, calculate the Jeans length under typical conditions in the interstellar medium. For example, in much of the ISM $c_s = 10\,\mathrm{km\ s^{-1}}$ and $\rho = 10^{-24}\,\mathrm{g\ cm^{-3}}$, and in dense clouds $c_s = 1\,\mathrm{km\ s^{-1}}$ and $\rho = 10^{-21}\,\mathrm{g\ cm^{-3}}$.

Chapter 5

Viscous fluids

We now turn our attention to the form of the viscous term in the momentum equation, $\mathbf{F}_{\rm visc}$. First we take a look at the equations and then use them to solve some simple problems.

5.1 The viscous stress tensor

First it is helpful to write the momentum equation in a more suitable form:

$$\frac{\partial}{\partial t}(\rho u_i) = -\frac{\partial T_{ij}}{\partial x_j} \tag{5.1}$$

where T_{ij} is the *momentum flux tensor*. It is easily demonstrated that it must always be possible to write the equation of motion in this form in situations without body forces such as gravity, by integrating over volume and using Gauss' theorem to express the right-hand side as a surface integral; the left-hand side then represents the rate of change of momentum of the volume and the right-hand side the forces acting on it at the boundaries. Also, note the similarity with the mass conservation equation. The momentum flux tensor is given by

$$T_{ij} = \rho u_i u_j + P\delta_{ij} - S_{ij}. \tag{5.2}$$

The first part of the momentum flux tensor is often called the *Reynolds stress*, while the second and third terms together are called the *stress tensor* – where the viscous part thereof, S_{ij}, is called the *viscous stress tensor*.

In finding the form of S_{ij}, the following axioms must be adhered to. First, it must vanish in the case of a uniform velocity, which means that terms containing the velocity must be absent, and that it must instead be made up of velocity gradients. Secondly, we know that the viscous stress is linear in these gradients. So, the tensor must consist of only terms like $\partial u_i/\partial x_j$. Now, a non-zero value of $\partial u_i/\partial x_j - \partial u_j/\partial x_i$ represents a uniform rotation of the fluid, in which case the viscous stress must also vanish; this means that only terms $\partial u_i/\partial x_j + \partial u_j/\partial x_i$ are permissible, which represent a change of size or shape of the fluid elements. It is common at this juncture to introduce the *rate of strain tensor*

$$e_{ij} \equiv \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right). \tag{5.3}$$

There are further properties which follow from the symmetry and rotational symmetry between the three dimensions: the three diagonal elements must all have the same form; the tensor is symmetrical, i.e. $S_{ij} = S_{ji}$, and the off-diagonal elements must also have the same form. Therefore we can write

$$S_{ij} = ae_{ij} + b\delta_{ij}\frac{\partial u_k}{\partial x_k}, \tag{5.4}$$

where a and b are properties of the fluid in question. Now, it is observed in many fluids, including monatomic gases that – to a good approximation – no energy is dissipated during an *isotropic* compression or expansion. In other words, a fluid element can be compressed while its shape is preserved, then expanded again back to its original size, and the work done by the element during the expansion is equal to that done on the element during the compression. During such a change, $\partial u_1/\partial x_1 = \partial u_2/\partial x_2 = \partial u_3/\partial x_3$ and bearing this in mind we can rewrite (5.4) thus:

$$S_{ij} = \mu\left(2e_{ij} - \frac{2}{3}\delta_{ij}\frac{\partial u_k}{\partial x_k}\right) + \zeta\delta_{ij}\frac{\partial u_k}{\partial x_k}, \tag{5.5}$$

where a and b have been replaced by μ and ζ which are called the *shear* and *bulk viscosities*, respectively; in monatomic gases the bulk viscosity is zero. To find the physical reason for this, it is necessary to make a brief digression from "classical" hydrodynamics and consider the constituent particles. We can consider the thermal energy in a gas in thermodynamic equilibrium as being divided equally (equipartition) between the various degrees of freedom, so that in a monatomic gas we have an energy per mole of $RT/2$ for the translational kinetic energy in each of the three dimensions and the total thermal energy per mole is $3RT/2$. During an isotropic expansion, energy is extracted at the same rate from kinetic energy in each of the three dimensions whereas the expansion of a gas in a cylinder-piston system extracts energy from just one dimension. In the latter case, the energies are brought out of equipartition and must gradually come back to equipartition; the finite time required to do this means that the pressure exerted on the piston during expansion is lower than it would be if the energy was redistributed instantly. During a compression the pressure on the piston is higher, therefore a net work must be done on the gas over a cycle consisting of expansion followed by compression; this work appears in the system as heat energy. This difference between the irreversibility of an isotropic and a non-isotropic change in volume is the origin of the second term inside the brackets in equation (5.5), ensuring that the stress tensor becomes zero in the isotropic expansion case. Now, a similar process occurs during the expansion of a gas made from diatomic or more complex molecules; at thermodynamic equilibrium, energy is split equally between not only the three translational kinetic energies but also the rotational kinetic energy (of which there are two degrees of freedom in the case of diatomic molecules such as those which make up the major fraction of the Earth's atmosphere). Even during an isotropic expansion, kinetic energy is extracted from the three translational degress of freedom but not from the rotational and the lag between the two gives rise to the same kind of dissipation as in the case of monatomic gas in a piston. This is the origin of bulk viscosity.

Recall that in section 1.1 we saw that the fluid approximation consists amongst other things in assuming that the mean-free path of particles is very much less than any other length scales of interest. This is because the idea of a local thermodynamic equilibrium is meaningful only in a fluid element at least as large as the mean-free path. Here, we have seen that the finite mean-free path, or rather the finite collision timescale, gives rise to a lag between energies and non-equipartition between different degrees of freedom. Therefore in some sense, the viscous terms in the fluid equations can be considered as first-order in the mean-free path.

In the case of an incompressible flow (i.e. the volume of each fluid element is not changing) with a velocity shear, the viscous stress acts to reduce the shear by transporting momentum across the fluid. Microscopically, this comes from individual particles transporting their momentum to another location where the mean velocity is different. Many undergraduate syllabuses include the calculation of the shear viscosity of a gas from consideration of momentum transport of particles in a shear flow. In applications where accuracy is not important (e.g. astrophysics) it is not important to know the detail of the calculation, but just that the dynamic viscosity of a gas is approximately equal to density × sound speed × mean-free path. In fact, this result can be obtained from a simple dimensional analysis. Remember that the sound speed is roughly equal to the thermal speed of the particles.

Back to the fluid picture: so far we have found that the viscous force per unit volume, i.e. the term $\mathbf{F}_{\rm visc}$ on the right-hand side of (1.3) can be written

$$F_{{\rm visc},i} = \frac{\partial S_{ij}}{\partial x_j} = \frac{\partial}{\partial x_j}\left\{\mu\left(2e_{ij} - \frac{2}{3}\delta_{ij}\frac{\partial u_k}{\partial x_k}\right) + \zeta\delta_{ij}\frac{\partial u_k}{\partial x_k}\right\}. \tag{5.6}$$

This is a rather complicated expression, and we can get a better intuitive understanding of the physics if we make some simplifications. First of all, in an incompressible flow we can drop the terms with the velocity divergence. Next, we can assume that the dynamic viscosity μ is a constant and can therefore be brought outside of the divergence, giving

$$F_{{\rm visc},i} = \mu\frac{\partial}{\partial x_j}(2e_{ij}) = \mu\frac{\partial}{\partial x_j}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right) = \mu\frac{\partial^2 u_i}{\partial x_j^2}, \tag{5.7}$$

where the zero-divergence of the velocity field was once again used to remove the second half of the rate of strain tensor. Defining the *kinematic viscosity* $\nu \equiv \mu/\rho$ we can write the momentum equation as

$$\frac{{\rm d}\mathbf{u}}{{\rm d}t} = -\frac{1}{\rho}\boldsymbol{\nabla}P + \mathbf{g} + \nu\nabla^2\mathbf{u}. \tag{5.8}$$

In this form it is easier to understand the action of viscosity – essentially it acts to smooth out variations in the velocity field. Where there is a local minimum in u_x, for example, $\nabla^2 u_x$ is positive and so the viscosity brings about an increase in u_x. We can see from the first and last terms in (5.8) that this smoothing happens on a timescale

$$\tau_{\rm visc} \sim \frac{L^2}{\nu} \tag{5.9}$$

where L is the characteristic length scale.

The extra viscous term in the momentum equation (5.8) fundamentally changes the nature of the equations. Without it, a problem can be entirely specified if the perpendicular component of the velocity is set to zero at the boundaries, as we shall see in section 5.5 when we calculate the inviscid flow past a sphere. However, with the viscous term, which contains a second order derivative of the velocity, this boundary condition is not sufficient and something else is needed to properly constrain the solution. What is needed is that not just the perpendicular component but also the parallel part of the velocity must go to zero at the boundary. In fact we already know this from everyday experience, for instance when trying to blow dust off a flat surface – some layer of dust always remains. As we shall see below, what is happening in that there is a thin boundary layer of strong velocity shear next to the surface where the velocity goes from zero at the surface to its value in the external flow.

5.2 Viscous heating

Having calculated the stress tensor and therefore the effect of viscosity on the velocity field, it is now necessary to calculate the energy dissipated as heat and add a term to the energy equation. First, we take the dot product of velocity with the momentum equation

$$\rho\mathbf{u}\cdot\frac{\partial\mathbf{u}}{\partial t} = -\rho\mathbf{u}\cdot(\mathbf{u}\cdot\boldsymbol{\nabla})\mathbf{u} - \mathbf{u}\cdot\boldsymbol{\nabla}P + u_i\frac{\partial S_{ij}}{\partial x_j} \tag{5.10}$$

which allows us to calculate the Eulerian rate of change of kinetic energy density:

$$\begin{aligned}\frac{\partial E_{\rm kin}}{\partial t} &= \frac{\partial}{\partial t}\left(\frac{1}{2}\rho\mathbf{u}^2\right) = \frac{u^2}{2}\frac{\partial\rho}{\partial t} + \rho\mathbf{u}\cdot\frac{\partial\mathbf{u}}{\partial t}\\ &= -\frac{u^2}{2}\boldsymbol{\nabla}\cdot(\rho\mathbf{u}) - \rho\mathbf{u}\cdot(\mathbf{u}\cdot\boldsymbol{\nabla})\mathbf{u} - \mathbf{u}\cdot\boldsymbol{\nabla}P + u_i\frac{\partial S_{ij}}{\partial x_j}\\ &= -\frac{u^2}{2}\boldsymbol{\nabla}\cdot(\rho\mathbf{u}) - \rho\mathbf{u}\cdot\boldsymbol{\nabla}\left(\frac{u^2}{2}\right) - \boldsymbol{\nabla}\cdot(P\mathbf{u}) + P\boldsymbol{\nabla}\cdot\mathbf{u} + \frac{\partial}{\partial x_j}(u_iS_{ij}) - S_{ij}\frac{\partial u_i}{\partial x_j}\\ &= -\boldsymbol{\nabla}\cdot\left(\rho\mathbf{u}\frac{u^2}{2} + P\mathbf{u} - \mathbf{u}\mathbf{S}\right) + P\boldsymbol{\nabla}\cdot\mathbf{u} - S_{ij}\frac{\partial u_i}{\partial x_j}.\end{aligned} \tag{5.11}$$

The term inside the bracket is an energy flux; integrating the equation over the volume of a fluid system one can convert this term to a surface integral of the flux over the boundary. The next term represents reversible conversion between thermal and kinetic energy; note that this term also appears in the energy equation (1.11), with the opposite sign. The last term is viscous conversion of kinetic into thermal energy. Taking the symmetry into account we can write this viscous heating as

$$Q_{\rm visc} = S_{ij}e_{ij} \tag{5.12}$$

Note that the viscous heating is not equal to the local rate at which work is being done against viscosity $-\mathbf{u}\cdot\mathbf{F}_{\rm visc}$, i.e. (minus) the last term in (5.10), which depends on the velocity rather than just on velocity gradients; part of this work is simply the transfer of momentum from one fluid element to its neighbours, and only part of it is actually dissipated. In addition, it is clear that the viscous heating must depend only on velocity gradients, and be positive; it is easily verified that $Q_{\rm visc}$ given in (5.12) satisfies both of these requirements as long as the viscosities μ and ζ are positive.

5.3 Examples of viscous flow

One much studied, presumably because of its practical engineering importance, example of viscous flow is that of flow through a pipe. Here, we shall look briefly at a physically similar but geometrically simpler case, that of flow between two planes.

The two parallel planes, separated by a distance a, are surfaces of constant y coordinate, and the flow is in the x direction (the x-component of the velocity is u), so that there is no dependence on z. The x and y components of the momentum equation (5.8) are

$$0 = -\frac{1}{\rho}\frac{\partial P}{\partial x} + \nu\frac{\partial^2 u}{\partial y^2} \tag{5.13}$$

$$0 = -\frac{1}{\rho}\frac{\partial P}{\partial y} \tag{5.14}$$

where it is assumed that the flow is incompressible, that ν is constant, and that a steady flow has been established between the two planes. Now, we see from the second of these that pressure depends only on x, and looking at the first equation we see that the pressure gradient must be constant. Integrating the x component equation twice in the y direction, we obtain

$$0 = -\frac{y^2}{2}\frac{\partial P}{\partial x} + \mu u + Ay + B \tag{5.15}$$

To find the constants A and B we need to specify the boundary conditions, i.e. the velocity u at $y = 0$ and $y = a$ (as mentioned at the end of section 5.1, all components of the velocity must go to zero at the boundaries) If we assume that the boundary plates are stationary and therefore set the velocity at both boundaries to zero (akin to flow through a pipe) – which is called plane Poiseuille flow – then we find that $A = (a/2)\mathrm{d}P/\mathrm{d}x$ and $B = 0$, so that

$$u = -\frac{a^2}{2\mu}\frac{\mathrm{d}P}{\mathrm{d}x}\left(\frac{y}{a} - \frac{y^2}{a^2}\right), \tag{5.16}$$

so the the velocity is positive if the pressure gradient is negative, and intuitively understandable result. Furthermore, we see that the velocity gradient $\partial u/\partial y$ is a constant, meaning that every fluid element experiences the same shear distortion as well as pressure drop and viscous heating rate.

It is informative to briefly mention the case of *plane Couette flow*, which is like the plane Poiseuille flow except that the pressure gradient is zero and the planes are moving relative to one another. Taking the upper plate to be moving with velocity U, the solution is

$$u = \frac{Uy}{a}. \tag{5.17}$$

The equations of viscous flow can be solved in some other, more complex situations. The first case above can be extended to the case of flow through a circular-cross-section pipe, and a popular case for investigation in astrophysics is the extension of the second case to flow between two concentric cylinders, which is called *circular Couette flow* or sometimes just *Couette flow*. In this situation, interesting effects can be seen at high rotation rates and/or low viscosity, in which case the flow is often called *Taylor-Couette flow*. Likewise, the nature of flow through a pipe changes at low viscosity; instability sets in and the flow changes from laminar to turbulent.

5.4 Similarity and dimensionless parameters

The value of ν can have an important effect on the properties of the flow. It is informative to compare the size of the viscous term in the momentum equation with the other terms, along the lines of (2.22), in a steady flow:

$$\begin{array}{ccccc} (\mathbf{u}\cdot\boldsymbol{\nabla})\mathbf{u} & = & -\dfrac{1}{\rho}\boldsymbol{\nabla}P & + & \nu\nabla^2\mathbf{u} \\ \dfrac{U^2}{L} & & \dfrac{\delta P}{\rho L} & & \dfrac{\nu U}{L^2}, \\ U^2 & & \dfrac{\delta P}{\rho} & & \dfrac{U^2}{\mathrm{Re}} \end{array} \qquad \text{where} \quad \mathrm{Re} \equiv \frac{UL}{\nu}, \tag{5.18}$$

where Re is called the *Reynolds number*. In the case of high Reynolds number the viscous term is small and the other two balance each other; conversely in the low Reynolds number case the viscous force balances the pressure gradient. Not surprisingly flows in these two regimes Re $\ll 1$ and Re $\gg 1$ have rather different properties which we explore in the following sections. Note however that a high Reynolds number does not generally mean that we can ignore viscous effects entirely; in fact as Re $\to \infty$ the behaviour of a flow only tends towards the behaviour of a perfectly inviscid fluid with $\nu = 0$ in some special cases such as small oscillations; generally the presence of even a small viscosity has a fundamental effect, as we shall see in the next section.

This brings us to an important property of the equations of hydrodynamics, namely that since they contain no fundamental constants they are scalable. For instance, in the simplest case of a steady incompressible (subsonic) flow without gravity or viscosity, the nature of the flow is determined only by the geometry and not by the magnitudes of the various parameters, which are L, U and ρ (we can consider the pressure variation δP as a function of these other parameters and so it cannot be set independently). We can set up two experiments with boundaries of the same geometry but with different densities, flow speeds and length scales, and the two flows will have identical geometry; the two flows are *similar*, hence the term *similarity flows*. This similarity is linked to the fact that it is impossible to make a dimensionless number out of combinations of L, U and ρ. In compressible flow, the sound speed c in the medium is an extra parameter and in order to make two similar flows with different L, U and ρ, we also need them to have the same Mach number $M = U/c$ so that the fractional variations in density are the same. In the same way, similar *viscous* flows must have the same Reynolds number, which is the only dimensionless number it is possible to make from combinations of L, U and ρ and ν (except trivial functions of Re). This is obviously of enormous practical value when testing for instance the aerodynamics of boats in miniature water tanks. Another good example is that jets from stellar-mass black holes look very similar to those from 'supermassive' black holes eight orders of magnitude more massive.

The set of dimensionless parameters grows with every additional component. Similar flows with gravity must have the same *Froude number*, which is the ratio of the inertia to gravity, and similar *unsteady* slows must have the same value of the *Strouhal number*, the ratio of the $(\mathbf{u} \cdot \boldsymbol{\nabla})\mathbf{u}$ and $\partial \mathbf{u}/\partial t$ terms. In summary we have

$$\text{Mach number} \qquad \frac{\text{velocity}}{\text{sound speed}} \qquad \mathrm{M} \equiv \frac{U}{c} \tag{5.19}$$

$$\text{Reynolds number} \qquad \frac{\text{inertia (steady)}}{\text{viscosity}} \qquad \mathrm{Re} \equiv \frac{UL}{\nu} \tag{5.20}$$

$$\text{Froude number} \qquad \frac{\text{inertia (steady)}}{\text{gravity}} \qquad \mathrm{Fr} \equiv \frac{U^2}{Lg} \tag{5.21}$$

$$\text{Strouhal number} \qquad \frac{\text{inertia (steady)}}{\text{inertia (unsteady)}} \qquad \mathrm{St} \equiv \frac{UT}{L} \tag{5.22}$$

Note that the Strouhal number is infinity in a steady flow, approaches zero in small-amplitude waves (where it is a function of the other numbers if the oscillations are excited within the fluid rather than by some external agent), but is often of order unity in a variety of flows.

5.5 Regimes of viscous flow: example of flow past a solid body

We looked at the difference between flows with low and high Mach numbers in sections 2.5 and 2.4; in this section we look at the effects of having low and high Reynolds numbers, i.e. high and low viscosity, using the context of a solid body moving through a fluid and examining the drag force on the body. First of all though, we look at the consequences of having no viscosity at all. Note that throughout this section, we assume the fluid is incompressible (equivalent to $M \ll 1$).

To calculate the drag force on a body, it is useful to go into the frame of reference in which the body is at rest because in this frame the flow is steady and we lose all $\partial/\partial t$ terms. Imagine a solid sphere or radius a moving at velocity v through an ideal ($\nu = 0$) incompressible fluid. Transferring to the inertial frame where the body is stationary, the fluid at a large distance from the sphere is irrotational, so the fluid must everywhere be irrotational, and we can express the velocity as the gradient of a scalar. In addition, the fluid is incompressible so that the continuity equation (1.8) reduces to $\nabla \cdot \mathbf{u} = 0$, so that the velocity potential ϕ must satisfy Laplace's equation $\nabla^2\phi = 0$. The solution of this equation is a boundary value problem. The velocity potential has to satisfy two boundary conditions – that the flow tends towards uniform at infinity and that the radial component of velocity is zero at the surface of the sphere – but we impose in this inviscid case no condition on the tangential velocity at the surface of the sphere:

$$\phi \to -vr\cos\theta \quad \text{as} \quad r \to \infty \qquad\qquad \text{and} \qquad\qquad \frac{\partial\phi}{\partial r} = 0 \quad \text{at} \quad r = a, \tag{5.23}$$

using spherical coordinates comoving with the body where r is the distance from the centre of the sphere and θ is the angle between the radius line and the direction of oncoming fluid. From undergraduate courses in electrostatics for instance, we know that the solutions to the Laplace equation in spherical coordinates are:

$$\phi = A + (Br + Cr^{-2})\cos\theta + (Dr^2 + Er^{-3})(3\cos^2\theta - 1) + ... \tag{5.24}$$

We can obviously ignore A and it follows from the boundary condition at infinity that D and all coefficients of higher positive powers of r are zero. The radial derivative of ϕ at $r = a$ is

$$\left(\frac{\partial\phi}{\partial r}\right)_{r=a} = (B - 2Ca^{-3})\cos\theta - 3Ea^{-4}(3\cos^2\theta - 1) - .... \tag{5.25}$$

and since this must be zero for all θ, E and higher coefficients must vanish. We also see of course that $B = 2Ca^{-3}$, leaving us with

$$\phi = -v\left(r + \frac{a^3}{2r^2}\right)\cos\theta. \tag{5.26}$$

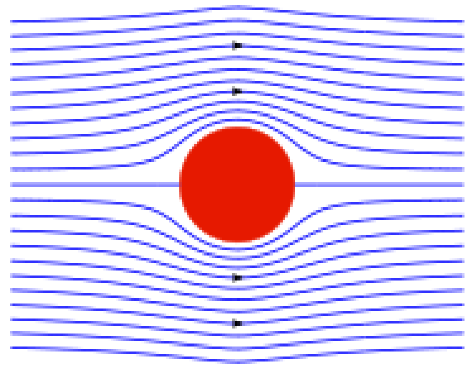
The appearance of this velocity field is illustrated in fig. 5.1.

To calculate the drag force on the sphere we need now to integrate $P\cos\theta$ over the surface. From Bernoulli's equation we have

$$\frac{1}{2}(\nabla\phi)^2 + \frac{P}{\rho} = \text{const.} \tag{5.27}$$

We can immediately see from (5.26) that the flow speed is symmetrical upwind and downwind, meaning that the pressure must also be symmetrical. The consequence of this is that the drag force vanishes! In fact we could have arrived at this conclusion much more quickly: in the steady state it is easy to see that no work is done on the fluid because after it passes by the sphere it returns exactly to its original state. There is no mechanism in an inviscid fluid to convert kinetic energy to thermal energy.

Figure 5.1: The flow of an inviscid incompressible fluid past a sphere.



If we take this solution for the velocity field and calculate the corresponding size of the viscous term in the momentum equation we find that for high Reynolds numbers this term should be negligible compared to the other terms, and yet we know from experience that the drag force remains very much non-zero even at very high Reynolds numbers ($> 10^{12}$). The drag coefficient (defined below) as a function of Reynolds number is plotted in fig. 5.2 (left panel). This shows that the assumption of zero viscosity can be an incredibly bad approximation for a fluid with low viscosity. The reason for this lies at the boundary between fluid and solid: the fluid velocity parallel to the surface of the sphere is in a fluid with a finite viscosity constrained to go to zero at the boundary. There is no solution including this extra boundary condition in the irrotational potential flow picture (we speak of an 'overconstrained' problem), so we must accept that there is at least some region in which the flow becomes rotational, i.e. develops a non-zero vorticity. In the case of low viscosity, this occurs only in a thin boundary layer near the surface of the object as well as sometimes in a larger volume behind the object, depending on its geometry. An inviscid irrotational solution applies elsewhere, but this boundary layer makes all the difference to the drag force.

