Plasma Physics

Richard Fitzpatrick

Professor of Physics

The University of Texas at Austin

Contents

1 Introduction

1.1 Sources

The major sources for this course are:

The Theory of Plasma Waves: T.H. Stix, 1st Ed. (McGraw-Hill, New York NY, 1962).

Plasma Physics: R.A. Cairns (Blackie, Glasgow UK, 1985).

The Framework of Plasma Physics: R.D. Hazeltine, and F.L. Waelbroeck (Westview, Boulder CO, 2004).

Other sources include:

The Mathematical Theory of Non-Uniform Gases: S. Chapman, and T.G. Cowling (Cambridge University Press, Cambridge UK, 1953).

Physics of Fully Ionized Gases: L. Spitzer, Jr., 1st Ed. (Interscience, New York NY, 1956).

Radio Waves in the Ionosphere: K.G. Budden (Cambridge University Press, Cambridge UK, 1961).

The Adiabatic Motion of Charged Particles: T.G. Northrop (Interscience, New York NY, 1963).

Coronal Expansion and the Solar Wind: A.J. Hundhausen (Springer-Verlag, Berlin, 1972).

Solar System Magnetic Fields: E.R. Priest, Ed. (D. Reidel Publishing Co., Dordrecht, Netherlands, 1985).

Lectures on Solar and Planetary Dynamos: M.R.E. Proctor, and A.D. Gilbert, Eds. (Cambridge University Press, Cambridge UK, 1994).

Introduction to Plasma Physics: R.J. Goldston, and P.H. Rutherford (Institute of Physics Publishing, Bristol UK, 1995).

Basic Space Plasma Physics: W. Baumjohann, and R. A. Treumann (Imperial College Press, London UK, 1996).

1.2 What is Plasma?

The electromagnetic force is generally observed to create *structure*: *e.g.*, stable atoms and molecules, crystalline solids. In fact, the most widely studied consequences of the electromagnetic force form the subject matter of Chemistry and Solid-State Physics, which are both disciplines developed to understand essentially static structures.

Structured systems have binding energies larger than the ambient thermal energy. Placed in a sufficiently hot environment, they decompose: *e.g.*, crystals melt, molecules disassociate. At temperatures near or exceeding atomic ionization energies, atoms similarly decompose into negatively charged electrons and positively charged ions. These charged particles are by no means free: in fact, they are strongly affected by each others' electromagnetic fields. Nevertheless, because the charges are no longer bound, their assemblage becomes capable of collective motions of great vigor and complexity. Such an assemblage is termed a *plasma*.

Of course, bound systems can display extreme complexity of structure: *e.g.*, a protein molecule. Complexity in a plasma is somewhat different, being expressed *temporally* as much as *spatially*. It is predominately characterized by the excitation of an enormous variety of *collective* dynamical modes.

Since thermal decomposition breaks interatomic bonds before ionizing, most terrestrial plasmas begin as gases. In fact, a plasma is sometimes defined as a gas that is sufficiently ionized to exhibit plasma-like behaviour. Note that plasma-like behaviour ensues after a remarkably small fraction of the gas has undergone ionization. Thus, fractionally ionized gases exhibit most of the exotic phenomena characteristic of fully ionized gases.

Plasmas resulting from ionization of neutral gases generally contain equal numbers of positive and negative charge carriers. In this situation, the oppositely charged fluids are strongly coupled, and tend to electrically neutralize one another on macroscopic length-scales. Such plasmas are termed *quasi-neutral* ("quasi" because the small deviations from exact neutrality have important dynamical consequences for certain types of plasma mode). Strongly *non-neutral* plasmas, which may even contain charges of only one sign, occur primarily in laboratory experiments: their equilibrium depends on the existence of intense magnetic fields, about which the charged fluid rotates.

It is sometimes remarked that 95% (or 99%, depending on whom you are trying to impress) of the baryonic content of the Universe consists of plasma. This statement has the double merit of being extremely flattering to Plasma Physics, and quite impossible to disprove (or verify). Nevertheless, it is worth pointing out the prevalence of the plasma state. In earlier epochs of the Universe, everything was plasma. In the present epoch, stars, nebulae, and even interstellar space, are filled with plasma. The Solar System is also permeated with plasma, in the form of the solar wind, and the Earth is completely surrounded by plasma trapped within its magnetic field.

Terrestrial plasmas are also not hard to find. They occur in lightning, fluorescent lamps, a variety of laboratory experiments, and a growing array of industrial processes. In fact, the glow discharge has recently become the mainstay of the micro-circuit fabrication industry. Liquid and even solid-state systems can occasionally display the collective electromagnetic effects that characterize plasma: *e.g.*, liquid mercury exhibits many dynamical modes, such as Alfvén waves, which occur in conventional plasmas.

1.3 Brief History of Plasma Physics

When blood is cleared of its various corpuscles there remains a transparent liquid, which was named *plasma* (after the Greek word $\pi\lambda\alpha\sigma\mu\alpha$, which means "moldable substance" or "jelly") by the great Czech medical scientist, Johannes Purkinje (1787-1869). The Nobel prize winning American chemist Irving Langmuir first used this term to describe an ionized gas in 1927—Langmuir was reminded of the way blood plasma carries red and white corpuscles by the way an electrified fluid carries electrons and ions. Langmuir, along with his colleague Lewi Tonks, was investigating the physics and chemistry of tungsten-filament light-bulbs, with a view to finding a way to greatly extend the lifetime of the filament (a goal which he eventually achieved). In the process, he developed the theory of *plasma sheaths*—the boundary layers which form between ionized plasmas and solid surfaces. He also discovered that certain regions of a plasma discharge tube exhibit periodic variations of the electron density, which we nowadays term *Langmuir waves*. This was the genesis of Plasma Physics. Interestingly enough, Langmuir's research nowadays forms the theoretical basis of most *plasma processing* techniques for fabricating integrated circuits. After Langmuir, plasma research gradually spread in other directions, of which *five* are particularly significant.

Firstly, the development of radio broadcasting led to the discovery of the Earth's *ionosphere*, a layer of partially ionized gas in the upper atmosphere which reflects radio waves, and is responsible for the fact that radio signals can be received when the transmitter is over the horizon. Unfortunately, the ionosphere also occasionally absorbs and distorts radio waves. For instance, the Earth's magnetic field causes waves with different polarizations (relative to the orientation of the magnetic field) to propagate at different velocities, an effect which can give rise to "ghost signals" (*i.e.*, signals which arrive a little before, or a little after, the main signal). In order to understand, and possibly correct, some of the deficiencies in radio communication, various scientists, such as E.V. Appleton and K.G. Budden, systematically developed the theory of electromagnetic wave propagation through non-uniform magnetized plasmas.

Secondly, astrophysicists quickly recognized that much of the Universe consists of plasma, and, thus, that a better understanding of astrophysical phenomena requires a better grasp of plasma physics. The pioneer in this field was Hannes Alfvén, who around 1940 developed the theory of *magnetohydrodyamics*, or MHD, in which plasma is treated essentially as a conducting fluid. This theory has been both widely and successfully employed to investigate sunspots, solar flares, the solar wind, star formation, and a host of other topics in astrophysics. Two topics of particular interest in MHD theory are *magnetic reconnection* and *dynamo theory*. Magnetic reconnection is a process by which magnetic field-lines suddenly change their topology: it can give rise to the sudden conversion of a great deal of magnetic energy into thermal energy, as well as the acceleration of some charged particles to extremely high energies, and is generally thought to be the basic mechanism behind solar flares. Dynamo theory studies how the motion of an MHD fluid can give rise to the generation of a macroscopic magnetic field. This process is important because both the terrestrial and solar magnetic fields would decay away comparatively rapidly (in astro-

physical terms) were they not maintained by dynamo action. The Earth's magnetic field is maintained by the motion of its molten core, which can be treated as an MHD fluid to a reasonable approximation.

Thirdly, the creation of the hydrogen bomb in 1952 generated a great deal of interest in *controlled thermonuclear fusion* as a possible power source for the future. At first, this research was carried out secretly, and independently, by the United States, the Soviet Union, and Great Britain. However, in 1958 thermonuclear fusion research was declassified, leading to the publication of a number of immensely important and influential papers in the late 1950's and the early 1960's. Broadly speaking, theoretical plasma physics first emerged as a mathematically rigorous discipline in these years. Not surprisingly, Fusion physicists are mostly concerned with understanding how a thermonuclear plasma can be trapped—in most cases by a magnetic field—and investigating the many plasma instabilities which may allow it to escape.

Fourthly, James A. Van Allen's discovery in 1958 of the Van Allen radiation belts surrounding the Earth, using data transmitted by the U.S. Explorer satellite, marked the start of the systematic exploration of the Earth's magnetosphere via satellite, and opened up the field of *space plasma physics*. Space scientists borrowed the theory of plasma trapping by a magnetic field from fusion research, the theory of plasma waves from ionospheric physics, and the notion of magnetic reconnection as a mechanism for energy release and particle acceleration from astrophysics.

Finally, the development of high powered lasers in the 1960's opened up the field of *laser plasma physics*. When a high powered laser beam strikes a solid target, material is immediately ablated, and a plasma forms at the boundary between the beam and the target. Laser plasmas tend to have fairly extreme properties (*e.g.*, densities characteristic of solids) not found in more conventional plasmas. A major application of laser plasma physics is the approach to fusion energy known as *inertial confinement fusion*. In this approach, tightly focused laser beams are used to implode a small solid target until the densities and temperatures characteristic of nuclear fusion (*i.e.*, the centre of a hydrogen bomb) are achieved. Another interesting application of laser plasma physics is the use of the extremely strong electric fields generated when a high intensity laser pulse passes through a plasma to accelerate particles. High-energy physicists hope to use plasma acceleration techniques to dramatically reduce the size and cost of particle accelerators.

1.4 Basic Parameters

Consider an idealized plasma consisting of an equal number of electrons, with mass m_e and charge $-e$ (here, e denotes the *magnitude* of the electron charge), and ions, with mass m_i and charge $+e$. We do not necessarily demand that the system has attained thermal equilibrium, but nevertheless use the symbol

$$T_s \equiv \frac{1}{3} m_s \langle v_s^2 \rangle \tag{1.1}$$

to denote a *kinetic temperature* measured in energy units (*i.e.*, joules). Here, v is a particle speed, and the angular brackets denote an ensemble average. The kinetic temperature of species s is essentially the average kinetic energy of particles of this species. In plasma physics, kinetic temperature is invariably measured in *electron-volts* (1 joule is equivalent to 6.24×10^{18} eV).

Quasi-neutrality demands that

$$n_i \simeq n_e \equiv n, \tag{1.2}$$

where n_s is the number density (*i.e.*, the number of particles per cubic meter) of species s.

Assuming that both ions and electrons are characterized by the same T (which is, by no means, always the case in plasmas), we can estimate typical particle speeds via the so-called *thermal speed*,

$$v_{ts} \equiv \sqrt{2\,T/m_s}. \tag{1.3}$$

Note that the ion thermal speed is usually far smaller than the electron thermal speed:

$$v_{ti} \sim \sqrt{m_e/m_i}\; v_{te}. \tag{1.4}$$

Of course, n and T are generally functions of position in a plasma.

1.5 Plasma Frequency

The *plasma frequency*,

$$\omega_p^{\,2} = \frac{n\,e^2}{\epsilon_0\,m}, \tag{1.5}$$

is the most fundamental time-scale in plasma physics. Clearly, there is a different plasma frequency for each species. However, the relatively fast electron frequency is, by far, the most important, and references to "the plasma frequency" in text-books invariably mean the *electron* plasma frequency.

It is easily seen that ω_p corresponds to the typical electrostatic oscillation frequency of a given species in response to a small charge separation. For instance, consider a one-dimensional situation in which a slab consisting entirely of one charge species is displaced from its quasi-neutral position by an infinitesimal distance δx. The resulting charge density which develops on the leading face of the slab is $\sigma = e\,n\,\delta x$. An equal and opposite charge density develops on the opposite face. The x-directed electric field generated inside the slab is of magnitude $E_x = -\sigma/\epsilon_0 = -e\,n\,\delta x/\epsilon_0$. Thus, Newton's law applied to an individual particle inside the slab yields

$$m\,\frac{d^2\delta x}{dt^2} = e\,E_x = -m\,\omega_p^{\,2}\,\delta x, \tag{1.6}$$

giving $\delta x = (\delta x)_0\,\cos(\omega_p\,t)$.

Note that plasma oscillations will only be observed if the plasma system is studied over time periods τ longer than the plasma period $\tau_p \equiv 1/\omega_p$, and if external actions change the

system at a rate no faster than ω_p. In the opposite case, one is clearly studying something other than plasma physics (*e.g.*, nuclear reactions), and the system cannot not usefully be considered to be a plasma. Likewise, observations over length-scales L shorter than the distance $v_t\,\tau_p$ traveled by a typical plasma particle during a plasma period will also not detect plasma behaviour. In this case, particles will exit the system before completing a plasma oscillation. This distance, which is the spatial equivalent to τ_p, is called the *Debye length*, and takes the form

$$\lambda_D \equiv \sqrt{T/m}\;\omega_p^{-1}. \tag{1.7}$$

Note that

$$\lambda_D = \sqrt{\frac{\epsilon_0\,T}{n\,e^2}} \tag{1.8}$$

is independent of mass, and therefore generally comparable for different species.

Clearly, our idealized system can only usefully be considered to be a plasma provided that

$$\frac{\lambda_D}{L} \ll 1, \tag{1.9}$$

and

$$\frac{\tau_p}{\tau} \ll 1. \tag{1.10}$$

Here, τ and L represent the typical time-scale and length-scale of the process under investigation.

It should be noted that, despite the conventional requirement (1.9), plasma physics *is* capable of considering structures on the Debye scale. The most important example of this is the *Debye sheath*: *i.e.*, the boundary layer which surrounds a plasma confined by a material surface.

1.6 Debye Shielding

Plasmas generally do not contain strong electric fields in their rest frames. The shielding of an external electric field from the interior of a plasma can be viewed as a result of high plasma conductivity: *i.e.*, plasma current generally flows freely enough to short out interior electric fields. However, it is more useful to consider the shielding as a *dielectric* phenomena: *i.e.*, it is the *polarization* of the plasma medium, and the associated redistribution of space charge, which prevents penetration by an external electric field. Not surprisingly, the length-scale associated with such shielding is the Debye length.

Let us consider the simplest possible example. Suppose that a quasi-neutral plasma is sufficiently close to thermal equilibrium that its particle densities are distributed according to the Maxwell-Boltzmann law,

$$n_s = n_0\,e^{-e_s\,\Phi/T}, \tag{1.11}$$

where $\Phi(\mathbf{r})$ is the electrostatic potential, and n_0 and T are constant. From $e_i = -e_e = e$, it is clear that quasi-neutrality requires the equilibrium potential to be a constant. Suppose

that this equilibrium potential is perturbed, by an amount $\delta\Phi$, by a small, localized charge density $\delta\rho_{\rm ext}$. The total perturbed charge density is written

$$\delta\rho = \delta\rho_{\rm ext} + e\,(\delta n_{\rm i} - \delta n_e) = \delta\rho_{\rm ext} - 2\,e^2\,n_0\,\delta\Phi/T. \tag{1.12}$$

Thus, Poisson's equation yields

$$\nabla^2\delta\Phi = -\frac{\delta\rho}{\epsilon_0} = -\left(\frac{\delta\rho_{\rm ext} - 2\,e^2\,n_0\,\delta\Phi/T}{\epsilon_0}\right), \tag{1.13}$$

which reduces to

$$\left(\nabla^2 - \frac{2}{\lambda_{\rm D}^2}\right)\delta\Phi = -\frac{\delta\rho_{\rm ext}}{\epsilon_0}. \tag{1.14}$$

If the perturbing charge density actually consists of a point charge q, located at the origin, so that $\delta\rho_{\rm ext} = q\,\delta(\mathbf{r})$, then the solution to the above equation is written

$$\delta\Phi(r) = \frac{q}{4\pi\epsilon_0\,r}\,e^{-\sqrt{2}\,r/\lambda_{\rm D}}. \tag{1.15}$$

Clearly, the Coulomb potential of the perturbing point charge q is shielded on distance scales longer than the Debye length by a *shielding cloud* of approximate radius $\lambda_{\rm D}$ consisting of charge of the opposite sign.

Note that the above argument, by treating n as a continuous function, implicitly assumes that there are many particles in the shielding cloud. Actually, Debye shielding remains statistically significant, and physical, in the opposite limit in which the cloud is barely populated. In the latter case, it is the *probability* of observing charged particles within a Debye length of the perturbing charge which is modified.

1.7 Plasma Parameter

Let us define the average distance between particles,

$$r_{\rm d} \equiv n^{-1/3}, \tag{1.16}$$

and the distance of closest approach,

$$r_{\rm c} \equiv \frac{e^2}{4\pi\epsilon_0\,T}. \tag{1.17}$$

Recall that $r_{\rm c}$ is the distance at which the Coulomb energy

$$U(r, v) = \frac{1}{2}\,m v^2 - \frac{e^2}{4\pi\epsilon_0\,r} \tag{1.18}$$

of one charged particle in the electrostatic field of another vanishes. Thus, $U(r_{\rm c}, v_{\rm t}) = 0$.

The significance of the ratio r_d/r_c is readily understood. When this ratio is small, charged particles are dominated by one another's electrostatic influence more or less continuously, and their kinetic energies are small compared to the interaction potential energies. Such plasmas are termed *strongly coupled*. On the other hand, when the ratio is large, strong electrostatic interactions between individual particles are occasional and relatively rare events. A typical particle is electrostatically influenced by all of the other particles within its Debye sphere, but this interaction very rarely causes any sudden change in its motion. Such plasmas are termed *weakly coupled*. It is possible to describe a weakly coupled plasma using a standard Fokker-Planck equation (*i.e.*, the same type of equation as is conventionally used to describe a neutral gas). Understanding the strongly coupled limit is far more difficult, and will not be attempted in this course. Actually, a strongly coupled plasma has more in common with a liquid than a conventional weakly coupled plasma.

Let us define the *plasma parameter*

$$\Lambda = 4\pi\, n\, \lambda_D^3. \tag{1.19}$$

This dimensionless parameter is obviously equal to the typical number of particles contained in a Debye sphere. However, Eqs. (1.8), (1.16), (1.17), and (1.19) can be combined to give

$$\Lambda = \frac{\lambda_D}{r_c} = \frac{1}{\sqrt{4\pi}}\left(\frac{r_d}{r_c}\right)^{3/2} = \frac{4\pi\,\epsilon_0^{\,3/2}}{e^3}\frac{T^{3/2}}{n^{1/2}}. \tag{1.20}$$

It can be seen that the case $\Lambda \ll 1$, in which the Debye sphere is sparsely populated, corresponds to a strongly coupled plasma. Likewise, the case $\Lambda \gg 1$, in which the Debye sphere is densely populated, corresponds to a weakly coupled plasma. It can also be appreciated, from Eq. (1.20), that strongly coupled plasmas tend to be cold and dense, whereas weakly coupled plasmas are diffuse and hot. Examples of strongly coupled plasmas include solid-density laser ablation plasmas, the very "cold" (*i.e.*, with kinetic temperatures similar to the ionization energy) plasmas found in "high pressure" arc discharges, and the plasmas which constitute the atmospheres of collapsed objects such as white dwarfs and neutron stars. On the other hand, the hot diffuse plasmas typically encountered in ionospheric physics, astrophysics, nuclear fusion, and space plasma physics are invariably weakly coupled. Table 1.1 lists the key parameters for some typical weakly coupled plasmas.

In conclusion, characteristic *collective* plasma behaviour is only observed on time-scales longer than the plasma period, and on length-scales larger than the Debye length. The statistical character of this behaviour is controlled by the plasma parameter. Although ω_p, λ_D, and Λ are the three most fundamental plasma parameters, there are a number of other parameters which are worth mentioning.

1.8 Collisions

Collisions between charged particles in a plasma differ fundamentally from those between molecules in a neutral gas because of the long range of the Coulomb force. In fact, it is

	$n(m^{-3})$	$T(eV)$	$\omega_p(sec^{-1})$	$\lambda_D(m)$	Λ
Interstellar	10^6	10^{-2}	6×10^4	0.7	4×10^6
Solar Chromosphere	10^{18}	2	6×10^{10}	5×10^{-6}	2×10^3
Solar Wind (1AU)	10^7	10	2×10^5	7	5×10^{10}
Ionosphere	10^{12}	0.1	6×10^7	2×10^{-3}	1×10^5
Arc discharge	10^{20}	1	6×10^{11}	7×10^{-7}	5×10^2
Tokamak	10^{20}	10^4	6×10^{11}	7×10^{-5}	4×10^8
Inertial Confinement	10^{28}	10^4	6×10^{15}	7×10^{-9}	5×10^4

Table 1.1: *Key parameters for some typical weakly coupled plasmas.*

clear from the discussion in Sect. 1.7 that *binary* collision processes can only be defined for weakly coupled plasmas. Note, however, that binary collisions in weakly coupled plasmas are still modified by collective effects—the many-particle process of Debye shielding enters in a crucial manner. Nevertheless, for large Λ we can speak of binary collisions, and therefore of a *collision frequency*, denoted by $\nu_{ss'}$. Here, $\nu_{ss'}$ measures the rate at which particles of species s are scattered by those of species s'. When specifying only a single subscript, one is generally referring to the *total* collision rate for that species, including impacts with all other species. Very roughly,

$$\nu_s \simeq \sum_{s'} \nu_{ss'}. \tag{1.21}$$

The species designations are generally important. For instance, the relatively small electron mass implies that, for unit ionic charge and comparable species temperatures,

$$\nu_e \sim \left(\frac{m_i}{m_e}\right)^{1/2} \nu_i. \tag{1.22}$$

Note that the collision frequency ν measures the frequency with which a particle trajectory undergoes a *major* angular change due to Coulomb interactions with other particles. Coulomb collisions are, in fact, predominately small angle scattering events, so the collision frequency is *not* the inverse of the typical time between collisions. Instead, it is the inverse of the typical time needed for enough collisions to occur that the particle trajectory is deviated through 90°. For this reason, the collision frequency is sometimes termed the "90° scattering rate."

It is conventional to define the *mean-free-path*,

$$\lambda_{mfp} \equiv v_t/\nu. \tag{1.23}$$

Clearly, the mean-free-path measures the typical distance a particle travels between "collisions" (*i.e.*, 90° scattering events). A collision-dominated, or *collisional*, plasma is simply one in which

$$\lambda_{mfp} \ll L, \tag{1.24}$$

where L is the observation length-scale. The opposite limit of large mean-free-path is said to correspond to a *collisionless* plasma. Collisions greatly simplify plasma behaviour by driving the system towards statistical equilibrium, characterized by Maxwell-Boltzmann distribution functions. Furthermore, short mean-free-paths generally ensure that plasma transport is *local* (*i.e.*, diffusive) in nature, which is a considerable simplification.

The typical magnitude of the collision frequency is

$$\nu \sim \frac{\ln \Lambda}{\Lambda}\,\omega_{\rm p}. \tag{1.25}$$

Note that $\nu \ll \omega_{\rm p}$ in a weakly coupled plasma. It follows that collisions do not seriously interfere with plasma oscillations in such systems. On the other hand, Eq. (1.25) implies that $\nu \gg \omega_{\rm p}$ in a strongly coupled plasma, suggesting that collisions effectively prevent plasma oscillations in such systems. This accords well with our basic picture of a strongly coupled plasma as a system dominated by Coulomb interactions which does not exhibit conventional plasma dynamics. It follows from Eqs. (1.5) and (1.20) that

$$\nu \sim \frac{e^4\,\ln\Lambda}{4\pi\epsilon_0^{\,2}\,m^{1/2}}\frac{n}{T^{3/2}}. \tag{1.26}$$

Thus, diffuse, high temperature plasmas tend to be collisionless, whereas dense, low temperature plasmas are more likely to be collisional.

Note that whilst collisions are crucial to the confinement and dynamics (*e.g.*, sound waves) of neutral gases, they play a far less important role in plasmas. In fact, in many plasmas the magnetic field effectively plays the role that collisions play in a neutral gas. In such plasmas, charged particles are constrained from moving perpendicular to the field by their small Larmor orbits, rather than by collisions. Confinement along the field-lines is more difficult to achieve, unless the field-lines form closed loops (or closed surfaces). Thus, it makes sense to talk about a "collisionless plasma," whereas it makes little sense to talk about a "collisionless neutral gas." Note that many plasmas are collisionless to a very good approximation, especially those encountered in astrophysics and space plasma physics contexts.

1.9 Magnetized Plasmas

A *magnetized* plasma is one in which the ambient magnetic field $\mathbf{B}$ is strong enough to significantly alter particle trajectories. In particular, magnetized plasmas are *anisotropic*, responding differently to forces which are parallel and perpendicular to the direction of $\mathbf{B}$. Note that a magnetized plasma moving with mean velocity $\mathbf{V}$ contains an electric field $\mathbf{E} = -\mathbf{V}\times\mathbf{B}$ which is *not affected* by Debye shielding. Of course, in the rest frame of the plasma the electric field is essentially zero.

As is well-known, charged particles respond to the Lorentz force,

$$\mathbf{F} = q\,\mathbf{v}\times\mathbf{B}, \tag{1.27}$$

by freely streaming in the direction of **B**, whilst executing circular Larmor orbits, or *gyro-orbits*, in the plane perpendicular to **B**. As the field-strength increases, the resulting helical orbits become more tightly wound, effectively tying particles to magnetic field-lines.

The typical Larmor radius, or *gyroradius*, of a charged particle gyrating in a magnetic field is given by

$$\rho \equiv \frac{v_t}{\Omega}, \tag{1.28}$$

where

$$\Omega = eB/m \tag{1.29}$$

is the cyclotron frequency, or *gyrofrequency*, associated with the gyration. As usual, there is a distinct gyroradius for each species. When species temperatures are comparable, the electron gyroradius is distinctly smaller than the ion gyroradius:

$$\rho_e \sim \left(\frac{m_e}{m_i}\right)^{1/2} \rho_i. \tag{1.30}$$

A plasma system, or process, is said to be *magnetized* if its characteristic length-scale L is large compared to the gyroradius. In the opposite limit, $\rho \gg L$, charged particles have essentially straight-line trajectories. Thus, the ability of the magnetic field to significantly affect particle trajectories is measured by the *magnetization parameter*

$$\delta \equiv \frac{\rho}{L}. \tag{1.31}$$

There are some cases of interest in which the electrons are magnetized, but the ions are not. However, a "magnetized" plasma conventionally refers to one in which both species are magnetized. This state is generally achieved when

$$\delta_i \equiv \frac{\rho_i}{L} \ll 1. \tag{1.32}$$

1.10 Plasma Beta

The fundamental measure of a magnetic field's effect on a plasma is the magnetization parameter δ. The fundamental measure of the inverse effect is called β, and is defined as the ratio of the thermal energy density $n\,T$ to the magnetic energy density $B^2/2\,\mu_0$. It is conventional to identify the plasma energy density with the pressure,

$$p \equiv n\,T, \tag{1.33}$$

as in an ideal gas, and to define a separate β_s for each plasma species. Thus,

$$\beta_s = \frac{2\,\mu_0\,p_s}{B^2}. \tag{1.34}$$

The total β is written

$$\beta = \sum_s \beta_s. \tag{1.35}$$

2 Charged Particle Motion

2.1 Introduction

All descriptions of plasma behaviour are based, ultimately, on the motions of the constituent particles. For the case of an unmagnetized plasma, the motions are fairly trivial, since the constituent particles move essentially in straight lines between collisions. The motions are also trivial in a magnetized plasma where the collision frequency ν greatly exceeds the gyrofrequency Ω: in this case, the particles are scattered after executing only a small fraction of a gyro-orbit, and, therefore, still move essentially in straight lines between collisions. The situation of primary interest in this section is that of a *collisionless* (*i.e.*, $\nu \ll \Omega$), *magnetized* plasma, where the gyroradius ρ is much smaller than the typical variation length-scale L of the **E** and **B** fields, and the gyroperiod Ω^{-1} is much less than the typical time-scale τ on which these fields change. In such a plasma, we expect the motion of the constituent particles to consist of a rapid gyration perpendicular to magnetic field-lines, combined with free-streaming parallel to the field-lines. We are particularly interested in calculating how this motion is affected by the spatial and temporal *gradients* in the **E** and **B** fields. In general, the motion of charged particles in spatially and temporally *non-uniform* electromagnetic fields is extremely complicated: however, we hope to considerably simplify this motion by exploiting the assumed smallness of the parameters ρ/L and $(\Omega\,\tau)^{-1}$. What we are really trying to understand, in this section, is how the *magnetic confinement* of an essentially collisionless plasma works at an individual particle level. Note that the type of collisionless, magnetized plasma considered in this section occurs primarily in magnetic fusion and space plasma physics contexts. In fact, in the following we shall be studying methods of analysis first developed by fusion physicists, and illustrating these methods primarily by investigating problems of interest in magnetospheric physics.

2.2 Motion in Uniform Fields

Let us, first of all, consider the motion of charged particles in spatially and temporally *uniform* electromagnetic fields. The equation of motion of an individual particle takes the form

$$m\,\frac{d\mathbf{v}}{dt} = e\,(\mathbf{E} + \mathbf{v}\times\mathbf{B}). \tag{2.1}$$

The component of this equation parallel to the magnetic field,

$$\frac{dv_\parallel}{dt} = \frac{e}{m}\,E_\parallel, \tag{2.2}$$

predicts uniform acceleration along magnetic field-lines. Consequently, plasmas near equilibrium generally have either small or vanishing $E_\parallel$.

As can easily be verified by substitution, the perpendicular component of Eq. (2.1) yields

$$\mathbf{v}_\perp = \frac{\mathbf{E}\times\mathbf{B}}{B^2} + \rho\,\Omega\,[\mathbf{e}_1\,\sin(\Omega\,t+\gamma_0) + \mathbf{e}_2\,\cos(\Omega\,t+\gamma_0)]\,, \tag{2.3}$$

where $\Omega = eB/m$ is the gyrofrequency, ρ is the gyroradius, $\mathbf{e}_1$ and $\mathbf{e}_2$ are unit vectors such that ($\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{B}$) form a right-handed, mutually orthogonal set, and γ_0 is the initial gyrophase of the particle. The motion consists of gyration around the magnetic field at frequency Ω, superimposed on a steady drift at velocity

$$\mathbf{v}_E = \frac{\mathbf{E}\times\mathbf{B}}{B^2}. \tag{2.4}$$

This drift, which is termed the *E-cross-B drift* by plasma physicists, is *identical* for all plasma species, and can be eliminated entirely by transforming to a new inertial frame in which $\mathbf{E}_\perp = \mathbf{0}$. This frame, which moves with velocity $\mathbf{v}_E$ with respect to the old frame, can properly be regarded as the *rest frame* of the plasma.

We complete the solution by integrating the velocity to find the particle position:

$$\mathbf{r}(t) = \mathbf{R}(t) + \boldsymbol{\rho}(t), \tag{2.5}$$

where

$$\boldsymbol{\rho}(t) = \rho\,[-\mathbf{e}_1\,\cos(\Omega\,t+\gamma_0) + \mathbf{e}_2\,\sin(\Omega\,t+\gamma_0)], \tag{2.6}$$

and

$$\mathbf{R}(t) = \left(v_{0\parallel}\,t + \frac{e}{m}\,E_\parallel\,\frac{t^2}{2}\right)\mathbf{b} + \mathbf{v}_E\,t. \tag{2.7}$$

Here, $\mathbf{b} \equiv \mathbf{B}/B$. Of course, the trajectory of the particle describes a *spiral*. The gyrocentre $\mathbf{R}$ of this spiral, termed the *guiding centre* by plasma physicists, drifts across the magnetic field with velocity $\mathbf{v}_E$, and also accelerates along the field at a rate determined by the parallel electric field.

The concept of a guiding centre gives us a clue as to how to proceed. Perhaps, when analyzing charged particle motion in *non-uniform* electromagnetic fields, we can somehow neglect the rapid, and relatively uninteresting, gyromotion, and focus, instead, on the far slower motion of the guiding centre? Clearly, what we need to do in order to achieve this goal is to somehow *average* the equation of motion over gyrophase, so as to obtain a reduced equation of motion for the guiding centre.

2.3 Method of Averaging

In many dynamical problems, the motion consists of a rapid oscillation superimposed on a slow secular drift. For such problems, the most efficient approach is to describe the evolution in terms of the average values of the dynamical variables. The method outlined below is adapted from a classic paper by Morozov and Solov'ev.[1]

[1] A.I. Morozov, and L.S. Solev'ev, *Motion of Charged Particles in Electromagnetic Fields*, in *Reviews of Plasma Physics*, Vol. 2 (Consultants Bureau, New York NY, 1966).

Consider the equation of motion

$$\frac{d\mathbf{z}}{dt} = \mathbf{f}(\mathbf{z}, t, \tau), \tag{2.8}$$

where $\mathbf{f}$ is a periodic function of its last argument, with period 2π, and

$$\tau = t/\epsilon. \tag{2.9}$$

Here, the small parameter ϵ characterizes the separation between the short oscillation period τ and the time-scale t for the slow secular evolution of the "position" $\mathbf{z}$.

The basic idea of the averaging method is to treat t and τ as *distinct* independent variables, and to look for solutions of the form $\mathbf{z}(t, \tau)$ which are *periodic* in τ. Thus, we replace Eq. (2.8) by

$$\frac{\partial \mathbf{z}}{\partial t} + \frac{1}{\epsilon}\frac{\partial \mathbf{z}}{\partial \tau} = \mathbf{f}(\mathbf{z}, t, \tau), \tag{2.10}$$

and reserve Eq. (2.9) for substitution in the final result. The indeterminacy introduced by increasing the number of variables is lifted by the requirement of periodicity in τ. All of the secular drifts are thereby attributed to the t-variable, whilst the oscillations are described entirely by the τ-variable.

Let us denote the τ-average of $\mathbf{z}$ by $\mathbf{Z}$, and seek a change of variables of the form

$$\mathbf{z}(t, \tau) = \mathbf{Z}(t) + \epsilon\, \zeta(\mathbf{Z}, t, \tau). \tag{2.11}$$

Here, ζ is a periodic function of τ with vanishing mean. Thus,

$$\langle \zeta(\mathbf{Z}, t, \tau)\rangle \equiv \frac{1}{2\pi}\oint \zeta(\mathbf{Z}, t, \tau)\, d\tau = 0, \tag{2.12}$$

where $\oint$ denotes the integral over a full period in τ.

The evolution of $\mathbf{Z}$ is determined by substituting the expansions

$$\zeta = \zeta_0(\mathbf{Z}, t, \tau) + \epsilon\, \zeta_1(\mathbf{Z}, t, \tau) + \epsilon^2\, \zeta_2(\mathbf{Z}, t, \tau) + \cdots, \tag{2.13}$$

$$\frac{d\mathbf{Z}}{dt} = \mathbf{F}_0(\mathbf{Z}, t) + \epsilon\, \mathbf{F}_1(\mathbf{Z}, t) + \epsilon^2\, \mathbf{F}_2(\mathbf{Z}, t) + \cdots, \tag{2.14}$$

into the equation of motion (2.10), and solving order by order in ϵ.

To lowest order, we obtain

$$\mathbf{F}_0(\mathbf{Z}, t) + \frac{\partial \zeta_0}{\partial \tau} = \mathbf{f}(\mathbf{Z}, t, \tau). \tag{2.15}$$

The solubility condition for this equation is

$$\mathbf{F}_0(\mathbf{Z}, t) = \langle \mathbf{f}(\mathbf{Z}, t, \tau)\rangle. \tag{2.16}$$

Integrating the oscillating component of Eq. (2.15) yields

$$\zeta_0(\mathbf{Z}, t, \tau) = \int_0^\tau (\mathbf{f} - \langle \mathbf{f} \rangle)\ d\tau'. \tag{2.17}$$

To first order, we obtain

$$\mathbf{F}_1 + \frac{\partial \zeta_0}{\partial t} + \mathbf{F}_0 \cdot \nabla \zeta_0 + \frac{\partial \zeta_1}{\partial \tau} = \zeta_0 \cdot \nabla \mathbf{f}. \tag{2.18}$$

The solubility condition for this equation yields

$$\mathbf{F}_1 = \langle \zeta_0 \cdot \nabla \mathbf{f} \rangle. \tag{2.19}$$

The final result is obtained by combining Eqs. (2.16) and (2.19):

$$\frac{d\mathbf{Z}}{dt} = \langle \mathbf{f} \rangle + \epsilon \langle \zeta_0 \cdot \nabla \mathbf{f} \rangle + O(\epsilon^2). \tag{2.20}$$

Note that $\mathbf{f} = \mathbf{f}(\mathbf{Z}, t)$ in the above equation. Evidently, the secular motion of the "guiding centre" position $\mathbf{Z}$ is determined to lowest order by the average of the "force" $\mathbf{f}$, and to next order by the correlation between the oscillation in the "position" $\mathbf{z}$ and the oscillation in the spatial gradient of the "force."

2.4 Guiding Centre Motion

Consider the motion of a charged particle in the limit in which the electromagnetic fields experienced by the particle do not vary much in a gyroperiod: *i.e.*,

$$\rho\, |\nabla \mathbf{B}| \ll B, \tag{2.21}$$

$$\frac{1}{\Omega} \frac{\partial B}{\partial t} \ll B. \tag{2.22}$$

The electric force is assumed to be comparable to the magnetic force. To keep track of the order of the various quantities, we introduce the parameter ϵ as a book-keeping device, and make the substitution $\rho \rightarrow \epsilon\, \rho$, as well as $(\mathbf{E}, \mathbf{B}, \Omega) \rightarrow \epsilon^{-1}(\mathbf{E}, \mathbf{B}, \Omega)$. The parameter ϵ is set to unity in the final answer.

In order to make use of the technique described in the previous section, we write the dynamical equations in first-order differential form,

$$\frac{d\mathbf{r}}{dt} = \mathbf{v}, \tag{2.23}$$

$$\frac{d\mathbf{v}}{dt} = \frac{e}{\epsilon\, m} (\mathbf{E} + \mathbf{v} \times \mathbf{B}), \tag{2.24}$$

and seek a change of variables,

$$\mathbf{r} = \mathbf{R} + \epsilon\, \boldsymbol{\rho}(\mathbf{R}, \mathbf{U}, t, \gamma), \tag{2.25}$$

$$\mathbf{v} = \mathbf{U} + \mathbf{u}(\mathbf{R}, \mathbf{U}, t, \gamma), \tag{2.26}$$

such that the new guiding centre variables $\mathbf{R}$ and $\mathbf{U}$ are free of oscillations along the particle trajectory. Here, γ is a new independent variable describing the phase of the gyrating particle. The functions $\boldsymbol{\rho}$ and $\mathbf{u}$ represent the gyration radius and velocity, respectively. We require periodicity of these functions with respect to their last argument, with period 2π, and with vanishing mean:

$$\langle \boldsymbol{\rho} \rangle = \langle \mathbf{u} \rangle = \mathbf{0}. \tag{2.27}$$

Here, the angular brackets refer to the average over a period in γ.

The equation of motion is used to determine the coefficients in the expansion of $\boldsymbol{\rho}$ and $\mathbf{u}$:

$$\boldsymbol{\rho} = \boldsymbol{\rho}_0(\mathbf{R}, \mathbf{U}, t, \gamma) + \epsilon\, \boldsymbol{\rho}_1(\mathbf{R}, \mathbf{U}, t, \gamma) + \cdots, \tag{2.28}$$

$$\mathbf{u} = \mathbf{u}_0(\mathbf{R}, \mathbf{U}, t, \gamma) + \epsilon\, \mathbf{u}_1(\mathbf{R}, \mathbf{U}, t, \gamma) + \cdots. \tag{2.29}$$

The dynamical equation for the gyrophase is likewise expanded, assuming that $\mathrm{d}\gamma/\mathrm{d}t \simeq \Omega = \mathrm{O}(\epsilon^{-1})$,

$$\frac{\mathrm{d}\gamma}{\mathrm{d}t} = \epsilon^{-1}\, \omega_{-1}(\mathbf{R}, \mathbf{U}, t) + \omega_0(\mathbf{R}, \mathbf{U}, t) + \cdots. \tag{2.30}$$

In the following, we suppress the subscripts on all quantities except the guiding centre velocity $\mathbf{U}$, since this is the only quantity for which the first-order corrections are calculated.

To each order in ϵ, the evolution of the guiding centre position $\mathbf{R}$ and velocity $\mathbf{U}$ are determined by the solubility conditions for the equations of motion (2.23)–(2.24) when expanded to that order. The oscillating components of the equations of motion determine the evolution of the gyrophase. Note that the velocity equation (2.23) is *linear*. It follows that, to all orders in ϵ, its solubility condition is simply

$$\frac{\mathrm{d}\mathbf{R}}{\mathrm{d}t} = \mathbf{U}. \tag{2.31}$$

To lowest order [*i.e.*, $\mathrm{O}(\epsilon^{-1})$], the momentum equation (2.24) yields

$$\omega\, \frac{\partial \mathbf{u}}{\partial \gamma} - \Omega\, \mathbf{u} \times \mathbf{b} = \frac{e}{m}\, (\mathbf{E} + \mathbf{U}_0 \times \mathbf{B})\,. \tag{2.32}$$

The solubility condition (*i.e.*, the gyrophase average) is

$$\mathbf{E} + \mathbf{U}_0 \times \mathbf{B} = \mathbf{0}. \tag{2.33}$$

This immediately implies that

$$E_{\parallel} \equiv \mathbf{E} \cdot \mathbf{b} \sim \epsilon\, E. \tag{2.34}$$

Clearly, the rapid acceleration caused by a large parallel electric field would invalidate the ordering assumptions used in this calculation. Solving for $\mathbf{U}_0$, we obtain

$$\mathbf{U}_0 = U_{0\parallel}\,\mathbf{b} + \mathbf{v}_E, \tag{2.35}$$

where all quantities are evaluated at the guiding-centre position $\mathbf{R}$. The perpendicular component of the velocity, $\mathbf{v}_E$, has the same form (2.4) as for uniform fields. Note that the parallel velocity is undetermined at this order.

The integral of the oscillating component of Eq. (2.32) yields

$$\mathbf{u} = \mathbf{c} + u_\perp\,[\mathbf{e}_1\,\sin(\Omega\,\gamma/\omega) + \mathbf{e}_2\,\cos(\Omega\,\gamma/\omega)]\,, \tag{2.36}$$

where $\mathbf{c}$ is a constant vector, and $\mathbf{e}_1$ and $\mathbf{e}_2$ are again mutually orthogonal unit vectors perpendicular to $\mathbf{b}$. All quantities in the above equation are functions of $\mathbf{R}$, $\mathbf{U}$, and t. The periodicity constraint, plus Eq. (2.27), require that $\omega = \Omega(\mathbf{R}, t)$ and $\mathbf{c} = \mathbf{0}$. The gyration velocity is thus

$$\mathbf{u} = u_\perp\,(\mathbf{e}_1\,\sin\gamma + \mathbf{e}_2\,\cos\gamma)\,, \tag{2.37}$$

and the gyrophase is given by

$$\gamma = \gamma_0 + \Omega\,t, \tag{2.38}$$

where γ_0 is the initial phase. Note that the amplitude $u_\perp$ of the gyration velocity is undetermined at this order.

The lowest order oscillating component of the velocity equation (2.23) yields

$$\Omega\,\frac{\partial\boldsymbol{\rho}}{\partial\gamma} = \mathbf{u}. \tag{2.39}$$

This is easily integrated to give

$$\boldsymbol{\rho} = \rho\,(-\mathbf{e}_1\,\cos\gamma + \mathbf{e}_2\,\sin\gamma), \tag{2.40}$$

where $\rho = u_\perp/\Omega$. It follows that

$$\mathbf{u} = \Omega\,\boldsymbol{\rho}\times\mathbf{b}. \tag{2.41}$$

The gyrophase average of the first-order [*i.e.*, $O(\epsilon^0)$] momentum equation (2.24) reduces to

$$\frac{d\mathbf{U}_0}{dt} = \frac{e}{m}\,\left[E_\parallel\,\mathbf{b} + \mathbf{U}_1\times\mathbf{B} + \langle\mathbf{u}\times(\boldsymbol{\rho}\cdot\nabla)\,\mathbf{B}\rangle\right]. \tag{2.42}$$

Note that all quantities in the above equation are functions of the guiding centre position $\mathbf{R}$, rather than the instantaneous particle position $\mathbf{r}$. In order to evaluate the last term, we make the substitution $\mathbf{u} = \Omega\,\boldsymbol{\rho}\times\mathbf{b}$ and calculate

$$\begin{aligned}\langle(\boldsymbol{\rho}\times\mathbf{b})\times(\boldsymbol{\rho}\cdot\nabla)\,\mathbf{B}\rangle &= \mathbf{b}\,\langle\boldsymbol{\rho}\cdot(\boldsymbol{\rho}\cdot\nabla)\,\mathbf{B}\rangle - \langle\boldsymbol{\rho}\,\mathbf{b}\cdot(\boldsymbol{\rho}\cdot\nabla)\,\mathbf{B}\rangle\\ &= \mathbf{b}\,\langle\boldsymbol{\rho}\cdot(\boldsymbol{\rho}\cdot\nabla)\,\mathbf{B}\rangle - \langle\boldsymbol{\rho}\,(\boldsymbol{\rho}\cdot\nabla B)\rangle.\end{aligned} \tag{2.43}$$

The averages are specified by

$$\langle \boldsymbol{\rho}\,\boldsymbol{\rho}\rangle = \frac{u_\perp^{\,2}}{2\,\Omega^2}\,(\mathbf{I}-\mathbf{b}\mathbf{b}), \tag{2.44}$$

where $\mathbf{I}$ is the identity tensor. Thus, making use of $\mathbf{I}\!:\!\nabla\mathbf{B} = \nabla\cdot\mathbf{B} = 0$, it follows that

$$-e\,\langle \mathbf{u}\times(\boldsymbol{\rho}\cdot\nabla)\,\mathbf{B}\rangle = \frac{m\,u_\perp^{\,2}}{2\,B}\,\nabla B. \tag{2.45}$$

This quantity is the secular component of the gyration induced fluctuations in the magnetic force acting on the particle.

The coefficient of ∇B in the above equation,

$$\mu = \frac{m\,u_\perp^{\,2}}{2\,B}, \tag{2.46}$$

plays a central role in the theory of magnetized particle motion. We can interpret this coefficient as a *magnetic moment* by drawing an analogy between a gyrating particle and a current loop. The (vector) magnetic moment of a current loop is simply

$$\boldsymbol{\mu} = I\,A\,\mathbf{n}, \tag{2.47}$$

where I is the current, A the area of the loop, and $\mathbf{n}$ the unit normal to the surface of the loop. For a circular loop of radius $\rho = u_\perp/\Omega$, lying in the plane perpendicular to $\mathbf{b}$, and carrying the current $e\,\Omega/2\pi$, we find

$$\boldsymbol{\mu} = I\,\pi\,\rho^2\,\mathbf{b} = \frac{m\,u_\perp^{\,2}}{2\,B}\,\mathbf{b}. \tag{2.48}$$

We shall demonstrate later on that the (scalar) magnetic moment μ is a *constant* of the particle motion. Thus, the guiding centre behaves exactly like a particle with a conserved magnetic moment μ which is always aligned with the magnetic field.

The first-order guiding centre equation of motion reduces to

$$m\,\frac{d\mathbf{U}_0}{dt} = e\,E_\parallel\,\mathbf{b} + e\,\mathbf{U}_1\times\mathbf{B} - \mu\,\nabla B. \tag{2.49}$$

The component of this equation along the magnetic field determines the evolution of the parallel guiding centre velocity:

$$m\,\frac{dU_{0\parallel}}{dt} = e\,E_\parallel - \boldsymbol{\mu}\cdot\nabla B - m\,\mathbf{b}\cdot\frac{d\mathbf{v}_E}{dt}. \tag{2.50}$$

Here, use has been made of Eq. (2.35) and $\mathbf{b}\cdot d\mathbf{b}/dt = 0$. The component of Eq. (2.49) perpendicular to the magnetic field determines the first-order perpendicular drift velocity:

$$\mathbf{U}_{1\perp} = \frac{\mathbf{b}}{\Omega}\times\left[\frac{d\mathbf{U}_0}{dt} + \frac{\mu}{m}\,\nabla B\right]. \tag{2.51}$$

Note that the first-order correction to the parallel velocity, the parallel drift velocity, is undetermined to this order. This is not generally a problem, since the first-order parallel drift is a small correction to a type of motion which already exists at zeroth-order, whereas the first-order perpendicular drift is a completely new type of motion. In particular, the first-order perpendicular drift differs fundamentally from the $\mathbf{E}\times\mathbf{B}$ drift, since it is not the same for different species, and, therefore, cannot be eliminated by transforming to a new inertial frame.

We can now understand the motion of a charged particle as it moves through slowly varying electric and magnetic fields. The particle always gyrates around the magnetic field at the local gyrofrequency $\Omega = eB/m$. The local perpendicular gyration velocity $u_\perp$ is determined by the requirement that the magnetic moment $\mu = m\,u_\perp^2/2\,B$ be a constant of the motion. This, in turn, fixes the local gyroradius $\rho = u_\perp/\Omega$. The parallel velocity of the particle is determined by Eq. (2.50). Finally, the perpendicular drift velocity is the sum of the $\mathbf{E}\times\mathbf{B}$ drift velocity $\mathbf{v}_E$ and the first-order drift velocity $\mathbf{U}_{1\perp}$.

2.5 Magnetic Drifts

Equations (2.35) and (2.51) can be combined to give

$$\mathbf{U}_{1\perp} = \frac{\mu}{m\,\Omega}\,\mathbf{b}\times\nabla B + \frac{U_{0\parallel}}{\Omega}\,\mathbf{b}\times\frac{d\mathbf{b}}{dt} + \frac{\mathbf{b}}{\Omega}\times\frac{d\mathbf{v}_E}{dt}. \tag{2.52}$$

The three terms on the right-hand side of the above expression are conventionally called the *magnetic*, or *grad-B*, *drift*, the *inertial drift*, and the *polarization drift*, respectively.

The magnetic drift,

$$\mathbf{U}_{\rm mag} = \frac{\mu}{m\,\Omega}\,\mathbf{b}\times\nabla B, \tag{2.53}$$

is caused by the slight variation of the gyroradius with gyrophase as a charged particle rotates in a non-uniform magnetic field. The gyroradius is reduced on the high-field side of the Larmor orbit, whereas it is increased on the low-field side. The net result is that the orbit does not quite close. In fact, the motion consists of the conventional gyration around the magnetic field combined with a slow drift which is perpendicular to both the local direction of the magnetic field and the local gradient of the field-strength.

Given that

$$\frac{d\mathbf{b}}{dt} = \frac{\partial\mathbf{b}}{\partial t} + (\mathbf{v}_E\cdot\nabla)\,\mathbf{b} + U_{0\parallel}\,(\mathbf{b}\cdot\nabla)\,\mathbf{b}, \tag{2.54}$$

the inertial drift can be written

$$\mathbf{U}_{\rm int} = \frac{U_{0\parallel}}{\Omega}\,\mathbf{b}\times\left[\frac{\partial\mathbf{b}}{\partial t} + (\mathbf{v}_E\cdot\nabla)\,\mathbf{b}\right] + \frac{U_{0\parallel}^2}{\Omega}\,\mathbf{b}\times(\mathbf{b}\cdot\nabla)\,\mathbf{b}. \tag{2.55}$$

In the important limit of stationary magnetic fields and weak electric fields, the above expression is dominated by the final term,

$$\mathbf{U}_{\rm curv} = \frac{U_{0\parallel}^2}{\Omega}\,\mathbf{b}\times(\mathbf{b}\cdot\nabla)\,\mathbf{b}, \tag{2.56}$$

which is called the *curvature drift*. As is easily demonstrated, the quantity $(\mathbf{b}\cdot\nabla)\,\mathbf{b}$ is a vector whose direction is towards the centre of the circle which most closely approximates the magnetic field-line at a given point, and whose magnitude is the inverse of the radius of this circle. Thus, the centripetal acceleration imposed by the curvature of the magnetic field on a charged particle following a field-line gives rise to a slow drift which is perpendicular to both the local direction of the magnetic field and the direction to the local centre of curvature of the field.

The polarization drift,

$$\mathbf{U}_{\rm polz} = \frac{\mathbf{b}}{\Omega}\times\frac{d\mathbf{v}_E}{dt}, \tag{2.57}$$

reduces to

$$\mathbf{U}_{\rm polz} = \frac{1}{\Omega}\frac{d}{dt}\left(\frac{\mathbf{E}_\perp}{B}\right) \tag{2.58}$$

in the limit in which the magnetic field is stationary but the electric field varies in time. This expression can be understood as a polarization drift by considering what happens when we suddenly impose an electric field on a particle at rest. The particle initially accelerates in the direction of the electric field, but is then deflected by the magnetic force. Thereafter, the particle undergoes conventional gyromotion combined with $\mathbf{E}\times\mathbf{B}$ drift. The time between the switch-on of the field and the magnetic deflection is approximately $\Delta t\sim\Omega^{-1}$. Note that there is no deflection if the electric field is directed parallel to the magnetic field, so this argument only applies to perpendicular electric fields. The initial displacement of the particle in the direction of the field is of order

$$\delta\sim\frac{e\,\mathbf{E}_\perp}{m}\,(\Delta t)^2\sim\frac{\mathbf{E}_\perp}{\Omega\,B}. \tag{2.59}$$

Note that, because $\Omega\propto m^{-1}$, the displacement of the ions *greatly exceeds* that of the electrons. Thus, when an electric field is suddenly switched on in a plasma, there is an initial polarization of the plasma medium caused, predominately, by a displacement of the ions in the direction of the field. If the electric field, in fact, varies *continuously* in time, then there is a slow drift due to the constantly changing polarization of the plasma medium. This drift is essentially the time derivative of Eq. (2.59) [*i.e.*, Eq. (2.58)].

2.6 Invariance of Magnetic Moment

Let us now demonstrate that the magnetic moment $\mu = m\,u_\perp^2/2\,B$ is indeed a constant of the motion, at least to lowest order. The scalar product of the equation of motion (2.24) with the velocity $\mathbf{v}$ yields

$$\frac{m}{2}\frac{dv^2}{dt} = e\,\mathbf{v}\cdot\mathbf{E}. \tag{2.60}$$

This equation governs the evolution of the particle energy during its motion. Let us make the substitution $\mathbf{v}=\mathbf{U}+\mathbf{u}$, as before, and then average the above equation over gyrophase.

To lowest order, we obtain

$$\frac{m}{2}\frac{d}{dt}(u_\perp^2 + U_0^2) = e\,U_{0\parallel}\,E_\parallel + e\,\mathbf{U}_1\cdot\mathbf{E} + e\,\langle\mathbf{u}\cdot(\boldsymbol{\rho}\cdot\nabla)\,\mathbf{E}\rangle. \tag{2.61}$$

Here, use has been made of the result

$$\frac{d}{dt}\langle f\rangle = \langle\frac{df}{dt}\rangle, \tag{2.62}$$

which is valid for any f. The final term on the right-hand side of Eq. (2.61) can be written

$$e\,\Omega\,\langle(\boldsymbol{\rho}\times\mathbf{b})\cdot(\boldsymbol{\rho}\cdot\nabla)\,\mathbf{E}\rangle = -\mu\,\mathbf{b}\cdot\nabla\times\mathbf{E} = \boldsymbol{\mu}\cdot\frac{\partial\mathbf{B}}{\partial t}. \tag{2.63}$$

Thus, Eq. (2.61) reduces to

$$\frac{dK}{dt} = e\,\mathbf{U}\cdot\mathbf{E} + \boldsymbol{\mu}\cdot\frac{\partial\mathbf{B}}{\partial t} = e\,\mathbf{U}\cdot\mathbf{E} + \mu\,\frac{\partial B}{\partial t}. \tag{2.64}$$

Here, $\mathbf{U}$ is the guiding centre velocity, evaluated to first order, and

$$K = \frac{m}{2}(U_{0\parallel}^2 + \mathbf{v}_E^2 + u_\perp^2) \tag{2.65}$$

is the kinetic energy of the particle. Evidently, the kinetic energy can change in one of two ways. Either by motion of the guiding centre along the direction of the electric field, or by the acceleration of the gyration due to the electromotive force generated around the Larmor orbit by a changing magnetic field.

Equations (2.35), (2.50), and (2.51) can be used to eliminate $U_{0\parallel}$ and $\mathbf{U}_1$ from Eq. (2.64). The final result is

$$\frac{d}{dt}\left(\frac{m\,u_\perp^2}{2\,B}\right) = \frac{d\mu}{dt} = 0. \tag{2.66}$$

Thus, the magnetic moment μ is a constant of the motion to lowest order. Kruskal[2] has shown that $m\,u_\perp^2/2\,B$ is the lowest order approximation to a quantity which is a constant of the motion to *all* orders in the perturbation expansion. Such a quantity is called an *adiabatic invariant*.

2.7 Poincaré Invariants

An adiabatic invariant is an approximation to a more fundamental type of invariant known as a *Poincaré invariant*. A Poincaré invariant takes the form

$$\mathcal{I} = \oint_{C(t)} \mathbf{p}\cdot d\mathbf{q}, \tag{2.67}$$

[2]M. Kruskal, J. Math. Phys. **3**, 806 (1962).

where all points on the closed curve $C(t)$ in phase-space move according to the equations of motion.

In order to demonstrate that $\mathcal{I}$ is a constant of the motion, we introduce a periodic variable s parameterizing the points on the curve C. The coordinates of a general point on C are thus written $q_i = q_i(s,t)$ and $p_i = p_i(s,t)$. The rate of change of $\mathcal{I}$ is then

$$\frac{d\mathcal{I}}{dt} = \oint \left(p_i \frac{\partial^2 q_i}{\partial t\, \partial s} + \frac{\partial p_i}{\partial t} \frac{\partial q_i}{\partial s} \right) ds. \tag{2.68}$$

We integrate the first term by parts, and then used Hamilton's equations of motion to simplify the result. We obtain

$$\frac{d\mathcal{I}}{dt} = \oint \left(-\frac{\partial q_i}{\partial t} \frac{\partial p_i}{\partial s} + \frac{\partial p_i}{\partial t} \frac{\partial q_i}{\partial s} \right) ds = -\oint \left(\frac{\partial H}{\partial p_i} \frac{\partial p_i}{\partial s} + \frac{\partial H}{\partial q_i} \frac{\partial q_i}{\partial s} \right) ds, \tag{2.69}$$

where $H(\mathbf{p}, \mathbf{q}, t)$ is the Hamiltonian for the motion. The integrand is now seen to be the total derivative of H along C. Since the Hamiltonian is a single-valued function, it follows that

$$\frac{d\mathcal{I}}{dt} = -\oint \frac{dH}{ds}\, ds = 0. \tag{2.70}$$

Thus, $\mathcal{I}$ is indeed a constant of the motion.

2.8 Adiabatic Invariants

Poincaré invariants are generally of little practical interest unless the curve C closely corresponds to the trajectories of actual particles. Now, for the motion of magnetized particles it is clear from Eqs. (2.25) and (2.38) that points having the same guiding centre at a certain time will continue to have approximately the same guiding centre at a later time. An approximate Poincaré invariant may thus be obtained by choosing the curve C to be a circle of points corresponding to a gyrophase period. In other words,

$$\mathcal{I} \simeq I = \oint \mathbf{p} \cdot \frac{\partial \mathbf{q}}{\partial \gamma}\, d\gamma. \tag{2.71}$$

Here, I is an *adiabatic invariant*.

To evaluate I for a magnetized plasma recall that the canonical momentum for charged particles is

$$\mathbf{p} = m\,\mathbf{v} + e\,\mathbf{A}, \tag{2.72}$$

where $\mathbf{A}$ is the vector potential. We express $\mathbf{A}$ in terms of its Taylor series about the guiding centre position:

$$\mathbf{A}(\mathbf{r}) = \mathbf{A}(\mathbf{R}) + (\boldsymbol{\rho} \cdot \nabla)\, \mathbf{A}(\mathbf{R}) + O(\rho^2). \tag{2.73}$$

The element of length along the curve $C(t)$ is [see Eq. (2.39)]

$$d\mathbf{r} = \frac{\partial \boldsymbol{\rho}}{\partial \gamma}\, d\gamma = \frac{\mathbf{u}}{\Omega}\, d\gamma. \tag{2.74}$$

The adiabatic invariant is thus

$$I = \oint \frac{\mathbf{u}}{\Omega} \cdot \{m\,(\mathbf{U}+\mathbf{u}) + e\,[\mathbf{A} + (\boldsymbol{\rho}\cdot\nabla)\,\mathbf{A}]\}\,d\gamma + O(\epsilon), \tag{2.75}$$

which reduces to

$$I = 2\pi\, m\, \frac{u_\perp^{\,2}}{\Omega} + 2\pi\, \frac{e}{\Omega}\,\langle \mathbf{u}\cdot(\boldsymbol{\rho}\cdot\nabla)\,\mathbf{A}\rangle + O(\epsilon). \tag{2.76}$$

The final term on the right-hand side is written [see Eq. (2.41)]

$$2\pi\, e\,\langle(\boldsymbol{\rho}\times\mathbf{b})\cdot(\boldsymbol{\rho}\cdot\nabla)\,\mathbf{A}\rangle = -2\pi\, e\,\frac{u_\perp^{\,2}}{2\,\Omega^2}\,\mathbf{b}\cdot\nabla\times\mathbf{A} = -\pi\, m\,\frac{u_\perp^{\,2}}{\Omega}. \tag{2.77}$$

It follows that

$$I = 2\pi\,\frac{m}{e}\,\mu + O(\epsilon). \tag{2.78}$$

Thus, to lowest order the adiabatic invariant is proportional to the magnetic moment μ.

2.9 Magnetic Mirrors

Consider the important case in which the electromagnetic fields do not vary in time. It immediately follows from Eq. (2.64) that

$$\frac{d\mathcal{E}}{dt} = 0, \tag{2.79}$$

where

$$\mathcal{E} = K + e\,\phi = \frac{m}{2}\,(U_\parallel^{\,2} + \mathbf{v}_E^{\,2}) + \mu\, B + e\,\phi \tag{2.80}$$

is the total particle energy, and ϕ is the electrostatic potential. Not surprisingly, a charged particle neither gains nor loses energy as it moves around in non-time-varying electromagnetic fields. Since both $\mathcal{E}$ and μ are constants of the motion, we can rearrange Eq. (2.80) to give

$$U_\parallel = \pm\sqrt{(2/m)[\mathcal{E} - \mu\, B - e\,\phi] - \mathbf{v}_E^{\,2}}. \tag{2.81}$$

Thus, in regions where $\mathcal{E} > \mu\, B + e\,\phi + m\,\mathbf{v}_E^{\,2}/2$ charged particles can drift in either direction along magnetic field-lines. However, particles are excluded from regions where $\mathcal{E} < \mu\, B + e\,\phi + m\,\mathbf{v}_E^{\,2}/2$ (since particles cannot have imaginary parallel velocities!). Evidently, charged particles must reverse direction at those points on magnetic field-lines where $\mathcal{E} = \mu\, B + e\,\phi + m\,\mathbf{v}_E^{\,2}/2$: such points are termed "bounce points" or "mirror points."

Let us now consider how we might construct a device to confine a collisionless (*i.e.*, very hot) plasma. Obviously, we cannot use conventional solid walls, because they would melt. However, it is possible to confine a hot plasma using a magnetic field (fortunately, magnetic fields do not melt!): this technique is called *magnetic confinement*. The electric field in confined plasmas is usually weak (*i.e.*, $E \ll B\,v$), so that the $\mathbf{E}\times\mathbf{B}$ drift is similar in

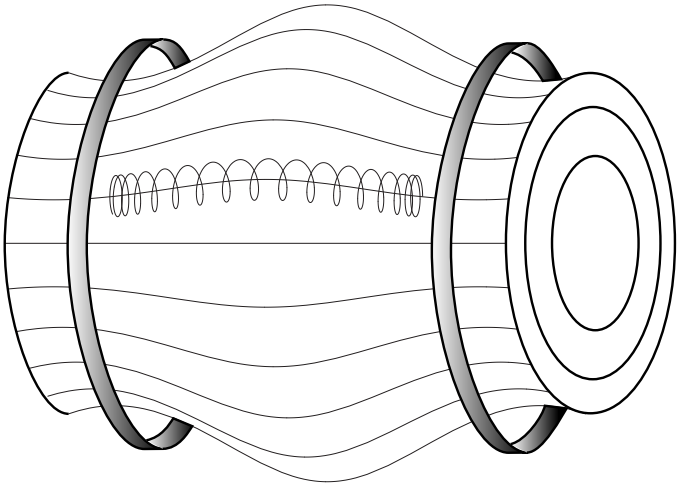

Figure 2.1: *Motion of a trapped particle in a mirror machine.*

magnitude to the magnetic and curvature drifts. In this case, the bounce point condition, $U_\parallel = 0$, reduces to

$$\mathcal{E} = \mu\, B. \tag{2.82}$$

Consider the magnetic field configuration shown in Fig. 1. This is most easily produced using two Helmholtz coils. Incidentally, this type of magnetic confinement device is called a *magnetic mirror machine*. The magnetic field configuration obviously possesses axial symmetry. Let z be a coordinate which measures distance along the axis of symmetry. Suppose that $z = 0$ corresponds to the mid-plane of the device (*i.e.*, halfway between the two field-coils).

It is clear from Fig. 2.1 that the magnetic field-strength $B(z)$ on a magnetic field-line situated close to the axis of the device attains a local minimum B_{min} at $z = 0$, increases symmetrically as $|z|$ increases until reaching a maximum value B_{max} at about the location of the two field-coils, and then decreases as $|z|$ is further increased. According to Eq. (2.82), any particle which satisfies the inequality

$$\mu > \mu_{trap} = \frac{\mathcal{E}}{B_{max}} \tag{2.83}$$

is *trapped* on such a field-line. In fact, the particle undergoes periodic motion along the field-line between two symmetrically placed (in z) mirror points. The magnetic field-strength at the mirror points is

$$B_{mirror} = \frac{\mu_{trap}}{\mu}\, B_{max} < B_{max}. \tag{2.84}$$

Now, on the mid-plane $\mu = m\, v_\perp^2/2\, B_{min}$ and $\mathcal{E} = m\,(v_\parallel^2 + v_\perp^2)/2$. (*n.b.* From now on, we shall write $\mathbf{v} = v_\parallel\, \mathbf{b} + \mathbf{v}_\perp$, for ease of notation.) Thus, the trapping condition (2.83) reduces to

$$\frac{|v_\parallel|}{|v_\perp|} < (B_{max}/B_{min} - 1)^{1/2}. \tag{2.85}$$

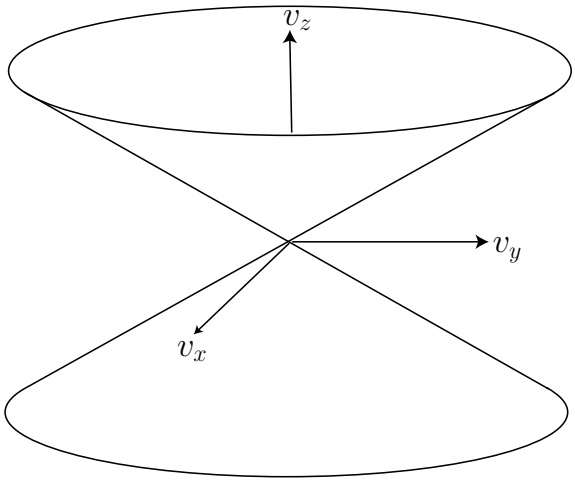

Figure 2.2: *Loss cone in velocity space. The particles lying inside the cone are not reflected by the magnetic field.*

Particles on the mid-plane which satisfy this inequality are trapped: particles which do not satisfy this inequality escape along magnetic field-lines. Clearly, a magnetic mirror machine is incapable of trapping charged particles which are moving parallel, or nearly parallel, to the direction of the magnetic field. In fact, the above inequality defines a *loss cone* in velocity space—see Fig. 2.2.

It is clear that if plasma is placed inside a magnetic mirror machine then all of the particles whose velocities lie in the loss cone promptly escape, but the remaining particles are confined. Unfortunately, that is not the end of the story. There is no such thing as an absolutely collisionless plasma. Collisions take place at a low rate even in very hot plasmas. One important effect of collisions is to cause *diffusion* of particles in velocity space. Thus, in a mirror machine collisions continuously scatter trapped particles into the loss cone, giving rise to a slow leakage of plasma out of the device. Even worse, plasmas whose distribution functions deviate strongly from an isotropic Maxwellian (*e.g.*, a plasma confined in a mirror machine) are prone to *velocity space instabilities*, which tend to relax the distribution function back to a Maxwellian. Clearly, such instabilities are likely to have a disastrous effect on plasma confinement in a mirror machine. For these reasons, magnetic mirror machines are not particularly successful plasma confinement devices, and attempts to achieve nuclear fusion using this type of device have mostly been abandoned.[3]

[3]This is not quite true. In fact, fusion scientists have developed advanced mirror concepts which do not suffer from the severe end-losses characteristic of standard mirror machines. Mirror research is still being carried out, albeit at a comparatively low level, in Russia and Japan.

2.10 Van Allen Radiation Belts

Plasma confinement via magnetic mirroring occurs in nature as well as in unsuccessful fusion devices. For instance, the Van Allen radiation belts, which surround the Earth, consist of energetic particles trapped in the Earth's dipole-like magnetic field. These belts were discovered by James A. Van Allen and co-workers using data taken from Geiger counters which flew on the early U.S. satellites, Explorer 1 (which was, in fact, the first U.S. satellite), Explorer 4, and Pioneer 3. Van Allen was actually trying to measure the flux of cosmic rays (high energy particles whose origin is outside the Solar System) in outer space, to see if it was similar to that measured on Earth. However, the flux of energetic particles detected by his instruments so greatly exceeded the expected value that it prompted one of his co-workers to exclaim, "My God, space is radioactive!" It was quickly realized that this flux was due to energetic particles trapped in the Earth's magnetic field, rather than to cosmic rays.

There are, in fact, two radiation belts surrounding the Earth. The *inner belt*, which extends from about 1–3 Earth radii in the equatorial plane is mostly populated by protons with energies exceeding 10 MeV. The origin of these protons is thought to be the decay of neutrons which are emitted from the Earth's atmosphere as it is bombarded by cosmic rays. The inner belt is fairly quiescent. Particles eventually escape due to collisions with neutral atoms in the upper atmosphere above the Earth's poles. However, such collisions are sufficiently uncommon that the lifetime of particles in the belt range from a few hours to 10 years. Clearly, with such long trapping times only a small input rate of energetic particles is required to produce a region of intense radiation.

The *outer belt*, which extends from about 3–9 Earth radii in the equatorial plane, consists mostly of electrons with energies below 10 MeV. The origin of these electrons is via injection from the outer magnetosphere. Unlike the inner belt, the outer belt is very dynamic, changing on time-scales of a few hours in response to perturbations emanating from the outer magnetosphere.

In regions not too far distant (*i.e.*, less than 10 Earth radii) from the Earth, the geomagnetic field can be approximated as a dipole field,

$$\mathbf{B} = \frac{\mu_0}{4\pi}\frac{M_E}{r^3}(-2\cos\theta, -\sin\theta, 0), \tag{2.86}$$

where we have adopted conventional spherical polar coordinates (r, θ, φ) aligned with the Earth's dipole moment, whose magnitude is $M_E = 8.05\times 10^{22}\,\mathrm{A\,m^2}$. It is usually convenient to work in terms of the *latitude*, $\vartheta = \pi/2 - \theta$, rather than the polar angle, θ. An individual magnetic field-line satisfies the equation

$$r = r_{eq}\,\cos^2\vartheta, \tag{2.87}$$

where r_{eq} is the radial distance to the field-line in the equatorial plane ($\vartheta = 0^\circ$). It is conventional to label field-lines using the *L-shell parameter*, $L = r_{eq}/R_E$. Here, $R_E = 6.37\times$

10^6 m is the Earth's radius. Thus, the variation of the magnetic field-strength along a field-line characterized by a given L-value is

$$B = \frac{B_E}{L^3}\frac{(1+3\,\sin^2\vartheta)^{1/2}}{\cos^6\vartheta}, \tag{2.88}$$

where $B_E = \mu_0 M_E/(4\pi\,R_E^3) = 3.11\times 10^{-5}$ T is the equatorial magnetic field-strength on the Earth's surface.

Consider, for the sake of simplicity, charged particles located on the equatorial plane ($\vartheta = 0^\circ$) whose velocities are predominately directed perpendicular to the magnetic field. The proton and electron gyrofrequencies are written[4]

$$\Omega_p = \frac{e\,B}{m_p} = 2.98\,L^{-3}\ \text{kHz}, \tag{2.89}$$

and

$$|\Omega_e| = \frac{e\,B}{m_e} = 5.46\,L^{-3}\ \text{MHz}, \tag{2.90}$$

respectively. The proton and electron gyroradii, expressed as fractions of the Earth's radius, take the form

$$\frac{\rho_p}{R_E} = \frac{\sqrt{2\,\mathcal{E}\,m_p}}{e\,B\,R_E} = \sqrt{\mathcal{E}(\text{MeV})}\left(\frac{L}{11.1}\right)^3, \tag{2.91}$$

and

$$\frac{\rho_e}{R_E} = \frac{\sqrt{2\,\mathcal{E}\,m_e}}{e\,B\,R_E} = \sqrt{\mathcal{E}(\text{MeV})}\left(\frac{L}{38.9}\right)^3, \tag{2.92}$$

respectively. It is clear that MeV energy charged particles in the inner magnetosphere (*i.e*, $L \ll 10$) gyrate at frequencies which are much greater than the typical rate of change of the magnetic field (which changes on time-scales which are, at most, a few minutes). Likewise, the gyroradii of such particles are much smaller than the typical variation length-scale of the magnetospheric magnetic field. Under these circumstances, we expect the magnetic moment to be a conserved quantity: *i.e.*, we expect the magnetic moment to be a good adiabatic invariant. It immediately follows that any MeV energy protons and electrons in the inner magnetosphere which have a sufficiently large magnetic moment are *trapped* on the dipolar field-lines of the Earth's magnetic field, bouncing back and forth between mirror points located just above the Earth's poles.

It is helpful to define the *pitch-angle*,

$$\alpha = \tan^{-1}(v_\perp/v_\parallel), \tag{2.93}$$

of a charged particle in the magnetosphere. If the magnetic moment is a conserved quantity then a particle of fixed energy drifting along a field-line satisfies

$$\frac{\sin^2\alpha}{\sin^2\alpha_{eq}} = \frac{B}{B_{eq}}, \tag{2.94}$$

[4]It is conventional to take account of the negative charge of electrons by making the electron gyrofrequency Ω_e negative. This approach is implicit in formulae such as Eq. (2.52).

where α_{eq} is the *equatorial pitch-angle* (*i.e.*, the pitch-angle on the equatorial plane) and $B_{eq} = B_E/L^3$ is the magnetic field-strength on the equatorial plane. It is clear from Eq. (2.88) that the pitch-angle increases (*i.e.*, the parallel component of the particle velocity decreases) as the particle drifts off the equatorial plane towards the Earth's poles.

The mirror points correspond to $\alpha = 90^\circ$ (*i.e.*, $v_\parallel = 0$). It follows from Eqs. (2.88) and (2.94) that

$$\sin^2\alpha_{eq} = \frac{B_{eq}}{B_m} = \frac{\cos^6\vartheta_m}{(1+3\,\sin^2\vartheta_m)^{1/2}}, \tag{2.95}$$

where B_m is the magnetic field-strength at the mirror points, and ϑ_m is the latitude of the mirror points. Clearly, the latitude of a particle's mirror point depends only on its equatorial pitch-angle, and is independent of the L-value of the field-line on which it is trapped.

Charged particles with large equatorial pitch-angles have small parallel velocities, and mirror points located at relatively low latitudes. Conversely, charged particles with small equatorial pitch-angles have large parallel velocities, and mirror points located at high latitudes. Of course, if the pitch-angle becomes too small then the mirror points enter the Earth's atmosphere, and the particles are lost via collisions with neutral particles. Neglecting the thickness of the atmosphere with respect to the radius of the Earth, we can say that all particles whose mirror points lie inside the Earth are lost via collisions. It follows from Eq. (2.95) that the *equatorial loss cone* is of approximate width

$$\sin^2\alpha_l = \frac{\cos^6\vartheta_E}{(1+3\sin^2\vartheta_E)^{1/2}}, \tag{2.96}$$

where ϑ_E is the latitude of the point where the magnetic field-line under investigation intersects the Earth. Note that all particles with $|\alpha_{eq}| < \alpha_l$ and $|\pi - \alpha_{eq}| < \alpha_l$ lie in the loss cone. It is easily demonstrated from Eq. (2.87) that

$$\cos^2\vartheta_E = L^{-1}. \tag{2.97}$$

It follows that

$$\sin^2\alpha_l = (4\,L^6 - 3\,L^5)^{-1/2}. \tag{2.98}$$

Thus, the width of the loss cone is independent of the charge, the mass, or the energy of the particles drifting along a given field-line, and is a function only of the field-line radius on the equatorial plane. The loss cone is surprisingly small. For instance, at the radius of a geostationary orbit ($6.6\,R_E$), the loss cone is less than 3° degrees wide. The smallness of the loss cone is a consequence of the very strong variation of the magnetic field-strength along field-lines in a dipole field—see Eqs. (2.85) and (2.88).

A dipole field is clearly a far more effective configuration for confining a collisionless plasma via magnetic mirroring than the more traditional linear configuration shown in Fig. 2.1. In fact, M.I.T. has recently constructed a dipole mirror machine. The dipole field is generated by a superconducting current loop levitating in a vacuum chamber.

The *bounce period*, τ_b, is the time it takes a particle to move from the equatorial plane to one mirror point, then to the other, and then return to the equatorial plane. It follows that

$$\tau_b = 4\int_0^{\vartheta_m} \frac{d\vartheta}{v_\parallel}\frac{ds}{d\vartheta}, \tag{2.99}$$

where ds is an element of arc length along the field-line under investigation, and $v_\parallel = v\,(1 - B/B_m)^{1/2}$. The above integral cannot be performed analytically. However, it can be solved numerically, and is conveniently approximated as

$$\tau_b \simeq \frac{L\,R_E}{(\mathcal{E}/m)^{1/2}}\,(3.7 - 1.6\sin\alpha_{eq}). \tag{2.100}$$

Thus, for protons

$$(\tau_b)_p \simeq 2.41\,\frac{L}{\sqrt{\mathcal{E}(\mathrm{MeV})}}\,(1 - 0.43\sin\alpha_{eq})\ \ \mathrm{secs}, \tag{2.101}$$

whilst for electrons

$$(\tau_b)_e \simeq 5.62\times 10^{-2}\,\frac{L}{\sqrt{\mathcal{E}(\mathrm{MeV})}}\,(1 - 0.43\sin\alpha_{eq})\ \ \mathrm{secs}. \tag{2.102}$$

It follows that MeV electrons typically have bounce periods which are less than a second, whereas the bounce periods for MeV protons usually lie in the range 1 to 10 seconds. The bounce period only depends weakly on equatorial pitch-angle, since particles with small pitch angles have relatively large parallel velocities but a comparatively long way to travel to their mirror points, and *vice versa*. Naturally, the bounce period is longer for longer field-lines (*i.e.*, for larger L).

2.11 Ring Current

Up to now, we have only considered the lowest order motion (*i.e.*, gyration combined with parallel drift) of charged particles in the magnetosphere. Let us now examine the higher order corrections to this motion. For the case of non-time-varying fields, and a weak electric field, these corrections consist of a combination of $\mathbf{E}\times\mathbf{B}$ drift, grad-B drift, and curvature drift:

$$\mathbf{v}_{1\perp} = \frac{\mathbf{E}\times\mathbf{B}}{B^2} + \frac{\mu}{m\,\Omega}\,\mathbf{b}\times\nabla B + \frac{v_\parallel^{\,2}}{\Omega}\,\mathbf{b}\times(\mathbf{b}\cdot\nabla)\,\mathbf{b}. \tag{2.103}$$

Let us neglect $\mathbf{E}\times\mathbf{B}$ drift, since this motion merely gives rise to the convection of plasma within the magnetosphere, without generating a current. By contrast, there is a net current associated with grad-B drift and curvature drift. In the limit in which this current does not strongly modify the ambient magnetic field (*i.e.*, $\nabla\times\mathbf{B}\simeq\mathbf{0}$), which is certainly the situation in the Earth's magnetosphere, we can write

$$(\mathbf{b}\cdot\nabla)\,\mathbf{b} = -\mathbf{b}\times(\nabla\times\mathbf{b}) \simeq \frac{\nabla_\perp B}{B}. \tag{2.104}$$

It follows that the higher order drifts can be combined to give

$$\mathbf{v}_{1\perp} = \frac{(v_\perp^2/2 + v_\parallel^2)}{\Omega\, B}\, \mathbf{b}\times\nabla B. \tag{2.105}$$

For the dipole field (2.86), the above expression yields

$$\mathbf{v}_{1\perp} \simeq -\mathrm{sgn}(\Omega)\, \frac{6\,\mathcal{E}\, L^2}{e\, B_E\, R_E}\, (1 - B/2B_m)\, \frac{\cos^5\vartheta\,(1+\sin^2\vartheta)}{(1+3\sin^2\vartheta)^2}\, \hat{\boldsymbol{\varphi}}. \tag{2.106}$$

Note that the drift is in the azimuthal direction. A positive drift velocity corresponds to eastward motion, whereas a negative velocity corresponds to westward motion. It is clear that, in addition to their gyromotion and periodic bouncing motion along field-lines, charged particles trapped in the magnetosphere also slowly *precess* around the Earth. The ions drift westwards and the electrons drift eastwards, giving rise to a net westward current circulating around the Earth. This current is known as the *ring current*.

Although the perturbations to the Earth's magnetic field induced by the ring current are small, they are still detectable. In fact, the ring current causes a slight *reduction* in the Earth's magnetic field in equatorial regions. The size of this reduction is a good measure of the number of charged particles contained in the Van Allen belts. During the development of so-called *geomagnetic storms*, charged particles are injected into the Van Allen belts from the outer magnetosphere, giving rise to a sharp increase in the ring current, and a corresponding decrease in the Earth's equatorial magnetic field. These particles eventually precipitate out of the magnetosphere into the upper atmosphere at high latitudes, giving rise to intense auroral activity, serious interference in electromagnetic communications, and, in extreme cases, disruption of electric power grids. The ring current induced reduction in the Earth's magnetic field is measured by the so-called *Dst index*, which is based on hourly averages of the northward horizontal component of the terrestrial magnetic field recorded at four low-latitude observatories; Honolulu (Hawaii), San Juan (Puerto Rico), Hermanus (South Africa), and Kakioka (Japan). Figure 2.3 shows the Dst index for the month of March 1989.[5] The very marked reduction in the index, centred about March 13th, corresponds to one of the most severe geomagnetic storms experienced in recent decades. In fact, this particular storm was so severe that it tripped out the whole Hydro Quebec electric distribution system, plunging more than 6 million customers into darkness. Most of Hydro Quebec's neighbouring systems in the United States came uncomfortably close to experiencing the same cascading power outage scenario. Note that a reduction in the Dst index by 600 nT corresponds to a 2% reduction in the terrestrial magnetic field at the equator.

According to Eq. (2.106), the precessional drift velocity of charged particles in the magnetosphere is a rapidly decreasing function of increasing latitude (*i.e.*, most of the ring current is concentrated in the equatorial plane). Since particles typically complete

[5]Dst data is freely availabel from the following web site in Kyoto (Japan): `http://swdcdb.kugi.kyoto-u.ac.jp/dstdir`

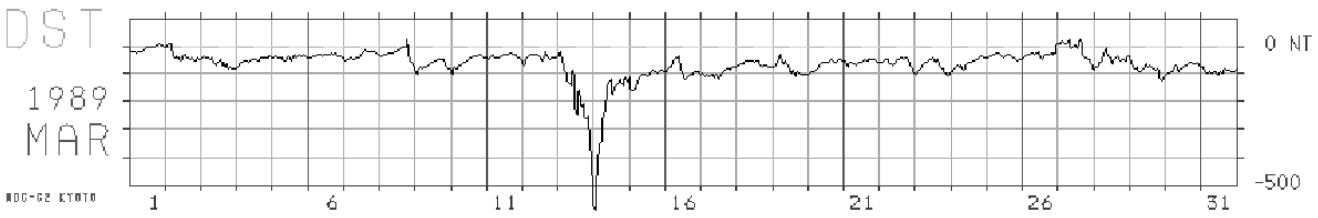

Figure 2.3: *Dst data for March 1989 showing an exceptionally severe geomagnetic storm on the 13th.*

many bounce orbits during a full rotation around the Earth, it is convenient to average Eq. (2.106) over a bounce period to obtain the *average drift velocity*. This averaging can only be performed numerically. The final answer is well approximated by

$$\langle v_{\rm d}\rangle \simeq \frac{6\,\mathcal{E}\,{\rm L}^2}{e\,{\rm B_E}\,{\rm R_E}}\,(0.35+0.15\,\sin\alpha_{\rm eq}). \tag{2.107}$$

The *average drift period* (*i.e.*, the time required to perform a complete rotation around the Earth) is simply

$$\langle \tau_{\rm d}\rangle = \frac{2\pi\,{\rm L}\,{\rm R_E}}{\langle v_{\rm d}\rangle} \simeq \frac{\pi\,e\,{\rm B_E}\,{\rm R_E}^2}{3\,\mathcal{E}\,{\rm L}}\,(0.35+0.15\,\sin\alpha_{\rm eq})^{-1}. \tag{2.108}$$

Thus, the drift period for protons and electrons is

$$\langle \tau_{\rm d}\rangle_{\rm p} = \langle \tau_{\rm d}\rangle_{e} \simeq \frac{1.05}{\mathcal{E}({\rm MeV})\,{\rm L}}\,(1+0.43\,\sin\alpha_{\rm eq})^{-1}\ \ {\rm hours}. \tag{2.109}$$

Note that MeV energy electrons and ions precess around the Earth with about the same velocity, only in opposite directions, because there is no explicit mass dependence in Eq. (2.107). It typically takes an hour to perform a full rotation. The drift period only depends weakly on the equatorial pitch angle, as is the case for the bounce period. Somewhat paradoxically, the drift period is shorter on more distant L-shells. Note, of course, that particles only get a chance to complete a full rotation around the Earth if the inner magnetosphere remains quiescent on time-scales of order an hour, which is, by no means, always the case.

Note, finally, that, since the rest mass of an electron is 0.51 MeV, most of the above formulae require relativistic correction when applied to MeV energy electrons. Fortunately, however, there is no such problem for protons, whose rest mass energy is 0.94 GeV.

2.12 Second Adiabatic Invariant

We have seen that there is an adiabatic invariant associated with the periodic gyration of a charged particle around magnetic field-lines. Thus, it is reasonable to suppose that there

is a second adiabatic invariant associated with the periodic bouncing motion of a particle trapped between two mirror points on a magnetic field-line. This is indeed the case.

Recall that an adiabatic invariant is the lowest order approximation to a Poincaré invariant:

$$\mathcal{J} = \oint_C \mathbf{p}\cdot \mathrm{d}\mathbf{q}. \tag{2.110}$$

In this case, let the curve C correspond to the trajectory of a guiding centre as a charged particle trapped in the Earth's magnetic field executes a bounce orbit. Of course, this trajectory does not quite close, because of the slow azimuthal drift of particles around the Earth. However, it is easily demonstrated that the azimuthal displacement of the end point of the trajectory, with respect to the beginning point, is of order the gyroradius. Thus, in the limit in which the ratio of the gyroradius, ρ, to the variation length-scale of the magnetic field, L, tends to zero, the trajectory of the guiding centre can be regarded as being approximately closed, and the actual particle trajectory conforms very closely to that of the guiding centre. Thus, the adiabatic invariant associated with the bounce motion can be written

$$\mathcal{J} \simeq \mathrm{J} = \oint p_\parallel \,\mathrm{d}s, \tag{2.111}$$

where the path of integration is along a field-line: from the equator to the upper mirror point, back along the field-line to the lower mirror point, and then back to the equator. Furthermore, $\mathrm{d}s$ is an element of arc-length along the field-line, and $p_\parallel \equiv \mathbf{p}\cdot\mathbf{b}$. Using $\mathbf{p} = m\,\mathbf{v} + e\,\mathbf{A}$, the above expression yields

$$\mathrm{J} = m\oint v_\parallel \,\mathrm{d}s + e\oint \mathbf{A}_\parallel \,\mathrm{d}s = m\oint v_\parallel \,\mathrm{d}s + e\,\Phi. \tag{2.112}$$

Here, Φ is the total magnetic flux enclosed by the curve—which, in this case, is obviously zero. Thus, the so-called *second adiabatic invariant* or *longitudinal adiabatic invariant* takes the form

$$\mathrm{J} = m\oint v_\parallel \,\mathrm{d}s. \tag{2.113}$$

In other words, the second invariant is proportional to the loop integral of the parallel (to the magnetic field) velocity taken over a bounce orbit. Actually, the above "proof" is not particularly rigorous: the rigorous proof that J is an adiabatic invariant was first given by Northrop and Teller.[6] It should be noted, of course, that J is only a constant of the motion for particles trapped in the inner magnetosphere provided that the magnetospheric magnetic field varies on time-scales much longer than the bounce time, τ_b. Since the bounce time for MeV energy protons and electrons is, at most, a few seconds, this is not a particularly onerous constraint.

The invariance of J is of great importance for charged particle dynamics in the Earth's inner magnetosphere. It turns out that the Earth's magnetic field is distorted from pure axisymmetry by the action of the solar wind, as illustrated in Fig. 2.4. Because of this

[6] T.G. Northrop, and E. Teller, Phys. Rev. **117**, 215 (1960).

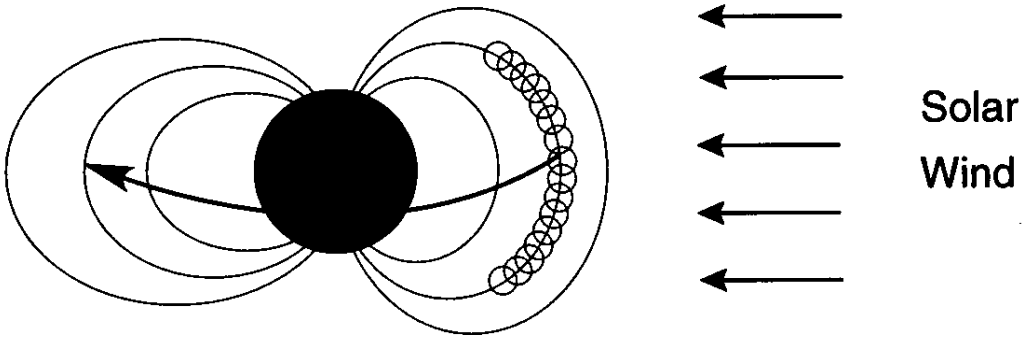

Figure 2.4: *The distortion of the Earth's magnetic field by the solar wind.*

asymmetry, there is no particular reason to believe that a particle will return to its earlier trajectory as it makes a full rotation around the Earth. In other words, the particle may well end up on a different field-line when it returns to the same azimuthal angle. However, at a given azimuthal angle, each field-line has a different length between mirror points, and a different variation of the field-strength B between the mirror points, for a particle with given energy $\mathcal{E}$ and magnetic moment μ. Thus, each field-line represents a different value of J for that particle. So, if J is conserved, as well as $\mathcal{E}$ and μ, then the particle *must* return to the same field-line after precessing around the Earth. In other words, the conservation of J prevents charged particles from spiraling radially in or out of the Van Allen belts as they rotate around the Earth. This helps to explain the persistence of these belts.

2.13 Third Adiabatic Invariant

It is clear, by now, that there is an adiabatic invariant associated with every periodic motion of a charged particle in an electromagnetic field. Now, we have just demonstrated that, as a consequence of J-conservation, the drift orbit of a charged particle precessing around the Earth is approximately closed, despite the fact that the Earth's magnetic field is non-axisymmetric. Thus, there must be a third adiabatic invariant associated with the precession of particles around the Earth. Just as we can define a guiding centre associated with a particle's gyromotion around field-lines, we can also define a *bounce centre* associated with a particle's bouncing motion between mirror points. The bounce centre lies on the equatorial plane, and orbits the Earth once every drift period, τ_d. We can write the third adiabatic invariant as

$$K \simeq \oint p_\phi \, ds, \tag{2.114}$$

where the path of integration is the trajectory of the bounce centre around the Earth. Note that the drift trajectory effectively collapses onto the trajectory of the bounce centre in the limit in which $\rho/L \rightarrow 0$—all of the particle's gyromotion and bounce motion averages to

zero. Now $p_\phi = m v_\phi + e A_\phi$ is dominated by its second term, since the drift velocity v_ϕ is very small. Thus,

$$K \simeq e \oint A_\phi \, ds = e \, \Phi, \tag{2.115}$$

where Φ is the total *magnetic flux* enclosed by the drift trajectory (*i.e.*, the flux enclosed by the orbit of the bounce centre around the Earth). The above "proof" is, again, not particularly rigorous—the invariance of Φ is demonstrated rigorously by Northrup.[7] Note, of course, that Φ is only a constant of the motion for particles trapped in the inner magnetosphere provided that the magnetospheric magnetic field varies on time-scales much longer than the drift period, τ_d. Since the drift period for MeV energy protons and electrons is of order an hour, this is only likely to be the case when the magnetosphere is relatively quiescent (*i.e.*, when there are no geomagnetic storms in progress).

The invariance of Φ has interesting consequences for charged particle dynamics in the Earth's inner magnetosphere. Suppose, for instance, that the strength of the solar wind were to increase slowly (*i.e.*, on time-scales significantly longer than the drift period), thereby, compressing the Earth's magnetic field. The invariance of Φ would cause the charged particles which constitute the Van Allen belts to move radially inwards, towards the Earth, in order to conserve the magnetic flux enclosed by their drift orbits. Likewise, a slow decrease in the strength of the solar wind would cause an outward radial motion of the Van Allen belts.

2.14 Motion in Oscillating Fields

We have seen that charged particles can be confined by a static magnetic field. A somewhat more surprising fact is that charged particles can also be confined by a rapidly oscillating, inhomogeneous electromagnetic wave-field. In order to demonstrate this, we again make use of our averaging technique. To lowest order, a particle executes simple harmonic motion in response to an oscillating wave-field. However, to higher order, any weak inhomogeneity in the field causes the restoring force at one turning point to exceed that at the other. On average, this yields a net force which acts on the *centre of oscillation* of the particle.

Consider a spatially inhomogeneous electromagnetic wave-field oscillating at frequency ω:

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{E}_0(\mathbf{r}) \, \cos \omega t. \tag{2.116}$$

The equation of motion of a charged particle placed in this field is written

$$m \frac{d\mathbf{v}}{dt} = e \, [\mathbf{E}_0(\mathbf{r}) \, \cos \omega t + \mathbf{v} \times \mathbf{B}_0(\mathbf{r}) \, \sin \omega t] \,, \tag{2.117}$$

where

$$\mathbf{B}_0 = -\omega^{-1} \, \nabla \times \mathbf{E}_0, \tag{2.118}$$

[7] T.G. Northrup, *The Adiabatic Motion of Charged Particles* (Interscience, New York NY, 1963).

according to Faraday's law.

In order for our averaging technique to be applicable, the electric field $\mathbf{E}_0$ experienced by the particle must remain approximately constant during an oscillation. Thus,

$$(\mathbf{v}\cdot\nabla)\,\mathbf{E} \ll \omega\,\mathbf{E}. \tag{2.119}$$

When this inequality is satisfied, Eq. (2.118) implies that the magnetic force experienced by the particle is smaller than the electric force by one order in the expansion parameter. In fact, Eq. (2.119) is equivalent to the requirement, $\Omega \ll \omega$, that the particle be unmagnetized.

We now apply the averaging technique. We make the substitution $t \to \tau$ in the oscillatory terms, and seek a change of variables,

$$\mathbf{r} = \mathbf{R} + \boldsymbol{\xi}(\mathbf{R}, \mathbf{U}\,t, \tau), \tag{2.120}$$

$$\mathbf{v} = \mathbf{U} + \mathbf{u}(\mathbf{R}, \mathbf{U}\,t, \tau), \tag{2.121}$$

such that $\boldsymbol{\xi}$ and $\mathbf{u}$ are periodic functions of τ with vanishing mean. Averaging $d\mathbf{r}/dt = \mathbf{v}$ again yields $d\mathbf{R}/dt = \mathbf{U}$ to all orders. To lowest order, the momentum evolution equation reduces to

$$\frac{\partial \mathbf{u}}{\partial \tau} = \frac{e}{m}\,\mathbf{E}_0(\mathbf{R})\,\cos\omega\tau. \tag{2.122}$$

The solution, taking into account the constraints $\langle\mathbf{u}\rangle = \langle\boldsymbol{\xi}\rangle = \mathbf{0}$, is

$$\mathbf{u} = \frac{e}{m\,\omega}\,\mathbf{E}_0\,\sin\omega\tau, \tag{2.123}$$

$$\boldsymbol{\xi} = -\frac{e}{m\,\omega^2}\,\mathbf{E}_0\,\cos\omega\tau. \tag{2.124}$$

Here, $\langle\cdots\rangle \equiv (2\pi)^{-2}\oint(\cdots)\,d(\omega\tau)$ represents an oscillation average.

Clearly, there is no motion of the centre of oscillation to lowest order. To first order, the oscillation average of Eq. (2.117) yields

$$\frac{d\mathbf{U}}{dt} = \frac{e}{m}\,\langle(\boldsymbol{\xi}\cdot\nabla)\mathbf{E} + \mathbf{u}\times\mathbf{B}\rangle\,, \tag{2.125}$$

which reduces to

$$\frac{d\mathbf{U}}{dt} = -\frac{e^2}{m^2\,\omega^2}\left[(\mathbf{E}_0\cdot\nabla)\mathbf{E}_0\,\langle\cos^2\omega\tau\rangle + \mathbf{E}_0\times(\nabla\times\mathbf{E}_0)\,\langle\sin^2\omega\tau\rangle\right]. \tag{2.126}$$

The oscillation averages of the trigonometric functions are both equal to $1/2$. Furthermore, we have $\nabla(|\mathbf{E}_0|^2/2) \equiv (\mathbf{E}_0\cdot\nabla)\mathbf{E}_0 + \mathbf{E}_0\times(\nabla\times\mathbf{E}_0)$. Thus, the equation of motion for the centre of oscillation reduces to

$$m\,\frac{d\mathbf{U}}{dt} = -e\,\nabla\Phi_{\rm pond}, \tag{2.127}$$

where

$$\Phi_{\rm pond} = \frac{1}{4}\frac{e}{m\,\omega^2}\,|\mathbf{E}_0|^2. \tag{2.128}$$

It is clear that the oscillation centre experiences a force, called the *ponderomotive force*, which is proportional to the gradient in the amplitude of the wave-field. The ponderomotive force is independent of the sign of the charge, so both electrons and ions can be confined in the same potential well.

The total energy of the oscillation centre,

$$\mathcal{E}_{\rm oc} = \frac{m}{2}\,U^2 + e\,\Phi_{\rm pond}, \tag{2.129}$$

is conserved by the equation of motion (2.126). Note that the ponderomotive potential energy is equal to the average kinetic energy of the oscillatory motion:

$$e\,\Phi_{\rm pond} = \frac{m}{2}\,\langle u^2\rangle. \tag{2.130}$$

Thus, the force on the centre of oscillation originates in a transfer of energy from the oscillatory motion to the average motion.

Most of the important applications of the ponderomotive force occur in laser plasma physics. For instance, a laser beam can propagate in a plasma provided that its frequency exceeds the plasma frequency. If the beam is sufficiently intense then plasma particles are repulsed from the centre of the beam by the ponderomotive force. The resulting variation in the plasma density gives rise to a cylindrical well in the index of refraction which acts as a wave-guide for the laser beam.

3 Plasma Fluid Theory

3.1 Introduction

In plasma fluid theory, a plasma is characterized by a few local parameters—such as the particle density, the kinetic temperature, and the flow velocity—the time evolution of which are determined by means of *fluid equations*. These equations are analogous to, but generally more complicated than, the equations of hydrodynamics.

Plasma physics can be viewed formally as a closure of Maxwell's equations by means of *constitutive relations*: *i.e.*, expressions for the charge density, ρ_c, and the current density, $\mathbf{j}$, in terms of the electric and magnetic fields, $\mathbf{E}$ and $\mathbf{B}$. Such relations are easily expressed in terms of the microscopic distribution functions, $\mathcal{F}_s$, for each plasma species. In fact,

$$\rho_c = \sum_s e_s \int \mathcal{F}_s(\mathbf{r},\mathbf{v},t)\, d^3\mathbf{v}, \tag{3.1}$$

$$\mathbf{j} = \sum_s e_s \int \mathbf{v}\, \mathcal{F}_s(\mathbf{r},\mathbf{v},t)\, d^3\mathbf{v}. \tag{3.2}$$

Here, $\mathcal{F}_s(\mathbf{r},\mathbf{v},t)$ is the exact, "microscopic" phase-space density of plasma species s (charge e_s, mass m_s) near point $(\mathbf{r},\mathbf{v})$ at time t. The distribution function $\mathcal{F}_s$ is normalized such that its velocity integral is equal to the particle density in coordinate space. Thus,

$$\int \mathcal{F}_s(\mathbf{r},\mathbf{v},t)\, d^3\mathbf{v} = n_s(\mathbf{r},t), \tag{3.3}$$

where $n_s(\mathbf{r},t)$ is the number (per unit volume) of species-s particles near point $\mathbf{r}$ at time t.

If we could determine each $\mathcal{F}_s(\mathbf{r},\mathbf{v},t)$ in terms of the electromagnetic fields, then Eqs. (3.1)–(3.2) would immediately give us the desired constitutive relations. Furthermore, it is easy to see, in principle, how each distribution function evolves. Phase-space conservation requires that

$$\frac{\partial \mathcal{F}_s}{\partial t} + \mathbf{v}\cdot\nabla\mathcal{F}_s + \mathbf{a}_s\cdot\nabla_v\mathcal{F}_s = 0, \tag{3.4}$$

where ∇_v is the velocity space grad-operator, and

$$\mathbf{a}_s = \frac{e_s}{m_s}(\mathbf{E} + \mathbf{v}\times\mathbf{B}) \tag{3.5}$$

is the species-s particle acceleration under the influence of the $\mathbf{E}$ and $\mathbf{B}$ fields.

It would appear that the distribution functions for the various plasma species, from which the constitutive relations are trivially obtained, are determined by a set of rather harmless looking first-order partial differential equations. At this stage, we might wonder

why, if plasma dynamics is apparently so simple when written in terms of distribution functions, we need a fluid description of plasma dynamics at all. It is not at all obvious that fluid theory represents an advance.

The above argument is misleading for several reasons. However, by far the most serious flaw is the view of Eq. (3.4) as a tractable equation. Note that this equation is easy to derive, because it is exact, taking into account all scales from the microscopic to the macroscopic. Note, in particular, that there is *no* statistical averaging involved in Eq. (3.4). It follows that the microscopic distribution function $\mathcal{F}_s$ is essentially a sum of Dirac delta-functions, each following the detailed trajectory of a single particle. Furthermore, the electromagnetic fields in Eq. (3.4) are horribly spiky and chaotic on microscopic scales. In other words, solving Eq. (3.4) amounts to nothing less than solving the classical electromagnetic many-body problem—a completely hopeless task.

A much more useful and tractable equation can be extracted from Eq. (3.4) by ensemble averaging. The average distribution function,

$$\bar{\mathcal{F}}_s \equiv \langle \mathcal{F}_s \rangle_{\rm ensemble}, \tag{3.6}$$

is sensibly smooth, and is closely related to actual experimental measurements. Similarly, the ensemble averaged electromagnetic fields are also smooth. Unfortunately, the extraction of an ensemble averaged equation from Eq. (3.4) is a mathematically challenging exercise, and always requires severe approximation. The problem is that, since the exact electromagnetic fields depend on particle trajectories, **E** and **B** are *not* statistically independent of $\mathcal{F}_s$. In other words, the nonlinear acceleration term in Eq. (3.4),

$$\langle \mathbf{a}_s \cdot \nabla_v \mathcal{F}_s \rangle_{\rm ensemble} \neq \bar{\mathbf{a}}_s \cdot \nabla_v \bar{\mathcal{F}}_s, \tag{3.7}$$

involves *correlations* which need to be evaluated explicitly. In the following, we introduce the short-hand

$$f_s \equiv \bar{\mathcal{F}}_s. \tag{3.8}$$

The traditional goal of kinetic theory is to analyze the correlations, using approximations tailored to the parameter regime of interest, and thereby express the average acceleration term in terms of f_s and the average electromagnetic fields alone. Let us assume that this ambitious task has already been completed, giving an expression of the form

$$\langle \mathbf{a}_s \cdot \nabla_v \mathcal{F}_s \rangle_{\rm ensemble} = \bar{\mathbf{a}}_s \cdot \nabla_v \bar{\mathcal{F}}_s - C_s(f), \tag{3.9}$$

where C_s is a generally extremely complicated operator which accounts for the correlations. Since the most important correlations result from close encounters between particles, C_s is called the *collision operator* (for species s). It is not necessarily a linear operator, and usually involves the distribution functions of both species (the subscript in the argument of C_s is omitted for this reason). Hence, the ensemble averaged version of Eq. (3.4) is written

$$\frac{\partial f_s}{\partial t} + \mathbf{v} \cdot \nabla f_s + \bar{\mathbf{a}}_s \cdot \nabla_v f_s = C_s(f). \tag{3.10}$$

In general, the above equation is very difficult to solve, because of the complexity of the collision operator. However, there are some situations where collisions can be completely neglected. In this case, the apparent simplicity of Eq. (3.4) is not deceptive. A useful kinetic description is obtained by just ensemble averaging this equation to give

$$\frac{\partial f_s}{\partial t} + \mathbf{v} \cdot \nabla f_s + \bar{\mathbf{a}}_s \cdot \nabla_v f_s = 0. \tag{3.11}$$

The above equation, which is known as the *Vlasov equation*, is tractable in sufficiently simple geometry. Nevertheless, the fluid approach has much to offer even in the Vlasov limit: it has intrinsic advantages that weigh decisively in its favour in almost every situation.

Firstly, fluid equations possess the key simplicity of involving fewer dimensions: three spatial dimensions instead of six phase-space dimensions. This advantage is especially important in computer simulations.

Secondly, the fluid description is intuitively appealing. We immediately understand the significance of fluid quantities such as density and temperature, whereas the significance of distribution functions is far less obvious. Moreover, fluid variables are relatively easy to measure in experiments, whereas, in most cases, it is extraordinarily difficult to measure a distribution function accurately. There seems remarkably little point in centering our theoretical description of plasmas on something that we cannot generally measure.

Finally, the kinetic approach to plasma physics is spectacularly inefficient. The species distribution functions f_s provide vastly more information than is needed to obtain the constitutive relations. After all, these relations only depend on the two lowest moments of the species distribution functions. Admittedly, fluid theory cannot generally compute ρ_c and $\mathbf{j}$ without reference to other higher moments of the distribution functions, but it can be regarded as an attempt to impose some efficiency on the task of dynamical closure.

3.2 Moments of the Distribution Function

The kth moment of the (ensemble averaged) distribution function $f_s(\mathbf{r}, \mathbf{v}, t)$ is written

$$\mathbf{M}_k(\mathbf{r}, t) = \int \mathbf{v}\mathbf{v}\cdots\mathbf{v}\, f_s(\mathbf{r}, \mathbf{v}, t)\, d^3\mathbf{v}, \tag{3.12}$$

with k factors of $\mathbf{v}$. Clearly, $\mathbf{M}_k$ is a tensor of rank k.

The set $\{\mathbf{M}_k, k = 0, 1, 2, \cdots\}$ can be viewed as an alternative description of the distribution function, which, indeed, uniquely specifies f_s when the latter is sufficiently smooth. For example, a (displaced) Gaussian distribution is uniquely specified by three moments: M_0, the vector $\mathbf{M}_1$, and the scalar formed by contracting $\mathbf{M}_2$.

The low-order moments all have names and simple physical interpretations. First, we have the (particle) *density*,

$$n_s(\mathbf{r}, t) = \int f_s(\mathbf{r}, \mathbf{v}, t)\, d^3\mathbf{v}, \tag{3.13}$$

and the particle *flux density*,

$$n_s \mathbf{V}_s(\mathbf{r}, t) = \int \mathbf{v}\, f_s(\mathbf{r}, \mathbf{v}, t)\, d^3\mathbf{v}. \tag{3.14}$$

The quantity $\mathbf{V}_s$ is, of course, the *flow velocity*. Note that the electromagnetic sources, (3.1)–(3.2), are determined by these lowest moments:

$$\rho_c = \sum_s e_s n_s, \tag{3.15}$$

$$\mathbf{j} = \sum_s e_s n_s \mathbf{V}_s. \tag{3.16}$$

The second-order moment, describing the flow of momentum in the laboratory frame, is called the *stress tensor*, and denoted by

$$\mathbf{P}_s(\mathbf{r}, t) = \int m_s \mathbf{v}\mathbf{v}\, f_s(\mathbf{r}, \mathbf{v}, t)\, d^3\mathbf{v}. \tag{3.17}$$

Finally, there is an important third-order moment measuring the *energy flux density*,

$$\mathbf{Q}_s(\mathbf{r}, t) = \int \frac{1}{2} m_s v^2 \mathbf{v}\, f_s(\mathbf{r}, \mathbf{v}, t)\, d^3\mathbf{v}. \tag{3.18}$$

It is often convenient to measure the second- and third-order moments in the rest-frame of the species under consideration. In this case, the moments assume different names: the stress tensor measured in the rest-frame is called the *pressure tensor*, $\mathbf{p}_s$, whereas the energy flux density becomes the *heat flux density*, $\mathbf{q}_s$. We introduce the relative velocity,

$$\mathbf{w}_s \equiv \mathbf{v} - \mathbf{V}_s, \tag{3.19}$$

in order to write

$$\mathbf{p}_s(\mathbf{r}, t) = \int m_s \mathbf{w}_s \mathbf{w}_s\, f_s(\mathbf{r}, \mathbf{v}, t)\, d^3\mathbf{v}, \tag{3.20}$$

and

$$\mathbf{q}_s(\mathbf{r}, t) = \int \frac{1}{2} m_s w_s^2 \mathbf{w}_s\, f_s(\mathbf{r}, \mathbf{v}, t)\, d^3\mathbf{v}. \tag{3.21}$$

The trace of the pressure tensor measures the ordinary (or "scalar") pressure,

$$p_s \equiv \frac{1}{3} \mathrm{Tr}\,(\mathbf{p}_s). \tag{3.22}$$

Note that $(3/2)\, p_s$ is the kinetic energy density of species s:

$$\frac{3}{2} p_s = \int \frac{1}{2} m_s w_s^2 f_s\, d^3\mathbf{v}. \tag{3.23}$$

In thermodynamic equilibrium, the distribution function becomes a Maxwellian characterized by some temperature T, and Eq. (3.23) yields $p = n\,T$. It is, therefore, natural to define the (kinetic) temperature as

$$T_s \equiv \frac{p_s}{n_s}. \tag{3.24}$$

Of course, the moments measured in the two different frames are related. By direct substitution, it is easily verified that

$$\mathbf{P}_s = \mathbf{p}_s + m_s n_s \mathbf{V}_s \mathbf{V}_s, \tag{3.25}$$

$$\mathbf{Q}_s = \mathbf{q}_s + \mathbf{p}_s \cdot \mathbf{V}_s + \frac{3}{2} p_s \mathbf{V}_s + \frac{1}{2} m_s n_s V_s^{\,2} \mathbf{V}_s. \tag{3.26}$$

3.3 Moments of the Collision Operator

Boltzmann's famous collision operator for a neutral gas considers only binary collisions, and is, therefore, bilinear in the distribution functions of the two colliding species:

$$C_s(f) = \sum_{s'} C_{ss'}(f_s, f_{s'}), \tag{3.27}$$

where $C_{ss'}$ is linear in each of its arguments. Unfortunately, such bilinearity is not strictly valid for the case of Coulomb collisions in a plasma. Because of the long-range nature of the Coulomb interaction, the closest analogue to ordinary two-particle interaction is mediated by Debye shielding, an intrinsically many-body effect. Fortunately, the departure from bilinearity is *logarithmic* in a weakly coupled plasma, and can, therefore, be neglected to a fairly good approximation (since a logarithm is a comparatively weakly varying function). Thus, from now on, $C_{ss'}$ is presumed to be bilinear.

It is important to realize that there is no simple relationship between the quantity $C_{ss'}$, which describes the effect *on* species s of collisions *with* species s', and the quantity $C_{s's}$. The two operators can have quite different mathematical forms (for example, where the masses m_s and $m_{s'}$ are disparate), and they appear in different equations.

Neutral particle collisions are characterized by Boltzmann's collisional conservation laws: the collisional process conserves particles, momentum, and energy at each point. We expect the same *local* conservation laws to hold for Coulomb collisions in a plasma: the maximum range of the Coulomb force in a plasma is the Debye length, which is assumed to be vanishingly small.

Collisional *particle conservation* is expressed by

$$\int C_{ss'}\, d^3\mathbf{v} = 0. \tag{3.28}$$

Collisional *momentum conservation* requires that

$$\int m_s\, \mathbf{v}\, C_{ss'}\, d^3\mathbf{v} = -\int m_{s'}\, \mathbf{v}\, C_{s's}\, d^3\mathbf{v}. \tag{3.29}$$

That is, the net momentum exchanged between species s and s' must vanish. It is useful to introduce the rate of collisional momentum exchange, called the collisional friction force, or simply the *friction force*:

$$\mathbf{F}_{ss'} \equiv \int m_s \, \mathbf{v} \, C_{ss'} \, d^3\mathbf{v}. \tag{3.30}$$

Clearly, $\mathbf{F}_{ss'}$ is the momentum-moment of the collision operator. The total friction force experienced by species s is

$$\mathbf{F}_s \equiv \sum_{s'} \mathbf{F}_{ss'}. \tag{3.31}$$

Momentum conservation is expressed in detailed form as

$$\mathbf{F}_{ss'} = -\mathbf{F}_{s's}, \tag{3.32}$$

and in non-detailed form as

$$\sum_s \mathbf{F}_s = \mathbf{0}. \tag{3.33}$$

Collisional *energy conservation* requires the quantity

$$W_{Lss'} \equiv \int \frac{1}{2} m_s \, v^2 \, C_{ss'} \, d^3\mathbf{v} \tag{3.34}$$

to be conserved in collisions: *i.e.*,

$$W_{Lss'} + W_{Ls's} = 0. \tag{3.35}$$

Here, the L-subscript indicates that the kinetic energy of both species is measured in the same "lab" frame. Because of Galilean invariance, the choice of this common reference frame does not matter.

An alternative collisional energy-moment is

$$W_{ss'} \equiv \int \frac{1}{2} m_s \, w_s^2 \, C_{ss'} \, d^3\mathbf{v} : \tag{3.36}$$

i.e., the kinetic energy change experienced by species s, due to collisions with species s', measured in the rest frame of species s. The total energy change for species s is, of course,

$$W_s \equiv \sum_{s'} W_{ss'}. \tag{3.37}$$

It is easily verified that

$$W_{Lss'} = W_{ss'} + \mathbf{V}_s \cdot \mathbf{F}_{ss'}. \tag{3.38}$$

Thus, the collisional energy conservation law can be written

$$W_{ss'} + W_{s's} + (\mathbf{V}_s - \mathbf{V}_{s'}) \cdot \mathbf{F}_{ss'} = 0, \tag{3.39}$$

or in non-detailed form

$$\sum_s (W_s + \mathbf{V}_s \cdot \mathbf{F}_s) = 0. \tag{3.40}$$

3.4 Moments of the Kinetic Equation

We obtain fluid equations by taking appropriate moments of the ensemble-average kinetic equation, (3.10). In the following, we suppress all ensemble-average over-bars for ease of notation. It is convenient to rearrange the acceleration term,

$$\mathbf{a}_s \cdot \nabla_v f_s = \nabla_v \cdot (\mathbf{a}_s f_s). \tag{3.41}$$

The two forms are equivalent because flow in velocity space under the Lorentz force is incompressible: *i.e.*,

$$\nabla_v \cdot \mathbf{a}_s = 0. \tag{3.42}$$

Thus, Eq. (3.10) becomes

$$\frac{\partial f_s}{\partial t} + \nabla \cdot (\mathbf{v} f_s) + \nabla_v \cdot (\mathbf{a}_s f_s) = C_s(f). \tag{3.43}$$

The rearrangement of the flow term is, of course, trivial, since $\mathbf{v}$ is independent of $\mathbf{r}$.

The kth moment of the ensemble-average kinetic equation is obtained by multiplying the above equation by k powers of $\mathbf{v}$ and integrating over velocity space. The flow term is simplified by pulling the divergence outside the velocity integral. The acceleration term is treated by partial integration. Note that these two terms couple the kth moment to the $(k+1)$th and $(k-1)$th moments, respectively.

Making use of the collisional conservation laws, the zeroth moment of Eq. (3.43) yields the *continuity equation* for species s:

$$\frac{\partial n_s}{\partial t} + \nabla \cdot (n_s \mathbf{V}_s) = 0. \tag{3.44}$$

Likewise, the first moment gives the *momentum conservation equation* for species s:

$$\frac{\partial (m_s n_s \mathbf{V}_s)}{\partial t} + \nabla \cdot \mathbf{P}_s - e_s n_s (\mathbf{E} + \mathbf{V}_s \times \mathbf{B}) = \mathbf{F}_s. \tag{3.45}$$

Finally, the contracted second moment yields the *energy conservation equation* for species s:

$$\frac{\partial}{\partial t}\left(\frac{3}{2} p_s + \frac{1}{2} m_s n_s V_s^2\right) + \nabla \cdot \mathbf{Q}_s - e_s n_s \mathbf{E} \cdot \mathbf{V}_s = W_s + \mathbf{V}_s \cdot \mathbf{F}_s. \tag{3.46}$$

The interpretation of Eqs. (3.44)–(3.46) as *conservation laws* is straightforward. Suppose that G is some physical quantity (*e.g.*, total number of particles, total energy, ...), and $g(\mathbf{r}, t)$ is its density:

$$G = \int g \, d^3\mathbf{r}. \tag{3.47}$$

If G is conserved then g must evolve according to

$$\frac{\partial g}{\partial t} + \nabla \cdot \mathbf{g} = \Delta g, \tag{3.48}$$

where $\mathbf{g}$ is the flux density of G, and Δg is the local rate per unit volume at which G is created or exchanged with other entities in the fluid. Thus, the density of G at some point changes because there is net flow of G towards or away from that point (measured by the divergence term), or because of local sources or sinks of G (measured by the right-hand side).

Applying this reasoning to Eq. (3.44), we see that $n_s \mathbf{V}_s$ is indeed the species-s particle flux density, and that there are no local sources or sinks of species-s particles.[1] From Eq. (3.45), we see that the stress tensor $\mathbf{P}_s$ is the species-s momentum flux density, and that the species-s momentum is changed locally by the Lorentz force and by collisional friction with other species. Finally, from Eq. (3.46), we see that $\mathbf{Q}_s$ is indeed the species-s energy flux density, and that the species-s energy is changed locally by electrical work, energy exchange with other species, and frictional heating.

3.5 Fluid Equations

It is conventional to rewrite our fluid equations in terms of the pressure tensor, $\mathbf{p}_s$, and the heat flux density, $\mathbf{q}_s$. Substituting from Eqs. (3.25)–(3.26), and performing a little tensor algebra, Eqs. (3.44)–(3.46) reduce to:

$$\frac{dn_s}{dt} + n_s \nabla\cdot\mathbf{V}_s = 0, \tag{3.49}$$

$$m_s n_s \frac{d\mathbf{V}_s}{dt} + \nabla\cdot\mathbf{p}_s - e_s n_s(\mathbf{E} + \mathbf{V}_s \times \mathbf{B}) = \mathbf{F}_s, \tag{3.50}$$

$$\frac{3}{2}\frac{dp_s}{dt} + \frac{3}{2} p_s \nabla\cdot\mathbf{V}_s + \mathbf{p}_s : \nabla\mathbf{V}_s + \nabla\cdot\mathbf{q}_s = W_s. \tag{3.51}$$

Here,

$$\frac{d}{dt} \equiv \frac{\partial}{\partial t} + \mathbf{V_s}\cdot\nabla \tag{3.52}$$

is the well-known *convective derivative*, and

$$\mathbf{p} : \nabla\mathbf{V}_s \equiv (p_s)_{\alpha\beta} \frac{\partial (V_s)_\beta}{\partial r_\alpha}. \tag{3.53}$$

In the above, α and β refer to Cartesian components, and repeated indices are summed (according to the Einstein summation convention). The convective derivative, of course, measures time variation in the local rest frame of the species-s fluid. Strictly speaking, we should include an s subscript with each convective derivative, since this operator is clearly different for different plasma species.

[1] In general, this is not true. Atomic or nuclear processes operating in a plasma can give rise to local sources and sinks of particles of various species. However, if a plasma is sufficiently hot to be completely ionized, but still cold enough to prevent nuclear reactions from occurring, then such sources and sinks are usually negligible.

There is one additional refinement to our fluid equations which is worth carrying out. We introduce the *generalized viscosity tensor*, $\boldsymbol{\pi}_s$, by writing

$$\mathbf{p}_s = p_s\,\mathbf{I} + \boldsymbol{\pi}_s, \tag{3.54}$$

where $\mathbf{I}$ is the unit (identity) tensor. We expect the scalar pressure term to dominate if the plasma is relatively close to thermal equilibrium. We also expect, by analogy with conventional fluid theory, the second term to describe viscous stresses. Indeed, this is generally the case in plasmas, although the generalized viscosity tensor can also include terms which are quite unrelated to conventional viscosity. Equations (3.49)–(3.51) can, thus, be rewritten:

$$\frac{dn_s}{dt} + n_s\,\nabla\cdot\mathbf{V}_s = 0, \tag{3.55}$$

$$m_s n_s\,\frac{d\mathbf{V}_s}{dt} + \nabla p_s + \nabla\cdot\boldsymbol{\pi}_s - e_s n_s(\mathbf{E} + \mathbf{V}_s\times\mathbf{B}) = \mathbf{F}_s, \tag{3.56}$$

$$\frac{3}{2}\frac{dp_s}{dt} + \frac{5}{2}\,p_s\,\nabla\cdot\mathbf{V}_s + \boldsymbol{\pi}_s : \nabla\mathbf{V}_s + \nabla\cdot\mathbf{q}_s = W_s. \tag{3.57}$$

According to Eq. (3.55), the species-s density is constant along a fluid trajectory unless the species-s flow is non-solenoidal. For this reason, the condition

$$\nabla\cdot\mathbf{V}_s = 0 \tag{3.58}$$

is said to describe *incompressible* species-s flow. According to Eq. (3.56), the species-s flow accelerates along a fluid trajectory under the influence of the scalar pressure gradient, the viscous stresses, the Lorentz force, and the frictional force due to collisions with other species. Finally, according to Eq. (3.57), the species-s energy density (*i.e.*, p_s) changes along a fluid trajectory because of the work done in compressing the fluid, viscous heating, heat flow, and the local energy gain due to collisions with other species. Note that the electrical contribution to plasma heating, which was explicit in Eq. (3.46), has now become entirely implicit.

3.6 Entropy Production

It is instructive to rewrite the species-s energy evolution equation (3.57) as an *entropy* evolution equation. The fluid definition of entropy density, which coincides with the thermodynamic entropy density in the limit in which the distribution function approaches a Maxwellian, is

$$s_s = n_s \log\left(\frac{T_s^{\,3/2}}{n_s}\right). \tag{3.59}$$

The corresponding entropy flux density is written

$$\mathbf{s}_s = s_s\,\mathbf{V}_s + \frac{\mathbf{q}_s}{T_s}. \tag{3.60}$$

Clearly, entropy is convected by the fluid flow, but is also carried by the flow of heat, in accordance with the second law of thermodynamics. After some algebra, Eq. (3.57) can be rearranged to give

$$\frac{\partial s_s}{\partial t} + \nabla\cdot\mathbf{s}_s = \Theta_s, \tag{3.61}$$

where the right-hand side is given by

$$\Theta_s = \frac{W_s}{T_s} - \frac{\boldsymbol{\pi}_s : \nabla\mathbf{V}_s}{T_s} - \frac{\mathbf{q}_s}{T_s}\cdot\frac{\nabla T_s}{T_s}. \tag{3.62}$$

It is clear, from our previous discussion of conservation laws, that the quantity Θ_s can be regarded as the *entropy production rate* per unit volume for species s. Note that entropy is produced by collisional heating, viscous heating, and heat flow down temperature gradients.

3.7 Fluid Closure

No amount of manipulation, or rearrangement, can cure our fluid equations of their most serious defect: the fact that they are *incomplete*. In their present form, (3.55)–(3.57), our equations relate interesting fluid quantities, such as the density, n_s, the flow velocity, $\mathbf{V}_s$, and the scalar pressure, p_s, to unknown quantities, such as the viscosity tensor, $\boldsymbol{\pi}_s$, the heat flux density, $\mathbf{q}_s$, and the moments of the collision operator, $\mathbf{F}_s$ and W_s. In order to complete our set of equations, we need to use some additional information to express the latter quantities in terms of the former. This process is known as *closure*.

Lack of closure is an endemic problem in fluid theory. Since each moment is coupled to the next higher moment (*e.g.*, the density evolution depends on the flow velocity, the flow velocity evolution depends on the viscosity tensor, *etc*.), any finite set of exact moment equations is bound to contain more unknowns than equations.

There are two basic types of fluid closure schemes. In *truncation* schemes, higher order moments are arbitrarily assumed to vanish, or simply prescribed in terms of lower moments. Truncation schemes can often provide quick insight into fluid systems, but always involve uncontrolled approximation. *Asymptotic* schemes depend on the rigorous exploitation of some small parameter. They have the advantage of being systematic, and providing some estimate of the error involved in the closure. On the other hand, the asymptotic approach to closure is mathematically very demanding, since it inevitably involves working with the kinetic equation.

The classic example of an asymptotic closure scheme is the Chapman-Enskog theory of a neutral gas dominated by collisions. In this case, the small parameter is the ratio of the mean-free-path between collisions to the macroscopic variation length-scale. It is instructive to briefly examine this theory, which is very well described in a classic monograph by Chapman and Cowling.[2]

[2]S. Chapman, and T.G. Cowling, *The Mathematical Theory of Non-Uniform Gases* (Cambridge University Press, Cambridge UK, 1953).

Consider a neutral gas consisting of identical hard-sphere molecules of mass m and diameter σ. Admittedly, this is not a particularly physical model of a neutral gas, but we are only considering it for illustrative purposes. The fluid equations for such a gas are similar to Eqs. (3.55)–(3.57):

$$\frac{dn}{dt} + n\,\nabla\cdot\mathbf{V} = 0, \tag{3.63}$$

$$mn\,\frac{d\mathbf{V}}{dt} + \nabla p + \nabla\cdot\boldsymbol{\pi} + mn\,\mathbf{g} = \mathbf{0}, \tag{3.64}$$

$$\frac{3}{2}\frac{dp}{dt} + \frac{5}{2}\,p\,\nabla\cdot\mathbf{V} + \boldsymbol{\pi} : \nabla\mathbf{V} + \nabla\cdot\mathbf{q} = 0. \tag{3.65}$$

Here, n is the (particle) density, $\mathbf{V}$ the flow velocity, p the scalar pressure, and $\mathbf{g}$ the acceleration due to gravity. We have dropped the subscript s because, in this case, there is only a single species. Note that there is no collisional friction or heating in a single species system. Of course, there are no electrical or magnetic forces in a neutral gas, so we have included gravitational forces instead. The purpose of the closure scheme is to express the viscosity tensor, $\boldsymbol{\pi}$, and the heat flux density, $\mathbf{q}$, in terms of n, $\mathbf{V}$, or p, and, thereby, complete the set of equations.

The mean-free-path l for hard-sphere molecules is given by

$$l = \frac{1}{\sqrt{2}\,\pi\,n\,\sigma^2}. \tag{3.66}$$

This formula is fairly easy to understand: the volume swept out by a given molecule in moving a mean-free-path must contain, on average, approximately one other molecule. Note that l is completely independent of the speed or mass of the molecules. The mean-free-path is assumed to be much smaller than the variation length-scale L of macroscopic quantities, so that

$$\epsilon = \frac{l}{L} \ll 1. \tag{3.67}$$

In the Chapman-Enskog scheme, the distribution function is expanded, order by order, in the small parameter ϵ:

$$f(\mathbf{r},\mathbf{v},t) = f_0(\mathbf{r},\mathbf{v},t) + \epsilon\, f_1(\mathbf{r},\mathbf{v},t) + \epsilon^2\, f_2(\mathbf{r},\mathbf{v},t) + \cdots. \tag{3.68}$$

Here, f_0, f_1, f_2, *etc.*, are all assumed to be of the same order of magnitude. In fact, only the first *two* terms in this expansion are ever calculated. To zeroth order in ϵ, the kinetic equation requires that f_0 be a Maxwellian:

$$f_0(\mathbf{r},\mathbf{v},t) = n(\mathbf{r})\left(\frac{m}{2\pi\,T(\mathbf{r})}\right)^{3/2}\exp\left[-\frac{m\,(\mathbf{v}-\mathbf{V})^2}{2\,T(\mathbf{r})}\right]. \tag{3.69}$$

Recall that $p = n\,T$. Note that there is zero heat flow or viscous stress associated with a Maxwellian distribution function. Thus, both the heat flux density, $\mathbf{q}$, and the viscosity

tensor, $\boldsymbol{\pi}$, depend on the first-order non-Maxwellian correction to the distribution function, f_1.

It is possible to *linearize* the kinetic equation, and then rearrange it so as to obtain an *integral equation* for f_1 in terms of f_0. This rearrangement depends crucially on the *bilinearity* of the collision operator. Incidentally, the equation is integral because the collision operator is an integral operator. The integral equation is solved by expanding f_1 in velocity space using Laguerre polynomials (sometime called Sonine polynomials). It is possible to reduce the integral equation to an infinite set of simultaneous algebraic equations for the coefficients in this expansion. If the expansion is truncated, after N terms, say, then these algebraic equations can be solved for the coefficients. It turns out that the Laguerre polynomial expansion converges very rapidly. Thus, it is conventional to only keep the first *two* terms in this expansion, which is usually sufficient to ensure an accuracy of about 1% in the final result. Finally, the appropriate moments of f_1 are taken, so as to obtain expression for the heat flux density and the viscosity tensor. Strictly speaking, after evaluating f_1, we should then go on to evaluate f_2, so as to ensure that f_2 really is negligible compared to f_1. In reality, this is never done because the mathematical difficulties involved in such a calculation are prohibitive.

The Chapman-Enskog method outlined above can be applied to *any* assumed force law between molecules, provided that the force is sufficiently short-range (*i.e.*, provided that it falls off faster with increasing separation than the Coulomb force). For all sensible force laws, the viscosity tensor is given by

$$\pi_{\alpha\beta} = -\eta\left(\frac{\partial V_\alpha}{\partial r_\beta} + \frac{\partial V_\beta}{\partial r_\alpha} - \frac{2}{3}\,\nabla\cdot\mathbf{V}\,\delta_{\alpha\beta}\right), \tag{3.70}$$

whereas the heat flux density takes the form

$$\mathbf{q} = -\kappa\,\nabla T. \tag{3.71}$$

Here, η is the *coefficient of viscosity*, and κ is the *coefficient of thermal conduction*. It is convenient to write

$$\eta = m n \chi_\nu, \tag{3.72}$$

$$\kappa = n \chi_t, \tag{3.73}$$

where χ_ν is the *viscous diffusivity* and χ_t is the *thermal diffusivity*. Note that both χ_ν and χ_t have the dimensions $m^2\,s^{-1}$ and are, effectively, *diffusion coefficients*. For the special case of hard-sphere molecules, Chapman-Enskog theory yields:

$$\chi_\nu = \frac{75\,\pi^{1/2}}{64}\left[1 + \frac{3}{202} + \cdots\right]\nu\,l^2 = A_\nu\,\nu\,l^2, \tag{3.74}$$

$$\chi_t = \frac{5\,\pi^{1/2}}{16}\left[1 + \frac{1}{44} + \cdots\right]\nu\,l^2 = A_t\,\nu\,l^2. \tag{3.75}$$

Here,

$$\nu \equiv \frac{\nu_t}{l} \equiv \frac{\sqrt{2\,T/m}}{l} \tag{3.76}$$

is the *collision frequency*. Note that the first two terms in the Laguerre polynomial expansion are shown explicitly (in the square brackets) in Eqs. (3.74)–(3.75).

Equations (3.74)–(3.75) have a simple physical interpretation: the viscous and thermal diffusivities of a neutral gas can be accounted for in terms of the *random-walk diffusion* of molecules with excess momentum and energy, respectively. Recall the standard result in stochastic theory that if particles jump an average distance l, in a random direction, ν times a second, then the diffusivity associated with such motion is $\chi \sim \nu\, l^2$. Chapman-Enskog theory basically allows us to calculate the numerical constants A_ν and A_t, multiplying $\nu\, l^2$ in the expressions for χ_ν and χ_t, for a given force law between molecules. Obviously, these coefficients are different for different force laws. The expression for the mean-free-path, l, is also different for different force laws.

Let $\bar{n}$, $\bar{v}_t$, and $\bar{l}$ be typical values of the particle density, the thermal velocity, and the mean-free-path, respectively. Suppose that the typical flow velocity is $\lambda\,\bar{v}_t$, and the typical variation length-scale is L. Let us define the following normalized quantities: $\hat{n} = n/\bar{n}$, $\hat{v}_t = v_t/\bar{v}_t$, $\hat{l} = l/\bar{l}$, $\hat{\mathbf{r}} = \mathbf{r}/L$, $\hat{\nabla} = L\,\nabla$, $\hat{t} = \lambda\,\bar{v}_t\,t/L$, $\hat{\mathbf{V}} = \mathbf{V}/\lambda\,\bar{v}_t$, $\hat{T} = T/m\,\bar{v}_t^{\,2}$, $\hat{\mathbf{g}} = L\,\mathbf{g}/(1+\lambda^2)\,\bar{v}_t^{\,2}$, $\hat{p} = p/m\,\bar{n}\,\bar{v}_t^{\,2}$, $\hat{\boldsymbol{\pi}} = \boldsymbol{\pi}/\lambda\,\epsilon\,m\,\bar{n}\,\bar{v}_t^{\,2}$, $\hat{\mathbf{q}} = \mathbf{q}/\epsilon\,m\,\bar{n}\,\bar{v}_t^{\,3}$. Here, $\epsilon = \bar{l}/L \ll 1$. Note that

$$\hat{\boldsymbol{\pi}} = -A_\nu\,\hat{n}\,\hat{v}_t\,\hat{l}\left(\frac{\partial \hat{V}_\alpha}{\partial \hat{r}_\beta} + \frac{\partial \hat{V}_\beta}{\partial \hat{r}_\alpha} - \frac{2}{3}\,\hat{\nabla}\cdot\hat{\mathbf{V}}\,\delta_{\alpha\beta}\right), \tag{3.77}$$

$$\hat{\mathbf{q}} = -A_t\,\hat{n}\,\hat{v}_t\,\hat{l}\,\hat{\nabla}\hat{T}. \tag{3.78}$$

All hatted quantities are designed to be $O(1)$. The normalized fluid equations are written:

$$\frac{d\hat{n}}{d\hat{t}} + \hat{n}\,\hat{\nabla}\cdot\hat{\mathbf{V}} = 0, \tag{3.79}$$

$$\lambda^2\,\hat{n}\,\frac{d\hat{\mathbf{V}}}{d\hat{t}} + \hat{\nabla}\hat{p} + \lambda\,\epsilon\,\hat{\nabla}\cdot\hat{\boldsymbol{\pi}} + (1+\lambda^2)\,\hat{n}\,\hat{\mathbf{g}} = \mathbf{0}, \tag{3.80}$$

$$\lambda\,\frac{3}{2}\frac{d\hat{p}}{d\hat{t}} + \lambda\,\frac{5}{2}\,\hat{p}\,\hat{\nabla}\cdot\hat{\mathbf{V}} + \lambda^2\,\epsilon\,\hat{\boldsymbol{\pi}} : \hat{\nabla}\hat{\mathbf{V}} + \epsilon\,\hat{\nabla}\cdot\hat{\mathbf{q}} = 0, \tag{3.81}$$

where

$$\frac{d}{d\hat{t}} \equiv \frac{\partial}{\partial \hat{t}} + \hat{\mathbf{V}}\cdot\hat{\nabla}. \tag{3.82}$$

Note that the only large or small quantities in the above equations are the parameters λ and ϵ.