It is possible to calculate the drag force for very low Reynolds number. To do the complete calculation is tedious and if we just want astrophysical accuracy we can make do with a dimensional argument. Ignoring the inertial term in the momentum equation and equating the pressure gradient to the viscous term gives

$$\frac{1}{\rho}\boldsymbol{\nabla} P = \nu\nabla^2\mathbf{u} \qquad \Rightarrow \qquad F_{\rm drag} \approx L^2\delta P \sim \rho\nu L U \tag{5.28}$$

since the drag force can be thought of as the integration over the surface area L^2 of the body of the pressure variation. This type of flow is called the *Stokes regime*. The full calculation introduces just a numerical factor (in the case of a spherical body, a factor of 6π is introduced if L is the radius of the sphere). We can now repeat the exercise for high Reynolds numbers:

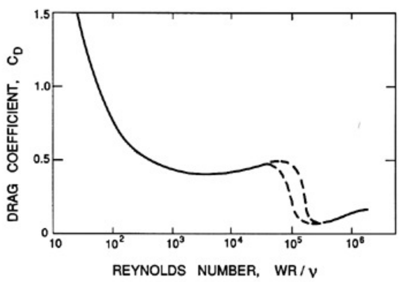
$$\frac{1}{\rho}\boldsymbol{\nabla} P = -(\mathbf{u}\cdot\boldsymbol{\nabla})\mathbf{u} \qquad \Rightarrow \qquad F_{\rm drag} \approx L^2\delta P \sim \rho L^2 U^2. \tag{5.29}$$

Here we are of course also missing a numerical factor. It is common in the literature to write the drag force as

$$F_{\rm drag} = \frac{1}{2}u^2 C_{\rm d}\rho A \tag{5.30}$$

where u is the speed of the object through a stationary medium, A is the cross-sectional area of the object as viewed from the direction of the oncoming fluid and $C_{\rm d}$ is the numerical factor (the 'drag coefficient') which depends on the geometry of the body as well as on the Reynolds number. The drag coefficient for a sphere is plotted in fig. 5.2; the shape of this curve is a consequence of various phenomena to do with boundary layers and ‘turbulence’, which it is not necessary to explore here in detail.

Figure 5.2: *Below:* Drag coefficient of a sphere as a function of the Reynolds number. *Right:* Flow past the ball at various Reynolds numbers (from Kundu & Cohen).



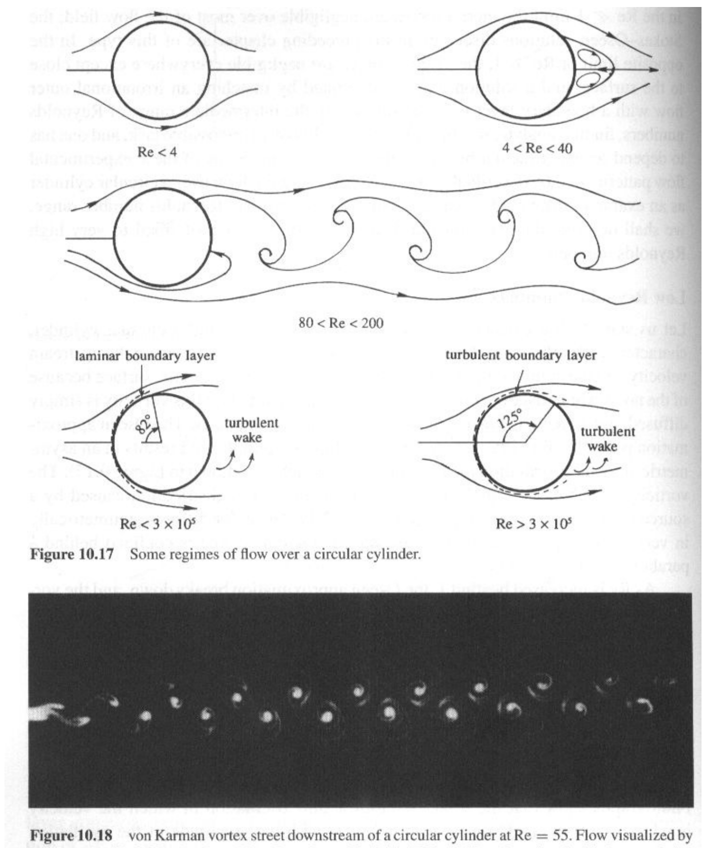


Figure 10.17 Some regimes of flow over a circular cylinder.

Figure 10.18 von Karman vortex street downstream of a circular cylinder at Re = 55. Flow visualized by condensed milk. S. Taneda, *Jour. Phys. Soc., Japan* **20**: 1714–1721, 1965, and reprinted with the permission of The Physical Society of Japan and Dr. Sadatoshi Taneda.

5.6 Boundary layers

Boundary layers are important in many engineering as well as geophysical contacts and occasionally appear in astrophysics. We take a brief look at this very important phenomenon here.

Consider the flow around a thin plate aligned parallel to the flow. In the inviscid case, the velocity field is completely unaffected by the presence of the plate; however in a real fluid we must have $\mathbf{u} = \mathbf{0}$ at the surface of the plate. From experiments we know that the inviscid solution does still apply in the bulk of the volume but that there is a thin boundary layer where the viscous force is comparable to the other forces. This layer contains a strong velocity shear, which mean that vorticity is generated. The thickness of this boundary layer depends on the viscosity – it is thicker if the viscosity is greater. Furthermore, the layer grows downstream, and it is possible to estimate this growth (see exercise below). Boundary layers may be either laminar or become turbulent if a shear instability develops.

5.7 Heat diffusion

We know from everyday experience that heat flows from hot to cold. Without going to deeply into the thermodynamics, we can describe the flow of heat as

$$\mathbf{F}_{\text{heat}} = -K\boldsymbol{\nabla}T, \tag{5.31}$$

where the heat flux has dimensions of energy per unit time per unit area and K is the thermal conductivity. We have made an implicit assumption here that the conductivity is isotropic, i.e. that the fluid conducts heat equally in all directions.[1] Now, to look at the effect of heat conduction on a body of fluid we need to know the net inflow/outflow of heat into/out of a fluid element; the net heat energy influx per unit volume per unit time is given by

$$Q_{\text{vol}} = \boldsymbol{\nabla}\cdot(K\boldsymbol{\nabla}T). \tag{5.32}$$

which fits into the heat equation in the following way

$$\frac{\mathrm{d}T}{\mathrm{d}t} = \ldots\ldots\ldots\ldots\frac{1}{\rho c_p}\boldsymbol{\nabla}\cdot(K\boldsymbol{\nabla}T) \tag{5.33}$$

and if K can be assumed constant throughout the fluid, then we can write the term above simply as $\kappa\nabla^2 T$ where we have defined a thermal diffusivity

$$\kappa \equiv K/(\rho c_p) \qquad \text{and} \qquad \mathrm{Pr} \equiv \nu/\kappa, \tag{5.34}$$

where we have also defined the Prandtl number as the ratio of the thermal and kinetic diffusivities (both have units of area per unit time). In some situations, such as convection in stars, it is thought that a flow can behave quite differently according to whether the Prandtl number is greater than or less than unity.

Exercises

5.1 Momentum equation

Verify that the momentum equation in Einstein summation notation (5.1) with (5.2) is equivalent to the vector-notation form (2.1), except for the gravity and viscous parts. Show that it is generally not possible to incorporate body forces, such as gravity, into the divergence-of-a-tensor form of the momentum equation.

5.2 Model testing

We are designing a boat which will sail at 4 m/s and is 8 m long. We shall assume that the drag on the boat will be entirely due to buoyancy effects, i.e. generation of gravity waves. If we construct a model 50 cm long, at what speed should the water in the testing tank be moving past the model? [Hint: the Froude numbers must be the same.]

[1] Isotropic conductivity is a valid assumption in most contexts of interest but in some cases, such as low-density plasmas where the mean-free path is greater than the gyration radius associated with the magnetic field present, the scalar K must be replaced by a tensor.

5.3 Growth of boundary layer

An infinitely thin solid sheet is inserted into a uniform flow such that sheet and flow are parallel. Argue that in the inviscid case, the flow is not affected. In the case of finite viscosity, show that a boundary layer forms and estimate the thickness of the boundary layer as a function of distance downstream. [Hint: look at the relative sizes of terms in the momentum equation.]

Chapter 6

Shocks

In this section we look at a phenomenon where a discontinuity in the density and pressure of a fluid appears when there is supersonic motion of some kind. This phenomenon is in some sense non-fluid in that the relevant length scale is of the order of the mean-free path of the particles, but we can still derive useful results without considering microscopic processes.

6.1 Viscous vs. pressure gradient force

Let us first examine in what situation the viscous force (normally first order in λ/L, the ratio of mean free path to characteristic length scale of the system under consideration) is comparable to the pressure gradient force. Looking at the momentum equation

$$\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t} = \frac{1}{\rho}\boldsymbol{\nabla} P + \nu \boldsymbol{\nabla}^2 \mathbf{u} \tag{6.1}$$

we see that the ratio of pressure to viscous force is

$$\frac{F_P}{F_{\mathrm{visc}}} \sim \frac{P}{\rho L}\left(\frac{\nu U}{L^2}\right)^{-1} \approx \frac{L c_s^2}{U \nu} \approx \frac{c_s}{U}\frac{L}{\lambda} \tag{6.2}$$

where U is the typical velocity. The gas relation $\nu \approx c_s \lambda$ has been used. Clearly, while $L \gg \lambda$ the viscous force is relatively unimportant; however if the motion becomes supersonic ($U > c_s$) and a discontinuity appears ($L \approx \lambda$) then the viscous force becomes important. In this situation we should expect kinetic energy to be dissipated into heat.

6.2 The jump conditions

To derive relations between the pressure, density and velocity on either side of a discontinuity it is easiest to go into the inertial frame in which the shock is at rest. In the following, quantities on either side of the discontinuity have the subscripts 0 and 1. The fluid has velocity component u perpendicular to the discontinuity. By consideration of mass conservation – mass entering and leaving the discontinuity per unit time per unit area – we have

$$\rho_0 u_0 = \rho_1 u_1. \tag{6.3}$$

Next we can consider the rate of change of momentum contained within a volume spanning the discontinuity, which in the steady state (in the case of unsteady flow, we consider a short interval of time) must vanish. Contributions to increase momentum come from the momentum of the fluid entering the volume as well as the pressure P_0 acting on the volume; these must be balanced by momentum loss on the other side:

$$\rho_0 u_0^2 + P_0 = \rho_1 u_1^2 + P_1. \tag{6.4}$$

We now need to use conservation of energy. Again considering a volume spanning the discontinuity and equating the rate of change of its energy, made up of internal and kinetic energy flux inwards and outwards plus '$p\mathrm{d}V$' work done, to zero, we have

$$\rho_0 u_0 \left(\epsilon_0 + \frac{u_0^2}{2} + \frac{P_0}{\rho_0} \right) = \rho_1 u_1 \left(\epsilon_1 + \frac{u_1^2}{2} + \frac{P_1}{\rho_1} \right) \tag{6.5}$$

where ϵ is the internal energy per unit mass, a function of pressure and density. Of course, this is just an expression of Bernouilli's equation (2.9).[1] Finally we can consider velocity parallel to the discontinuity, where conservation of momentum gives us

$$\rho_0 u_0 \mathbf{v}_0 = \rho_1 u_1 \mathbf{v}_1, \tag{6.6}$$

where $\mathbf{v}$ is the component of velocity parallel to the discontinuity.

Now, there are two possible types of solution, assuming both densities and pressures are non-zero. First, we can have $u_0 = u_1 = 0$. In this case, we see from (6.3) and (6.6) that ρ and $\mathbf{v}$ are unconstrained on both sides. However from (6.4) we have $P_0 = P_1$. This kind of discontinuity is called a *tangential discontinuity*, and in many contexts (particularly if the fluids on either side are of a different type or origin) it is called a *contact discontinuity*. Note that if the parallel velocities are not equal, i.e. if $\mathbf{v}_0 \neq \mathbf{v}_1$, then the flow is generally unstable (see section 4.3).

The second solution has non-zero perpendicular velocities, and is called a *shock* or *shock wave*. In this case, we define the axes in such a way that both u_0 and u_1 are positive. From (6.3) and (6.6) we see that $\mathbf{v}_0 = \mathbf{v}_1$. Applying mass conservation to (2.9) gives

$$h_0 + \frac{u_0^2}{2} = h_1 + \frac{u_1^2}{2}, \tag{6.7}$$

where internal energy ϵ and P/ρ have been brought together into enthalpy h. Now substituting (6.3) into (6.6) we have

$$\rho_0 u_0 (u_0 - u_1) = P_1 - P_0. \tag{6.8}$$

Given that $\rho_0 u_0$ is positive, there are now two possibilities: either $u_0 > u_1$, $P_0 < P_1$ and $\rho_0 < \rho_1$ or $u_0 < u_1$, $P_0 > P_1$ and $\rho_0 > \rho_1$. It is left as an exercise for the student to show formally that these two possibilities represent an increase and a decrease in entropy; since we know that entropy must increase, only one set of solutions is permissible. If we define u_0 as the velocity upstream of the shock and u_1

[1] Here we are simply equating the inward and outward energy fluxes on two surfaces of a volume. The energy flux through the sides of the volume is made to vanish by making the volume infinitesimally flat whilst still containing the discontinuity.

as the velocity downstream, an increase in entropy requires that $u_0 > u_1$. This means that the gas is compressed and heated on its passage through the shock. The energy to do this can only come from the decrease in kinetic energy. In fact, the opposite of this would apparently violate the second law of thermodynamics, according to which it is impossible to construct a system whose sole result is the conversion of energy from heat to kinetic energy.

The change of state of the gas happens via microscopic processes in the shock, which generally has a thickness of order the mean free path of the particles. An interesting feature is that the change in the thermodynamic state of the gas is determined entirely by the macroscopic quantities on either side; the shock itself can be thought of as adjusting itself to meet the external requirements placed upon it, regardless of the fluid's microscopic properties. For instance, a shock can be passed at a given speed through two fluids which are identical except for their viscosities; in the less viscous fluid the shock discontinuity will become thinner to allow it to dissipate the same energy as in the more viscous fluid, and it is impossible to tell the difference between the two if one just measures the macroscopic quantities. This behaviour is also seen for instance in magnetic reconnection.

To look at the interdependence of the variables in (6.3), (6.4) and (6.7) it is first helpful to express the enthalpy in terms of pressure and density: in an ideal gas, $h = (P/\rho)\gamma/(\gamma - 1)$. After some algebra, we arrive at:

$$\frac{P_1}{P_0} = 1 + \frac{2\gamma}{\gamma + 1}\left(M_0^2 - 1\right), \tag{6.9}$$

where $M \equiv u/c$ is the Mach number, the velocity as a fraction of the sound speed. Note that the Mach numbers on either side of the shock are the velocities as fractions of the sound speed on the respective side. The fluid enters the shock supersonically (remember that we defined the directions such that $P_1 > P_0$). After more algebra we have

$$M_1^2 = \frac{(\gamma - 1)M_0^2 + 2}{2\gamma M_0^2 + 1 - \gamma}. \tag{6.10}$$

It is easily verified that $M_1 < 1$ (except in the trivial solution where all quantities are the same on both sides) and that it tends towards unity and M_0 tends also towards unity. Therefore, the material enters supersonically and exits subsonically. Furthermore, as $M_0 \to \infty$, $M_1 \to (\gamma - 1)/2\gamma$, so there is a limit to the conversion of kinetic energy into heat. The density and velocity ratios are

$$\frac{\rho_1}{\rho_0} = \frac{(\gamma + 1)M_0^2}{(\gamma - 1)M_0^2 + 2} = \frac{u_0}{u_1}. \tag{6.11}$$

A very important point to note at this juncture is the limit on the compression factor of $(\gamma + 1)/(\gamma - 1)$, which is equal to 4 for a monatomic gas, in which case no more than 15/16 of the kinetic energy can be converted into heat. Finally, the temperature ratio is

$$\frac{T_1}{T_0} = 1 + \frac{2(\gamma - 1)}{(\gamma + 1)^2} \cdot \frac{(\gamma M_0^2 + 1)(M_0^2 - 1)}{M_0^2}. \tag{6.12}$$

As with pressure, there is no limit on the temperature ratio.

In many situations a shock is in fact 'stationary' in some sense, for instance where material falls onto a compact star it becomes supersonic under the influence of gravity and then reaches a shock near the stellar surface. In this example, gravitational energy is released and converted into kinetic energy, and then in the shock this energy is converted to heat, and then the heat energy converted into electromagnetic radiation. Another example of a stationary shock is where the solar wind meets the boundary of the Earth's ionosphere. On Earth, a more familiar example would be a stationary hydraulic shock, which is looked at in section 6.3. In other contexts, it makes sense to think of the shock as moving into a fluid at rest, for instance after an explosion such as a supernova.

6.3 Hydraulic jumps

It was shown in section 4.1 that shallow water waves propagate with a speed $\sqrt{gh}$ where h is the depth of the water, and short-wavelength waves propagate more slowly. This is therefore the maximum speed at which disturbances can propagate in the shallow water system. Now imagine a discontinuity in the depth of water parallel to the line $x = 0$, with depth h_1 on the left and h_2 on the right. Relative to the discontinuity, the water is flowing into the discontinuity from the left at speed u_1 and away from it on the right at speed u_2. Mass conservation gives

$$h_1 u_1 = h_2 u_2 \tag{6.13}$$

while momentum conservation gives

$$h_1 u_1^2 + \frac{1}{2} g h_1^2 = h_2 u_2^2 + \frac{1}{2} g h_2^2 \tag{6.14}$$

where the first term on each side represents momentum advected into and out of a volume containing the shock and the second term on each side represents the pressure exerted on that volume. [The density of the water obviously drops out of both of these equations.] If we know h_1 and u_1 then we can calculate h_2 and u_2. If we now considered the flow of kinetic and potential energy $(hu^2/2 + gh^2/2)u$ into and out of the volume, we would find the problem had become overconstrained. This is because our water has only one equivalent of a thermodynamic variable, h, compared to two degrees of freedom in the gas considered in the previous section. But does this mean energy is not conserved? The answer is that energy is converted into a form we have not considered, i.e. disordered kinetic and heat. The rate at which the energy is converted is

$$\begin{aligned} q &= \frac{u_1}{2}(h_1 u_1^2 + g h_1^2) - \frac{u_2}{2}(h_2 u_2^2 + g h_2^2) \\ &= \frac{\rho u_1 h_1 g (h_1^2 + h_2^2)(h_2 - h_1)}{4 h_1 h_2} \end{aligned} \tag{6.15}$$

where (6.13) and (6.14) have been used. Since this energy must be positive, we can see that $h_2 > h_1$. With the help of some algebra it is also possible to show that $u_1 > \sqrt{gh_1}$ and $u_2 < \sqrt{gh_2}$.

Although this kind of shock is of little importance in astrophysics, it is helpful to study it because it is literally a kitchen sink experiment.

Chapter 7

Vorticity and rotating fluids

In this section we look at vorticity in more detail as well as various phenomena in rotating systems. First, recall that in section 2.6 vorticity was defined as $\boldsymbol{\omega} \equiv \boldsymbol{\nabla} \times \mathbf{u}$ and that circulation Γ was defined as the integral of velocity around a (comoving) closed loop. Recall also that circulation is conserved in an inviscid, barotropic flow where body forces are conservative.

Before looking at vortices and their behaviour, consider a small spherical fluid element of radius a which is rotating with angular velocity Ω. The circulation of a loop around the 'equator' of this element is then $2\pi a^2\Omega$. From (2.31) we see that the this can be expressed in terms of the vorticity in the following way

$$\Gamma = \oint \mathbf{u} \cdot \mathrm{d}\mathbf{s} = \int \boldsymbol{\omega} \cdot \mathrm{d}\mathbf{S} \tag{7.1}$$

$$2\pi a^2 \Omega = \pi a^2 \omega \tag{7.2}$$

$$2\boldsymbol{\Omega} = \boldsymbol{\omega} \tag{7.3}$$

where the surface integral is taken over the equatorial plane, which is perpendicular to the vorticity. The vorticity is simply double the angular velocity of a fluid element.

7.1 Vortices

Here, we take a look at vorticity generation, vortex tubes, and the interaction of vortices. Imagine two basic velocity fields, with

$$u_\theta = \frac{1}{2}\omega\varpi \qquad \text{and} \qquad u_\theta = \frac{\Gamma}{2\pi\varpi} \tag{7.4}$$

where the former represents solid-body rotation and the latter is irrotational (i.e. zero vorticity) everywhere except for a singularity on the axis. An everyday example of solid-body rotation would be a cup of coffee which has come into a rotational equilibrium inside a microwave. The latter kind, called a *line vortex*, does not really exist in nature; rather, the singularity is replaced with an inner cylinder of solid-body rotation, surrounded as before by irrotational flow. This is called a *Rankine vortex*; it can alternatively be described as a cylinder of constant vorticity (a *vortex tube*) surrounded by zero vorticity. A tornado can well be approximated by this kind of vortex; note that the radius of the inner region is very small compared to the sky and cloud from which the tornado forms, so if one is only interested in the longer length scales one can make the approximation of a perfect line vortex.

We can draw vorticity lines in the same way as we can draw streamlines. Note that the divergence of vorticity is zero. Like magnetic fields in a conducting fluid, vortex lines can be thought of as being 'frozen' into the fluid. Imagine a patch S on the surface of a vortex tube. Since the vorticity is everywhere parallel to this surface, the circulation around its perimeter is zero. As the fluid moves around, Kelvin's circulation theorem tells us that the circulation must always remain zero; therefore so must the vorticity perpendicular to the surface, so that the comoving patch must remain on the surface of the vortex tube. We conclude that vortex lines are frozen into the fluid. This is analogous to the freezing of magnetic field lines into a conducting fluid (section 8.7).

Figure 7.1: Tornadoes.

7.2 Vorticity generation

Let us take the curl of the momentum equation (5.8), taking care first to write out the Lagrangian time derivative as Eulerian time derivative and advective term:

$$\begin{aligned}
\frac{\partial \boldsymbol{\omega}}{\partial t} + \boldsymbol{\nabla} \times (\mathbf{u} \cdot \boldsymbol{\nabla} \mathbf{u}) &= -\boldsymbol{\nabla} \times \left(\frac{1}{\rho} \boldsymbol{\nabla} P\right) + \boldsymbol{\nabla} \times \mathbf{g} + \boldsymbol{\nabla} \times (\nu \boldsymbol{\nabla}^2 \mathbf{u}), \\
\frac{\partial \boldsymbol{\omega}}{\partial t} + \boldsymbol{\nabla} \times (\boldsymbol{\omega} \times \mathbf{u}) &= -\boldsymbol{\nabla} \left(\frac{1}{\rho}\right) \times \boldsymbol{\nabla} P + \nu \nabla^2 \boldsymbol{\omega}, \\
\frac{\mathrm{d}\boldsymbol{\omega}}{\mathrm{d}t} &= \frac{1}{\rho^2} \boldsymbol{\nabla} \rho \times \boldsymbol{\nabla} P - \boldsymbol{\omega}(\boldsymbol{\nabla} \cdot \mathbf{u}) + (\boldsymbol{\omega} \cdot \boldsymbol{\nabla})\mathbf{u} + \nu \nabla^2 \boldsymbol{\omega}, \qquad (7.5)
\end{aligned}$$

where we have used the fact that the gravitational force is conservative (and assumed that any other body forces present are also conservative) as well as some vector calculus identities. The first term on the right vanishes in a barotropic flow, the next two terms on the right contain vorticity, and the last term is the viscosity, so that we recover the conclusion (section 2.6) that if vorticity is zero everywhere at some point in time, it is zero at all other times, provided that the flow is barotropic and inviscid and that the body forces are conservative.