Suppose that $\lambda \gg 1$. In other words, the flow velocity is much greater than the thermal speed. Retaining only the largest terms in Eqs. (3.79)–(3.81), our system of fluid equations

reduces to (in unnormalized form):

$$\frac{dn}{dt} + n\,\nabla\cdot\mathbf{V} = 0, \tag{3.83}$$

$$\frac{d\mathbf{V}}{dt} + \mathbf{g} \simeq \mathbf{0}. \tag{3.84}$$

These are called the *cold-gas* equations, because they can also be obtained by formally taking the limit $T \rightarrow 0$. The cold-gas equations describe externally driven, highly supersonic, gas dynamics. Note that the gas pressure (*i.e.*, energy density) can be neglected in the cold-gas limit, since the thermal velocity is much smaller than the flow velocity, and so there is no need for an energy evolution equation. Furthermore, the viscosity can also be neglected, since the viscous diffusion velocity is also far smaller than the flow velocity.

Suppose that $\lambda \sim O(1)$. In other words, the flow velocity is of order the thermal speed. Again, retaining only the largest terms in Eqs. (3.79)–(3.81), our system of fluid equations reduces to (in unnormalized form):

$$\frac{dn}{dt} + n\,\nabla\cdot\mathbf{V} = 0, \tag{3.85}$$

$$mn\,\frac{d\mathbf{V}}{dt} + \nabla p + mn\,\mathbf{g} \simeq \mathbf{0}, \tag{3.86}$$

$$\frac{3}{2}\frac{dp}{dt} + \frac{5}{2}\,p\,\nabla\cdot\mathbf{V} \simeq 0. \tag{3.87}$$

The above equations can be rearranged to give:

$$\frac{dn}{dt} + n\,\nabla\cdot\mathbf{V} = 0, \tag{3.88}$$

$$mn\,\frac{d\mathbf{V}}{dt} + \nabla p + mn\,\mathbf{g} \simeq \mathbf{0}, \tag{3.89}$$

$$\frac{d}{dt}\left(\frac{p}{n^{5/3}}\right) \simeq 0. \tag{3.90}$$

These are called the *hydrodynamic* equations, since they are similar to the equations governing the dynamics of water. The hydrodynamic equations govern relatively fast, internally driven, gas dynamics: in particular, the dynamics of *sound waves*. Note that the gas pressure is non-negligible in the hydrodynamic limit, since the thermal velocity is of order the flow speed, and so an energy evolution equation is needed. However, the energy equation takes a particularly simple form, because Eq. (3.90) is immediately recognizable as the *adiabatic* equation of state for a monatomic gas. This is not surprising, since the flow velocity is still much faster than the viscous and thermal diffusion velocities (hence, the absence of viscosity and thermal conductivity in the hydrodynamic equations), in which case the gas acts effectively like a perfect thermal insulator.

Suppose, finally, that $\lambda \sim \epsilon$. In other words, the flow velocity is of order the viscous and thermal diffusion velocities. Our system of fluid equations now reduces to a force balance

criterion,

$$\nabla p + m n \, \mathbf{g} \simeq \mathbf{0}, \tag{3.91}$$

to lowest order. To next order, we obtain a set of equations describing the relatively slow viscous and thermal evolution of the gas:

$$\frac{dn}{dt} + n \, \nabla\cdot\mathbf{V} = 0, \tag{3.92}$$

$$m n \frac{d\mathbf{V}}{dt} + \nabla\cdot\boldsymbol{\pi} \simeq \mathbf{0}, \tag{3.93}$$

$$\frac{3}{2}\frac{dp}{dt} + \frac{5}{2} p \, \nabla\cdot\mathbf{V} + \nabla\cdot\mathbf{q} \simeq 0. \tag{3.94}$$

Clearly, this set of equations is only appropriate to relatively quiescent, quasi-equilibrium, gas dynamics. Note that virtually all of the terms in our original fluid equations, (3.63)–(3.65), must be retained in this limit.

The above investigation reveals an important truth in gas dynamics, which also applies to plasma dynamics. Namely, the form of the fluid equations depends crucially on the typical fluid *velocity* associated with the type of dynamics under investigation. As a general rule, the equations get simpler as the typical velocity get faster, and *vice versa*.

3.8 Braginskii Equations

Let now consider the problem of closure in *plasma* fluid equations. There are, in fact, two possible small parameters in plasmas upon which we could base an asymptotic closure scheme. The first is the ratio of the mean-free-path, l, to the macroscopic length-scale, L. This is only appropriate to *collisional* plasmas. The second is the ratio of the Larmor radius, ρ, to the macroscopic length-scale, L. This is only appropriate to *magnetized* plasmas. There is, of course, no small parameter upon which to base an asymptotic closure scheme in a collisionless, unmagnetized plasma. However, such systems occur predominately in accelerator physics contexts, and are not really "plasmas" at all, since they exhibit virtually no collective effects. Let us investigate Chapman-Enskog-like closure schemes in a *collisional*, quasi-neutral plasma consisting of equal numbers of electrons and ions. We shall treat the unmagnetized and magnetized cases separately.

The first step in our closure scheme is to approximate the actual collision operator for Coulomb interactions by an operator which is strictly *bilinear* in its arguments (see Sect. 3.3). Once this has been achieved, the closure problem is formally of the type which can be solved using the Chapman-Enskog method.

The electrons and ions collision times, $\tau = l/v_t = \nu^{-1}$, are written

$$\tau_e = \frac{6\sqrt{2}\,\pi^{3/2}\,\epsilon_0^{\,2}\,\sqrt{m_e}\;T_e^{\,3/2}}{\ln\Lambda\, e^4\, n}, \tag{3.95}$$

and

$$\tau_i = \frac{12\,\pi^{3/2}\,\epsilon_0^{\,2}\,\sqrt{m_i}\;T_i^{\,3/2}}{\ln\Lambda\, e^4\, n}, \tag{3.96}$$

respectively. Here, $n = n_e = n_i$ is the number density of particles, and $\ln\Lambda$ is a quantity called the *Coulomb logarithm* whose origin is the slight modification to the collision operator mentioned above. The Coulomb logarithm is equal to the natural logarithm of the ratio of the maximum to minimum impact parameters for Coulomb "collisions." In other words, $\ln\Lambda = \ln(d_{max}/d_{min})$. The minimum parameter is simply the distance of closest approach, $d_{min} \simeq r_c = e^2/4\pi\epsilon_0 T_e$ [see Eq. (1.17)]. The maximum parameter is the Debye length, $d_{max} \simeq \lambda_D = \sqrt{\epsilon_0 T_e/n e^2}$, since the Coulomb potential is shielded over distances greater than the Debye length. The Coulomb logarithm is a *very slowly varying* function of the plasma density and the electron temperature, and is well approximated by

$$\ln\Lambda \simeq 6.6 - 0.5\ln n + 1.5\ln T_e, \tag{3.97}$$

where n is expressed in units of $10^{20}\,m^{-3}$, and T_e is expressed in electron volts.

The basic forms of Eqs. (3.95) and (3.96) are not hard to understand. From Eq. (3.66), we expect

$$\tau \sim \frac{l}{\nu_t} \sim \frac{1}{n\,\sigma^2\,\nu_t}, \tag{3.98}$$

where σ^2 is the typical "cross-section" of the electrons or ions for Coulomb "collisions." Of course, this cross-section is simply the square of the distance of closest approach, r_c, defined in Eq. (1.17). Thus,

$$\tau \sim \frac{1}{n\,r_c^{\,2}\,\nu_t} \sim \frac{\epsilon_0^{\,2}\sqrt{m}\;T^{3/2}}{e^4\,n}. \tag{3.99}$$

The most significant feature of Eqs. (3.95) and (3.96) is the strong variation of the collision times with *temperature*. As the plasma gets hotter, the distance of closest approach gets smaller, so that both electrons and ions offer much smaller cross-sections for Coulomb collisions. The net result is that such collisions become far less frequent, and the collision times (*i.e.*, the mean times between 90° degree scattering events) get much longer. It follows that as plasmas are heated they become less collisional very rapidly.

The electron and ion fluid equations in a collisional plasma take the form [see Eqs. (3.55)–(3.57)]:

$$\frac{dn}{dt} + n\,\nabla\cdot\mathbf{V}_e = 0, \tag{3.100}$$

$$m_e n\,\frac{d\mathbf{V}_e}{dt} + \nabla p_e + \nabla\cdot\boldsymbol{\pi}_e + e n\,(\mathbf{E} + \mathbf{V}_e\times\mathbf{B}) = \mathbf{F}, \tag{3.101}$$

$$\frac{3}{2}\frac{dp_e}{dt} + \frac{5}{2}\,p_e\,\nabla\cdot\mathbf{V}_e + \boldsymbol{\pi}_e : \nabla\mathbf{V}_e + \nabla\cdot\mathbf{q}_e = W_e, \tag{3.102}$$

and

$$\frac{dn}{dt} + n\,\nabla\cdot\mathbf{V}_i = 0, \tag{3.103}$$

$$m_i n \frac{d\mathbf{V}_i}{dt} + \nabla p_i + \nabla\cdot\boldsymbol{\pi}_i - e n (\mathbf{E} + \mathbf{V}_i \times \mathbf{B}) = -\mathbf{F}, \tag{3.104}$$

$$\frac{3}{2}\frac{dp_i}{dt} + \frac{5}{2} p_i \nabla\cdot\mathbf{V}_i + \boldsymbol{\pi}_i : \nabla\mathbf{V}_i + \nabla\cdot\mathbf{q}_i = W_i, \tag{3.105}$$

respectively. Here, use has been made of the momentum conservation law (3.33). Equations (3.100)–(3.102) and (3.103)–(3.105) are called the *Braginskii equations*, since they were first obtained in a celebrated article by S.I. Braginskii.[3]

In the *unmagnetized* limit, which actually corresponds to

$$\Omega_i \tau_i, \;\; \Omega_e \tau_e \ll 1, \tag{3.106}$$

the standard two-Laguerre-polynomial Chapman-Enskog closure scheme yields

$$\mathbf{F} = \frac{n e \mathbf{j}}{\sigma_\parallel} - 0.71\, n \nabla T_e, \tag{3.107}$$

$$W_i = \frac{3\, m_e}{m_i} \frac{n (T_e - T_i)}{\tau_e}, \tag{3.108}$$

$$W_e = -W_i + \frac{\mathbf{j}\cdot\mathbf{F}}{n e} = -W_i + \frac{j^2}{\sigma_\parallel} - 0.71 \frac{\mathbf{j}\cdot\nabla T_e}{e}. \tag{3.109}$$

Here, $\mathbf{j} = -n e (\mathbf{V}_e - \mathbf{V}_i)$ is the net plasma current, and the *electrical conductivity* $\sigma_\parallel$ is given by

$$\sigma_\parallel = 1.96 \frac{n e^2 \tau_e}{m_e}. \tag{3.110}$$

In the above, use has been made of the conservation law (3.40).

Let us examine each of the above collisional terms, one by one. The first term on the right-hand side of Eq. (3.107) is a friction force due to the relative motion of electrons and ions, and obviously controls the electrical conductivity of the plasma. The form of this term is fairly easy to understand. The electrons lose their ordered velocity with respect to the ions, $\mathbf{U} = \mathbf{V}_e - \mathbf{V}_i$, in an electron collision time, τ_e, and consequently lose momentum $m_e \mathbf{U}$ per electron (which is given to the ions) in this time. This means that a frictional force $(m_e n/\tau_e)\,\mathbf{U} \sim n e \mathbf{j}/(n e^2 \tau_e/m_e)$ is exerted on the electrons. An equal and opposite force is exerted on the ions. Note that, since the Coulomb cross-section diminishes with increasing electron energy (*i.e.*, $\tau_e \sim T_e^{3/2}$), the conductivity of the fast electrons in the distribution function is higher than that of the slow electrons (since, $\sigma_\parallel \sim \tau_e$). Hence, electrical current in plasmas is carried predominately by the *fast* electrons. This effect has some important and interesting consequences.

One immediate consequence is the second term on the right-hand side of Eq. (3.107), which is called the *thermal force*. To understand the origin of a frictional force proportional to minus the gradient of the electron temperature, let us assume that the electron and ion

[3] S.I. Braginskii, *Transport Processes in a Plasma*, in Reviews of Plasma Physics (Consultants Bureau, New York NY, 1965), Vol. 1, p. 205.

fluids are at rest (*i.e.*, $V_e = V_i = 0$). It follows that the number of electrons moving from left to right (along the x-axis, say) and from right to left per unit time is exactly the same at a given point (coordinate x_0, say) in the plasma. As a result of electron-ion collisions, these fluxes experience frictional forces, $\mathbf{F}_-$ and $\mathbf{F}_+$, respectively, of order $m_e\, n\, \nu_e/\tau_e$, where ν_e is the electron thermal velocity. In a completely homogeneous plasma these forces balance exactly, and so there is zero net frictional force. Suppose, however, that the electrons coming from the right are, on average, hotter than those coming from the left. It follows that the frictional force $\mathbf{F}_+$ acting on the fast electrons coming from the right is less than the force $\mathbf{F}_-$ acting on the slow electrons coming from the left, since τ_e increases with electron temperature. As a result, there is a net frictional force acting to the left: *i.e.*, in the direction of $-\nabla T_e$.

Let us estimate the magnitude of the frictional force. At point x_0, collisions are experienced by electrons which have traversed distances of order a mean-free-path, $l_e \sim \nu_e\, \tau_e$. Thus, the electrons coming from the right originate from regions in which the temperature is approximately $l_e\, \partial T_e/\partial x$ greater than the regions from which the electrons coming from the left originate. Since the friction force is proportional to T_e^{-1}, the net force $\mathbf{F}_+ - \mathbf{F}_-$ is of order

$$\mathbf{F}_T \sim -\frac{l_e}{T_e}\frac{\partial T_e}{\partial x}\frac{m_e\, n\, \nu_e}{\tau_e} \sim -\frac{m_e\, {\nu_e}^2}{T_e}\, n\, \frac{\partial T_e}{\partial x} \sim -n\, \frac{\partial T_e}{\partial x}. \tag{3.111}$$

It must be emphasized that the thermal force is a direct consequence of *collisions*, despite the fact that the expression for the thermal force does not contain τ_e explicitly.

The term W_i, specified by Eq. (3.108), represents the rate at which energy is acquired by the ions due to collisions with the electrons. The most striking aspect of this term is its *smallness* (note that it is proportional to an inverse mass ratio, m_e/m_i). The smallness of W_i is a direct consequence of the fact that electrons are considerably lighter than ions. Consider the limit in which the ion mass is infinite, and the ions are at rest on average: *i.e.*, $V_i = 0$. In this case, collisions of electrons with ions take place *without any* exchange of energy. The electron velocities are randomized by the collisions, so that the energy associated with their ordered velocity, $\mathbf{U} = \mathbf{V}_e - \mathbf{V}_i$, is converted into heat energy in the electron fluid [this is represented by the second term on the extreme right-hand side of Eq. (3.109)]. However, the ion energy remains unchanged. Let us now assume that the ratio m_i/m_e is large, but finite, and that $U = 0$. If $T_e = T_i$, the ions and electrons are in thermal equilibrium, so no heat is exchanged between them. However, if $T_e > T_i$, heat is transferred from the electrons to the ions. As is well known, when a light particle collides with a heavy particle, the order of magnitude of the transferred energy is given by the mass ratio m_1/m_2, where m_1 is the mass of the lighter particle. For example, the mean fractional energy transferred in isotropic scattering is $2m_1/m_2$. Thus, we would expect the energy per unit time transferred from the electrons to the ions to be roughly

$$W_i \sim \frac{n}{\tau_e}\,\frac{2m_e}{m_i}\,\frac{3}{2}\,(T_e - T_i). \tag{3.112}$$

In fact, τ_e is defined so as to make the above estimate exact.

The term W_e, specified by Eq. (3.109), represents the rate at which energy is acquired by the electrons due to collisions with the ions, and consists of three terms. Not surprisingly, the first term is simply minus the rate at which energy is acquired by the ions due to collisions with the electrons. The second term represents the conversion of the ordered motion of the electrons, relative to the ions, into random motion (*i.e.*, heat) via collisions with the ions. Note that this term is positive definite, indicating that the randomization of the electron ordered motion gives rise to *irreversible* heat generation. Incidentally, this term is usually called the *ohmic heating* term. Finally, the third term represents the work done against the thermal force. Note that this term can be either positive or negative, depending on the direction of the current flow relative to the electron temperature gradient. This indicates that work done against the thermal force gives rise to *reversible* heat generation. There is an analogous effect in metals called the *Thomson effect*.

The electron and ion heat flux densities are given by

$$\mathbf{q}_e = -\kappa_\parallel^e \nabla T_e - 0.71 \frac{T_e \mathbf{j}}{e}, \tag{3.113}$$

$$\mathbf{q}_i = -\kappa_\parallel^i \nabla T_i, \tag{3.114}$$

respectively. The electron and ion *thermal conductivities* are written

$$\kappa_\parallel^e = 3.2 \frac{n \tau_e T_e}{m_e}, \tag{3.115}$$

$$\kappa_\parallel^i = 3.9 \frac{n \tau_i T_i}{m_i}, \tag{3.116}$$

respectively.

It follows, by comparison with Eqs. (3.71)–(3.76), that the first term on the right-hand side of Eq. (3.113) and the expression on the right-hand side of Eq. (3.114) represent straightforward random-walk heat diffusion, with frequency ν, and step-length l. Recall, that $\nu = \tau^{-1}$ is the collision frequency, and $l = \tau v_t$ is the mean-free-path. Note that the electron heat diffusivity is generally much greater than that of the ions, since $\kappa_\parallel^e/\kappa_\parallel^i \sim \sqrt{m_i/m_e}$, assuming that $T_e \sim T_i$.

The second term on the right-hand side of Eq. (3.113) describes a convective heat flux due to the motion of the electrons relative to the ions. To understand the origin of this flux, we need to recall that electric current in plasmas is carried predominately by the fast electrons in the distribution function. Suppose that U is non-zero. In the coordinate system in which V_e is zero, more fast electron move in the direction of $\mathbf{U}$, and more slow electrons move in the opposite direction. Although the electron fluxes are balanced in this frame of reference, the energy fluxes are not (since a fast electron possesses more energy than a slow electron), and heat flows in the direction of $\mathbf{U}$: *i.e.*, in the opposite direction to the electric current. The net heat flux density is of order $n T_e U$: *i.e.*, there is no near cancellation of the fluxes due to the fast and slow electrons. Like the thermal force, this effect depends on collisions despite the fact that the expression for the convective heat flux does not contain τ_e explicitly.

Finally, the electron and ion viscosity tensors take the form

$$(\pi_e)_{\alpha\beta} = -\eta_0^e \left(\frac{\partial V_\alpha}{\partial r_\beta} + \frac{\partial V_\beta}{\partial r_\alpha} - \frac{2}{3} \nabla\cdot\mathbf{V}\, \delta_{\alpha\beta} \right), \tag{3.117}$$

$$(\pi_i)_{\alpha\beta} = -\eta_0^i \left(\frac{\partial V_\alpha}{\partial r_\beta} + \frac{\partial V_\beta}{\partial r_\alpha} - \frac{2}{3} \nabla\cdot\mathbf{V}\, \delta_{\alpha\beta} \right), \tag{3.118}$$

respectively. Obviously, V_α refers to a Cartesian component of the electron fluid velocity in Eq. (3.117) and the ion fluid velocity in Eq. (3.118). Here, the electron and ion *viscosities* are given by

$$\eta_0^e = 0.73\, n\, \tau_e\, T_e, \tag{3.119}$$

$$\eta_0^i = 0.96\, n\, \tau_i\, T_i, \tag{3.120}$$

respectively. It follows, by comparison with Eqs. (3.70)–(3.76), that the above expressions correspond to straightforward random-walk diffusion of momentum, with frequency ν, and step-length l. Again, the electron diffusivity exceeds the ion diffusivity by the square root of a mass ratio (assuming $T_e \sim T_i$). However, the ion viscosity exceeds the electron viscosity by the same factor (recall that $\eta \sim n m \chi_\nu$): *i.e.*, $\eta_0^i/\eta_0^e \sim \sqrt{m_i/m_e}$. For this reason, the viscosity of a plasma is determined essentially by the ions. This is not surprising, since viscosity is the diffusion of momentum, and the ions possess nearly all of the momentum in a plasma by virtue of their large masses.

Let us now examine the *magnetized* limit,

$$\Omega_i\, \tau_i, \;\; \Omega_e\, \tau_e \gg 1, \tag{3.121}$$

in which the electron and ion gyroradii are much *smaller* than the corresponding mean-free-paths. In this limit, the two-Laguerre-polynomial Chapman-Enskog closure scheme yields

$$\mathbf{F} = n e \left(\frac{\mathbf{j}_\parallel}{\sigma_\parallel} + \frac{\mathbf{j}_\perp}{\sigma_\perp} \right) - 0.71\, n\, \nabla_\parallel T_e - \frac{3\, n}{2\, |\Omega_e|\, \tau_e}\, \mathbf{b} \times \nabla_\perp T_e, \tag{3.122}$$

$$W_i = \frac{3\, m_e}{m_i} \frac{n\, (T_e - T_i)}{\tau_e}, \tag{3.123}$$

$$W_e = -W_i + \frac{\mathbf{j}\cdot\mathbf{F}}{n\, e}. \tag{3.124}$$

Here, the *parallel electrical conductivity*, $\sigma_\parallel$, is given by Eq. (3.110), whereas the *perpendicular electrical conductivity*, $\sigma_\perp$, takes the form

$$\sigma_\perp = 0.51\, \sigma_\parallel = \frac{n\, e^2\, \tau_e}{m_e}. \tag{3.125}$$

Note that $\nabla_\parallel \cdots \equiv \mathbf{b}\, (\mathbf{b}\cdot\nabla \cdots)$ denotes a gradient parallel to the magnetic field, whereas $\nabla_\perp \equiv \nabla - \nabla_\parallel$ denotes a gradient perpendicular to the magnetic field. Likewise, $\mathbf{j}_\parallel \equiv \mathbf{b}\, (\mathbf{b}\cdot\mathbf{j})$

represents the component of the plasma current flowing parallel to the magnetic field, whereas $\mathbf{j}_\perp \equiv \mathbf{j} - \mathbf{j}_\parallel$ represents the perpendicular component of the plasma current.

We expect the presence of a strong magnetic field to give rise to a marked *anisotropy* in plasma properties between directions parallel and perpendicular to $\mathbf{B}$, because of the completely different motions of the constituent ions and electrons parallel and perpendicular to the field. Thus, not surprisingly, we find that the electrical conductivity perpendicular to the field is approximately half that parallel to the field [see Eqs. (3.122) and (3.125)]. The thermal force is unchanged (relative to the unmagnetized case) in the parallel direction, but is radically modified in the perpendicular direction. In order to understand the origin of the last term in Eq. (3.122), let us consider a situation in which there is a strong magnetic field along the z-axis, and an electron temperature gradient along the x-axis—see Fig. 3.1. The electrons gyrate in the x-y plane in circles of radius $\rho_e \sim v_e/|\Omega_e|$. At a given point, coordinate x_0, say, on the x-axis, the electrons that come from the right and the left have traversed distances of order ρ_e. Thus, the electrons from the right originate from regions where the electron temperature is of order $\rho_e\, \partial T_e/\partial x$ greater than the regions from which the electrons from the left originate. Since the friction force is proportional to T_e^{-1}, an unbalanced friction force arises, directed along the $-y$-axis—see Fig. 3.1. This direction corresponds to the direction of $-\mathbf{b}\times\nabla T_e$. Note that there is no friction force along the x-axis, since the x-directed fluxes are due to electrons which originate from regions where $x = x_0$. By analogy with Eq. (3.111), the magnitude of the perpendicular thermal force is

$$\mathbf{F}_{T\perp} \sim \frac{\rho_e}{T_e}\frac{\partial T_e}{\partial x}\frac{m_e\, n\, \nu_e}{\tau_e} \sim \frac{n}{|\Omega_e|\,\tau_e}\frac{\partial T_e}{\partial x}. \tag{3.126}$$

Note that the effect of a strong magnetic field on the perpendicular component of the thermal force is directly analogous to a well-known phenomenon in metals, called the *Nernst effect*.

In the magnetized limit, the electron and ion heat flux densities become

$$\begin{aligned}\mathbf{q}_e &= -\kappa_\parallel^e\,\nabla_\parallel T_e - \kappa_\perp^e\,\nabla_\perp T_e - \kappa_\times^e\,\mathbf{b}\times\nabla_\perp T_e\\ &\quad -0.71\,\frac{T_e\,\mathbf{j}_\parallel}{e} - \frac{3\,T_e}{2\,|\Omega_e|\,\tau_e\,e}\,\mathbf{b}\times\mathbf{j}_\perp,\end{aligned} \tag{3.127}$$

$$\mathbf{q}_i = -\kappa_\parallel^i\,\nabla_\parallel T_i - \kappa_\perp^i\,\nabla_\perp T_i + \kappa_\times^i\,\mathbf{b}\times\nabla_\perp T_i, \tag{3.128}$$

respectively. Here, the *parallel thermal conductivities* are given by Eqs. (3.115)–(3.116), and the *perpendicular thermal conductivities* take the form

$$\kappa_\perp^e = 4.7\,\frac{n\,T_e}{m_e\,\Omega_e^{\,2}\,\tau_e}, \tag{3.129}$$

$$\kappa_\perp^i = 2\,\frac{n\,T_i}{m_i\,\Omega_i^{\,2}\,\tau_i}. \tag{3.130}$$

Finally, the *cross thermal conductivities* are written

$$\kappa_\times^e = \frac{5\,n\,T_e}{2\,m_e\,|\Omega_e|}, \tag{3.131}$$

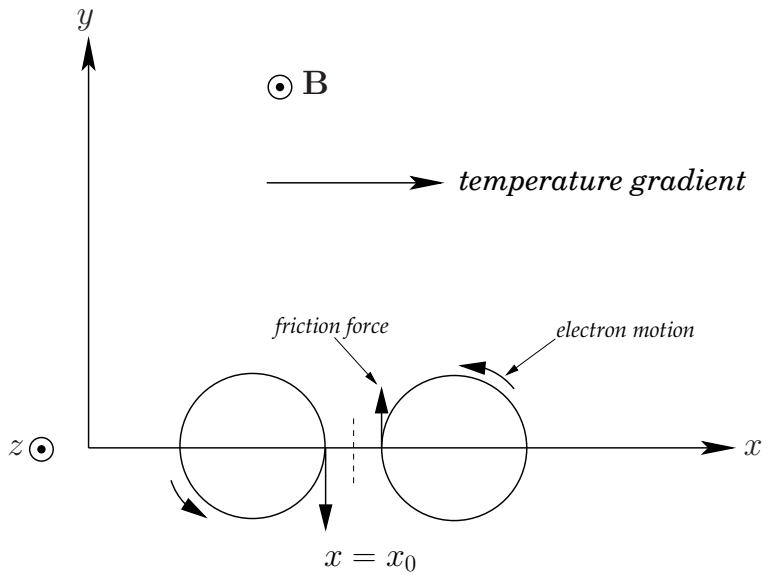

Figure 3.1: *Origin of the perpendicular thermal force in a magnetized plasma.*

$$\kappa_\times^i = \frac{5\,n\,T_i}{2\,m_i\,\Omega_i}. \tag{3.132}$$

The first two terms on the right-hand sides of Eqs. (3.127) and (3.128) correspond to *diffusive* heat transport by the electron and ion fluids, respectively. According to the first terms, the diffusive transport in the direction parallel to the magnetic field is exactly the same as that in the unmagnetized case: *i.e.*, it corresponds to collision-induced random-walk diffusion of the ions and electrons, with frequency ν, and step-length l. According to the second terms, the diffusive transport in the direction perpendicular to the magnetic field is *far smaller* than that in the parallel direction. In fact, it is smaller by a factor $(\rho/l)^2$, where ρ is the gyroradius, and l the mean-free-path. Note, that the perpendicular heat transport also corresponds to collision-induced random-walk diffusion of charged particles, but with frequency ν, and step-length ρ. Thus, it is the greatly reduced step-length in the perpendicular direction, relative to the parallel direction, which ultimately gives rise to the strong reduction in the perpendicular heat transport. If $T_e \sim T_i$, then the ion perpendicular heat diffusivity actually *exceeds* that of the electrons by the square root of a mass ratio: $\kappa_\perp^i/\kappa_\perp^e \sim \sqrt{m_i/m_e}$.

The third terms on the right-hand sides of Eqs. (3.127) and (3.128) correspond to heat fluxes which are perpendicular to both the magnetic field and the direction of the temperature gradient. In order to understand the origin of these terms, let us consider the ion flux. Suppose that there is a strong magnetic field along the z-axis, and an ion temperature gradient along the x-axis—see Fig. 3.2. The ions gyrate in the x-y plane in circles of radius $\rho_i \sim v_i/\Omega_i$, where v_i is the ion thermal velocity. At a given point, coordinate x_0, say, on the x-axis, the ions that come from the right and the left have traversed distances of order ρ_i. The ions from the right are clearly somewhat hotter than

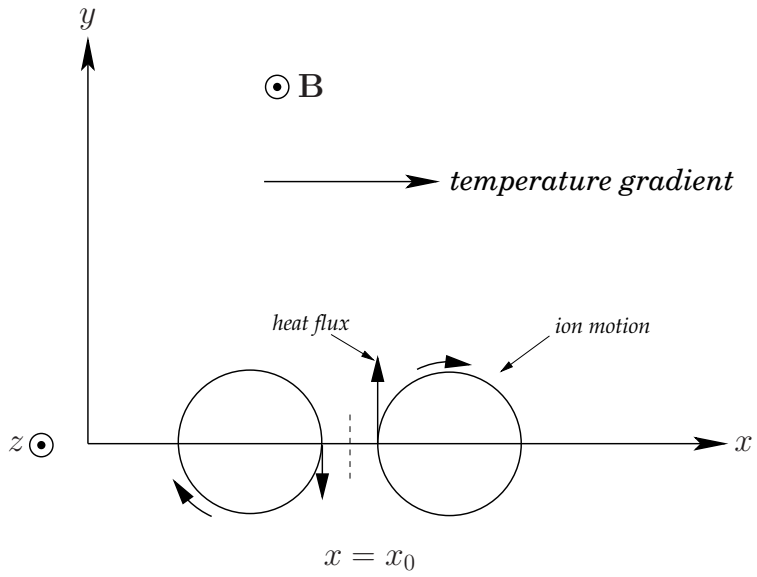

Figure 3.2: *Origin of the convective perpendicular heat flux in a magnetized plasma.*

those from the left. If the unidirectional particle fluxes, of order $n\,v_i$, are balanced, then the unidirectional heat fluxes, of order $n\,T_i\,v_i$, will have an unbalanced component of fractional order $(\rho_i/T_i)\partial T_i/\partial x$. As a result, there is a net heat flux in the $+y$-direction (*i.e.*, the direction of $\mathbf{b}\times\nabla T_i$). The magnitude of this flux is

$$q^i_\times \sim n\,v_i\,\rho_i\,\frac{\partial T_i}{\partial x} \sim \frac{n\,T_i}{m_i\,|\Omega_i|}\,\frac{\partial T_i}{\partial x}. \tag{3.133}$$

There is an analogous expression for the electron flux, except that the electron flux is in the opposite direction to the ion flux (because the electrons gyrate in the opposite direction to the ions). Note that both ion and electron fluxes transport heat *along isotherms*, and do not, therefore, give rise to any plasma heating.

The fourth and fifth terms on the right-hand side of Eq. (3.127) correspond to the convective component of the electron heat flux density, driven by motion of the electrons relative to the ions. It is clear from the fourth term that the convective flux parallel to the magnetic field is exactly the same as in the unmagnetized case [see Eq. (3.113)]. However, according to the fifth term, the convective flux is radically modified in the perpendicular direction. Probably the easiest method of explaining the fifth term is via an examination of Eqs. (3.107), (3.113), (3.122), and (3.127). There is clearly a very close connection between the electron thermal force and the convective heat flux. In fact, starting from general principles of the thermodynamics of irreversible processes, the so-called *Onsager principles*, it is possible to demonstrate that an electron frictional force of the form $\alpha\,(\nabla T_e)_\beta\,\mathbf{i}$ necessarily gives rise to an electron heat flux of the form $\alpha\,(T_e\,j_\beta/ne)\,\mathbf{i}$, where the subscript β corresponds to a general Cartesian component, and $\mathbf{i}$ is a unit vector. Thus, the fifth term on the right-hand side of Eq. (3.127) follows by *Onsager symmetry* from the

third term on the right-hand side of Eq. (3.122). This is one of many Onsager symmetries which occur in plasma transport theory.

In order to describe the viscosity tensor in a magnetized plasma, it is helpful to define the *rate-of-strain tensor*

$$W_{\alpha\beta} = \frac{\partial V_\alpha}{\partial r_\beta} + \frac{\partial V_\beta}{\partial r_\alpha} - \frac{2}{3}\,\nabla\cdot\mathbf{V}\,\delta_{\alpha\beta}. \tag{3.134}$$

Obviously, there is a separate rate-of-strain tensor for the electron and ion fluids. It is easily demonstrated that this tensor is zero if the plasma translates or rotates as a rigid body, or if it undergoes isotropic compression. Thus, the rate-of-strain tensor measures the *deformation* of plasma volume elements.

In a magnetized plasma, the viscosity tensor is best described as the sum of *five* component tensors,

$$\boldsymbol{\pi} = \sum_{n=0}^{4} \boldsymbol{\pi}_n, \tag{3.135}$$

where

$$\boldsymbol{\pi}_0 = -3\,\eta_0\,\left(\mathbf{bb} - \frac{1}{3}\,\mathbf{I}\right)\left(\mathbf{bb} - \frac{1}{3}\,\mathbf{I}\right) : \nabla\mathbf{V}, \tag{3.136}$$

with

$$\boldsymbol{\pi}_1 = -\eta_1\left[\mathbf{I}_\perp\cdot\mathbf{W}\cdot\mathbf{I}_\perp + \frac{1}{2}\,\mathbf{I}_\perp\,(\mathbf{b}\cdot\mathbf{W}\cdot\mathbf{b})\right], \tag{3.137}$$

and

$$\boldsymbol{\pi}_2 = -4\,\eta_1\,\left[\mathbf{I}_\perp\cdot\mathbf{W}\cdot\mathbf{bb} + \mathbf{bb}\cdot\mathbf{W}\cdot\mathbf{I}_\perp\right]. \tag{3.138}$$

plus

$$\boldsymbol{\pi}_3 = \frac{\eta_3}{2}\,\left[\mathbf{b}\times\mathbf{W}\cdot\mathbf{I}_\perp - \mathbf{I}_\perp\cdot\mathbf{W}\times\mathbf{b}\right], \tag{3.139}$$

and

$$\boldsymbol{\pi}_4 = 2\,\eta_3\,\left[\mathbf{b}\times\mathbf{W}\cdot\mathbf{bb} - \mathbf{bb}\cdot\mathbf{W}\times\mathbf{b}\right]. \tag{3.140}$$

Here, $\mathbf{I}$ is the identity tensor, and $\mathbf{I}_\perp = \mathbf{I} - \mathbf{bb}$. The above expressions are valid for both electrons and ions.

The tensor $\boldsymbol{\pi}_0$ describes what is known as *parallel viscosity*. This is a viscosity which controls the variation along magnetic field-lines of the velocity component parallel to field-lines. The parallel viscosity coefficients, η_0^e and η_0^i are specified in Eqs. (3.119)–(3.120). Note that the parallel viscosity is unchanged from the unmagnetized case, and is due to the collision-induced random-walk diffusion of particles, with frequency ν, and step-length l.

The tensors $\boldsymbol{\pi}_1$ and $\boldsymbol{\pi}_2$ describe what is known as *perpendicular viscosity*. This is a viscosity which controls the variation perpendicular to magnetic field-lines of the velocity components perpendicular to field-lines. The perpendicular viscosity coefficients are given by

$$\eta_1^e \;=\; 0.51\,\frac{n\,T_e}{\Omega_e^{\,2}\,\tau_e}, \tag{3.141}$$

$$\eta_1^i = \frac{3\,n\,T_i}{10\,{\Omega_i}^2\,\tau_i}. \tag{3.142}$$

Note that the perpendicular viscosity is *far smaller* than the parallel viscosity. In fact, it is smaller by a factor $(\rho/l)^2$. The perpendicular viscosity corresponds to collision-induced random-walk diffusion of particles, with frequency ν, and step-length ρ. Thus, it is the greatly reduced step-length in the perpendicular direction, relative to the parallel direction, which accounts for the smallness of the perpendicular viscosity compared to the parallel viscosity.

Finally, the tensors $\boldsymbol{\pi}_3$ and $\boldsymbol{\pi}_4$ describe what is known as *gyroviscosity*. This is not really viscosity at all, since the associated viscous stresses are always perpendicular to the velocity, implying that there is no dissipation (*i.e.*, viscous heating) associated with this effect. The gyroviscosity coefficients are given by

$$\eta_3^e = -\frac{n\,T_e}{2\,|\Omega_e|}, \tag{3.143}$$

$$\eta_3^i = \frac{n\,T_i}{2\,\Omega_i}. \tag{3.144}$$

The origin of gyroviscosity is very similar to the origin of the cross thermal conductivity terms in Eqs. (3.127)–(3.128). Note that both cross thermal conductivity and gyroviscosity are *independent* of the collision frequency.

3.9 Normalization of the Braginskii Equations

As we have just seen, the Braginskii equations contain terms which describe a very wide range of physical phenomena. For this reason, they are extremely complicated. Fortunately, however, it is not generally necessary to retain all of the terms in these equations when investigating a particular problem in plasma physics: *e.g.*, electromagnetic wave propagation through plasmas. In this section, we shall attempt to construct a systematic normalization scheme for the Braginskii equations which will, hopefully, enable us to determine which terms to keep, and which to discard, when investigating a particular aspect of plasma physics.

Let us consider a magnetized plasma. It is convenient to split the friction force $\mathbf{F}$ into a component $\mathbf{F}_U$ due to resistivity, and a component $\mathbf{F}_T$ corresponding to the thermal force. Thus,

$$\mathbf{F} = \mathbf{F}_U + \mathbf{F}_T, \tag{3.145}$$

where

$$\mathbf{F}_U = n e \left(\frac{\mathbf{j}_\parallel}{\sigma_\parallel} + \frac{\mathbf{j}_\perp}{\sigma_\perp} \right), \tag{3.146}$$

$$\mathbf{F}_T = -0.71\,n\,\nabla_\parallel T_e - \frac{3\,n}{2\,|\Omega_e|\,\tau_e}\,\mathbf{b}\times\nabla_\perp T_e. \tag{3.147}$$

Likewise, the electron collisional energy gain term W_e is split into a component $-W_i$ due to the energy lost to the ions (in the ion rest frame), a component W_U due to work done by the friction force $\mathbf{F}_U$, and a component W_T due to work done by the thermal force $\mathbf{F}_T$. Thus,

$$W_e = -W_i + W_U + W_T, \tag{3.148}$$

where

$$W_U = \frac{\mathbf{j}\cdot\mathbf{F}_U}{ne}, \tag{3.149}$$

$$W_T = \frac{\mathbf{j}\cdot\mathbf{F}_T}{ne}. \tag{3.150}$$

Finally, it is helpful to split the electron heat flux density $\mathbf{q}_e$ into a diffusive component $\mathbf{q}_{Te}$ and a convective component $\mathbf{q}_{Ue}$. Thus,

$$\mathbf{q}_e = \mathbf{q}_{Te} + \mathbf{q}_{Ue}, \tag{3.151}$$

where

$$\mathbf{q}_{Te} = -\kappa_\parallel^e\,\nabla_\parallel T_e - \kappa_\perp^e\,\nabla_\perp T_e - \kappa_\times^e\,\mathbf{b}\times\nabla_\perp T_e, \tag{3.152}$$

$$\mathbf{q}_{Ue} = 0.71\,\frac{T_e\,\mathbf{j}_\parallel}{e} - \frac{3\,T_e}{2\,|\Omega_e|\,\tau_e\,e}\,\mathbf{b}\times\mathbf{j}_\perp. \tag{3.153}$$

Let us, first of all, consider the electron fluid equations, which can be written:

$$\frac{dn}{dt} + n\,\nabla\cdot\mathbf{V}_e = 0, \tag{3.154}$$

$$m_e n\,\frac{d\mathbf{V}_e}{dt} + \nabla p_e + \nabla\cdot\boldsymbol{\pi}_e + e n\,(\mathbf{E}+\mathbf{V}_e\times\mathbf{B}) = \mathbf{F}_U + \mathbf{F}_T, \tag{3.155}$$

$$\frac{3}{2}\frac{dp_e}{dt} + \frac{5}{2}\,p_e\,\nabla\cdot\mathbf{V}_e + \boldsymbol{\pi}_e : \nabla\mathbf{V}_e + \nabla\cdot\mathbf{q}_{Te} + \nabla\cdot\mathbf{q}_{Ue} = -W_i + W_U + W_T. \tag{3.156}$$

Let $\bar{n}$, $\bar{v}_e$, $\bar{l}_e$, $\bar{B}$, and $\bar{\rho}_e = \bar{v}_e/(e\bar{B}/m_e)$, be typical values of the particle density, the electron thermal velocity, the electron mean-free-path, the magnetic field-strength, and the electron gyroradius, respectively. Suppose that the typical electron flow velocity is $\lambda_e\,\bar{v}_e$, and the typical variation length-scale is L. Let

$$\delta_e = \frac{\bar{\rho}_e}{L}, \tag{3.157}$$

$$\zeta_e = \frac{\bar{\rho}_e}{\bar{l}_e}, \tag{3.158}$$

$$\mu = \sqrt{\frac{m_e}{m_i}}. \tag{3.159}$$

All three of these parameters are assumed to be *small* compared to unity.

We define the following normalized quantities: $\hat{n} = n/\bar{n}$, $\hat{v}_e = v_e/\bar{v}_e$, $\hat{\mathbf{r}} = \mathbf{r}/L$, $\hat{\nabla} = L\,\nabla$, $\hat{t} = \lambda_e\,\bar{\nu}_e\,t/L$, $\hat{\mathbf{V}}_e = \mathbf{V}_e/\lambda_e\,\bar{\nu}_e$, $\hat{\mathbf{B}} = \mathbf{B}/\bar{B}$, $\hat{\mathbf{E}} = \mathbf{E}/\lambda_e\,\bar{\nu}_e\,\bar{B}$, $\hat{\mathbf{U}} = \mathbf{U}/(1+\lambda_e^{\,2})\,\delta_e\,\bar{\nu}_e$, plus $\hat{p}_e = p_e/m_e\,\bar{n}\,\bar{\nu}_e^{\,2}$, $\hat{\boldsymbol{\pi}}_e = \boldsymbol{\pi}_e/\lambda_e\,\delta_e\,\zeta_e^{-1}\,m_e\,\bar{n}\,\bar{\nu}_e^{\,2}$, $\hat{\mathbf{q}}_{Te} = \mathbf{q}_{Te}/\delta_e\,\zeta_e^{-1}\,m_e\,\bar{n}\,\bar{\nu}_e^{\,3}$, $\hat{\mathbf{q}}_{Ue} = \mathbf{q}_{Ue}/(1+\lambda_e^{\,2})\,\delta_e\,m_e\,\bar{n}\,\bar{\nu}_e^{\,3}$, $\hat{\mathbf{F}}_U = \mathbf{F}_U/(1+\lambda_e^{\,2})\,\zeta_e\,m_e\,\bar{n}\,\bar{\nu}_e^{\,2}/L$, $\hat{\mathbf{F}}_T = \mathbf{F}_T/m_e\,\bar{n}\,\bar{\nu}_e^{\,2}/L$, $\hat{W}_i = W_i/\delta_e^{-1}\,\zeta_e\,\mu^2\,m_e\,\bar{n}\,\bar{\nu}_e^{\,3}/L$, $\hat{W}_U = W_U/(1+\lambda_e^{\,2})^2\,\delta_e\,\zeta_e\,m_e\,\bar{n}\,\bar{\nu}_e^{\,3}/L$, $\hat{W}_T = W_T/(1+\lambda_e^{\,2})\,\delta_e\,m_e\,\bar{n}\,\bar{\nu}_e^{\,3}/L$.

The normalization procedure is designed to make all hatted quantities $O(1)$. The normalization of the electric field is chosen such that the $\mathbf{E}\times\mathbf{B}$ velocity is of order the electron fluid velocity. Note that the parallel viscosity makes an $O(1)$ contribution to $\hat{\boldsymbol{\pi}}_e$, whereas the gyroviscosity makes an $O(\zeta_e)$ contribution, and the perpendicular viscosity only makes an $O(\zeta_e^{\,2})$ contribution. Likewise, the parallel thermal conductivity makes an $O(1)$ contribution to $\hat{\mathbf{q}}_{Te}$, whereas the cross conductivity makes an $O(\zeta_e)$ contribution, and the perpendicular conductivity only makes an $O(\zeta_e^{\,2})$ contribution. Similarly, the parallel components of $\mathbf{F}_T$ and $\mathbf{q}_{Ue}$ are $O(1)$, whereas the perpendicular components are $O(\zeta_e)$.

The normalized electron fluid equations take the form:

$$\frac{d\hat{n}}{d\hat{t}} + \hat{n}\,\hat{\nabla}\cdot\hat{\mathbf{V}}_e = 0, \tag{3.160}$$

$$\begin{aligned}\lambda_e^{\,2}\,\delta_e\,\hat{n}\,\frac{d\hat{\mathbf{V}}_e}{d\hat{t}} + \delta_e\,\hat{\nabla}\hat{p}_e + \lambda_e\,\delta_e^{\,2}\,\zeta_e^{-1}\,\hat{\nabla}\cdot\hat{\boldsymbol{\pi}}_e & \\ +\lambda_e\,\hat{n}\,(\hat{\mathbf{E}} + \hat{\mathbf{V}}_e\times\hat{\mathbf{B}}) &= (1+\lambda_e^{\,2})\,\delta_e\,\zeta_e\,\hat{\mathbf{F}}_U + \delta_e\,\hat{\mathbf{F}}_T,\end{aligned} \tag{3.161}$$

$$\begin{aligned}\lambda_e\,\frac{3}{2}\frac{d\hat{p}_e}{d\hat{t}} + \lambda_e\,\frac{5}{2}\,\hat{p}_e\,\hat{\nabla}\cdot\hat{\mathbf{V}}_e + \lambda_e^{\,2}\,\delta_e\,\zeta_e^{-1}\,\hat{\boldsymbol{\pi}}_e : \hat{\nabla}\cdot\hat{\mathbf{V}}_e & \\ +\delta_e\,\zeta_e^{-1}\,\hat{\nabla}\cdot\hat{\mathbf{q}}_{Te} + (1+\lambda_e^{\,2})\,\delta_e\,\hat{\nabla}\cdot\hat{\mathbf{q}}_{Ue} &= -\delta_e^{-1}\,\zeta_e\,\mu^2\,\hat{W}_i \\ &\quad +(1+\lambda_e^{\,2})^2\,\delta_e\,\zeta_e\,\hat{W}_U \\ &\quad +(1+\lambda_e^{\,2})\,\delta_e\,\hat{W}_T.\end{aligned} \tag{3.162}$$

Note that the only large or small quantities in these equations are the parameters λ_e, δ_e, ζ_e, and μ. Here, $d/d\hat{t} \equiv \partial/\partial\hat{t} + \hat{\mathbf{V}}_e\cdot\hat{\nabla}$. It is assumed that $T_e \sim T_i$.

Let us now consider the ion fluid equations, which can be written:

$$\frac{dn}{dt} + n\,\nabla\cdot\mathbf{V}_i = 0, \tag{3.163}$$

$$m_i n\,\frac{d\mathbf{V}_i}{dt} + \nabla p_i + \nabla\cdot\boldsymbol{\pi}_i - e n\,(\mathbf{E} + \mathbf{V}_i\times\mathbf{B}) = -\mathbf{F}_U - \mathbf{F}_T, \tag{3.164}$$

$$\frac{3}{2}\frac{dp_i}{dt} + \frac{5}{2}\,p_i\,\nabla\cdot\mathbf{V}_i + \boldsymbol{\pi}_i : \nabla\mathbf{V}_i + \nabla\cdot\mathbf{q}_i = W_i. \tag{3.165}$$

It is convenient to adopt a normalization scheme for the ion equations which is similar to, but independent of, that employed to normalize the electron equations. Let $\bar{n}$, $\bar{v}_i$, $\bar{l}_i$, $\bar{B}$, and $\bar{\rho}_i = \bar{v}_i/(e\bar{B}/m_i)$, be typical values of the particle density, the ion thermal velocity, the ion mean-free-path, the magnetic field-strength, and the ion gyroradius, respectively.

Suppose that the typical ion flow velocity is $\lambda_i\,\bar{v}_i$, and the typical variation length-scale is L. Let

$$\delta_i = \frac{\bar{\rho}_i}{L}, \tag{3.166}$$

$$\zeta_i = \frac{\bar{\rho}_i}{\bar{l}_i}, \tag{3.167}$$

$$\mu = \sqrt{\frac{m_e}{m_i}}. \tag{3.168}$$

All three of these parameters are assumed to be *small* compared to unity.

We define the following normalized quantities: $\hat{n} = n/\bar{n}$, $\hat{v}_i = v_i/\bar{v}_i$, $\hat{\mathbf{r}} = \mathbf{r}/L$, $\hat{\nabla} = L\,\nabla$, $\hat{t} = \lambda_i\,\bar{v}_i\,t/L$, $\hat{\mathbf{V}}_i = \mathbf{V}_i/\lambda_i\,\bar{v}_i$, $\hat{\mathbf{B}} = \mathbf{B}/\bar{B}$, $\hat{\mathbf{E}} = \mathbf{E}/\lambda_i\,\bar{v}_i\,\bar{B}$, $\hat{\mathbf{U}} = \mathbf{U}/(1+\lambda_i^{\,2})\,\delta_i\,\bar{v}_i$, $\hat{p}_i = p_i/m_i\,\bar{n}\,\bar{v}_i^{\,2}$, $\hat{\boldsymbol{\pi}}_i = \boldsymbol{\pi}_i/\lambda_i\,\delta_i\,\zeta_i^{-1}\,m_i\,\bar{n}\,\bar{v}_i^{\,2}$, $\hat{\mathbf{q}}_i = \mathbf{q}_i/\delta_i\,\zeta_i^{-1}\,m_i\,\bar{n}\,\bar{v}_i^{\,3}$, $\hat{\mathbf{F}}_U = \mathbf{F}_U/(1+\lambda_i^{\,2})\,\zeta_i\,\mu\,m_i\,\bar{n}\,\bar{v}_i^{\,2}/L$, $\hat{\mathbf{F}}_T = \mathbf{F}_T/m_i\,\bar{n}\,\bar{v}_i^{\,2}/L$, $\hat{W}_i = W_i/\delta_i^{-1}\,\zeta_i\,\mu\,m_i\,\bar{n}\,\bar{v}_i^{\,3}/L$.

As before, the normalization procedure is designed to make all hatted quantities O(1). The normalization of the electric field is chosen such that the $\mathbf{E}\times\mathbf{B}$ velocity is of order the ion fluid velocity. Note that the parallel viscosity makes an O(1) contribution to $\hat{\boldsymbol{\pi}}_i$, whereas the gyroviscosity makes an $O(\zeta_i)$ contribution, and the perpendicular viscosity only makes an $O(\zeta_i^{\,2})$ contribution. Likewise, the parallel thermal conductivity makes an O(1) contribution to $\hat{\mathbf{q}}_i$, whereas the cross conductivity makes an $O(\zeta_i)$ contribution, and the perpendicular conductivity only makes an $O(\zeta_i^{\,2})$ contribution. Similarly, the parallel component of $\mathbf{F}_T$ is O(1), whereas the perpendicular component is $O(\zeta_i\,\mu)$.

The normalized ion fluid equations take the form:

$$\frac{d\hat{n}}{d\hat{t}} + \hat{n}\,\hat{\nabla}\cdot\hat{\mathbf{V}}_i = 0, \tag{3.169}$$

$$\lambda_i^{\,2}\,\delta_i\,\hat{n}\,\frac{d\hat{\mathbf{V}}_i}{d\hat{t}} + \delta_i\,\hat{\nabla}\hat{p}_i + \lambda_i\,\delta_i^{\,2}\,\zeta_i^{-1}\,\hat{\nabla}\cdot\hat{\boldsymbol{\pi}}_i - \lambda_i\,\hat{n}\,(\hat{\mathbf{E}} + \hat{\mathbf{V}}_i\times\hat{\mathbf{B}}) = -(1+\lambda_i^{\,2})\,\delta_i\,\zeta_i\,\mu\,\hat{\mathbf{F}}_U - \delta_i\,\hat{\mathbf{F}}_T, \tag{3.170}$$

$$\lambda_i\,\frac{3}{2}\frac{d\hat{p}_i}{d\hat{t}} + \lambda_i\,\frac{5}{2}\,\hat{p}_i\,\hat{\nabla}\cdot\hat{\mathbf{V}}_i + \lambda_i^{\,2}\,\delta_i\,\zeta_i^{-1}\,\hat{\boldsymbol{\pi}}_i : \hat{\nabla}\cdot\hat{\mathbf{V}}_i + \delta_i\,\zeta_i^{-1}\,\hat{\nabla}\cdot\hat{\mathbf{q}}_i = \delta_i^{-1}\,\zeta_i\,\mu\,\hat{W}_i. \tag{3.171}$$

Note that the only large or small quantities in these equations are the parameters λ_i, δ_i, ζ_i, and μ. Here, $d/d\hat{t} \equiv \partial/\partial\hat{t} + \hat{\mathbf{V}}_i\cdot\hat{\nabla}$.

Let us adopt the ordering

$$\delta_e, \delta_i \ll \zeta_e, \zeta_i, \quad \mu \ll 1, \tag{3.172}$$

which is appropriate to a collisional, *highly magnetized* plasma. In the first stage of our ordering procedure, we shall treat δ_e and δ_i as small parameters, and ζ_e, ζ_i, and μ as O(1). In the second stage, we shall take note of the smallness of ζ_e, ζ_i, and μ. Note that the parameters λ_e and λ_i are "free ranging:" *i.e.*, they can be either large, small, or O(1). In

the initial stage of the ordering procedure, the ion and electron normalization schemes we have adopted become essentially identical [since $\mu \sim O(1)$], and it is convenient to write

$$\lambda_e \sim \lambda_i \sim \lambda, \tag{3.173}$$

$$\delta_e \sim \delta_i \sim \delta, \tag{3.174}$$

$$V_e \sim V_i \sim V, \tag{3.175}$$

$$\nu_e \sim \nu_i \sim \nu_t, \tag{3.176}$$

$$\Omega_e \sim \Omega_i \sim \Omega. \tag{3.177}$$

There are *three fundamental orderings* in plasma fluid theory. These are analogous to the three orderings in neutral gas fluid theory discussed in Sect. 3.7.

The first ordering is

$$\lambda \sim \delta^{-1}. \tag{3.178}$$

This corresponds to

$$V \gg \nu_t. \tag{3.179}$$

In other words, the fluid velocities are much greater than the thermal velocities. We also have

$$\frac{V}{L} \sim \Omega. \tag{3.180}$$

Here, V/L is conventionally termed the *transit frequency*, and is the frequency with which fluid elements traverse the system. It is clear that the transit frequencies are of order the gyrofrequencies in this ordering. Keeping only the largest terms in Eqs. (3.160)–(3.162) and (3.169)–(3.171), the Braginskii equations reduce to (in unnormalized form):

$$\frac{dn}{dt} + n\,\nabla\cdot\mathbf{V}_e = 0, \tag{3.181}$$

$$m_e n \frac{d\mathbf{V}_e}{dt} + e n\,(\mathbf{E} + \mathbf{V}_e \times \mathbf{B}) = [\zeta]\,\mathbf{F}_U, \tag{3.182}$$

and

$$\frac{dn}{dt} + n\,\nabla\cdot\mathbf{V}_i = 0, \tag{3.183}$$

$$m_i n \frac{d\mathbf{V}_i}{dt} - e n\,(\mathbf{E} + \mathbf{V}_i \times \mathbf{B}) = -[\zeta]\,\mathbf{F}_U. \tag{3.184}$$

The factors in square brackets are just to remind us that the terms they precede are smaller than the other terms in the equations (by the corresponding factors inside the brackets).

Equations (3.181)–(3.182) and (3.183)–(3.184) are called the *cold-plasma equations*, because they can be obtained from the Braginskii equations by formally taking the limit $T_e, T_i \rightarrow 0$. Likewise, the ordering (3.178) is called the *cold-plasma approximation*. Note that the cold-plasma approximation applies not only to cold plasmas, but also to very *fast*

disturbances which propagate through conventional plasmas. In particular, the cold-plasma equations provide a good description of the propagation of *electromagnetic waves* through plasmas. After all, electromagnetic waves generally have very high velocities (*i.e.*, $V \sim c$), which they impart to plasma fluid elements, so there is usually no difficulty satisfying the inequality (3.179).

Note that the electron and ion pressures can be neglected in the cold-plasma limit, since the thermal velocities are much smaller than the fluid velocities. It follows that there is no need for an electron or ion energy evolution equation. Furthermore, the motion of the plasma is so fast, in this limit, that relatively slow "transport" effects, such as viscosity and thermal conductivity, play no role in the cold-plasma fluid equations. In fact, the only collisional effect which appears in these equations is *resistivity*.

The second ordering is

$$\lambda \sim 1, \tag{3.185}$$

which corresponds to

$$V \sim v_t. \tag{3.186}$$

In other words, the fluid velocities are of order the thermal velocities. Keeping only the largest terms in Eqs. (3.160)–(3.162) and (3.169)–(3.171), the Braginskii equations reduce to (in unnormalized form):

$$\frac{dn}{dt} + n\,\nabla\cdot\mathbf{V}_e = 0, \tag{3.187}$$

$$m_e n \frac{d\mathbf{V}_e}{dt} + \nabla p_e + [\delta^{-1}]\, e n\,(\mathbf{E} + \mathbf{V}_e \times \mathbf{B}) = [\zeta]\,\mathbf{F}_U + \mathbf{F}_T, \tag{3.188}$$

$$\frac{3}{2}\frac{dp_e}{dt} + \frac{5}{2}\,p_e\,\nabla\cdot\mathbf{V}_e = -[\delta^{-1}\,\zeta\,\mu^2]\,W_i, \tag{3.189}$$

and

$$\frac{dn}{dt} + n\,\nabla\cdot\mathbf{V}_i = 0, \tag{3.190}$$

$$m_i n \frac{d\mathbf{V}_i}{dt} + \nabla p_i - [\delta^{-1}]\, e n\,(\mathbf{E} + \mathbf{V}_i \times \mathbf{B}) = -[\zeta]\,\mathbf{F}_U - \mathbf{F}_T, \tag{3.191}$$

$$\frac{3}{2}\frac{dp_i}{dt} + \frac{5}{2}\,p_i\,\nabla\cdot\mathbf{V}_i = [\delta^{-1}\,\zeta\,\mu^2]\,W_i. \tag{3.192}$$

Again, the factors in square brackets remind us that the terms they precede are larger, or smaller, than the other terms in the equations.

Equations (3.187)–(3.189) and (3.190)–(3.191) are called the *magnetohydrodynamical equations*, or *MHD equations*, for short. Likewise, the ordering (3.185) is called the *MHD approximation*. The MHD equations are conventionally used to study macroscopic plasma instabilities possessing relatively fast growth-rates: *e.g.*, "sausage" modes, "kink" modes.

Note that the electron and ion pressures cannot be neglected in the MHD limit, since the fluid velocities are of order the thermal velocities. Thus, electron and ion energy

evolution equations are needed in this limit. However, MHD motion is sufficiently fast that "transport" effects, such as viscosity and thermal conductivity, are too slow to play a role in the MHD equations. In fact, the only collisional effects which appear in these equations are resistivity, the thermal force, and electron-ion collisional energy exchange.

The final ordering is

$$\lambda \sim \delta, \tag{3.193}$$

which corresponds to

$$V \sim \delta\, \nu_t \sim \nu_d, \tag{3.194}$$

where ν_d is a typical drift (*e.g.*, a curvature or grad-B drift—see Sect. 2) velocity. In other words, the fluid velocities are of order the drift velocities. Keeping only the largest terms in Eqs. (3.113) and (3.116), the Braginskii equations reduce to (in unnormalized form):

$$\frac{dn}{dt} + n\,\nabla\cdot\mathbf{V}_e = 0, \tag{3.195}$$

$$\begin{aligned} m_e n \frac{d\mathbf{V}_e}{dt} + [\delta^{-2}]\,\nabla p_e + [\zeta^{-1}]\,\nabla\cdot\boldsymbol{\pi}_e & \\ +[\delta^{-2}]\, en\,(\mathbf{E} + \mathbf{V}_e\times\mathbf{B}) &= [\delta^{-2}\,\zeta]\,\mathbf{F}_U + [\delta^{-2}]\,\mathbf{F}_T, \end{aligned} \tag{3.196}$$

$$\begin{aligned} \frac{3}{2}\frac{dp_e}{dt} + \frac{5}{2}\,p_e\,\nabla\cdot\mathbf{V}_e + [\zeta^{-1}]\,\nabla\cdot\mathbf{q}_{Te} + \nabla\cdot\mathbf{q}_{Ue} &= -[\delta^{-2}\,\zeta\,\mu^2]\,W_i \\ &\quad +[\zeta]\,W_U + W_T, \end{aligned} \tag{3.197}$$

and

$$\frac{dn}{dt} + n\,\nabla\cdot\mathbf{V}_i = 0, \tag{3.198}$$

$$\begin{aligned} m_i n \frac{d\mathbf{V}_i}{dt} + [\delta^{-2}]\,\nabla p_i + [\zeta^{-1}]\,\nabla\cdot\boldsymbol{\pi}_i & \\ -[\delta^{-2}]\, en\,(\mathbf{E} + \mathbf{V}_i\times\mathbf{B}) &= -[\delta^{-2}\,\zeta]\,\mathbf{F}_U - [\delta^{-2}]\,\mathbf{F}_T, \end{aligned} \tag{3.199}$$

$$\frac{3}{2}\frac{dp_i}{dt} + \frac{5}{2}\,p_i\,\nabla\cdot\mathbf{V}_i + [\zeta^{-1}]\,\nabla\cdot\mathbf{q}_i = [\delta^{-2}\,\zeta\,\mu^2]\,W_i. \tag{3.200}$$

As before, the factors in square brackets remind us that the terms they precede are larger, or smaller, than the other terms in the equations.

Equations (3.195)–(3.198) and (3.198)–(3.200) are called the *drift equations*. Likewise, the ordering (3.193) is called the *drift approximation*. The drift equations are conventionally used to study equilibrium evolution, and the slow growing "microinstabilities" which are responsible for turbulent transport in tokamaks. It is clear that virtually all of the original terms in the Braginskii equations must be retained in this limit.

In the following sections, we investigate the cold-plasma equations, the MHD equations, and the drift equations, in more detail.

3.10 Cold-Plasma Equations

Previously, we used the smallness of the magnetization parameter δ to derive the cold-plasma equations:

$$\frac{\partial n}{\partial t} + \nabla\cdot(n\,\mathbf{V}_e) = 0, \tag{3.201}$$

$$m_e n\,\frac{\partial \mathbf{V}_e}{\partial t} + m_e n\,(\mathbf{V}_e\cdot\nabla)\mathbf{V}_e + e n\,(\mathbf{E}+\mathbf{V}_e\times\mathbf{B}) = [\zeta]\,\mathbf{F}_U, \tag{3.202}$$

and

$$\frac{\partial n}{\partial t} + \nabla\cdot(n\,\mathbf{V}_i) = 0, \tag{3.203}$$

$$m_i n\,\frac{\partial \mathbf{V}_i}{\partial t} + m_i n\,(\mathbf{V}_i\cdot\nabla)\mathbf{V}_i - e n\,(\mathbf{E}+\mathbf{V}_i\times\mathbf{B}) = -[\zeta]\,\mathbf{F}_U. \tag{3.204}$$

Let us now use the smallness of the mass ratio m_e/m_i to further simplify these equations. In particular, we would like to write the electron and ion fluid velocities in terms of the centre-of-mass velocity,

$$\mathbf{V} = \frac{m_i\,\mathbf{V}_i + m_e\,\mathbf{V}_e}{m_i + m_e}, \tag{3.205}$$

and the plasma current

$$\mathbf{j} = -n e\,\mathbf{U}, \tag{3.206}$$

where $\mathbf{U} = \mathbf{V}_e - \mathbf{V}_i$. According to the ordering scheme adopted in the previous section, $U \sim V_e \sim V_i$ in the cold-plasma limit. We shall continue to regard the mean-free-path parameter ζ as $O(1)$.

It follows from Eqs. (3.205) and (3.206) that

$$\mathbf{V}_i \simeq \mathbf{V} + O(m_e/m_i), \tag{3.207}$$

and

$$\mathbf{V}_e \simeq \mathbf{V} - \frac{\mathbf{j}}{n e} + O\left(\frac{m_e}{m_i}\right). \tag{3.208}$$

Equations (3.201), (3.203), (3.207), and (3.208) yield the *continuity equation*:

$$\frac{dn}{dt} + n\,\nabla\cdot\mathbf{V} = 0, \tag{3.209}$$

where $d/dt \equiv \partial/\partial t + \mathbf{V}\cdot\nabla$. Here, use has been made of the fact that $\nabla\cdot\mathbf{j} = 0$ in a quasi-neutral plasma.

Equations (3.202) and (3.204) can be summed to give the *equation of motion*:

$$m_i n\,\frac{d\mathbf{V}}{dt} - \mathbf{j}\times\mathbf{B} \simeq 0. \tag{3.210}$$

Finally, Eqs. (3.202), (3.207), and (3.208) can be combined and to give a modified *Ohm's law*:

$$\mathbf{E}+\mathbf{V}\times\mathbf{B} \simeq \frac{\mathbf{F}_U}{ne}+\frac{\mathbf{j}\times\mathbf{B}}{ne}+\frac{m_e}{ne^2}\frac{d\mathbf{j}}{dt} +\frac{m_e}{ne^2}(\mathbf{j}\cdot\nabla)\mathbf{V}-\frac{m_e}{n^2e^3}(\mathbf{j}\cdot\nabla)\mathbf{j}. \tag{3.211}$$

The first term on the right-hand side of the above equation corresponds to *resistivity*, the second corresponds to the *Hall effect*, the third corresponds to the effect of *electron inertia*, and the remaining terms are usually negligible.

3.11 MHD Equations

The MHD equations take the form:

$$\frac{\partial n}{\partial t}+\nabla\cdot(n\,\mathbf{V}_e) = 0, \tag{3.212}$$

$$m_e n\frac{\partial \mathbf{V}_e}{\partial t}+m_e n(\mathbf{V}_e\cdot\nabla)\mathbf{V}_e+\nabla p_e \tag{3.213}$$

$$+[\delta^{-1}]\,en(\mathbf{E}+\mathbf{V}_e\times\mathbf{B}) = [\zeta]\,\mathbf{F}_U+\mathbf{F}_T,$$

$$\frac{3}{2}\frac{\partial p_e}{\partial t}+\frac{3}{2}(\mathbf{V}_e\cdot\nabla)\,p_e+\frac{5}{2}\,p_e\,\nabla\cdot\mathbf{V}_e = -[\delta^{-1}\,\zeta\,\mu^2]\,W_i, \tag{3.214}$$

and

$$\frac{\partial n}{\partial t}+\nabla\cdot(n\,\mathbf{V}_i) = 0, \tag{3.215}$$

$$m_i n\frac{\partial \mathbf{V}_i}{\partial t}+m_i n(\mathbf{V}_i\cdot\nabla)\mathbf{V}_i+\nabla p_i \tag{3.216}$$

$$-[\delta^{-1}]\,en(\mathbf{E}+\mathbf{V}_i\times\mathbf{B}) = -[\zeta]\,\mathbf{F}_U-\mathbf{F}_T,$$

$$\frac{3}{2}\frac{\partial p_i}{\partial t}+\frac{3}{2}(\mathbf{V}_i\cdot\nabla)\,p_i+\frac{5}{2}\,p_i\,\nabla\cdot\mathbf{V}_i = [\delta^{-1}\,\zeta\,\mu^2]\,W_i. \tag{3.217}$$

These equations can also be simplified by making use of the smallness of the mass ratio m_e/m_i. Now, according to the ordering adopted in Sect. 3.9, $U\sim\delta\,V_e\sim\delta\,V_i$ in the MHD limit. It follows from Eqs. (3.207) and (3.208) that

$$\mathbf{V}_i\simeq\mathbf{V}+O(m_e/m_i), \tag{3.218}$$

and

$$\mathbf{V}_e\simeq\mathbf{V}-[\delta]\,\frac{\mathbf{j}}{ne}+O\left(\frac{m_e}{m_i}\right). \tag{3.219}$$

The main point, here, is that in the MHD limit the velocity difference between the electron and ion fluids is relatively small.

Equations (3.212) and (3.215) yield the *continuity equation*:

$$\frac{dn}{dt} + n\,\nabla\cdot\mathbf{V} = 0, \tag{3.220}$$

where $d/dt \equiv \partial/\partial t + \mathbf{V}\cdot\nabla$.

Equations (3.213) and (3.216) can be summed to give the *equation of motion*:

$$m_i n \frac{d\mathbf{V}}{dt} + \nabla p - \mathbf{j}\times\mathbf{B} \simeq 0. \tag{3.221}$$

Here, $p = p_e + p_i$ is the total pressure. Note that all terms in the above equation are the same order in δ.

The $O(\delta^{-1})$ components of Eqs. (3.213) and (3.216) yield the *Ohm's law*:

$$\mathbf{E} + \mathbf{V}\times\mathbf{B} \simeq 0. \tag{3.222}$$

This is sometimes called the *perfect conductivity equation*, since it is identical to the Ohm's law in a perfectly conducting liquid.

Equations (3.214) and (3.217) can be summed to give the *energy evolution equation*:

$$\frac{3}{2}\frac{dp}{dt} + \frac{5}{2}\,p\,\nabla\cdot\mathbf{V} \simeq 0. \tag{3.223}$$

Equations (3.220) and (3.223) can be combined to give the more familiar *adiabatic equation of state:*

$$\frac{d}{dt}\left(\frac{p}{n^{5/3}}\right) \simeq 0. \tag{3.224}$$

Finally, the $O(\delta^{-1})$ components of Eqs. (3.214) and (3.217) yield

$$W_i \simeq 0, \tag{3.225}$$

or $T_e \simeq T_i$ [see Eq. (3.108)]. Thus, we expect equipartition of the thermal energy between electrons and ions in the MHD limit.

3.12 Drift Equations

The drift equations take the form:

$$\frac{\partial n}{\partial t} + \nabla\cdot(n\,\mathbf{V}_e) = 0, \tag{3.226}$$

$$m_e n \frac{\partial \mathbf{V}_e}{\partial t} + m_e n\,(\mathbf{V}_e\cdot\nabla)\mathbf{V}_e + [\delta^{-2}]\,\nabla p_e + [\zeta^{-1}]\,\nabla\cdot\boldsymbol{\pi}_e + [\delta^{-2}]\,e n\,(\mathbf{E} + \mathbf{V}_e\times\mathbf{B}) = [\delta^{-2}\,\zeta]\,\mathbf{F}_U + [\delta^{-2}]\,\mathbf{F}_T, \tag{3.227}$$

$$\frac{3}{2}\frac{\partial p_e}{\partial t} + \frac{3}{2}\,(\mathbf{V}_e\cdot\nabla)\,p_e + \frac{5}{2}\,p_e\,\nabla\cdot\mathbf{V}_e + [\zeta^{-1}]\,\nabla\cdot\mathbf{q}_{Te} + \nabla\cdot\mathbf{q}_{Ue} = -[\delta^{-2}\,\zeta\,\mu^2]\,W_i + [\zeta]\,W_U + W_T, \tag{3.228}$$

and

$$\frac{\partial n}{\partial t} + \nabla\cdot(n\,\mathbf{V}_i) = 0, \tag{3.229}$$

$$m_i n \frac{\partial \mathbf{V}_i}{\partial t} + m_i n\,(\mathbf{V}_i\cdot\nabla)\mathbf{V}_i + [\delta^{-2}]\,\nabla p_i + [\zeta^{-1}]\,\nabla\cdot\boldsymbol{\pi}_i \tag{3.230}$$

$$[0.5ex] - [\delta^{-2}]\,e n\,(\mathbf{E} + \mathbf{V}_i\times\mathbf{B}) = -[\delta^{-2}\,\zeta]\,\mathbf{F}_U - [\delta^{-2}]\,\mathbf{F}_T,$$

$$\frac{3}{2}\frac{\partial p_i}{\partial t} + \frac{3}{2}\,(\mathbf{V}_i\cdot\nabla)\,p_i + \frac{5}{2}\,p_i\,\nabla\cdot\mathbf{V}_i \tag{3.231}$$

$$+[\zeta^{-1}]\,\nabla\cdot\mathbf{q}_i = [\delta^{-2}\,\zeta\,\mu^2]\,W_i.$$

In the drift limit, the motions of the electron and ion fluids are sufficiently different that there is little to be gained in rewriting the drift equations in terms of the centre of mass velocity and the plasma current. Instead, let us consider the $O(\delta^{-2})$ components of Eqs. (3.227) and (3.231):

$$\mathbf{E} + \mathbf{V}_e\times\mathbf{B} \simeq -\frac{\nabla p_e}{e n} - \frac{0.71\,\nabla_\parallel T_e}{e}, \tag{3.232}$$

$$\mathbf{E} + \mathbf{V}_i\times\mathbf{B} \simeq +\frac{\nabla p_i}{e n} - \frac{0.71\,\nabla_\parallel T_e}{e}. \tag{3.233}$$

In the above equations, we have neglected all $O(\zeta)$ terms for the sake of simplicity. Equations (3.232)–(3.233) can be inverted to give

$$\mathbf{V}_{\perp e} \simeq \mathbf{V}_E + \mathbf{V}_{*e}, \tag{3.234}$$

$$\mathbf{V}_{\perp i} \simeq \mathbf{V}_E + \mathbf{V}_{*i}. \tag{3.235}$$

Here, $\mathbf{V}_E \equiv \mathbf{E}\times\mathbf{B}/B^2$ is the $\mathbf{E}\times\mathbf{B}$ velocity, whereas

$$\mathbf{V}_{*e} \equiv \frac{\nabla p_e\times\mathbf{B}}{e n\,B^2}, \tag{3.236}$$

and

$$\mathbf{V}_{*i} \equiv -\frac{\nabla p_i\times\mathbf{B}}{e n\,B^2}, \tag{3.237}$$

are termed the *electron diamagnetic velocity* and the *ion diamagnetic velocity*, respectively.

According to Eqs. (3.234)–(3.235), in the drift approximation the velocity of the electron fluid perpendicular to the magnetic field is the sum of the $\mathbf{E}\times\mathbf{B}$ velocity and the electron diamagnetic velocity. Similarly, for the ion fluid. Note that in the MHD approximation the perpendicular velocities of the two fluids consist of the $\mathbf{E}\times\mathbf{B}$ velocity alone, and are, therefore, identical to lowest order. The main difference between the two ordering lies in the assumed magnitude of the electric field. In the MHD limit

$$\frac{E}{B} \sim v_t, \tag{3.238}$$

whereas in the drift limit

$$\frac{E}{B} \sim \delta\, v_t \sim v_d. \tag{3.239}$$

Thus, the MHD ordering can be regarded as a *strong* (in the sense used in Sect. 2) electric field ordering, whereas the drift ordering corresponds to a *weak* electric field ordering.

The diamagnetic velocities are so named because the *diamagnetic current*,

$$\mathbf{j}_* \equiv -e n\, (\mathbf{V}_{*e} - \mathbf{V}_{*i}) = -\frac{\nabla p \times \mathbf{B}}{B^2}, \tag{3.240}$$

generally acts to *reduce* the magnitude of the magnetic field inside the plasma.

The electron diamagnetic velocity can be written

$$\mathbf{V}_{*e} = \frac{T_e\, \nabla n \times \mathbf{b}}{e n\, B} + \frac{\nabla T_e \times \mathbf{b}}{e\, B}. \tag{3.241}$$

In order to account for this velocity, let us consider a simplified case in which the electron temperature is uniform, there is a uniform *density* gradient running along the x-direction, and the magnetic field is parallel to the z-axis—see Fig. 3.3. The electrons gyrate in the x-y plane in circles of radius $\rho_e \sim v_e/|\Omega_e|$. At a given point, coordinate x_0, say, on the x-axis, the electrons that come from the right and the left have traversed distances of order ρ_e. Thus, the electrons from the right originate from regions where the particle density is of order $\rho_e\, \partial n/\partial x$ greater than the regions from which the electrons from the left originate. It follows that the y-directed particle flux is unbalanced, with slightly more particles moving in the $-y$-direction than in the $+y$-direction. Thus, there is a net particle flux in the $-y$-direction: *i.e.*, in the direction of $\nabla n \times \mathbf{b}$. The magnitude of this flux is

$$n\, V_{*e} \sim \rho_e\, \frac{\partial n}{\partial x}\, v_e \sim \frac{T_e}{e\, B}\, \frac{\partial n}{\partial x}. \tag{3.242}$$

Note that there is no unbalanced particle flux in the x-direction, since the x-directed fluxes are due to electrons which originate from regions where $x = x_0$. We have now accounted for the first term on the right-hand side of the above equation. We can account for the second term using similar arguments. The ion diamagnetic velocity is similar in magnitude to the electron diamagnetic velocity, but is *oppositely* directed, since ions gyrate in the *opposite* direction to electrons.

The most curious aspect of diamagnetic flows is that they represent fluid flows for which there is *no corresponding motion* of the particle guiding centres. Nevertheless, the diamagnetic velocities are *real* fluid velocities, and the associated diamagnetic current is a *real* current. For instance, the diamagnetic current contributes to force balance inside the plasma, and also gives rise to ohmic heating.

3.13 Closure in Collisionless Magnetized Plasmas

Up to now, we have only considered fluid closure in *collisional* magnetized plasmas. Unfortunately, most magnetized plasmas encountered in nature—in particular, fusion, space,

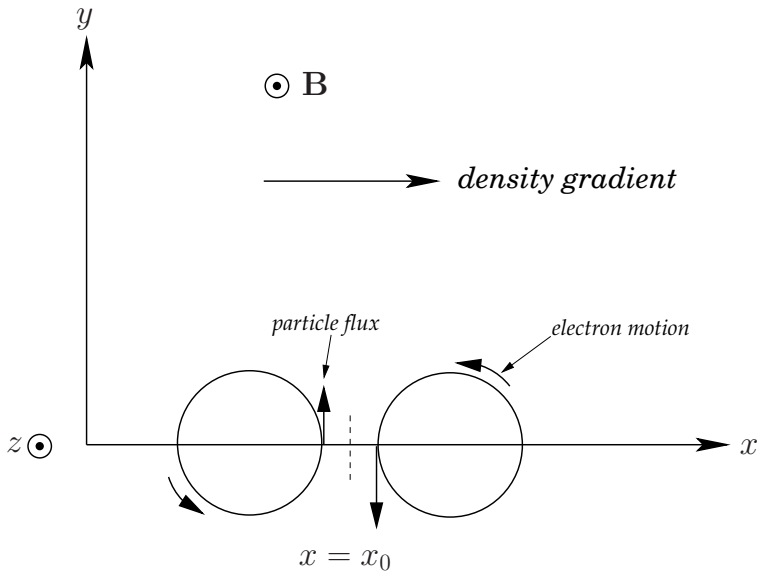

Figure 3.3: *Origin of the diamagnetic velocity in a magnetized plasma.*

and astrophysical plasmas—are *collisionless*. Let us consider what happens to the cold-plasma equations, the MHD equations, and the drift equations, in the limit in which the mean-free-path goes to infinity (*i.e.*, $\zeta \rightarrow 0$).

In the limit $\zeta \rightarrow 0$, the cold-plasma equations reduce to

$$\frac{dn}{dt} + n\,\nabla\cdot\mathbf{V} = 0, \tag{3.243}$$

$$m_i n \frac{d\mathbf{V}}{dt} - \mathbf{j}\times\mathbf{B} = 0, \tag{3.244}$$

$$\mathbf{E} + \mathbf{V}\times\mathbf{B} = \frac{\mathbf{j}\times\mathbf{B}}{ne} + \frac{m_e}{ne^2}\frac{d\mathbf{j}}{dt} + \frac{m_e}{ne^2}(\mathbf{j}\cdot\nabla)\mathbf{V} - \frac{m_e}{n^2e^3}(\mathbf{j}\cdot\nabla)\mathbf{j}. \tag{3.245}$$

Here, we have neglected the resistivity term, since it is $O(\zeta)$. Note that none of the remaining terms in these equations depend explicitly on collisions. Nevertheless, the absence of collisions poses a serious problem. Whereas the magnetic field effectively confines charged particles in directions perpendicular to magnetic field-lines, by forcing them to execute tight Larmor orbits, we have now lost all confinement along field-lines. But, does this matter?

The typical frequency associated with fluid motion is the transit frequency, V/L. However, according to Eq. (3.180), the cold-plasma ordering implies that the transit frequency is of order a typical gyrofrequency:

$$\frac{V}{L} \sim \Omega. \tag{3.246}$$

So, how far is a charged particle likely to drift along a field-line in an inverse transit frequency? The answer is

$$\Delta l_\parallel \sim \frac{\nu_t\, L}{V} \sim \frac{\nu_t}{\Omega} \sim \rho. \tag{3.247}$$

In other words, the fluid motion in the cold-plasma limit is so fast that charged particles only have time to drift a Larmor radius along field-lines on a typical dynamical time-scale. Under these circumstances, it does not really matter that the particles are not localized along field-lines—the lack of parallel confinement manifests itself *too slowly* to affect the plasma dynamics. We conclude, therefore, that the cold-plasma equations *remain valid in the collisionless limit*, provided, of course, that the plasma dynamics are sufficiently rapid for the basic cold-plasma ordering (3.246) to apply. In fact, the only difference between the collisional and collisionless cold-plasma equations is the absence of the resistivity term in Ohm's law in the latter case.

Let us now consider the MHD limit. In this case, the typical transit frequency is

$$\frac{V}{L} \sim \delta\,\Omega. \tag{3.248}$$

Thus, charged particles typically drift a distance

$$\Delta l_\parallel \sim \frac{\nu_t\, L}{V} \sim \frac{\nu_t}{\delta\,\Omega} \sim L \tag{3.249}$$

along field-lines in an inverse transit frequency. In other words, the fluid motion in the MHD limit is sufficiently slow that changed particles have time to drift along field-lines all the way across the system on a typical dynamical time-scale. Thus, strictly speaking, the MHD equations are *invalidated* by the lack of particle confinement along magnetic field-lines.

In fact, in collisionless plasmas, MHD theory is replaced by a theory known as *kinetic-MHD*.[4] The latter theory is a combination of a one-dimensional kinetic theory, describing particle motion along magnetic field-lines, and a two-dimensional fluid theory, describing perpendicular motion. As can well be imagined, the equations of kinetic-MHD are considerably more complicated that the conventional MHD equations. Is there any situation in which we can salvage the simpler MHD equations in a collisionless plasma? Fortunately, there is one case in which this is possible.

It turns out that in both varieties of MHD the motion of the plasma parallel to magnetic field-lines is associated with the dynamics of *sound waves*, whereas the motion perpendicular to field-lines is associated with the dynamics of a new type of wave called an *Alfvén wave*. As we shall see, later on, Alfvén waves involve the "twanging" motion of magnetic field-lines—a bit like the twanging of guitar strings. It is only the sound wave dynamics which are significantly modified when we move from a collisional to a collisionless plasma. It follows, therefore, that the MHD equations remain a reasonable approximation in a collisionless plasma in situations where the dynamics of sound waves, parallel to the magnetic

[4]Kinetic-MHD is described in the following two classic papers: M.D. Kruskal, and C.R. Oberman, Phys. Fluids **1**, 275 (1958): M.N. Rosenbluth, and N. Rostoker, Phys. Fluids **2**, 23 (1959).

field, are unimportant compared to the dynamics of Alfvén waves, perpendicular to the field. This situation arises whenever the parameter

$$\beta = \frac{2\,\mu_0\,p}{B^2} \tag{3.250}$$

(see Sect. 1.10) is much less than unity. In fact, it is easily demonstrated that

$$\beta \sim \left(\frac{V_S}{V_A}\right)^2, \tag{3.251}$$

where V_S is the sound speed (*i.e.*, thermal velocity), and V_A is the speed of an Alfvén wave. Thus, the inequality

$$\beta \ll 1 \tag{3.252}$$

ensures that the collisionless parallel plasma dynamics are *too slow* to affect the perpendicular dynamics.

We conclude, therefore, that in a low-β, collisionless, magnetized plasma the MHD equations,

$$\frac{dn}{dt} + n\,\nabla\cdot\mathbf{V} = 0, \tag{3.253}$$

$$m_i n \frac{d\mathbf{V}}{dt} = \mathbf{j}\times\mathbf{B} - \nabla p, \tag{3.254}$$

$$\mathbf{E} + \mathbf{V}\times\mathbf{B} = 0, \tag{3.255}$$

$$\frac{d}{dt}\left(\frac{p}{n^{5/3}}\right) = 0, \tag{3.256}$$

fairly well describe plasma dynamics which satisfy the basic MHD ordering (3.248).

Let us, finally, consider the drift limit. In this case, the typical transit frequency is

$$\frac{V}{L} \sim \delta^2\,\Omega. \tag{3.257}$$

Thus, charged particles typically drift a distance

$$\Delta l_\parallel \sim \frac{\nu_t\,L}{V} \sim \frac{L}{\delta} \tag{3.258}$$

along field-lines in an inverse transit frequency. In other words, the fluid motion in the drift limit is so slow that charged particles drifting along field-lines have time to traverse the system *very many times* on a typical dynamical time-scale. In fact, in this limit we have to draw a distinction between those particles which always drift along field-lines in the same direction, and those particles which are trapped between magnetic mirror points and, therefore, continually reverse their direction of motion along field-lines. The former are termed *passing particles*, whereas the latter are termed *trapped particles*.

Now, in the drift limit, the perpendicular drift velocity of charged particles, which is a combination of $\mathbf{E}\times\mathbf{B}$ drift, grad-B drift, and curvature drift (see Sect. 2), is of order

$$v_d \sim \delta\, v_t. \tag{3.259}$$

Thus, charged particles typically drift a distance

$$\Delta l_\perp \sim \frac{v_d\, L}{V} \sim L \tag{3.260}$$

across field-lines in an inverse transit time. In other words, the fluid motion in the drift limit is so slow that charged particles have time to drift perpendicular to field-lines all the way across the system on a typical dynamical time-scale. It is, thus, clear that in the drift limit the absence of collisions implies lack of confinement both parallel and perpendicular to the magnetic field. This means that the collisional drift equations, (3.226)–(3.229) and (3.229)–(3.232), are *completely invalid* in the long mean-free-path limit.

In fact, in collisionless plasmas, Braginskii-type transport theory—conventionally known as *classical transport theory*—is replaced by a new theory—known as *neoclassical transport theory*[5] —which is a combination of a two-dimensional kinetic theory, describing particle motion on *drift surfaces*, and a one-dimensional fluid theory, describing motion perpendicular to the drift surfaces. Here, a drift surface is a closed surface formed by the locus of a charged particle's drift orbit (including drifts parallel and perpendicular to the magnetic field). Of course, the orbits only form closed surfaces if the plasma is *confined*, but there is little point in examining transport in an unconfined plasma. Unlike classical transport theory, which is strictly *local* in nature, neoclassical transport theory is *nonlocal*, in the sense that the transport coefficients depend on the *average* values of plasma properties taken over drift surfaces. Needless to say, neoclassical transport theory is horribly complicated!

3.14 Langmuir Sheaths

Virtually all terrestrial plasmas are contained inside solid vacuum vessels. So, an obvious question is: what happens to the plasma in the immediate vicinity of the vessel wall? Actually, to a first approximation, when ions and electrons hit a solid surface they recombine and are lost to the plasma. Hence, we can treat the wall as a perfect sink of particles. Now, given that the electrons in a plasma generally move much faster than the ions, the initial electron flux into the wall greatly exceeds the ion flux, assuming that the wall starts off unbiased with respect to the plasma. Of course, this flux imbalance causes the wall to charge up negatively, and so generates a potential barrier which repels the electrons, and thereby reduces the electron flux. Debye shielding confines this barrier to a thin layer of plasma, whose thickness is a few Debye lengths, coating the inside surface of the wall. This

[5]Neoclassical transport theory in axisymmetric systems is described in the following classic papers: I.B. Bernstein, Phys. Fluids **17**, 547 (1974): F.L. Hinton, and R.D. Hazeltine, Rev. Mod. Phys. **48**, 239 (1976).

layer is known as a *plasma sheath* or a *Langmuir sheath*. The height of the potential barrier continues to grow as long as there is a net flux of negative charge into the wall. This process presumably comes to an end, and a steady-state is attained, when the potential barrier becomes sufficiently large to make electron flux equal to the ion flux.

Let us construct a one-dimensional model of an unmagnetized, steady-state, Langmuir sheath. Suppose that the wall lies at $x = 0$, and that the plasma occupies the region $x > 0$. Let us treat the ions and the electrons inside the sheath as *collisionless fluids*. The ion and electron equations of motion are thus written

$$m_i\, n_i\, V_i \frac{dV_i}{dx} = -T_i \frac{dn_i}{dx} - e\, n_i \frac{d\phi}{dx}, \tag{3.261}$$

$$m_e\, n_e\, V_e \frac{dV_e}{dx} = -T_e \frac{dn_e}{dx} + e\, n_e \frac{d\phi}{dx}, \tag{3.262}$$

respectively. Here, $\phi(x)$ is the electrostatic potential. Moreover, we have assumed *uniform* ion and electron temperatures, T_i and T_e, respectively, for the sake of simplicity. We have also neglected any off-diagonal terms in the ion and electron stress-tensors, since these terms are comparatively small. Note that quasi-neutrality does not apply inside the sheath, and so the ion and electron number densities, n_e and n_i, respectively, are not necessarily equal to one another.

Consider the ion fluid. Let us assume that the mean ion velocity, V_i, is *much greater* than the ion thermal velocity, $(T_i/m_i)^{1/2}$. Since, as will become apparent, $V_i \sim (T_e/m_i)^{1/2}$, this ordering necessarily implies that $T_i \ll T_i$: *i.e.*, that the ions are *cold* with respect to the electrons. It turns out that plasmas in the immediate vicinity of solid walls often have comparatively cold ions, so our ordering assumption is fairly reasonable. In the cold ion limit, the pressure term in Eq. (3.261) is negligible, and the equation can be integrated to give

$$\frac{1}{2}\, m_i\, V_i^2(x) + e\, \phi(x) = \frac{1}{2}\, m_i\, V_s^2 + e\, \phi_s. \tag{3.263}$$

Here, V_s and ϕ_s are the mean ion velocity and electrostatic potential, respectively, at the edge of the sheath (*i.e.*, $x \rightarrow \infty$). Now, ion fluid continuity requires that

$$n_i(x)\, V_i(x) = n_s\, V_s, \tag{3.264}$$

where n_s is the ion number density at the sheath boundary. Incidentally, since we expect quasi-neutrality to hold in the plasma outside the sheath, the electron number density at the edge of the sheath must also be n_s (assuming singly charged ions). The previous two equations can be combined to give

$$V_i = V_s \left[1 - \frac{2\,e}{m_i\, V_s^2}\,(\phi - \phi_s)\right]^{1/2}, \tag{3.265}$$

$$n_i = n_s \left[1 - \frac{2\,e}{m_i\, V_s^2}\,(\phi - \phi_s)\right]^{-1/2}. \tag{3.266}$$

Consider the electron fluid. Let us assume that the mean electron velocity, V_e, is *much less* than the electron thermal velocity, $(m_e/T_e)^{1/2}$. In fact, this must be the case, otherwise, the electron flux to the wall would greatly exceed the ion flux. Now, if the electron fluid is essentially stationary then the left-hand side of Eq. (3.262) is negligible, and the equation can be integrated to give

$$n_e = n_s \exp\left[\frac{e\,(\phi-\phi_s)}{T_e}\right]. \tag{3.267}$$

Here, we have made use of the fact that $n_e = n_s$ at the edge of the sheath.

Now, Poisson's equation is written

$$\epsilon_0 \frac{d^2\phi}{dx^2} = e\,(n_e - n_i). \tag{3.268}$$

It follows that

$$\epsilon_0 \frac{d^2\phi}{dx^2} = e\,n_s\left(\exp\left[\frac{e\,(\phi-\phi_s)}{T_e}\right] - \left[1-\frac{2\,e}{m_i\,V_s^2}\,(\phi-\phi_s)\right]^{-1/2}\right). \tag{3.269}$$

Let $\Phi = -e\,(\phi-\phi_s)/T_e$, $y = \sqrt{2}\,x/\lambda_D$, and

$$K = \frac{m_i\,V_s^2}{2\,T_e}, \tag{3.270}$$

where $\lambda_D = (\epsilon_0\,T_e/e^2\,n_s)^{1/2}$ is the Debye length. Equation (3.269) transforms to

$$2\,\frac{d^2\Phi}{dy^2} = -e^{-\Phi} + \left(1+\frac{\Phi}{K}\right)^{-1/2}, \tag{3.271}$$

subject to the boundary condition $\Phi \rightarrow 0$ as $y \rightarrow \infty$. Multiplying through by $d\Phi/dy$, integrating with respect to y, and making use of the boundary condition, we obtain

$$\left(\frac{d\Phi}{dy}\right)^2 = e^{-\Phi} - 1 + 2\,K\left[\left(1+\frac{\Phi}{K}\right)^{1/2} - 1\right]. \tag{3.272}$$

Unfortunately, the above equation is highly nonlinear, and can only be solved numerically. However, it is not necessary to attempt this to see that a physical solution can only exist if the right-hand side of the equation is *positive* for all $y \geq 0$. Consider the the limit $y \rightarrow \infty$. It follows from the boundary condition that $\Phi \rightarrow 0$. Expanding the right-hand side of Eq. (3.272) in powers of Φ, we find that the zeroth- and first-order terms cancel, and we are left with

$$\left(\frac{d\Phi}{dy}\right)^2 \simeq \frac{\Phi^2}{2}\left(1-\frac{1}{2\,K}\right) + \frac{\Phi^3}{3}\left(\frac{3}{8\,K^2}-1\right) + O(\Phi^4). \tag{3.273}$$

Now, the purpose of the sheath is to *shield* the plasma from the wall potential. It can be seen, from the above expression, that the physical solution with maximum possible

shielding corresponds to $K = 1/2$, since this choice eliminates the first term on the right-hand side (thereby making Φ as small as possible at large y) leaving the much smaller, but positive (note that Φ is positive), second term. Hence, we conclude that

$$V_s = \left(\frac{T_e}{m_i}\right)^{1/2}. \tag{3.274}$$

This result is known as the *Bohm sheath criterion*. It is a somewhat surprising result, since it indicates that ions at the edge of the sheath are already moving toward the wall at a considerable velocity. Of course, the ions are further accelerated as they pass through the sheath. Since the ions are presumably at rest in the interior of the plasma, it is clear that there must exist a region sandwiched between the sheath and the main plasma in which the ions are accelerated from rest to the Bohm velocity, $V_s = (T_e/m_i)^{1/2}$. This region is called the *pre-sheath*, and is both quasi-neutral and much wider than the sheath (the actual width depends on the nature of the ion source).

The ion current density at the wall is

$$j_i = -e\,n_i(0)\,V_i(0) = -e\,n_s\,V_s = -e\,n_s\left(\frac{T_e}{m_i}\right)^{1/2}. \tag{3.275}$$

This current density is negative because the ions are moving in the negative x-direction. What about the electron current density? Well, the number density of electrons at the wall is $n_e(0) = n_s\,\exp[e\,(\phi_w - \phi_s)/T_e)]$, where $\phi_w = \phi(0)$ is the wall potential. Let us assume that the electrons have a Maxwellian velocity distribution peaked at zero velocity (since the electron fluid velocity is much less than the electron thermal velocity). It follows that half of the electrons at $x = 0$ are moving in the negative-x direction, and half in the positive-x direction. Of course, the former electrons hit the wall, and thereby constitute an electron current to the wall. This current is $j_e = (1/4)\,e\,n_e(0)\,\bar{V}_e$, where the $1/4$ comes from averaging over solid angle, and $\bar{V}_e = (8\,T_e/\pi\,m_e)^{1/2}$ is the mean electron speed corresponding to a Maxwellian velocity distribution. Thus, the electron current density at the wall is

$$j_e = e\,n_s\left(\frac{T_e}{2\,\pi\,m_e}\right)^{1/2}\exp\left[\frac{e\,(\phi_w - \phi_s)}{T_e}\right]. \tag{3.276}$$

Now, in order to replace the electrons lost to the wall, the electrons must have a mean velocity

$$V_{e\,s} = \frac{j_e}{e\,n_s} = \left(\frac{T_e}{2\pi\,m_e}\right)^{1/2}\exp\left[\frac{e\,(\phi_w - \phi_s)}{T_e}\right] \tag{3.277}$$

at the edge of the sheath. However, we previously assumed that any electron fluid velocity was much less than the electron thermal velocity, $(T_e/m_e)^{1/2}$. As is clear from the above equation, this is only possible provided that

$$\exp\left[\frac{e\,(\phi_w - \phi_s)}{T_e}\right] \ll 1. \tag{3.278}$$

i.e., provided that the wall potential is sufficiently negative to strongly reduce the electron number density at the wall. The net current density at the wall is

$$j = e\,n_s\left(\frac{T_e}{m_i}\right)^{1/2}\left\{\left(\frac{m_i}{2\pi\,m_e}\right)^{1/2}\exp\left[\frac{e\,(\phi_w-\phi_s)}{T_e}\right]-1\right\}. \tag{3.279}$$

Of course, we require $j = 0$ in a steady-state sheath, in order to prevent wall charging, and so we obtain

$$e\,(\phi_w-\phi_s) = -T_e\,\ln\left(\frac{m_i}{2\pi\,m_e}\right)^{1/2}. \tag{3.280}$$

We conclude that, in a steady-state sheath, the wall is biased *negatively* with respect to the sheath edge by an amount which is proportional to the electron temperature.

For a hydrogen plasma, $\ln(m_i/2\pi\,m_e)\simeq 2.8$. Thus, hydrogen ions enter the sheath with an initial energy $(1/2)\,m_i\,V_s^2 = 0.5\,T_e$ eV, fall through the sheath potential, and so impact the wall with energy $3.3\,T_e$ eV.

A *Langmuir probe* is a device used to determine the electron temperature and electron number density of a plasma. It works by inserting an electrode which is biased with respect to the vacuum vessel into the plasma. Provided that the bias voltage is not too positive, we would expect the probe current to vary as

$$I = A\,e\,n_s\left(\frac{T_e}{m_i}\right)^{1/2}\left[\left(\frac{m_i}{2\pi\,m_e}\right)^{1/2}\exp\left(\frac{e\,V}{T_e}\right)-1\right], \tag{3.281}$$

where A is the surface area of the probe, and V its bias with respect to the vacuum vessel—see Eq. (3.279). For strongly negative biases, the probe current saturates in the ion (negative) direction. The characteristic current which flows in this situation is called the *ion saturation current*, and is of magnitude

$$I_s = A\,e\,n_s\left(\frac{T_e}{m_i}\right)^{1/2}. \tag{3.282}$$

For less negative biases, the current-voltage relation of the probe has the general form

$$\ln I = C + \frac{e\,V}{T_e}, \tag{3.283}$$

where C is a constant. Thus, a plot of $\ln I$ versus V gives a *straight-line* from whose slope the electron temperature can be deduced. Note, however, that if the bias voltage becomes too positive then electrons cease to be effectively repelled from the probe surface, and the current-voltage relation (3.281) breaks down. Given the electron temperature, a measurement of the ion saturation current allows the electron number density at the sheath edge, n_s, to be calculated from Eq. (3.282). Now, in order to accelerate ions to the Bohm velocity, the potential drop across the pre-sheath needs to be $e\,(\phi_p-\phi_s) = -T_e/2$, where ϕ_p is the electric potential in the interior of the plasma. It follows from Eq. (3.267)

that the relationship between the electron number density at the sheath boundary, n_s, and the number density in the interior of the plasma, n_p, is

$$n_s = n_p \, e^{-0.5} \simeq 0.61 \, n_p. \tag{3.284}$$

Thus, n_p can also be determined from the probe.

4 Waves in Cold Plasmas

4.1 Introduction

The cold-plasma equations describe waves, and other perturbations, which propagate through a plasma *much faster* than a typical thermal velocity. It is instructive to consider the relationship between the collective motions described by the cold-plasma model and the motions of individual particles that we studied in Sect. 2. The key observation is that in the cold-plasma model all particles (of a given species) at a given position effectively move with the same velocity. It follows that the fluid velocity is identical to the particle velocity, and is, therefore, governed by the same equations. However, the cold-plasma model goes beyond the single-particle description because it determines the electromagnetic fields *self-consistently* in terms of the charge and current densities generated by the motions of the constituent particles of the plasma.

What role, if any, does the geometry of the plasma equilibrium play in determining the properties of plasma waves? Clearly, geometry plays a key role for modes whose wave-lengths are comparable to the dimensions of the plasma. However, we shall show that modes whose wave-lengths are *much smaller* than the plasma dimensions have properties which are, in a local sense, *independent* of the geometry. Thus, the local properties of small-wave-length oscillations are *universal* in nature. To investigate these properties, we may, to a first approximation, represent the plasma as a homogeneous equilibrium (corresponding to the limit $k\,L \rightarrow 0$, where k is the magnitude of the wave-vector, and L is the characteristic equilibrium length-scale).

4.2 Plane Waves in a Homogeneous Plasma

The propagation of small amplitude waves is described by *linearized equations*. These are obtained by expanding the equations of motion in powers of the wave amplitude, and neglecting terms of order higher than unity. In the following, we use the subscript 0 to distinguish equilibrium quantities from perturbed quantities, for which we retain the previous notation.

Consider a homogeneous, quasi-neutral plasma, consisting of equal numbers of electrons and ions, in which both plasma species are at rest. It follows that $\mathbf{E}_0 = \mathbf{0}$, and $\mathbf{j}_0 = \nabla \times \mathbf{B}_0 = \mathbf{0}$. In a homogeneous medium, the general solution of a system of linear equations can be constructed as a superposition of plane wave solutions:

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{E}_{\mathbf{k}}\, \exp[\,i\,(\mathbf{k}\cdot\mathbf{r} - \omega\, t)], \tag{4.1}$$

with similar expressions for $\mathbf{B}$ and $\mathbf{V}$. The surfaces of constant phase,

$$\mathbf{k}\cdot\mathbf{r} - \omega\, t = \text{constant}, \tag{4.2}$$

are planes perpendicular to $\mathbf{k}$, traveling at the velocity

$$\mathbf{v}_{\rm ph} = \frac{\omega}{k}\,\hat{\mathbf{k}}, \tag{4.3}$$

where $k \equiv |\mathbf{k}|$, and $\hat{\mathbf{k}}$ is a unit vector pointing in the direction of $\mathbf{k}$. Here, $\mathbf{v}_{\rm ph}$ is termed the *phase-velocity*. Henceforth, we shall omit the subscript $\mathbf{k}$ from field variables, for ease of notation.

Substitution of the plane wave solution (4.1) into Maxwell's equations yields:

$$\mathbf{k}\times\mathbf{B} = -{\rm i}\,\mu_0\,\mathbf{j} - \frac{\omega}{c^2}\,\mathbf{E}, \tag{4.4}$$

$$\mathbf{k}\times\mathbf{E} = \omega\,\mathbf{B}. \tag{4.5}$$

In linear theory, the current is related to the electric field via

$$\mathbf{j} = \boldsymbol{\sigma}\cdot\mathbf{E}, \tag{4.6}$$

where the *conductivity tensor* $\boldsymbol{\sigma}$ is a function of both $\mathbf{k}$ and ω. Note that the conductivity tensor is *anisotropic* in the presence of a non-zero equilibrium magnetic field. Furthermore, $\boldsymbol{\sigma}$ completely specifies the plasma response.

Substitution of Eq. (4.6) into Eq. (4.4) yields

$$\mathbf{k}\times\mathbf{B} = -\frac{\omega}{c^2}\,\mathbf{K}\cdot\mathbf{E}, \tag{4.7}$$

where we have introduced the *dielectric permittivity tensor*,

$$\mathbf{K} = \mathbf{I} + \frac{{\rm i}\,\boldsymbol{\sigma}}{\epsilon_0\,\omega}. \tag{4.8}$$

Here, $\mathbf{I}$ is the identity tensor. Eliminating the magnetic field between Eqs. (4.5) and (4.7), we obtain

$$\mathbf{M}\cdot\mathbf{E} = \mathbf{0}, \tag{4.9}$$

where

$$\mathbf{M} = \mathbf{k}\mathbf{k} - k^2\,\mathbf{I} + \frac{\omega^2}{c^2}\,\mathbf{K}. \tag{4.10}$$

The solubility condition for Eq. (4.10),

$$\mathcal{M}(\omega,\mathbf{k}) \equiv \det(\mathbf{M}) = 0, \tag{4.11}$$

is called the *dispersion relation*. The dispersion relation relates the frequency, ω, to the wave-vector, $\mathbf{k}$. Also, as the name "dispersion relation" indicates, it allows us to determine the rate at which the different Fourier components in a wave-train *disperse* due to the variation of their phase-velocity with wave-length.

4.3 Cold-Plasma Dielectric Permittivity

In a collisionless plasma, the linearized cold-plasma equations are written [see Eqs. (3.243)–(3.246)]:

$$m_i n \frac{\partial \mathbf{V}}{\partial t} = \mathbf{j} \times \mathbf{B}_0, \tag{4.12}$$

$$\mathbf{E} = -\mathbf{V} \times \mathbf{B}_0 + \frac{\mathbf{j} \times \mathbf{B}_0}{n e} + \frac{m_e}{n e^2} \frac{\partial \mathbf{j}}{\partial t}. \tag{4.13}$$

Substitution of plane wave solutions of the type (4.1) into the above equations yields

$$-\mathrm{i}\,\omega\, m_i n \mathbf{V} = \mathbf{j} \times \mathbf{B}_0, \tag{4.14}$$

$$\mathbf{E} = -\mathbf{V} \times \mathbf{B}_0 + \frac{\mathbf{j} \times \mathbf{B}_0}{n e} - \mathrm{i}\,\omega \frac{m_e}{n e^2} \mathbf{j}. \tag{4.15}$$

Let

$$\Pi_e = \sqrt{\frac{n e^2}{\epsilon_0 m_e}}, \tag{4.16}$$

$$\Pi_i = \sqrt{\frac{n e^2}{\epsilon_0 m_i}}, \tag{4.17}$$

$$\Omega_e = -\frac{e B_0}{m_e}, \tag{4.18}$$

$$\Omega_i = \frac{e B_0}{m_i}, \tag{4.19}$$

be the *electron plasma frequency*, the *ion plasma frequency*, the *electron cyclotron frequency*, and the *ion cyclotron frequency*, respectively. The "plasma frequency," ω_p, mentioned in Sect. 1, is identical to the electron plasma frequency, Π_e. Eliminating the fluid velocity $\mathbf{V}$ between Eqs. (4.14) and (4.15), and making use of the above definitions, we obtain

$$\mathrm{i}\,\omega\,\epsilon_0 \mathbf{E} = \frac{\omega^2 \mathbf{j} - \mathrm{i}\,\omega\,\Omega_e \mathbf{j} \times \mathbf{b} + \Omega_e\,\Omega_i \mathbf{j}_\perp}{\Pi_e^{\,2}}. \tag{4.20}$$

The parallel component of the above equation is readily solved to give

$$j_\parallel = \frac{\Pi_e^{\,2}}{\omega^2} (\mathrm{i}\,\omega\,\epsilon_0 E_\parallel). \tag{4.21}$$

In solving for $\mathbf{j}_\perp$, it is helpful to define the vectors:

$$\mathbf{e}_+ = \frac{\mathbf{e}_1 + \mathrm{i}\,\mathbf{e}_2}{\sqrt{2}}, \tag{4.22}$$

$$\mathbf{e}_- = \frac{\mathbf{e}_1 - \mathrm{i}\,\mathbf{e}_2}{\sqrt{2}}. \tag{4.23}$$

Here, $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{b})$ are a set of mutually orthogonal, right-handed unit vectors. Note that

$$\mathbf{b} \times \mathbf{e}_\pm = \mp \mathrm{i}\, \mathbf{e}_\pm. \tag{4.24}$$

It is easily demonstrated that

$$\mathrm{j}_\pm = \frac{\Pi_e^{\,2}}{\omega^2 \pm \omega\, \Omega_e + \Omega_e\, \Omega_\mathrm{i}}\, \mathrm{i}\, \omega\, \epsilon_0\, \mathrm{E}_\pm, \tag{4.25}$$

where $\mathrm{j}_\pm = \mathbf{j} \cdot \mathbf{e}_\pm$, *etc.*

The conductivity tensor is *diagonal* in the basis $(\mathbf{e}_+, \mathbf{e}_-, \mathbf{b})$. Its elements are given by the coefficients of $\mathrm{E}_\pm$ and $\mathrm{E}_\parallel$ in Eqs. (4.25) and (4.21), respectively. Thus, the dielectric permittivity (4.8) takes the form

$$\mathbf{K}_\mathrm{circ} = \begin{pmatrix} \mathrm{R} & 0 & 0 \\ 0 & \mathrm{L} & 0 \\ 0 & 0 & \mathrm{P} \end{pmatrix}, \tag{4.26}$$

where

$$\mathrm{R} \simeq 1 - \frac{\Pi_e^{\,2}}{\omega^2 + \omega\, \Omega_e + \Omega_e\, \Omega_\mathrm{i}}, \tag{4.27}$$

$$\mathrm{L} \simeq 1 - \frac{\Pi_e^{\,2}}{\omega^2 - \omega\, \Omega_e + \Omega_e\, \Omega_\mathrm{i}}, \tag{4.28}$$

$$\mathrm{P} \simeq 1 - \frac{\Pi_e^{\,2}}{\omega^2}. \tag{4.29}$$

Here, R and L represent the permittivities for right- and left-handed circularly polarized waves, respectively. The permittivity parallel to the magnetic field, P, is identical to that of an unmagnetized plasma.

In fact, the above expressions are only approximate, because the small mass-ratio ordering $m_e/m_\mathrm{i} \ll 1$ has already been folded into the cold-plasma equations. The exact expressions, which are most easily obtained by solving the individual charged particle equations of motion, and then summing to obtain the fluid response, are:

$$\mathrm{R} = 1 - \frac{\Pi_e^{\,2}}{\omega^2}\left(\frac{\omega}{\omega + \Omega_e}\right) - \frac{\Pi_\mathrm{i}^{\,2}}{\omega^2}\left(\frac{\omega}{\omega + \Omega_\mathrm{i}}\right), \tag{4.30}$$

$$\mathrm{L} = 1 - \frac{\Pi_e^{\,2}}{\omega^2}\left(\frac{\omega}{\omega - \Omega_e}\right) - \frac{\Pi_\mathrm{i}^{\,2}}{\omega^2}\left(\frac{\omega}{\omega - \Omega_\mathrm{i}}\right), \tag{4.31}$$

$$\mathrm{P} = 1 - \frac{\Pi_e^{\,2}}{\omega^2} - \frac{\Pi_\mathrm{i}^{\,2}}{\omega^2}. \tag{4.32}$$

Equations (4.27)–(4.29) and (4.30)–(4.32) are equivalent in the limit $m_e/m_\mathrm{i} \rightarrow 0$. Note that Eqs. (4.30)–(4.32) generalize in a fairly obvious manner in plasmas consisting of more than two particle species.

In order to obtain the actual dielectric permittivity, it is necessary to transform back to the Cartesian basis $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{b})$. Let $\mathbf{b} \equiv \mathbf{e}_3$, for ease of notation. It follows that the components of an arbitrary vector $\mathbf{W}$ in the Cartesian basis are related to the components in the "circular" basis via

$$\begin{pmatrix} W_1 \\ W_2 \\ W_3 \end{pmatrix} = \mathbf{U} \begin{pmatrix} W_+ \\ W_- \\ W_3 \end{pmatrix}, \tag{4.33}$$

where the unitary matrix $\mathbf{U}$ is written

$$\mathbf{U} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 & 0 \\ \mathrm{i} & -\mathrm{i} & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix}. \tag{4.34}$$

The dielectric permittivity in the Cartesian basis is then

$$\mathbf{K} = \mathbf{U}\,\mathbf{K}_{\rm circ}\,\mathbf{U}^\dagger. \tag{4.35}$$

We obtain

$$\mathbf{K} = \begin{pmatrix} S & -\mathrm{i}\,D & 0 \\ \mathrm{i}\,D & S & 0 \\ 0 & 0 & P \end{pmatrix}, \tag{4.36}$$

where

$$S = \frac{R + L}{2}, \tag{4.37}$$

and

$$D = \frac{R - L}{2}, \tag{4.38}$$

represent the sum and difference of the right- and left-handed dielectric permittivities, respectively.

4.4 Cold-Plasma Dispersion Relation

It is convenient to define a vector

$$\mathbf{n} = \frac{\mathbf{k}\,c}{\omega}, \tag{4.39}$$

which points in the same direction as the wave-vector, $\mathbf{k}$, and whose magnitude n is the *refractive index* (*i.e.*, the ratio of the velocity of light in vacuum to the phase-velocity). Note that n should not be confused with the particle density. Equation (4.9) can be rewritten

$$\mathbf{M}\cdot\mathbf{E} = (\mathbf{n}\cdot\mathbf{E})\,\mathbf{n} - n^2\,\mathbf{K}\cdot\mathbf{E} = \mathbf{0}. \tag{4.40}$$

We may, without loss of generality, assume that the equilibrium magnetic field is directed along the z-axis, and that the wave-vector, $\mathbf{k}$, lies in the xz-plane. Let θ be the angle

subtended between $\mathbf{k}$ and $\mathbf{B}_0$. The eigenmode equation (4.40) can be written

$$\begin{pmatrix} S - n^2\cos^2\theta & -i\,D & n^2\cos\theta\sin\theta \\ i\,D & S - n^2 & 0 \\ n^2\cos\theta\sin\theta & 0 & P - n^2\sin^2\theta \end{pmatrix} \begin{pmatrix} E_x \\ E_y \\ E_z \end{pmatrix} = \mathbf{0}. \tag{4.41}$$

The condition for a nontrivial solution is that the determinant of the square matrix be zero. With the help of the identity

$$S^2 - D^2 \equiv R\,L, \tag{4.42}$$

we find that

$$\mathcal{M}(\omega,\mathbf{k}) \equiv A\,n^4 - B\,n^2 + C = 0, \tag{4.43}$$

where

$$A = S\sin^2\theta + P\cos^2\theta, \tag{4.44}$$

$$B = R\,L\sin^2\theta + P\,S\,(1+\cos^2\theta), \tag{4.45}$$

$$C = P\,R\,L. \tag{4.46}$$

The dispersion relation (4.43) is evidently a quadratic in n^2, with two roots. The solution can be written

$$n^2 = \frac{B \pm F}{2\,A}, \tag{4.47}$$

where

$$F^2 = (R\,L - P\,S)^2\sin^4\theta + 4\,P^2\,D^2\cos^2\theta. \tag{4.48}$$

Note that $F^2 \geq 0$. It follows that n^2 is always real, which implies that n is either purely real or purely imaginary. In other words, the cold-plasma dispersion relation describes waves which either propagate without evanescense, or decay without spatial oscillation. The two roots of opposite sign for n, corresponding to a particular root for n^2, simply describe waves of the same type propagating, or decaying, in opposite directions.

The dispersion relation (4.43) can also be written

$$\tan^2\theta = -\frac{P\,(n^2 - R)\,(n^2 - L)}{(S\,n^2 - R\,L)\,(n^2 - P)}. \tag{4.49}$$

For the special case of wave propagation *parallel* to the magnetic field (*i.e.*, $\theta = 0$), the above expression reduces to

$$P = 0, \tag{4.50}$$

$$n^2 = R, \tag{4.51}$$

$$n^2 = L. \tag{4.52}$$

Likewise, for the special case of propagation *perpendicular* to the field (*i.e.*, $\theta = \pi/2$), Eq. (4.49) yields

$$n^2 = \frac{R\,L}{S}, \tag{4.53}$$

$$n^2 = P. \tag{4.54}$$

4.5 Polarization

A pure right-handed circularly polarized wave propagating along the z-axis takes the form

$$E_x = A\cos(k\,z-\omega\,t), \tag{4.55}$$

$$E_y = -A\sin(k\,z-\omega\,t). \tag{4.56}$$

In terms of complex amplitudes, this becomes

$$\frac{i\,E_x}{E_y} = 1. \tag{4.57}$$

Similarly, a left-handed circularly polarized wave is characterized by

$$\frac{i\,E_x}{E_y} = -1. \tag{4.58}$$

The polarization of the transverse electric field is obtained from the middle line of Eq. (4.41):

$$\frac{i\,E_x}{E_y} = \frac{n^2-S}{D} = \frac{2n^2-(R+L)}{R-L}. \tag{4.59}$$

For the case of parallel propagation, with $n^2 = R$, the above formula yields $i\,E_x/E_y = 1$. Similarly, for the case of parallel propagation, with $n^2 = L$, we obtain $i\,E_x/E_y = -1$. Thus, it is clear that the roots $n^2 = R$ and $n^2 = L$ in Eqs. (4.50)–(4.52) correspond to right- and left-handed circularly polarized waves, respectively.

4.6 Cutoff and Resonance

For certain values of the plasma parameters, n^2 goes to zero or infinity. In both cases, a transition is made from a region of propagation to a region of evanescense, or *vice versa*. It will be demonstrated later on that *reflection* occurs wherever n^2 goes through zero, and that *absorption* takes place wherever n^2 goes through infinity. The former case is called a wave *cutoff*, whereas the latter case is termed a wave *resonance*.

According to Eqs. (4.43) and (4.44)–(4.46), cutoff occurs when

$$P = 0, \tag{4.60}$$

or

$$R = 0, \tag{4.61}$$

or

$$L = 0. \tag{4.62}$$

Note that the cutoff points are independent of the direction of propagation of the wave relative to the magnetic field.

According to Eq. (4.49), resonance takes place when

$$\tan^2\theta = -\frac{P}{S}. \tag{4.63}$$

Evidently, resonance points do depend on the direction of propagation of the wave relative to the magnetic field. For the case of parallel propagation, resonance occurs whenever $S \rightarrow \infty$. In other words, when

$$R \rightarrow \infty, \tag{4.64}$$

or

$$L \rightarrow \infty. \tag{4.65}$$

For the case of perpendicular propagation, resonance takes place when

$$S = 0. \tag{4.66}$$

4.7 Waves in an Unmagnetized Plasma

Let us now investigate the cold-plasma dispersion relation in detail. It is instructive to first consider the limit in which the equilibrium magnetic field goes to zero. In the absence of the magnetic field, there is no preferred direction, so we can, without loss of generality, assume that $\mathbf{k}$ is directed along the z-axis (*i.e.*, $\theta = 0$). In the zero magnetic field limit (*i.e.*, $\Omega_e, \Omega_i \rightarrow 0$), the eigenmode equation (4.41) reduces to

$$\begin{pmatrix} P - n^2 & 0 & 0 \\ 0 & P - n^2 & 0 \\ 0 & 0 & P \end{pmatrix} \begin{pmatrix} E_x \\ E_y \\ E_z \end{pmatrix} = \mathbf{0}, \tag{4.67}$$

where

$$P \simeq 1 - \frac{\Pi_e^{\,2}}{\omega^2}. \tag{4.68}$$

Here, we have neglected Π_i with respect to Π_e.

It is clear from Eq. (4.67) that there are two types of wave. The first possesses the eigenvector $(0, 0, E_z)$, and has the dispersion relation

$$1 - \frac{\Pi_e^{\,2}}{\omega^2} = 0. \tag{4.69}$$

The second possesses the eigenvector $(E_x, E_y, 0)$, and has the dispersion relation

$$1 - \frac{\Pi_e^{\,2}}{\omega^2} - \frac{k^2 c^2}{\omega^2} = 0. \tag{4.70}$$

Here, E_x, E_y, and E_z are arbitrary non-zero quantities.

The first wave has $\mathbf{k}$ parallel to $\mathbf{E}$, and is, thus, a *longitudinal* wave. This wave is know as the *plasma wave*, and possesses the fixed frequency $\omega = \Pi_e$. Note that if $\mathbf{E}$ is parallel to

$\mathbf{k}$ then it follows from Eq. (4.5) that $\mathbf{B} = \mathbf{0}$. In other words, the wave is purely *electrostatic* in nature. In fact, a plasma wave is an electrostatic oscillation of the type discussed in Sect. 1.5. Since ω is independent of $\mathbf{k}$, the *group velocity*,

$$\mathbf{v}_g = \frac{\partial \omega}{\partial \mathbf{k}}, \tag{4.71}$$

associated with a plasma wave, is zero. As we shall demonstrate later on, the group velocity is the propagation velocity of localized wave packets. It is clear that the plasma wave is *not* a propagating wave, but instead has the property than an oscillation set up in one region of the plasma remains localized in that region. It should be noted, however, that in a "warm" plasma (*i.e.*, a plasma with a finite thermal velocity) the plasma wave acquires a non-zero, albeit very small, group velocity (see Sect. 6.2).

The second wave is a *transverse* wave, with $\mathbf{k}$ perpendicular to $\mathbf{E}$. There are two independent linear polarizations of this wave, which propagate at identical velocities, just like a vacuum electromagnetic wave. The dispersion relation (4.70) can be rearranged to give

$$\omega^2 = \Pi_e^{\,2} + k^2 c^2, \tag{4.72}$$

showing that this wave is just the conventional electromagnetic wave, whose vacuum dispersion relation is $\omega^2 = k^2 c^2$, modified by the presence of the plasma. An important property, which follows immediately from the above expression, is that for the propagation of this wave we need $\omega \geq \Pi_e$. Since Π_e is proportional to the square root of the plasma density, it follows that electromagnetic radiation of a given frequency will only propagate through a plasma when the plasma density falls below a critical value.

4.8 Low-Frequency Wave Propagation

Let us now consider wave propagation through a magnetized plasma at frequencies far below the ion cyclotron or plasma frequencies, which are, in turn, well below the corresponding electron frequencies. In the low-frequency regime (*i.e.*, $\omega \ll \Omega_i, \Pi_i$), we have [see Eqs. (4.27)–(4.29)]

$$S \simeq 1 + \frac{\Pi_i^{\,2}}{\Omega_i^2}, \tag{4.73}$$

$$D \simeq 0, \tag{4.74}$$

$$P \simeq -\frac{\Pi_e^{\,2}}{\omega^2}. \tag{4.75}$$

Here, use has been made of $\Pi_e^{\,2}/\Omega_e\Omega_i = -\Pi_i^{\,2}/\Omega_i^{\,2}$. Thus, the eigenmode equation (4.41) reduces to

$$\begin{pmatrix} 1+\Pi_i^{\,2}/\Omega_i^2 - n^2\cos^2\theta & 0 & n^2\cos\theta\sin\theta \\ 0 & 1+\Pi_i^{\,2}/\Omega_i^2 - n^2 & 0 \\ n^2\cos\theta\sin\theta & 0 & -\Pi_e^{\,2}/\omega^2 - n^2\sin^2\theta \end{pmatrix} \begin{pmatrix} E_x \\ E_y \\ E_z \end{pmatrix} = \mathbf{0}. \tag{4.76}$$

The solubility condition for Eq. (4.76) yields the dispersion relation

$$\begin{vmatrix} 1+\Pi_i^2/\Omega_i^2 - n^2\cos^2\theta & 0 & n^2\cos\theta\sin\theta \\ 0 & 1+\Pi_i^2/\Omega_i^2 - n^2 & 0 \\ n^2\cos\theta\sin\theta & 0 & -\Pi_e^2/\omega^2 - n^2\sin^2\theta \end{vmatrix} = 0. \tag{4.77}$$

Note that in the low-frequency ordering, $\Pi_e^2/\omega^2 \gg \Pi_i^2/\Omega_i^2$. Thus, we can see that the bottom right-hand element of the above determinant is far larger than any of the other elements, so to a good approximation the roots of the dispersion relation are obtained by equating the term multiplying this large factor to zero. In this manner, we obtain two roots:

$$n^2\cos^2\theta = 1 + \frac{\Pi_i^2}{\Omega_i^2}, \tag{4.78}$$

and

$$n^2 = 1 + \frac{\Pi_i^2}{\Omega_i^2}. \tag{4.79}$$

It is fairly easy to show, from the definitions of the plasma and cyclotron frequencies [see Eqs. (4.16)–(4.19], that

$$\frac{\Pi_i^2}{\Omega_i^2} = \frac{c^2}{B_0^2/\mu_0\rho} = \frac{c^2}{V_A^2}. \tag{4.80}$$

Here, $\rho \simeq n\, m_i$ is the plasma mass density, and

$$V_A = \sqrt{\frac{B_0^2}{\mu_0\,\rho}} \tag{4.81}$$

is called the *Alfvén velocity*. Thus, the dispersion relations of the two low-frequency waves can be written

$$\omega = \frac{k\,V_A\,\cos\theta}{\sqrt{1+V_A^2/c^2}} \simeq k\,V_A\,\cos\theta \equiv k_\parallel\,V_A, \tag{4.82}$$

and

$$\omega = \frac{k\,V_A}{\sqrt{1+V_A^2/c^2}} \simeq k\,V_A. \tag{4.83}$$

Here, we have made use of the fact that $V_A \ll c$ in conventional plasmas.

The dispersion relation (4.82) corresponds to the *slow* or *shear* Alfvén wave, whereas the dispersion relation (4.83) corresponds to the *fast* or *compressional* Alfvén wave. The fast/slow terminology simply refers to the ordering of the phase velocities of the two waves. The shear/compressional terminology refers to the velocity fields associated with the waves. In fact, it is clear from Eq. (4.76) that $E_z = 0$ for both waves, whereas $E_y = 0$ for the shear wave, and $E_x = 0$ for the compressional wave. Both waves are, in fact, MHD modes which satisfy the linearized MHD Ohm's law [see Eq. (3.222)]

$$\mathbf{E} + \mathbf{V}\times\mathbf{B}_0 = 0. \tag{4.84}$$

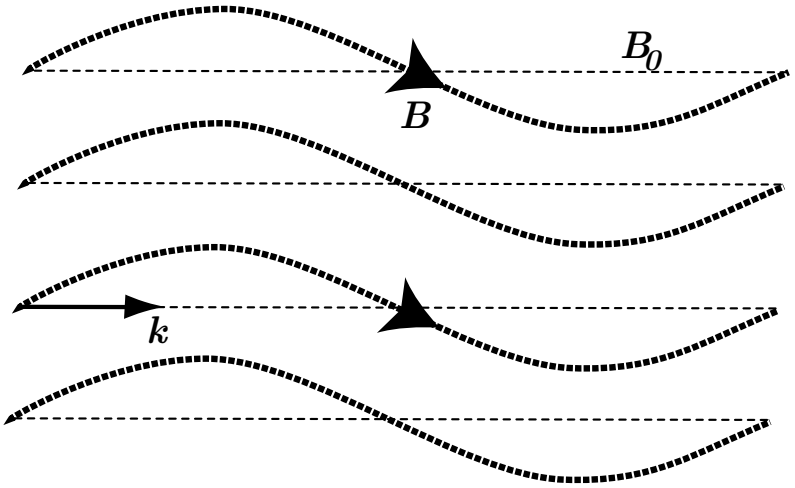

Figure 4.1: *Magnetic field perturbation associated with a shear-Alfvén wave.*

Thus, for the shear wave

$$V_y = -\frac{E_x}{B_0}, \tag{4.85}$$

and $V_x = V_z = 0$, whereas for the compressional wave

$$V_x = \frac{E_y}{B_0}, \tag{4.86}$$

and $V_y = V_z = 0$. Now $\nabla\cdot\mathbf{V} = \mathrm{i}\,\mathbf{k}\cdot\mathbf{V} = \mathrm{i}\,k\,V_x\,\sin\theta$. Thus, the shear-Alfvén wave is a torsional wave, with zero divergence of the flow, whereas the compressional wave involves a non-zero divergence of the flow. It is important to realize that the thing which is resisting compression in the compressional wave is the magnetic field, not the plasma, since there is negligible plasma pressure in the cold-plasma approximation.

Figure 4.1 shows the characteristic distortion of the magnetic field associated with a shear-Alfvén wave propagating parallel to the equilibrium field. Clearly, this wave bends magnetic field-lines without compressing them. Figure 4.2 shows the characteristic distortion of the magnetic field associated with a compressional-Alfvén wave propagating perpendicular to the equilibrium field. Clearly, this wave compresses magnetic field-lines without bending them.

It should be noted that the thermal velocity is not necessarily negligible compared to the Alfvén velocity in conventional plasmas. Thus, we can expect the dispersion relations (4.82) and (4.83) to undergo considerable modification in a "warm" plasma (see Sect. 5.4).

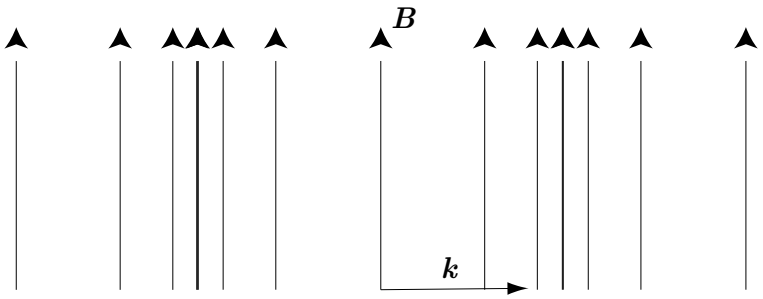

Figure 4.2: *Magnetic field perturbation associated with a compressional Alfvén-wave.*

4.9 Parallel Wave Propagation

Let us now consider wave propagation, at arbitrary frequencies, *parallel* to the equilibrium magnetic field. When $\theta = 0$, the eigenmode equation (4.41) simplifies to

$$\begin{pmatrix} S - n^2 & -i\,D & 0 \\ i\,D & S - n^2 & 0 \\ 0 & 0 & P \end{pmatrix} \begin{pmatrix} E_x \\ E_y \\ E_z \end{pmatrix} = \mathbf{0}. \tag{4.87}$$

One obvious way of solving this equation is to have

$$P \simeq 1 - \frac{\Pi_e^{\,2}}{\omega^2} = 0, \tag{4.88}$$

with the eigenvector $(0, 0, E_z)$. This is just the *electrostatic plasma wave* which we found previously in an unmagnetized plasma. This mode is longitudinal in nature, and, therefore, causes particles to oscillate *parallel* to $\mathbf{B}_0$. It follows that the particles experience zero Lorentz force due to the presence of the equilibrium magnetic field, with the result that this field has no effect on the mode dynamics.

The other two solutions to Eq. (4.87) are obtained by setting the 2×2 determinant involving the x- and y- components of the electric field to zero. The first wave has the dispersion relation

$$n^2 = R \simeq 1 - \frac{\Pi_e^{\,2}}{(\omega + \Omega_e)(\omega + \Omega_i)}, \tag{4.89}$$

and the eigenvector $(E_x, i\,E_x, 0)$. This is evidently a *right-handed* circularly polarized wave. The second wave has the dispersion relation

$$n^2 = L \simeq 1 - \frac{\Pi_e^{\,2}}{(\omega - \Omega_e)(\omega - \Omega_i)}, \tag{4.90}$$

and the eigenvector $(E_x, -i\,E_x, 0)$. This is evidently a *left-handed* circularly polarized wave. At low frequencies (*i.e.*, $\omega \ll \Omega_i$), both waves tend to the Alfvén wave found previously. Note that the fast and slow Alfvén waves are indistinguishable for parallel propagation. Let us now examine the high-frequency behaviour of the right- and left-handed waves.

For the right-handed wave, it is evident, since Ω_e is negative, that $n^2 \rightarrow \infty$ as $\omega \rightarrow |\Omega_e|$. This resonance, which corresponds to $R \rightarrow \infty$, is termed the *electron cyclotron resonance*. At the electron cyclotron resonance the transverse electric field associated with a right-handed wave rotates at the same velocity, and in the same direction, as electrons gyrating around the equilibrium magnetic field. Thus, the electrons experience a *continuous* acceleration from the electric field, which tends to increase their perpendicular energy. It is, therefore, not surprising that right-handed waves, propagating parallel to the equilibrium magnetic field, and oscillating at the frequency Ω_e, are *absorbed* by electrons.

When ω is just above $|\Omega_e|$, we find that n^2 is negative, and so there is no wave propagation in this frequency range. However, for frequencies much greater than the electron cyclotron or plasma frequencies, the solution to Eq. (4.89) is approximately $n^2 = 1$. In other words, $\omega^2 = k^2c^2$: the dispersion relation of a right-handed vacuum electromagnetic wave. Evidently, at some frequency above $|\Omega_e|$ the solution for n^2 must pass through zero, and become positive again. Putting $n^2 = 0$ in Eq. (4.89), we find that the equation reduces to

$$\omega^2 + \Omega_e\,\omega - \Pi_e^{\,2} \simeq 0, \tag{4.91}$$

assuming that $V_A \ll c$. The above equation has only one positive root, at $\omega = \omega_1$, where

$$\omega_1 \simeq |\Omega_e|/2 + \sqrt{\Omega_e^{\,2}/4 + \Pi_e^{\,2}} > |\Omega_e|. \tag{4.92}$$

Above this frequency, the wave propagates once again.

The dispersion curve for a right-handed wave propagating parallel to the equilibrium magnetic field is sketched in Fig. 4.3. The continuation of the Alfvén wave above the ion cyclotron frequency is called the *electron cyclotron wave*, or sometimes the *whistler wave*. The latter terminology is prevalent in ionospheric and space plasma physics contexts. The wave which propagates above the cutoff frequency, ω_1, is a standard right-handed circularly polarized electromagnetic wave, somewhat modified by the presence of the plasma. Note that the low-frequency branch of the dispersion curve differs fundamentally from the high-frequency branch, because the former branch corresponds to a wave which can only propagate through the plasma in the presence of an equilibrium magnetic field, whereas the high-frequency branch corresponds to a wave which can propagate in the absence of an equilibrium field.

The curious name "whistler wave" for the branch of the dispersion relation lying between the ion and electron cyclotron frequencies is originally derived from ionospheric physics. Whistler waves are a very characteristic type of audio-frequency radio interference, most commonly encountered at high latitudes, which take the form of brief, intermittent pulses, starting at high frequencies, and rapidly descending in pitch. Figure 4.4 shows the power spectra of some typical whistler waves.