It is informative to examine the physical meaning of each of the terms on the right-hand side of (7.5). On the right-hand side we have:

(a) The baroclinic term. The pressure acts normal to the surface of a fluid element; imagine that the fluid element is spherical: the total pressure force, equal to the sum of the pressure force on surface elements, acts through the centre of the fluid element and is parallel to the gradient of pressure. The fluid element can be made to rotate if the centre of mass does not lie on the line of net pressure gradient force; however if $\rho = \rho(P)$ then this is impossible. Therefore the pressure gradient force can only bring about a linear acceleration of the fluid element and cannot act as a torque thereupon: hence the constancy of circulation. Note that vorticity is an expression of the rotation of a fluid element about its own centre of mass, not about any other point – fluid elements with zero vorticity are still free to move in circles around each other. This is the reason that gravity cannot make the fluid element rotate, as it acts through the centre of mass.

   A good example of where the baroclinic term can generate torque is after heating in the atmosphere, for instance after an explosion has created a hot bubble. If the bubble has uniform temperature then vorticity is created at its boundary when the bubble rises, and if the bubble has a temperature profile which gradually decreases outwards then vorticity is generated throughout the volume; the result of this is that the bubble deforms into a rising vortex ring.

(b) The second term simply represents conservation of angular momentum as the fluid is compressed or expands; clearly the vorticity should increase if the density increases, which happens if the divergence of the velocity is negative, hence the minus sign above. One can also think in terms of the vortex lines being squashed together when the density rises, increasing the density of vortex lines and hence the vorticity.

(c) The third term represents conservation of angular momentum during 'stretching' of vortices. If a fluid element is stretched out along the direction of the vorticity whilst its volume remains constant, then the moment of inertia of the fluid element has decreased and so the angular velocity and vorticity must increase. Again one can think in terms of vortex lines being squashed closer together and the vorticity, which is the density of vortex lines, increasing. Note here an important difference between two- and three-dimensional flow – it is impossible in two dimensions to stretch a vortex.

(d) The viscous term, which has the same form as that in the original momentum equation, simply causes vorticity to leak from local maxima to minima. It causes an isolated vortex tube of the kind described in the previous section to become broader. It also generates vorticity in boundary layers.

In addition, non-conservative body forces can also generate vorticity; we examine the role of the Coriolis force in section 7.4.

7.3 Behaviour of vortices

We now look very briefly at the interaction of vortices with each other and with solid boundaries. First we approximate a vortex with a line vortex as described above. Now, since the flow around a line vortex is irrotational, we can write $\mathbf{u} = \nabla\phi$ where ϕ is a scalar potential. When two or more vortices are present, we can simply add together the scalars belonging to each vortex in isolation to produce the resultant velocity field. Recalling from above that vortex lines are 'frozen' into the fluid, we see that each individual vortex must move according to the sum of the velocity fields belonging to each of the other vortices in isolation.

First let us consider two parallel vortices of equal magnitude and sense. Associated with each one is a velocity field in which the other moves around; in this case they encircle each other. Far away from the vortices, the two velocity fields cancel each other out. If, on the other hand, the two vortices have a circulation of the opposite directions, the pair will move together in a straight line perpendicular to the line between them. Anyone who has tried rowing will be familiar with vortex pairs created by an oar travelling surprisingly large distances behind the boat.

Now imagine a vortex ring such as that created by a skillful smoker or alternative a ring which forms after a bomb explodes in the atmosphere. Each section of the ring moves according to the velocity field associated with the rest of the ring, with the result that the ring propagates forwards much further than one would otherwise expect.

As a vortex approaches a boundary, we can predict what will happen with the method of images where one imagines removing the boundary and placing an image vortex on the other side. In the case of a single vortex near a boundary, we can recreate the flow with an image vortex of opposite spin, since the velocity field where the boundary once was is now constrained to be parallel to it. Of course in reality there will be a small difference between the two flows, namely that there will be a thin boundary layer, but that can ignored in the bulk of the volume. So, a single vortex near a wall will move parallel to the wall.

7.4 The momentum equation in a rotating frame of reference

A frame of reference which is rotating with angular velocity $\boldsymbol{\Omega}$ with respect to the non-rotating, inertial frame, the Lagrangian time derivatives of scalar quantities such as density and temperature must of course be the same. However, velocity is frame dependent and so generally the Lagrangian derivative of velocity is different in the two frames. For instance, a fluid element which is stationary in the rotating frame must be experiencing an acceleration in the inertial frame. It turns out that the comoving derivatives of velocity in the two frames are related by

$$\left(\frac{\mathrm{d}\mathbf{u}_\mathrm{I}}{\mathrm{d}t}\right)_\mathrm{I} = \left(\frac{\mathrm{d}\mathbf{u}_\mathrm{R}}{\mathrm{d}t}\right)_\mathrm{R} + \boldsymbol{\Omega}\times\boldsymbol{\Omega}\times\mathbf{r}_\mathrm{R} + 2\boldsymbol{\Omega}\times\mathbf{u}_\mathrm{R} + \frac{\mathrm{d}\boldsymbol{\Omega}}{\mathrm{d}t}\times\mathbf{r}_\mathrm{R} \tag{7.6}$$

where the comoving derivative on the left in the inertial frame is equal to the comoving derivative of the velocity in the rotating frame plus three acceleration terms, where the subscripts I and R denote inertial and rotating frames. A proper derivation of this can be found in any textbook, but the origin of these terms can be understood intuitively in the following way:

(a) The first of the three terms is the centrifugal acceleration: to make anything move in a circle (i.e. accelerate towards the centre of the circle) it is necessary to provide an inwards-directed force (for instance from gravity) of magnitude $\varpi\Omega^2$ per unit mass where ϖ is the distance from the axis of rotation. In the rotating frame in which the fluid element is stationary, there is no longer any acceleration towards the centre, The original inwards force therefore must be balanced by an extra force directed outwards, the centrifugal force. Note that the centrifugal force can be written $\hat{\boldsymbol{\varpi}}\varpi\Omega^2$ where $\hat{\boldsymbol{\varpi}}$ is the cylindrical radius unit vector.

(b) Unlike the centrifugal force, the Coriolis force arises only when fluid is moving within the rotating frame; it can be thought of as accounting for conservation of angular momentum. When a fluid element moves to a larger radius whilst preserving its angular momentum it must attain a smaller angular velocity and therefore begin to drift backwards in the azimuthal direction relative to the rotating frame. Similarly, if a fluid element moves in the azimuthal direction, i.e. with a different angular velocity to the frame of reference, it experiences a different centrifugal acceleration from the frame and so accelerates in the (cylindrical) radial direction relative to the frame. Note that there is no component of the Coriolis force in the direction of the angular velocity $\boldsymbol{\Omega}$. Alternatively, one might think of standing on the north pole and firing a projectile horizontally – in the inertial frame the object moves in a straight line but since the Earth is rotating, an observer on the ground will see the object curve towards the west.

(c) Finally, the third term represents the apparent acceleration a fluid element experiences when the underlying rotation of its reference frame changes. In most situations we use a reference frame with constant $\boldsymbol{\Omega}$ and can ignore this term, as we do in all of the following.

Putting this together into a rotating-frame momentum equation, we have

$$\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t} = -\frac{1}{\rho}\boldsymbol{\nabla} P - \boldsymbol{\nabla}\Phi_{\mathrm{eff}} - 2\boldsymbol{\Omega}\times\mathbf{u} + \nu\nabla^2\mathbf{u}. \tag{7.7}$$

The viscous term has been simplified in the usual way and $\mathbf{r}$ and $\mathbf{u}$ are the position and velocity vectors in the rotating frame, the subscript R having been dropped. The centrifugal term has been absorbed into the gravitational potential, which is possible because it is conservative:

$$\mathbf{g}_{\mathrm{eff}} \equiv \mathbf{g} - \boldsymbol{\Omega}\times\boldsymbol{\Omega}\times\mathbf{r} = -\boldsymbol{\nabla}\Phi_{\mathrm{grav}} + \boldsymbol{\nabla}\left(\frac{1}{2}\Omega^2\varpi^2\right) \tag{7.8}$$

$$= -\boldsymbol{\nabla}\Phi_{\mathrm{eff}} \qquad \text{where} \qquad \Phi_{\mathrm{eff}} \equiv \Phi_{\mathrm{grav}} - \frac{1}{2}\varpi^2\Omega^2. \tag{7.9}$$

7.5 The centrifugal force and the von Zeipel paradox

Since it can be absorbed into the gravitational potential, the centrifugal force merely produces a change in the shape of equipotential surfaces. This is the reason that rotating stars and planets are flattened. In fact rotational flattening is clearly visible with a small telescope trained on Jupiter or Saturn (see fig. 7.2).

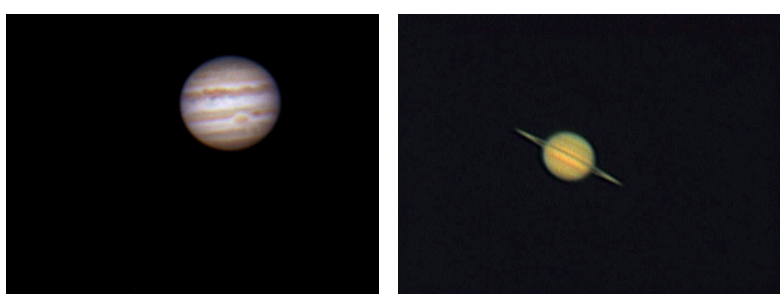


Figure 7.2: Rotational flattening of Jupiter and Saturn is visible in these images taken with a 28cm telescope. *Pictures taken by Tim Kent.*

An interesting effect of the centrifugal force in stars was discovered in 1924 by von Zeipel. Now, in a static equilibrium in a rotating frame of reference we must have

$$\frac{1}{\rho}\boldsymbol{\nabla} P = -\boldsymbol{\nabla}\Phi_{\text{eff}} \tag{7.10}$$

from which we see, by taking the curl, that

$$\boldsymbol{\nabla}\rho \times \boldsymbol{\nabla} P = \mathbf{0}. \tag{7.11}$$

So we must have a barotropic relation $\rho = \rho(P)$, and so the contours of pressure, density, temperature and effective potential Φ_{eff} must all coincide. However, the flux of heat is proportional to $-\boldsymbol{\nabla} T$, so the flux must be greater at the poles of the star than around the equator, and the divergence of this flux must be equal to the rate of nuclear energy generation, which is a function of the local thermodynamic state but not of its gradient. Thus the rotation places demands on the nuclear energy generation which in general cannot be met; there is therefore no static equilibrium in a rotating star. This is called the *von Zeipel paradox*. The solution is to have a large-scale circulation, where motion adds a Coriolis term to (7.10) as well as advecting heat.

7.6 Vorticity equation in a rotating frame

In section 7.2 we derived an equation for the Eulerian and Lagrangian time derivatives of vorticity. Extending this procedure to include the Coriolis force, we obtain

$$\begin{aligned}
\frac{\partial \boldsymbol{\omega}}{\partial t} + \boldsymbol{\nabla} \times (\mathbf{u}\cdot\boldsymbol{\nabla}\mathbf{u}) &= -\boldsymbol{\nabla} \times \left(\frac{1}{\rho}\boldsymbol{\nabla} P\right) + \boldsymbol{\nabla} \times \mathbf{g} - \boldsymbol{\nabla} \times \mathbf{2}\Omega \times \mathbf{u} \\
\frac{\partial}{\partial t}(\boldsymbol{\omega} + 2\boldsymbol{\Omega}) + \boldsymbol{\nabla} \times ((\boldsymbol{\omega} + 2\boldsymbol{\Omega}) \times \mathbf{u}) &= -\boldsymbol{\nabla}\left(\frac{1}{\rho}\right) \times \boldsymbol{\nabla} P, \\
\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{\omega} + 2\boldsymbol{\Omega}) &= -\boldsymbol{\nabla}\left(\frac{1}{\rho}\right) \times \boldsymbol{\nabla} P - (\boldsymbol{\omega} + 2\boldsymbol{\Omega})(\boldsymbol{\nabla}\cdot\mathbf{u}) + [(\boldsymbol{\omega} + 2\boldsymbol{\Omega})\cdot\boldsymbol{\nabla}]\mathbf{u}
\end{aligned} \tag{7.12}$$

where on the second and third lines Ω has been taken into the time derivative on the left-hand side on the condition that its Eulerian and Lagrangian time derivatives, respectively, are zero. [The diffusion term included in (7.5) has been ignored here.] So, vorticity has simply been replaced by the *absolute vorticity* $\boldsymbol{\omega} + 2\boldsymbol{\Omega}$; in rotating systems ω is referred to as the *relative vorticity*. Furthermore, by comparison with section 2.6 we find the equivalent of Kelvin's circulation theorem, that if we define

$$\Gamma \equiv \oint (\mathbf{u} + \boldsymbol{\Omega} \times \mathbf{r}) \cdot \delta\mathbf{s} \tag{7.13}$$

that $\mathrm{d}\Gamma/\mathrm{d}t = 0$ in the absence of baroclinicity, viscosity and non-conservative body forces, as before. Applying Stokes' theorem we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\int (\boldsymbol{\omega} + 2\boldsymbol{\Omega}) \cdot \mathrm{d}\mathbf{S} = 0. \tag{7.14}$$

We can see then a mechanism for generating (relative) vorticity where there was none before; if a fluid element is changes its extent in the plane perpendicular to $\boldsymbol{\Omega}$ it will start to spin relative to the rotating frame even if it did not do so initially. An example of this is when heating on the Earth causes the air over a region to warm up and expand laterally.

7.7 Inertial waves

We have seen in previous sections how the pressure gradient force and gravity can both provide a restoring force for waves; in this section we look at waves with the Coriolis force as the restoring agent. To simplify the equations (avoiding solving also for sound and gravity waves at the same time as the inertial waves we are investigating) we make the assumption of constant density, meaning that we can automatically ignore gravity since gravity has no interesting effect on a constant-density fluid except at the surface; we look here at waves internal to the fluid. The momentum equation is

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \boldsymbol{\nabla})\mathbf{u} = -\frac{1}{\rho}\boldsymbol{\nabla} P - 2\boldsymbol{\Omega} \times \mathbf{u} + \hat{\boldsymbol{\varpi}}\varpi\Omega^2. \tag{7.15}$$

We now make the usual assumption when looking at waves of small amplitude, so that we can drop the second term above. Taking the curl of the remaining terms we have

$$\frac{\partial}{\partial t}\boldsymbol{\nabla} \times \mathbf{u} = -2\boldsymbol{\nabla} \times (\boldsymbol{\Omega} \times \mathbf{u}) \tag{7.16}$$

since the curls of the centrifugal and pressure gradients forces are zero. In fact, in many situations the centrifugal force can be ignored since it can be expressed as the gradient of a scalar; as we saw above, in situations with gravity it can simply be added to the gravitational potential which often can be removed completely by an adjustment to the definition of the vertical axis, in a local analysis. Now, the term on the right of (7.16) can be rewritten in a more convenient form; with the help of a vector calculus identity $\boldsymbol{\nabla} \times (\boldsymbol{\Omega} \times \mathbf{u}) = \boldsymbol{\Omega}(\boldsymbol{\nabla} \cdot \mathbf{u}) - (\boldsymbol{\Omega} \cdot \boldsymbol{\nabla})\mathbf{u}$, and noting that the incompressible continuity equation means that the first of these two terms vanishes. Finally we set the rotation axis along the z-axis and write

$$\frac{\partial}{\partial t}\boldsymbol{\nabla} \times \mathbf{u} = 2\Omega\frac{\partial \mathbf{u}}{\partial z}. \tag{7.17}$$

As before, we now assume a solution of the form $\exp[i(\mathbf{k} \cdot \mathbf{r} - \omega t)]$; the continuity equation gives

$$\mathbf{k} \cdot \mathbf{u} = 0 \tag{7.18}$$

and the momentum equation (7.17) becomes

$$\omega\mathbf{k} \times \mathbf{u} = 2\Omega i k_z \mathbf{u}. \tag{7.19}$$

We now take the curl of both sides and noting that $\mathbf{k}$ and $\mathbf{u}$ are perpendicular we can write $\mathbf{k} \times \mathbf{k} \times \mathbf{u} = -k^2\mathbf{u}$:

$$-\omega k^2\mathbf{u} = 2\Omega i k_z \mathbf{k} \times \mathbf{u} \tag{7.20}$$

which we can compare with the previous form to give

$$\omega = 2\Omega\frac{k_z}{k} = 2\Omega\cos\theta \tag{7.21}$$

where k is the magnitude of the wavevector $\mathbf{k}$ and θ is the angle between the wavevector and the rotation axis Ω. This is a similar dispersion relation to that of internal gravity waves (4.49) in that the frequency of the oscillations depends only on the direction of the wavevector and not on its magnitude. Here, the frequency of the oscillations goes to zero when the wavevector is perpendicular to the rotation axis, i.e. when the velocity field is parallel to it. It is of course obvious from the form of the Coriolis force that motion parallel to the rotation axis experiences no restoring force.

Another interesting feature of these waves is that the energy is entirely kinetic, rather than being converted back and forth between two different forms. Instead one has to think here of energy conversion back and forth between kinetic energy of motion in two different directions.

7.8 The Taylor-Proudman theorem

Imagine motions in a rotating fluid with characteristic length-scale, time-scale and velocity L, T and U. Looking at the sizes of the various terms in the momentum equation we have

$$\frac{\partial \mathbf{u}}{\partial t} \quad + \quad (\mathbf{u}\cdot\boldsymbol{\nabla})\mathbf{u} \quad = \quad -\frac{1}{\rho}\boldsymbol{\nabla}P \quad - \quad 2\boldsymbol{\Omega}\times\mathbf{u} \tag{7.22}$$

$$\frac{U}{T} \qquad \frac{U^2}{L} \qquad \frac{\delta P}{\rho} \qquad \Omega U \tag{7.23}$$

where the centrifugal force has been ignored as it can be absorbed into the equilibrium pressure gradient force. Furthermore we assume a relation $T \sim L/U$, typical of flows rather than waves, meaning that the two terms on the left-hand side are of comparable size. The ratio of the Coriolis to inertial terms is ΩT; the inverse of this number is called the *Rossby number* after the Swedish physicist. At low Rossby number, i.e. $T^{-1} \ll \Omega$, the momentum equation can be approximated to

$$\frac{1}{\rho}\boldsymbol{\nabla}P = -2\boldsymbol{\Omega}\times\mathbf{u}. \tag{7.24}$$

We see from this that the gradient of P in the direction of the rotation axis vanishes; furthermore we see that this equation, which relates the velocity perpendicular to the rotation axis to the pressure gradient perpendicular to it, demonstrates that the gradient of the velocity field along the rotation axis also vanishes. Finally, taking the curl of this equation, we lose the left hand side if the fluid is incompressible, and using a vector identity on the remaining term (again assuming incompressibility and therefore $\boldsymbol{\nabla}\cdot\mathbf{u} = 0$) gives $(\Omega\cdot\boldsymbol{\nabla})\mathbf{u} = 0$. In summary, if the rotation axis is parallel to the z axis, we have

$$\frac{\partial w}{\partial z} = 0 \qquad \text{and} \qquad \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = 0 \qquad \text{and in general} \qquad \frac{\partial}{\partial z} = 0 \tag{7.25}$$

This is called the *Taylor-Proudman theorem*. In the other extreme, in the limit of high Rossby number we may ignore the Coriolis force altogether.

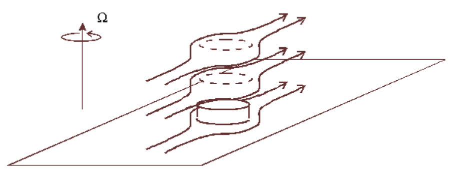
Taylor-Proudman ‘columns’ are thought to exist in rotating astrophysical bodies which are convective. Instead of moving up and down in the radial direction, convective cells move up and down parallel to the rotation axis. The best understood example of this is the Earth’s mantle.

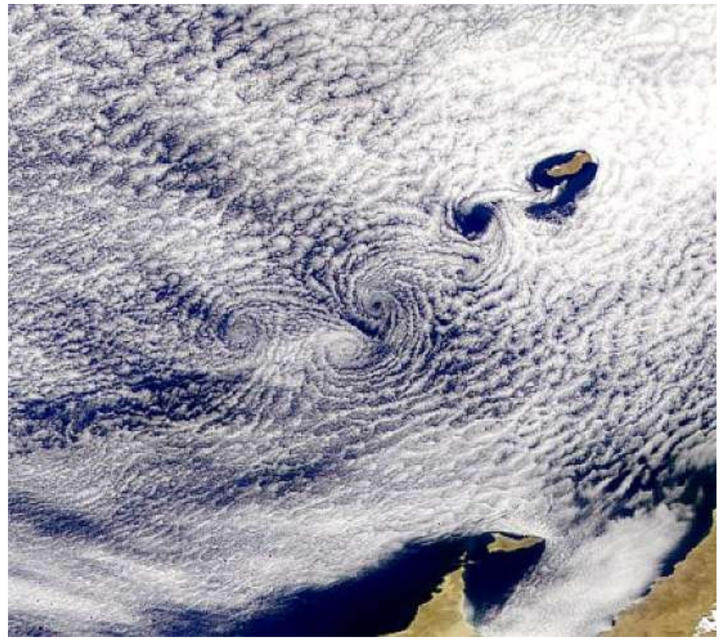
7.9 The geostrophic approximation

As we have done before, by comparing sizes of various terms in the momentum equation we can, according to the context, simplify it by dropping all but the largest terms. In the following we look at the example of the Earth’s atmosphere and oceans, but the same principles are applicable in many other astrophysical objects.

The Earth rotates considerably more slowly than the break-up spin ($\sim$ 2 hours), so that we can ignore the centrifugal force. Furthermore, the Earth’s atmosphere and oceans are very thin compared to its horizontal extent, and the gravitational force is much stronger than inertia or the Coriolis force. Thus, for most purposes we can assume hydrostatic equilibrium. The two horizontal components of the momentum equation now contain only the vertical component of the angular velocity of the Earth’s rotation. If

Figure 7.3: *Below:* The flow of fluid around and above an obstacle in a rotating frame with low Rossby number. *Right:* A von Karmán vortex street forms in the wake of an obstacle, but the clouds are at much higher altitude than the island – an illustration of the Taylor-Proudman theorem.





the colatitude (defined as 0° at the north pole, 90° at the equator and 180° at the south pole) is θ then we define $f \equiv 2\Omega \cos\theta$ and the equations of motion are

$$\frac{\mathrm{d}u}{\mathrm{d}t} = -\frac{1}{\rho}\frac{\partial P}{\partial x} + fv \tag{7.26}$$

$$\frac{\mathrm{d}v}{\mathrm{d}t} = -\frac{1}{\rho}\frac{\partial P}{\partial y} - fu \tag{7.27}$$

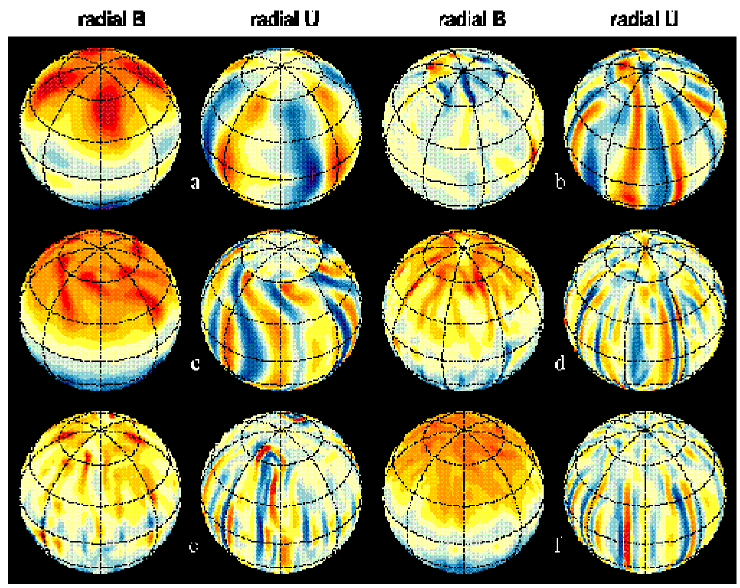
$$\partial P/\partial z = -\rho g \tag{7.28}$$

where u and v are the x and y components of velocity, where x is east and y is north. Here we have transformed to a local coordinate system where the curvature of the Earth's surface is ignored, along with the latitudinal variation in the Coriolis parameter f. Now, for sufficiently large-scale motions (or equivalently at low Rossby number) the acceleration terms on the left-hand side of the x and y parts above can be dropped; this is called the *geostrophic approximation*. In situations where this approximation applies, we can calculate the velocity field if we know the pressure field; we can see from these equations that the velocity is directed along the contours of pressure (isobars). This is the reason for the familiar patterns of anticlockwise winds around a low-pressure area (cyclone) and clockwise winds around a high-pressure area (anticyclone); in the southern hemisphere the directions are reversed along with the sign of f.