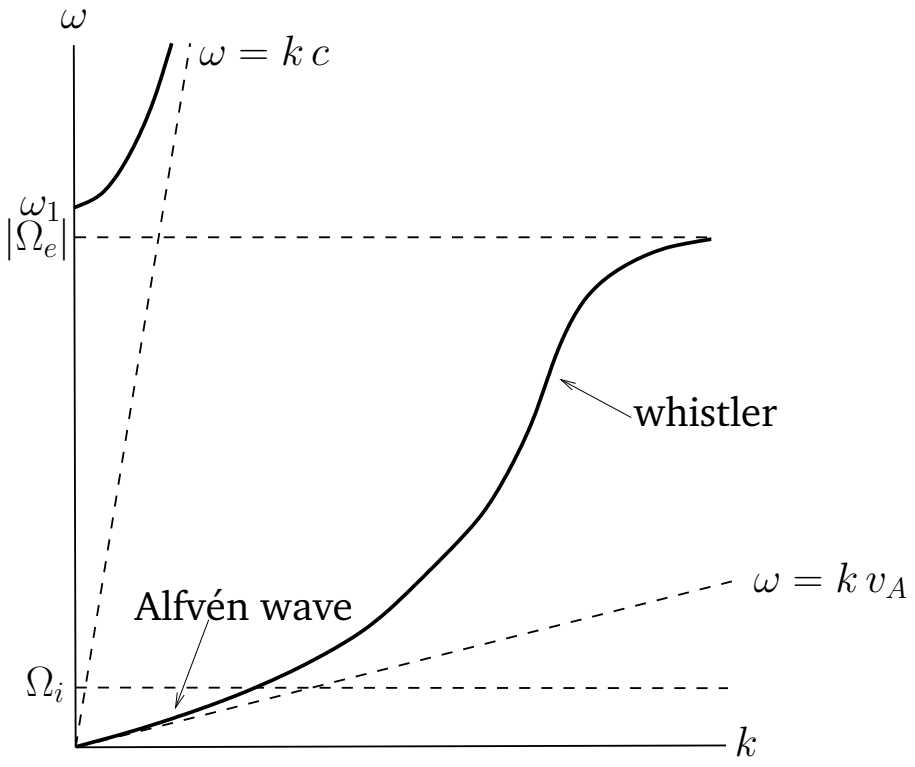

Figure 4.3: *Dispersion relation for a right-handed wave propagating parallel to the magnetic field in a magnetized plasma.*

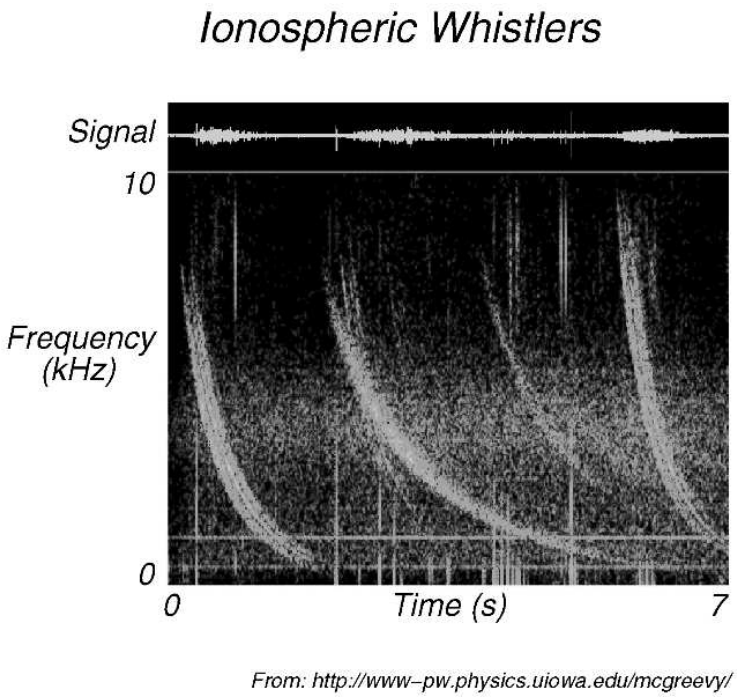

Figure 4.4: *Power spectrum of a typical whistler wave.*

Whistlers were discovered in the early days of radio communication, but were not explained until much later. Whistler waves start off as "instantaneous" radio pulses, generated by lightning flashes at high latitudes. The pulses are channeled along the Earth's dipolar magnetic field, and eventually return to ground level in the opposite hemisphere. Fig. 4.5 illustrates the typical path of a whistler wave. Now, in the frequency range $\Omega_i \ll \omega \ll |\Omega_e|$, the dispersion relation (4.89) reduces to

$$n^2 = \frac{k^2 c^2}{\omega^2} \simeq \frac{{\Pi_e}^2}{\omega\,|\Omega_e|}. \tag{4.93}$$

As is well-known, pulses propagate at the group-velocity,

$$v_g = \frac{d\omega}{dk} = 2c\,\frac{\sqrt{\omega\,|\Omega_e|}}{\Pi_e}. \tag{4.94}$$

Clearly, the low-frequency components of a pulse propagate *more slowly* than the high-frequency components. It follows that by the time a pulse returns to ground level it has been stretched out temporally, because the high-frequency components of the pulse arrive slightly before the low-frequency components. This also accounts for the characteristic whistling-down effect observed at ground level.

The shape of whistler pulses, and the way in which the pulse frequency varies in time, can yield a considerable amount of information about the regions of the Earth's magnetosphere through which they have passed. For this reason, many countries maintain

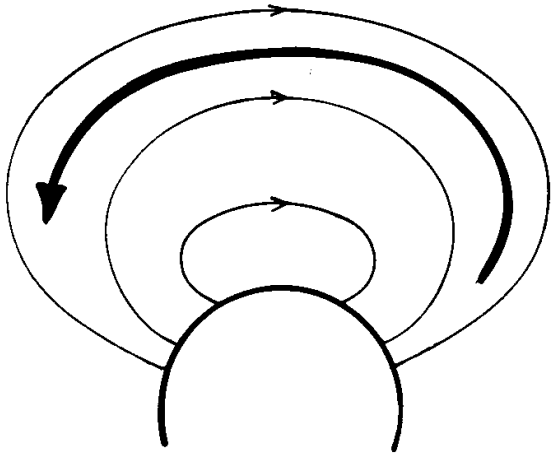

Figure 4.5: *Typical path of a whistler wave through the Earth's magnetosphere.*

observatories in polar regions, especially Antarctica, which monitor and collect whistler data: *e.g.*, the Halley research station, maintained by the British Antarctic Survey, which is located on the edge of the Antarctic mainland.

For a left-handed circularly polarized wave, similar considerations to the above give a dispersion curve of the form sketched in Fig. 4.6. In this case, n^2 goes to infinity at the ion cyclotron frequency, Ω_i, corresponding to the so-called *ion cyclotron resonance* (at $L \rightarrow \infty$). At this resonance, the rotating electric field associated with a left-handed wave resonates with the gyromotion of the ions, allowing wave energy to be converted into perpendicular kinetic energy of the ions. There is a band of frequencies, lying above the ion cyclotron frequency, in which the left-handed wave does not propagate. At very high frequencies a propagating mode exists, which is basically a standard left-handed circularly polarized electromagnetic wave, somewhat modified by the presence of the plasma. The cutoff frequency for this wave is

$$\omega_2 \simeq -|\Omega_e|/2 + \sqrt{\Omega_e^2/4 + \Pi_e^2}. \tag{4.95}$$

As before, the lower branch in Fig. 4.6 describes a wave that can only propagate in the presence of an equilibrium magnetic field, whereas the upper branch describes a wave that can propagate in the absence an equilibrium field. The continuation of the Alfvén wave to just below the ion cyclotron frequency is generally called the *ion cyclotron wave*.

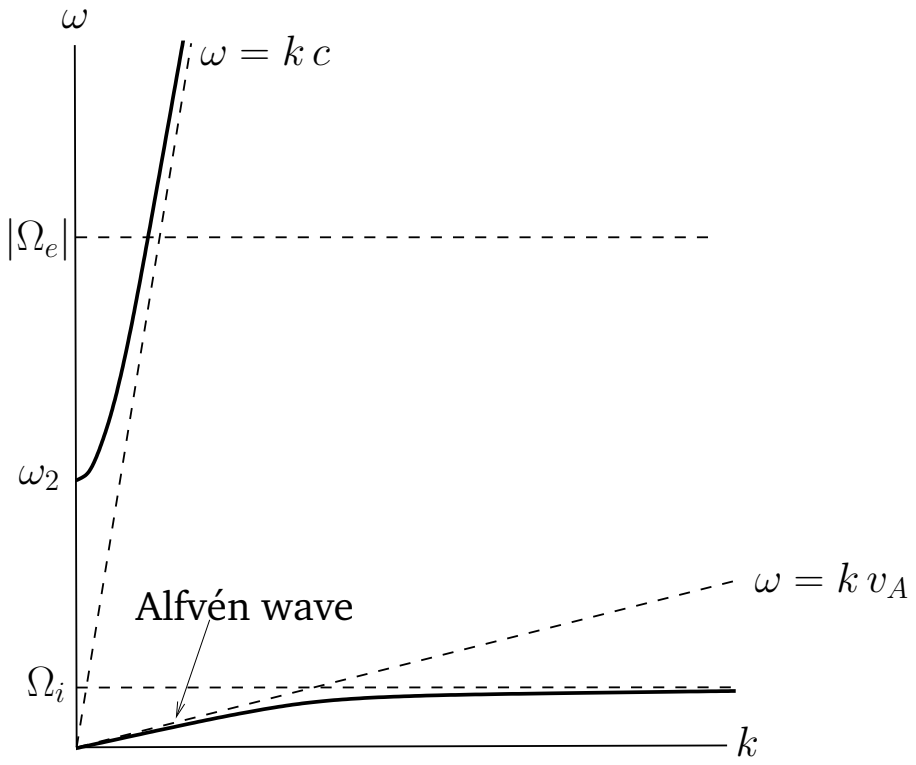

Figure 4.6: *Dispersion relation for a left-handed wave propagating parallel to the magnetic field in a magnetized plasma.*

4.10 Perpendicular Wave Propagation

Let us now consider wave propagation, at arbitrary frequencies, *perpendicular* to the equilibrium magnetic field. When $\theta = \pi/2$, the eigenmode equation (4.41) simplifies to

$$\begin{pmatrix} S & -i\,D & 0 \\ i\,D & S-n^2 & 0 \\ 0 & 0 & P-n^2 \end{pmatrix} \begin{pmatrix} E_x \\ E_y \\ E_z \end{pmatrix} = \mathbf{0}. \tag{4.96}$$

One obvious way of solving this equation is to have $P - n^2 = 0$, or

$$\omega^2 = \Pi_e^{\,2} + k^2 c^2, \tag{4.97}$$

with the eigenvector $(0, 0, E_z)$. Since the wave-vector now points in the x-direction, this is clearly a transverse wave polarized with its electric field parallel to the equilibrium magnetic field. Particle motions are along the magnetic field, so the mode dynamics are completely unaffected by this field. Thus, the wave is identical to the *electromagnetic plasma wave* found previously in an unmagnetized plasma. This wave is known as the *ordinary*, or O-, mode.

The other solution to Eq. (4.96) is obtained by setting the 2×2 determinant involving the x- and y- components of the electric field to zero. The dispersion relation reduces to

$$n^2 = \frac{R\,L}{S}, \tag{4.98}$$

with the associated eigenvector $E_x\,(1, -i\,S/D, 0)$.

Let us, first of all, search for the cutoff frequencies, at which n^2 goes to zero. According to Eq. (4.98), these frequencies are the roots of $R = 0$ and $L = 0$. In fact, we have already solved these equations (recall that cutoff frequencies do not depend on θ). There are two cutoff frequencies, ω_1 and ω_2, which are specified by Eqs. (4.92) and (4.95), respectively.

Let us, next, search for the resonant frequencies, at which n^2 goes to infinity. According to Eq. (4.98), the resonant frequencies are solutions of

$$S = 1 - \frac{\Pi_e^{\,2}}{\omega^2 - \Omega_e^{\,2}} - \frac{\Pi_i^{\,2}}{\omega^2 - \Omega_i^{\,2}} = 0. \tag{4.99}$$

The roots of this equations can be obtained as follows. First, we note that if the first two terms are equated to zero, we obtain $\omega = \omega_{UH}$, where

$$\omega_{UH} = \sqrt{\Pi_e^{\,2} + \Omega_e^{\,2}}. \tag{4.100}$$

If this frequency is substituted into the third term, the result is far less than unity. We conclude that ω_{UH} is a good approximation to one of the roots of Eq. (4.99). To obtain the second root, we make use of the fact that the product of the square of the roots is

$$\Omega_e^{\,2}\,\Omega_i^{\,2} + \Pi_e^{\,2}\,\Omega_i^{\,2} + \Pi_i^{\,2}\,\Omega_e^{\,2} \simeq \Omega_e^{\,2}\,\Omega_i^{\,2} + \Pi_i^{\,2}\,\Omega_e^{\,2}. \tag{4.101}$$

We, thus, obtain $\omega = \omega_{\rm LH}$, where

$$\omega_{\rm LH} = \sqrt{\frac{\Omega_e^{\,2}\,\Omega_i^{\,2} + \Pi_i^{\,2}\,\Omega_e^{\,2}}{\Pi_e^{\,2} + \Omega_e^{\,2}}}. \tag{4.102}$$

The first resonant frequency, $\omega_{\rm UH}$, is greater than the electron cyclotron or plasma frequencies, and is called the *upper hybrid frequency*. The second resonant frequency, $\omega_{\rm LH}$, lies between the electron and ion cyclotron frequencies, and is called the *lower hybrid frequency*.

Unfortunately, there is no simple explanation of the origins of the two hybrid resonances in terms of the motions of individual particles.

At low frequencies, the mode in question reverts to the compressional-Alfvén wave discussed previously. Note that the shear-Alfvén wave does not propagate perpendicular to the magnetic field.

Using the above information, and the easily demonstrated fact that

$$\omega_{\rm LH} < \omega_2 < \omega_{\rm UH} < \omega_1, \tag{4.103}$$

we can deduce that the dispersion curve for the mode in question takes the form sketched in Fig. 4.7. The lowest frequency branch corresponds to the compressional-Alfvén wave. The other two branches constitute the *extraordinary*, or X-, wave. The upper branch is basically a linearly polarized (in the y-direction) electromagnetic wave, somewhat modified by the presence of the plasma. This branch corresponds to a wave which propagates in the absence of an equilibrium magnetic field. The lowest branch corresponds to a wave which does not propagate in the absence of an equilibrium field. Finally, the middle branch corresponds to a wave which converts into an electrostatic plasma wave in the absence of an equilibrium magnetic field.

Wave propagation at oblique angles is generally more complicated than propagation parallel or perpendicular to the equilibrium magnetic field, but does not involve any new physical effects.

4.11 Wave Propagation Through Inhomogeneous Plasmas

Up to now, we have only analyzed wave propagation through homogeneous plasmas. Let us now broaden our approach to take into account the far more realistic case of wave propagation through *inhomogeneous* plasmas.

Let us start off by examining a very simple case. Consider a plane electromagnetic wave, of frequency ω, propagating along the z-axis in an unmagnetized plasma whose refractive index, n, is a function of z. We assume that the wave normal is initially aligned along the z-axis, and, furthermore, that the wave starts off polarized in the y-direction. It is easily demonstrated that the wave normal subsequently remains aligned along the z-axis, and also that the polarization state of the wave does not change. Thus, the wave is fully described by

$$E_y(z,t) \equiv E_y(z)\,\exp(-{\rm i}\,\omega\,t), \tag{4.104}$$

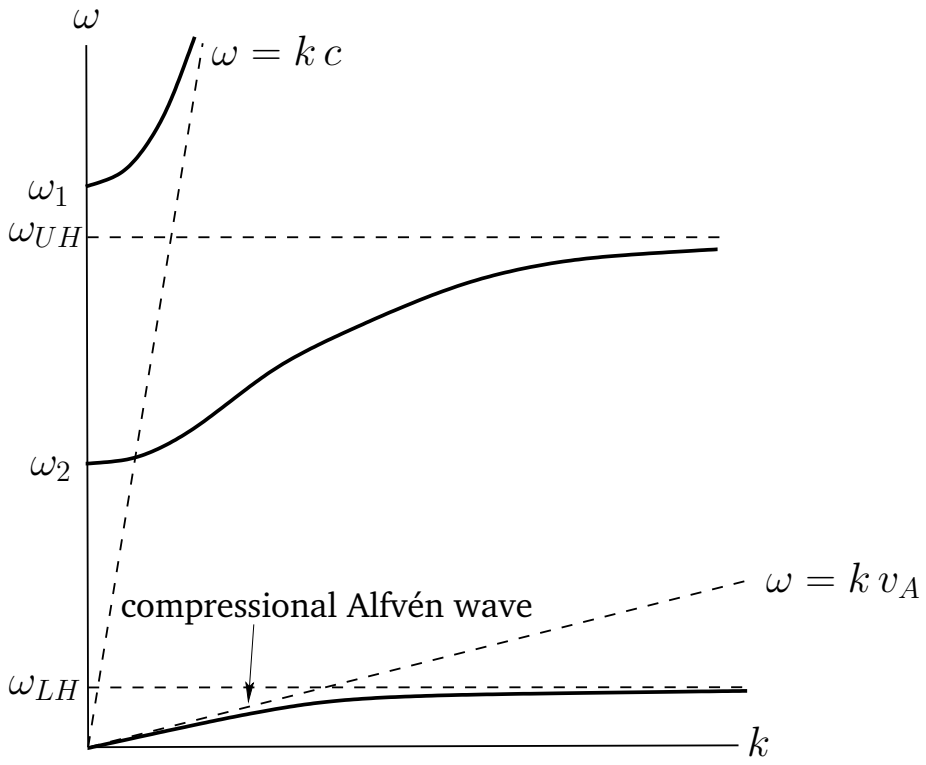

Figure 4.7: *Dispersion relation for a wave propagating perpendicular to the magnetic field in a magnetized plasma.*

and

$$B_x(z,t) \equiv B_x(z)\,\exp(-i\,\omega\,t). \tag{4.105}$$

It can easily be shown that $E_y(z)$ and $B_x(z)$ satisfy the differential equations

$$\frac{d^2E_y}{dz^2} + k_0^{\,2}\,n^2\,E_y = 0, \tag{4.106}$$

and

$$\frac{d\,cB_x}{dz} = -i\,k_0\,n^2\,E_y, \tag{4.107}$$

respectively. Here, $k_0 = \omega/c$ is the wave-number in free space. Of course, the actual wave-number is $k = k_0\,n$.

The solution to Eq. (4.106) for the case of a homogeneous plasma, for which n is constant, is straightforward:

$$E_y = A\,e^{\,i\,\phi(z)}, \tag{4.108}$$

where A is a constant, and

$$\phi = \pm k_0\,n\,z. \tag{4.109}$$

The solution (4.108) represents a wave of constant amplitude, A, and phase, $\phi(z)$. According to Eq. (4.109), there are, in fact, two independent waves which can propagate through the plasma. The upper sign corresponds to a wave which propagates in the $+z$-direction, whereas the lower sign corresponds to a wave which propagates in the $-z$-direction. Both waves propagate with the constant phase velocity c/n.

In general, if $n = n(z)$ then the solution of Eq. (4.106) does not remotely resemble the wave-like solution (4.108). However, in the limit in which $n(z)$ is a "slowly varying" function of z (exactly how slowly varying is something which will be established later on), we expect to recover wave-like solutions. Let us suppose that $n(z)$ is indeed a "slowly varying" function, and let us try substituting the wave solution (4.108) into Eq. (4.106). We obtain

$$\left(\frac{d\phi}{dz}\right)^2 = k_0^{\,2}n^2 + i\,\frac{d^2\phi}{dz^2}. \tag{4.110}$$

This is a non-linear differential equation which, in general, is very difficult to solve. However, we note that if n is a constant then $d^2\phi/dz^2 = 0$. It is, therefore, reasonable to suppose that if $n(z)$ is a "slowly varying" function then the last term on the right-hand side of the above equation can be regarded as being small. Thus, to a first approximation Eq. (4.91) yields

$$\frac{d\phi}{dz} \simeq \pm k_0\,n, \tag{4.111}$$

and

$$\frac{d^2\phi}{dz^2} \simeq \pm k_0\,\frac{dn}{dz}. \tag{4.112}$$

It is clear from a comparison of Eqs. (4.110) and (4.112) that $n(z)$ can be regarded as a "slowly varying" function of z as long as its variation length-scale is far longer than the wave-length of the wave. In other words, provided that $(dn/dz)/(k_0\,n^2) \ll 1$.

The second approximation to the solution is obtained by substituting Eq. (4.112) into the right-hand side of Eq. (4.110):

$$\frac{d\phi}{dz} \simeq \pm \left({k_0}^2 n^2 \pm i\, k_0 \frac{dn}{dz} \right)^{1/2}. \tag{4.113}$$

This gives

$$\frac{d\phi}{dz} \simeq \pm k_0\, n \left(1 \pm \frac{i}{k_0\, n^2} \frac{dn}{dz} \right)^{1/2} \simeq \pm k_0\, n + \frac{i}{2n} \frac{dn}{dz}, \tag{4.114}$$

where use has been made of the binomial expansion. The above expression can be integrated to give

$$\phi \sim \pm k_0 \int^z n\, dz + i\, \log(n^{1/2}). \tag{4.115}$$

Substitution of Eq. (4.115) into Eq. (4.108) yields the final result

$$E_y \simeq A\, n^{-1/2}\, \exp\left(\pm i\, k_0 \int^z n\, dz \right). \tag{4.116}$$

It follows from Eq. (4.107) that

$$cB_x \simeq \mp A\, n^{1/2}\, \exp\left(\pm i\, k_0 \int^z n\, dz \right) - \frac{i\, A}{2k_0\, n^{3/2}} \frac{dn}{dz}\, \exp\left(\pm i\, k_0 \int^z n\, dz \right). \tag{4.117}$$

Note that the second term is small compared to the first, and can usually be neglected.

Let us test to what extent the expression (4.116) is a good solution of Eq. (4.106) by substituting this expression into the left-hand side of the equation. The result is

$$\frac{A}{n^{1/2}} \left\{ \frac{3}{4} \left(\frac{1}{n} \frac{dn}{dz} \right)^2 - \frac{1}{2n} \frac{d^2 n}{dz^2} \right\} \exp\left(\pm i\, k_0 \int^z n\, dz \right). \tag{4.118}$$

This must be small compared with either term on the left-hand side of Eq. (4.106). Hence, the condition for Eq. (4.116) to be a good solution of Eq. (4.106) becomes

$$\frac{1}{{k_0}^2} \left| \frac{3}{4} \left(\frac{1}{n^2} \frac{dn}{dz} \right)^2 - \frac{1}{2n^3} \frac{d^2 n}{dz^2} \right| \ll 1. \tag{4.119}$$

The solutions

$$E_y \simeq A\, n^{-1/2}\, \exp\left(\pm i\, k_0 \int^z n\, dz \right), \tag{4.120}$$

$$cB_x \simeq \mp A\, n^{1/2}\, \exp\left(\pm i\, k_0 \int^z n\, dz \right), \tag{4.121}$$

to the non-uniform wave equations (4.106) and (4.107) are most commonly referred to as the *WKB solutions*, in honour of G. Wentzel, H.A. Kramers, and L. Brillouin, who are credited with independently discovering these solutions (in a quantum mechanical context) in 1926. Actually, H. Jeffries wrote a paper on the WKB solutions (in a wave propagation context) in 1923. Hence, some people call them the WKBJ solutions (or even the JWKB solutions). To be strictly accurate, the WKB solutions were first discussed by Liouville and Green in 1837, and again by Rayleigh in 1912. In the following, we refer to Eqs. (4.120)–(4.121) as the WKB solutions, since this is what they are most commonly known as. However, it should be understand that, in doing so, we are not making any definitive statement as to the credit due to various scientists in discovering them.

Recall, that when a propagating wave is normally incident on an *interface*, where the refractive index suddenly changes (for instance, when a light wave propagating through air is normally incident on a glass slab), there is generally significant reflection of the wave. However, according to the WKB solutions, (4.120)–(4.121), when a propagating wave is normally incident on a medium in which the refractive index changes *slowly* along the direction of propagation of the wave then the wave is not reflected at all. This is true even if the refractive index varies *very substantially* along the path of propagation of the wave, as long as it varies *slowly*. The WKB solutions imply that as the wave propagates through the medium its wave-length gradually changes. In fact, the wave-length at position z is approximately $\lambda(z) = 2\pi/k_0\, n(z)$. Equations (4.120)–(4.121) also imply that the amplitude of the wave gradually changes as it propagates. In fact, the amplitude of the electric field component is inversely proportional to $n^{1/2}$, whereas the amplitude of the magnetic field component is directly proportional to $n^{1/2}$. Note, however, that the energy flux in the z-direction, given by the the Poynting vector $-(E_y B_x^* + E_y^* B_x)/(4\mu_0)$, remains constant (assuming that n is predominately real).

Of course, the WKB solutions (4.120)–(4.121) are only *approximations*. In reality, a wave propagating into a medium in which the refractive index is a slowly varying function of position is subject to a small amount of reflection. However, it is easily demonstrated that the ratio of the reflected amplitude to the incident amplitude is of order $(dn/dz)/(k_0\, n^2)$. Thus, as long as the refractive index varies on a much longer length-scale than the wave-length of the radiation, the reflected wave is negligibly small. This conclusion remains valid as long as the inequality (4.119) is satisfied. This inequality obviously breaks down in the vicinity of a point where $n^2 = 0$. We would, therefore, expect strong reflection of the incident wave from such a point. Furthermore, the WKB solutions also break down at a point where $n^2 \rightarrow \infty$, since the amplitude of B_x becomes infinite.

4.12 Cutoffs

We have seen that electromagnetic wave propagation (in one dimension) through an inhomogeneous plasma, in the physically relevant limit in which the variation length-scale of the plasma is much greater than the wave-length of the wave, is well described by the WKB solutions, (4.120)–(4.121). However, these solutions break down in the immediate

vicinity of a *cutoff*, where $n^2 = 0$, or a *resonance*, where $n^2 \rightarrow \infty$. Let us now examine what happens to electromagnetic waves propagating through a plasma when they encounter a cutoff or a resonance.

Suppose that a cutoff is located at $z = 0$, so that

$$n^2 = a\, z + O(z^2) \tag{4.122}$$

in the immediate vicinity of this point, where $a > 0$. It is evident, from the WKB solutions, (4.120)–(4.121), that the cutoff point lies at the boundary between a region ($z > 0$) in which electromagnetic waves propagate, and a region ($z < 0$) in which the waves are evanescent. In a physically realistic solution, we would expect the wave amplitude to decay (as z decreases) in the evanescent region $z < 0$. Let us search for such a wave solution.

In the immediate vicinity of the cutoff point, $z = 0$, Eqs. (4.106) and (4.122) yield

$$\frac{d^2 E_y}{d\hat{z}^2} + \hat{z}\, E_y = 0, \tag{4.123}$$

where

$$\hat{z} = (k_0^2\, a)^{1/3}\, z. \tag{4.124}$$

Equation (4.123) is a standard equation, known as *Airy's equation*, and possesses two independent solutions, denoted $\mathrm{Ai}(-\hat{z})$ and $\mathrm{Bi}(-\hat{z})$.[1] The second solution, $\mathrm{Bi}(-\hat{z})$, is unphysical, since it blows up as $\hat{z} \rightarrow -\infty$. The physical solution, $\mathrm{Ai}(-\hat{z})$, has the asymptotic behaviour

$$\mathrm{Ai}(-\hat{z}) \sim \frac{1}{2\sqrt{\pi}}\, |\hat{z}|^{-1/4}\, \exp\left(-\frac{2}{3}\, |\hat{z}|^{3/2}\right) \tag{4.125}$$

in the limit $\hat{z} \rightarrow -\infty$, and

$$\mathrm{Ai}(-\hat{z}) \sim \frac{1}{\sqrt{\pi}}\, \hat{z}^{-1/4}\, \sin\left(\frac{2}{3}\, \hat{z}^{3/2} + \frac{\pi}{4}\right) \tag{4.126}$$

in the limit $\hat{z} \rightarrow +\infty$.

Suppose that a unit amplitude plane electromagnetic wave, polarized in the y-direction, is launched from an antenna, located at large positive z, towards the cutoff point at $z = 0$. It is assumed that $n = 1$ at the launch point. In the non-evanescent region, $z > 0$, the wave can be represented as a linear combination of propagating WKB solutions:

$$E_y(z) = n^{-1/2}\, \exp\left(-i\, k_0 \int_0^z n\, dz\right) + R\, n^{-1/2}\, \exp\left(+i\, k_0 \int_0^z n\, dz\right). \tag{4.127}$$

The first term on the right-hand side of the above equation represents the incident wave, whereas the second term represents the reflected wave. The complex constant R is the

[1] M. Abramowitz, and I.A. Stegun, *Handbook of Mathematical Functions* (Dover, New York NY, 1964), p. 446.

coefficient of reflection. In the vicinity of the cutoff point (*i.e.*, z small and positive, or $\hat{z}$ large and positive) the above expression reduces to

$$E_y(\hat{z}) = (k_0/a)^{1/6}\left[\hat{z}^{-1/4}\exp\left(-i\frac{2}{3}\hat{z}^{3/2}\right) + R\,\hat{z}^{-1/4}\exp\left(+i\frac{2}{3}\hat{z}^{3/2}\right)\right]. \tag{4.128}$$

However, we have another expression for the wave in this region. Namely,

$$E_y(\hat{z}) = C\,Ai(-\hat{z}) \simeq \frac{C}{\sqrt{\pi}}\,\hat{z}^{-1/4}\sin\left(\frac{2}{3}\hat{z}^{3/2}+\frac{\pi}{4}\right), \tag{4.129}$$

where C is an arbitrary constant. The above equation can be written

$$E_y(\hat{z}) = \frac{C}{2}\sqrt{\frac{i}{\pi}}\left[\hat{z}^{-1/4}\exp\left(-i\frac{2}{3}\hat{z}^{3/2}\right) - i\,\hat{z}^{-1/4}\exp\left(+i\frac{2}{3}\hat{z}^{3/2}\right)\right]. \tag{4.130}$$

A comparison of Eqs. (4.128) and (4.130) yields

$$R = -i. \tag{4.131}$$

In other words, at a cutoff point there is *total reflection*, since $|R| = 1$, with a $-\pi/2$ phase-shift.

4.13 Resonances

Suppose, now, that a resonance is located at $z = 0$, so that

$$n^2 = \frac{b}{z + i\,\epsilon} + O(1) \tag{4.132}$$

in the immediate vicinity of this point, where $b > 0$. Here, ϵ is a small real constant. We introduce ϵ at this point principally as a mathematical artifice to ensure that E_y remains single-valued and finite. However, as will become clear later on, ϵ has a physical significance in terms of damping or spontaneous excitation.

In the immediate vicinity of the resonance point, $z = 0$, Eqs. (4.106) and (4.132) yield

$$\frac{d^2E_y}{d\hat{z}^2} + \frac{E_y}{\hat{z} + i\,\hat{\epsilon}} = 0, \tag{4.133}$$

where

$$\hat{z} = (k_0^{\,2}\,b)\,z, \tag{4.134}$$

and $\hat{\epsilon} = (k_0^{\,2}\,b)\,\epsilon$. This equation is singular at the point $\hat{z} = -i\,\hat{\epsilon}$. Thus, it is necessary to introduce a branch-cut into the complex-$\hat{z}$ plane in order to ensure that $E_y(\hat{z})$ is single-valued. If $\epsilon > 0$ then the branch-cut lies in the lower half-plane, whereas if $\epsilon < 0$ then the branch-cut lies in the upper half-plane—see Fig. 4.8. Suppose that the argument of $\hat{z}$ is 0

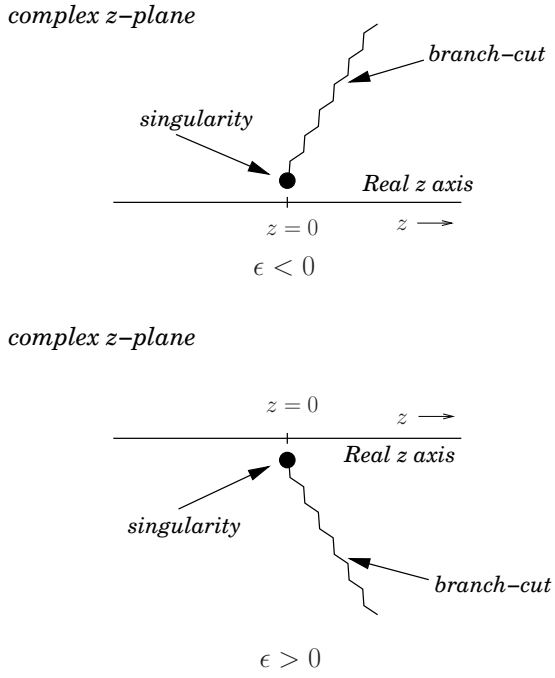

Figure 4.8: *Branch-cuts in the z-plane close to a wave resonance.*

on the positive real $\hat{z}$-axis. It follows that the argument of $\hat{z}$ on the negative real $\hat{z}$-axis is $+\pi$ when $\epsilon > 0$ and $-\pi$ when $\epsilon < 0$.

Let

$$y = 2\sqrt{\hat{z}}, \tag{4.135}$$

$$E_y = y\,\psi(y). \tag{4.136}$$

In the limit $\epsilon \rightarrow 0$, Eq. (4.133) transforms into

$$\frac{d^2\psi}{dy^2} + \frac{1}{y}\frac{d\psi}{dy} + \left(1 - \frac{1}{y^2}\right)\psi = 0. \tag{4.137}$$

This is a standard equation, known as *Bessel's equation* of order one,[2] and possesses two independent solutions, denoted $J_1(y)$ and $Y_1(y)$, respectively. Thus, on the positive real $\hat{z}$-axis we can write the most general solution to Eq. (4.133) in the form

$$E_y(\hat{z}) = A\,\sqrt{\hat{z}}\,J_1(2\sqrt{\hat{z}}) + B\,\sqrt{\hat{z}}\,Y_1(2\sqrt{\hat{z}}), \tag{4.138}$$

where A and B are two arbitrary constants.

Let

$$y = 2\sqrt{a\hat{z}}, \tag{4.139}$$

$$E_y = y\,\psi(y), \tag{4.140}$$

[2]M. Abramowitz, and I.A. Stegun, *Handbook of Mathematical Functions* (Dover, New York NY, 1964), p. 358.

where

$$a = e^{-i\,\pi\,\mathrm{sgn}(\epsilon)}. \tag{4.141}$$

Note that the argument of $a\hat{z}$ is zero on the negative real $\hat{z}$-axis. In the limit $\epsilon \rightarrow 0$, Eq. (4.133) transforms into

$$\frac{d^2\psi}{dy^2} + \frac{1}{y}\frac{d\psi}{dy} - \left(1 + \frac{1}{y^2}\right)\psi = 0. \tag{4.142}$$

This is a standard equation, known as *Bessel's modified equation* of order one,[3] and possesses two independent solutions, denoted $I_1(y)$ and $K_1(y)$, respectively. Thus, on the negative real $\hat{z}$-axis we can write the most general solution to Eq. (4.133) in the form

$$E_y(\hat{z}) = C\,\sqrt{a\hat{z}}\,I_1(2\,\sqrt{a\hat{z}}) + D\,\sqrt{a\hat{z}}\,K_1(2\,\sqrt{a\hat{z}}), \tag{4.143}$$

where C and D are two arbitrary constants.

Now, the Bessel functions J_1, Y_1, I_1, and K_1 are all perfectly well-defined for complex arguments, so the two expressions (4.138) and (4.143) must, in fact, be *identical*. In particular, the constants C and D must somehow be related to the constants A and B. In order to establish this relationship, it is convenient to investigate the behaviour of the expressions (4.138) and (4.143) in the limit of small $\hat{z}$: *i.e.*, $|\hat{z}| \ll 1$. In this limit,

$$\sqrt{\hat{z}}\,J_1(2\,\sqrt{\hat{z}}) = \hat{z} + O(\hat{z}^2), \tag{4.144}$$

$$\sqrt{a\hat{z}}\,I_1(2\,\sqrt{a\hat{z}}) = -\hat{z} + O(\hat{z}^2), \tag{4.145}$$

$$\sqrt{\hat{z}}\,Y_1(2\,\sqrt{\hat{z}}) = -\frac{1}{\pi}\,[1 - \{\ln|\hat{z}| + 2\,\gamma - 1\}\,\hat{z}] + O(\hat{z}^2), \tag{4.146}$$

$$\sqrt{a\hat{z}}\,K_1(2\,\sqrt{a\hat{z}}) = \frac{1}{2}\,[1 - \{\ln|\hat{z}| + 2\,\gamma - 1\}\,\hat{z} - i\arg(a)\,\hat{z}] + O(\hat{z}^2), \tag{4.147}$$

where γ is Euler's constant, and $\hat{z}$ is assumed to lie on the positive real $\hat{z}$-axis. It follows, by a comparison of Eqs. (4.138), (4.143), and (4.144)–(4.147), that the choice

$$C = -A + i\,\frac{\pi}{2}\,\mathrm{sgn}(\epsilon)\,D = -A - i\,\mathrm{sgn}(\epsilon)\,B, \tag{4.148}$$

$$D = -\frac{2}{\pi}\,B, \tag{4.149}$$

ensures that the expressions (4.138) and (4.143) are indeed identical.

[3] M. Abramowitz, and I.A. Stegun, *Handbook of Mathematical Functions* (Dover, New York NY, 1964), p. 374.

Now, in the limit $|\hat{z}| \gg 1$,

$$\sqrt{a\hat{z}}\, I_1(2\sqrt{a\hat{z}}) \sim \frac{|\hat{z}|^{1/4}}{2\sqrt{\pi}}\, e^{+2\sqrt{|\hat{z}|}}, \tag{4.150}$$

$$\sqrt{a\hat{z}}\, K_1(2\sqrt{a\hat{z}}) \sim \frac{\sqrt{\pi}\,|\hat{z}|^{1/4}}{2}\, e^{-2\sqrt{|\hat{z}|}}, \tag{4.151}$$

where $\hat{z}$ is assumed to lie on the negative real $\hat{z}$-axis. It is clear that the I_1 solution is unphysical, since it blows up in the evanescent region ($\hat{z} < 0$). Thus, the coefficient C in expression (4.143) must be set to zero in order to prevent $E_y(\hat{z})$ from blowing up as $\hat{z} \to -\infty$. According to Eq. (4.148), this constraint implies that

$$A = -i\,\mathrm{sgn}(\epsilon)\, B. \tag{4.152}$$

In the limit $|\hat{z}| \gg 1$,

$$\sqrt{\hat{z}}\, J_1(2\sqrt{\hat{z}}) \sim \frac{\hat{z}^{1/4}}{\sqrt{\pi}} \cos\left(2\sqrt{z} - \frac{3}{4}\pi\right), \tag{4.153}$$

$$\sqrt{\hat{z}}\, Y_1(2\sqrt{\hat{z}}) \sim \frac{\hat{z}^{1/4}}{\sqrt{\pi}} \sin\left(2\sqrt{z} - \frac{3}{4}\pi\right), \tag{4.154}$$

where $\hat{z}$ is assumed to lie on the positive real $\hat{z}$-axis. It follows from Eqs. (4.138), (4.152), and (4.153)–(4.154) that in the non-evanescent region ($\hat{z} > 0$) the most general *physical* solution takes the form

$$\begin{aligned} E_y(\hat{z}) &= A'\,[\mathrm{sgn}(\epsilon)+1]\;\hat{z}^{1/4}\;\exp\left[+i\,2\sqrt{\hat{z}} - \frac{3}{4}\pi\right] \\ &\quad + A'\,[\mathrm{sgn}(\epsilon)-1]\;\hat{z}^{1/4}\;\exp\left[-i\,2\sqrt{\hat{z}} + \frac{3}{4}\pi\right], \end{aligned} \tag{4.155}$$

where A' is an arbitrary constant.

Suppose that a plane electromagnetic wave, polarized in the y-direction, is launched from an antenna, located at large positive z, towards the resonance point at $z = 0$. It is assumed that $n = 1$ at the launch point. In the non-evanescent region, $z > 0$, the wave can be represented as a linear combination of propagating WKB solutions:

$$E_y(z) = E\, n^{-1/2}\, \exp\left(-i\,k_0 \int_0^z n\, dz\right) + F\, n^{-1/2}\, \exp\left(+i\,k_0 \int_0^z n\, dz\right). \tag{4.156}$$

The first term on the right-hand side of the above equation represents the incident wave, whereas the second term represents the reflected wave. Here, E is the amplitude of the incident wave, and F is the amplitude of the reflected wave. In the vicinity of the resonance point (*i.e.*, z small and positive, or $\hat{z}$ large and positive) the above expression reduces to

$$E_y(\hat{z}) \simeq (k_0 b)^{-1/2} \left[E\,\hat{z}^{1/4} \exp\left(-i\,2\sqrt{\hat{z}}\right) + F\,\hat{z}^{1/4} \exp\left(+i\,2\sqrt{\hat{z}}\right)\right]. \tag{4.157}$$

A comparison of Eqs. (4.155) and (4.157) shows that if $\epsilon > 0$ then $E = 0$. In other words, there is a reflected wave, but no incident wave. This corresponds to the *spontaneous excitation* of waves in the vicinity of the resonance. On the other hand, if $\epsilon < 0$ then $F = 0$. In other words, there is an incident wave, but no reflected wave. This corresponds to the *total absorption* of incident waves in the vicinity of the resonance. It is clear that if $\epsilon > 0$ then ϵ represents some sort of spontaneous wave excitation mechanism, whereas if $\epsilon < 0$ then ϵ represents a wave absorption, or damping, mechanism. We would normally expect plasmas to absorb incident wave energy, rather than spontaneously emit waves, so we conclude that, under most circumstances, $\epsilon < 0$, and resonances *absorb* incident waves *without reflection*.

4.14 Resonant Layers

Consider the situation under investigation in the preceding section, in which a plane wave, polarized in the y-direction, is launched along the z-axis, from an antenna located at large positive z, and absorbed at a resonance located at $z = 0$. In the vicinity of the resonant point, the electric component of the wave satisfies

$$\frac{d^2E_y}{dz^2} + \frac{k_0^2\, b}{z + i\,\epsilon} E_y = 0, \tag{4.158}$$

where $b > 0$ and $\epsilon < 0$.

The time-averaged Poynting flux in the z-direction is written

$$P_z = -\frac{(E_y\, B_x^* + E_y^*\, B_x)}{4\mu_0}. \tag{4.159}$$

Now, the Faraday-Maxwell equation yields

$$i\,\omega\, B_x = -\frac{dE_y}{dz}. \tag{4.160}$$

Thus, we have

$$P_z = -\frac{i}{4\,\mu_0\,\omega}\left(\frac{dE_y}{dz}\,E_y^* - \frac{dE_y^*}{dz}\,E_y\right). \tag{4.161}$$

Let us ascribe any variation of P_z with z to the wave energy emitted by the plasma. We then have

$$\frac{dP_z}{dz} = W, \tag{4.162}$$

where W is the power emitted by the plasma per unit volume. It follows that

$$W = -\frac{i}{4\,\mu_0\,\omega}\left(\frac{d^2E_y}{dz^2}\,E_y^* - \frac{d^2E_y^*}{dz^2}\,E_y\right). \tag{4.163}$$

Equations (4.158) and (4.163) yield

$$W = \frac{k_0^{\,2}\, b}{2\,\mu_0\,\omega}\,\frac{\epsilon}{z^2+\epsilon^2}\,|E_y|^2. \tag{4.164}$$

Note that $W < 0$, since $\epsilon < 0$, so wave energy is *absorbed* by the plasma. It is clear from the above formula that the absorption takes place in a narrow layer, of thickness $|\epsilon|$, centred on the resonance point, $z = 0$.

4.15 Collisional Damping

Let us now consider a real-life damping mechanism. Equation (4.15) specifies the linearized Ohm's law in the *collisionless* cold-plasma approximation. However, in the presence of *collisions* this expression acquires an extra term (see Sect. 3), such that

$$\mathbf{E} = -\mathbf{V}\times\mathbf{B}_0 + \frac{\mathbf{j}\times\mathbf{B}_0}{n e} - i\,\omega\,\frac{m_e}{n e^2}\,\mathbf{j} + \nu\,\frac{m_e}{n e^2}\,\mathbf{j}, \tag{4.165}$$

where $\nu \equiv \tau_e^{-1}$ is the *collision frequency*. Here, we have neglected the small difference between the parallel and perpendicular plasma electrical conductivities, for the sake of simplicity. When Eq. (4.165) is used to calculate the dielectric permittivity for a right-handed wave, in the limit $\omega \gg \Omega_i$, we obtain

$$R \simeq 1 - \frac{\Pi_e^{\,2}}{\omega\,(\omega + i\,\nu - |\Omega_e|)}. \tag{4.166}$$

A right-handed circularly polarized wave, propagating parallel to the magnetic field, is governed by the dispersion relation

$$n^2 = R \simeq 1 + \frac{\Pi_e^{\,2}}{\omega\,(|\Omega_e| - \omega - i\,\nu)}. \tag{4.167}$$

Suppose that $n = n(z)$. Furthermore, let

$$|\Omega_e| = \omega + |\Omega_e|'\,z, \tag{4.168}$$

so that the electron cyclotron resonance is located at $z = 0$. We also assume that $|\Omega_e|' > 0$, so that the evanescent region corresponds to $z < 0$. It follows that in the immediate vicinity of the resonance

$$n^2 \simeq \frac{b}{z + i\,\epsilon}, \tag{4.169}$$

where

$$b = \frac{\Pi_e^{\,2}}{\omega\,|\Omega_e|'}, \tag{4.170}$$

and

$$\epsilon = -\frac{\nu}{|\Omega_e|'}. \tag{4.171}$$

It can be seen that $\epsilon < 0$, which is consistent with the *absorption* of incident wave energy by the resonant layer. The approximate width of the resonant layer is

$$\delta \sim |\epsilon| = \frac{\nu}{|\Omega_e|'}. \tag{4.172}$$

Note that the damping mechanism, in this case collisions, controls the *thickness* of the resonant layer, but does not control the amount of wave energy absorbed by the layer. In fact, in the simple theory outlined above, *all* of the incident wave energy is absorbed by the layer.

4.16 Pulse Propagation

Consider the situation under investigation in Sect. 4.12, in which a plane wave, polarized in the y-direction, is launched along the z-axis, from an antenna located at large positive z, and reflected from a cutoff located at $z = 0$. Up to now, we have only considered *infinite* wave-trains, characterized by a discrete frequency, ω. Let us now consider the more realistic case in which the antenna emits a finite *pulse* of radio waves.

The pulse structure is conveniently represented as

$$E_y(t) = \int_{-\infty}^{\infty} F(\omega)\, e^{-i\,\omega\, t}\, d\omega, \tag{4.173}$$

where $E_y(t)$ is the electric field produced by the antenna, which is assumed to lie at $z = a$. Suppose that the pulse is a signal of roughly constant (angular) frequency ω_0, which lasts a time T, where T is long compared to $1/\omega_0$. It follows that $F(\omega)$ possesses narrow maxima around $\omega = \pm\omega_0$. In other words, only those frequencies which lie very close to the central frequency ω_0 play a significant role in the propagation of the pulse.

Each component frequency of the pulse yields a wave which propagates *independently* along the z-axis, in a manner specified by the appropriate WKB solution [see Eqs. (4.120)–(4.121)]. Thus, if Eq. (4.173) specifies the signal at the antenna (*i.e.*, at $z = a$), then the signal at coordinate z (where $z < a$) is given by

$$E_y(z,t) = \int_{-\infty}^{\infty} \frac{F(\omega)}{n^{1/2}(\omega,z)}\; e^{\,i\,\phi(\omega,z,t)}\, d\omega, \tag{4.174}$$

where

$$\phi(\omega,z,t) = \frac{\omega}{c}\int_z^a n(\omega,z)\, dz - \omega\, t. \tag{4.175}$$

Here, we have used $k_0 = \omega/c$.

Equation (4.174) can be regarded as a contour integral in ω-space. The quantity $F/n^{1/2}$ is a relatively slowly varying function of ω, whereas the phase, ϕ, is a large and rapidly

varying function of ω. The rapid oscillations of $\exp(\mathrm{i}\,\phi)$ over most of the path of integration ensure that the integrand averages almost to zero. However, this cancellation argument does not apply to places on the integration path where the phase is *stationary*: *i.e.*, places where $\phi(\omega)$ has an extremum. The integral can, therefore, be estimated by finding those points where $\phi(\omega)$ has a vanishing derivative, evaluating (approximately) the integral in the neighbourhood of each of these points, and summing the contributions. This procedure is called the *method of stationary phase*.

Suppose that $\phi(\omega)$ has a vanishing first derivative at $\omega = \omega_s$. In the neighbourhood of this point, $\phi(\omega)$ can be expanded as a Taylor series,

$$\phi(\omega) = \phi_s + \frac{1}{2}\phi_s''(\omega - \omega_s)^2 + \cdots. \tag{4.176}$$

Here, the subscript s is used to indicate ϕ or its second derivative evaluated at $\omega = \omega_s$. Since $F(\omega)/n^{1/2}(\omega, z)$ is slowly varying, the contribution to the integral from this stationary phase point is approximately

$$E_{y\,s} \simeq \frac{F(\omega_s)\,\mathrm{e}^{\,\mathrm{i}\,\phi_s}}{n^{1/2}(\omega_s, z)} \int_{-\infty}^{\infty} \mathrm{e}^{\,(\mathrm{i}/2)\phi_s''(\omega-\omega_s)^2}\,\mathrm{d}\omega. \tag{4.177}$$

The above expression can be written in the form

$$E_{y\,s} \simeq \frac{F(\omega_s)\,\mathrm{e}^{\,\mathrm{i}\,\phi_s}}{n^{1/2}(\omega_s, z)} \sqrt{\frac{4\pi}{\phi_s''}} \int_0^{\infty} \left[\cos(\pi\,t^2/2) + \mathrm{i}\,\sin(\pi\,t^2/2)\right]\,\mathrm{d}t, \tag{4.178}$$

where

$$\frac{\pi}{2}\,t^2 = \frac{1}{2}\,\phi_s''\,(\omega - \omega_s)^2. \tag{4.179}$$

The integrals in the above expression are *Fresnel integrals*,[4] and can be shown to take the values

$$\int_0^{\infty} \cos(\pi\,t^2/2)\,\mathrm{d}t = \int_0^{\infty} \sin(\pi\,t^2/2)\,\mathrm{d}t = \frac{1}{2}. \tag{4.180}$$

It follows that

$$E_{y\,s} \simeq \sqrt{\frac{2\pi\,\mathrm{i}}{\phi_s''}}\,\frac{F(\omega_s)}{n^{1/2}(\omega_s, z)}\,\mathrm{e}^{\,\mathrm{i}\,\phi_s}. \tag{4.181}$$

If there is more than one point of stationary phase in the range of integration then the integral is approximated as a sum of terms like the above.

Integrals of the form (4.174) can be calculated exactly using the *method of steepest decent*.[5] The stationary phase approximation (4.181) agrees with the leading term of the method of steepest decent (which is far more difficult to implement than the method of stationary phase) provided that $\phi(\omega)$ is real (*i.e.*, provided that the stationary point lies on

[4]M. Abramowitz, and I.A. Stegun, *Handbook of Mathematical Functions*, (Dover, New York NY, 1965), Sect. 7.3.

[5]Léon Brillouin, *Wave Propagation and Group Velocity*, (Academic press, New York NY, 1960).

the real axis). If ϕ is complex, however, the stationary phase method can yield erroneous results.

It follows, from the above discussion, that the right-hand side of Eq. (4.174) averages to a very small value, expect for those special values of z and t at which one of the points of stationary phase in ω-space coincides with one of the peaks of $F(\omega)$. The locus of these special values of z and t can obviously be regarded as the equation of motion of the pulse as it propagates along the z-axis. Thus, the equation of motion is specified by

$$\left(\frac{\partial\phi}{\partial\omega}\right)_{\omega=\omega_0} = 0, \tag{4.182}$$

which yields

$$t = \frac{1}{c}\int_z^a \left[\frac{\partial(\omega\, n)}{\partial\omega}\right]_{\omega=\omega_0} dz. \tag{4.183}$$

Suppose that the z-velocity of a pulse of central frequency ω_0 at coordinate z is given by $-u_z(\omega_0, z)$. The differential equation of motion of the pulse is then $dt = -dz/u_z$. This can be integrated, using the boundary condition $z = a$ at $t = 0$, to give the full equation of motion:

$$t = \int_z^a \frac{dz}{u_z}. \tag{4.184}$$

A comparison of Eqs. (4.183) and (4.184) yields

$$u_z(\omega_0, z) = c \Bigg/ \left\{\frac{\partial[\omega\, n(\omega, z)]}{\partial\omega}\right\}_{\omega=\omega_0}. \tag{4.185}$$

The velocity u_z is usually called the *group velocity*. It is easily demonstrated that the above expression for the group velocity is entirely consistent with that given previously [see Eq. (4.71)].

The dispersion relation for an electromagnetic plasma wave propagating through an unmagnetized plasma is

$$n(\omega, z) = \left(1 - \frac{\Pi_e{}^2(z)}{\omega^2}\right)^{1/2}. \tag{4.186}$$

Here, we have assumed that equilibrium quantities are functions of z only, and that the wave propagates along the z-axis. The phase velocity of waves of frequency ω propagating along the z-axis is given by

$$v_z(\omega, z) = \frac{c}{n(\omega, z)} = c\,\left(1 - \frac{\Pi_e{}^2(z)}{\omega^2}\right)^{-1/2}. \tag{4.187}$$

According to Eqs. (4.185) and (4.186), the corresponding group velocity is

$$u_z(\omega, z) = c\,\left(1 - \frac{\Pi_e{}^2(z)}{\omega^2}\right)^{1/2}. \tag{4.188}$$

It follows that

$$v_z\, u_z = c^2. \tag{4.189}$$

It is assumed that $\Pi_e(0) = \omega$, and $\Pi_e(z) < \omega$ for $z > 0$, which implies that the reflection point corresponds to $z = 0$. Note that the phase velocity is always greater than the velocity of light in vacuum, whereas the group velocity is always less than this velocity. Note, also, that as the reflection point, $z = 0$, is approached from positive z, the phase velocity tends to infinity, whereas the group velocity tends to zero.

Although we have only analyzed the motion of the pulse as it travels from the antenna to the reflection point, it is easily demonstrated that the speed of the reflected pulse at position z is the same as that of the incident pulse. In other words, the group velocities of pulses traveling in opposite directions are of equal magnitude.

4.17 Ray Tracing

Let us now generalize the preceding analysis so that we can deal with pulse propagation though a three-dimensional magnetized plasma.

A general wave problem can be written as a set of n coupled, linear, homogeneous, first-order, partial-differential equations, which take the form

$$\mathbf{M}(\,i\,\partial/\partial t, -i\,\nabla, \mathbf{r}, t)\,\boldsymbol{\psi} = \mathbf{0}. \tag{4.190}$$

The vector-field $\boldsymbol{\psi}(\mathbf{r}, t)$ has n components (e.g., $\boldsymbol{\psi}$ might consist of $\mathbf{E}$, $\mathbf{B}$, $\mathbf{j}$, and $\mathbf{V}$) characterizing some small disturbance, and $\mathbf{M}$ is an $n \times n$ matrix characterizing the undisturbed plasma.

The lowest order WKB approximation is premised on the assumption that $\mathbf{M}$ depends so weakly on $\mathbf{r}$ and t that all of the spatial and temporal dependence of the components of $\psi(\mathbf{r}, t)$ is specified by a common factor $\exp(\,i\,\phi)$. Thus, Eq. (4.190) reduces to

$$\mathbf{M}(\omega, \mathbf{k}, \mathbf{r}, t)\,\boldsymbol{\psi} = \mathbf{0}, \tag{4.191}$$

where

$$\mathbf{k} \equiv \nabla\phi, \tag{4.192}$$

$$\omega \equiv -\frac{\partial\phi}{\partial t}. \tag{4.193}$$

In general, Eq. (4.191) has many solutions, corresponding to the many different types and polarizations of wave which can propagate through the plasma in question, all of which satisfy the dispersion relation

$$\mathcal{M}(\omega, \mathbf{k}, \mathbf{r}, t) = 0, \tag{4.194}$$

where $\mathcal{M} \equiv \det(\mathbf{M})$. As is easily demonstrated (see Sect. 4.11), the WKB approximation is valid provided that the characteristic variation length-scale and variation time-scale of the

plasma are much longer than the wave-length, $2\pi/k$, and the period, $2\pi/\omega$, respectively, of the wave in question.

Let us concentrate on one particular solution of Eq. (4.191) (*e.g.*, on one particular type of plasma wave). For this solution, the dispersion relation (4.194) yields

$$\omega = \Omega(\mathbf{k}, \mathbf{r}, t) : \tag{4.195}$$

i.e., the dispersion relation yields a unique frequency for a wave of a given wave-vector, $\mathbf{k}$, located at a given point, $(\mathbf{r}, t)$, in space and time. There is also a unique ψ associated with this frequency, which is obtained from Eq. (4.191). To lowest order, we can neglect the variation of ψ with $\mathbf{r}$ and t. A general pulse solution is written

$$\psi(\mathbf{r}, t) = \int F(\mathbf{k})\, \psi\, e^{i\phi}\, d^3\mathbf{k}, \tag{4.196}$$

where (locally)

$$\phi = \mathbf{k}\cdot\mathbf{r} - \Omega\, t, \tag{4.197}$$

and F is a function which specifies the initial structure of the pulse in $\mathbf{k}$-space.

The integral (4.196) averages to zero, except at a point of *stationary phase*, where $\nabla_{\mathbf{k}}\phi = 0$ (see Sect. 4.16). Here, $\nabla_{\mathbf{k}}$ is the $\mathbf{k}$-space gradient operator. It follows that the (instantaneous) trajectory of the pulse matches that of a point of stationary phase: *i.e.*,

$$\nabla_{\mathbf{k}}\phi = \mathbf{r} - \mathbf{v}_g\, t = 0, \tag{4.198}$$

where

$$\mathbf{v}_g = \frac{\partial\Omega}{\partial\mathbf{k}} \tag{4.199}$$

is the *group velocity*. Thus, the instantaneous velocity of a pulse is always equal to the local group velocity.

Let us now determine how the wave-vector, $\mathbf{k}$, and frequency, ω, of a pulse evolve as the pulse propagates through the plasma. We start from the cross-differentiation rules [see Eqs. (4.192)–(4.193)]:

$$\frac{\partial k_i}{\partial t} + \frac{\partial\omega}{\partial r_i} = 0, \tag{4.200}$$

$$\frac{\partial k_j}{\partial r_i} - \frac{\partial k_i}{\partial r_j} = 0. \tag{4.201}$$

Equations (4.195) and (4.200)–(4.201) yield (making use of the Einstein summation convention)

$$\frac{\partial k_i}{\partial t} + \frac{\partial\Omega}{\partial k_j}\frac{\partial k_j}{\partial r_i} + \frac{\partial\Omega}{\partial r_i} = \frac{\partial k_i}{\partial t} + \frac{\partial\Omega}{\partial k_j}\frac{\partial k_i}{\partial r_j} + \frac{\partial\Omega}{\partial r_i} = 0, \tag{4.202}$$

or

$$\frac{d\mathbf{k}}{dt} \equiv \frac{\partial\mathbf{k}}{\partial t} + (\mathbf{v}_g\cdot\nabla)\,\mathbf{k} = -\nabla\Omega. \tag{4.203}$$

In other words, the variation of $\mathbf{k}$, as seen in a frame co-moving with the pulse, is determined by the spatial gradients in Ω.

Partial differentiation of Eq. (4.195) with respect to t gives

$$\frac{\partial \omega}{\partial t} = \frac{\partial \Omega}{\partial k_j}\frac{\partial k_j}{\partial t} + \frac{\partial \Omega}{\partial t} = -\frac{\partial \Omega}{\partial k_j}\frac{\partial \omega}{\partial r_j} + \frac{\partial \Omega}{\partial t}, \tag{4.204}$$

which can be written

$$\frac{d\omega}{dt} \equiv \frac{\partial \omega}{\partial t} + (\mathbf{v}_g\cdot\nabla)\,\omega = \frac{\partial \Omega}{\partial t}. \tag{4.205}$$

In other words, the variation of ω, as seen in a frame co-moving with the pulse, is determined by the time variation of Ω.

According to the above analysis, the evolution of a pulse propagating though a spatially and temporally non-uniform plasma can be determined by solving the *ray equations*:

$$\frac{d\mathbf{r}}{dt} = \frac{\partial \Omega}{\partial \mathbf{k}}, \tag{4.206}$$

$$\frac{d\mathbf{k}}{dt} = -\nabla\Omega, \tag{4.207}$$

$$\frac{d\omega}{dt} = \frac{\partial \Omega}{\partial t}. \tag{4.208}$$

The above equations are conveniently rewritten in terms of the dispersion relation (4.194):

$$\frac{d\mathbf{r}}{dt} = -\frac{\partial \mathcal{M}/\partial \mathbf{k}}{\partial \mathcal{M}/\partial \omega}, \tag{4.209}$$

$$\frac{d\mathbf{k}}{dt} = \frac{\partial \mathcal{M}/\partial \mathbf{r}}{\partial \mathcal{M}/\partial \omega}, \tag{4.210}$$

$$\frac{d\omega}{dt} = -\frac{\partial \mathcal{M}/\partial t}{\partial \mathcal{M}/\partial \omega}. \tag{4.211}$$

Note, finally, that the variation in the amplitude of the pulse, as it propagates through though the plasma, can only be determined by expanding the WKB solutions to higher order (see Sect. 4.11).

4.18 Radio Wave Propagation Through the Ionosphere

To a first approximation, the Earth's ionosphere consists of an unmagnetized, horizontally stratified, partially ionized gas. The dispersion relation for the electromagnetic plasma wave takes the form [see Eq. (4.97)]

$$\mathcal{M} = \omega^2 - k^2c^2 - \Pi_e^{\,2} = 0, \tag{4.212}$$

where

$$\Pi_e = \sqrt{\frac{N\, e^2}{\epsilon_0\, m_e}}. \tag{4.213}$$

Here, $N = N(z)$ is the density of free electrons in the ionosphere, and z is a coordinate which measures height above the surface of the Earth. (*N.B.*, The curvature of the Earth is neglected in the following analysis.)

Now,

$$\frac{\partial \mathcal{M}}{\partial \omega} = 2\,\omega, \tag{4.214}$$

$$\frac{\partial \mathcal{M}}{\partial \mathbf{k}} = -2\,\mathbf{k}\,c^2, \tag{4.215}$$

$$\frac{\partial \mathcal{M}}{\partial \mathbf{r}} = -\nabla \Pi_e^{\,2}, \tag{4.216}$$

$$\frac{\partial \mathcal{M}}{\partial t} = 0. \tag{4.217}$$

Thus, the ray equations, (4.209)–(4.211), yield

$$\frac{d\mathbf{r}}{dt} = \frac{\mathbf{k}\,c^2}{\omega}, \tag{4.218}$$

$$\frac{d\mathbf{k}}{dt} = -\frac{\nabla \Pi_e^{\,2}}{2\,\omega}, \tag{4.219}$$

$$\frac{d\omega}{dt} = 0. \tag{4.220}$$

Note that the frequency of a radio pulse does not change as it propagates through the ionosphere, provided that $N(z)$ does not vary in time. It is clear, from Eqs. (4.218)–(4.220), and the fact that $\Pi_e = \Pi_e(z)$, that a radio pulse which starts off at ground level propagating in the x-z plane, say, will continue to propagate in this plane.

For pulse propagation in the x-z plane, we have

$$\frac{dx}{dt} = \frac{k_x\,c^2}{\omega}, \tag{4.221}$$

$$\frac{dz}{dt} = \frac{k_z\,c^2}{\omega}, \tag{4.222}$$

$$\frac{dk_x}{dt} = 0. \tag{4.223}$$

The dispersion relation (4.212) yields

$$n^2 = \frac{(k_x^{\,2} + k_z^{\,2})\,c^2}{\omega^2} = 1 - \frac{\Pi_e^{\,2}}{\omega^2}, \tag{4.224}$$

where $n(z)$ is the refractive index.

We assume that $n = 1$ at $z = 0$, which is equivalent to the reasonable assumption that the atmosphere is non-ionized at ground level. It follows from Eq. (4.223) that

$$k_x = k_x(z = 0) = \frac{\omega}{c}\, S, \tag{4.225}$$

where S is the sine of the angle of incidence of the pulse, with respect to the vertical axis, at ground level. Equations (4.224) and (4.225) yield

$$k_z = \pm\frac{\omega}{c}\,\sqrt{n^2 - S^2}. \tag{4.226}$$

According to Eq. (4.222), the plus sign corresponds to the upward trajectory of the pulse, whereas the minus sign corresponds to the downward trajectory. Finally, Eqs. (4.221), (4.222), (4.225), and (4.226) yield the equations of motion of the pulse:

$$\frac{dx}{dt} = c\, S, \tag{4.227}$$

$$\frac{dz}{dt} = \pm c\,\sqrt{n^2 - S^2}. \tag{4.228}$$

The pulse attains its maximum altitude, $z = z_0$, when

$$n(z_0) = |S|. \tag{4.229}$$

The total distance traveled by the pulse (*i.e.*, the distance from its launch point to the point where it intersects the Earth's surface again) is

$$x_0 = 2\, S \int_0^{z_0(S)} \frac{dz}{\sqrt{n^2(z) - S^2}}. \tag{4.230}$$

In the limit in which the radio pulse is launched vertically (*i.e.*, $S = 0$) into the ionosphere, the turning point condition (4.229) reduces to that characteristic of a cutoff (*i.e.*, $n = 0$). The WKB turning point described in Eq. (4.229) is a generalization of the conventional turning point, which occurs when k^2 changes sign. Here, k_z^2 changes sign, whilst k_x^2 and k_y^2 are constrained by symmetry (*i.e.*, k_x is constant, and k_y is zero).

According to Eqs. (4.218)–(4.220) and (4.224), the equation of motion of the pulse can also be written

$$\frac{d^2\mathbf{r}}{dt^2} = \frac{c^2}{2}\,\nabla n^2. \tag{4.231}$$

It follows that the trajectory of the pulse is the same as that of a particle moving in the gravitational potential $-c^2\, n^2/2$. Thus, if n^2 decreases *linearly* with increasing height above the ground [which is the case if $N(z)$ increases linearly with z] then the trajectory of the pulse is a *parabola*.

5 Magnetohydrodynamic Fluids

5.1 Introduction

As we have seen in Sect. 3, the MHD equations take the form

$$\frac{d\rho}{dt} + \rho\,\nabla\cdot\mathbf{V} = 0, \tag{5.1}$$

$$\rho\,\frac{d\mathbf{V}}{dt} + \nabla p - \mathbf{j}\times\mathbf{B} = \mathbf{0}, \tag{5.2}$$

$$\mathbf{E} + \mathbf{V}\times\mathbf{B} = \mathbf{0}, \tag{5.3}$$

$$\frac{d}{dt}\left(\frac{p}{\rho^{\Gamma}}\right) = 0, \tag{5.4}$$

where $\rho \simeq m_i\, n$ is the plasma mass density, and $\Gamma = 5/3$ is the ratio of specific heats.

It is often observed that the above set of equations are identical to the equations governing the motion of an inviscid, adiabatic, perfectly conducting, electrically neutral liquid. Indeed, this observation is sometimes used as the sole justification for the MHD equations. After all, a hot, tenuous, quasi-neutral plasma is highly conducting, and if the motion is sufficiently fast then both viscosity and heat conduction can be plausibly neglected. However, we can appreciate, from Sect. 3, that this is a highly oversimplified and misleading argument. The problem is, of course, that a weakly coupled plasma is a far more complicated dynamical system than a conducting liquid.