7.10 Rossby waves

We conclude this section on rotating fluids with a brief look at a wave which is ubiquitous in the Earth's atmosphere and ocean, but which is also possibly of great importance in fast-spinning neutron stars where it may be responsible for the emission of gravitational radiation.

Figure 7.4: Simulations of the Earth's dynamo. Note the prominent Taylor-Proudman convective columns.



First of all, let us assume that the Rossby number of a system is small: Ro $\ll$ 1. We shall look at the purest kind of Rossby wave, namely that where the motion takes place on a plane, and where the motion on that plane is incompressible. Using the shallow water equations in the linear regime we have

$$\frac{\partial u}{\partial t} - fv = -g\frac{\partial \zeta}{\partial x} \tag{7.29}$$

$$\frac{\partial v}{\partial t} + fu = -g\frac{\partial \zeta}{\partial y} \tag{7.30}$$

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = 0, \tag{7.31}$$

where the Coriolis parameter f now depends on latitudinal coordinate y. To be more precise, $f = f_0 + \beta y$, which is valid as long as we are not looking at too large a region in latitude. This is called the β-plane approximation, as opposed to the f-plane approximation where f is a constant. Cross differentiating and subtracting from one another the horizontal momentum equations, and assuming as usual a solution of the form $e^{i(kx+ly-i\omega)}$, we have

$$l\omega u - \beta v - flv - k\omega v - fku \tag{7.32}$$

$$ku + lv = 0 \tag{7.33}$$

which can easily be rearranged to give the dispersion relation

$$\omega = -\frac{k\beta}{k^2 + l^2}. \tag{7.34}$$

We see from this that the phase speed ω/k is has the same sign as β, meaning that the waveform moves always in the opposite direction to the rotation of the system. On Earth, the waves move to the west. However, note that the group velocity $\partial\omega/\partial\mathbf{k}$ can be in either direction.

Alternatively we can take the curl of the momentum equation:

$$\nabla \times \frac{\partial \mathbf{u}}{\partial t} + \nabla \times ((\mathbf{u} \cdot \nabla)\mathbf{u}) = -\nabla \times \left(f\hat{\mathbf{k}} \times \mathbf{u}\right) \tag{7.35}$$

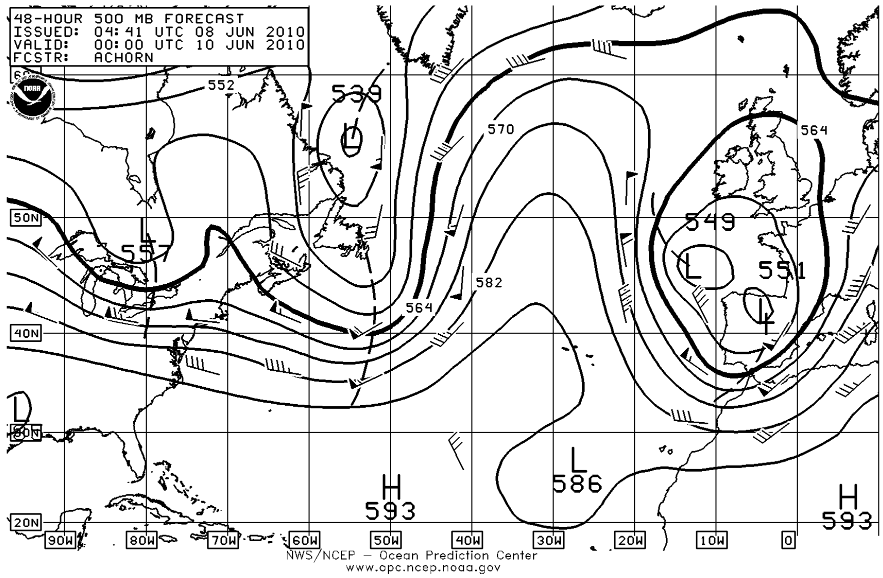


Figure 7.5: A weather forecast for 10th June 2010. Note that the wind is parallel to the isobars.

where $\hat{\mathbf{k}}$ is the vertical unit vector. In this two-dimensional velocity field $\mathbf{u}$ the vorticity is parallel to $\hat{\mathbf{k}}$ and we shall call it ψ, a scalar. After a little algebra we can rearrange to the shallow-water vorticity equation

$$\frac{\mathrm{d}}{\mathrm{d}t}(f+\psi) = -(f+\psi)\boldsymbol{\nabla}\cdot\mathbf{u} \tag{7.36}$$

which is analogous to (7.12); note the different symbol used here for vorticity to avoid confusion with oscillation frequency ω. To obtain the simplest modes we could again take $\boldsymbol{\nabla}\cdot\mathbf{u} = 0$, write out the Lagrangian derivative in its various parts and then assume a solution of the form $e^{i(kx+ly-i\omega)}$. However, the purpose of looking at the vorticity equation is to obtain an intuitive understanding of the wave. Since the absolute vorticity $f+\psi$ is conserved, if material which initially has $\psi = 0$ moves northwards or southwards (i.e. to a region of different f) it will start to rotate. Therefore any perturbation to the latitudinal velocity field will propagate to the west.

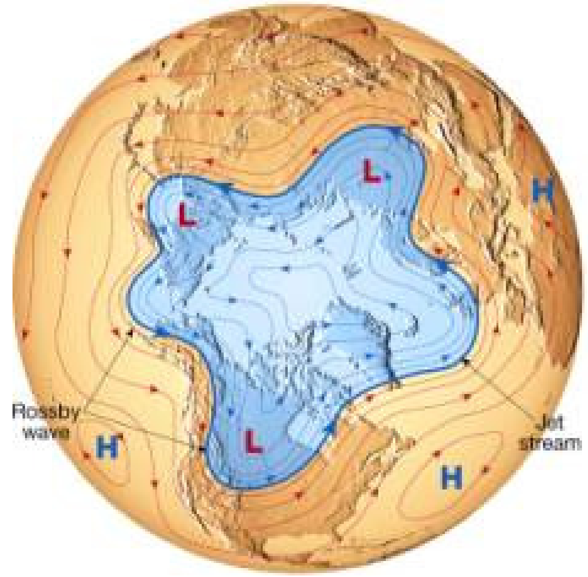
Exercises

7.1 Conservation of angular momentum

Show that the two terms $-\boldsymbol{\omega}(\boldsymbol{\nabla}\cdot\mathbf{u})$ and $(\boldsymbol{\omega}\cdot\boldsymbol{\nabla})\mathbf{u}$ in (7.5) can be written together as $-\boldsymbol{\omega}(\boldsymbol{\nabla}_\perp\cdot\mathbf{u})$ where $\boldsymbol{\nabla}_\perp$ is the divergence in a place perpendicular to $\boldsymbol{\omega}$. By considering a small rotating spheroidal fluid element of equatorial and polar (i.e. normal and parallel to $\boldsymbol{\omega}$) radius a and b in a velocity field, show that this represents the conservation of angular momentum as a and b change. Now take (7.5) and using the continuity equation turn it into an expression for the Lagrangian derivative of $\boldsymbol{\omega}/\rho$.

7.2 Behaviour of smoke rings

Figure 7.6: A Rossby wave propagating around the Earth.



Using the method of images, describe what happens when a vortex ring (e.g. smoke ring) approaches a wall. In addition, describe the interaction of two vortex rings, both of the same and opposite senses of rotation, where the line joining the centres of the two rings is perpendicular to the planes of the rings.

7.3 Tornadoes

A typical tornado can be approximated by a Rankine vortex with a core radius of 50 m, with a wind speed at this radius of 50 m s^{-1}. Write down an expression for the wind speed as a function of radius, and find an expression for the pressure as a function of radius. Assuming no external driving, find the speed and direction the tornado moves when it is a horizontal distance 200m below a high cliff.

7.4 Rotational flattening of stars and planets

Find an approximate expression for the extent of the rotational flattening of a star as a function of rotation period, and show that there is a lower limit to the rotation period at which the centrifugal force is comparable to gravity. Write this limit in terms of the mean density of the star.

7.5 Water in spinning tank

A tank of water is made to rotate at a constant angular velocity. Show that the surface of the water is a parabola shape.

7.6 Shallow-water vorticity equation

Show that it is possible to rewrite (7.36) as

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{f+\psi}{\xi}\right) = 0 \tag{7.37}$$

where ξ is now the total depth of the fluid including equilibrium and perturbation depths; above it was simply assumed that the Lagrangian derivative of ξ was zero. Comment on the consequences of the conservation of this quantity in terms of mechanisms for generating vorticity ψ in an ocean.

Chapter 8

Magnetohydrodynamics: equations and basic concepts

In this section the (non-relativistic) MHD equations are derived, starting with the non-magnetic fluid equations and then using Maxwell's equations to add the magnetohydrodynamic terms. Following that, some basic ideas in MHD are described, which are useful to building up an intuitive understanding of the subject. Students who are just using these notes to learn about MHD might wish to look first at section 1.2 to revise the derivation of the hydrodynamic equations.

8.1 The MHD equations

The MHD equations are essentially an extension of the HD equations with one extra variable: the magnetic field. There is one extra term in the momentum equation and a new partial differential equation called the induction equation.

We shall temporarily abandon the fluid picture, going back to individual particles with which the student may be more familiar from previous courses. The force on a single particle of charge q moving with velocity $\mathbf{v}$ in an electromagnetic field is given by:

$$\mathbf{F} = q\left(\mathbf{E} + \frac{\mathbf{v}}{c} \times \mathbf{B}\right), \tag{8.1}$$

where $\mathbf{E}$ and $\mathbf{B}$ are the electric and magnetic fields and c is the speed of light.[1] This force is normally called the Lorentz force. Now, in MHD we are interested in the force on the fluid as a whole rather than on individual particles – the total force per unit volume (also called the Lorentz force, confusingly) is therefore

$$\mathbf{F}_{\rm Lor} = \mathbf{F}_{\rm i} + \mathbf{F}_{\rm e} = (n_{\rm i}q_{\rm i} + n_{\rm e}q_{\rm e})\mathbf{E} + \left(n_{\rm i}q_{\rm i}\frac{\bar{\mathbf{v}}_{\rm i}}{c} + n_{\rm e}q_{\rm e}\frac{\bar{\mathbf{v}}_{\rm e}}{c}\right) \times \mathbf{B}, \tag{8.2}$$

where $\mathbf{F}_{\rm i}$ is the total force on the ions, $n_{\rm i}$, $q_{\rm i}$ and $\bar{\mathbf{v}}_{\rm i}$ are the number density, charge and mean velocity of ions and the quantities with subscript e refer to electrons. We have assumed here that all ions have the

[1] In some sense we can use this as the *definition* of $\mathbf{E}$ and $\mathbf{B}$, by saying that the force on a particle is some function of its charge and velocity which can be characterised by two vector fields in the form of (8.1).

same charge and that there are no neutral particles, but it is trivial to show that a generalisation does not affect the end result. We now define the fractional charge imbalance ratio $\epsilon \equiv (n_i q_i + n_e q_e)/n_e q_e$, as well as the drift velocity as the mean velocity of the electrons relative to the fluid, i.e. $\mathbf{v}_{\rm drift} \equiv \bar{\mathbf{v}}_e - \mathbf{u}$, where the fluid velocity $\mathbf{u} \approx \bar{\mathbf{v}}_i$ since the ions carry almost all of the momentum. Let us now rewrite (8.2) as

$$\mathbf{F}_{\rm Lor} = n_e q_e \left[\epsilon \mathbf{E} + \left(\epsilon \frac{\mathbf{u}}{c} + \frac{\mathbf{v}_{\rm drift}}{c} \right) \times \mathbf{B} \right]. \tag{8.3}$$

Noting that in the Earth we need $\epsilon < 10^{-36}$ so that the electric field does not overcome gravity and cause it to explode and that in almost all astrophysical contexts ϵ is negligible[2] we drop terms with ϵ (despite the fact that normally $\mathbf{v}_{\rm drift} \ll \mathbf{u}$; more on this in section 8.2). It is now convenient to introduce the concept of electric current density $\mathbf{J} = n_e q_e \mathbf{v}_{\rm drift}$, which with we simplify the Lorentz force (a force per unit volume) to

$$\mathbf{F}_{\rm Lor} = \frac{1}{c} \mathbf{J} \times \mathbf{B}. \tag{8.4}$$

We still need to answer the question of the origin of this relative velocity of the electrons to the ions; this comes essentially from the difference in force on the two species. The electrons experience a force relative to the fluid given by

$$\mathbf{F}_e - \mathbf{F}_{\rm Lor} = n_e q_e \left(\mathbf{E} + \frac{\mathbf{u}}{c} \times \mathbf{B} \right). \tag{8.5}$$

This force will accelerate the electrons relative to the ions, and there will quickly be a balance established between this acceleration and losses through collisions between electrons and ions. In fluids with normal conductivity properties, the drift velocity (and therefore current) established will be proportional to this acceleration. This gives us Ohm's law:

$$\mathbf{J} = \sigma \left(\mathbf{E} + \frac{\mathbf{u}}{c} \times \mathbf{B} \right), \tag{8.6}$$

where σ is the conductivity of the fluid, which will depend on mean free path, temperature, etc.

So far, we have added a term (8.4) to the momentum equation containing two new variables $\mathbf{B}$ and $\mathbf{J}$. Ohm's law (8.6) introduces yet another new variable $\mathbf{E}$, so that we need an additional two equations to close the set. Maxwell's equations are

$$\nabla \cdot \mathbf{E} = 4\pi \rho_e, \tag{8.7}$$
$$\nabla \cdot \mathbf{B} = 0, \tag{8.8}$$
$$\nabla \times \mathbf{E} = -\frac{1}{c} \frac{\partial \mathbf{B}}{\partial t}, \tag{8.9}$$
$$\nabla \times \mathbf{B} = \frac{1}{c} \frac{\partial \mathbf{E}}{\partial t} + \frac{4\pi}{c} \mathbf{J}, \tag{8.10}$$

where ρ_e is the net charge density.[3] First of all, note that if (8.8) is satisfied at some point in time, (8.9) ensures that it is satisfied at all other times, since the divergence of the curl of any vector field is zero. In standard magnetohydrodynamics we now make the approximation that the charge density ρ_e is small;

[2]Except where density is low and velocities are relativistic, e.g. pulsar magnetospheres.

[3]Note the similarity between (8.7) and the equation relating the gravitational field to density $\nabla \cdot \mathbf{g} = -4\pi G \rho$. No constant is required in the electromagnetic equivalent because it is built into the unit of charge. The other difference of course is that ρ_e can be either positive or negative.

also that the displacement current in (8.10) can be neglected, i.e. that $4\pi\mathbf{J} \gg \partial\mathbf{E}/\partial t$. See section 8.2 for a justification.

From Ohm's law (8.6) we obtain an expression for the electric field $\mathbf{E} = (1/\sigma)\mathbf{J} - (\mathbf{u}/c) \times \mathbf{B}$, which we can use in conjunction with (8.9) to obtain

$$\frac{\partial\mathbf{B}}{\partial t} = -c\boldsymbol{\nabla} \times \mathbf{E} = \boldsymbol{\nabla} \times \left(\mathbf{u} \times \mathbf{B} - \frac{c}{\sigma}\mathbf{J}\right), \tag{8.11}$$

and dropping the displacement current from (8.10) and substituting for $\mathbf{J}$ gives the induction equation

$$\frac{\partial\mathbf{B}}{\partial t} = \boldsymbol{\nabla} \times (\mathbf{u} \times \mathbf{B} - \eta\boldsymbol{\nabla} \times \mathbf{B}), \tag{8.12}$$

where magnetic diffusivity has been defined $\eta \equiv c^2/4\pi\sigma$, with units cm s^{-1}. Finally we can substitute for $\mathbf{J}$ into the Lorentz force (8.4) to give force per unit volume

$$\mathbf{F}_{\mathrm{Lor}} = \frac{1}{4\pi}(\boldsymbol{\nabla} \times \mathbf{B}) \times \mathbf{B}. \tag{8.13}$$

Thus $\mathbf{E}$ and $\mathbf{J}$ have been eliminated. In summary, compared to the original hydrodynamics equations we have one additional variable $\mathbf{B}$, one additional equation (8.12) and one additional term (8.13) in the momentum equation. Note that nothing changes if we reverse the direction of the magnetic field.

8.2 The MHD approximation

In addition to the standard conditions under which the fluid approximation is valid, e.g. collision frequency, etc. we have made further approximations. We assumed that the magnetic permeability and dielectric permittivity of the plasma can be ignored (refractive index equal to unity). We also assumed the flow is non-relativistic. Finally we made what is known as the 'MHD approximation', according to which the conductivity of the material is high enough so that the charge density ρ_{e} is low and (8.7) can be ignored.

In the rest frame of a test particle (in which quantities are denoted with a prime), the force experienced is $\mathbf{F}' = q\mathbf{E}'$. In transforming to an inertial frame, we ignore terms in v^2/c^2 (so that the Lorentz factor $\Gamma \approx 1$) to obtain the lab frame relation (8.1). Returning to the fluid picture, a high conductivity σ ensures that current can flow in order to almost neutralise the rest frame electric field $\mathbf{E}' = \mathbf{E} + (\mathbf{u}/c) \times \mathbf{B}$ so that $E' \ll E$ and $E \sim (u/c)B$. We can now justify neglecting the displacement current in (8.10), because it is of order $E(u/c)/L \sim B(u/c)^2/L$ where L is a typical length scale of the flow. This is smaller by a factor v^2/c^2 than the curl of the magnetic field; meaning that $J \sim cB/L$ (dropping factors of 4π).

Looking at the first of Maxwell's equations (8.7), also known as Gauss' law, we see that $\rho_{\mathrm{e}} \sim E/L \sim (u/c)B/L \sim (u/c)J/c$. Since this law holds in every reference frame it follows from $E' \ll E$ that $\rho'_{\mathrm{e}} \ll \rho_{\mathrm{e}}$. Current density $\mathbf{J}$ can be considered equal in lab and co-moving frames since assuming that $\Gamma \approx 1$ gives $\mathbf{J}' = \mathbf{J} - \rho_{\mathrm{e}}\mathbf{u}$; the ratio of the two terms is $(u/c)^2$ so that $\mathbf{J}' = \mathbf{J}$. Magnetic field can also be considered frame-independent, since a transformation assuming $\Gamma \approx 1$ gives $\mathbf{B}' = \mathbf{B} - (\mathbf{u}/c) \times \mathbf{E}$, and the electric field $\mathbf{E}$ is itself smaller than the magnetic field by a factor u/c so we are left with a ratio u^2/c^2 between the two terms so that $\mathbf{B} = \mathbf{B}'$.

There are obviously astrophysical contexts in which these approximations do not hold, for instance relativistic flows such as GRB jets, or situations where plasma effects become important or where the fluid approximation breaks down; these are outside the scope of this course.

8.3 The magnetic and other fields

The MHD equations in the form of (8.12) and (8.13) contain magnetic field but not current density or electric field. In fact there are only a few contexts in astrophysics where it is necessary to think about these extra fields. In the comoving frame, i.e. the frame 'felt' by the fluid, we have seen that the electric field vanishes in the case of infinite conductivity and other, non-relativistic contexts is very small compared to the magnetic field. This, combined with the fact that the fluid is normally almost perfectly neutral, renders its effect totally insignificant, except in special contexts where densities are very low and velocities are relativistic, such as the magnetospheres of neutron stars. Likewise, we need only think about current density in special situations, such as that where a limited density of charge carriers forces the drift velocity to become relativistic and the effective conductivity becomes small.

It is therefore surprising that the literature on MHD often refers to these additional fields, including the situations where the MHD approximation can be considered to hold perfectly. The reason for this can historically be traced back to a transfer of understanding of terrestrial phenomena, particularly that of electric circuits, coils, inductances and so on to the astrophysical context. However, the two contexts are rather different. For example, consider using a battery to pass a current through a wire and measuring the magnetic field it produces in the insulating fluid (air) surrounding it. None of these exists in astrophysics, and it makes no more sense in MHD to say that a magnetic field is produced by a current than to say that a current is produced by a magnetic field – the two are related by $(4\pi/c)\mathbf{J} = \boldsymbol{\nabla}\times\mathbf{B}$ and that is the end of the story. In fact, the magnetic field can in some sense be considered the more fundamental (or rather, useful) of the two, as its evolution is governed by conservation laws which do not apply to the current. The student is advised to avoid thinking in terms of currents, electromotive forces and circuits as these will distract from an understanding of the subject; only **B** is required!

8.4 A brief note concerning units

At this juncture it is worth commenting on the difference between the c.g.s. units employed here and the S.I. units often taught in undergraduate courses. The reader will notice that the equations above contain only one constant of nature: the speed of light c. In contrast, a glance at some of the text books reveals that the S.I. system is burdened not only with c but also with the rather nineteenth century concepts of the permittivity and permeability of the ether, ϵ_0 and μ_0. Another advantage of c.g.s. is that the electric and magnetic fields have the same units. However, a word of caution: there exist variations of c.g.s. units; here we use 'Gaussian c.g.s.' units, which are fairly standard in astrophysics. In this system the unit of charge, called the statcoulomb, is defined from Coulomb's law of the force between two point charges $F = q_1q_2/r^2$ such that two unit charges at a separation of 1 cm experience a repulsion of 1 dyne. It can be written in terms of the other units: 1 statC = 1 $\mathrm{g}^{1/2}\ \mathrm{cm}^{3/2}\ \mathrm{s}^{-1}$. In S.I. units the unit of charge is pre-defined and Coulomb's law requires a constant $1/4\pi\epsilon_0$. Occasionally in astrophysics one comes across Heaviside-Lorentz units which differ from Gaussian only in factors of 4π.

The unit of magnetic field in the c.g.s. system is the gauss (abbreviation G), which is equal to 10^{-4} tesla, the unit in the SI system. Magnetic fields strengths in the universe range from 10^{-6} G in the intergalactic medium, $0.3 - 0.6$ G on the surface of the Earth, 10^4 G on some magnetic main-sequence stars, $9 \cdot 10^5$ G in a laboratory in Dresden recently, up to 10^9 G in white dwarfs and up to 10^{15} G in neutron stars.