According to the discussion in Sect. 3, the MHD equations are only valid when

$$\delta^{-1}\, v_t \gg V \gg \delta\, v_t. \tag{5.5}$$

Here, V is the typical velocity associated with the plasma dynamics under investigation, v_t is the typical thermal velocity, and δ is the typical magnetization parameter (*i.e.*, the typical ratio of a particle gyro-radius to the scale-length of the motion). Clearly, the above inequality is most likely to be satisfied in a *highly magnetized* (*i.e.*, $\delta \rightarrow 0$) plasma.

If the plasma dynamics becomes too fast (*i.e.*, $V \sim \delta^{-1}\, v_t$) then resonances occur with the motions of individual particles (*e.g.*, the cyclotron resonances) which invalidate the MHD equations. Furthermore, effects, such as electron inertia and the Hall effect, which are not taken into account in the MHD equations, become important.

MHD is essentially a *single-fluid* plasma theory. A single-fluid approach is justified because the perpendicular motion is dominated by $\mathbf{E}\times\mathbf{B}$ drifts, which are the same for both plasma species. Furthermore, the relative streaming velocity, $U_\parallel$, of both species parallel to the magnetic field is strongly constrained by the fundamental MHD ordering (see Sect. 3.9)

$$U \sim \delta\, V. \tag{5.6}$$

Note, however, that if the plasma dynamics becomes too slow (*i.e.*, $V \sim \delta v_t$) then the motions of the electron and ion fluids become sufficiently different that a single-fluid approach is no longer tenable. This occurs whenever the diamagnetic velocities, which are quite different for different plasma species, become comparable to the $\mathbf{E}\times\mathbf{B}$ velocity (see Sect. 3.12). Furthermore, effects such as plasma resistivity, viscosity, and thermal conductivity, which are not taken into account in the MHD equations, become important in this limit.

Broadly speaking, the MHD equations describe relatively *violent*, *large-scale* motions of highly *magnetized* plasmas.

Strictly speaking, the MHD equations are only valid in *collisional* plasmas (*i.e.*, plasmas in which the mean-free-path is much smaller than the typical variation scale-length). However, as discussed in Sect. 3.13, the MHD equations also fairly well describe the *perpendicular* (but not the *parallel*!) motions of collisionless plasmas.

Assuming that the MHD equations are valid, let us now investigate their properties.

5.2 Magnetic Pressure

The MHD equations can be combined with Maxwell's equations,

$$\nabla\times\mathbf{B} = \mu_0\,\mathbf{j}, \tag{5.7}$$

$$\nabla\times\mathbf{E} = -\frac{\partial\mathbf{B}}{\partial t}, \tag{5.8}$$

to form a closed set. The displacement current is neglected in Eq. (5.7) on the reasonable assumption that MHD motions are slow compared to the velocity of light. Note that Eq. (5.8) guarantees that $\nabla\cdot\mathbf{B} = 0$, provided that this relation is presumed to hold initially. Similarly, the assumption of quasi-neutrality renders the Poisson-Maxwell equation, $\nabla\cdot\mathbf{E} = \rho_c/\epsilon_0$, irrelevant.

Equations (5.2) and (5.7) can be combined to give the MHD equation of motion:

$$\rho\,\frac{d\mathbf{V}}{dt} = -\nabla p + \nabla\cdot\mathbf{T}, \tag{5.9}$$

where

$$T_{ij} = \frac{B_i\,B_j - \delta_{ij}\,B^2/2}{\mu_0}. \tag{5.10}$$

Suppose that the magnetic field is approximately uniform, and directed along the z-axis. In this case, the above equation of motion reduces to

$$\rho\,\frac{d\mathbf{V}}{dt} = -\nabla\cdot\mathbf{P}, \tag{5.11}$$

where

$$\mathbf{P} = \begin{pmatrix} p + B^2/2\mu_0 & 0 & \\ 0 & p + B^2/2\mu_0 & 0 \\ 0 & 0 & p - B^2/2\mu_0 \end{pmatrix}. \tag{5.12}$$

Note that the magnetic field *increases* the plasma pressure, by an amount $B^2/2\mu_0$, in directions *perpendicular* to the magnetic field, and *decreases* the plasma pressure, by the same amount, in the *parallel* direction. Thus, the magnetic field gives rise to a *magnetic pressure*, $B^2/2\,\mu_0$, acting *perpendicular* to field-lines, and a *magnetic tension*, $B^2/2\,\mu_0$, acting *along* field-lines. Since, as we shall see presently, the plasma is tied to magnetic field-lines, it follows that magnetic field-lines embedded in an MHD plasma act rather like *mutually repulsive elastic bands*.

5.3 Flux Freezing

The MHD Ohm's law,

$$\mathbf{E} + \mathbf{V} \times \mathbf{B} = \mathbf{0}, \tag{5.13}$$

is sometimes referred to as the *perfect conductivity* equation, for obvious reasons, and sometimes as the *flux freezing* equation. The latter nomenclature comes about because Eq. (5.13) implies that the magnetic flux through any closed contour in the plasma, each element of which moves with the local plasma velocity, is a *conserved quantity*.

In order to verify the above assertion, let us consider the magnetic flux, Ψ, through a contour, C, which is co-moving with the plasma:

$$\Psi = \int_S \mathbf{B}\cdot d\mathbf{S}. \tag{5.14}$$

Here, $\mathbf{S}$ is some surface which spans C. The time rate of change of Ψ is made up of two parts. Firstly, there is the part due to the time variation of $\mathbf{B}$ over the surface $\mathbf{S}$. This can be written

$$\left(\frac{\partial \Psi}{\partial t}\right)_1 = \int_S \frac{\partial \mathbf{B}}{\partial t}\cdot d\mathbf{S}. \tag{5.15}$$

Using the Faraday-Maxwell equation, this reduces to

$$\left(\frac{\partial \Psi}{\partial t}\right)_1 = -\int_S \nabla \times \mathbf{E}\cdot d\mathbf{S}. \tag{5.16}$$

Secondly, there is the part due to the motion of C. If $d\mathbf{l}$ is an element of C then $\mathbf{V} \times d\mathbf{l}$ is the area swept out by $d\mathbf{l}$ per unit time. Hence, the flux crossing this area is $\mathbf{B}\cdot\mathbf{V} \times d\mathbf{l}$. It follows that

$$\left(\frac{\partial \Psi}{\partial t}\right)_2 = \int_C \mathbf{B}\cdot\mathbf{V} \times d\mathbf{l} = \int_C \mathbf{B} \times \mathbf{V}\cdot d\mathbf{l}. \tag{5.17}$$

Using Stokes's theorem, we obtain

$$\left(\frac{\partial \Psi}{\partial t}\right)_2 = \int_S \nabla \times (\mathbf{B} \times \mathbf{V})\cdot d\mathbf{S}. \tag{5.18}$$

Hence, the total time rate of change of Ψ is given by

$$\frac{d\Psi}{dt} = -\int_S \nabla \times (\mathbf{E} + \mathbf{V} \times \mathbf{B}) \cdot d\mathbf{S}. \tag{5.19}$$

The condition

$$\mathbf{E} + \mathbf{V} \times \mathbf{B} = \mathbf{0} \tag{5.20}$$

clearly implies that Ψ remains constant in time for any arbitrary contour. This, in turn, implies that magnetic field-lines must move with the plasma. In other words, the field-lines are *frozen* into the plasma.

A *flux-tube* is defined as a topologically cylindrical volume whose sides are defined by magnetic field-lines. Suppose that, at some initial time, a flux-tube is embedded in the plasma. According to the flux-freezing constraint,

$$\frac{d\Psi}{dt} = 0, \tag{5.21}$$

the subsequent motion of the plasma and the magnetic field is always such as to maintain the integrity of the flux-tube. Since magnetic field-lines can be regarded as infinitely thin flux-tubes, we conclude that MHD plasma motion also maintains the integrity of field-lines. In other words, magnetic field-lines embedded in an MHD plasma can never break and reconnect: *i.e.*, MHD forbids any change in *topology* of the field-lines. It turns out that this is an extremely restrictive constraint. Later on, we shall discuss situations in which this constraint is relaxed.

5.4 MHD Waves

Let us investigate the small amplitude waves which propagate through a spatially uniform MHD plasma. We start by combining Eqs. (5.1)–(5.4) and (5.7)–(5.8) to form a closed set of equations:

$$\frac{d\rho}{dt} + \rho \nabla \cdot \mathbf{V} = 0, \tag{5.22}$$

$$\rho \frac{d\mathbf{V}}{dt} + \nabla p - \frac{(\nabla \times \mathbf{B}) \times \mathbf{B}}{\mu_0} = \mathbf{0}, \tag{5.23}$$

$$-\frac{\partial \mathbf{B}}{\partial t} + \nabla \times (\mathbf{V} \times \mathbf{B}) = \mathbf{0}, \tag{5.24}$$

$$\frac{d}{dt}\left(\frac{p}{\rho^\Gamma}\right) = 0. \tag{5.25}$$

Next, we linearize these equations (assuming, for the sake of simplicity, that the equilibrium flow velocity and equilibrium plasma current are both zero) to give

$$\frac{\partial \rho}{\partial t} + \rho_0 \nabla \cdot \mathbf{V} = 0, \tag{5.26}$$

$$\rho_0 \frac{\partial \mathbf{V}}{\partial t} + \nabla p - \frac{(\nabla \times \mathbf{B}) \times \mathbf{B}_0}{\mu_0} = \mathbf{0}, \tag{5.27}$$

$$-\frac{\partial \mathbf{B}}{\partial t} + \nabla \times (\mathbf{V} \times \mathbf{B}_0) = \mathbf{0}, \tag{5.28}$$

$$\frac{\partial}{\partial t}\left(\frac{p}{p_0} - \frac{\Gamma \rho}{\rho_0}\right) = 0. \tag{5.29}$$

Here, the subscript 0 denotes an equilibrium quantity. Perturbed quantities are written without subscripts. Of course, ρ_0, p_0, and $\mathbf{B}_0$ are constants in a spatially uniform plasma.

Let us search for wave-like solutions of Eqs. (5.26)–(5.29) in which perturbed quantities vary like $\exp[\,i\,(\mathbf{k}\cdot\mathbf{r} - \omega t)]$. It follows that

$$-\omega\,\rho + \rho_0\,\mathbf{k}\cdot\mathbf{V} = 0, \tag{5.30}$$

$$-\omega\,\rho_0\,\mathbf{V} + \mathbf{k}\,p - \frac{(\mathbf{k}\times\mathbf{B})\times\mathbf{B}_0}{\mu_0} = \mathbf{0}, \tag{5.31}$$

$$\omega\,\mathbf{B} + \mathbf{k}\times(\mathbf{V}\times\mathbf{B}_0) = \mathbf{0}, \tag{5.32}$$

$$-\omega\left(\frac{p}{p_0} - \frac{\Gamma\rho}{\rho_0}\right) = 0. \tag{5.33}$$

Assuming that $\omega \neq 0$, the above equations yield

$$\rho = \rho_0\,\frac{\mathbf{k}\cdot\mathbf{V}}{\omega}, \tag{5.34}$$

$$p = \Gamma\,p_0\,\frac{\mathbf{k}\cdot\mathbf{V}}{\omega}, \tag{5.35}$$

$$\mathbf{B} = \frac{(\mathbf{k}\cdot\mathbf{V})\,\mathbf{B}_0 - (\mathbf{k}\cdot\mathbf{B}_0)\,\mathbf{V}}{\omega}. \tag{5.36}$$

Substitution of these expressions into the linearized equation of motion, Eq. (5.31), gives

$$\left[\omega^2 - \frac{(\mathbf{k}\cdot\mathbf{B}_0)^2}{\mu_0\,\rho_0}\right]\mathbf{V} = \left\{\left[\frac{\Gamma\,p_0}{\rho_0} + \frac{{B_0}^2}{\mu_0\,\rho_0}\right]\mathbf{k} - \frac{(\mathbf{k}\cdot\mathbf{B}_0)}{\mu_0\,\rho_0}\,\mathbf{B}_0\right\}(\mathbf{k}\cdot\mathbf{V}) - \frac{(\mathbf{k}\cdot\mathbf{B}_0)\,(\mathbf{V}\cdot\mathbf{B}_0)}{\mu_0\,\rho_0}\,\mathbf{k}. \tag{5.37}$$

We can assume, without loss of generality, that the equilibrium magnetic field $\mathbf{B}_0$ is directed along the z-axis, and that the wave-vector $\mathbf{k}$ lies in the x-z plane. Let θ be the angle subtended between $\mathbf{B}_0$ and $\mathbf{k}$. Equation (5.37) reduces to the eigenvalue equation

$$\begin{pmatrix} \omega^2 - k^2\,{V_A}^2 - k^2\,{V_S}^2\,\sin^2\theta & 0 & -k^2\,{V_S}^2\,\sin\theta\cos\theta \\ 0 & \omega^2 - k^2\,{V_A}^2\,\cos^2\theta & 0 \\ -k^2\,{V_S}^2\,\sin\theta\,\cos\theta & 0 & \omega^2 - k^2\,{V_S}^2\,\cos^2\theta \end{pmatrix} \begin{pmatrix} V_x \\ V_y \\ V_z \end{pmatrix} = \mathbf{0}. \tag{5.38}$$

Here,

$$V_A = \sqrt{\frac{{B_0}^2}{\mu_0\,\rho_0}} \tag{5.39}$$

is the *Alfvén speed*, and

$$V_S = \sqrt{\frac{\Gamma\,p_0}{\rho_0}} \tag{5.40}$$

is the *sound speed*. The solubility condition for Eq. (5.38) is that the determinant of the square matrix is zero. This yields the dispersion relation

$$(\omega^2 - k^2\,{V_A}^2\,\cos^2\theta)\left[\omega^4 - \omega^2\,k^2\,({V_A}^2 + {V_S}^2) + k^4\,{V_A}^2\,{V_S}^2\,\cos^2\theta\right] = 0. \tag{5.41}$$

There are *three* independent roots of the above dispersion relation, corresponding to the three different types of wave that can propagate through an MHD plasma. The first, and most obvious, root is

$$\omega = k\,V_A\,\cos\theta, \tag{5.42}$$

which has the associated eigenvector $(0, V_y, 0)$. This root is characterized by both $\mathbf{k}\cdot\mathbf{V} = 0$ and $\mathbf{V}\cdot\mathbf{B}_0 = 0$. It immediately follows from Eqs. (5.34) and (5.35) that there is *zero* perturbation of the plasma density or pressure associated with this root. In fact, this root can easily be identified as the *shear-Alfvén wave*, which was introduced in Sect. 4.8. Note that the properties of the shear-Alfvén wave in a warm (*i.e.*, non-zero pressure) plasma are unchanged from those we found earlier in a cold plasma. Note, finally, that since the shear-Alfvén wave only involves plasma motion *perpendicular* to the magnetic field, we can expect the dispersion relation (5.42) to hold good in a *collisionless*, as well as a collisional, plasma.

The remaining two roots of the dispersion relation (5.41) are written

$$\omega = k\,V_+, \tag{5.43}$$

and

$$\omega = k\,V_-, \tag{5.44}$$

respectively. Here,

$$V_\pm = \left\{\frac{1}{2}\left[{V_A}^2 + {V_S}^2 \pm \sqrt{({V_A}^2 + {V_S}^2)^2 - 4\,{V_A}^2\,{V_S}^2\,\cos^2\theta}\right]\right\}^{1/2}. \tag{5.45}$$

Note that $V_+ \geq V_-$. The first root is generally termed the *fast magnetosonic wave*, or fast wave, for short, whereas the second root is usually called the *slow magnetosonic wave*, or slow wave. The eigenvectors for these waves are $(V_x, 0, V_z)$. It follows that $\mathbf{k}\cdot\mathbf{V} \neq 0$ and $\mathbf{V}\cdot\mathbf{B}_0 \neq 0$. Hence, these waves are associated with non-zero perturbations in the plasma density and pressure, and also involve plasma motion parallel, as well as perpendicular, to the magnetic field. The latter observation suggests that the dispersion relations (5.43) and (5.44) are likely to undergo significant modification in collisionless plasmas.

In order to better understand the nature of the fast and slow waves, let us consider the *cold-plasma* limit, which is obtained by letting the sound speed V_S tend to zero. In this limit, the slow wave ceases to exist (in fact, its phase velocity tends to zero) whereas the dispersion relation for the fast wave reduces to

$$\omega = k\, V_A. \tag{5.46}$$

This can be identified as the dispersion relation for the *compressional-Alfvén wave*, which was introduced in Sect. 4.8. Thus, we can identify the fast wave as the compressional-Alfvén wave modified by a non-zero plasma pressure.

In the limit $V_A \gg V_S$, which is appropriate to low-β plasmas (see Sect. 3.13), the dispersion relation for the slow wave reduces to

$$\omega \simeq k\, V_S\, \cos\theta. \tag{5.47}$$

This is actually the dispersion relation of a sound wave propagating *along* magnetic field-lines. Thus, in low-β plasmas the slow wave is a sound wave modified by the presence of the magnetic field.

The distinction between the fast and slow waves can be further understood by comparing the signs of the wave induced fluctuations in the plasma and magnetic pressures: p and $\mathbf{B}_0\cdot\mathbf{B}/\mu_0$, respectively. It follows from Eq. (5.36) that

$$\frac{\mathbf{B}_0\cdot\mathbf{B}}{\mu_0} = \frac{\mathbf{k}\cdot\mathbf{V}\, B_0^{\,2} - (\mathbf{k}\cdot\mathbf{B}_0)\,(\mathbf{B}_0\cdot\mathbf{V})}{\mu_0\,\omega}. \tag{5.48}$$

Now, the z- component of Eq. (5.31) yields

$$\omega\,\rho_0\, V_z = k\,\cos\theta\, p. \tag{5.49}$$

Combining Eqs. (5.35), (5.39), (5.40), (5.48), and (5.49), we obtain

$$\frac{\mathbf{B}_0\cdot\mathbf{B}}{\mu_0} = \frac{V_A^{\,2}}{V_S^{\,2}}\left(1 - \frac{k^2\, V_S^{\,2}\,\cos^2\theta}{\omega^2}\right) p. \tag{5.50}$$

Hence, p and $\mathbf{B}_0\cdot\mathbf{B}/\mu_0$ have the same sign if $V > V_S\cos\theta$, and the opposite sign if $V < V_S\cos\theta$. Here, $V = \omega/k$ is the phase velocity. It is straightforward to show that $V_+ > V_S\cos\theta$, and $V_- < V_S\cos\theta$. Thus, we conclude that in the fast magnetosonic wave the pressure and magnetic energy fluctuations reinforce one another, whereas the fluctuations oppose one another in the slow magnetosonic wave.

Figure 5.1 shows the phase velocities of the three MHD waves plotted in the x-z plane for a low-β plasma in which $V_S < V_A$. It can be seen that the slow wave always has a smaller phase velocity than the shear-Alfvén wave, which, in turn, always has a smaller phase velocity than the fast wave.

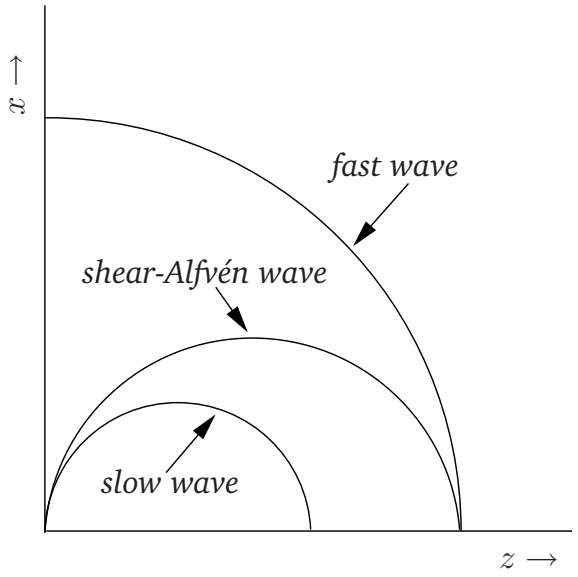

Figure 5.1: *Phase velocities of the three MHD waves in the* x*-z plane.*

5.5 The Solar Wind

The *solar wind* is a high-speed particle stream continuously blown out from the Sun into interplanetary space. It extends far beyond the orbit of the Earth, and terminates in a shock front, called the *heliopause*, where it interfaces with the weakly ionized interstellar medium. The heliopause is predicted to lie between 110 and 160 AU (1 Astronomical Unit is 1.5×10^{11} m) from the centre of the Sun. Voyager 1 is expected to pass through the heliopause sometime in the next decade: hopefully, it will still be functional at that time !

In the vicinity of the Earth, (*i.e.*, at about 1 AU from the Sun) the solar wind velocity typically ranges between 300 and 1400 $\mathrm{km\,s^{-1}}$. The average value is approximately $500\,\mathrm{km\,s^{-1}}$, which corresponds to about a 4 day time of flight from the Sun. Note that the solar wind is both *super-sonic* and *super-Alfvénic*.

The solar wind is predominately composed of protons and electrons.

Amazingly enough, the solar wind was predicted theoretically by Eugine Parker[1] a number of years *before* its existence was confirmed using satellite data.[2] Parker's prediction of a super-sonic outflow of gas from the Sun is a fascinating scientific detective story, as well as a wonderful application of plasma physics.

The solar wind originates from the *solar corona*. The solar corona is a hot, tenuous plasma surrounding the Sun, with characteristic temperatures and particle densities of about 10^6 K and $10^{14}\,\mathrm{m^{-3}}$, respectively. Note that the corona is far hotter than the solar atmosphere, or *photosphere*. In fact, the temperature of the photosphere is only about 6000 K. It is thought that the corona is heated by Alfvén waves emanating from the pho-

[1] E.N. Parker, Astrophys. J. **128**, 664 (1958).

[2] M. Neugebauer, C.W. Snyder, J. Geophys. Res. **71**, 4469 (1966).

tosphere. The solar corona is most easily observed during a total solar eclipse, when it is visible as a white filamentary region immediately surrounding the Sun.

Let us start, following Chapman,[3] by attempting to construct a model for a *static* solar corona. The equation of hydrostatic equilibrium for the corona takes the form

$$\frac{dp}{dr} = -\rho \frac{G M_\odot}{r^2}, \tag{5.51}$$

where $G = 6.67 \times 10^{-11}\,m^3\,s^{-2}\,kg^{-1}$ is the gravitational constant, and $M_\odot = 2 \times 10^{30}\,kg$ is the solar mass. The plasma density is written

$$\rho \simeq n\,m_p, \tag{5.52}$$

where n is the number density of protons. If both protons and electrons are assumed to possess a common temperature, $T(r)$, then the coronal pressure is given by

$$p = 2\,n\,T. \tag{5.53}$$

The thermal conductivity of the corona is dominated by the electron thermal conductivity, and takes the form [see Eqs. (3.95) and (3.115)]

$$\kappa = \kappa_0\,T^{5/2}, \tag{5.54}$$

where κ_0 is a relatively weak function of density and temperature. For typical coronal conditions this conductivity is extremely high: *i.e.*, it is about twenty times the thermal conductivity of copper at room temperature. The coronal heat flux density is written

$$\mathbf{q} = -\kappa\,\nabla T. \tag{5.55}$$

For a static corona, in the absence of energy sources or sinks, we require

$$\nabla\cdot\mathbf{q} = 0. \tag{5.56}$$

Assuming spherical symmetry, this expression reduces to

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\,\kappa_0\,T^{5/2}\,\frac{dT}{dr}\right) = 0. \tag{5.57}$$

Adopting the sensible boundary condition that the coronal temperature must tend to zero at large distances from the Sun, we obtain

$$T(r) = T(a)\left(\frac{a}{r}\right)^{2/7}. \tag{5.58}$$

The reference level $r = a$ is conveniently taken to be the base of the corona, where $a \sim 7 \times 10^5\,km$, $n \sim 2 \times 10^{14}\,m^{-3}$, and $T \sim 2 \times 10^6\,K$.

[3] S. Chapman, Smithsonian Contrib. Astrophys. **2**, 1 (1957).

Equations (5.51), (5.52), (5.53), and (5.58) can be combined and integrated to give

$$p(r) = p(a) \exp\left\{\frac{7}{5}\,\frac{G\,M_\odot\,m_p}{2\,T(a)\,a}\left[\left(\frac{a}{r}\right)^{5/7} - 1\right]\right\}. \tag{5.59}$$

Note that as $r \rightarrow \infty$ the coronal pressure tends towards a finite constant value:

$$p(\infty) = p(a)\ \exp\left\{-\frac{7}{5}\,\frac{G\,M_\odot\,m_p}{2\,T(a)\,a}\right\}. \tag{5.60}$$

There is, of course, nothing at large distances from the Sun which could contain such a pressure (the pressure of the interstellar medium is negligibly small). Thus, we conclude, with Parker, that the static coronal model is *unphysical*.

Since we have just demonstrated that a static model of the solar corona is unsatisfactory, let us now attempt to construct a *dynamic* model in which material flows outward from the Sun.

5.6 Parker Model of Solar Wind

By symmetry, we expect a purely radial coronal outflow. The radial momentum conservation equation for the corona takes the form

$$\rho\,u\,\frac{du}{dr} = -\frac{dp}{dr} - \rho\,\frac{G\,M_\odot}{r^2}, \tag{5.61}$$

where u is the radial expansion speed. The continuity equation reduces to

$$\frac{1}{r^2}\frac{d(r^2\,\rho\,u)}{dr} = 0. \tag{5.62}$$

In order to obtain a closed set of equations, we now need to adopt an equation of state for the corona, relating the pressure, p, and the density, ρ. For the sake of simplicity, we adopt the simplest conceivable equation of state, which corresponds to an *isothermal* corona. Thus, we have

$$p = \frac{2\,\rho\,T}{m_p}, \tag{5.63}$$

where T is a constant. Note that more realistic equations of state complicate the analysis, but do not significantly modify any of the physics results.

Equation (5.62) can be integrated to give

$$r^2\,\rho\,u = I, \tag{5.64}$$

where I is a constant. The above expression simply states that the mass flux per unit solid angle, which takes the value I, is independent of the radius, r. Equations (5.61), (5.63), and (5.64) can be combined together to give

$$\frac{1}{u}\,\frac{du}{dr}\left(u^2 - \frac{2\,T}{m_p}\right) = \frac{4\,T}{m_p\,r} - \frac{G\,M_\odot}{r^2}. \tag{5.65}$$

Let us restrict our attention to coronal temperatures which satisfy

$$T < T_c \equiv \frac{G\, M_\odot\, m_p}{4\, a}, \tag{5.66}$$

where a is the radius of the base of the corona. For typical coronal parameters (see above), $T_c \simeq 5.8 \times 10^6\,$K, which is certainly greater than the temperature of the corona at $r = a$. For $T < T_c$, the right-hand side of Eq. (5.65) is negative for $a < r < r_c$, where

$$\frac{r_c}{a} = \frac{T_c}{T}, \tag{5.67}$$

and positive for $r_c < r < \infty$. The right-hand side of (5.65) is zero at $r = r_c$, implying that the left-hand side is also zero at this radius, which is usually termed the "critical radius." There are two ways in which the left-hand side of (5.65) can be zero at the critical radius. Either

$$u^2(r_c) = {u_c}^2 \equiv \frac{2\,T}{m_p}, \tag{5.68}$$

or

$$\frac{du(r_c)}{dr} = 0. \tag{5.69}$$

Note that u_c is the coronal *sound speed*.

As is easily demonstrated, if Eq. (5.68) is satisfied then du/dr has the same sign for all r, and $u(r)$ is either a monotonically increasing, or a monotonically decreasing, function of r. On the other hand, if Eq. (5.69) is satisfied then $u^2 - {u_c}^2$ has the same sign for all r, and $u(r)$ has an extremum close to $r = r_c$. The flow is either super-sonic for all r, or sub-sonic for all r. These possibilities lead to the existence of *four* classes of solutions to Eq. (5.65), with the following properties:

1. $u(r)$ is sub-sonic throughout the domain $a < r < \infty$. $u(r)$ increases with r, attains a maximum value around $r = r_c$, and then decreases with r.

2. a unique solution for which $u(r)$ increases monotonically with r, and $u(r_c) = u_c$.

3. a unique solution for which $u(r)$ decreases monotonically with r, and $u(r_c) = u_c$.

4. $u(r)$ is super-sonic throughout the domain $a < r < \infty$. $u(r)$ decreases with r, attains a minimum value around $r = r_c$, and then increases with r.

These four classes of solutions are illustrated in Fig. 5.2.

Each of the classes of solutions described above fits a different set of boundary conditions at $r = a$ and $r \rightarrow \infty$. The *physical* acceptability of these solutions depends on these boundary conditions. For example, both Class 3 and Class 4 solutions can be ruled out as plausible models for the solar corona since they predict *super-sonic* flow at the base of the corona, which is not observed, and is also not consistent with a static solar photosphere. Class 1 and Class 2 solutions remain acceptable models for the solar corona on the basis

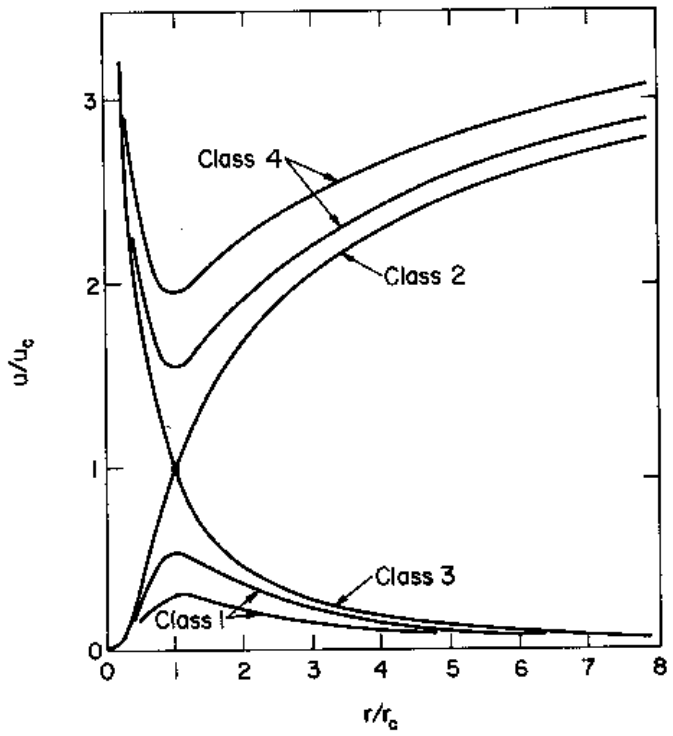

Figure 5.2: *The four classes of Parker outflow solutions for the solar wind.*

of their properties around $r = a$, since they both predict sub-sonic flow in this region. However, the Class 1 and Class 2 solutions behave quite differently as $r \rightarrow \infty$, and the physical acceptability of these two classes hinges on this difference.

Equation (5.65) can be rearranged to give

$$\frac{du^2}{dr}\left(1-\frac{u_c^{\,2}}{u^2}\right) = \frac{4\,u_c^{\,2}}{r}\left(1-\frac{r_c}{r}\right), \tag{5.70}$$

where use has been made of Eqs. (5.66) and (5.67). The above expression can be integrated to give

$$\left(\frac{u}{u_c}\right)^2 - \ln\left(\frac{u}{u_c}\right)^2 = 4\,\ln r + 4\,\frac{r_c}{r} + C, \tag{5.71}$$

where C is a constant of integration.

Let us consider the behaviour of Class 1 solutions in the limit $r \rightarrow \infty$. It is clear from Fig. 5.2 that, for Class 1 solutions, u/u_c is less than unity and monotonically decreasing as $r \rightarrow \infty$. In the large-r limit, Eq. (5.71) reduces to

$$\ln\frac{u}{u_c} \simeq -2\,\ln r, \tag{5.72}$$

so that

$$u \propto \frac{1}{r^2}. \tag{5.73}$$

It follows from Eq. (5.64) that the coronal density, ρ, approaches a finite, constant value, ρ_∞, as $r \rightarrow \infty$. Thus, the Class 1 solutions yield a finite pressure,

$$p_\infty = \frac{2\,\rho_\infty\,T}{m_p}, \tag{5.74}$$

at large r, which cannot be matched to the much smaller pressure of the interstellar medium. Clearly, Class 1 solutions are unphysical.

Let us consider the behaviour of the Class 2 solution in the limit $r \rightarrow \infty$. It is clear from Fig. 5.2 that, for the Class 2 solution, u/u_c is greater than unity and monotonically increasing as $r \rightarrow \infty$. In the large-r limit, Eq. (5.71) reduces to

$$\left(\frac{u}{u_c}\right)^2 \simeq 4\,\ln r, \tag{5.75}$$

so that

$$u \simeq 2\,u_c\,(\ln r)^{1/2}. \tag{5.76}$$

It follows from Eq. (5.64) that $\rho \rightarrow 0$ and $r \rightarrow \infty$. Thus, the Class 2 solution yields $p \rightarrow 0$ at large r, and can, therefore, be matched to the low pressure interstellar medium.

We conclude that the only solution to Eq. (5.65) which is consistent with physical boundary conditions at $r = a$ and $r \rightarrow \infty$ is the Class 2 solution. This solution predicts that the solar corona expands radially outward at relatively modest, sub-sonic velocities

close to the Sun, and gradually accelerates to super-sonic velocities as it moves further away from the Sun. Parker termed this continuous, super-sonic expansion of the corona the *solar wind*.

Equation (5.71) can be rewritten

$$\left[\frac{u^2}{{u_c}^2} - 1\right] - \ln \frac{u^2}{{u_c}^2} = 4 \ln \frac{r}{r_c} + 4 \left[\frac{r_c}{r} - 1\right], \tag{5.77}$$

where the constant C is determined by demanding that $u = u_c$ when $r = r_c$. Note that both u_c and r_c can be evaluated in terms of the coronal temperature T via Eqs. (5.67) and (5.68). Figure 5.3 shows $u(r)$ calculated from Eq. (5.77) for various values of the coronal temperature. It can be seen that plausible values of T (*i.e.*, $T \sim 1$–2×10^6 K) yield expansion speeds of several hundreds of kilometers per second at 1 AU, which accords well with satellite observations. The critical surface at which the solar wind makes the transition from sub-sonic to super-sonic flow is predicted to lie a few solar radii away from the Sun (*i.e.*, $r_c \sim 5\,R_\odot$). Unfortunately, the Parker model's prediction for the density of the solar wind at the Earth is significantly too high compared to satellite observations. Consequently, there have been many further developments of this model. In particular, the unrealistic assumption that the solar wind plasma is isothermal has been relaxed, and two-fluid effects have been incorporated into the analysis.[4]

5.7 Interplanetary Magnetic Field

Let us now investigate how the solar wind and the interplanetary magnetic field affect one another.

The hot coronal plasma making up the solar wind possesses an extremely high electrical conductivity. In such a plasma, we expect the concept of "frozen-in" magnetic field-lines, discussed in Sect. 5.3, to be applicable. The continuous flow of coronal material into interplanetary space must, therefore, result in the transport of the solar magnetic field into the interplanetary region. If the Sun did not rotate, the resulting magnetic configuration would be very simple. The radial coronal expansion considered above (with the neglect of any magnetic forces) would produce magnetic field-lines extending radially outward from the Sun.

Of course, the Sun does rotate, with a (latitude dependent) period of about 25 days.[5] Since the solar photosphere is an excellent electrical conductor, the magnetic field at the base of the corona is frozen into the rotating frame of reference of the Sun. A magnetic field-line starting from a given location on the surface of the Sun is drawn out along the path followed by the element of the solar wind emanating from that location. As before, let us suppose that the coronal expansion is purely radial in a stationary frame of reference. Consider a spherical polar coordinate system (r, θ, ϕ) which *co-rotates* with the Sun. Of

[4] *Solar Magnetohydrodynamics*, E.R. Priest, (D. Reidel Publishing Co., Dordrecht, Netherlands, 1987).

[5] To an observer orbiting with the Earth, the rotation period appears to be about 27 days.

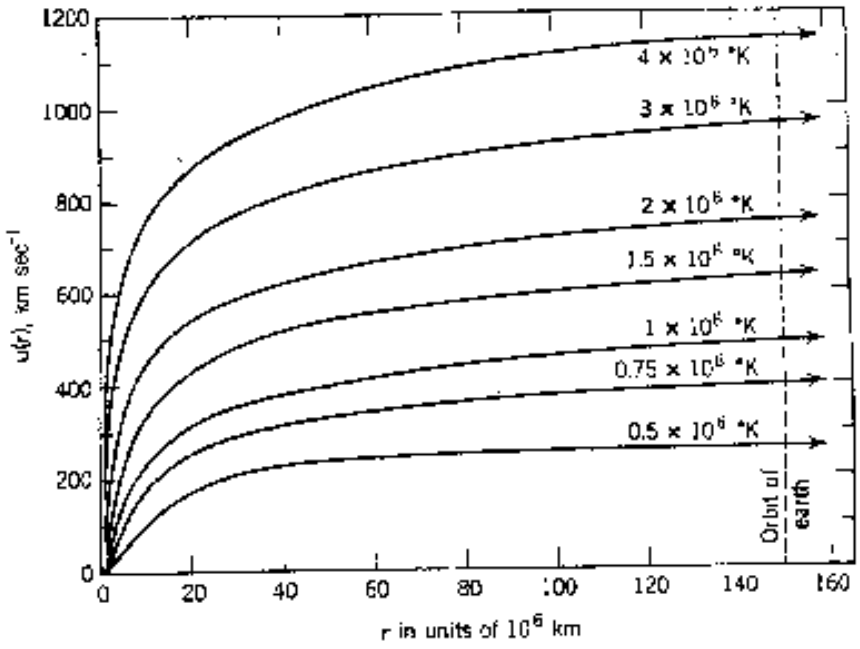

Figure 5.3: *Parker outflow solutions for the solar wind.*

course, the symmetry axis of the coordinate system is assumed to coincide with the axis of the Sun's rotation. In the rotating coordinate system, the velocity components of the solar wind are written

$$u_r = u, \tag{5.78}$$

$$u_\theta = 0, \tag{5.79}$$

$$u_\phi = -\Omega\, r\, \sin\theta, \tag{5.80}$$

where $\Omega = 2.7 \times 10^{-6}\,\mathrm{rad\,sec^{-1}}$ is the angular velocity of solar rotation. The azimuthal velocity u_ϕ is entirely due to the transformation to the rotating frame of reference. The stream-lines of the flow satisfy the differential equation

$$\frac{1}{r\,\sin\theta}\frac{dr}{d\phi} \simeq \frac{u_r}{u_\phi} = -\frac{u}{\Omega\, r\, \sin\theta} \tag{5.81}$$

at constant θ. The stream-lines are also magnetic field-lines, so Eq. (5.81) can also be regarded as the differential equation of a magnetic field-line. For radii r greater than several times the critical radius, r_c, the solar wind solution (5.77) predicts that $u(r)$ is almost constant (see Fig. 5.3). Thus, for $r \gg r_c$ it is reasonable to write $u(r) = u_s$, where u_s is a constant. Equation (5.81) can then be integrated to give the equation of a magnetic field-line:

$$r - r_0 = -\frac{u_s}{\Omega}\,(\phi - \phi_0), \tag{5.82}$$

where the field-line is assumed to pass through the point (r_0, θ, ϕ_0). Maxwell's equation $\nabla\cdot\mathbf{B} = 0$, plus the assumption of a spherically symmetric magnetic field, easily yields the following expressions for the components of the interplanetary magnetic field:

$$B_r(r,\theta,\phi) = B(r_0,\theta,\phi_0)\left(\frac{r_0}{r}\right)^2, \tag{5.83}$$

$$B_\theta(r,\theta,\phi) = 0, \tag{5.84}$$

$$B_\phi(r,\theta,\phi) = -B(r_0,\theta,\phi_0)\,\frac{\Omega\, r_0}{u_s}\,\frac{r_0}{r}\,\sin\theta. \tag{5.85}$$

Figure 5.4 illustrates the interplanetary magnetic field close to the ecliptic plane. The magnetic field-lines of the Sun are drawn into spirals (Archemedian spirals, to be more exact) by the solar rotation. Transformation to a stationary frame of reference give the same magnetic field configuration, with the addition of an electric field

$$\mathbf{E} = -\mathbf{u}\times\mathbf{B} = -u_s\, B_\phi\, \hat{\theta}. \tag{5.86}$$

The latter field arises because the radial plasma flow is no longer parallel to magnetic field-lines in the stationary frame.

The interplanetary magnetic field at 1 AU is observed to lie in the ecliptic plane, and is directed at an angle of approximately 45° from the radial direction to the Sun. This is in basic agreement with the spiral configuration predicted above.

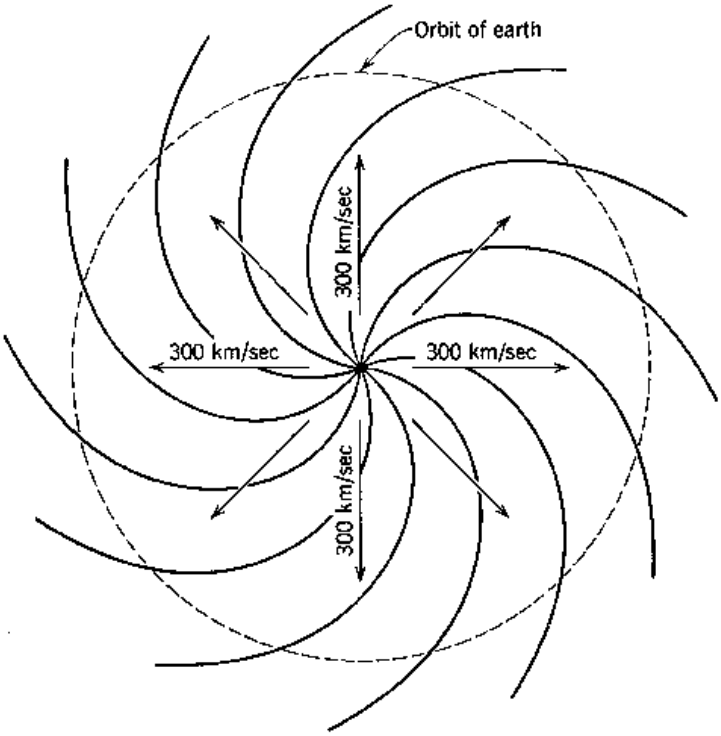

Figure 5.4: *The interplanetary magnetic field.*

The analysis presented above is premised on the assumption that the interplanetary magnetic field is too weak to affect the coronal outflow, and is, therefore, passively convected by the solar wind. In fact, this is only likely to be the case if the interplanetary magnetic energy density, $B^2/2\,\mu_0$, is much less that the kinetic energy density, $\rho\,u^2/2$, of the solar wind. Rearrangement yields the condition

$$u > V_A, \tag{5.87}$$

where V_A is the Alfvén speed. It turns out that $u \sim 10\,V_A$ at 1 AU. On the other hand, $u \ll V_A$ close to the base of the corona. In fact, the solar wind becomes super-Alfvénic at a radius, denoted r_A, which is typically $50\,R_\odot$, or 1/4 of an astronomical unit. We conclude that the previous analysis is only valid well outside the Alfvén radius: *i.e.*, in the region $r \gg r_A$.

Well inside the Alfvén radius (*i.e.*, in the region $r \ll r_A$), the solar wind is too weak to modify the structure of the solar magnetic field. In fact, in this region we expect the solar magnetic field to force the solar wind to *co-rotate* with the Sun. Note that flux-freezing is a two-way-street: if the energy density of the flow greatly exceeds that of the magnetic field then the magnetic field is passively convected by the flow, but if the energy density of the magnetic field greatly exceeds that of the flow then the flow is forced to conform to the magnetic field.

The above discussion leads us to the following rather crude picture of the interaction of the solar wind and the interplanetary magnetic field. We expect the interplanetary magnetic field to be simply the undistorted continuation of the Sun's magnetic field for $r < r_A$. On the other hand, we expect the interplanetary field to be dragged out into a spiral pattern for $r > r_A$. Furthermore, we expect the Sun's magnetic field to impart a non-zero azimuthal velocity $u_\phi(r)$ to the solar wind. In the ecliptic plane, we expect

$$u_\phi = \Omega\, r \tag{5.88}$$

for $r < r_A$, and

$$u_\phi = \Omega\, r_A \left(\frac{r_A}{r}\right) \tag{5.89}$$

for $r > r_A$. This corresponds to co-rotation with the Sun inside the Alfvén radius, and outflow at constant angular velocity outside the Alfvén radius. We, therefore, expect the solar wind at 1 AU to possess a small azimuthal velocity component. This is indeed the case. In fact, the direction of the solar wind at 1 AU deviates from purely radial outflow by about 1.5°.

5.8 Mass and Angular Momentum Loss

Since the Sun is the best observed of any star, it is interesting to ask what impact the solar wind has as far as solar, and stellar, evolution are concerned. The most obvious question is whether the mass loss due to the wind is significant, or not. Using typical measured

values (*i.e.*, a typical solar wind velocity and particle density at 1 AU of $500\,\mathrm{km\,s^{-1}}$ and $7\times10^6\ \mathrm{m^{-3}}$, respectively), the Sun is apparently losing mass at a rate of $3\times10^{-14}\,M_\odot$ per year, implying a time-scale for significant mass loss of 3×10^{13} years, or some $6,000$ times longer than the estimated 5×10^9 year age of the Sun. Clearly, the mass carried off by the solar wind has a negligible effect on the Sun's evolution. Note, however, that many other stars in the Galaxy exhibit significant mass loss via stellar winds. This is particularly the case for late-type stars.

Let us now consider the angular momentum carried off by the solar wind. Angular momentum loss is a crucially important topic in astrophysics, since only by losing angular momentum can large, diffuse objects, such as interstellar gas clouds, collapse under the influence of gravity to produce small, compact objects, such as stars and proto-stars. Magnetic fields generally play a crucial role in angular momentum loss. This is certainly the case for the solar wind, where the solar magnetic field enforces co-rotation with the Sun out to the Alfvén radius, r_A. Thus, the angular momentum carried away by a particle of mass m is $\Omega\, r_A^2\, m$, rather than $\Omega\, R_\odot^2\, m$. The angular momentum loss time-scale is, therefore, shorter than the mass loss time-scale by a factor $(R_\odot/r_A)^2 \simeq 1/2500$, making the angular momentum loss time-scale comparable to the solar lifetime. It is clear that magnetized stellar winds represent a very important vehicle for angular momentum loss in the Universe. Let us investigate angular momentum loss via stellar winds in more detail.

Under the assumption of spherical symmetry and steady flow, the azimuthal momentum evolution equation for the solar wind, taking into account the influence of the interplanetary magnetic field, is written

$$\rho\,\frac{u_r}{r}\,\frac{d(r\,u_\phi)}{dr} = (\mathbf{j}\times\mathbf{B})_\phi = \frac{B_r}{\mu_0\,r}\,\frac{d(rB_\phi)}{dr}. \tag{5.90}$$

The constancy of the mass flux [see Eq. (5.64)] and the $1/r^2$ dependence of B_r [see Eq. (5.83)] permit the immediate integration of the above equation to give

$$r\,u_\phi - \frac{r\,B_r\,B_\phi}{\mu_0\,\rho\,u_r} = L, \tag{5.91}$$

where L is the angular momentum per unit mass carried off by the solar wind. In the presence of an azimuthal wind velocity, the magnetic field and velocity components are related by an expression similar to Eq. (5.81):

$$\frac{B_r}{B_\phi} = \frac{u_r}{u_\phi - \Omega\, r\,\sin\theta}. \tag{5.92}$$

The fundamental physics assumption underlying the above expression is the absence of an electric field in the frame of reference co-rotating with the Sun. Using Eq. (5.92) to eliminate B_ϕ from Eq. (5.91), we obtain (in the ecliptic plane, where $\sin\theta = 1$)

$$r\,u_\phi = \frac{L\,M_A^2 - \Omega\, r^2}{M_A^2 - 1}, \tag{5.93}$$

where

$$M_A = \sqrt{\frac{u_r^2}{B_r^2/\mu_0\,\rho}} \tag{5.94}$$

is the *radial Alfvén Mach number*. The radial Alfvén Mach number is small near the base of the corona, and about 10 at 1 AU: it passes through unity at the Alfvén radius, r_A, which is about 0.25 AU from the Sun. The zero denominator on the right-hand side of Eq. (5.93) at $r = r_A$ implies that u_ϕ is finite and continuous only if the numerator is also zero at the Alfvén radius. This condition then determines the angular momentum content of the outflow via

$$L = \Omega\, r_A^2. \tag{5.95}$$

Note that the angular momentum carried off by the solar wind is indeed equivalent to that which would be carried off were coronal plasma to co-rotate with the Sun out to the Alfvén radius, and subsequently outflow at constant angular velocity. Of course, the solar wind does not actually rotate rigidly with the Sun in the region $r < r_A$: much of the angular momentum in this region is carried in the form of electromagnetic stresses.

It is easily demonstrated that the quantity $M_A^2/u_r\,r^2$ is a constant, and can, therefore, be evaluated at $r = r_A$ to give

$$M_A^2 = \frac{u_r\,r^2}{u_{rA}\,r_A^2}, \tag{5.96}$$

where $u_{rA} \equiv u_r(r_A)$. Equations (5.93), (5.95), and (5.96) can be combined to give

$$u_\phi = \frac{\Omega\, r}{u_{rA}}\frac{u_{rA} - u_r}{1 - M_A^2}. \tag{5.97}$$

In the limit $r \rightarrow \infty$, we have $M_A \gg 1$, so the above expression yields

$$u_\phi \rightarrow \Omega\, r_A \left(\frac{r_A}{r}\right)\left(1 - \frac{u_{rA}}{u_r}\right) \tag{5.98}$$

at large distances from the Sun. Recall, from Sect. 5.7, that if the coronal plasma were to simply co-rotate with the Sun out to $r = r_A$, and experience no torque beyond this radius, then we would expect

$$u_\phi \rightarrow \Omega\, r_A \left(\frac{r_A}{r}\right) \tag{5.99}$$

at large distances from the Sun. The difference between the above two expressions is the factor $1 - u_{rA}/u_r$, which is a correction for the angular momentum retained by the magnetic field at large r.

The analysis presented above was first incorporated into a quantitative coronal expansion model by Weber and Davis.[6] The model of Weber and Davis is very complicated. For instance, the solar wind is required to flow smoothly through no less than *three* critical points. These are associated with the sound speed (as in Parker's original model), the

[6]E.J. Weber, and L. Davis Jr., Astrophys. J. **148**, 217 (1967).

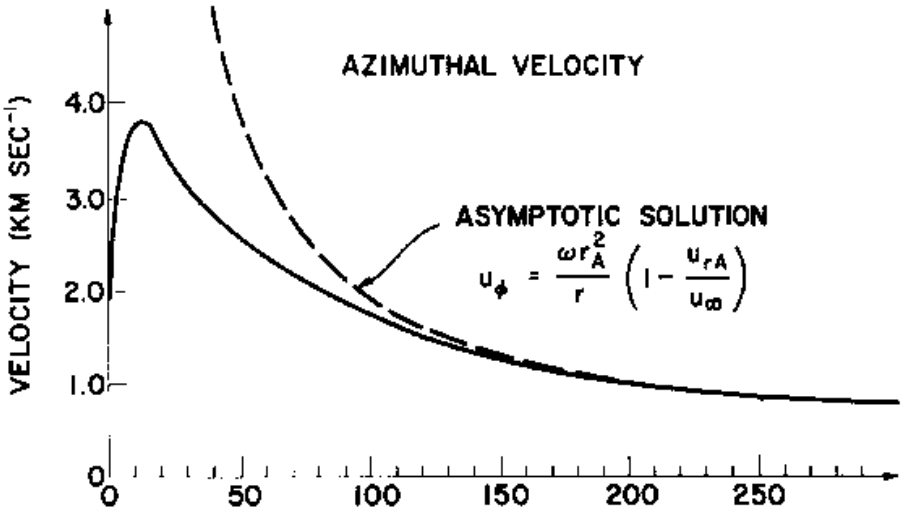

Figure 5.5: *Comparison of asymptotic form for azimuthal flow velocity of solar wind with Weber-Davis solution.*

radial Alfvén speed, $B_r/\sqrt{\mu_0\,\rho}$, (as described above), and the total Alfvén speed, $B/\sqrt{\mu_0\,\rho}$. Nevertheless, the simplified analysis outlined above captures most of the essential features of the outflow. For instance, Fig. 5.5 shows a comparison between the large-r asymptotic form for the azimuthal flow velocity predicted above [see Eq. (5.98)] and that calculated by Weber and Davis, showing the close agreement between the two.

5.9 MHD Dynamo Theory

Many stars, planets, and galaxies possess magnetic fields whose origins are not easily explained. Even the "solid" planets could not possibly be sufficiently ferromagnetic to account for their magnetism, since the bulk of their interiors are above the Curie temperature at which permanent magnetism disappears. It goes without saying that stars and galaxies cannot be ferromagnetic at all. Magnetic fields cannot be dismissed as transient phenomena which just happen to be present today. For instance, *paleomagnetism*, the study of magnetic fields "fossilized" in rocks at the time of their formation in the remote geological past, shows that the Earth's magnetic field has existed at much its present strength for at least the past 3×10^9 years. The problem is that, in the absence of an internal source of electric currents, magnetic fields contained in a conducting body *decay* ohmically on a time-scale

$$\tau_{\rm ohm} = \mu_0\,\sigma\,L^2, \tag{5.100}$$

where σ is the typical electrical conductivity, and L is the typical length-scale of the body, and this decay time-scale is generally very *small* compared to the inferred lifetimes of astronomical magnetic fields. For instance, the Earth contains a highly conducting region,

namely, its molten core, of radius $L \sim 3.5 \times 10^6\,\mathrm{m}$, and conductivity $\sigma \sim 4 \times 10^5\,\mathrm{S\,m^{-1}}$. This yields an ohmic decay time for the terrestrial magnetic field of only $\tau_{\rm ohm} \sim 2 \times 10^5$ years, which is obviously far shorter than the inferred lifetime of this field. Clearly, some process inside the Earth must be actively maintaining the terrestrial magnetic field. Such a process is conventionally termed a *dynamo*. Similar considerations lead us to postulate the existence of dynamos acting inside stars and galaxies, in order to account for the persistence of stellar and galactic magnetic fields over cosmological time-scales.

The basic premise of dynamo theory is that all astrophysical bodies which contain anomalously long-lived magnetic fields also contain highly conducting fluids (*e.g.*, the Earth's molten core, the ionized gas which makes up the Sun), and it is the electric currents associated with the motions of these fluids which maintain the observed magnetic fields. At first sight, this proposal, first made by Larmor in 1919,[7] sounds suspiciously like pulling yourself up by your own shoelaces. However, there is really no conflict with the demands of energy conservation. The magnetic energy irreversibly lost via ohmic heating is replenished at the rate (per unit volume) $\mathbf{V}\cdot(\mathbf{j}\times\mathbf{B})$: *i.e.*, by the rate of work done against the Lorentz force. The flow field, $\mathbf{V}$, is assumed to be driven via thermal convention. If the flow is sufficiently vigorous then it is, at least, plausible that the energy input to the magnetic field can overcome the losses due to ohmic heating, thus permitting the field to persist over time-scales far longer than the characteristic ohmic decay time.

Dynamo theory involves two vector fields, $\mathbf{V}$ and $\mathbf{B}$, coupled by a rather complicated force: *i.e.*, the Lorentz force. It is not surprising, therefore, that dynamo theory tends to be extremely complicated, and is, at present, far from completely understood. Fig. 5.6 shows paleomagnetic data illustrating the variation of the polarity of the Earth's magnetic field over the last few million years, as deduced from marine sediment cores. It can be seen that the Earth's magnetic field is quite variable, and actually reversed polarity about $700,000$ years ago. In fact, more extensive data shows that the Earth's magnetic field reverses polarity about once every ohmic decay time-scale (*i.e.*, a few times every million years). The Sun's magnetic field exhibits similar behaviour, reversing polarity about once every 11 years. It is clear from examining this type of data that dynamo magnetic fields (and velocity fields) are essentially *chaotic* in nature, exhibiting strong random variability superimposed on more regular quasi-periodic oscillations.

Obviously, we are not going to attempt to tackle full-blown dynamo theory in this course: that would be far too difficult. Instead, we shall examine a far simpler theory, known as *kinematic dynamo theory*, in which the velocity field, $\mathbf{V}$, is *prescribed*. In order for this approach to be self-consistent, the magnetic field must be assumed to be sufficiently small that it does not affect the velocity field. Let us start from the MHD Ohm's law, modified by resistivity:

$$\mathbf{E} + \mathbf{V}\times\mathbf{B} = \eta\,\mathbf{j}. \tag{5.101}$$

Here, the resistivity η is assumed to be a constant, for the sake of simplicity. Taking the

[7] J. Larmor, Brit. Assoc. Reports, **159** (1919).

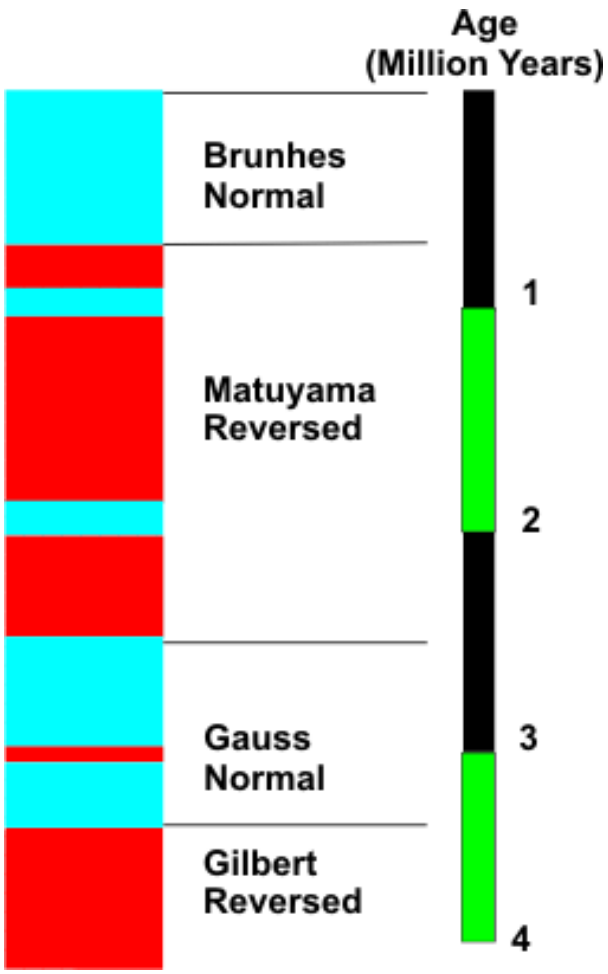

Figure 5.6: *Polarity of the Earth's magnetic field as a function of time, as deduced from marine sediment cores.*

curl of the above equation, and making use of Maxwell's equations, we obtain

$$\frac{\partial \mathbf{B}}{\partial t} - \nabla \times (\mathbf{V} \times \mathbf{B}) = \frac{\eta}{\mu_0} \nabla^2 \mathbf{B}. \tag{5.102}$$

If the velocity field, **V**, is prescribed, and unaffected by the presence of the magnetic field, then the above equation is essentially a *linear eigenvalue* equation for the magnetic field, **B**. The question we wish to address is as follows: for what sort of velocity fields, if any, does the above equation possess solutions where the magnetic field grows exponentially? In trying to answer this question, we hope to learn what type of motion of an MHD fluid is capable of self-generating a magnetic field.

5.10 Homopolar Generators

Some of the peculiarities of dynamo theory are well illustrated by the prototype example of self-excited dynamo action, which is the *homopolar disk dynamo*. As illustrated in Fig. 5.7, this device consists of a conducting disk which rotates at angular frequency Ω about its axis under the action of an applied torque. A wire, twisted about the axis in the manner shown, makes sliding contact with the disc at A, and with the axis at B, and carries a current $I(t)$. The magnetic field **B** associated with this current has a flux $\Phi = M\,I$ across the disc, where M is the mutual inductance between the wire and the rim of the disc. The rotation of the disc in the presence of this flux generates a radial electromotive force

$$\frac{\Omega}{2\pi}\,\Phi = \frac{\Omega}{2\pi}\,M\,I, \tag{5.103}$$

since a radius of the disc cuts the magnetic flux Φ once every $2\pi/\Omega$ seconds. According to this simplistic description, the equation for I is written

$$L\,\frac{dI}{dt} + R\,I = \frac{M}{2\pi}\,\Omega\,I, \tag{5.104}$$

where R is the total resistance of the circuit, and L is its self-inductance.

Suppose that the angular velocity Ω is maintained by suitable adjustment of the driving torque. It follows that Eq. (5.104) possesses an exponential solution $I(t) = I(0)\,\exp(\gamma\,t)$, where

$$\gamma = L^{-1}\left[\frac{M}{2\pi}\,\Omega - R\right]. \tag{5.105}$$

Clearly, we have exponential growth of $I(t)$, and, hence, of the magnetic field to which it gives rise (*i.e.*, we have dynamo action), provided that

$$\Omega > \frac{2\pi\,R}{M}: \tag{5.106}$$

i.e., provided that the disk rotates rapidly enough. Note that the homopolar generator depends for its success on its built-in axial *asymmetry*. If the disk rotates in the opposite

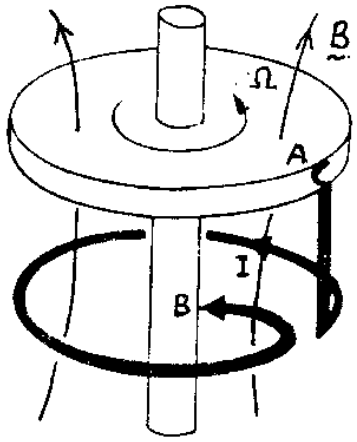

Figure 5.7: *The homopolar generator.*

direction to that shown in Fig. 5.7 then $\Omega < 0$, and the electromotive force generated by the rotation of the disk always acts to reduce I. In this case, dynamo action is impossible (*i.e.*, γ is always negative). This is a troubling observation, since most astrophysical objects, such as stars and planets, possess very good axial symmetry. We conclude that if such bodies are to act as dynamos then the asymmetry of their internal motions must somehow compensate for their lack of built-in asymmetry. It is far from obvious how this is going to happen.

Incidentally, although the above analysis of a homopolar generator (which is the standard analysis found in most textbooks) is very appealing in its simplicity, it cannot be entirely correct. Consider the limiting situation of a *perfectly conducting* disk and wire, in which $R = 0$. On the one hand, Eq. (5.105) yields $\gamma = M\,\Omega/2\pi\,L$, so that we still have dynamo action. But, on the other hand, the rim of the disk is a closed circuit embedded in a perfectly conducting medium, so the flux freezing constraint requires that the flux, Φ, through this circuit must remain a constant. There is an obvious contradiction. The problem is that we have neglected the currents that flow azimuthally in the disc: *i.e.*, the very currents which control the diffusion of magnetic flux across the rim of the disk. These currents become particularly important in the limit $R \rightarrow \infty$.

The above paradox can be resolved by supposing that the azimuthal current $J(t)$ is constrained to flow around the rim of the disk (*e.g.*, by a suitable distribution of radial insulating strips). In this case, the fluxes through the I and J circuits are

$$\Phi_1 = L\,I + M\,J, \tag{5.107}$$

$$\Phi_2 = M\,I + L'\,J, \tag{5.108}$$

and the equations governing the current flow are

$$\frac{d\Phi_1}{dt} = \frac{\Omega}{2\pi}\,\Phi_2 - R\,I, \tag{5.109}$$

$$\frac{d\Phi_2}{dt} = -R' J, \tag{5.110}$$

where R', and L' refer to the J circuit. Let us search for exponential solutions, $(I, J) \propto \exp(\gamma t)$, of the above system of equations. It is easily demonstrated that

$$\gamma = \frac{-[L R' + L' R] \pm \sqrt{[L R' + L' R]^2 + 4 R' [L L' - M^2] [M\Omega/2\pi - R]}}{2 [L L' - M^2]}. \tag{5.111}$$

Recall the standard result in electromagnetic theory that $L L' > M^2$ for two non-coincident circuits. It is clear, from the above expression, that the condition for dynamo action (*i.e.*, $\gamma > 0$) is

$$\Omega > \frac{2\pi R}{M}, \tag{5.112}$$

as before. Note, however, that $\gamma \rightarrow 0$ as $R' \rightarrow 0$. In other words, if the rotating disk is a perfect conductor then dynamo action is impossible. The above system of equations can transformed into the well-known Lorenz system, which exhibits *chaotic* behaviour in certain parameter regimes.[8] It is noteworthy that this simplest prototype dynamo system already contains the seeds of chaos (provided that the formulation is self-consistent).

It is clear from the above discussion that, whilst dynamo action requires the resistance of the circuit, R, to be low, we lose dynamo action altogether if we go to the perfectly conducting limit, $R \rightarrow 0$, because magnetic fields are unable to diffuse into the region in which magnetic induction is operating. Thus, an efficient dynamo requires a conductivity that is large, but not too large.

5.11 Slow and Fast Dynamos

Let us search for solutions of the MHD kinematic dynamo equation,

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{V} \times \mathbf{B}) + \frac{\eta}{\mu_0} \nabla^2 \mathbf{B}, \tag{5.113}$$

for a prescribed *steady-state* velocity field, $\mathbf{V}(\mathbf{r})$, subject to certain practical constraints. Firstly, we require a *self-contained* solution: *i.e.*, a solution in which the magnetic field is maintained by the motion of the MHD fluid, rather than by currents at infinity. This suggests that $V, B \rightarrow 0$ as $r \rightarrow \infty$. Secondly, we require an exponentially growing solution: *i.e.*, a solution for which $\mathbf{B} \propto \exp(\gamma t)$, where $\gamma > 0$.

In most MHD fluids occurring in astrophysics, the resistivity, η, is extremely small. Let us consider the perfectly conducting limit, $\eta \rightarrow 0$. In this limit, Vainshtein and Zel'dovich, in 1978, introduced an important distinction between two fundamentally different classes

[8] E. Knobloch, Phys. Lett. **82A**, 439 (1981).

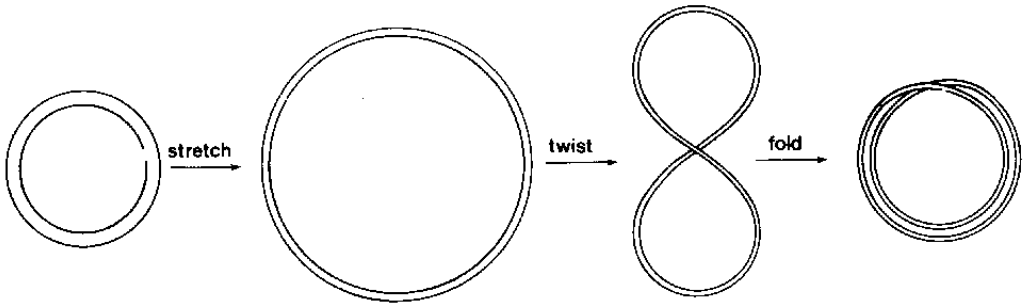

Figure 5.8: *The stretch-twist-fold cycle of a fast dynamo.*

of dynamo solutions.[9] Suppose that we solve the eigenvalue equation (5.113) to obtain the growth-rate, γ, of the magnetic field in the limit $\eta \rightarrow 0$. We expect that

$$\lim_{\eta\rightarrow 0} \gamma \propto \eta^{\alpha}, \tag{5.114}$$

where $0 \leq \alpha \leq 1$. There are two possibilities. Either $\alpha > 0$, in which case the growth-rate depends on the resistivity, or $\alpha = 0$, in which case the growth-rate is independent of the resistivity. The former case is termed a *slow dynamo*, whereas the latter case is termed a *fast dynamo*. By definition, slow dynamos are unable to operate in the perfectly conducting limit, since $\gamma \rightarrow 0$ as $\eta \rightarrow 0$. On the other hand, fast dynamos can, in principle, operate when $\eta = 0$.