8.5 Magnetic diffusivity

If we assume that the electrical conductivity σ is uniform, we can rearrange the induction equation (8.12) using the constraint $\boldsymbol{\nabla} \cdot \mathbf{B} = 0$ and the vector identity $\boldsymbol{\nabla}\times\boldsymbol{\nabla}\times\mathbf{A} = \mathbf{A}(\boldsymbol{\nabla}\cdot\mathbf{A}) - \boldsymbol{\nabla}^2\mathbf{A}$, to

$$\frac{\partial \mathbf{B}}{\partial t} = \boldsymbol{\nabla} \times (\mathbf{u} \times \mathbf{B}) + \eta \boldsymbol{\nabla}^2 \mathbf{B}, \tag{8.14}$$

where the magnetic diffusivity η, like the kinetic and thermal diffusivities ν and κ, has units cm^2 s^{-1}. We give names to the ratios of diffusivities: the Prandtl number and magnetic Prandtl number are defined as $\mathrm{Pr} \equiv \nu/\kappa$ and $\mathrm{Pr_m} \equiv \nu/\eta$ respectively. Note the similarity in the three diffusive terms in the equations – all contain a diffusive coefficient multiplied by $\boldsymbol{\nabla}^2$ of the relevant variable. As with the other diffusivities, from analysis of units we see that there is a characteristic magnetic diffusive timescale, equal to $\mathcal{L}^2/\eta$ where $\mathcal{L}$ is the characteristic length scale of the system. This is called the *Ohmic timescale*. In many contexts the Ohmic timescale is very much longer than other timescales of interest and it is possible to ignore the diffusive term in the inductive equation. This regime is called *ideal MHD*. In many, if not most, applications we can use ideal MHD.

Finally, it is worth mentioning the heating from Ohmic dissipation, known as *Joule heating*: per unit volume this is equal to $\mathbf{J} \cdot \mathbf{E}'$ where $\mathbf{E}'$ is the electric field in the comoving frame (recalling that current density is the same in both frames). Expressed differently, this means that

$$Q_{\mathrm{Joule}} = \frac{1}{\sigma} J^2 = \frac{\eta}{4\pi} (\boldsymbol{\nabla} \times \mathbf{B})^2. \tag{8.15}$$

8.6 Different regimes in MHD

First let us summarise the set of equations we have so far, including the approximations of constant viscosities, assuming an ideal gas equation of state $P = \rho RT/\mu$ and ignoring gravity:

$$\begin{aligned}
\rho \frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t} &= -\boldsymbol{\nabla} P + \frac{1}{4\pi} (\boldsymbol{\nabla} \times \mathbf{B}) \times \mathbf{B} + \rho \nu \boldsymbol{\nabla}^2 \mathbf{u}, && (8.16)\\
\frac{\partial \rho}{\partial t} &= -\boldsymbol{\nabla}\cdot(\rho \mathbf{u}), && (8.17)\\
\frac{\mathrm{d}P}{\mathrm{d}t} &= (\gamma - 1) Q - \gamma P\, \boldsymbol{\nabla} \cdot \mathbf{u}, && (8.18)\\
\frac{\partial \mathbf{B}}{\partial t} &= \boldsymbol{\nabla} \times (\mathbf{u} \times \mathbf{B} - \eta \boldsymbol{\nabla} \times \mathbf{B}). && (8.19)
\end{aligned}$$

The heating rate is given by

$$Q = e_{ij} S_{ij} + \boldsymbol{\nabla} \cdot (K \boldsymbol{\nabla} T) + \frac{\eta}{4\pi} (\boldsymbol{\nabla} \times \mathbf{B})^2 + Q_{\mathrm{other}}, \tag{8.20}$$

where e_{ij} and S_{ij} are the rate of strain and viscous stress tensors respectively and K is the thermal conductivity.

Assigning typical flow parameters $\mathcal{L}$, $\mathcal{U}$ and $\mathcal{T} = \mathcal{L}/\mathcal{U}$ – typical length scale, velocity and timescale – and comparing the size of various terms in the equations, we can rewrite equations (8.16) and (8.19), dropping gravity as well as factors of order unity such as γ, in terms of non-dimensional variables and gradients, such as $\mathbf{u}' = \mathcal{U}\mathbf{u}$ (where we drop the prime below), as:

$$\begin{aligned}
\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t} &= -\frac{c_\mathrm{s}^2}{\mathcal{U}^2}\nabla P + \frac{v_\mathrm{A}^2}{\mathcal{U}^2}(\nabla\times\mathbf{B})\times\mathbf{B} + \frac{1}{\mathrm{Re}}\nabla^2\mathbf{u}, \\
&= \frac{1}{M^2}\left[-\nabla P + \frac{1}{\beta}(\nabla\times\mathbf{B})\times\mathbf{B}\right] + \frac{1}{\mathrm{Re}}\nabla^2\mathbf{u}, \qquad (8.21)\\
\frac{\partial\mathbf{B}}{\partial t} &= \nabla\times(\mathbf{u}\times\mathbf{B}) + \frac{1}{\mathrm{Re_m}}\nabla^2\mathbf{B}, \qquad (8.22)
\end{aligned}$$

where $c_\mathrm{s}^2 = \mathrm{d}P/\mathrm{d}\rho = \gamma P/\rho$ and $v_\mathrm{A}^2 = B^2/4\pi\rho$ are the sound and Alfvén speeds and $\mathrm{Re} \equiv \mathcal{U}\mathcal{L}/\nu$ and $\mathrm{Re_m} \equiv \mathcal{U}\mathcal{L}/\eta$ are the Reynolds number and magnetic Reynolds number respectively, measures of the ratio of inertia to diffusivity of two kinds. Note that it has been assumed that the Lagrangian derivative on the left-hand side has size of order $\mathcal{U}/\mathcal{T} \approx \mathcal{U}^2/\mathcal{L}$, which is generally true in flows; see section 5.4. We can also define a plasma $\beta \equiv 8\pi P/B^2 \approx c_\mathrm{s}^2/v_\mathrm{A}^2$, a (likely) ratio of the first and second terms on the right hand side of the momentum equation. The Mach number is the ratio of flow speed to sound speed $M \equiv \mathcal{U}/c_\mathrm{s}$. Sometimes one also talks in terms of the Alfvénic Mach number $M_\mathrm{A} \equiv \mathcal{U}/v_\mathrm{A}$. Also, note that $\mathrm{Re_m}/\mathrm{Re} = \mathrm{Pr_m}$.

In an unmagnetised fluid, we can describe the flow with M and Re (and possibly also the Strouhal number; see section 5.4). We have already seen that if $M \ll 1$ the flow is roughly incompressible, i.e. that $\mathrm{d}\rho/\mathrm{d}t \approx 0$, which gives the simplified continuity equation $\nabla\cdot\mathbf{u} = 0$. This is the regime we assumed in simplifying the viscous force in the momentum equation. The Reynolds number characterises the importance of viscosity.

In a magnetised medium we now have two extra parameters $\mathrm{Re_m}$ and β; the former describes the importance of diffusivite (finite conductivity) and the latter describes the relative importance of gas and magnetic pressures. The value of the 'plasma β' is very important in MHD. If $\beta \ll 1$, one expects from looking at the momentum equation that the Lorentz force will be much larger than the pressure gradient force. This means that velocities of order the Alfvén speed (which can be very high) will result unless the current $\nabla\times\mathbf{B}$ is almost parallel to $\mathbf{B}$, the so-called 'force-free' regime. Conversely, if $\beta \gg 1$ then the magnetic field will only have much effect on the flow in directions where the Lorentz force is not opposed by the pressure gradient or other stronger forces such as gravity.

8.7 Field lines, flux conservation and flux freezing

In the remainder of this chapter we derive some results for the case of a fluid with infinite conductivity ($\eta = 0$) which can be considered approximately true in the more realistic case of finite conductivity. Taking $\eta = 0$ gives the ideal MHD induction equation

$$\frac{\partial\mathbf{B}}{\partial t} = \nabla\times(\mathbf{u}\times\mathbf{B}). \qquad (8.23)$$

Let us define a magnetic flux ϕ as an integral of the normal component of $\mathbf{B}$ on a surface S

$$\phi = \int_S \mathbf{B}\cdot\mathrm{d}\mathbf{S} \qquad (8.24)$$

and then calculate the change of flux through that surface as it moves with the flow of the fluid:

$$\frac{\mathrm{d}\phi}{\mathrm{d}t} = \int_S \frac{\partial \mathbf{B}}{\partial t} \cdot \mathrm{d}\mathbf{S} + \oint_l (\mathbf{u} \times \mathrm{d}\mathbf{l}) \cdot \mathbf{B}, \tag{8.25}$$

where the first term comes from the rate of change of flux through the surface if it were fixed in space and the second comes from the movement of the surface from the fluid velocity $\mathbf{u}$. The surface is bounded by a line l. Substituting (8.23) into this equation and using Stokes' theorem, the first term becomes $\oint_l \mathbf{u} \times \mathbf{B} \cdot \mathrm{d}\mathbf{l}$. From the triple vector product rule we now see that the two terms cancel and that the flux through the co-moving surface is constant in time.

If we now imagine the fluid being composed of small co-moving fluid elements, each threaded by a constant flux, it becomes clear that the concept of field lines and of their being 'frozen' into the fluid are useful tools in understanding MHD. This will be discussed below at greater length.

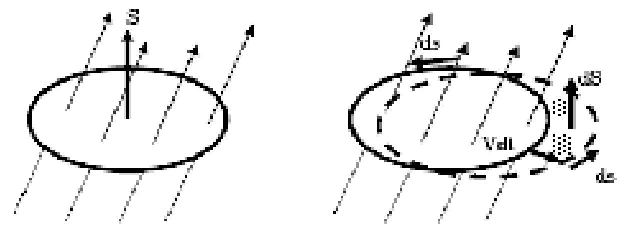


Figure 8.1: A surface S moving with the flow, threaded by a magnetic flux ϕ.

Note that this result is analogous to that of freezing of vorticity in hydrodynamics. The vorticity equation in an inviscid barotropic flow with conservative body forces can be written

$$\frac{\partial \boldsymbol{\omega}}{\partial t} = \boldsymbol{\nabla} \times (\mathbf{u} \times \boldsymbol{\omega}), \tag{8.26}$$

which has the same form as (8.23).

8.8 Magnetic pressure, tension and energy density

The Lorentz force can be written in an alternative form, making use of a vector identity and the solenoidal constraint $\boldsymbol{\nabla} \cdot \mathbf{B} = 0$

$$\mathbf{F}_{\mathrm{Lor}} = \frac{1}{4\pi}(\boldsymbol{\nabla} \times \mathbf{B}) \times \mathbf{B} = -\boldsymbol{\nabla}\left(\frac{B^2}{8\pi}\right) + \frac{1}{4\pi}(\mathbf{B} \cdot \boldsymbol{\nabla})\mathbf{B}. \tag{8.27}$$

The first term looks like the pressure gradient term $-\boldsymbol{\nabla} P$, so the quantity $B^2/8\pi$ is called the *magnetic pressure*. The second term on the right is often called the magnetic tension or curvature force. However, we need to remind ourselves that the total Lorentz force is always perpendicular to the magnetic field, so that the components of these two terms parallel to the field must cancel. After some manipulation we can rewrite the Lorentz force without these cancelling components as

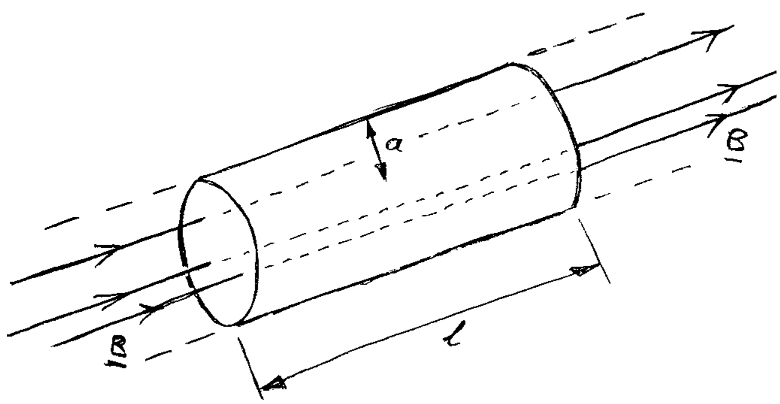
$$\mathbf{F}_{\mathrm{Lor}} = -\boldsymbol{\nabla}_\perp \frac{B^2}{8\pi} + \boldsymbol{\kappa}\frac{B^2}{4\pi}, \tag{8.28}$$

where $\boldsymbol{\kappa} = [(\mathbf{B}/B) \cdot \boldsymbol{\nabla}](\mathbf{B}/B)$ is the curvature vector (directed along the radius of curvature of the field and equal in magnitude to the reciprocal of that radius) and $\boldsymbol{\nabla}_\perp$ is the part of the gradient perpendicular to the field. The curvature force resembles the tension in a string in that it will tend to restore a perturbed straight field line to its original shape.

It can be shown that the magnetic pressure $B^2/8\pi$ is also the energy density of the magnetic field; to demonstrate this in a non-rigorous way is reasonably straightforward. Imagine a straight tube of infinite length with cylindrical cross section of radius a. It contains a uniform field B parallel to its length and is surrounded by unmagnetised fluid. The curvature force vanishes everywhere and the only place where the other part of the Lorentz force does not vanish is the boundary, where it is a delta function directed normal to the boundary. Assuming equilibrium, this must be balanced by a delta function in the gas pressure gradient force, i.e. a discontinuity in pressure, with the gas pressure outside the tube is greater than that inside by a quantity $B^2/8\pi$. Now imagine making an adiabatic change in a so that the flux $\phi = \pi a^2 B = \text{const.}$ If the magnetic field has energy e per unit volume, then the magnetic energy of the tube is $E = \pi a^2 e$ per unit length. From the $\mathrm{d}U = -p\,\mathrm{d}V$ relation in thermodynamics (doing work on the magnetic field by pushing against the Lorentz force) we therefore have $\mathrm{d}E = -(B^2/8\pi)2\pi a\mathrm{d}a = -\phi^2/(4\pi^2 a^3)\mathrm{d}a$ and so assuming that the energy E goes to zero at $a = \infty$ we can integrate from $a' = \infty$ to a to give

$$\int_0^E \mathrm{d}E' = -\int_\infty^a \frac{\phi^2}{4\pi^2 a'^3}\mathrm{d}a' \qquad \Longrightarrow \qquad E = \frac{\phi^2}{8\pi^2 a^2} = \frac{B^2}{8\pi}\pi a^2 \qquad \Longrightarrow \qquad e = \frac{B^2}{8\pi}. \quad (8.29)$$

Figure 8.2: A section of a flux tube.



The magnetic pressure is different from the gas pressure in that it is not isotropic. To demonstrate this, imagine a section of the aforementioned flux tube of length l, which contains magnetic energy $E = \pi a^2 l B^2/8\pi$. Stretching this section of the tube whilst keeping the cross section a fixed will not change B so the increase in energy is simply $\mathrm{d}E = \pi a^2 \mathrm{d}l B^2/8\pi$. Equating this to the energy $\mathrm{d}E = -P\,\mathrm{d}V$ again, we have $-P_{\mathrm{mag}}\pi a^2 dl = (B^2/8\pi)\pi a^2 \mathrm{d}l$ and so the tube has a tension per unit area of $B^2/8\pi$, which can be thought of as a negative pressure. Of course, no infinite tube exists in reality – however, we can imagine a tube being connected to itself in a loop, which helps us to understand how the tension comes about even though the Lorentz force is always perpendicular to the field.

If we repeat the above thought experiment with not a stretch or compression in one direction, but an *isotropic* expansion or compression of a magnetised fluid element, it is easily shown that the resistance of the magnetic field to such a compression is equivalent to an isotropic pressure of $B^2/24\pi$. This is simply the average of a pressure of $B^2/8\pi$ in the two directions perpendicular to the field and tension of $B^2/8\pi$ parallel to it. Furthermore, it can be shown that any self-contained magnetic feature (i.e. one without field lines crossing its boundary) exerts a mean pressure $B^2/24\pi$ on its surroundings, meaning that in equilibrium, the gas pressure outside the feature must be greater than the average pressure inside by this quantity. This should not be surprising since magnetic field is a relativistic phenomenon and relativistic fluids (e.g. photons, relativistic particles, gravity) exert a pressure equal to one third of their energy density – equivalent to an adiabatic index of $4/3$. However, one must always be careful when simplifying the effect of a magnetic field to an isotropic magnetic pressure since the geometry of the field is often crucial.

8.9 Waves

The restoring force from bending field lines and from squeezing field lines together allows the propagation of waves in a magnetised fluid. The simplest form of waves propagates in a medium of initially constant pressure and density threaded by a uniform magnetic field, and since magnetic pressure is not isotropic it is necessary in the general case to consider the angle between the wavevector and the magnetic field; however the equivalence of the two dimensions perpendicular to the field allows us to drop one and consider only the remaining two dimensions. It turns out that there are three kinds of wave in a compressible magnetised fluid: the Alfvén wave and the fast and slow magnetoacoustic (or magnetosonic) waves. First of all we write the linearised momentum, continuity, induction and energy equations:

$$\frac{\partial \mathbf{u}}{\partial t} = -\frac{1}{\rho}\boldsymbol{\nabla}\delta P + \frac{1}{4\pi\rho}\left(\boldsymbol{\nabla}\times\delta\mathbf{B}\right)\times\mathbf{B} \tag{8.30}$$

$$\frac{\partial \delta\rho}{\partial t} = -\rho\boldsymbol{\nabla}\cdot\mathbf{u} \tag{8.31}$$

$$\frac{\partial \delta\mathbf{B}}{\partial t} = \boldsymbol{\nabla}\times(\mathbf{u}\times\mathbf{B}) \tag{8.32}$$

$$\delta P = c_{\mathrm{s}}^2\delta\rho \tag{8.33}$$

where $\mathbf{B}$ is the equilibrium field which is parallel to the y-axis and quantities with δ are the perturbations which, as well as the x and y components of the velocity field u and v, are small. Writing $\tilde{\rho} = \delta\rho/\rho$ and $\mathbf{b} = \delta\mathbf{B}/B$ and substituting from (8.33) we can write out the equations as

$$\frac{\partial u}{\partial t} = -c_{\mathrm{s}}^2\frac{\partial\tilde{\rho}}{\partial x} + v_{\mathrm{A}}^2\left(\frac{\partial b_x}{\partial y} - \frac{\partial b_y}{\partial x}\right) \tag{8.34}$$

$$\frac{\partial v}{\partial t} = -c_{\mathrm{s}}^2\frac{\partial\tilde{\rho}}{\partial y} \tag{8.35}$$

$$\frac{\partial \tilde{\rho}}{\partial t} = -\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} \tag{8.36}$$

$$\frac{\partial b_x}{\partial t} = \frac{\partial u}{\partial y} \tag{8.37}$$

$$\frac{\partial b_y}{\partial t} = -\frac{\partial u}{\partial x} \tag{8.38}$$

where the Alfvén speed has been defined as $v_{\mathrm{A}} \equiv B/\sqrt{4\pi\rho}$. To work through the derivation of the dispersion relation is a little lengthy so we restrict ourselves here to two special cases, where the wavevector is parallel and perpendicular to the field. In the parallel case $\partial/\partial x = 0$ and the equations split into two sets, one set made of (8.35) and (8.36) containing v and $\tilde{\rho}$ and the other set made of (8.34) and (8.37), containing just u and b_x. The first set describes a longitudinal wave, since the motion is parallel to the wavevector. Motion parallel to a magnetic field has no effect on it (since $\mathbf{u} \times \mathbf{B}$ vanishes), and this wave is simply a sound wave. The other set describes transverse wave called an *Alfvén wave* which is non-compressional (since $\tilde{\rho}$ is absent). It is left as an exercise for the student to derive the dispersion relation for these waves, and to show that the propagation speed is v_{A}.

The other special case is that of perpendicular wavevector and field, where $\partial/\partial y = 0$. Here, v and b_x drop out of the equations and there is only one wave, a longitudinal wave called the *fast magnetoacoustic wave*. It is left as an exercise to derive the dispersion relation and to show that the propagation speed is $(c_{\mathrm{s}}^2 + v_{\mathrm{A}}^2)^{1/2}$. These waves are similar to sound waves, but the magnetic field provides an extra restoring force.

If the wavevector and field are neither parallel nor perpendicular, two waves are possible (the fast and slow magnetoacoustic waves), and the phase and group velocities are no longer parallel. These waves, which are not easy to visualise, are compressional and involve a mixture of gas and magnetic restoring forces.

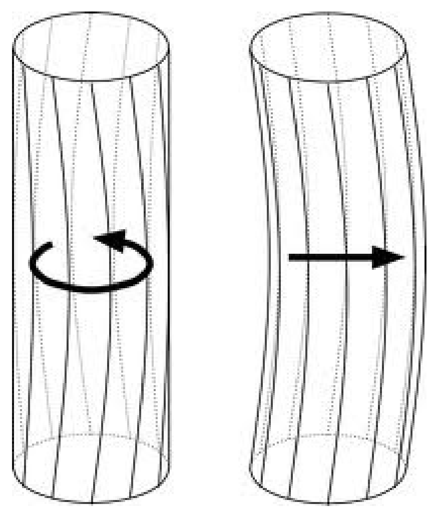


Figure 8.3: Torsional and plane Alfvén waves propagating along a flux tube.

Finally, a particular kind of Alfvén wave deserves a mention: the torsional Alfvén wave. Imagine a flux tube which is perturbed not by a sideways motion but by a twisting motion – the twist propagates along the tube.[4] This kind of wave is very important in various astrophysical situations, for instance in differentially rotating stars and in star formation, since it carries angular momentum.

8.10 Magnetic helicity

In section 8.7 we saw that the magnetic flux through a co-moving fluid surface is constant in the limit of high conductivity. We can now look at an additional quantity which is also conserved in this limit.

[4]According to MHD folklore Alfvén got the idea from growing sunflowers in the Swedish arctic, where the flowers rotate once every day during the summer to follow the Sun; at the end of the season the sun finally sets and the sunflowers 'unwind'.

Magnetic helicity is a global quantity defined as

$$H \equiv \int_V \mathbf{A} \cdot \mathbf{B} \, \mathrm{d}V \tag{8.39}$$

where $\mathbf{A}$ is the vector potential defined from $\mathbf{B} = \nabla \times \mathbf{A}$. Now, since the curl of the divergence of a scalar is zero, we can add any gradient of a scalar $\nabla\phi$ to the vector potential without changing the magnetic field; however this will in general affect the magnetic helicity. It can be shown in the following way though that magnetic helicity is gauge invariant provided that no magnetic field lines pass through the boundary of the volume of integration. Consider some new vector potential $\mathbf{A}' = \mathbf{A} + \nabla\phi$. The helicity is now

$$\begin{aligned} H' &= \int_V [\mathbf{A} \cdot \mathbf{B} + (\nabla\phi) \cdot \mathbf{B}] \, \mathrm{d}V &(8.40) \\ &= H + \int_V [\nabla \cdot (\phi\mathbf{B}) - \phi(\nabla \cdot \mathbf{B})] \, \mathrm{d}V &(8.41) \\ &= H + \oint_S \phi \, \mathbf{B} \cdot \mathrm{d}\mathbf{S}, &(8.42) \end{aligned}$$

since the first term on the first line is simply equal to the original helicity; the second term was expanded with a standard vector identity to give the expression on the second line. Since the magnetic field is solenoidal, the second of the new terms vanishes, and the first can be rewritten with the aid of Gauss' theorem to give a surface integral. Therefore if $\mathbf{B} \cdot \mathrm{d}\mathbf{S} = 0$ everywhere on the boundary, helicity is gauge invariant.

Helicity is a useful concept because of its conservation properties. It can be shown that it is perfectly conserved in the limit of infinite conductivity. First note that by integrating the ideal MHD induction equation (8.23) we have $\partial\mathbf{A}/\partial t = \mathbf{u} \times \mathbf{B}$. Now

$$\begin{aligned} \frac{\partial H}{\partial t} &= \int_V \mathrm{d}V \, [\mathbf{A}_t \cdot \mathbf{B} + \mathbf{A} \cdot \mathbf{B}_t] = \int_V \mathrm{d}V \, [\mathbf{u} \times \mathbf{B} \cdot \mathbf{B} + \mathbf{A} \cdot \nabla \times (\mathbf{u} \times \mathbf{B})] \\ &= \int_V \mathrm{d}V \, [(\mathbf{u} \times \mathbf{B}) \cdot \nabla \times \mathbf{A} - \nabla \cdot (\mathbf{A} \times (\mathbf{u} \times \mathbf{B}))] \\ &= -\oint_S \mathrm{d}\mathbf{S} \cdot \mathbf{A} \times (\mathbf{u} \times \mathbf{B}). \end{aligned} \tag{8.43}$$

It is sufficient then that the velocity goes to zero on the boundary of the domain.

In a fluid with finite conductivity helicity is still approximately conserved – we can see this from an argument with units. Now, the diffusive timescale on which the magnetic field decays due to finite conductivity, as we saw above, is $\tau \sim \mathcal{L}^2/\eta$. This is shorter on shorter length scales so that when magnetic energy is converted to heat via Ohmic dissipation, it is mainly the small-scale structure where the energy is converted. Helicity however has units of length times energy and is therefore present more in the large scale components of the magnetic field than the magnetic energy, and less is therefore lost due to diffusive processes on small scales. Often then during MHD processes where the flow contains a range of length scales, energy is lost at the smallest scales while helicity is roughly conserved.

Helicity also has units of flux squared, and can in fact be thought of in some sense of the product of two fluxes of different components of the magnetic field. It is often said that helicity is a measure of the 'twist' of the magnetic field, because a twisted field must contain at least two components. As we have seen, that twist is conserved even as magnetic energy is lost.

8.11 MHD equilibria

In many astrophysical contexts, we are interested in equilibrium situations where the forces are balanced. However, before we proceed, it is important to clarify what we mean by equilibrium. If we simply set the velocity to zero, the momentum equation (8.16) gives us a relation between P, $\mathbf{B}$ and $\rho\mathbf{g}$, so that any combination of the three which satisfy that relation will be an equilibrium. Now, both sides of the continuity equation (8.17) go to zero, but the heat equation (8.18) contains Q, which does generally not go to zero when $\mathbf{u} = \mathbf{0}$ due to its terms with κ and η. Likewise, the term with η in the induction equation (8.19) will not in general vanish. The result of this is that the magnetic field and pressure field will evolve, giving rise to a non-zero velocity field – a truly stationary state is in general achievable only where $\kappa = \eta = 0$. However, provided that these diffusivities are small, we can still find a *dynamic equilibrium* by setting $\mathbf{u} = \mathbf{0}$ and balancing forces. This equilibrium will not change appreciably on a dynamic timescale, i.e. the time taken for a sound or Alfvén wave to travel across the domain; rather, it will evolve over a longer timescale due to the diffusive terms.

So, finding a (dynamic) equilibrium is simply a matter of finding a solution to the following equation:

$$-\nabla P + \frac{1}{4\pi}(\nabla \times \mathbf{B}) \times \mathbf{B} + \rho\mathbf{g} = \mathbf{0}, \tag{8.44}$$

together with the constraint $\nabla \cdot \mathbf{B} = 0$. It is interesting to explore some of the properties of this equation. First of all, note that in the non-magnetised case it reduces to the equation of hydrostatic equilibrium which one often sees in the form $\partial P/\partial z = -\rho g$ where gravity is directed downwards along the z-axis. In the magnetised case, taking the dot product with $\hat{\mathbf{B}}$, the unit vector in the direction of the magnetic field, gives

$$(\hat{\mathbf{B}} \cdot \nabla)P = \rho(\mathbf{g} \cdot \hat{\mathbf{B}}) \qquad \text{or} \qquad \frac{dP}{ds} = \rho g_s, \tag{8.45}$$

where dP/ds is the derivative along a field line and g_s is the component of gravity along the field line. In other words, in an MHD equilibrium there is hydrostatic balance *along field lines*.

In a situation where $\beta \gg 1$ and the Lorentz term in (8.44) is much smaller than the pressure and gravity terms, we can imagine first constructing a non-magnetic equilibrium where $\nabla P = \rho\mathbf{g}$, adding a weak magnetic field and then making small adjustments to the pressure and density fields to balance the Lorentz force. In principle this should be possible, because an arbitrary magnetic field and its associated Lorentz force have two degrees of freedom – three dimensions minus the one constraint ($\nabla \cdot \mathbf{B} = 0$) – and we also have two degrees of freedom in balancing the Lorentz force since we can adjust both the pressure and density fields independently of each other.[5] Note that in a fluid with a barotropic equation of state $\rho = \rho(P)$ we only have one degree of freedom in adjusting the pressure and density fields, so that depending on the context it may be either more difficult or impossible to construct an equilibrium.

In the special case without gravity we see from (8.45) that pressure is constant along field lines. Also, the gradients in thermal pressure P and magnetic pressure $B^2/8\pi$ must be comparable unless we can construct an approximately force-free field where current and magnetic field are almost parallel. Also note that density is no longer relevant for the structure of the equilibrium, meaning that we now only have one scalar field P to balance the Lorentz force with its two degrees of freedom, just as in the case above with gravity and a barotropic E.O.S.

[5] Strictly speaking, adjusting the density field will affect the gravitational field $\mathbf{g}$, but if only small adjustments to the non-magnetic equilibrium are needed it is hard to imagine that changes in $\mathbf{g}$ will prevent the existence of an equilibrium.

Exercises

8.1 Field amplification

This problem examines how a magnetic field can be amplified in a given velocity field. The effect of the magnetic field on the velocity field is ignored, which is called the *kinematic regime.* Assume ideal MHD, i.e. perfect flux freezing.

(a) An initially uniform field $B_0\hat{\mathbf{y}}$ evolves in a shear flow where $\mathbf{u} = ay\hat{\mathbf{x}}$ where a is a constant. Find an expression for the field at time t.

(b) Consider a magnetic field in a volume with some velocity field bounded by a stationary surface of fixed magnetic field, outside which the velocity field is zero. Initially, the field is in the lowest energy state possible, i.e. it is a potential field. Furthermore assume that the initial field is uniform in strength and direction, and ignore one of the dimensions perpendicular to the field. By relating the strength of the field to the distance between neighbouring field lines, argue that the lengthening of the field lines which results from 'stirring' of the fluid inevitably leads to higher energy.

(c) With the help of the continuity equation and some vector identities, write the induction equation in terms of $\mathrm{d}(\mathbf{B}/\rho)/\mathrm{d}t$. Comment on the physical meaning. [*Hint: the same was done before for the vorticity equation.*]

8.2 Tension of a flux tube

Imagine a flux tube of circular cross section with radius a and length l containing a uniform magnetic field B.

(a) Whilst holding a constant, by considering the increase in magnetic energy whilst increasing the length by a small quantity, calculate the tension T_{mag} of the tube (in units of energy per unit length, i.e. force).

(b) The tube is in equilibrium in the lateral direction with its unmagnetised surroundings, giving rise to a difference in gas pressure of magnitude B^2/π. Now calculate the total energy required to stretch the tube (again at constant a) by considering not only the increase in magnetic energy but also the $P\,\mathrm{d}V$ work done against the external gas in increasing the volume of the tube, minus that work done by the gas in the tube, showing that the tension is now double that found in part (a). Note that this is often neglected in the literature and that the tension calculated in part (a) is often used.

(c) Again assuming equilibrium in the lateral direction, calculate the tension in the tube again by considering the change in magnetic energy during a stretch at *constant volume*, so that the gas does no work, and constant magnetic flux.

(d) The tube is connected to itself in a circular loop of radius r where we can assume $r \gg a$. The Lorentz force is now non-zero in the interior of the tube and points towards the centre of the circle. By integrating this curvature force, the second term in (8.28), over the entire volume of the tube, one finds the magnetic energy change during an infinitesimal change in r. Using that fact that $\mathrm{d}l = 2\pi\mathrm{d}r$, show that this gives the same tension in the tube as that found in parts (b) and (c), arguing that the Lorentz force at the surface of the tube can be neglected if the volume remains constant.

8.3 MHD waves

Starting from equations (8.34) to (8.38), derive the dispersion relations of waves where the wavevector is both parallel and perpendicular to the magnetic field, finding the phase and group velocities in each case and showing that the waves are non-dispersive (that is, speed does not depend on frequency). In addition, find the relation between the energy associated with the perturbation to the magnetic field $\delta\mathbf{B} \cdot \mathbf{B}/4\pi$, the perturbation to the gas, and the kinetic energy, showing that the magnetic and kinetic are in equal in the parallel case, and that in the perpendicular case the kinetic energy accounts for one half of the energy and the other two forms account for the other half.

Chapter 9

MHD: astrophysical contexts

We now illustrate the principles introduced in the last chapter and look at some processes and phenomena in the astrophysical contexts of the solar corona, jets and accretion discs.

9.1 The solar corona

Figure 9.1: The solar corona photographed during the eclipse of 11 August 1999.

On the photosphere of the Sun, we observe magnetic fields in the quiescent regions which are structured on the granulation scale ($\sim$ 1000 km) and hundreds of gauss in strength, and in addition we see active regions with sunspots of sizes 10 to 100 times the granulation scale in which the magnetic field is in the range $1 - 3$ kG. The thermal energy density at the photosphere is about the same as a magnetic field of $\sim$ kG, which we call the 'equipartition field strength'. Since the flow speeds at and just below the photosphere are roughly sonic, the kinetic energy density is about the same. This explains why a field of at least around 1 kG is required to have much effect on the appearance of the photosphere. The thermal energy density increases rapidly below the photosphere and decreases rapidly above it – more rapidly than the magnetic energy density (pressure falls exponentially whereas the magnetic field tends to fall geometrically on large scales) so we can think of a $\beta = 1$ surface which lies at or just above the photosphere. Below this surface, the magnetic field has a rather subtle effect on the flow of gas; above this surface the magnetic field dominates. This region above the photosphere is called the corona, although strictly speaking the two are separated by the chromosphere and the transition region. The corona has a temperature of around $1-2$ million kelvin, which contrasts to the photospheric temperature of 5800 K – the origin of this high temperature, the 'coronal heating problem', is one of the best-known unsolved problems in astrophysics. It is generally agreed that the magnetic field transports energy through the photosphere and that it is converted from magnetic to thermal form in the corona; what is not understood is how the magnetic energy is dissipated. Theories normally invoke either reconnection or excitation and dissipation of magnetic waves.

Now let us look at some of the parameters in the corona: $T \sim 10^6$ K, $\rho \sim 10^{-15}$ g cm^{-3}, $P = \rho T R/\mu \sim 10^{-1}$ erg cm^{-3}, $|\mathbf{g}| \sim 3\times10^4$ cm s^{-2}, $B \sim 10$ G. Assuming hydrostatic equilibrium this means that the scale height $H = P/\rho g \sim 3\times10^9$ cm, plasma $\beta = 8\pi P/B^2 \sim 3\times10^{-2}$, sound speed $c_s = \sqrt{\gamma P/\rho} \sim 10^7$ cm s^{-1}, Alfvén speed $v_A = B/\sqrt{4\pi\rho} \sim 10^8$ cm s^{-1}.

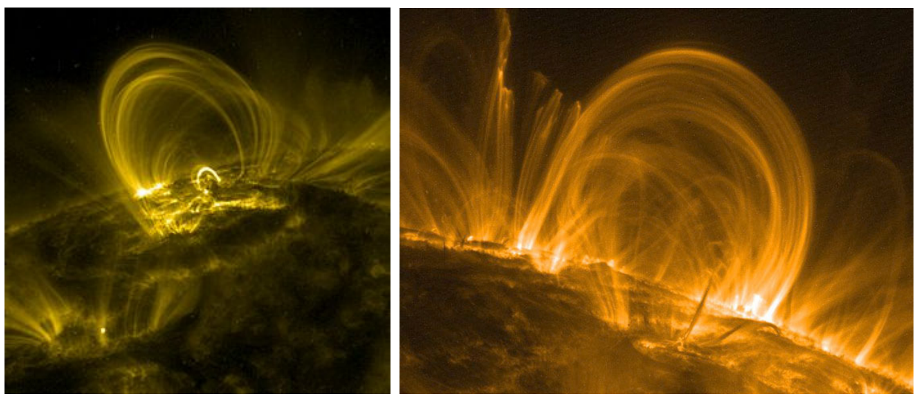


Figure 9.2: Images of coronal loops taken in the Fe IX line at 171Å by TRACE.

Looking at (8.21) it appears that the Lorentz force is much greater than the pressure gradient force. In addition, the gravitational force is comparable to the pressure force and acts only in the vertical direction. The Lorentz force must therefore be balanced by the inertia term on the left hand side of the momentum equation, meaning that flow speeds are comparable to the Alfvén speed. However, structures such as those in fig. 9.2 are observed to last for anything up to weeks, much greater than the Alfvén timescale $\tau_A = H/v_A \sim 30$ s. The only way out of this (as mentioned briefly in section 8.6) is for the Lorentz force to be reduced by having the current and magnetic field almost parallel to each other – we call this a 'force-free' field. The properties of these fields are explored in the next section. Often however we observe loop structures in the corona which, having apparently been in such a force-free equilibrium for some time, suddenly depart from equilibrium and convert much of their magnetic energy into heat on a timescale comparable to the Alfvén timescale. This process is explored in section 9.1.2.

9.1.1 Force-free and potential fields

In any low β plasma we see from (8.21) that the Lorentz force cannot apparently be balanced by the pressure gradient and it is difficult to imagine a gravitational field with the necessary geometry; balancing the Lorentz force with inertia, the term on the left-hand side of the momentum equation (8.16), would mean Alfvénic flow speeds and nothing approaching an equilibrium. We therefore speak of a 'force-free' field, where the current and magnetic field are almost parallel so that $|(\nabla\times\mathbf{B})\times\mathbf{B}| \ll B^2/\mathcal{L}$. This gives:

$$\nabla \times \mathbf{B} = \alpha \mathbf{B}, \tag{9.1}$$

$$\mathbf{B} \cdot \nabla \alpha = 0, \tag{9.2}$$

where the second equality comes from taking the divergence of the first and using the solenoidal condition $\nabla \cdot \mathbf{B} = 0$; it means that α is constant along field lines. In other words, as we follow a field line we see that the neighbouring lines curve around it in the same sense all the way along the line – force-free fields are 'twisted' in some sense. If α is a constant everywhere, we can write the Helmholtz equation $(\alpha^2 + \nabla^2)\mathbf{B} = 0$ by taking the curl of (9.1).

A special case is where $\alpha = 0$, which we call a *potential* or *curl-free* field, where the current vanishes. This is obviously the case in a vacuum, and is a good approximation in some other astrophysical contexts such as above the surface of some magnetic main-sequence stars. We call it a potential field because being curl-free we can express it as the gradient of a scalar potential $\mathbf{B} = \nabla\phi$. Since the divergence of the field is zero, we have the Laplace equation $\nabla^2\phi = 0$. This we can solve if we know the normal component $\mathbf{n} \cdot \mathbf{B} = \mathbf{n} \cdot \nabla\phi$ everywhere on the boundary of the domain.

Energy minima

Imagine a fixed volume with a given normal field component at the boundary; outside the volume there is no motion and no change in the magnetic field, and no energy is injected into the volume in mechanical or other form. One can solve for a potential field in the volume and the solution is unique. Furthermore, it can be shown that this field has the lowest energy of all which satisfy the boundary conditions. This can be seen by the following argument. Magnetic energy can be converted into other forms in two ways: into kinetic energy via the Lorentz force (this can work in both ways) and into heat energy via Joule heating (one-way conversion, see section 8.5). Any magnetic field with non-zero current will continuously be losing energy into heat, and this cannot be replaced by conversion from kinetic, since we know from the second law of thermodynamics that the heat energy cannot entirely be converted back in kinetic. There can be some back-and-forth flow of energy between magnetic and kinetic, i.e. oscillations, but these necessarily involve the magnetic field being always (except perhaps fleetingly) in a non-potential state with non-zero Joule heating. The energy of any non-potential field must therefore drop until the Joule heating vanishes entirely, i.e. when the current is zero everywhere – quod erat demonstrandum. If there is no means to convert kinetic energy directly into heat (viscosity $\nu = 0$) then all that can remain is acoustic oscillations propagating parallel to the field lines (a special and unlikely situation), otherwise all kinetic and all *free* magnetic energy is ultimately converted into heat. In summary, any field in a medium of finite conductivity in a volume with fixed boundaries will relax to this potential field. Once this state is reached, there is no way to extract the remaining $B^2/8\pi$ energy without changing the boundaries.

The situation is different if flux-freezing holds perfectly, i.e. the medium has infinite conductivity. If we also assume zero kinetic diffusivity there is no conversion of either magnetic or kinetic energy into heat and oscillations will not be damped. If viscosity is present, however, kinetic energy will be converted to heat and any oscillations or 'sloshing' will therefore result in energy loss from magnetic/kinetic into heat. This can only result in a stationary final state, which is presumably an energy minimum, i.e. minimum magnetic energy and zero kinetic energy. Perturbing this state with a displacement field $\boldsymbol{\xi}$ should therefore not affect the magnetic energy, since the energy should rise at fastest quadratically away from the minimum. The only way this can be possible is if the energy minimum is force-free. Force-free states are no longer unique in the same way as the potential states – here, unlike the finite conductivity case, flux freezing holds and so the final state now depends not only on the normal component of the field at the boundary but on the topology of the field, i.e. how the field lines entering the volume connect to those leaving.[1]

Astrophysically, both of these cases occur in many situations. An example of a potential field would be the volume outside an intermediate mass main-sequence star – inside the star is the conductivity is high and the Ohmic timescale is very long, so the star can contain a long-lived equilibrium which gives an essentially fixed normal component at the surface. Outside the star the conductivity and therefore the Ohmic timescale are much lower than in the interior so that the field relaxes to a potential field. In the *solar* corona, on the other hand, the temperature and therefore conductivity are higher than in the intermediate mass star. Moreover, the field at the surface is not static but moves on timescales of minutes to weeks, much shorter than the Ohmic timescale, so that the field does not have time to relax to potential. In addition, since the mean free path is so large, the kinetic diffusivity is large – very much larger than the magnetic diffusivity, so that kinetic energy can be removed but flux freezing holds. The field in the solar corona is indeed observed to be very close to force-free, which of course it has to be to avoid velocities on the order of the Alfvén velocity since this is a low-β plasma where gas pressure and gravity are not able to balance the Lorentz force, as mentioned above.

Vanishing force-free field theorem

There is a theorem which states that no equilibrium can be force-free everywhere. Imagine a force-free equilibrium in a region of volume V surrounded by an unmagnetised region, and imagine an isotropic expansion or contraction of the region. Under such a change, the position vector of any fluid element changes from $\mathbf{r}$ to $\mathbf{r}'$ and the field from $\mathbf{B}$ to $\mathbf{B}'$. In a uniform expansion by a factor a we have $\mathbf{r}' = a\mathbf{r}$ and from flux conservation ($\phi = Br^2$ is constant) we see that $r'^2\mathbf{B}'(\mathbf{r}') = r^2\mathbf{B}(\mathbf{r})$. The energy of the field after the expansion is

$$E' = \int_{V'} \frac{B'^2}{8\pi} \mathrm{d}V' = \frac{1}{a} \int_V \frac{B^2}{8\pi} \mathrm{d}V \tag{9.3}$$

since $\mathrm{d}V' = a^3\mathrm{d}V$ and $\mathbf{B}' = a^{-2}\mathbf{B}$. The region will therefore expand until either some force opposes it, at which point it is no longer force-free, or it reaches infinite extent and $E \to 0$. [More generally, anything with positive energy will tend to expand.] A force-free region must be subject to forces on its boundary.

9.1.2 Reconnection

We saw in above in section 8.6 and from equation (8.19) that the timescale over which the magnetic diffusivity acts is $\tau_{\rm diff} \sim \mathcal{L}^2/\eta$. However, we see in many astrophysical contexts such as the solar corona that changes in global magnetic topology, i.e. deviations from flux-freezing, and the associated dissipation of magnetic energy can occur on much shorter timescales. For instance, energy is released during solar flares over timescales of seconds and minutes although $\tau_{\rm diff} \gtrsim 10^6$yr, assuming a standard Spitzer conductivity. There are two possible reasons for this. The first is that somehow the diffusion is locally brought to work on shorter length scales than the global length scale $\mathcal{L}$, the second is that there is

[1] In fact even given both the boundary and topological constraints, the uniqueness of the force-free field is not obvious since local energy minima seem plausible, but that is outside the scope of this course.

some 'anomalous resistivity' – higher than the standard resistivity – perhaps when the current density is particularly high; plasma instabilities may also be involved. It seems likely that in some situations both mechanisms must be invoked.

Producing structure on small length scales from a global configuration initially lacking such small scales is a common phenomenon in physics. For instance, to mix two paints together so that they combine on the microscopic scale, it is sufficient to stir with a large spoon. In a turbulent flow, large-scale driving leads to the appearance of structure on a scale sufficiently small that the viscous dissipation timescale is equal to the flow timescale.[2] Similarly, in a shock a small length scale is produced to allow fast diffusive conversion of kinetic to heat energy. In the example of solar flares, we do not see turbulence; rather, the field is thought to dissipate in thin current sheets or otherwise localised features which separate regions with different magnetic field. This process is known as 'reconnection'.

The first model of reconnection, the *Sweet-Parker* mechanism, is illustrated in fig. 9.3. Material flows perpendicular to the field lines at speed v_0 towards the current sheet. This speed is equated to the 'diffusion speed' within the current sheet, which from the induction equation is equal to $\eta/\delta \approx v_0$. Assuming incompressibility, conservation of mass gives $v_0 L \approx v_* \delta$. Now, if we consider the force balance in the y direction it is clear that there is an excess thermal pressure at the centre of the current sheet, where the magnetic field vanishes, equal to the magnetic pressure at the boundaries of the sheet, i.e. $B^2/8\pi$. The material is accelerated in the x direction by this thermal pressure and escapes at the ends of the sheet; we have from the force balance that $\rho v_*^2/2 \approx B^2/8\pi$ which can be rearranged to $v_* \approx v_A = B/\sqrt{4\pi\rho}$. Now we can solve for the reconnection velocity v_0:

$$v_0 \approx \frac{\eta}{\delta} \approx \frac{v_A \eta}{v_0 L} \qquad \Longrightarrow \qquad \frac{v_0}{v_A} \approx \sqrt{\frac{\eta}{v_A L}} = \mathrm{Re_A}^{-1/2}, \tag{9.4}$$

where $\mathrm{Re_m}$ is the magnetic Reynolds number. This model produces reconnection speed ratios of $v_0/v_A \sim 10^{-6}$ in the solar corona and other astrophysical plasmas, which is unfortunately rather less than the generally observed value of ~ 0.1. One way out of this is the Petschek reconnection model in which most of the energy is dissipated in standing shocks attached to a small central Sweet-Parker-like diffusion region. Some kind of anomalous resistivity probably also plays a role, as well as 3-D effects. Finally, note that reconnection does not only convert magnetic energy into kinetic and thermal, but it also accelerates particles up to relativistic speeds – this is observed in the solar corona. In general, magnetic fields are ingredients in most cosmic-ray acceleration mechanisms.

9.2 Jets: launching, collimation and instabilities

Jets are found in many astrophysical accretion settings, for example protostars, neutron stars, AGN. In this section we examine the magneto-centrifugal model of jet launching and collimation.