It is clear, from the above discussion, that a homopolar disk generator is an example of a slow dynamo. In fact, it is easily seen that any dynamo which depends on the motion of a *rigid* conductor for its operation is bound to be a slow dynamo: in the perfectly conducting limit, the magnetic flux linking the conductor could never change, so there would be no magnetic induction. So, why do we believe that fast dynamo action is even a possibility for an MHD fluid? The answer is, of course, that an MHD fluid is a *non-rigid* body, and, thus, its motion possesses degrees of freedom not accessible to rigid conductors.

We know that in the perfectly conducting limit ($\eta \rightarrow 0$) magnetic field-lines are frozen into an MHD fluid. If the motion is incompressible (*i.e.*, $\nabla\cdot\mathbf{V} = 0$) then the stretching of field-lines implies a proportionate intensification of the field-strength. The simplest heuristic fast dynamo, first described by Vainshtein and Zel'dovich, is based on this effect. As illustrated in Fig. 5.8, a magnetic flux-tube can be *doubled* in intensity by taking it around a stretch-twist-fold cycle. The doubling time for this process clearly does not depend on the resistivity: in this sense, the dynamo is a fast dynamo. However, under repeated application of this cycle the magnetic field develops increasingly fine-scale structure. In fact, in the limit $\eta \rightarrow 0$ both the $\mathbf{V}$ and $\mathbf{B}$ fields eventually become chaotic and non-differentiable. A little resistivity is always required to smooth out the fields on small length-scales: even in this case the fields remain *chaotic*.

[9] S. Vainshtein, and Ya. B. Zel'dovich, Sov. Phys. Usp. **15**, 159 (1978).

At present, the physical existence of fast dynamos has not been conclusively established, since most of the literature on this subject is based on mathematical paradigms rather than actual solutions of the dynamo equation. It should be noted, however, that the need for fast dynamo solutions is fairly acute, especially in stellar dynamo theory. For instance, consider the Sun. The ohmic decay time for the Sun is about 10^{12} years, whereas the reversal time for the solar magnetic field is only 11 years. It is obviously a little difficult to believe that resistivity is playing any significant role in the solar dynamo.

In the following, we shall restrict our analysis to slow dynamos, which undoubtably exist in nature, and which are characterized by *non-chaotic* **V** and **B** fields.

5.12 Cowling Anti-Dynamo Theorem

One of the most important results in slow, kinematic dynamo theory is credited to Cowling.[10] The so-called *Cowling anti-dynamo theorem* states that:

> An axisymmetric magnetic field cannot be maintained via dynamo action.

Let us attempt to prove this proposition.

We adopt standard cylindrical polar coordinates: (ϖ, θ, z). The system is assumed to possess axial symmetry, so that $\partial/\partial\theta \equiv 0$. For the sake of simplicity, the plasma flow is assumed to be incompressible, which implies that $\nabla\cdot\mathbf{V} = 0$.

It is convenient to split the magnetic and velocity fields into *poloidal* and *toroidal* components:

$$\mathbf{B} = \mathbf{B}_p + \mathbf{B}_t, \tag{5.115}$$

$$\mathbf{V} = \mathbf{V}_p + \mathbf{V}_t. \tag{5.116}$$

Note that a poloidal vector only possesses non-zero ϖ- and z-components, whereas a toroidal vector only possesses a non-zero θ-component.

The poloidal components of the magnetic and velocity fields are written:

$$\mathbf{B}_p = \nabla\times\left(\frac{\psi}{\varpi}\,\hat{\boldsymbol{\theta}}\right) \equiv \frac{\nabla\psi\times\hat{\boldsymbol{\theta}}}{\varpi}, \tag{5.117}$$

$$\mathbf{V}_p = \nabla\times\left(\frac{\phi}{\varpi}\,\hat{\boldsymbol{\theta}}\right) \equiv \frac{\nabla\phi\times\hat{\boldsymbol{\theta}}}{\varpi}, \tag{5.118}$$

where $\psi = \psi(\varpi, z, t)$ and $\phi = \phi(\varpi, z, t)$. The toroidal components are given by

$$\mathbf{B}_t = B_t(\varpi, z, t)\,\hat{\boldsymbol{\theta}}, \tag{5.119}$$

$$\mathbf{V}_t = V_t(\varpi, z, t)\,\hat{\boldsymbol{\theta}}. \tag{5.120}$$

[10] T.G. Cowling, Mon. Not. Roy. Astr. Soc. **94**, 39 (1934); T.G. Cowling, Quart. J. Mech. Appl. Math. **10**, 129 (1957).

Note that by writing the **B** and **V** fields in the above form we ensure that the constraints $\nabla\cdot\mathbf{B} = 0$ and $\nabla\cdot\mathbf{V} = 0$ are *automatically* satisfied. Note, further, that since $\mathbf{B}\cdot\nabla\psi = 0$ and $\mathbf{V}\cdot\nabla\phi = 0$, we can regard ψ and ϕ as *stream-functions* for the magnetic and velocity fields, respectively.

The condition for the magnetic field to be maintained by dynamo currents, rather than by currents at infinity, is

$$\psi \rightarrow \frac{1}{r} \qquad \text{as } r \rightarrow \infty, \tag{5.121}$$

where $r = \sqrt{\varpi^2 + z^2}$. We also require the flow stream-function, ϕ, to remain bounded as $r \rightarrow \infty$.

Consider the MHD Ohm's law for a resistive plasma:

$$\mathbf{E} + \mathbf{V}\times\mathbf{B} = \eta\,\mathbf{j}. \tag{5.122}$$

Taking the toroidal component of this equation, we obtain

$$E_t + (\mathbf{V}_p\times\mathbf{B}_p)\cdot\hat{\boldsymbol{\theta}} = \eta\, j_t. \tag{5.123}$$

It is easily demonstrated that

$$E_t = -\frac{1}{\varpi}\frac{\partial\psi}{\partial t}. \tag{5.124}$$

Furthermore,

$$(\mathbf{V}_p\times\mathbf{B}_p)\cdot\hat{\boldsymbol{\theta}} = \frac{(\nabla\phi\times\nabla\psi)\cdot\hat{\boldsymbol{\theta}}}{\varpi^2} = \frac{1}{\varpi^2}\left(\frac{\partial\psi}{\partial\varpi}\frac{\partial\phi}{\partial z} - \frac{\partial\phi}{\partial\varpi}\frac{\partial\psi}{\partial z}\right), \tag{5.125}$$

and

$$\mu_0\, j_t = \nabla\times\mathbf{B}_p\cdot\hat{\boldsymbol{\theta}} = -\left[\nabla^2\left(\frac{\psi}{\varpi}\right) - \frac{\psi}{\varpi^3}\right] = -\frac{1}{\varpi}\left(\frac{\partial^2\psi}{\partial\varpi^2} - \frac{1}{\varpi}\frac{\partial\psi}{\partial\varpi} + \frac{\partial^2\psi}{\partial z^2}\right). \tag{5.126}$$

Thus, Eq. (5.123) reduces to

$$\frac{\partial\psi}{\partial t} - \frac{1}{\varpi}\left(\frac{\partial\psi}{\partial\varpi}\frac{\partial\phi}{\partial z} - \frac{\partial\phi}{\partial\varpi}\frac{\partial\psi}{\partial z}\right) = \frac{\eta}{\mu_0}\left(\frac{\partial^2\psi}{\partial\varpi^2} - \frac{1}{\varpi}\frac{\partial\psi}{\partial\varpi} + \frac{\partial^2\psi}{\partial z^2}\right). \tag{5.127}$$

Multiplying the above equation by ψ and integrating over all space, we obtain

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\int\psi^2\,dV \;&-\; \iint 2\pi\,\psi\left(\frac{\partial\psi}{\partial\varpi}\frac{\partial\phi}{\partial z} - \frac{\partial\phi}{\partial\varpi}\frac{\partial\psi}{\partial z}\right)d\varpi\,dz \\ &= \frac{\eta}{\mu_0}\iint 2\pi\,\varpi\,\psi\left(\frac{\partial^2\psi}{\partial\varpi^2} - \frac{1}{\varpi}\frac{\partial\psi}{\partial\varpi} + \frac{\partial^2\psi}{\partial z^2}\right)d\varpi\,dz.\end{aligned} \tag{5.128}$$

The second term on the left-hand side of the above expression can be integrated by parts to give

$$-\iint 2\pi\left[-\phi\,\frac{\partial}{\partial z}\left(\psi\,\frac{\partial\psi}{\partial\varpi}\right) + \phi\,\frac{\partial}{\partial\varpi}\left(\psi\,\frac{\partial\psi}{\partial z}\right)\right]d\varpi\,dz = 0, \tag{5.129}$$

where surface terms have been neglected, in accordance with Eq. (5.121). Likewise, the term on the right-hand side of Eq. (5.128) can be integrated by parts to give

$$\frac{\eta}{\mu_0}\iint 2\pi\left[-\frac{\partial(\varpi\,\psi)}{\partial\varpi}\,\frac{\partial\psi}{\partial\varpi}-\varpi\left(\frac{\partial\psi}{\partial z}\right)^2\right]\,d\varpi\,dz=$$
$$-\frac{\eta}{\mu_0}\iint 2\pi\,\varpi\left[\left(\frac{\partial\psi}{\partial\varpi}\right)^2+\left(\frac{\partial\psi}{\partial z}\right)^2\right]\,d\varpi\,dz. \tag{5.130}$$

Thus, Eq. (5.128) reduces to

$$\frac{d}{dt}\int\psi^2\,dV=-2\,\frac{\eta}{\mu_0}\int|\nabla\psi|^2\,dV. \tag{5.131}$$

It is clear from the above expression that the poloidal stream-function, ψ, and, hence, the poloidal magnetic field, $\mathbf{B}_p$, decays to zero under the influence of resistivity. We conclude that the poloidal magnetic field *cannot* be maintained via dynamo action.

Of course, we have not ruled out the possibility that the toroidal magnetic field can be maintained via dynamo action. In the absence of a poloidal field, the curl of the poloidal component of Eq. (5.122) yields

$$-\frac{\partial\mathbf{B}_t}{\partial t}+\nabla\times(\mathbf{V}_p\times\mathbf{B}_t)=\eta\,\nabla\times\mathbf{j}_p, \tag{5.132}$$

which reduces to

$$-\frac{\partial B_t}{\partial t}+\nabla\times(\mathbf{V}_p\times\mathbf{B}_t)\cdot\hat{\theta}=-\frac{\eta}{\mu_0}\,\nabla^2(B_t\,\hat{\theta})\cdot\hat{\theta}. \tag{5.133}$$

Now

$$\nabla^2(B_t\,\hat{\theta})\cdot\hat{\theta}=\frac{\partial^2 B_t}{\partial\varpi^2}+\frac{1}{\varpi}\frac{\partial B_t}{\partial\varpi}+\frac{\partial^2 B_t}{\partial z^2}-\frac{B_t}{\varpi^2}, \tag{5.134}$$

and

$$\nabla\times(\mathbf{V}_p\times\mathbf{B}_t)\cdot\hat{\theta}=\frac{\partial}{\partial\varpi}\left(\frac{B_t}{\varpi}\right)\frac{\partial\phi}{\partial z}-\frac{\partial}{\partial z}\left(\frac{B_t}{\varpi}\right)\frac{\partial\phi}{\partial\varpi}. \tag{5.135}$$

Thus, Eq. (5.133) yields

$$\frac{\partial\chi}{\partial t}-\frac{1}{\varpi}\left(\frac{\partial\chi}{\partial\varpi}\frac{\partial\phi}{\partial z}-\frac{\partial\phi}{\partial\varpi}\frac{\partial\chi}{\partial z}\right)=\frac{\eta}{\mu_0}\left(\frac{\partial^2\chi}{\partial\varpi^2}+\frac{3}{\varpi}\frac{\partial\chi}{\partial\varpi}+\frac{\partial^2\chi}{\partial z^2}\right), \tag{5.136}$$

where

$$B_t=\varpi\,\chi. \tag{5.137}$$

Multiply Eq. (5.136) by χ, integrating over all space, and then integrating by parts, we obtain

$$\frac{d}{dt}\int\chi^2\,dV=-2\,\frac{\eta}{\mu_0}\int|\nabla\chi|^2\,dV. \tag{5.138}$$

It is clear from this formula that χ, and, hence, the toroidal magnetic field, $\mathbf{B}_t$, decay to zero under the influence of resistivity. We conclude that no axisymmetric magnetic field, either poloidal or toroidal, can be maintained by dynamo action, which proves Cowling's theorem.

Cowling's theorem is the earliest and most significant of a number of *anti-dynamo theorems* which severely restrict the types of magnetic fields which can be maintained via dynamo action. For instance, it is possible to prove that a two-dimensional magnetic field cannot be maintained by dynamo action. Here, "two-dimensional" implies that in some Cartesian coordinate system, (x, y, z), the magnetic field is independent of z. The suite of anti-dynamo theorems can be summed up by saying that successful dynamos possess a rather low degree of symmetry.

5.13 Ponomarenko Dynamos

The simplest known kinematic dynamo is that of Ponomarenko.[11] Consider a conducting fluid of resistivity η which fills all space. The motion of the fluid is confined to a cylinder of radius a. Adopting cylindrical polar coordinates (r, θ, z) aligned with this cylinder, the flow field is written

$$\mathbf{V} = \begin{cases} (0, r\,\Omega, U) & \text{for } r \leq a \\ \mathbf{0} & \text{for } r > a \end{cases}, \tag{5.139}$$

where Ω and U are constants. Note that the flow is *incompressible*: *i.e.*, $\nabla \cdot \mathbf{V} = 0$.

The dynamo equation can be written

$$\frac{\partial \mathbf{B}}{\partial t} = (\mathbf{B}\cdot\nabla)\mathbf{V} - (\mathbf{V}\cdot\nabla)\mathbf{B} + \frac{\eta}{\mu_0}\nabla^2\mathbf{B}. \tag{5.140}$$

Let us search for solutions to this equation of the form

$$\mathbf{B}(r, \theta, z, t) = \mathbf{B}(r)\, \exp[i\,(m\,\theta - k\,z) + \gamma\, t]. \tag{5.141}$$

The r- and θ- components of Eq. (5.140) are written

$$\begin{aligned} \gamma\, B_r &= -i\,(m\,\Omega - k\,U)\, B_r \\ &\quad + \frac{\eta}{\mu_0}\left[\frac{d^2 B_r}{dr^2} + \frac{1}{r}\frac{dB_r}{dr} - \frac{(m^2 + k^2 r^2 + 1)\, B_r}{r^2} - \frac{i\,2\,m\, B_\theta}{r^2}\right], \end{aligned} \tag{5.142}$$

and

$$\begin{aligned} \gamma\, B_\theta &= r\,\frac{d\Omega}{dr}\, B_r - i\,(m\,\Omega - k\,U)\, B_\theta \\ &\quad + \frac{\eta}{\mu_0}\left[\frac{d^2 B_\theta}{dr^2} + \frac{1}{r}\frac{dB_\theta}{dr} - \frac{(m^2 + k^2 r^2 + 1)\, B_\theta}{r^2} + \frac{i\,2\,m\, B_r}{r^2}\right], \end{aligned} \tag{5.143}$$

[11] Yu. B. Ponomarenko, J. Appl. Mech. Tech. Phys. **14**, 775 (1973).

respectively. In general, the term involving $d\Omega/dr$ is zero. In fact, this term is only included in the analysis to enable us to evaluate the correct matching conditions at $r = a$. Note that we do not need to write the z-component of Eq. (5.140), since B_z can be obtained more directly from B_r and B_θ via the constraint $\nabla\cdot\mathbf{B} = 0$.

Let

$$B_\pm = B_r \pm i\, B_\theta, \tag{5.144}$$

$$y = \frac{r}{a}, \tag{5.145}$$

$$\tau_R = \frac{\mu_0\, a^2}{\eta}, \tag{5.146}$$

$$q^2 = k^2 a^2 + \gamma\tau_R + i\,(m\,\Omega - k\,U)\,\tau_R, \tag{5.147}$$

$$s^2 = k^2 a^2 + \gamma\tau_R. \tag{5.148}$$

Here, τ_R is the typical time for magnetic flux to diffuse a distance a under the action of resistivity. Equations (5.142)–(5.148) can be combined to give

$$y^2\,\frac{d^2B_\pm}{dy^2} + y\,\frac{dB_\pm}{dy} - \left[(m\pm1)^2 + q^2\,y^2\right] B_\pm = 0 \tag{5.149}$$

for $y \leq 1$, and

$$y^2\,\frac{d^2B_\pm}{dy^2} + y\,\frac{dB_\pm}{dy} - \left[(m\pm1)^2 + s^2\,y^2\right] B_\pm = 0 \tag{5.150}$$

for $y > 1$. The above equations are immediately recognized as modified Bessel's equations of order $m \pm 1$.[12] Thus, the physical solutions of Eqs. (5.149) and (5.150), which are well behaved as $y \rightarrow 0$ and $y \rightarrow \infty$, can be written

$$B_\pm = C_\pm\,\frac{I_{m\pm1}(q\,y)}{I_{m\pm1}(q)} \tag{5.151}$$

for $y \leq 1$, and

$$B_\pm = D_\pm\,\frac{K_{m\pm1}(s\,y)}{K_{m\pm1}(s)} \tag{5.152}$$

for $y > 1$. Here, $C_\pm$ and $D_\pm$ are arbitrary constants. Note that the arguments of q and s are both constrained to lie in the range $-\pi/2$ to $+\pi/2$.

The first set of matching conditions at $y = 1$ are, obviously, that $B_\pm$ are continuous, which yields

$$C_\pm = D_\pm. \tag{5.153}$$

[12] M. Abramowitz, and I.A. Stegun, *Handbook of Mathematical Functions* (Dover, New York NY, 1964), p. 374.

The second set of matching conditions are obtained by integrating Eq. (5.143) from $r = a - \delta$ to $r = a - \delta$, where δ is an infinitesimal quantity, and making use of the fact that the angular velocity Ω jumps discontinuously to zero at $r = a$. It follows that

$$a\,\Omega\,B_r = \frac{\eta}{\mu_0}\left[\frac{dB_\theta}{dr}\right]_{r=a_-}^{r=a_+}. \tag{5.154}$$

Furthermore, integration of Eq. (5.142) tells us that dB_r/dr is continuous at $r = a$. We can combine this information to give the matching condition

$$\left[\frac{dB_\pm}{dy}\right]_{y=1_-}^{y=1_+} = \pm i\,\Omega\,\tau_R\,\frac{B_+ + B_-}{2}. \tag{5.155}$$

Equations (5.151)–(5.155) can be combined to give the dispersion relation

$$G_+\,G_- = \frac{i}{2}\,\Omega\,\tau_R\,(G_+ - G_-), \tag{5.156}$$

where

$$G_\pm = q\,\frac{I'_{m\pm1}(q)}{I_{m\pm1}(q)} - s\,\frac{K'_{m\pm1}(s)}{K_{m\pm1}(s)}. \tag{5.157}$$

Here, $'$ denotes a derivative.

Unfortunately, despite the fact that we are investigating the simplest known dynamo, the dispersion relation (5.156) is sufficiently complicated that it can only be solved numerically. We can simplify matters considerably taking the limit $|q|, |s| \gg 1$, which corresponds either to that of small wave-length (*i.e.*, $k\,a \gg 1$), or small resistivity (*i.e.*, $\Omega\,\tau_R \gg 1$). The large argument asymptotic behaviour of the Bessel functions is specified by [13]

$$\sqrt{\frac{2\,z}{\pi}}\,K_m(z) = e^{-z}\left(1 + \frac{4\,m^2 - 1}{8\,z} + \cdots\right), \tag{5.158}$$

$$\sqrt{2\,z\,\pi}\,I_m(z) = e^{+z}\left(1 - \frac{4\,m^2 - 1}{8\,z} + \cdots\right), \tag{5.159}$$

where $|\arg(z)| < \pi/2$. It follows that

$$G_\pm = q + s + (m^2/2 \pm m + 3/8)(q^{-1} + s^{-1}) + O(q^{-2} + s^{-2}). \tag{5.160}$$

Thus, the dispersion relation (5.156) reduces to

$$(q + s)\,q\,s = i\,m\,\Omega\,\tau_R, \tag{5.161}$$

where $|\arg(q)|, |\arg(s)| < \pi/2$.

[13] M. Abramowitz, and I.A. Stegun, *Handbook of Mathematical Functions* (Dover, New York NY, 1964), p. 377.

In the limit $\mu \rightarrow 0$, where

$$\mu = (m\,\Omega - k\,U)\,\tau_R, \tag{5.162}$$

which corresponds to $(\mathbf{V}\cdot\nabla)\mathbf{B} \rightarrow 0$, the simplified dispersion relation (5.161) can be solved to give

$$\gamma\,\tau_R \simeq e^{\,i\pi/3}\left(\frac{m\,\Omega\,\tau_R}{2}\right)^{2/3} - k^2\,a^2 - i\,\frac{\mu}{2}. \tag{5.163}$$

Dynamo behaviour [*i.e.*, $\mathrm{Re}(\gamma) > 0$] takes place when

$$\Omega\,\tau_R > \frac{2^{5/2}\,(ka)^3}{m}. \tag{5.164}$$

Note that $\mathrm{Im}(\gamma) \neq 0$, implying that the dynamo mode *oscillates*, or rotates, as well as growing exponentially in time. The dynamo generated magnetic field is both non-axisymmetric [note that dynamo activity is impossible, according to Eq. (5.163), if $m = 0$] and three-dimensional, and is, thus, not subject to either of the anti-dynamo theorems mentioned in the preceding section.

It is clear from Eq. (5.164) that dynamo action occurs whenever the flow is made sufficiently rapid. But, what is the minimum amount of flow which gives rise to dynamo action? In order to answer this question we have to solve the full dispersion relation, (5.156), for various values of m and k in order to find the dynamo mode which grows exponentially in time for the smallest values of Ω and U. It is conventional to parameterize the flow in terms of the *magnetic Reynolds number*

$$S = \frac{\tau_R}{\tau_H}, \tag{5.165}$$

where

$$\tau_H = \frac{L}{V} \tag{5.166}$$

is the typical time-scale for convective motion across the system. Here, V is a typical flow velocity, and L is the scale-length of the system. Taking $V = |\mathbf{V}(a)| = \sqrt{\Omega^2\,a^2 + U^2}$, and $L = a$, we have

$$S = \frac{\tau_R\,\sqrt{\Omega^2\,a^2 + U^2}}{a} \tag{5.167}$$

for the Ponomarenko dynamo. The critical value of the Reynolds number above which dynamo action occurs is found to be

$$S_c = 17.7. \tag{5.168}$$

The most unstable dynamo mode is characterized by $m = 1$, $U/\Omega\,a = 1.3$, $k\,a = 0.39$, and $\mathrm{Im}(\gamma)\,\tau_R = 0.41$. As the magnetic Reynolds number, S, is increased above the critical value, S_c, other dynamo modes are eventually destabilized.

Interestingly enough, an attempt was made in the late 1980's to construct a Ponomarenko dynamo by rapidly pumping liquid sodium through a cylindrical pipe equipped

with a set of twisted vanes at one end to induce helical flow. Unfortunately, the experiment failed due to mechanical vibrations, after achieving a Reynolds number which was 80% of the critical value required for self-excitation of the magnetic field, and was not repaired due to budgetary problems.[14] More recently, there has been renewed interest worldwide in the idea of constructing a liquid metal dynamo, and two such experiments (one in Riga, and one in Karlsruhe) have demonstrated self-excited dynamo action in a controlled laboratory setting.

5.14 Magnetic Reconnection

Magnetic reconnection is a phenomenon which is of particular importance in solar system plasmas. In the solar corona, it results in the rapid release to the plasma of energy stored in the large-scale structure of the coronal magnetic field, an effect which is thought to give rise to *solar flares*. Small-scale reconnection may play a role in heating the corona, and, thereby, driving the outflow of the solar wind. In the Earth's magnetosphere, magnetic reconnection in the magnetotail is thought to be the precursor for *auroral sub-storms*.

The evolution of the magnetic field in a resistive-MHD plasma is governed by the following well-known equation:

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{V} \times \mathbf{B}) + \frac{\eta}{\mu_0} \nabla^2 \mathbf{B}. \tag{5.169}$$

The first term on the right-hand side of this equation describes the *convection* of the magnetic field by the plasma flow. The second term describes the resistive *diffusion* of the field through the plasma. If the first term dominates then magnetic flux is frozen into the plasma, and the topology of the magnetic field cannot change. On the other hand, if the second term dominates then there is little coupling between the field and the plasma flow, and the topology of the magnetic field is free to change.

The relative magnitude of the two terms on the right-hand side of Eq. (5.169) is conventionally measured in terms of *magnetic Reynolds number*, or *Lundquist number*:

$$S = \frac{\mu_0 V L}{\eta} \simeq \frac{|\nabla \times (\mathbf{V} \times \mathbf{B})|}{|(\eta/\mu_0) \nabla^2 \mathbf{B}|}, \tag{5.170}$$

where V is the characteristic flow speed, and L the characteristic length-scale of the plasma. If S is much larger than unity then convection dominates, and the *frozen flux* constraint prevails, whilst if S is much less than unity then diffusion dominates, and the coupling between the plasma flow and the magnetic field is relatively weak.

It turns out that in the solar system very large S-values are virtually guaranteed by the the extremely large scale-lengths of solar system plasmas. For instance, $S \sim 10^8$ for solar flares, whilst $S \sim 10^{11}$ is appropriate for the solar wind and the Earth's magnetosphere. Of

[14] A. Gailitis, *Topological Fluid Dynamics*, edited by H.K. Moffatt, and A. Tsinober (Cambridge University Press, Cambridge UK, 1990), p. 147.

course, in calculating these values we have identified the scale-length L with the overall size of the plasma under investigation.

On the basis of the above discussion, it seems reasonable to neglect diffusive processes altogether in solar system plasmas. Of course, this leads to very strong constraints on the behaviour of such plasmas, since all cross-field mixing of plasma elements is suppressed in this limit. Particles may freely mix along field-lines (within limitations imposed by magnetic mirroring, *etc.*), but are completely ordered perpendicular to the field, since they always remain tied to the same field-lines as they convect in the plasma flow.

Let us consider what happens when two initially separate plasma regions come into contact with one another, as occurs, for example, in the interaction between the solar wind and the Earth's magnetic field. Assuming that each plasma is frozen to its own magnetic field, and that cross-field diffusion is absent, we conclude that the two plasmas will not mix, but, instead, that a thin *boundary layer* will form between them, separating the two plasmas and their respective magnetic fields. In equilibrium, the location of the boundary layer will be determined by pressure balance. Since, in general, the frozen fields on either side of the boundary will have differing strengths, and orientations tangential to the boundary, the layer must also constitute a *current sheet*. Thus, flux freezing leads inevitably to the prediction that in plasma systems space becomes divided into separate cells, wholly containing the plasma and magnetic field from individual sources, and separated from each other by thin current sheets.

The "separate cell" picture constitutes an excellent zeroth-order approximation to the interaction of solar system plasmas, as witnessed, for example, by the well defined planetary magnetospheres. It must be noted, however, that the large S-values upon which the applicability of the frozen flux constraint was justified were derived using the *large* overall spatial scales of the systems involved. However, strict application of this constraint to the problem of the interaction of separate plasma systems leads to the inevitable conclusion that structures will form having *small* spatial scales, at least in one dimension: *i.e.*, the thin current sheets constituting the cell boundaries. It is certainly not guaranteed, from the above discussion, that the effects of diffusion can be neglected in these boundary layers. In fact, we shall demonstrate that the localized breakdown of the flux freezing constraint in the boundary regions, due to diffusion, not only has an impact on the properties of the boundary regions themselves, but can also have a decisive impact on the large length-scale plasma regions where the flux freezing constraint remains valid. This observation illustrates both the subtlety and the significance of the magnetic reconnection process.

5.15 Linear Tearing Mode Theory

Consider the interface between two plasmas containing magnetic fields of different orientations. The simplest imaginable field configuration is that illustrated in Fig. 5.9. Here, the field varies only in the x-direction, and points only in the y-direction. The field is directed in the $-y$-direction for $x < 0$, and in the $+y$-direction for $x > 0$. The interface is situated at $x = 0$. The sudden reversal of the field direction across the interface gives rise to a

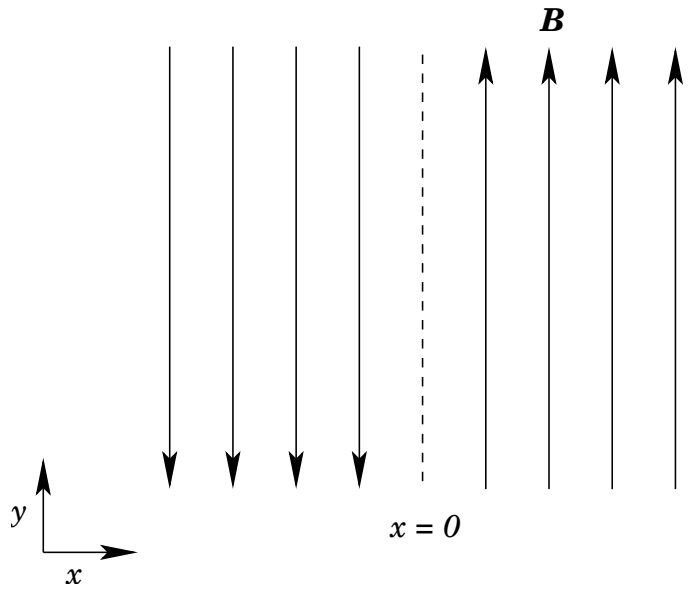

Figure 5.9: *A reconnecting magnetic field configuration.*

z-directed current sheet at $x = 0$.

With the neglect of plasma resistivity, the field configuration shown in Fig. 5.9 represents a *stable* equilibrium state, assuming, of course, that we have normal pressure balance across the interface. But, does the field configuration remain stable when we take resistivity into account? If not, we expect an instability to develop which relaxes the configuration to one possessing lower magnetic energy. As we shall see, this type of relaxation process inevitably entails the breaking and reconnection of magnetic field lines, and is, therefore, termed *magnetic reconnection*. The magnetic energy released during the reconnection process eventually appears as plasma thermal energy. Thus, magnetic reconnection also involves plasma heating.

In the following, we shall outline the standard method for determining the *linear* stability of the type of magnetic field configuration shown in Fig. 26, taking into account the effect of plasma resistivity. We are particularly interested in plasma instabilities which are stable in the absence of resistivity, and only grow when the resistivity is non-zero. Such instabilities are conventionally termed *tearing modes*. Since magnetic reconnection is, in fact, a *nonlinear* process, we shall then proceed to investigate the nonlinear development of tearing modes.

The equilibrium magnetic field is written

$$\mathbf{B}_0 = B_{0y}(x)\,\hat{\mathbf{y}}, \tag{5.171}$$

where $B_{0y}(-x) = -B_{0y}(x)$. There is assumed to be no equilibrium plasma flow. The linearized equations of resistive-MHD, assuming incompressible flow, take the form

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{V} \times \mathbf{B}_0) + \frac{\eta}{\mu_0}\nabla^2 \mathbf{B}, \tag{5.172}$$

$$\rho_0 \frac{\partial \mathbf{V}}{\partial t} = -\nabla p + \frac{(\nabla \times \mathbf{B}) \times \mathbf{B}_0}{\mu_0} + \frac{(\nabla \times \mathbf{B}_0) \times \mathbf{B}}{\mu_0} \tag{5.173}$$

$$\nabla \cdot \mathbf{B} = 0, \tag{5.174}$$

$$\nabla \cdot \mathbf{V} = 0. \tag{5.175}$$

Here, ρ_0 is the equilibrium plasma density, $\mathbf{B}$ the perturbed magnetic field, $\mathbf{V}$ the perturbed plasma velocity, and p the perturbed plasma pressure. The assumption of incompressible plasma flow is valid provided that the plasma velocity associated with the instability remains significantly smaller than both the Alfvén velocity and the sonic velocity.

Suppose that all perturbed quantities vary like

$$A(x, y, z, t) = A(x)\, e^{i k y + \gamma t}, \tag{5.176}$$

where γ is the instability growth-rate. The x-component of Eq. (5.172) and the z-component of the curl of Eq. (5.173) reduce to

$$\gamma B_x = i k B_{0y} V_x + \frac{\eta}{\mu_0} \left(\frac{d^2}{dx^2} - k^2 \right) B_x, \tag{5.177}$$

$$\gamma \rho_0 \left(\frac{d^2}{dx^2} - k^2 \right) V_x = \frac{i k B_{0y}}{\mu_0} \left(\frac{d^2}{dx^2} - k^2 - \frac{B_{0y}''}{B_{0y}} \right) B_x, \tag{5.178}$$

respectively, where use has been made of Eqs. (5.174) and (5.175). Here, $'$ denotes d/dx.

It is convenient to normalize Eqs. (5.177)–(5.178) using a typical magnetic field-strength, B_0, and a typical scale-length, a. Let us define the *Alfvén time-scale*

$$\tau_A = \frac{a}{V_A}, \tag{5.179}$$

where $V_A = B_0/\sqrt{\mu_0 \rho_0}$ is the Alfvén velocity, and the *resistive diffusion* time-scale

$$\tau_R = \frac{\mu_0 a^2}{\eta}. \tag{5.180}$$

The ratio of these two time-scales is the Lundquist number:

$$S = \frac{\tau_R}{\tau_A}. \tag{5.181}$$

Let $\psi = B_x/B_0$, $\phi = i k V_y/\gamma$, $\bar{x} = x/a$, $F = B_{0y}/B_0$, $F' \equiv dF/d\bar{x}$, $\bar{\gamma} = \gamma \tau_A$, and $\bar{k} = k a$. It follows that

$$\bar{\gamma} (\psi - F \phi) = S^{-1} \left(\frac{d^2}{d\bar{x}^2} - \bar{k}^2 \right) \psi, \tag{5.182}$$

$$\bar{\gamma}^2 \left(\frac{d^2}{d\bar{x}^2} - \bar{k}^2 \right) \phi = -\bar{k}^2 F \left(\frac{d^2}{d\bar{x}^2} - \bar{k}^2 - \frac{F''}{F} \right) \psi. \tag{5.183}$$

The term on the right-hand side of Eq. (5.182) represents plasma *resistivity*, whilst the term on the left-hand side of Eq. (5.183) represents plasma *inertia*.

It is assumed that the tearing instability grows on a *hybrid* time-scale which is much less than τ_R but much greater than τ_A. It follows that

$$\bar{\gamma} \ll 1 \ll S\,\bar{\gamma}. \tag{5.184}$$

Thus, throughout most of the plasma we can neglect the right-hand side of Eq. (5.182) and the left-hand side of Eq. (5.183), which is equivalent to the neglect of plasma resistivity and inertia. In this case, Eqs. (5.182)–(5.183) reduce to

$$\phi = \frac{\psi}{F}, \tag{5.185}$$

$$\frac{d^2\psi}{d\bar{x}^2} - \bar{k}^2\,\psi - \frac{F''}{F}\,\psi = 0. \tag{5.186}$$

Equation (5.185) is simply the flux freezing constraint, which requires the plasma to move with the magnetic field. Equation (5.186) is the linearized, static force balance criterion: $\nabla\times(\mathbf{j}\times\mathbf{B}) = \mathbf{0}$. Equations (5.185)–(5.186) are known collectively as the equations of *ideal-MHD*, and are valid throughout virtually the whole plasma. However, it is clear that these equations *break down* in the immediate vicinity of the interface, where $F = 0$ (*i.e.*, where the magnetic field reverses direction). Witness, for instance, the fact that the normalized "radial" velocity, ϕ, becomes infinite as $F \rightarrow 0$, according to Eq. (5.185).

The ideal-MHD equations break down close to the interface because the neglect of plasma resistivity and inertia becomes untenable as $F \rightarrow 0$. Thus, there is a thin layer, in the immediate vicinity of the interface, $\bar{x} = 0$, where the behaviour of the plasma is governed by the full MHD equations, (5.182)–(5.183). We can simplify these equations, making use of the fact that $\bar{x} \ll 1$ and $d/d\bar{x} \gg 1$ in a thin layer, to obtain the following layer equations:

$$\bar{\gamma}\,(\psi - \bar{x}\,\phi) = S^{-1}\frac{d^2\psi}{d\bar{x}^2}, \tag{5.187}$$

$$\bar{\gamma}^2\,\frac{d^2\phi}{d\bar{x}^2} = -\bar{x}\,\frac{d^2\psi}{d\bar{x}^2}. \tag{5.188}$$

Note that we have redefined the variables ϕ, $\bar{\gamma}$, and S, such that $\phi \rightarrow F'(0)\,\phi$, $\bar{\gamma} \rightarrow \gamma\,\tau_H$, and $S \rightarrow \tau_R/\tau_H$. Here,

$$\tau_H = \frac{\tau_A}{k\,a\,F'(0)} \tag{5.189}$$

is the *hydromagnetic time-scale*.

The tearing mode stability problem reduces to solving the non-ideal-MHD layer equations, (5.187)–(5.188), in the immediate vicinity of the interface, $\bar{x} = 0$, solving the ideal-MHD equations, (5.185)–(5.186), everywhere else in the plasma, matching the two solutions at the edge of the layer, and applying physical boundary conditions as $|\bar{x}| \rightarrow \infty$.

This method of solution was first described in a classic paper by Furth, Killeen, and Rosenbluth.[15]

Let us consider the solution of the ideal-MHD equation (5.186) throughout the bulk of the plasma. We could imagine launching a solution $\psi(\bar{x})$ at large positive $\bar{x}$, which satisfies physical boundary conditions as $\bar{x} \rightarrow \infty$, and integrating this solution to the right-hand boundary of the non-ideal-MHD layer at $\bar{x} = 0_+$. Likewise, we could also launch a solution at large negative $\bar{x}$, which satisfies physical boundary conditions as $\bar{x} \rightarrow -\infty$, and integrate this solution to the left-hand boundary of the non-ideal-MHD layer at $\bar{x} = 0_-$. Maxwell's equations demand that ψ must be continuous on either side of the layer. Hence, we can multiply our two solutions by appropriate factors, so as to ensure that ψ matches to the left and right of the layer. This leaves the function $\psi(\bar{x})$ undetermined to an overall arbitrary multiplicative constant, just as we would expect in a linear problem. In general, $d\psi/d\bar{x}$ is *not* continuous to the left and right of the layer. Thus, the ideal solution can be characterized by the real number

$$\Delta' = \left[\frac{1}{\psi}\frac{d\psi}{d\bar{x}}\right]_{\bar{x}=0_-}^{\bar{x}=0_+}: \tag{5.190}$$

i.e., by the jump in the logarithmic derivative of ψ to the left and right of the layer. This parameter is known as the *tearing stability index*, and is solely a property of the plasma equilibrium, the wave-number, k, and the boundary conditions imposed at infinity.

The layer equations (5.187)–(5.188) possess a trivial solution ($\phi = \phi_0$, $\psi = \bar{x}\,\phi_0$, where ϕ_0 is independent of $\bar{x}$), and a nontrivial solution for which $\psi(-\bar{x}) = \psi(\bar{x})$ and $\phi(-\bar{x}) = -\phi(\bar{x})$. The asymptotic behaviour of the nontrivial solution at the edge of the layer is

$$\psi(x) \rightarrow \left(\frac{\Delta}{2}|\bar{x}| + 1\right)\Psi, \tag{5.191}$$

$$\phi(x) \rightarrow \frac{\psi}{\bar{x}}, \tag{5.192}$$

where the parameter $\Delta(\bar{\gamma}, S)$ is determined by solving the layer equations, subject to the above boundary conditions. Finally, the growth-rate, γ, of the tearing instability is determined by the matching criterion

$$\Delta(\bar{\gamma}, S) = \Delta'. \tag{5.193}$$

The layer equations (5.187)–(5.188) can be solved in a fairly straightforward manner in Fourier transform space. Let

$$\phi(\bar{x}) = \int_{-\infty}^{\infty} \hat{\phi}(t)\, e^{\,i S^{1/3}\bar{x}\,t}\, dt, \tag{5.194}$$

$$\psi(\bar{x}) = \int_{-\infty}^{\infty} \hat{\psi}(t)\, e^{\,i S^{1/3}\bar{x}\,t}\, dt, \tag{5.195}$$

[15] H.P. Furth, J. Killeen, and M.N. Rosenbluth, Phys. Fluids **6**, 459 (1963).

where $\hat{\phi}(-t) = -\hat{\phi}(t)$. Equations (5.187)–(5.188) can be Fourier transformed, and the results combined, to give

$$\frac{d}{dt}\left(\frac{t^2}{Q+t^2}\frac{d\hat{\phi}}{dt}\right) - Q\,t^2\,\hat{\phi} = 0, \tag{5.196}$$

where

$$Q = \gamma\,\tau_H^{2/3}\,\tau_R^{1/3}. \tag{5.197}$$

The most general small-t asymptotic solution of Eq. (5.196) is written

$$\hat{\phi}(t) \rightarrow \frac{a_{-1}}{t} + a_0 + O(t), \tag{5.198}$$

where a_{-1} and a_0 are independent of t, and it is assumed that $t > 0$. When inverse Fourier transformed, the above expression leads to the following expression for the asymptotic behaviour of ϕ at the edge of the non-ideal-MHD layer:

$$\phi(\bar{x}) \rightarrow a_{-1}\,\frac{\pi}{2}\,S^{1/3}\,\mathrm{sgn}(x) + \frac{a_0}{\bar{x}} + O(|\bar{x}|^{-2}). \tag{5.199}$$

It follows from a comparison with Eqs. (5.191)–(5.192) that

$$\Delta = \pi\,\frac{a_{-1}}{a_0}\,S^{1/3}. \tag{5.200}$$

Thus, the matching parameter Δ is determined from the small-t asymptotic behaviour of the Fourier transformed layer solution.

Let us search for an unstable tearing mode, characterized by $Q > 0$. It is convenient to assume that

$$Q \ll 1. \tag{5.201}$$

This ordering, which is known as the *constant-ψ approximation* [since it implies that $\psi(\bar{x})$ is approximately constant across the layer] will be justified later on.

In the limit $t \gg Q^{1/2}$, Eq. (5.196) reduces to

$$\frac{d^2\hat{\phi}}{dt^2} - Q\,t^2\,\hat{\phi} = 0. \tag{5.202}$$

The solution to this equation which is well behaved in the limit $t \rightarrow \infty$ is written $U(0, \sqrt{2}\,Q^{1/4}\,t)$, where $U(a,x)$ is a standard parabolic cylinder function.[16] In the limit

$$Q^{1/2} \ll t \ll Q^{-1/4} \tag{5.203}$$

we can make use of the standard small argument asymptotic expansion of $U(a,x)$ to write the most general solution to Eq. (5.196) in the form

$$\hat{\phi}(t) = A\left[1 - 2\,\frac{\Gamma(3/4)}{\Gamma(1/4)}\,Q^{1/4}\,t + O(t^2)\right]. \tag{5.204}$$

[16] M. Abramowitz, and I.A. Stegun, *Handbook of Mathematical Functions* (Dover, New York NY, 1964), p. 686.

Here, A is an arbitrary constant.

In the limit

$$t \ll Q^{-1/4}, \tag{5.205}$$

Eq. (5.196) reduces to

$$\frac{d}{dt}\left(\frac{t^2}{Q+t^2}\,\frac{d\hat{\phi}}{dt}\right) = 0. \tag{5.206}$$

The most general solution to this equation is written

$$\hat{\phi}(t) = B\left(-\frac{Q}{t}+t\right) + C + O(t^2), \tag{5.207}$$

where B and C are arbitrary constants. Matching coefficients between Eqs. (5.204) and (5.207) in the range of t satisfying the inequality (5.203) yields the following expression for the most general solution to Eq. (5.196) in the limit $t \ll Q^{1/2}$:

$$\hat{\phi} = A\,\left[2\,\frac{\Gamma(3/4)}{\Gamma(1/4)}\,\frac{Q^{5/4}}{t} + 1 + O(t)\right]. \tag{5.208}$$

Finally, a comparison of Eqs. (5.198), (5.200), and (5.208) yields the result

$$\Delta = 2\pi\,\frac{\Gamma(3/4)}{\Gamma(1/4)}\,S^{1/3}\,Q^{5/4}. \tag{5.209}$$

The asymptotic matching condition (5.193) can be combined with the above expression for Δ to give the tearing mode dispersion relation

$$\gamma = \left[\frac{\Gamma(1/4)}{2\pi\,\Gamma(3/4)}\right]^{4/5}\,\frac{(\Delta')^{4/5}}{\tau_H^{2/5}\,\tau_R^{3/5}}. \tag{5.210}$$

Here, use has been made of the definitions of S and Q. According to the above dispersion relation, the tearing mode is unstable whenever $\Delta' > 0$, and grows on the hybrid time-scale $\tau_H^{2/5}\,\tau_R^{3/5}$. It is easily demonstrated that the tearing mode is stable whenever $\Delta' < 0$. According to Eqs. (5.193), (5.201), and (5.209), the constant-ψ approximation holds provided that

$$\Delta' \ll S^{1/3}: \tag{5.211}$$

i.e., provided that the tearing mode does not become too unstable.

From Eq. (5.202), the thickness of the non-ideal-MHD layer in t-space is

$$\delta_t \sim \frac{1}{Q^{1/4}}. \tag{5.212}$$

It follows from Eqs. (5.194)–(5.195) that the thickness of the layer in $\bar{x}$-space is

$$\bar{\delta} \sim \frac{1}{S^{1/3}\,\delta_t} \sim \left(\frac{\bar{\gamma}}{S}\right)^{1/4}. \tag{5.213}$$

When $\Delta' \sim 0(1)$ then $\bar{\gamma} \sim S^{-3/5}$, according to Eq. (5.210), giving $\bar{\delta} \sim S^{-2/5}$. It is clear, therefore, that if the Lundquist number, S, is very large then the non-ideal-MHD layer centred on the interface, $\bar{x} = 0$, is *extremely narrow*.

The time-scale for magnetic flux to diffuse across a layer of thickness $\bar{\delta}$ (in $\bar{x}$-space) is [*cf.*, Eq. (5.180)]

$$\tau \sim \tau_R \, \bar{\delta}^{\,2}. \tag{5.214}$$

If

$$\gamma \, \tau \ll 1, \tag{5.215}$$

then the tearing mode grows on a time-scale which is far longer than the time-scale on which magnetic flux diffuses across the non-ideal layer. In this case, we would expect the normalized "radial" magnetic field, ψ, to be approximately *constant* across the layer, since any non-uniformities in ψ would be smoothed out via resistive diffusion. It follows from Eqs. (5.213) and (5.214) that the constant-ψ approximation holds provided that

$$\bar{\gamma} \ll S^{-1/3} \tag{5.216}$$

(*i.e.*, $Q \ll 1$), which is in agreement with Eq. (5.201).

5.16 Nonlinear Tearing Mode Theory

We have seen that if $\Delta' > 0$ then a magnetic field configuration of the type shown in Fig. 5.9 is unstable to a tearing mode. Let us now investigate how a tearing instability affects the field configuration as it develops.

It is convenient to write the magnetic field in terms of a flux-function:

$$\mathbf{B} = B_0 \, a \, \nabla\psi \times \hat{\mathbf{z}}. \tag{5.217}$$

Note that $\mathbf{B}\cdot\nabla\psi = 0$. It follows that magnetic field-lines run along contours of $\psi(x, y)$.

We can write

$$\psi(\bar{x}, \bar{y}) \simeq \psi_0(\bar{x}) + \psi_1(\bar{x}, \bar{y}), \tag{5.218}$$

where ψ_0 generates the equilibrium magnetic field, and ψ_1 generates the perturbed magnetic field associated with the tearing mode. Here, $\bar{y} = y/a$. In the vicinity of the interface, we have

$$\psi \simeq -\frac{F'(0)}{2} \, \bar{x}^{\,2} + \Psi \cos \bar{k} \, \bar{y}, \tag{5.219}$$

where Ψ is a constant. Here, we have made use of the fact that $\psi_1(\bar{x}, \bar{y}) \simeq \psi_1(\bar{y})$ if the constant-ψ approximation holds good (which is assumed to be the case).

Let $\chi = -\psi/\Psi$ and $\theta = \bar{k} \, \bar{y}$. It follows that the normalized perturbed magnetic flux function, χ, in the vicinity of the interface takes the form

$$\chi = 8 \, X^2 - \cos\theta, \tag{5.220}$$

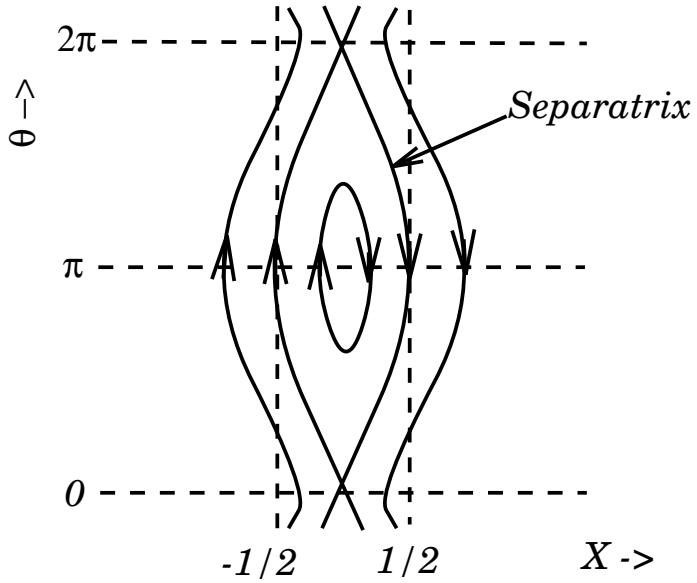

Figure 5.10: *Magnetic field-lines in the vicinity of a magnetic island.*

where $X = \bar{x}/\bar{W}$, and

$$\bar{W} = 4\sqrt{\frac{\Psi}{F'(0)}}. \tag{5.221}$$

Figure 5.10 shows the contours of χ plotted in X-θ space. It can be seen that the tearing mode gives rise to the formation of a *magnetic island* centred on the interface, $X = 0$. Magnetic field-lines situated outside the separatrix are displaced by the tearing mode, but still retain their original topology. By contrast, field-lines inside the separatrix have been broken and reconnected, and now possess quite different topology. The reconnection obviously takes place at the "X-points," which are located at $X = 0$ and $\theta = j\,2\pi$, where j is an integer. The maximum width of the reconnected region (in $\bar{x}$-space) is given by the *island width*, $a\,\bar{W}$. Note that the island width is proportional to the square root of the perturbed "radial" magnetic field at the interface (*i.e.*, $\bar{W} \propto \sqrt{\Psi}$).

According to a result first established in a very elegant paper by Rutherford,[17] the nonlinear evolution of the island width is governed by

$$0.823\,\tau_R\,\frac{d\bar{W}}{dt} = \Delta'(\bar{W}), \tag{5.222}$$

where

$$\Delta'(\bar{W}) = \left[\frac{1}{\psi}\frac{d\psi}{d\bar{x}}\right]_{-\bar{W}/2}^{+\bar{W}/2} \tag{5.223}$$

is the jump in the logarithmic derivative of ψ *taken across the island*. It is clear that once the tearing mode enters the nonlinear regime (*i.e.*, once the normalized island width, $\bar{W}$,

[17] P.H. Rutherford, Phys. Fluids **16**, 1903 (1973).

exceeds the normalized linear layer width, $S^{-2/5}$), the growth-rate of the instability slows down considerably, until the mode eventually ends up growing on the extremely slow resistive time-scale, τ_R. The tearing mode stops growing when it has attained a saturated island width $\bar{W}_0$, satisfying

$$\Delta'(\bar{W}_0) = 0. \tag{5.224}$$

The saturated width is a function of the original plasma equilibrium, but is independent of the resistivity. Note that there is no particular reason why $\bar{W}_0$ should be small: *i.e.*, in general, the saturated island width is comparable with the scale-length of the magnetic field configuration. We conclude that, although ideal-MHD only breaks down in a narrow region of width $S^{-2/5}$, centered on the interface, $\bar{x} = 0$, the reconnection of magnetic field-lines which takes place in this region is capable of significantly modifying the whole magnetic field configuration.

5.17 Fast Magnetic Reconnection

Up to now, we have only considered *spontaneous magnetic reconnection*, which develops from an instability of the plasma. As we have seen, such reconnection takes place at a fairly leisurely pace. Let us now consider *forced magnetic reconnection* in which the reconnection takes place as a consequence of an externally imposed flow or magnetic perturbation, rather than developing spontaneously. The principle difference between forced and spontaneous reconnection is the development of extremely large, positive Δ' values in the former case. Generally speaking, we expect Δ' to be $O(1)$ for spontaneous reconnection. By analogy with the previous analysis, we would expect forced reconnection to proceed *faster* than spontaneous reconnection (since the reconnection rate increases with increasing Δ'). The question is, how much faster? To be more exact, if we take the limit $\Delta' \rightarrow \infty$, which corresponds to the limit of extreme forced reconnection, just how fast can we make the magnetic field reconnect? At present, this is a *very controversial* question, which is far from being completely resolved. In the following, we shall content ourselves with a discussion of the two "classic" fast reconnection models. These models form the starting point of virtually all recent research on this subject.

Let us first consider the *Sweet-Parker* model, which was first proposed by Sweet[18] and Parker.[19] The main features of the envisioned magnetic and plasma flow fields are illustrated in Fig. 5.11. The system is two dimensional and steady-state (*i.e.*, $\partial/\partial z \equiv 0$ and $\partial/\partial t \equiv 0$). The reconnecting magnetic fields are anti-parallel, and of equal strength, B_*. We imagine that these fields are being forcibly pushed together via the action of some external agency. We expect a strong current sheet to form at the boundary between the two fields, where the direction of **B** suddenly changes. This current sheet is assumed to be of thickness δ and length L.

[18] P.A. Sweet, *Electromagnetic Phenomena in Cosmical Physics*, (Cambridge University Press, Cambridge UK, 1958).

[19] E.N. Parker, J. Geophys. Res. **62**, 509 (1957).

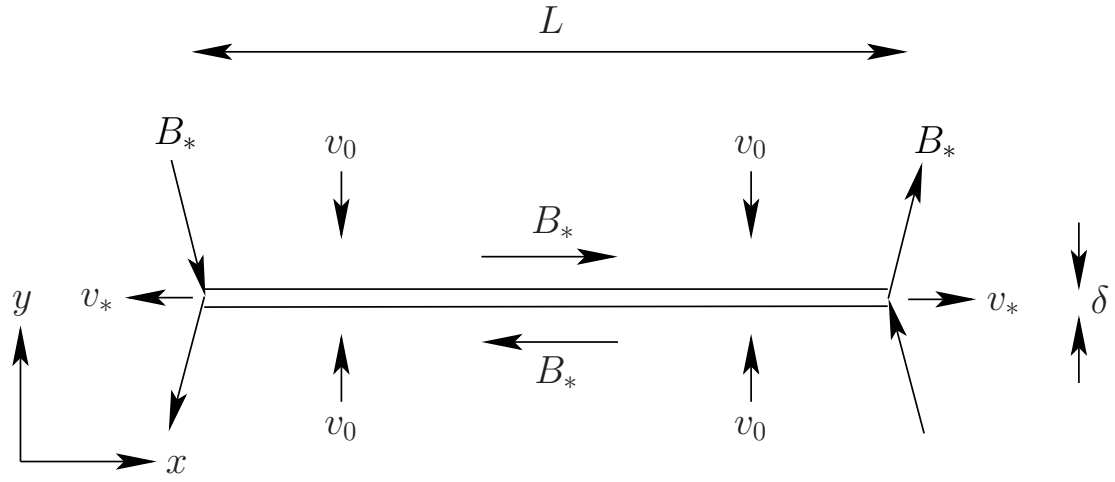

Figure 5.11: *The Sweet-Parker magnetic reconnection scenario.*

Plasma is assumed to diffuse into the current layer, along its whole length, at some relatively small inflow velocity, v_0. The plasma is accelerated along the layer, and eventually expelled from its two ends at some relatively large exit velocity, v_*. The inflow velocity is simply an $\mathbf{E}\times\mathbf{B}$ velocity, so

$$v_0 \sim \frac{E_z}{B_*}. \tag{5.225}$$

The z-component of Ohm's law yields

$$E_z \sim \frac{\eta\, B_*}{\mu_0\, \delta}. \tag{5.226}$$

Continuity of plasma flow inside the layer gives

$$L\, v_0 \sim \delta\, v_*, \tag{5.227}$$

assuming incompressible flow. Finally, pressure balance along the length of the layer yields

$$\frac{B_*^{\,2}}{\mu_0} \sim \rho\, v_*^{\,2}. \tag{5.228}$$

Here, we have balanced the magnetic pressure at the centre of the layer against the dynamic pressure of the outflowing plasma at the ends of the layer. Note that η and ρ are the plasma resistivity and density, respectively.

We can measure the rate of reconnection via the inflow velocity, v_0, since all of the magnetic field-lines which are convected into the layer, with the plasma, are eventually reconnected. The Alfvén velocity is written

$$V_A = \frac{B_*}{\sqrt{\mu_0\,\rho}}. \tag{5.229}$$

Likewise, we can write the Lundquist number of the plasma as

$$S = \frac{\mu_0 \, L \, V_A}{\eta}, \tag{5.230}$$

where we have assumed that the length of the reconnecting layer, L, is commensurate with the macroscopic length-scale of the system. The reconnection rate is parameterized via the Alfvénic Mach number of the inflowing plasma, which is defined

$$M_0 = \frac{v_0}{V_A}. \tag{5.231}$$

The above equations can be rearranged to give

$$v_* \sim V_A : \tag{5.232}$$

i.e., the plasma is squirted out of the ends of the reconnecting layer at the Alfvén velocity. Furthermore,

$$\delta \sim M_0 \, L, \tag{5.233}$$

and

$$M_0 \sim S^{-1/2}. \tag{5.234}$$

We conclude that the reconnecting layer is extremely narrow, assuming that the Lundquist number of the plasma is very large. The magnetic reconnection takes place on the hybrid time-scale $\tau_A^{1/2} \tau_R^{1/2}$, where τ_A is the Alfvén transit time-scale across the plasma, and τ_R is the resistive diffusion time-scale across the plasma.

The Sweet-Parker reconnection ansatz is undoubtedly correct. It has been simulated numerically innumerable times, and was recently confirmed experimentally in the Magnetic Reconnection Experiment (MRX) operated by Princeton Plasma Physics Laboratory.[20] The problem is that Sweet-Parker reconnection takes place *far too slowly* to account for many reconnection processes which are thought to take place in the solar system. For instance, in solar flares $S \sim 10^8$, $V_A \sim 100 \, \mathrm{km \, s^{-1}}$, and $L \sim 10^4 \, \mathrm{km}$. According to the Sweet-Parker model, magnetic energy is released to the plasma via reconnection on a typical time-scale of a few tens of days. In reality, the energy is released in a few minutes to an hour. Clearly, we can only hope to account for solar flares using a reconnection mechanism which operates *far faster* than the Sweet-Parker mechanism.

One, admittedly rather controversial, resolution of this problem was suggested by Petschek.[21] He pointed out that magnetic energy can be converted into plasma thermal energy as a result of shock waves being set up in the plasma, in addition to the conversion due to the action of resistive diffusion. The configuration envisaged by Petschek is sketched in Fig. 5.12. Two waves (slow mode shocks) stand in the flow on either side of the interface, where the direction of **B** reverses, marking the boundaries of the plasma

[20] H. Ji, M. Yamada, S. Hsu, and R. Kulsrud, Phys. Rev. Lett. **80**, 3256 (1998).

[21] H.E. Petschek, *AAS-NASA Symposium on the Physics of Solar Flares* (NASA Spec. Publ. Sp-50, 1964), p. 425.

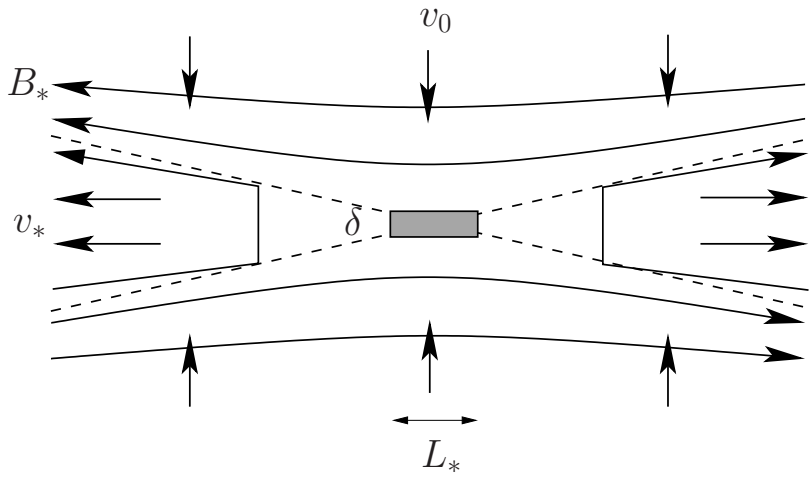

Figure 5.12: *The Petschek magnetic reconnection scenario.*

outflow regions. A small diffusion region still exists on the interface, but now constitutes a miniature (in length) Sweet-Parker system. The width of the reconnecting layer is given by

$$\delta = \frac{L}{M_0 S}, \tag{5.235}$$

just as in the Sweet-Parker model. However, we do not now assume that the length, L_*, of the layer is comparable to the scale-size, L, of the system. Rather, the length may be considerably smaller than L, and is determined self-consistently from the continuity condition

$$L_* = \frac{\delta}{M_0}, \tag{5.236}$$

where we have assumed incompressible flow, and an outflow speed of order the Alfvén speed, as before. Thus, if the inflow speed, v_0, is much less than V_A then the length of the reconnecting layer is much larger than its width, as assumed by Sweet and Parker. On the other hand, if we allow the inflow velocity to approach the Alfvén velocity then the layer shrinks in length, so that L_* becomes comparable with δ.

It follows that for reasonably large reconnection rates (*i.e.*, $M_0 \rightarrow 1$) the length of the diffusion region becomes much smaller than the scale-size of the system, L, so that most of the plasma flowing into the boundary region does so across the standing waves, rather than through the central diffusion region. The angle θ that the shock waves make with the interface is given approximately by

$$\tan\theta \sim M_0. \tag{5.237}$$

Thus, for small inflow speeds the outflow is confined to a narrow wedge along the interface, but as the inflow speed increases the angle of the outflow wedges increases to accommodate the increased flow.

It turns out that there is a maximum inflow speed beyond which Petschek-type solutions cease to exist. The corresponding maximum Alfvénic Mach number,

$$(M_0)_{\rm max} = \frac{\pi}{8\ln S}, \tag{5.238}$$

can be regarded as specifying the maximum allowable rate of magnetic reconnection according to the Petschek model. Clearly, since the maximum reconnection rate depends inversely on the logarithm of the Lundquist number, rather than its square root, it is much larger than that predicted by the Sweet-Parker model.

It must be pointed out that the Petschek model is *very* controversial. Many physicists think that it is completely wrong, and that the maximum rate of magnetic reconnection allowed by MHD is that predicted by the Sweet-Parker model. In particular, Biskamp[22] wrote an influential and widely quoted paper reporting the results of a numerical experiment which appeared to disprove the Petschek model. When the plasma inflow exceeded that allowed by the Sweet-Parker model, there was no acceleration of the reconnection rate. Instead, magnetic flux "piled up" in front of the reconnecting layer, and the rate of reconnection never deviated significantly from that predicted by the Sweet-Parker model. Priest and Forbes[23] later argued that Biskamp imposed boundary conditions in his numerical experiment which precluded Petschek reconnection. Probably the most powerful argument against the validity of the Petschek model is the fact that, more than 30 years after it was first proposed, nobody has ever managed to simulate Petschek reconnection numerically (except by artificially increasing the resistivity in the reconnecting region—which is not a legitimate approach).

5.18 MHD Shocks

Consider a subsonic disturbance moving through a conventional neutral fluid. As is well-known, *sound waves* propagating ahead of the disturbance give advance warning of its arrival, and, thereby, allow the response of the fluid to be both smooth and adiabatic. Now, consider a supersonic distrurbance. In this case, sound waves are unable to propagate ahead of the disturbance, and so there is no advance warning of its arrival, and, consequently, the fluid response is sharp and non-adiabatic. This type of response is generally known as a *shock*.

Let us investigate shocks in MHD fluids. Since information in such fluids is carried via three different waves—namely, *fast* or compressional-Alfvén waves, *intermediate* or shear-Alfvén waves, and *slow* or magnetosonic waves (see Sect. 5.4)—we might expect MHD fluids to support three different types of shock, corresponding to disturbances traveling faster than each of the aforementioned waves. This is indeed the case.

In general, a shock propagating through an MHD fluid produces a significant difference in plasma properties on either side of the shock front. The thickness of the front is

[22] D. Biskamp, Phys. Fluids **29**, 1520 (1986).

[23] E.R. Priest, and T.G. Forbes, J. Geophys. Res. **97**, 16757 (1992).

determined by a balance between convective and dissipative effects. However, dissipative effects in high temperature plasmas are only comparable to convective effects when the spatial gradients in plasma variables become extremely large. Hence, MHD shocks in such plasmas tend to be *extremely narrow*, and are well-approximated as *discontinuous* changes in plasma parameters. The MHD equations, and Maxwell's equations, can be integrated across a shock to give a set of *jump conditions* which relate plasma properties on each side of the shock front. If the shock is sufficiently narrow then these relations become *independent* of its detailed structure. Let us derive the jump conditions for a narrow, planar, steady-state, MHD shock.

Maxwell's equations, and the MHD equations, (5.1)–(5.4), can be written in the following convenient form:

$$\nabla\cdot\mathbf{B} = 0, \tag{5.239}$$

$$\frac{\partial\mathbf{B}}{\partial t} - \nabla\times(\mathbf{V}\times\mathbf{B}) = \mathbf{0}, \tag{5.240}$$

$$\frac{\partial\rho}{\partial t} + \nabla\cdot(\rho\,\mathbf{V}) = 0, \tag{5.241}$$

$$\frac{\partial(\rho\,\mathbf{V})}{\partial t} + \nabla\cdot\mathbf{T} = \mathbf{0}, \tag{5.242}$$

$$\frac{\partial U}{\partial t} + \nabla\cdot\mathbf{u} = 0, \tag{5.243}$$

where

$$\mathbf{T} = \rho\,\mathbf{V}\mathbf{V} + \left(p + \frac{B^2}{2\mu_0}\right)\mathbf{I} - \frac{\mathbf{B}\,\mathbf{B}}{\mu_0} \tag{5.244}$$

is the total (*i.e.*, including electromagnetic, as well as plasma, contributions) *stress tensor*, $\mathbf{I}$ the identity tensor,

$$U = \frac{1}{2}\,\rho\,V^2 + \frac{p}{\Gamma-1} + \frac{B^2}{2\mu_0} \tag{5.245}$$

the total *energy density*, and

$$\mathbf{u} = \left(\frac{1}{2}\,\rho\,V^2 + \frac{\Gamma}{\Gamma-1}\,p\right)\mathbf{V} + \frac{\mathbf{B}\times(\mathbf{V}\times\mathbf{B})}{\mu_0} \tag{5.246}$$

the total *energy flux density*.

Let us move into the *rest frame* of the shock. Suppose that the shock front coincides with the y-z plane. Furthermore, let the regions of the plasma upstream and downstream of the shock, which are termed regions 1 and 2, respectively, be *spatially uniform* and *non-time-varying*. It follows that $\partial/\partial t = \partial/\partial y = \partial/\partial z = 0$. Moreover, $\partial/\partial x = 0$, except in the immediate vicinity of the shock. Finally, let the velocity and magnetic fields upstream and downstream of the shock all lie in the x-y plane. The situation under discussion is illustrated in Fig. 5.13. Here, ρ_1, p_1, $\mathbf{V}_1$, and $\mathbf{B}_1$ are the downstream mass density, pressure,

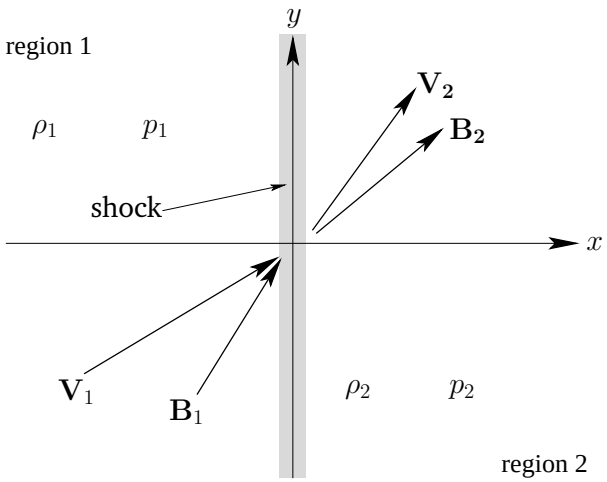

Figure 5.13: *A planar shock.*

velocity, and magnetic field, respectively, whereas ρ_2, p_2, $\mathbf{V}_2$, and $\mathbf{B}_2$ are the corresponding upstream quantities.