9.2.1 Launching

Imagine a Keplerian disc around a central object threaded by a magnetic field. The field which emerges from the disc is 'ordered' in some sense. If, for simplicity, we assume that the field component emerging

[2] Opinions differ as to how this works.

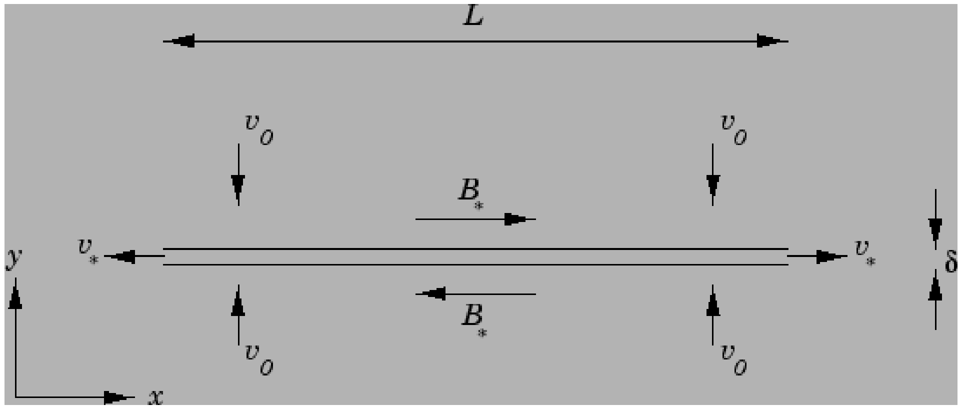


Figure 9.3: The Sweet-Parker reconnection mechanism. Regions of opposing magnetic field B_* are brought together, separated by a thin sheet of thickness δ.

normal to the disc is of uniform sign and its strength varies with cylindrical radius as $B_z = B_0(\varpi^2/\varpi_0^2 + 1)^{-1/2}$, and then assume that the field above the disc is curl-free (i.e. zero-current, force-free with $\alpha = 0$) we have the field illustrated in fig. 9.4. Near the disc the lines are inclined away from the centre because of the greater field strength at the centre, but further from the disc they tend towards the vertical because of the fact that flux per unit cylindrical radius increases outwards – in other words, most of the flux is threaded through the outer disc and so from a distance the inner part is of lesser importance. We expect that $B_z \propto \Sigma$ where Σ is the column density of the disc, so this picture of flux increasing outwards is realistic as long as $\varpi\Sigma$ increases outwards, i.e. $\partial \ln \Sigma / \partial \ln \varpi > -1$.

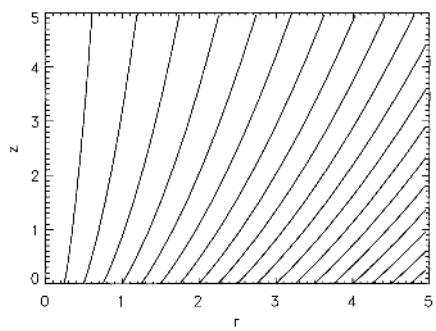


Figure 9.4: The field above a disc from which a vertical field component $B_z = B_0(\varpi^2/\varpi_0^2 + 1)^{-1/2}$ emerges, assuming the field is curl-free. Note that the field lines curve towards the vertical. [From Spruit et al. 1997.]

We now allow material to evaporate from the disc. Just above the disc this material will have a low density in the sense that the magnetic energy density $B^2/8\pi$ is much greater than the thermal (i.e. $\beta \ll 1$) and also than the kinetic $\rho u^2/2$. This means that the field must be force-free and that it is little affected by the material. Flux freezing requires that the material flows along the field lines, and since the field lines point away from the central object, the material is centrifugally accelerated away from the disc; it can be shown that if the angle between the field and the vertical exceeds some threshold then this centrifugal acceleration exceeds the downwards gravitational acceleration. The material is forced to co-rotate with the magnetic field.

As the material is accelerated its kinetic energy density eventually exceeds the magnetic, i.e. $\rho u^2/2 > B^2/8\pi$ or alternatively $u > v_{\mathrm{A}}$; we say that this transition happens at the 'Alfvén surface' which is analogous to the sonic point in non-magnetised flows such as in the nozzle of a rocket engine. Flux freezing holds on both sides, but whereas inside the Alfvén surface the flow follows the field lines, outside the Alfvén surface the field lines follow the flow. This means that the material ceases to co-rotate with the foot points of the field lines; rather, the field lines are 'wound up' so that a significant toroidal component B_t is produced.

9.2.2 Collimation

We saw above that the poloidal field may have a tendency to collimate the flow (fig. 9.4). In this section we examine the collimation after the initial acceleration phase, when the energy in the z component of the motion dominates and therefore $u_z \approx$ const.

Imagine a jet with circular cross-section of radius $a(z)$ at a distance z from the central object (outside the Alfvén surface). It contains a spiral magnetic field with toroidal and poloidal components B_t and B_p. If the toroidal and poloidal fluxes are both conserved as the material moves away from the source, then as the radius of the jet changes we have $B_t \propto a^{-1}$ and $B_p \propto a^{-2}$. After the initial acceleration phase the jet will expand ballistically with $a \propto z$ in the absence of significant pressure or magnetic forces. Obviously in the non-magnetic case, if the thermal pressure in the jet is greater than that in the surroundings, the jet will expand faster than $a \propto z$ – the jet will be 'flared'. The poloidal component of the field will also tend to make the jet flare as it exerts a pressure $B_p^2/8\pi$ on the jet's surroundings; the toroidal field exerts no pressure because its pressure and tension forces are equal and opposite. Another way of thinking about this is in terms of the energy per unit length of the jet in the toroidal and poloidal components of the field: $E_t = \pi a^2 B_t^2 = \pi \Phi_t^2$ and $E_p = \pi a^2 B_p^2 = \Phi_p^2/(\pi a^2)$. It is the drop in E_p which drives the expansion of the jet. It is therefore necessary to have some external pressure for collimation to occur – one can imagine that with some constant pressure in the ambient medium the jet would be flared near the source and then further away would settle at constant a when its internal pressure $P_{\rm jet} + B_p^2/8\pi$ is equal to the external.

It is worth looking in more detail at the collimating effect of the toroidal field. A jet with $B_t = B_t(\varpi)$ carries a current $\mathbf{J} = \hat{\mathbf{z}}(c/4\pi\varpi)\partial(\varpi B_t)\partial\varpi$ and so the Lorentz force is

$$\mathbf{F}_{\rm Lor} = -\hat{\boldsymbol{\varpi}} \frac{B_t}{4\pi\varpi} \frac{\partial(\varpi B_t)}{\partial\varpi} \tag{9.5}$$

where $\hat{\boldsymbol{\varpi}}$ is the unit vector in the ϖ direction. At least near the axis this force must be directed inwards: this effect is often referred to as 'hoop stress'. This has lead to the misleading concept of 'self collimation', according to which a jet can be collimated by its own toroidal magnetic field. The problem with this is that to avoid having the energy diverge towards infinity, the partial derivative must change sign at some radius outside which the Lorentz force is directed outwards, requiring in effect some external pressure support.

9.2.3 Instability

We saw above that as a jet expands the poloidal field falls faster than the toroidal, which leads eventually to an instability driven by the free energy in the toroidal field. It can be shown that the dominant modes have azimuthal wavenumbers $m = 0$ (sausage mode) and $m = 1$ (kink mode); here we look at a simple derivation of the instability criterion for the $m = 0$ mode.

A purely toroidal field is some function of cylindrical radius, $B = B(\varpi)$. Imagine two thin annuli at radii ϖ and $\varpi + \delta\varpi$ with magnetic fields B and $B + \delta B$, each of area A and therefore of thicknesses $A/(2\pi\varpi)$

and $A/(2\pi(\varpi + \delta\varpi))$. The energy per unit length jet of the magnetic field in the annuli is

$$E = \frac{A}{8\pi}\left[B^2 + (B + \delta B)^2\right]. \tag{9.6}$$

We now exchange adiabatically the positions of the two annuli, keeping the volume of each constant. In general, the most unstable modes of any instability will be incompressible (density unchanged), since compressing the gas will require work to be done by the magnetic field. Since the volumes of the annuli remain the same, the total thermal energy is unchanged. Since flux is conserved, the new fields at locations ϖ and $\varpi + \delta\varpi$ are $(B + \delta B)\varpi/(\varpi + \delta\varpi)$ and $B(\varpi + \delta\varpi)/\varpi$, so that the new energy is

$$E + \delta E = \frac{A}{8\pi}\left[\left(\frac{(B + \delta B)\varpi}{\varpi + \delta\varpi}\right)^2 + \left(\frac{B(\varpi + \delta\varpi)}{\varpi}\right)^2\right]. \tag{9.7}$$

For stability we need the exchange to have increased the energy, i.e. $\delta E > 0$. Subtracting (9.6) from (9.7) and dividing by $A/8\pi$ we have

$$\left(\frac{(B + \delta B)\varpi}{\varpi + \delta\varpi}\right)^2 + \left(\frac{B(\varpi + \delta\varpi)}{\varpi}\right)^2 - B^2 - (B + \delta B)^2 > 0 \tag{9.8}$$

$$\left(1 + 2\frac{\delta B}{B} + \frac{\delta B^2}{B^2}\right) + \left(1 + 2\frac{\delta\varpi}{\varpi} + \frac{\delta\varpi^2}{\varpi^2}\right)\left[1 + 2\frac{\delta\varpi}{\varpi} + \frac{\delta\varpi^2}{\varpi^2} - 1 - \left(1 + 2\frac{\delta B}{B} + \frac{\delta B^2}{B^2}\right)\right] > 0, \tag{9.9}$$

where the zeroth and first order terms cancel; keeping only the second order terms we have

$$\frac{\delta\varpi}{\varpi} - \frac{\delta B}{B} > 0 \qquad \text{or} \qquad \frac{\partial \ln B}{\partial \ln \varpi} < 1. \tag{9.10}$$

This corresponds to the result of Tayler (1957). For obvious reasons, this is known as an 'interchange' mode. A more general treatment including the non-axisymmetric modes reveals that $m \geq 1$ modes are stable if $\partial \ln B/\partial \ln \varpi < m^2/2 - 1$, meaning that the $m = 1$ mode (the 'kink mode') sets in first. Note that to avoid a current singularity we need $\partial \ln B/\partial \ln \varpi \geq 1$ on the axis, so it is impossible in practice to construct a toroidal field which is stable everywhere. The growth timescale of all modes is comparable to the dynamical timescale, i.e. the Alfvén timescale ϖ/v_A.

A jet contains not only a toroidal component, of course, but also an axial component B_z which can help to stabilise the jet against this instability. We can see approximately how strong this component needs to be by means of the following energy argument. It is clear that as the instability grows, work needs to be done against the axial component of the field and for stability this must be greater than the energy released from the toroidal part of the field via the instability. Now, since $\ddot{\boldsymbol{\xi}} = \sigma^2\boldsymbol{\xi}$ where the growth rate $\sigma = v_A/\varpi$, the $m = 1$ mode releases an energy per unit volume equal to $\frac{1}{2}\rho\ddot{\xi}\xi = \frac{1}{2}\rho\sigma^2\xi^2$. As the axial field lines are stretched, they exert a restoring force $\frac{1}{4\pi}B_z^2\xi/l_z^2$ where l_z is the length scale of the instability in the z direction, which is equal to $1/k_z = \lambda_z/2\pi$. The work to be done therefore is $\frac{1}{8\pi}B_z^2\xi^2/l_z^2$, which for stability needs to be greater than the energy released so that

$$\frac{1}{8\pi}B_z^2\xi^2/l_z^2 > \frac{1}{2}\rho\sigma^2\xi^2 \tag{9.11}$$

which can be rewritten as

$$\frac{B_z^2}{4\pi\rho}\frac{1}{l_z^2} > \sigma^2 \tag{9.12}$$

$$\frac{B_z^2}{4\pi\rho}\frac{\varpi^2}{l_z^2} > \frac{B_\varpi^2}{4\pi\rho} \tag{9.13}$$

$$\frac{B_z}{B_\varpi} > \frac{\lambda_z}{2\pi\varpi}. \tag{9.14}$$

The axial field therefore stabilises the shortest wavelengths, because it is the shortest wavelengths which have to bend the axial field lines to a greater degree for a given $|\boldsymbol{\xi}|$. Another way of expressing this result is that instability sets in for wavelengths greater than the distance over which a field line makes one full circle around the jet. This is often referred to in the literature as the *Kruskal-Shafranov* condition. Unfortunately, the ratio λ_z/ϖ could be very large in a narrow jet so this instability is expected to be present in essentially all collimated jets; work continues therefore on its non-linear development.

9.3 Angular momentum transport in discs

Accretion discs are found in many astrophysical contexts, such as during star formation, mass transfer in binary systems, and accretion of gas onto supermassive black holes. When matter is accreted, it falls deeper into a gravitational potential well and energy must be lost from the system, generally via radiation. However, it is not possible for a disc of material in isolation to accrete entirely onto the central object because of angular momentum conservation – angular momentum must be removed from the accreting material by transfer to other material which does not accrete. Therefore the lowest-energy end-state of a disc in isolation is to have an infinitesimally small amount of mass move outwards towards infinity and infinite specific angular momentum and the rest of the mass accreted onto the central object. This requires transport of angular momentum outwards.[3] Since all systems like to relax to energy minima, we should expect to find some mechanism operating in discs for the outward transport of angular momentum. According to the model of Shakura & Sunyaev (1973), the observational properties of accretion discs can be reproduced well by imagining there is some viscous stress equal in magnitude to some fraction α of the thermal pressure P. However, microscopic viscosity is far too small to account for values $\alpha \sim 0.1$ inferred from the observations; we turn our attention therefore to possible instabilities which could generate turbulence and the resultant 'turbulent viscosity'.

An example of a shear instability in a differentially rotating flow is the Rayleigh instability, an interchange instability. Imagine exchanging two annuli of equal volume and density between two radii ϖ and $\varpi + \delta\varpi$ which are initially moving with angular velocities Ω and $\Omega + \delta\Omega$. The kinetic energies before and after the exchange are E and $E + \delta E$:

$$E = \frac{\rho V}{2}\left[\varpi^2\Omega^2 + (\varpi+\delta\varpi)^2(\Omega+\delta\Omega)^2\right], \tag{9.15}$$

$$E + \delta E = \frac{\rho V}{2}\left[\varpi^2(\Omega+\delta\Omega)^2\left(\frac{\varpi+\delta\varpi}{\varpi}\right)^4 + (\varpi+\delta\varpi)^2\Omega^2\left(\frac{\varpi}{\varpi+\delta\varpi}\right)^4\right]. \tag{9.16}$$

$$\delta E \approx 2\rho V\varpi^2\Omega^2\delta\ln\varpi^2\left[2 + \frac{\delta\ln\Omega}{\delta\ln\varpi}\right] \tag{9.17}$$

[3] Transport is outwards in the Langrangian sense of each fluid element transfering its angular momentum to its outside neighbour. Imagining the rate of change of angular momentum inside a fixed volume containing the central object (whose angular momentum is increasing as it is spun up and becomes more massive) and part of a steady-state disc, we see that net transport must be *inwards*; in other words, the inwards advection of angular momentum exceeds the transport though turbulent stress, albeit only by some small amount.

so that the stability condition is $q \equiv \partial \ln \Omega / \partial \ln \varpi > -2$, or in other words that the specific angular momentum $\varpi^2 \Omega$ increases outwards. An accretion disc ($q = -3/2$) is therefore stable to this mechanism. However, another kind of shear instability, the *magneto-rotational instability* has the stability condition that $q > 0$.

9.3.1 Magneto-rotational instability: physical mechanism and stability condition

The physical mechanism can be thought of in the following way. Consider two fluid elements, initially at the same radius, one above the other so that they are threaded by the same field line. They are given a perturbation in the radial direction, in opposite senses. Initially their angular momenta remain unchanged so that the element which has been perturbed inwards moves faster than the other. As it moves forwards with respect to the other, the field line connecting them is stretched so that the inner element pulls on the outer, transferring angular momentum to it. This causes the inner element to move further inwards and the outer element to move further outwards. In addition, in some sense the field line can be thought of as a spring which oscillates at a frequency $v_{\rm A}/\lambda$ where λ is the wavelength. Clearly the spring has to have a lower intrinsic frequency than that at which it is driven in order that it can be stretched instead of oscillating, meaning that $v_{\rm A}/\lambda < \Omega/2\pi$. The field has therefore to be 'weak' in some sense for the instability to proceed.

Another way of imagining the instability is the following. Consider a fluid element at radius ϖ threaded by a vertical magnetic field B. In the rotating frame, in equilibrium the centrifugal force per unit mass at any radius $\Omega^2\varpi$ is balanced by some net inward-pointing force which in a disc is a combination of a pressure gradient and a gravitational force. The fluid element is now displaced a distance $\delta\varpi$ to a radius $\varpi + \delta\varpi$, while the magnetic field couples it to the other fluid elements on the same field line so that it retains its original angular velocity (angular momentum is transferred to it). The inward-pointing force at this new position is $(\Omega+\delta\Omega)^2(\varpi+\delta\varpi)$ but the centrifugal force is now $\Omega^2(\varpi+\delta\varpi)$. In addition there is a magnetic restoring force, giving the total (radial) force

$$\begin{aligned}
F &= \Omega^2(\varpi + \delta\varpi) - (\Omega + \delta\Omega)^2(\varpi + \delta\varpi) - \frac{B^2}{4\pi\rho}\frac{\delta\varpi}{(\lambda_z/2\pi)^2}, \\
&= \Omega^2\varpi\left[(1 + \delta\ln\varpi) - (1 + \delta\ln\Omega)^2(1 + \delta\ln\varpi) - \frac{v_{\rm A}^2}{\Omega^2}\frac{\delta\ln\varpi}{(\lambda_z/2\pi)^2}\right], \\
&= \Omega^2\varpi\left[-2\delta\ln\Omega - \frac{v_{\rm A}^2}{\Omega^2}\frac{\delta\ln\varpi}{(\lambda_z/2\pi)^2}\right], \\
&= -\Omega^2\delta\varpi\left(2q + \kappa_z^2\right),
\end{aligned} \tag{9.18}$$

where λ_z is the wavelength of the perturbation in the vertical direction, $v_{\rm A}^2 = B^2/4\pi\rho$ is the Alfvén speed and $\kappa_z = k_z v_{\rm A}/\Omega$ is a dimensionless wavenumber where $k_z = 2\pi/\lambda$ is the vertical wavenumber. For stability we need $F < 0$ and therefore

$$2q + \kappa_z^2 > 0, \tag{9.19}$$

remembering that $v_{\rm A}^2 = B^2/4\pi\rho$. For stability at all vertical wavenumbers, we clearly need $q > 0$, and in the unstable case the maximum unstable wavelength is given by $\kappa_z^2 = 2|q|$; as we shall see below, these estimates agree with the more rigorous treatment. Note that F is equal to the acceleration $\partial^2(\delta\varpi)/\partial t^2$; we can replace $\partial^2/\partial t^2$ with $-\omega^2$ where ω is the oscillation frequency (if real) or the growth rate (if imaginary). This gives, from (9.18),

$$-\omega^2 = -\Omega^2 \left(2q + \kappa_z^2\right), \tag{9.20}$$

so that the growth rate will generally be comparable to the rotation frequency. This not surprising, because the free energy source is the rotation. In this sense the instability is fundamentally different from the instability of toroidal fields described in section 9.2.3, whose free energy source is the magnetic field and whose growth time is comparable to the Alfvén timescale.

9.3.2 The dispersion relation

In our rotating fluid, locally we can look at a small corotating Cartesian volume at distance ϖ_0 from the centre, rotating at Ω_0, and change variables to $x = \varpi - \varpi_0$; the azimuthal direction is y and the vertical direction z. Since $\Omega(x) \approx \Omega_0 + qx\Omega_0/\varpi_0$, the bulk velocity in the rotating frame is $U \approx q\Omega_0 x$ (where $q = -3/2$ in the Keplerian case). In a rotating frame of reference, we must generally add to the momentum equation (8.16) both a Coriolis force $-2\rho\mathbf{\Omega}_0 \times \mathbf{v}$ (where $\mathbf{v}$ is the total velocity field) and a centrifugal force $\hat{\varpi}\rho\varpi\Omega_0^2$, but here we can drop the centrifugal term, the steady part of the Coriolis force coming from the basic flow U_y and the x component of the gravitational term because they cancel each other, leaving just the Coriolis force associated with any additional velocity field $\mathbf{u} = \mathbf{v} - U\hat{\mathbf{y}}$ on top of the basic flow, $-2\rho\mathbf{\Omega}\times\mathbf{u}$. In an accretion disc we also have the vertical part of the gravitational acceleration $-\rho GMz\varpi_0^{-3}\hat{\mathbf{z}} = -\rho z\mathbf{\Omega}_0^2\hat{\mathbf{z}}$, where the assumption is made that $z \ll \varpi_0$. This vertical stratification may be important in realistic discs, but we shall ignore it for the time being.

A general linear analysis of the MRI is quite involved, so we make two further simplifications: the initial magnetic field is of uniform strength and in the z direction so that $\mathbf{B} = B\hat{\mathbf{z}}$, and we consider only axisymmetric modes, meaning that $\partial/\partial y = 0$. The most unstable modes, at least for a weak field, will be incompressible and we assume that in the following. The perturbation to the magnetic field is $B\mathbf{b}$, so that $\mathbf{b}$ is dimensionless. We now linearise the MHD equations by subtracting equilibrium (zero-order) terms in the momentum equation, and keeping terms to first order in the perturbed quantities, ignoring the diffusive terms. Beginning with the momentum equation we have

$$\begin{aligned}
\partial_t\mathbf{u} + \mathbf{U}\cdot\mathbf{\nabla}\mathbf{u} + \mathbf{u}\cdot\mathbf{\nabla}\mathbf{U} &= -\frac{1}{\rho}\mathbf{\nabla}\delta P + \frac{1}{4\pi\rho}\left[(\mathbf{\nabla}\times(B\mathbf{b}))\times\mathbf{B} + (\mathbf{\nabla}\times\mathbf{B})\times(B\mathbf{b})\right] - 2\mathbf{\Omega}\times\mathbf{u},\\
\partial_t\mathbf{u} + \hat{\mathbf{y}}u_x\partial_x U &= -\frac{1}{\rho}\mathbf{\nabla}\delta P + v_{\mathrm{A}}^2(\mathbf{\nabla}\times\mathbf{b})\times\hat{\mathbf{z}} - 2\Omega\hat{\mathbf{z}}\times\mathbf{u},\\
\partial_t\mathbf{u} &= -\frac{1}{\rho}\mathbf{\nabla}\delta P + v_{\mathrm{A}}^2(\mathbf{\nabla}\times\mathbf{b})\times\hat{\mathbf{z}} - \Omega(2\hat{\mathbf{z}}\times\mathbf{u} + q\hat{\mathbf{y}}u_x),
\end{aligned} \tag{9.21}$$

noting that $\mathbf{U}\cdot\mathbf{\nabla}\mathbf{u} = \mathbf{0}$ and $\mathbf{\nabla}\times\mathbf{B} = \mathbf{0}$. We can now consider perturbations which vary in space and time as $e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}$ so that $\partial_x = ik_x$ and so on:

$$-i\omega\mathbf{u} = -\frac{i\mathbf{k}}{\rho}\delta P + iv_{\mathrm{A}}^2(\mathbf{k}\times\mathbf{b})\times\hat{\mathbf{z}} - \Omega(2\hat{\mathbf{z}}\times\mathbf{u} + q\hat{\mathbf{y}}u_x), \tag{9.22}$$

$$-i\psi\mathbf{u} = -i\boldsymbol{\kappa}\frac{\delta P}{\rho v_{\mathrm{A}}} + iv_{\mathrm{A}}(\boldsymbol{\kappa}\times\mathbf{b})\times\hat{\mathbf{z}} - 2\hat{\mathbf{z}}\times\mathbf{u} - q\hat{\mathbf{y}}u_x, \tag{9.23}$$

where we have introduced a dimensionless wavenumber $\boldsymbol{\kappa} \equiv \mathbf{k}v_{\mathrm{A}}/\Omega$ and a dimensionless frequency $\psi \equiv \omega/\Omega$. The induction equation becomes (assuming $\mathbf{\nabla}\cdot\mathbf{u} = 0$)

$$\partial_t(B\mathbf{b}) = \mathbf{\nabla}\times(\mathbf{u}\times\mathbf{B} + \mathbf{U}\times(B\mathbf{b})). \tag{9.24}$$

$$
\begin{aligned}
\partial_t \mathbf{b} &= \boldsymbol{\nabla} \times (\mathbf{u} \times \hat{\mathbf{z}} + U\hat{\mathbf{y}} \times \mathbf{b}). && (9.25)\\
\partial_t \mathbf{b} &= \partial_z \mathbf{u} + q\Omega\hat{\mathbf{y}} b_x, && (9.26)\\
-i\psi\mathbf{b} &= i\kappa_z \mathbf{u}/v_{\mathrm{A}} + q\hat{\mathbf{y}} b_x, && (9.27)\\
i\psi v_{\mathrm{A}} \mathbf{b} &= -\kappa_z \left(i\mathbf{u} - q\hat{\mathbf{y}} u_x/\psi\right), && (9.28)
\end{aligned}
$$

where the last line was obtained by substituting for b_x back into the y component of the equation. This can be substituted into the momentum equation (multiplied by ψ) to give:

$$
-i\psi^2 \mathbf{u} = -i\boldsymbol{\kappa}\psi \frac{\delta P}{\rho v_{\mathrm{A}}} - \kappa_z \left[\boldsymbol{\kappa} \times (i\mathbf{u} - q\hat{\mathbf{y}} u_x/\psi)\right] \times \hat{\mathbf{z}} - \psi(2\hat{\mathbf{z}} \times \mathbf{u} + q\hat{\mathbf{y}} u_x). \tag{9.29}
$$

Along with the incompressibility condition we have four equations and four quantities to be eliminated, δP and the three components of $\mathbf{u}$. Therefore:

$$
\begin{aligned}
-i\psi^2 u_x &= -i\kappa_x \psi \frac{\delta P}{\rho v_{\mathrm{A}}} - i\kappa_z(\kappa_z u_x - \kappa_x u_z) + 2\psi u_y, && (9.30)\\
-i\psi^2 u_y &= -\kappa_z^2 (iu_y - qu_x/\psi) - (2+q)\psi u_x. && (9.31)\\
-i\psi^2 u_z &= -i\kappa_z \psi \frac{\delta P}{\rho v_{\mathrm{A}}}. && (9.32)
\end{aligned}
$$

From the third of these, we see that $-i\kappa_x\psi\delta P/\rho v_{\mathrm{A}} = -i\psi^2 u_z \kappa_x/\kappa_z$ which, using the incompressibility condition $\kappa_x u_x + \kappa_z u_z = 0$, replaces the first term on the right hand side of the x-component equation and reduces the set to two equations and two variables u_x and u_y:

$$
\begin{aligned}
-i\psi^2 u_x &= i\psi^2 u_x \frac{\kappa_x^2}{\kappa_z^2} - iu_x(\kappa_z^2 + \kappa_x^2) + 2\psi u_y, && (9.33)\\
-i\psi^2 u_y &= -\kappa_z^2 (iu_y - qu_x/\psi) - (2+q)\psi u_x. && (9.34)
\end{aligned}
$$

Collecting terms with u_x and u_y gives

$$
\begin{aligned}
\left(i\psi^2 \frac{\kappa^2}{\kappa_z^2} - i\kappa^2\right) u_x + 2\psi u_y &= 0, && (9.35)\\
\left(\frac{q\kappa_z^2}{\psi} - (2+q)\psi\right) u_x + \left(i\psi^2 - i\kappa_z^2\right) u_y &= 0, && (9.36)
\end{aligned}
$$

where $\kappa^2 = \kappa_x^2 + \kappa_z^2$. Taking the determinant of $\mathbf{A}$ in $\mathbf{A} \cdot \mathbf{u} = \mathbf{0}$ to be zero we have the following quadratic in ψ^2:

$$
\left(\psi^2 \frac{\kappa^2}{\kappa_z^2} - \kappa^2\right)\left(\psi^2 - \kappa_z^2\right) + 2\psi\left(\frac{q\kappa_z^2}{\psi} - (2+q)\psi\right) = 0, \tag{9.37}
$$

which we rearrange to:

$$
\frac{\kappa^2}{\kappa_z^2}\psi^4 - 2\left(\kappa^2 + 2 + q\right)\psi^2 + \kappa_z^2\left(\kappa^2 + 2q\right) = 0. \tag{9.38}
$$

Solving the quadratic in ψ^2 we have

$$
\psi^2 = \frac{\kappa_z^2}{\kappa^2}\left[\kappa^2 + 2 + q \pm \sqrt{4\kappa^2 + (2+q)^2}\right]. \tag{9.39}
$$

It is straightforward to show from this that the stability condition, i.e. the condition that both roots are postive, is $2q + \kappa^2 > 0$ which is identical to (9.19) derived above. For stability at all wavenumbers we require $q > 0$, i.e. an angular velocity increasing with radius. For $q < 0$, there is instability for a range of wavenumbers $0 < \kappa^2 < 2|q|$ and the maximum growth rate is $|\psi_{\rm max}| = |q|/2$ at a wavenumber $\kappa^2_{z,\,\rm max} = (1 + q/4)|q|$ (and $\kappa_x = 0$).

9.3.3 MRI: remarks

We saw above that an accretion disc with a weak vertical magnetic field suffers an instability with some minimum wavelength. From vertical force balance, we see that $H \sim \varpi c_{\rm s}/v_{\rm Kep}$ where H is the thickness of the disc, $c_{\rm s}$ is the sound speed and $v_{\rm Kep}$ is the Keplerian orbit speed. Evidently, if this instability is to be effective the minimum wavelength must be less than H, so that since $q = -3/2$ we see from (9.19) that

$$\frac{2\pi v_{\rm A}}{\sqrt{3}\Omega} \lesssim \frac{\varpi c_{\rm s}}{v_{\rm Kep}} \qquad \Longrightarrow \qquad \beta \gtrsim \frac{4\pi^2}{3\gamma}. \tag{9.40}$$

The magnetic field must not therefore become too strong relative to the thermal pressure. Also, a strong magnetic field would be buoyantly unstable.

We looked here at the axisymmetric modes where the unperturbed field is parallel to the rotation axis; in reality we would expect the field to be dominated by its azimuthal component since this is the direction of the shear. The instability does operate on a purely azimuthal field but the modes are non-axisymmetric and the dispersion relation is more complex. Numerical nonlinear analysis of this instability shows a steady-state dynamo effect as well as the desired angular momentum transport with a Shakura-Sunyaev α parameter of roughly the right magnitude, 0.01 to 0.1. However, many issues remain and the instability and how it works in discs is still far from fully understood.

Finally, note that in the non-magnetic limit $v_{\rm A} \to 0$ the dispersion relation becomes

$$\frac{\omega^2}{\Omega^2} = 2(2+q)\frac{k_z^2}{k^2}, \tag{9.41}$$

so that we recover the Rayleigh stability condition $q > -2$ derived at the beginning of this section.

Exercises

9.1 Equilibria in non-convective stars

A non-rotating upper main-sequence star is radiative apart from a small convective core, which we can ignore here.

(a) Use the Spitzer conductivity to make an order-of-magnitude estimate of the diffusive timescale on which any magnetic field present will evolve, and compare this to the main-sequence lifetime of the star and to the dynamic (Alfvén) timescale for a magnetic field of 1kG.

(b) The Ohmic timescale in the interior is much greater than the main-sequence lifetime of the star, but the Ohmic timescale in the exterior is very short. Assuming that the kinetic viscosity is large, show that after some time any magnetic field present settles into an equilibrium (that is, in an energy minimum) and argue that in equilibrium, the field is potential outside the star and non-force free in the interior. [Hints: note that the plasma-β is high inside and low outside the star, and use the vanishing force-free field theorem. Once formed, this equilibrium evolves quasi-statically on the interior Ohmic timescale which is much longer than other timescales of interest.]

(c) The equilibrium is axisymmetric. Using cylindrical coordinates (ϖ, ϕ, z) show that the field can be expressed as the sum of poloidal and toroidal components as

$$\varpi\mathbf{B} = \nabla\psi \times \hat{\phi} + F\hat{\phi} \tag{9.42}$$

where $\hat{\phi}$ is the azimuthal unit vector and ψ and F are functions of ϖ and z, and that the poloidal and toroidal fields are associated with toroidal and poloidal currents, respectively. By considering azimuthal force balance, show that contours of F in the meridional plane are parallel to those of ψ, i.e. that $F = F(\psi)$. [Hint: show that $(\nabla\psi) \times (\nabla F) = \mathbf{0}$.]

(d) The interiors of stars have $\beta \gg 1$. Show that an arbitrary axisymmetric magnetic field configuration can be added to a star and an equilibrium constructed by making small adjustments to the pressure and density fields. Ignore surface effects as well as any changes in the gravitational potential. Argue that this result can be generalised to non-axisymmetric fields. Note: an equilibrium does not necessarily need to be stable; it can also be unstable.

9.2 Kink instability in solar coronal loops

A coronal loop of magnetic field links two sunspots of opposite polarity. According to one theory of solar flares, reconnection events are triggered when a loop crosses the kink instability threshold.

(a) In a straight flux tube, the growth rate of the kink $m = 1$ instability is comparable to the Alfvén frequency, defined as $\omega_A \equiv v_A^\phi/\varpi$ where v_A^ϕ is the Alfvén speed associated with the azimuthal component B_ϕ of the magnetic field and ϖ is the cylindrical radius. An axial field component B_z can stabilise the field by providing an extra tension against which the instability must do work. In a tube where $B_\phi \propto \varpi$, by consideration of the force balance perpendicular to the tube axis show that the stability criterion is

$$k_z B_z > \frac{B_\phi}{\varpi}, \tag{9.43}$$

where $k_z = 2\pi/\lambda_z$ is the wavenumber of the instability. [Hint: first show what force is required to produce the growth rate ω_A and then equate this to the restoring Lorentz force from the axial field.]

(b) We can make the approximation that the stability criterion for a curved flux tube does not differ enormously from that in a straight tube. The field between the two sunspots is initially untwisted, i.e. $B_\phi = 0$ and then one of the spots slowly rotates. Calculate the energy required to twist one sunspot up to the instability threshold by consideration of the Lorentz force in a shallow layer in the sunspot where B_ϕ changes from to its value in the coronal loop. As the tube is twisted, it passes quasi-statically through a series of force-free equilibria; show that the α parameter in the force-free equation $\nabla \times \mathbf{B} = \alpha\mathbf{B}$ increases from 0 up to some value. When the instability threshold is passed, the field in the corona relaxes back to the lowest energy state, i.e. the curl-free field $\alpha = 0$; make an estimate of the energy released in the flare and equate this to the sunspot-rotating energy calculated earlier.

Appendix A

Useful information

A.1 Physical constants

speed of light	c	3×10^{10} cm s^{-1}
gravitational constant	G	$2/3 \times 10^{-7}$ cm^3 g^{-1} s^{-2}
Planck constant	h	$2/3 \times 10^{-26}$ erg s
	$\hbar = h/2\pi$	10^{-27} erg s
Boltzmann constant	$k_{\rm B}$	1.4×10^{-16} erg K^{-1}
Avogadro's number	$N_{\rm A}$	6×10^{23} mol^{-1}
gas constant	$R = k_{\rm B} N_{\rm A}$	8.31×10^8 erg mol^{-1} K^{-1}
Stefan-Boltzmann constant	$\sigma_{\rm SB} = \pi^2 k_{\rm B}^4/60\hbar^3 c^2$	5.67×10^{-5} erg cm^{-2} s^{-1} K^{-4}
radiation constant	$a = 4\sigma/c$	7.6×10^{-15} erg cm^{-3} K^{-4}
fine structure constant	$\alpha = e^2/\hbar c$	$1/137$
electron charge	e	4.8×10^{-10} esu
	e^2	1.44×10^{-7} eV cm
electron volt	eV	1.6×10^{-12} erg
electron mass	m_e	9×10^{-28} g
		511 keV
proton mass	$m_p \approx 1$ g mol$^{-1}/N_{\rm A}$	$5/3 \times 10^{-24}$ g
		938 MeV
proton/electron mass ratio	m_p/m_e	1836
proton-neutron mass difference	$m_n - m_p$	1.3 MeV
Rydberg constant	$R_\infty = \alpha m_e c/2h$	1.1×10^5 cm^{-1}
		13.6 eV
Bohr radius	$a_0 = \hbar^2/m_e e^2$	5.3×10^{-9} cm
classical electron radius	$r_0 = e^2/m_e c^2$	2.8×10^{-13} cm
Thompson cross section	$\sigma_{\rm T} = (8\pi/3) r_0^2$	$2/3 \times 10^{-24}$ cm^2
Compton electron wavelength	$h/m_e c$	2.4×10^{-10} cm
	$\hbar/m_e c$	3.9×10^{-11} cm

A.2 Astrophysical constants

Solar luminosity	$L_\odot$	4×10^{33} erg s^{-1}
Solar mass	$M_\odot$	2×10^{33} g
Solar radius	$R_\odot$	7×10^{10} cm
Jupiter mass	M_J	$10^{-3}\, M_\odot$
Jupiter radius	R_J	$0.1\, R_\odot$
Earth mass	$M_\oplus$	$3 \times 10^{-6}\, M_\odot$
Earth radius	$R_\oplus$	$0.009\, R_\odot$
astronomical unit	AU	1.5×10^{13} cm
parsec	pc	3×10^{18} cm
light year	ly	10^{18} cm
Hubble constant	H_0	≈ 71 km s^{-1} Mpc^{-1}
Eddington luminosity	$L_{\rm Edd} = 4\pi c G M m/\sigma_{\rm T}$	$3.3 \times 10^4\, (M/M_\odot)\, L_\odot$
Schwarzschild radius	$r_{\rm S} = 2GM/c^2$	$3\, (M/M_\odot)$ km

A.3 Vector identities and vector calculus identities

$$\mathbf{a} \cdot \mathbf{b} \times \mathbf{c} = \mathbf{c} \cdot \mathbf{a} \times \mathbf{b} = \mathbf{b} \cdot \mathbf{c} \times \mathbf{a} \tag{A.1}$$

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c} \tag{A.2}$$

$$\boldsymbol{\nabla} \times \boldsymbol{\nabla}\phi = \mathbf{0} \tag{A.3}$$

$$\boldsymbol{\nabla} \cdot \boldsymbol{\nabla} \times \mathbf{a} = 0 \tag{A.4}$$

$$\nabla^2 = \boldsymbol{\nabla} \cdot \boldsymbol{\nabla}, \quad \text{i.e.} \quad \nabla^2\phi = \boldsymbol{\nabla} \cdot \boldsymbol{\nabla}\phi \quad \text{and} \quad \nabla^2\mathbf{a} = (\nabla^2 a_x, \nabla^2 a_y, \nabla^2 a_z) \tag{A.5}$$

$$\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \mathbf{a} = \boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot \mathbf{a}) - \nabla^2\mathbf{a} \tag{A.6}$$

$$\boldsymbol{\nabla}(\mathbf{a} \cdot \mathbf{b}) = (\mathbf{a} \cdot \boldsymbol{\nabla})\mathbf{b} + (\mathbf{b} \cdot \boldsymbol{\nabla})\mathbf{a} + \mathbf{a} \times \boldsymbol{\nabla} \times \mathbf{b} + \mathbf{b} \times \boldsymbol{\nabla} \times \mathbf{a} \tag{A.7}$$

$$\frac{1}{2}\boldsymbol{\nabla} a^2 = (\mathbf{a} \cdot \boldsymbol{\nabla})\mathbf{a} + \mathbf{a} \times \boldsymbol{\nabla} \times \mathbf{a} \tag{A.8}$$

$$\boldsymbol{\nabla} \cdot (\mathbf{a} \times \mathbf{b}) = \mathbf{b} \cdot \boldsymbol{\nabla} \times \mathbf{a} - \mathbf{a} \cdot \boldsymbol{\nabla} \times \mathbf{b} \tag{A.9}$$

$$\boldsymbol{\nabla} \times (\mathbf{a} \times \mathbf{b}) = \mathbf{a}(\boldsymbol{\nabla} \cdot \mathbf{b}) - \mathbf{b}(\boldsymbol{\nabla} \cdot \mathbf{a}) + (\mathbf{b} \cdot \boldsymbol{\nabla})\mathbf{a} - (\mathbf{a} \cdot \boldsymbol{\nabla})\mathbf{b} \tag{A.10}$$

$$\boldsymbol{\nabla} \cdot (\phi\mathbf{a}) = \mathbf{a} \cdot \boldsymbol{\nabla}\phi + \phi\boldsymbol{\nabla} \cdot \mathbf{a} \tag{A.11}$$

$$\boldsymbol{\nabla} \times (\phi\mathbf{a}) = \phi\boldsymbol{\nabla} \times \mathbf{a} + \mathbf{a} \times \boldsymbol{\nabla}\phi \tag{A.12}$$

$$\boldsymbol{\nabla}(\psi\phi) = \psi\boldsymbol{\nabla}\phi + \phi\boldsymbol{\nabla}\psi \tag{A.13}$$

A.4 Symbols used in these notes

To avoid confusion I've tried to avoid symbol overlap.

Time	t	s
Density	ρ	g cm^{-3}
Pressure	P	erg cm^{-3}
Fluid velocity	$\mathbf{u}$	cm s^{-1}
components thereof	u, v, w	
Magnetic field	$\mathbf{B}$	gauss = $erg^{1/2}$ $cm^{-3/2}$
Dynamic viscosity	μ	g cm^{-1} s^{-1}
Kinetic diffusivity	ν	cm^2 s^{-1}
Thermal diffusivity	χ	cm^2 s^{-1}
Magnetic diffusivity	η	cm^2 s^{-1}
Charge density	ρ_e	esu cm^{-3} =$erg^{1/2}$ $cm^{-5/2}$
Current density	$\mathbf{J}$	esu cm^{-2} s^{-1} =$erg^{1/2}$ $cm^{-3/2}$ s^{-1}
Coordinates: (spherical) radius	r	
cylindrical radius	ϖ	
colatitude	θ	
azimuthal angle	ϕ	

A.5 A collection of useful results from thermodynamics

I list here a series of useful relations and concepts without going into the detail of their origin. First of all, the zeroth law is

> *There are three bodies A, B and C. If A and B are both separately in equilibrium with C, then A and B are in equilibrium with each other.*

From this we can define some property of a body, temperature. If two bodies have the same temperature then they are in thermal equilibrium with each other. Furthermore, we know from experience that the state of a given mass of fluid can be completely specified by a number of parameters and that all other parameters can be worked out from the equation of state. For instance, the state of a given mass of air is completely specified by its pressure P and its volume V. Note that pressure is an *intensive* variable as it can be measured at a particular point in space and does not depend on the size of the system, while volume is an *extensive* variable which obviously does depend on size. We could alternatively specify the state of air by pressure and specific volume $v = 1/\rho$ and then calculate other variables such as temperature from the equation of state. In the case of an ideal gas, the equation of state is

$$P = \rho RT \tag{A.14}$$

where R is a constant with units erg g^{-1} K^{-1} whose value is simply the universal gas constant divided by the mean molecular weight. In some fluids we might need more than two variables to completely describe the state of a fluid, for instance where the mean molecular weight is not uniform, or in some cases we might need only one variable, in which case we speak of a *barotropic* equation of state $\rho = \rho(P)$. The first law of thermodynamics is

> *If the state of an otherwise isolated system is changed by the performance of work, the amount of work needed depends solely on the change accomplished, and not on the means*

by which the work is performed, nor on the intermediate stages through which the system passes between its initial and final states.

which can alternatively be expressed in the simpler form

Energy is conserved if heat is taken into account.

From this it is possible to demonstrate the existence of a quantity called the internal energy U which is a function of state, i.e. it can be expressed as a function of the variables which describe the state of the system, for instance $U = U(P, V)$, and that changes in U are given by

$$\mathrm{d}U = \mathrm{d}Q + \mathrm{d}W \tag{A.15}$$

where $\mathrm{d}Q$ and $\mathrm{d}W$ are the heat added to and the work done on the system.

The second law states that

It is impossible to devise an engine which, working in a cycle, shall produce no effect other than the transfer of heat from a colder to a hotter body.

but note that there are several popular ways of expressing this law and that their equivalence is not always obvious at first glance. In a fluid where P and V are the only two independent variables

$$\mathrm{d}U = T\mathrm{d}S - P\mathrm{d}V, \tag{A.16}$$

where entropy S is a new function of state. There are more terms on the right-hand side for systems with extra degrees of freedom; a favourite of textbooks is magnetisable systems where $-P\mathrm{d}V$ is replaced or joined by $\mathbf{H} \cdot \mathrm{d}\mathbf{M}$. In addition, it is often useful to use a new function of state – *enthalpy*, defined thus:

$$H \equiv U + PV \qquad \text{so that} \qquad \mathrm{d}H = T\mathrm{d}S + V\mathrm{d}P, \tag{A.17}$$

whereby we know that it must be a function of state because it is defined as a function only of other functions of state.

For *reversible* changes, it is possible to equate the terms of (A.15) and (A.16) and write

$$\mathrm{d}Q = T\mathrm{d}S \qquad \text{and} \qquad \mathrm{d}W = -P\mathrm{d}V. \tag{A.18}$$

The heat capacities are defined as $\mathrm{d}Q/\mathrm{d}T$ in a reversible change under various conditions. From (A.16) and (A.18) we have

$$\mathrm{d}Q = \mathrm{d}U + P\mathrm{d}V, \tag{A.19}$$

from which we see that the heat capacites at constant volume C_V and constant pressure C_P are

$$C_V = \left(\frac{\mathrm{d}U}{\mathrm{d}T}\right)_V \qquad \text{and} \qquad C_P = \left(\frac{\mathrm{d}U}{\mathrm{d}T}\right)_P + P\left(\frac{\mathrm{d}V}{\mathrm{d}T}\right)_P. \tag{A.20}$$

It is fairly straightforward to prove that $C_P \geq C_V$, or in other words, that the ratio of the two $\gamma \equiv C_P/C_V \geq 1$. Furthermore it is possible to show that

$$\left(\frac{\mathrm{d}P}{\mathrm{d}V}\right)_S = \gamma\left(\frac{\mathrm{d}P}{\mathrm{d}V}\right)_T \tag{A.21}$$

which is useful because we often need to calculate the relation between pressure and volume in an adiabatic change, and it is usually simple to calculate the two quantities on the right-hand side. For instance, for an ideal gas γ is a constant, not depending on the state of the gas, and from the equation of state we have $(\mathrm{d}P/\mathrm{d}V)_T = P/V$, giving $PV^\gamma = \text{const}$ during an adiabatic change.

Extensive quantities such as V and C_P can be made intensive by dividing them by the mass of the system, giving in these cases the *specific volume* and *specific heat capacity*, meaning volume and heat capacity per unit mass. Of course, the specific volume is simply $1/\rho$.

Finally, we quote here the law of entropy increase

> *The entropy of an isolated system can never diminish.*