In the immediate vicinity of the shock, Eqs. (5.239)–(5.243) reduce to

$$\frac{dB_x}{dx} = 0, \tag{5.247}$$

$$\frac{d}{dx}(V_x B_y - V_y B_x) = 0, \tag{5.248}$$

$$\frac{d(\rho V_x)}{dx} = 0, \tag{5.249}$$

$$\frac{dT_{xx}}{dx} = 0, \tag{5.250}$$

$$\frac{dT_{xy}}{dx} = 0, \tag{5.251}$$

$$\frac{du_x}{dx} = 0. \tag{5.252}$$

Integration across the shock yields the desired jump conditions:

$$[B_x]_1^2 = 0, \tag{5.253}$$

$$[V_x B_y - V_y B_x]_1^2 = 0, \tag{5.254}$$

$$[\rho V_x]_1^2 = 0, \tag{5.255}$$

$$[\rho V_x^2 + p + B_y^2/2\mu_0]_1^2 = 0, \tag{5.256}$$

$$[\rho V_x V_y - B_x B_y/\mu_0]_1^2 = 0, \tag{5.257}$$

$$\left[\frac{1}{2}\rho V^2 V_x + \frac{\Gamma}{\Gamma-1} p V_x + \frac{B_y (V_x B_y - V_y B_x)}{\mu_0}\right]_1^2 = 0, \tag{5.258}$$

where $[A]_1^2 \equiv A_2 - A_1$. These relations are often called the *Rankine-Hugoniot* relations for MHD. Assuming that all of the upstream plasma parameters are known, there are six unknown parameters in the problem—namely, $B_{x\,2}$, $B_{y\,2}$, $V_{x\,2}$, $V_{y\,2}$, ρ_2, and p_2. These six unknowns are fully determined by the six jump conditions. Unfortunately, the general case is very complicated. So, before tackling it, let us examine a couple of relatively simple special cases.

5.19 Parallel Shocks

The first special case is the so-called *parallel shock* in which both the upstream and downstream plasma flows are parallel to the magnetic field, as well as perpendicular to the shock front. In other words,

$$\mathbf{V}_1 = (V_1, 0, 0), \qquad \mathbf{V}_2 = (V_2, 0, 0), \tag{5.259}$$

$$\mathbf{B}_1 = (B_1, 0, 0), \qquad \mathbf{B}_2 = (B_2, 0, 0). \tag{5.260}$$

Substitution into the general jump conditions (5.253)–(5.258) yields

$$\frac{B_2}{B_1} = 1, \tag{5.261}$$

$$\frac{\rho_2}{\rho_1} = r, \tag{5.262}$$

$$\frac{V_2}{V_1} = r^{-1}, \tag{5.263}$$

$$\frac{p_2}{p_1} = R, \tag{5.264}$$

with

$$r = \frac{(\Gamma+1)\,M_1^2}{2+(\Gamma-1)\,M_1^2}, \tag{5.265}$$

$$R = 1+\Gamma\,M_1^2\,(1-r^{-1}) = \frac{(\Gamma+1)\,r-(\Gamma-1)}{(\Gamma+1)-(\Gamma-1)\,r}. \tag{5.266}$$

Here, $M_1 = V_1/V_{S\,1}$, where $V_{S\,1} = (\Gamma\,p_1/\rho_1)^{1/2}$ is the upstream sound speed. Thus, the upstream flow is supersonic if $M_1 > 1$, and subsonic if $M_1 < 1$. Incidentally, as is clear from the above expressions, a parallel shock is *unaffected* by the presence of a magnetic field. In fact, this type of shock is identical to that which occurs in neutral fluids, and is, therefore, usually called a *hydrodynamic shock*.

It is easily seen from Eqs. (5.261)–(5.264) that there is no shock (*i.e.*, no jump in plasma parameters across the shock front) when the upstream flow is exactly sonic: *i.e.*, when $M_1 = 1$. In other words, $r = R = 1$ when $M_1 = 1$. However, if $M_1 \neq 1$ then the

upstream and downstream plasma parameters become different (*i.e.*, $r \neq 1$, $R \neq 1$) and a true shock develops. In fact, it is easily demonstrated that

$$\frac{\Gamma-1}{\Gamma+1} \leq \quad r \quad \leq \frac{\Gamma+1}{\Gamma-1}, \tag{5.267}$$

$$0 \leq \quad R \quad \leq \infty, \tag{5.268}$$

$$\frac{\Gamma-1}{2\,\Gamma} \leq \quad M_1^2 \quad \leq \infty. \tag{5.269}$$

Note that the upper and lower limits in the above inequalities are all attained simultaneously.

The previous discussion seems to imply that a parallel shock can be either compressive (*i.e.*, $r > 1$) or expansive (*i.e.*, $r < 1$). However, there is one additional physics principle which needs to be factored into our analysis—namely, the *second law of thermodynamics*. This law states that the *entropy* of a closed system can spontaneously increase, but can never spontaneously decrease. Now, in general, the entropy per particle is different on either side of a hydrodynamic shock front. Accordingly, the second law of thermodynamics mandates that the downstream entropy must *exceed* the upstream entropy, so as to ensure that the shock generates a net increase, rather than a net decrease, in the overall entropy of the system, as the plasma flows through it.

The (suitably normalized) entropy per particle of an ideal plasma takes the form [see Eq. (3.59)]

$$S = \ln(p/\rho^\Gamma). \tag{5.270}$$

Hence, the difference between the upstream and downstream entropies is

$$[S]_1^2 = \ln R - \Gamma\,\ln r. \tag{5.271}$$

Now, using (5.265),

$$r\,\frac{d[S]_1^2}{dr} = \frac{r}{R}\,\frac{dR}{dr} - \Gamma = \frac{\Gamma\,(\Gamma^2-1)\,(r-1)^2}{[(\Gamma+1)\,r-(\Gamma-1)]\,[(\Gamma+1)-(\Gamma-1)\,r]}. \tag{5.272}$$

Furthermore, it is easily seen from Eqs. (5.267)–(5.269) that $d[S]_1^2/dr \geq 0$ in all situations of physical interest. However, $[S]_1^2 = 0$ when $r = 1$, since, in this case, there is no discontinuity in plasma parameters across the shock front. We conclude that $[S]_1^2 < 0$ for $r < 1$, and $[S]_1^2 > 0$ for $r > 1$. It follows that the second law of thermodynamics requires hydrodynamic shocks to be *compressive*: *i.e.*, $r > 1$. In other words, the plasma density must always *increase* when a shock front is crossed in the direction of the relative plasma flow. It turns out that this is a general rule which applies to all three types of MHD shock.

The upstream Mach number, M_1, is a good measure of shock strength: *i.e.*, if $M_1 = 1$ then there is no shock, if $M_1 - 1 \ll 1$ then the shock is weak, and if $M_1 \gg 1$ then the shock is strong. We can define an analogous downstream Mach number, $M_2 = V_2/(\Gamma\,p_2/\rho_2)^{1/2}$. It is easily demonstrated from the jump conditions that if $M_1 > 1$ then $M_2 < 1$. In other words, in the shock rest frame, the shock is associated with an irreversible (since

the entropy suddenly increases) transition from supersonic to subsonic flow. Note that $r \equiv \rho_2/\rho_1 \rightarrow (\Gamma+1)/(\Gamma-1)$, whereas $R \equiv p_2/p_1 \rightarrow \infty$, in the limit $M_1 \rightarrow \infty$. In other words, as the shock strength increases, the compression ratio, r, asymptotes to a finite value, whereas the pressure ratio, P, increases without limit. For a conventional plasma with $\Gamma = 5/3$, the limiting value of the compression ratio is 4: *i.e.*, the downstream density can never be more than four times the upstream density. We conclude that, in the strong shock limit, $M_1 \gg 1$, the large jump in the plasma pressure across the shock front must be predominately a consequence of a large jump in the plasma *temperature*, rather than the plasma density. In fact, Eqs. (5.265)–(5.266) imply that

$$\frac{T_2}{T_1} \equiv \frac{R}{r} \rightarrow \frac{2\,\Gamma\,(\Gamma-1)\,M_1^2}{(\Gamma+1)^2} \gg 1 \tag{5.273}$$

as $M_1 \rightarrow \infty$. Thus, a strong parallel, or hydrodynamic, shock is associated with intense plasma heating.

As we have seen, the condition for the existence of a hydrodynamic shock is $M_1 > 1$, or $V_1 > V_{S\,1}$. In other words, in the shock frame, the upstream plasma velocity, V_1, must be supersonic. However, by Galilean invariance, V_1 can also be interpreted as the *propagation velocity* of the shock through an initially *stationary* plasma. It follows that, in a stationary plasma, a parallel, or hydrodynamic, shock propagates along the magnetic field with a *supersonic* velocity.

5.20 Perpendicular Shocks

The second special case is the so-called *perpendicular shock* in which both the upstream and downstream plasma flows are perpendicular to the magnetic field, as well as the shock front. In other words,

$$\mathbf{V}_1 = (V_1,\,0,\,0), \qquad \mathbf{V}_2 = (V_2,\,0,\,0), \tag{5.274}$$

$$\mathbf{B}_1 = (0,\,B_1,\,0), \qquad \mathbf{B}_2 = (0,\,B_2,\,0). \tag{5.275}$$

Substitution into the general jump conditions (5.253)–(5.258) yields

$$\frac{B_2}{B_1} = r, \tag{5.276}$$

$$\frac{\rho_2}{\rho_1} = r, \tag{5.277}$$

$$\frac{V_2}{V_1} = r^{-1}, \tag{5.278}$$

$$\frac{p_2}{p_1} = R, \tag{5.279}$$

where

$$R = 1 + \Gamma\,M_1^2\,(1 - r^{-1}) + \beta_1^{-1}\,(1 - r^2), \tag{5.280}$$

and r is a real positive root of the quadratic

$$F(r) = 2\,(2-\Gamma)\,r^2 + \Gamma\,[2\,(1+\beta_1) + (\Gamma-1)\,\beta_1\,M_1^2]\,r - \Gamma\,(\Gamma+1)\,\beta_1\,M_1^2 = 0. \tag{5.281}$$

Here, $\beta_1 = 2\mu_0\,p_1/B_1^2$.

Now, if r_1 and r_2 are the two roots of Eq. (5.281) then

$$r_1\,r_2 = -\frac{\Gamma\,(\Gamma+1)\,\beta_1\,M_1^2}{2\,(2-\Gamma)}. \tag{5.282}$$

Assuming that $\Gamma < 2$, we conclude that one of the roots is negative, and, hence, that Eq. (5.281) only possesses *one* physical solution: *i.e.*, there is only one type of MHD shock which is consistent with Eqs. (5.274) and (5.275). Now, it is easily demonstrated that $F(0) < 0$ and $F(\Gamma+1/\Gamma-1) > 0$. Hence, the physical root lies between $r = 0$ and $r = (\Gamma+1)/(\Gamma-1)$.

Using similar analysis to that employed in the previous subsection, it is easily demonstrated that the second law of thermodynamics requires a perpendicular shock to be compressive: *i.e.*, $r > 1$. It follows that a physical solution is only obtained when $F(1) < 0$, which reduces to

$$M_1^2 > 1 + \frac{2}{\Gamma\,\beta_1}. \tag{5.283}$$

This condition can also be written

$$V_1^2 > V_{S1}^2 + V_{A1}^2, \tag{5.284}$$

where $V_{A1} = B_1/(\mu_0\,\rho_1)^{1/2}$ is the upstream Alfvén velocity. Now, $V_{+1} = (V_{S1}^2 + V_{A1}^2)^{1/2}$ can be recognized as the velocity of a *fast wave* propagating perpendicular to the magnetic field—see Sect. 5.4. Thus, the condition for the existence of a perpendicular shock is that the relative upstream plasma velocity must be *greater* than the upstream fast wave velocity. Incidentally, it is easily demonstrated that if this is the case then the downstream plasma velocity is *less* than the downstream fast wave velocity. We can also deduce that, in a stationary plasma, a perpendicular shock propagates across the magnetic field with a velocity which exceeds the fast wave velocity.

In the strong shock limit, $M_1 \gg 1$, Eqs. (5.280) and (5.281) become identical to Eqs. (5.265) and (5.266). Hence, a strong perpendicular shock is very similar to a strong hydrodynamic shock (except that the former shock propagates perpendicular, whereas the latter shock propagates parallel, to the magnetic field). In particular, just like a hydrodynamic shock, a perpendicular shock cannot compress the density by more than a factor $(\Gamma+1)/(\Gamma-1)$. However, according to Eq. (5.276), a perpendicular shock compresses the magnetic field by the same factor that it compresses the plasma density. It follows that there is also an upper limit to the factor by which a perpendicular shock can compress the magnetic field.

5.21 Oblique Shocks

Let us now consider the general case in which the plasma velocities and the magnetic fields on each side of the shock are neither parallel nor perpendicular to the shock front. It is convenient to transform into the so-called *de Hoffmann-Teller* frame in which $|\mathbf{V}_1 \times \mathbf{B}_1| = 0$, or

$$V_{x\,1}\, B_{y\,1} - V_{y\,1}\, B_{x\,1} = 0. \tag{5.285}$$

In other words, it is convenient to transform to a frame which moves at the local $\mathbf{E} \times \mathbf{B}$ velocity of the plasma. It immediately follows from the jump condition (5.254) that

$$V_{x\,2}\, B_{y\,2} - V_{y\,2}\, B_{x\,2} = 0, \tag{5.286}$$

or $|\mathbf{V}_2 \times \mathbf{B}_2| = 0$. Thus, in the de Hoffmann-Teller frame, the upstream plasma flow is *parallel* to the upstream magnetic field, and the downstream plasma flow is also parallel to the downstream magnetic field. Furthermore, the magnetic contribution to the jump condition (5.258) becomes identically zero, which is a considerable simplification.

Equations (5.285) and (5.286) can be combined with the general jump conditions (5.253)–(5.258) to give

$$\frac{\rho_2}{\rho_1} = r, \tag{5.287}$$

$$\frac{B_{x\,2}}{B_{x\,1}} = 1, \tag{5.288}$$

$$\frac{B_{y\,2}}{B_{y\,1}} = r\left(\frac{v_1^2 - \cos^2\theta_1\, V_{A\,1}^2}{v_1^2 - r\,\cos^2\theta_1\, V_{A\,1}^2}\right), \tag{5.289}$$

$$\frac{V_{x\,2}}{V_{x\,1}} = \frac{1}{r}, \tag{5.290}$$

$$\frac{V_{y\,2}}{V_{y\,1}} = \frac{v_1^2 - \cos^2\theta_1\, V_{A\,1}^2}{v_1^2 - r\,\cos^2\theta_1\, V_{A\,1}^2}, \tag{5.291}$$

$$\frac{p_2}{p_1} = 1 + \frac{\Gamma\, v_1^2\,(r-1)}{V_{S\,1}^2\, r}\left[1 - \frac{r\, V_{A\,1}^2\,[(r+1)\, v_1^2 - 2\, r\, V_{A\,1}^2\,\cos^2\theta_1]}{2\,(v_1^2 - r\, V_{A\,1}^2\,\cos^2\theta_1)^2}\right]. \tag{5.292}$$

where $v_1 = V_{x\,1} = V_1\,\cos\theta_1$ is the component of the upstream velocity normal to the shock front, and θ_1 is the angle subtended between the upstream plasma flow and the shock front normal. Finally, given the compression ratio, r, the square of the normal upstream velocity, v_1^2, is a real root of a cubic equation known as the *shock adiabatic*:

$$\begin{aligned} 0 = {} & (v_1^2 - r\,\cos^2\theta_1\, V_{A\,1}^2)^2\left\{[(\Gamma+1) - (\Gamma-1)\, r]\, v_1^2 - 2\, r\, V_{S\,1}^2\right\} \\ & - r\,\sin^2\theta_1\, v_1^2\, V_{A\,1}^2\left\{[\Gamma + (2-\Gamma)\, r]\, v_1^2 - [(\Gamma+1) - (\Gamma-1)\, r]\, r\,\cos^2\theta_1\, V_{A\,1}^2\right]\}. \end{aligned} \tag{5.293}$$

As before, the second law of thermodynamics mandates that $r > 1$.

Let us first consider the weak shock limit $r \rightarrow 1$. In this case, it is easily seen that the three roots of the shock adiabatic reduce to

$$v_1^2 = V_{-1}^2 \equiv \frac{V_{A1}^2 + V_{S1}^2 - [(V_{A1} + V_{S1})^2 - 4\,\cos^2\theta_1\,V_{S1}^2\,V_{A1}^2]^{1/2}}{2}, \tag{5.294}$$

$$v_1^2 = \cos^2\theta_1\,V_{A1}^2, \tag{5.295}$$

$$v_1^2 = V_{+1}^2 \equiv \frac{V_{A1}^2 + V_{S1}^2 + [(V_{A1} + V_{S1})^2 - 4\,\cos^2\theta_1\,V_{S1}^2\,V_{A1}^2]^{1/2}}{2}. \tag{5.296}$$

However, from Sect. 5.4, we recognize these velocities as belonging to slow, intermediate (or Shear-Alfvén), and fast waves, respectively, propagating in the normal direction to the shock front. We conclude that slow, intermediate, and fast MHD shocks degenerate into the associated MHD waves in the limit of small shock amplitude. Conversely, we can think of the various MHD shocks as *nonlinear* versions of the associated MHD waves. Now it is easily demonstrated that

$$V_{+1} > \cos\theta_1\,V_{A1} > V_{-1}. \tag{5.297}$$

In other words, a fast wave travels faster than an intermediate wave, which travels faster than a slow wave. It is reasonable to suppose that the same is true of the associated MHD shocks, at least at relatively low shock strength. It follows from Eq. (5.289) that $B_{y2} > B_{y1}$ for a fast shock, whereas $B_{y2} < B_{y1}$ for a slow shock. For the case of an intermediate shock, we can show, after a little algebra, that $B_{y2} \rightarrow -B_{y1}$ in the limit $r \rightarrow 1$. We conclude that (in the de Hoffmann-Teller frame) *fast* shocks refract the magnetic field and plasma flow (recall that they are parallel in our adopted frame of the reference) *away* from the normal to the shock front, whereas *slow* shocks refract these quantities *toward* the normal. Moreover, the tangential magnetic field and plasma flow generally *reverse* across an *intermediate* shock front. This is illustrated in Fig. 5.14.

When r is slightly larger than unity it is easily demonstrated that the conditions for the existence of a slow, intermediate, and fast shock are $v_1 > V_{-1}$, $v_1 > \cos\theta_1\,V_{A1}$, and $v_1 > V_{+1}$, respectively.

Let us now consider the strong shock limit, $v_1^2 \gg 1$. In this case, the shock adiabatic yields $r \rightarrow r_m = (\Gamma+1)/(\Gamma-1)$, and

$$v_1^2 \simeq \frac{r_m}{\Gamma-1}\,\frac{2\,V_{S1}^2 + \sin^2\theta_1\,[\Gamma + (2-\Gamma)\,r_m]\,V_{A1}^2}{r_m - r}. \tag{5.298}$$

There are no other real roots. The above root is clearly a type of fast shock. The fact that there is only one real root suggests that there exists a critical shock strength above which the slow and intermediate shock solutions cease to exist. (In fact, they merge and annihilate one another.) In other words, there is a limit to the strength of a slow or an intermediate shock. On the other hand, there is no limit to the strength of a fast shock. Note, however, that the plasma density and tangential magnetic field cannot be compressed by more than a factor $(\Gamma+1)/(\Gamma-1)$ by any type of MHD shock.

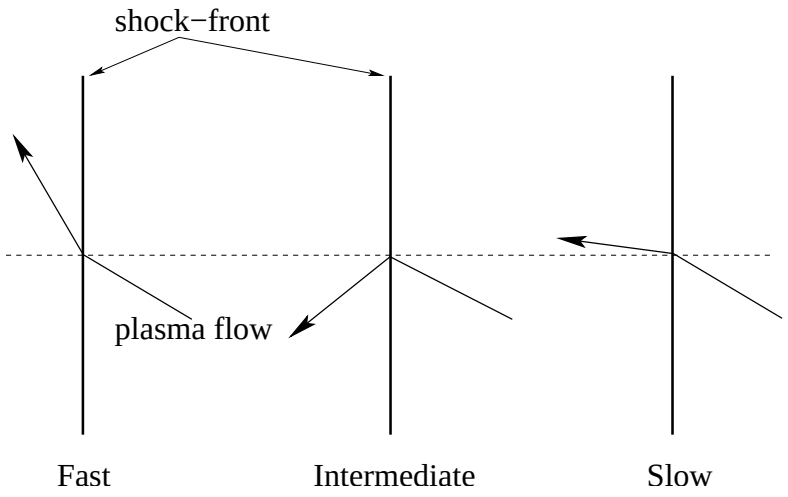

Figure 5.14: *Characteristic plasma flow patterns across the three different types of MHD shock in the shock rest frame.*

Consider the special case $\theta_1 = 0$ in which both the plasma flow and the magnetic field are normal to the shock front. In this case, the three roots of the shock adiabatic are

$$v_1^2 = \frac{2\,r\,V_{S1}^2}{(\Gamma+1)-(\Gamma-1)\,r}, \tag{5.299}$$

$$v_1^2 = r\,V_{A1}^2, \tag{5.300}$$

$$v_1^2 = r\,V_{A1}^2. \tag{5.301}$$

We recognize the first of these roots as the hydrodynamic shock discussed in Sect. 5.19—*cf.* Eq. (5.265). This shock is classified as a slow shock when $V_{S1} < V_{A1}$, and as a fast shock when $V_{S1} > V_{A1}$. The other two roots are identical, and correspond to shocks which propagate at the velocity $v_1 = \sqrt{r}\,V_{A1}$ and "switch-on" the tangential components of the plasma flow and the magnetic field: *i.e.*, it can be seen from Eqs. (5.289) and (5.291) that $V_{y1} = B_{y1} = 0$ whilst $V_{y2} \neq 0$ and $B_{y2} \neq 0$ for these types of shock. Incidentally, it is also possible to have a "switch-off" shock which eliminates the tangential components of the plasma flow and the magnetic field. According to Eqs. (5.289) and (5.291), such a shock propagates at the velocity $v_1 = \cos\theta_1\,V_{A1}$. Switch-on and switch-off shocks are illustrated in Fig. 5.15.

Let us, finally, consider the special case $\theta = \pi/2$. As is easily demonstrated, the three roots of the shock adiabatic are

$$v_1^2 = r\left(\frac{2\,V_{S1}^2 + [\Gamma + (2-\Gamma)\,r]\,V_{A1}^2}{(\Gamma+1)-(\Gamma-1)\,r}\right), \tag{5.302}$$

$$v_1^2 = 0, \tag{5.303}$$

$$v_1^2 = 0. \tag{5.304}$$

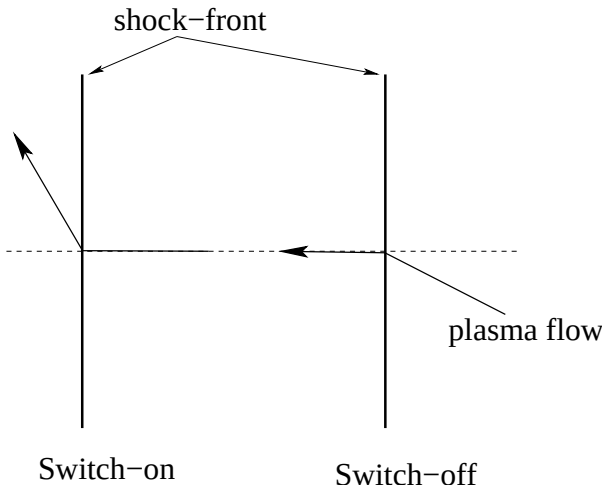

Figure 5.15: *Characteristic plasma flow patterns across switch-on and switch-off shocks in the shock rest frame.*

The first of these roots is clearly a fast shock, and is identical to the perpendicular shock discussed in Sect. 5.20, except that there is no plasma flow across the shock front in this case. The fact that the two other roots are zero indicates that, like the corresponding MHD waves, slow and intermediate MHD shocks do not propagate perpendicular to the magnetic field.

MHD shocks have been observed in a large variety of situations. For instance, shocks are known to be formed by supernova explosions, by strong stellar winds, by solar flares, and by the solar wind upstream of planetary magnetospheres.[24]

[24]D.A. Gurnett, and A. Bhattacharjee, *Introduction to Plasma Physics*, Cambridge University Press, Cambridge UK, 2005.

6 Waves in Warm Plasmas

6.1 Introduction

In this section we shall investigate wave propagation in a warm collisionless plasma, extending the discussion given in Sect. 4 to take thermal effects into account. It turns out that thermal modifications to wave propagation are not very well described by fluid equations. We shall, therefore, adopt a *kinetic* description of the plasma. The appropriate kinetic equation is, of course, the *Vlasov equation*, which is described in Sect. 3.1.

6.2 Landau Damping

Let us begin our study of the Vlasov equation by examining what appears, at first sight, to be a fairly simple and straight-forward problem. Namely, the propagation of small amplitude plasma waves through a uniform plasma with no equilibrium magnetic field. For the sake of simplicity, we shall only consider electron motion, assuming that the ions form an immobile, neutralizing background. The ions are also assumed to be singly-charged. We shall look for electrostatic plasma waves of the type discussed in Sect. 4.7. Such waves are longitudinal in nature, and possess a perturbed electric field, but no perturbed magnetic field.

Our starting point is the Vlasov equation for an unmagnetized, collisionless plasma:

$$\frac{\partial f_e}{\partial t} + \mathbf{v}\cdot\nabla f_e - \frac{e}{m_e}\,\mathbf{E}\cdot\nabla_v f_e = 0, \tag{6.1}$$

where $f_e(\mathbf{r},\mathbf{v},t)$ is the ensemble averaged electron distribution function. The electric field satisfies

$$\mathbf{E} = -\nabla\phi. \tag{6.2}$$

where

$$\nabla^2\phi = -\frac{e}{\epsilon_0}\left(n - \int f_e\, d^3\mathbf{v}\right). \tag{6.3}$$

Here, n is the number density of ions (which is the same as the number density of electrons).

Since we are dealing with small amplitude waves, it is appropriate to *linearize* the Vlasov equation. Suppose that the electron distribution function is written

$$f_e(\mathbf{r},\mathbf{v},t) = f_0(\mathbf{v}) + f_1(\mathbf{r},\mathbf{v},t). \tag{6.4}$$

Here, f_0 represents the equilibrium electron distribution, whereas f_1 represents the small perturbation due to the wave. Note that $\int f_0\, d^3\mathbf{v} = n$, otherwise the equilibrium state is

not quasi-neutral. The electric field is assumed to be zero in the unperturbed state, so that **E** can be regarded as a small quantity. Thus, linearization of Eqs. (6.1) and (6.3) yields

$$\frac{\partial f_1}{\partial t} + \mathbf{v}\cdot\nabla f_1 - \frac{e}{m_e}\,\mathbf{E}\cdot\nabla_v f_0 = 0, \tag{6.5}$$

and

$$\nabla^2\phi = \frac{e}{\epsilon_0}\int f_1\, d^3\mathbf{v}, \tag{6.6}$$

respectively.

Let us now follow the standard procedure for analyzing small amplitude waves, by assuming that all perturbed quantities vary with **r** and t like $\exp[i\,(\mathbf{k}\cdot\mathbf{r} - \omega\, t)]$. Equations (6.5) and (6.6) reduce to

$$-i\,(\omega - \mathbf{k}\cdot\mathbf{v}) f_1 + i\,\frac{e}{m_e}\,\phi\,\mathbf{k}\cdot\nabla_v f_0 = 0, \tag{6.7}$$

and

$$-k^2\,\phi = \frac{e}{\epsilon_0}\int f_1\, d^3\mathbf{v}, \tag{6.8}$$

respectively. Solving the first of these equations for f_1, and substituting into the integral in the second, we conclude that if ϕ is non-zero then we must have

$$1 + \frac{e^2}{\epsilon_0\, m_e\, k^2}\int \frac{\mathbf{k}\cdot\nabla_v f_0}{\omega - \mathbf{k}\cdot\mathbf{v}}\, d^3\mathbf{v} = 0. \tag{6.9}$$

We can interpret Eq. (6.9) as the dispersion relation for electrostatic plasma waves, relating the wave-vector, **k**, to the frequency, ω. However, in doing so, we run up against a serious problem, since the integral has a *singularity* in velocity space, where $\omega = \mathbf{k}\cdot\mathbf{v}$, and is, therefore, not properly defined.

The way around this problem was first pointed out by Landau[1] in a very influential paper which laid the basis of much subsequent research on plasma oscillations and instabilities. Landau showed that, instead of simply assuming that f_1 varies in time as $\exp(-i\,\omega\, t)$, the problem must be regarded as an initial value problem in which f_1 is given at $t = 0$ and found at later times. We may still Fourier analyze with respect to **r**, so we write

$$f_1(\mathbf{r}, \mathbf{v}, t) = f_1(\mathbf{v}, t)\, e^{\,i\,\mathbf{k}\cdot\mathbf{r}}. \tag{6.10}$$

It is helpful to define u as the velocity component along **k** (*i.e.*, $u = \mathbf{k}\cdot\mathbf{v}/k$) and to define $F_0(u)$ and $F_1(u, t)$ to be the integrals of $f_0(\mathbf{v})$ and $f_1(\mathbf{v}, t)$ over the velocity components perpendicular to **k**. Thus, we obtain

$$\frac{\partial F_1}{\partial t} + i\,k\,u\,F_1 - \frac{e}{m_e}\,E\,\frac{\partial F_0}{\partial u} = 0, \tag{6.11}$$

[1] L.D. Landau, Sov. Phys.–JETP **10**, 25 (1946).

and

$$\mathrm{i}\,k\,E = -\frac{e}{\epsilon_0}\int_{-\infty}^{\infty} F_1(u)\,du. \tag{6.12}$$

In order to solve Eqs. (6.11) and (6.12) as an initial value problem, we introduce the Laplace transform of F_1 with respect to t:

$$\bar{F}_1(u,p) = \int_0^{\infty} F_1(u,t)\,e^{-p\,t}\,dt. \tag{6.13}$$

If the growth of F_1 with t is no faster than exponential then the above integral converges and defines $\bar{F}_1$ as an analytic function of p, provided that the real part of p is sufficiently large.

Noting that the Laplace transform of $\partial F_1/\partial t$ is $p\,\bar{F}_1 - F_1(u, t=0)$ (as is easily shown by integration by parts), we can Laplace transform Eqs. (6.11) and (6.12) to obtain

$$p\,\bar{F}_1 + \mathrm{i}\,k\,u\,\bar{F}_1 = \frac{e}{m_e}\,\bar{E}\,\frac{\partial F_0}{\partial u} + F_1(u, t=0), \tag{6.14}$$

and

$$\mathrm{i}\,k\,\bar{E} = -\frac{e}{\epsilon_0}\int_{-\infty}^{\infty} \bar{F}_1(u)\,du, \tag{6.15}$$

respectively. The above two equations can be combined to give

$$\mathrm{i}\,k\,\bar{E} = -\frac{e}{\epsilon_0}\int_{-\infty}^{\infty}\left[\frac{e}{m_e}\,\bar{E}\,\frac{\partial F_0/\partial u}{p+\mathrm{i}\,k\,u} + \frac{F_1(u, t=0)}{p+\mathrm{i}\,k\,u}\right]du, \tag{6.16}$$

yielding

$$\bar{E} = -\frac{(e/\epsilon_0)}{\mathrm{i}\,k\,\epsilon(k,p)}\int_{-\infty}^{\infty}\frac{F_1(u, t=0)}{p+\mathrm{i}\,k\,u}\,du, \tag{6.17}$$

where

$$\epsilon(k,p) = 1 + \frac{e^2}{\epsilon_0\,m_e\,k}\int_{-\infty}^{\infty}\frac{\partial F_0/\partial u}{\mathrm{i}\,p - k\,u}\,du. \tag{6.18}$$

The function $\epsilon(k,p)$ is known as the *plasma dielectric function*. Note that if p is replaced by $-\mathrm{i}\,\omega$ then the dielectric function becomes equivalent to the left-hand side of Eq. (6.9). However, since p possesses a positive real part, the above integral is well defined.

The Laplace transform of the distribution function is written

$$\bar{F}_1 = \frac{e}{m_e}\,\bar{E}\,\frac{\partial F_0/\partial u}{p+\mathrm{i}\,k\,u} + \frac{F_1(u, t=0)}{p+\mathrm{i}\,k\,u}, \tag{6.19}$$

or

$$\bar{F}_1 = -\frac{e^2}{\epsilon_0\,m_e\,\mathrm{i}\,k}\,\frac{\partial F_0/\partial u}{\epsilon(k,p)\,(p+\mathrm{i}\,k\,u)}\int_{-\infty}^{\infty}\frac{F_1(u', t=0)}{p+\mathrm{i}\,k\,u'}\,du' + \frac{F_1(u, t=0)}{p+\mathrm{i}\,k\,u}. \tag{6.20}$$

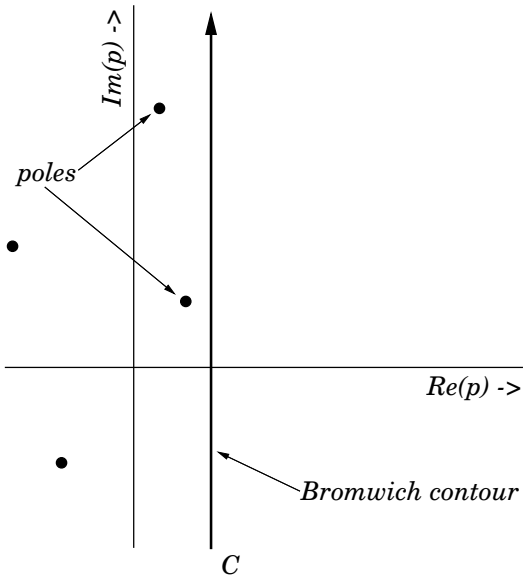

Figure 6.1: *The Bromwich contour.*

Having found the Laplace transforms of the electric field and the perturbed distribution function, we must now invert them to obtain E and F_1 as functions of time. The inverse Laplace transform of the distribution function is given by

$$F_1(u,t) = \frac{1}{2\pi i}\int_C \bar{F}_1(u,p)\, e^{pt}\, dp, \tag{6.21}$$

where C, the so-called *Bromwich contour*, is a contour running parallel to the imaginary axis, and lying to the right of all singularities of $\bar{F}_1$ in the complex-p plane (see Fig. 6.1). There is an analogous expression for the parallel electric field, $E(t)$.

Rather than trying to obtain a general expression for $F_1(u,t)$, from Eqs. (6.20) and (6.21), we shall concentrate on the behaviour of the perturbed distribution function at *large* times. Looking at Fig. 6.1, we note that if $\bar{F}_1(u,p)$ has only a finite number of simple poles in the region $\mathrm{Re}(p) > -\sigma$, then we may deform the contour as shown in Fig. 6.2, with a loop around each of the singularities. A pole at p_0 gives a contribution going as $e^{p_0 t}$, whilst the vertical part of the contour goes as $e^{-\sigma t}$. For sufficiently long times this latter contribution is negligible, and the behaviour is dominated by contributions from the poles furthest to the right.

Equations (6.17)–(6.20) all involve integrals of the form

$$\int_{-\infty}^{\infty} \frac{G(u)}{u - i\,p/k}\, du, \tag{6.22}$$

which become singular as p approaches the imaginary axis. In order to distort the contour C, in the manner shown in Fig. 31, we need to continue these integrals smoothly across the imaginary p-axis. By virtue of the way in which the Laplace transform was originally

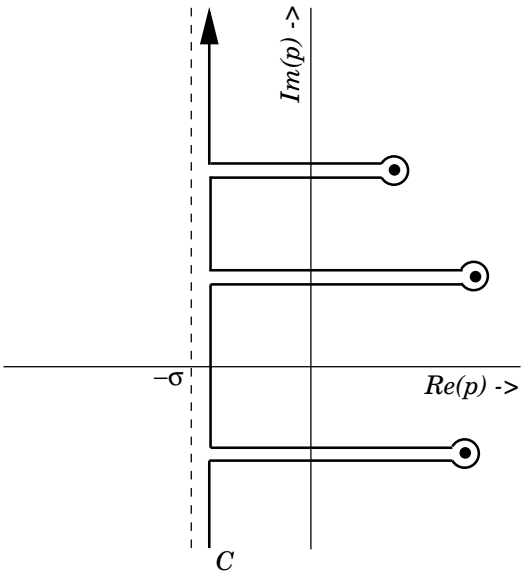

Figure 6.2: *The distorted Bromwich contour.*

defined, for $\mathrm{Re}(p)$ sufficiently large, the appropriate way to do this is to take the values of these integrals when p is in the right-hand half-plane, and find the analytic continuation into the left-hand half-plane.

If $G(u)$ is sufficiently well-behaved that it can be continued off the real axis as an analytic function of a complex variable u then the continuation of (6.22) as the singularity crosses the real axis in the complex u-plane, from the upper to the lower half-plane, is obtained by letting the singularity take the contour with it, as shown in Fig. 6.3.

Note that the ability to deform the contour C into that of Fig. 6.2, and find a dominant contribution to $E(t)$ and $F_1(u,t)$ from a few poles, depends on $F_0(u)$ and $F_1(u,t=0)$ having *smooth* enough velocity dependences that the integrals appearing in Eqs. (6.17)–(6.20) can be continued sufficiently far into the left-hand half of the complex p-plane.

If we consider the electric field given by the inversion of Eq. (6.17), we see that its behaviour at large times is dominated by the zero of $\epsilon(k,p)$ which lies furthest to the right in the complex p-plane. According to Eqs. (6.20) and (6.21), F_1 has a similar contribution, as well as a contribution going as e^{-ikut}. Thus, for sufficiently long times after the initiation of the wave, the electric field depends only on the positions of the roots of $\epsilon(k,p)=0$ in the complex p-plane. The distribution function has a corresponding contribution from the poles, as well as a component going as e^{-ikut}. For large times, the latter component of the distribution function is a rapidly oscillating function of velocity, and its contribution to the charge density, obtained by integrating over u, is negligible.

As we have already noted, the function $\epsilon(k,p)$ is equivalent to the left-hand side of Eq. (6.9), provided that p is replaced by $-i\omega$. Thus, the dispersion relation, (6.9), obtained via Fourier transformation of the Vlasov equation, gives the correct behaviour at large times as long as the singular integral is treated correctly. Adapting the procedure

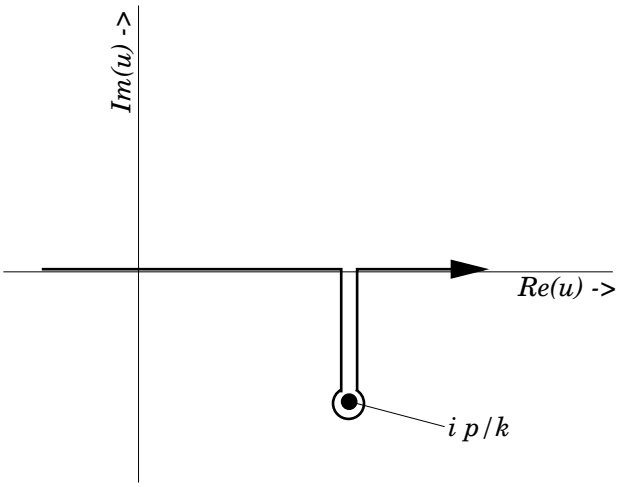

Figure 6.3: *The Bromwich contour for Landau damping.*

which we found using the variable p, we see that the integral is defined as it is written for $\mathrm{Im}(\omega) > 0$, and analytically continued, by deforming the contour of integration in the u-plane (as shown in Fig. 6.3), into the region $\mathrm{Im}(\omega) < 0$. The simplest way to remember how to do the analytic continuation is to note that the integral is continued from the part of the ω-plane corresponding to *growing* perturbations, to that corresponding to *damped* perturbations. Once we know this rule, we can obtain kinetic dispersion relations in a fairly direct manner via Fourier transformation of the Vlasov equation, and there is no need to attempt the more complicated Laplace transform solution.

In Sect. 4, where we investigated the cold-plasma dispersion relation, we found that for any given k there were a finite number of values of ω, say ω_1, ω_2, $\cdots$, and a general solution was a linear superposition of functions varying in time as $e^{-i\,\omega_1\,t}$, $e^{-i\,\omega_2\,t}$, *etc*. This set of values of ω is called the *spectrum*, and the cold-plasma equations yield a *discrete* spectrum. On the other hand, in the kinetic problem we obtain contributions to the distribution function going as $e^{-i\,k\,u\,t}$, with u taking any real value. All of the mathematical difficulties of the kinetic problem arise from the existence of this *continuous spectrum*. At short times, the behaviour is very complicated, and depends on the details of the initial perturbation. It is only asymptotically that a mode varying as $e^{-i\,\omega\,t}$ is obtained, with ω determined by a dispersion relation which is solely a function of the unperturbed state. As we have seen, the emergence of such a mode depends on the initial velocity disturbance being sufficiently smooth.

Suppose, for the sake of simplicity, that the background plasma state is a Maxwellian distribution. Working in terms of ω, rather than p, the kinetic dispersion relation for electrostatic waves takes the form

$$\epsilon(k,\omega) = 1 + \frac{e^2}{\epsilon_0\, m_e\, k}\int_{-\infty}^{\infty} \frac{\partial F_0/\partial u}{\omega - k\,u}\,du = 0, \tag{6.23}$$

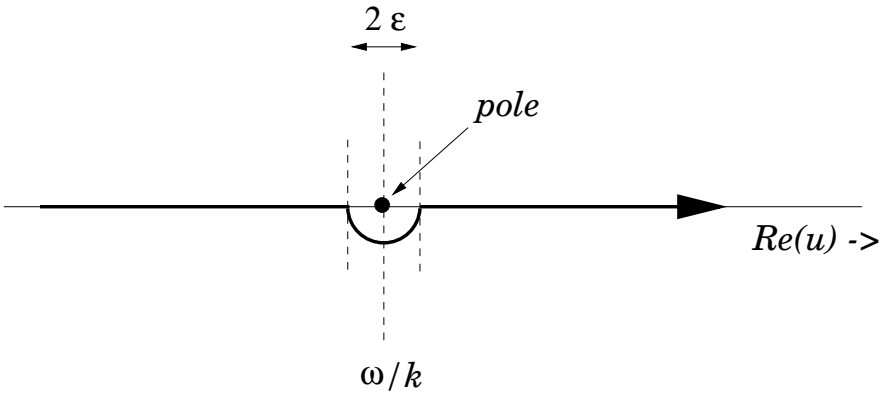

Figure 6.4: *Integration path about a pole.*

where

$$F_0(u) = \frac{n}{(2\pi T_e/m_e)^{1/2}}\,\exp(-m_e\,u^2/2\,T_e). \tag{6.24}$$

Suppose that, to a first approximation, ω is real. Letting ω tend to the real axis from the domain $\mathrm{Im}(\omega) > 0$, we obtain

$$\int_{-\infty}^{\infty} \frac{\partial F_0/\partial u}{\omega - k\,u}\,du = P\int_{-\infty}^{\infty} \frac{\partial F_0/\partial u}{\omega - k\,u}\,du - \frac{i\,\pi}{k}\left(\frac{\partial F_0}{\partial u}\right)_{u=\omega/k}, \tag{6.25}$$

where P denotes the *principal part* of the integral. The origin of the two terms on the right-hand side of the above equation is illustrated in Fig. 6.4. The first term—the principal part—is obtained by removing an interval of length $2\,\epsilon$, symmetrical about the pole, $u = \omega/k$, from the range of integration, and then letting $\epsilon \rightarrow 0$. The second term comes from the small semi-circle linking the two halves of the principal part integral. Note that the semi-circle deviates *below* the real u-axis, rather than above, because the integral is calculated by letting the pole approach the axis from the *upper* half-plane in u-space.

Suppose that k is sufficiently small that $\omega \gg k\,u$ over the range of u where $\partial F_0/\partial u$ is non-negligible. It follows that we can expand the denominator of the principal part integral in a Taylor series:

$$\frac{1}{\omega - k\,u} \simeq \frac{1}{\omega}\left(1 + \frac{k\,u}{\omega} + \frac{k^2\,u^2}{\omega^2} + \frac{k^3 u^3}{\omega^3} + \cdots\right). \tag{6.26}$$

Integrating the result term by term, and remembering that $\partial F_0/\partial u$ is an odd function, Eq. (6.23) reduces to

$$1 - \frac{{\omega_p}^2}{\omega^2} - 3\,k^2\,\frac{T_e\,{\omega_p}^2}{m_e\,\omega^4} - \frac{e^2}{\epsilon_0\,m_e}\frac{i\,\pi}{k^2}\left(\frac{\partial F_0}{\partial u}\right)_{u=\omega/k} \simeq 0, \tag{6.27}$$

where $\omega_p = \sqrt{n\,e^2/\epsilon_0\,m_e}$ is the electron plasma frequency. Equating the real part of the above expression to zero yields

$$\omega^2 \simeq {\omega_p}^2(1 + 3\,k^2\,\lambda_D^2), \tag{6.28}$$

where $\lambda_D = \sqrt{T_e/m_e\,\omega_p^2}$ is the Debye length, and it is assumed that $k\lambda_D \ll 1$. We can regard the imaginary part of ω as a small perturbation, and write $\omega = \omega_0 + \delta\omega$, where ω_0 is the root of Eq. (6.28). It follows that

$$2\,\omega_0\,\delta\omega \simeq {\omega_0}^2\,\frac{e^2}{\epsilon_0\,m_e}\frac{i\,\pi}{k^2}\left(\frac{\partial F_0}{\partial u}\right)_{u=\omega/k}, \tag{6.29}$$

and so

$$\delta\omega \simeq \frac{i\,\pi}{2}\frac{e^2\,\omega_p}{\epsilon_0\,m_e\,k^2}\left(\frac{\partial F_0}{\partial u}\right)_{u=\omega/k}, \tag{6.30}$$

giving

$$\delta\omega \simeq -\frac{i}{2}\sqrt{\frac{\pi}{2}}\frac{\omega_p}{(k\lambda_D)^3}\exp\left[-\frac{1}{2\,(k\lambda_D)^2}\right]. \tag{6.31}$$

If we compare the above results with those for a cold-plasma, where the dispersion relation for an electrostatic plasma wave was found to be simply $\omega^2 = {\omega_p}^2$, we see, firstly, that ω now depends on k, according to Eq. (6.28), so that in a warm plasma the electrostatic plasma wave is a *propagating* mode, with a non-zero group velocity. Secondly, we now have an imaginary part to ω, given by Eq. (6.31), corresponding, since it is negative, to the *damping* of the wave in time. This damping is generally known as *Landau damping*. If $k\lambda_D \ll 1$ (*i.e.*, if the wave-length is much larger than the Debye length) then the imaginary part of ω is small compared to the real part, and the wave is only lightly damped. However, as the wave-length becomes comparable to the Debye length, the imaginary part of ω becomes comparable to the real part, and the damping becomes strong. Admittedly, the approximate solution given above is not very accurate in the short wave-length case, but it is sufficient to indicate the existence of very strong damping.

There are *no* dissipative effects included in the collisionless Vlasov equation. Thus, it can easily be verified that if the particle velocities are reversed at any time then the solution up to that point is simply reversed in time. At first sight, this reversible behaviour does not seem to be consistent with the fact that an initial perturbation dies out. However, we should note that it is only the electric field which decays. The distribution function contains an undamped term going as $e^{-iku\,t}$. Furthermore, the decay of the electric field depends on there being a sufficiently smooth initial perturbation in velocity space. The presence of the $e^{-iku\,t}$ term means that as time advances the velocity space dependence of the perturbation becomes more and more convoluted. It follows that if we reverse the velocities after some time then we are not starting with a smooth distribution. Under these circumstances, there is no contradiction in the fact that under time reversal the electric field will grow initially, until the smooth initial state is recreated, and subsequently decay away.

6.3 Physics of Landau Damping

We have explained Landau damping in terms of mathematics. Let us now consider the physical explanation for this effect. The motion of a charged particle situated in a one-

dimensional electric field varying as $E_0 \exp[i(kx - \omega t)]$ is determined by

$$\frac{d^2x}{dt^2} = \frac{e}{m} E_0 e^{i(kx-\omega t)}. \tag{6.32}$$

Since we are dealing with a linearized theory in which the perturbation due to the wave is small, it follows that if the particle starts with velocity u_0 at position x_0 then we may substitute $x_0 + u_0 t$ for x in the electric field term. This is actually the position of the particle on its unperturbed trajectory, starting at $x = x_0$ at $t = 0$. Thus, we obtain

$$\frac{du}{dt} = \frac{e}{m} E_0 e^{i(kx_0+ku_0 t-\omega t)}, \tag{6.33}$$

which yields

$$u - u_0 = \frac{e}{m} E_0 \left[\frac{e^{i(kx_0+ku_0 t-\omega t)} - e^{ikx_0}}{i(ku_0 - \omega)}\right]. \tag{6.34}$$

As $ku_0 - \omega \rightarrow 0$, the above expression reduces to

$$u - u_0 = \frac{e}{m} E_0 t e^{ikx_0}, \tag{6.35}$$

showing that particles with u_0 close to ω/k, that is with velocity components along the x-axis close to the phase velocity of the wave, have velocity perturbations which *grow* in time. These so-called *resonant particles* gain energy from, or lose energy to, the wave, and are responsible for the damping. This explains why the damping rate, given by Eq. (6.30), depends on the slope of the distribution function calculated at $u = \omega/k$. The remainder of the particles are non-resonant, and have an oscillatory response to the wave field.

To understand why energy should be transferred from the electric field to the resonant particles requires more detailed consideration. Whether the speed of a resonant particle increases or decreases depends on the phase of the wave at its initial position, and it is not the case that all particles moving slightly faster than the wave lose energy, whilst all particles moving slightly slower than the wave gain energy. Furthermore, the density perturbation is out of phase with the wave electric field, so there is no initial wave generated excess of particles gaining or losing energy. However, if we consider those particles which start off with velocities slightly above the phase velocity of the wave then if they gain energy they move away from the resonant velocity whilst if they lose energy they approach the resonant velocity. The result is that the particles which lose energy interact more effectively with the wave, and, on average, there is a transfer of energy from the particles to the electric field. Exactly the opposite is true for particles with initial velocities lying just below the phase velocity of the wave. In the case of a Maxwellian distribution there are more particles in the latter class than in the former, so there is a net transfer of energy from the electric field to the particles: *i.e.*, the electric field is damped. In the limit as the wave amplitude tends to zero, it is clear that the gradient of the distribution function at the wave speed is what determines the damping rate.

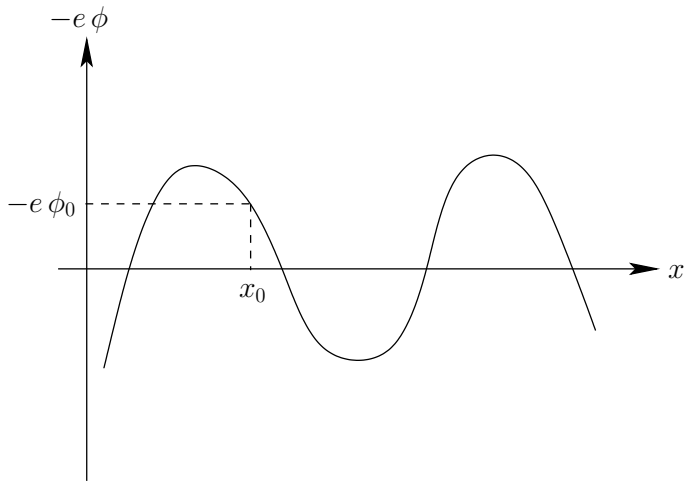

Figure 6.5: *Wave-particle interaction.*

It is of some interest to consider the limitations of the above result, in terms of the magnitude of the initial electric field above which it is seriously in error and *nonlinear* effects become important. The basic requirement for the validity of the linear result is that a resonant particle should maintain its position relative to the phase of the electric field over a sufficiently long time for the damping to take place. To obtain a condition that this be the case, let us consider the problem in the frame of reference in which the wave is at rest, and the potential $-e\,\phi$ seen by an electron is as sketched in Fig. 6.5.

If the electron starts at rest (*i.e.*, in resonance with the wave) at x_0 then it begins to move towards the potential minimum, as shown. The time for the electron to shift its position relative to the wave may be estimated as the period with which it bounces back and forth in the potential well. Near the bottom of the well the equation of motion of the electron is written

$$\frac{d^2x}{dt^2} = -\frac{e}{m_e}\,k^2\,x\,\phi_0, \tag{6.36}$$

where k is the wave-number, and so the bounce time is

$$\tau_b \sim 2\pi\sqrt{\frac{m_e}{e\,k^2\,\phi_0}} = 2\pi\sqrt{\frac{m_e}{e\,k\,E_0}}, \tag{6.37}$$

where E_0 is the amplitude of the electric field. We may expect the wave to damp according to linear theory if the bounce time, τ_b, given above, is much greater than the damping time. Since the former varies inversely with the square root of the electric field amplitude, whereas the latter is amplitude independent, this criterion gives us an estimate of the maximum allowable initial perturbation which is consistent with linear damping.

If the initial amplitude is large enough for the resonant electrons to bounce back and forth in the potential well a number of times before the wave is damped, then it can be demonstrated that the result to be expected is a *non-monotonic* decrease in the amplitude

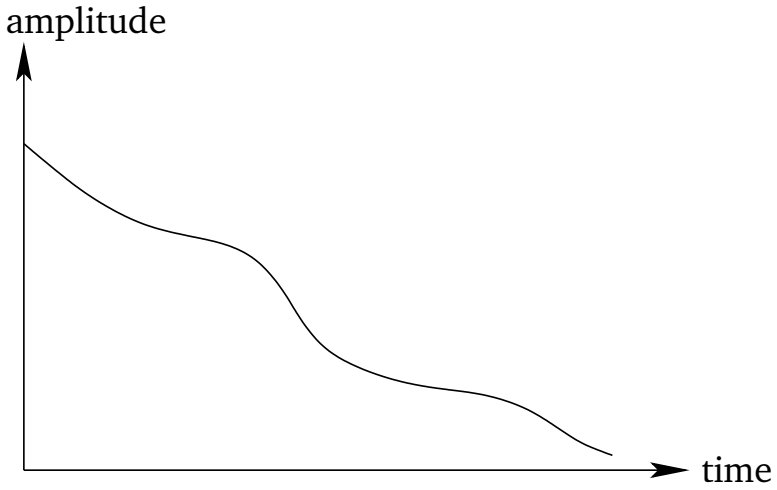

Figure 6.6: *Nonlinear Landau damping.*

of the electric field, as shown in Fig. 6.6. The period of the amplitude oscillations is similar to the bounce time, τ_b.

6.4 Plasma Dispersion Function

If the unperturbed distribution function, F_0, appearing in Eq. (6.23), is a Maxwellian then it is readily seen that, with a suitable scaling of the variables, the dispersion relation for electrostatic plasma waves can be expressed in terms of the function

$$Z(\zeta) = \pi^{-1/2} \int_{-\infty}^{\infty} \frac{e^{-t^2}}{t-\zeta}\, dt, \tag{6.38}$$

which is defined as it is written for $\mathrm{Im}(\zeta) > 0$, and is analytically continued for $\mathrm{Im}(\zeta) \leq 0$. This function is known as the *plasma dispersion function*, and very often crops up in problems involving small-amplitude waves propagating through warm plasmas. Incidentally, $Z(\zeta)$ is the Hilbert transform of a Gaussian.

In view of the importance of the plasma dispersion function, and its regular occurrence in the literature of plasma physics, let us briefly examine its main properties. We first of all note that if we differentiate $Z(\zeta)$ with respect to ζ we obtain

$$Z'(\zeta) = \pi^{-1/2} \int_{-\infty}^{\infty} \frac{e^{-t^2}}{(t-\zeta)^2}\, dt, \tag{6.39}$$

which yields, on integration by parts,

$$Z'(\zeta) = -\pi^{-1/2} \int_{-\infty}^{\infty} \frac{2\,t}{t-\zeta}\, e^{-t^2}\, dt = -2\,[1 + \zeta\, Z]. \tag{6.40}$$

If we let ζ tend to zero from the upper half of the complex plane, we get

$$Z(0) = \pi^{-1/2}\,\mathrm{P}\int_{-\infty}^{\infty} \frac{\mathrm{e}^{-t^2}}{t}\,\mathrm{d}t + \mathrm{i}\,\pi^{1/2} = \mathrm{i}\,\pi^{1/2}. \tag{6.41}$$

Note that the principle part integral is zero because its integrand is an odd function of t.

Integrating the linear differential equation (6.40), which possesses an integrating factor e^{ζ^2}, and using the boundary condition (6.41), we obtain an alternative expression for the plasma dispersion function:

$$Z(\zeta) = \mathrm{e}^{-\zeta^2}\left(\mathrm{i}\,\pi^{1/2} - 2\int_0^{\zeta} \mathrm{e}^{x^2}\,\mathrm{d}x\right). \tag{6.42}$$

Making the substitution $t = \mathrm{i}\,x$ in the integral, and noting that

$$\int_{-\infty}^{0} \mathrm{e}^{-t^2}\,\mathrm{d}t = \frac{\pi^{1/2}}{2}, \tag{6.43}$$

we finally arrive at the expression

$$Z(\zeta) = 2\,\mathrm{i}\,\mathrm{e}^{-\zeta^2}\int_{-\infty}^{\mathrm{i}\,\zeta} \mathrm{e}^{-t^2}\,\mathrm{d}t. \tag{6.44}$$

This formula, which relates the plasma dispersion function to an error function of imaginary argument, is valid for all values of ζ.

For small ζ we have the expansion

$$Z(\zeta) = \mathrm{i}\,\pi^{1/2}\,\mathrm{e}^{-\zeta^2} - 2\,\zeta\left[1 - \frac{2\,\zeta^2}{3} + \frac{4\,\zeta^4}{15} - \frac{8\,\zeta^6}{105} + \cdots\right]. \tag{6.45}$$

For large ζ, where $\zeta = x + \mathrm{i}\,y$, the asymptotic expansion for $x > 0$ is written

$$Z(\zeta) \sim \mathrm{i}\,\pi^{1/2}\,\sigma\,\mathrm{e}^{-\zeta^2} - \zeta^{-1}\left[1 + \frac{1}{2\,\zeta^2} + \frac{3}{4\,\zeta^4} + \frac{15}{8\,\zeta^6} + \cdots\right]. \tag{6.46}$$

Here,

$$\sigma = \begin{cases} 0 & y > 1/|x| \\ 1 & |y| < 1/|x| \\ 2 & y < -1/|x| \end{cases}. \tag{6.47}$$

In deriving our expression for the Landau damping rate we have, in effect, used the first few terms of the above asymptotic expansion.

The properties of the plasma dispersion function are specified in exhaustive detail in a well-known book by Fried and Conte.[2]

[2]B.D. Fried, and S.D. Conte, *The Plasma Dispersion Function* (Academic Press, New York NY, 1961.)

6.5 Ion Sound Waves

If we now take ion dynamics into account then the dispersion relation (6.23), for electrostatic plasma waves, generalizes to

$$1+\frac{e^2}{\epsilon_0\,m_e\,k}\int_{-\infty}^{\infty}\frac{\partial F_{0\,e}/\partial u}{\omega-k\,u}\,du+\frac{e^2}{\epsilon_0\,m_i\,k}\int_{-\infty}^{\infty}\frac{\partial F_{0\,i}/\partial u}{\omega-k\,u}\,du=0: \tag{6.48}$$

i.e., we simply add an extra term for the ions which has an analogous form to the electron term. Let us search for a wave with a phase velocity, ω/k, which is much less than the electron thermal velocity, but much greater than the ion thermal velocity. We may assume that $\omega\gg k\,u$ for the ion term, as we did previously for the electron term. It follows that, to lowest order, this term reduces to $-\omega_{p\,i}^2/\omega^2$. Conversely, we may assume that $\omega\ll k\,u$ for the electron term. Thus, to lowest order we may neglect ω in the velocity space integral. Assuming $F_{0\,e}$ to be a Maxwellian with temperature T_e, the electron term reduces to

$$\frac{\omega_{p\,e}^2}{k^2}\frac{m_e}{T_e}=\frac{1}{(k\,\lambda_D)^2}. \tag{6.49}$$

Thus, to a first approximation, the dispersion relation can be written

$$1+\frac{1}{(k\,\lambda_D)^2}-\frac{\omega_{p\,i}^2}{\omega^2}=0, \tag{6.50}$$

giving

$$\omega^2=\frac{\omega_{p\,i}^2\,k^2\,\lambda_D^2}{1+k^2\,\lambda_D^2}=\frac{T_e}{m_i}\frac{k^2}{1+k^2\,\lambda_D^2}. \tag{6.51}$$

For $k\,\lambda_D\ll 1$, we have $\omega=(T_e/m_i)^{1/2}\,k$, a dispersion relation which is like that of an ordinary sound wave, with the pressure provided by the electrons, and the inertia by the ions. As the wave-length is reduced towards the Debye length, the frequency levels off and approaches the ion plasma frequency.

Let us check our original assumptions. In the long wave-length limit, we see that the wave phase velocity $(T_e/m_i)^{1/2}$ is indeed much less than the electron thermal velocity [by a factor $(m_e/m_i)^{1/2}$], but that it is only much greater than the ion thermal velocity if the ion temperature, T_i, is much less than the electron temperature, T_e. In fact, if $T_i\ll T_e$ then the wave phase velocity can lie on almost flat portions of the ion and electron distribution functions, as shown in Fig. 6.7, implying that the wave is subject to very little Landau damping. Indeed, an ion sound wave can only propagate a distance of order its wave-length without being strongly damped provided that T_e is at least five to ten times greater than T_i.

Of course, it is possible to obtain the ion sound wave dispersion relation, $\omega^2/k^2=T_e/m_i$, using fluid theory. The kinetic treatment used here is an improvement on the fluid theory to the extent that no equation of state is assumed, and it makes it clear to us that ion sound waves are subject to strong Landau damping (*i.e.*, they cannot be considered normal modes of the plasma) unless $T_e\gg T_i$.

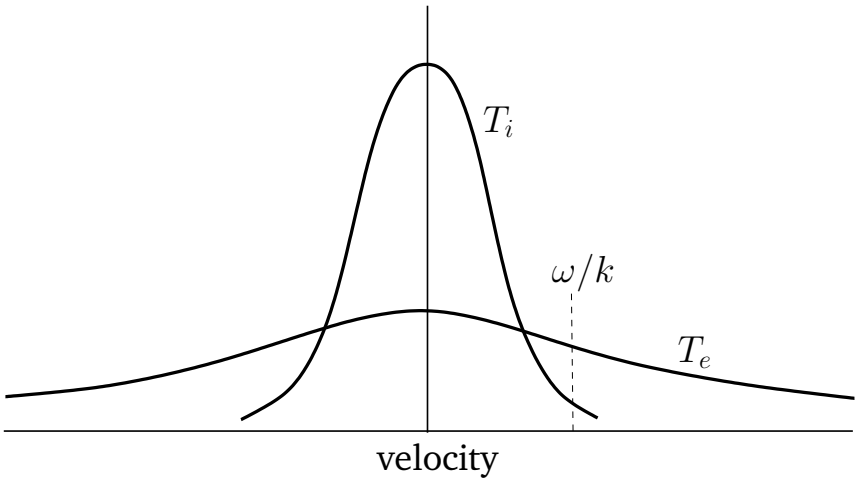

Figure 6.7: *Ion and electron distribution functions with* $T_i \ll T_e$.

6.6 Waves in Magnetized Plasmas

Consider waves propagating through a plasma placed in a uniform magnetic field, $\mathbf{B}_0$. Let us take the perturbed magnetic field into account in our calculations, in order to allow for electromagnetic, as well as electrostatic, waves. The linearized Vlasov equation takes the form

$$\frac{\partial f_1}{\partial t} + \mathbf{v}\cdot\nabla f_1 + \frac{e}{m}\,(\mathbf{v}\times\mathbf{B}_0)\cdot\nabla_v f_1 = -\frac{e}{m}\,(\mathbf{E}+\mathbf{v}\times\mathbf{B})\cdot\nabla_v f_0 \tag{6.52}$$

for both ions and electrons, where $\mathbf{E}$ and $\mathbf{B}$ are the perturbed electric and magnetic fields, respectively. Likewise, f_1 is the perturbed distribution function, and f_0 the equilibrium distribution function.

In order to have an equilibrium state at all, we require that

$$(\mathbf{v}\times\mathbf{B}_0)\cdot\nabla_v f_0 = 0. \tag{6.53}$$

Writing the velocity, $\mathbf{v}$, in cylindrical polar coordinates, $(v_\perp, \theta, v_z)$, aligned with the equilibrium magnetic field, the above expression can easily be shown to imply that $\partial f_0/\partial\theta = 0$: *i.e.*, f_0 is a function only of $v_\perp$ and v_z.

Let the trajectory of a particle be $\mathbf{r}(t)$, $\mathbf{v}(t)$. In the unperturbed state

$$\frac{d\mathbf{r}}{dt} = \mathbf{v}, \tag{6.54}$$

$$\frac{d\mathbf{v}}{dt} = \frac{e}{m}\,(\mathbf{v}\times\mathbf{B}_0). \tag{6.55}$$

It follows that Eq. (6.52) can be written

$$\frac{Df_1}{Dt} = -\frac{e}{m}\,(\mathbf{E}+\mathbf{v}\times\mathbf{B})\cdot\nabla_v f_0, \tag{6.56}$$

where Df_1/Dt is the total rate of change of f_1, following the unperturbed trajectories. Under the assumption that f_1 vanishes as $t \rightarrow -\infty$, the solution to Eq. (6.56) can be written

$$f_1(\mathbf{r},\mathbf{v},t) = -\frac{e}{m}\int_{-\infty}^{t} [\mathbf{E}(\mathbf{r}',t') + \mathbf{v}'\times\mathbf{B}(\mathbf{r}',t')]\cdot\nabla_v f_0(\mathbf{v}')\,dt', \tag{6.57}$$

where $(\mathbf{r}',\ \mathbf{v}')$ is the unperturbed trajectory which passes through the point $(\mathbf{r},\ \mathbf{v})$ when $t' = t$.

It should be noted that the above method of solution is valid for any set of equilibrium electromagnetic fields, not just a uniform magnetic field. However, in a uniform magnetic field the unperturbed trajectories are merely helices, whilst in a general field configuration it is difficult to find a closed form for the particle trajectories which is sufficiently simple to allow further progress to be made.

Let us write the velocity in terms of its Cartesian components:

$$\mathbf{v} = (v_\perp\cos\theta, v_\perp\sin\theta, v_z). \tag{6.58}$$

It follows that

$$\mathbf{v}' = (v_\perp\cos[\Omega\,(t-t')+\theta\,], v_\perp\sin[\Omega\,(t-t')+\theta\,], v_z)\,, \tag{6.59}$$

where $\Omega = e\,B_0/m$ is the cyclotron frequency. The above expression can be integrated to give

$$x'-x = -\frac{v_\perp}{\Omega}\,(\sin[\Omega\,(t-t')+\theta\,]-\sin\theta)\,, \tag{6.60}$$

$$y'-y = \frac{v_\perp}{\Omega}\,(\cos[\Omega\,(t-t')+\theta\,]-\cos\theta)\,, \tag{6.61}$$

$$z'-z = v_z\,(t'-t). \tag{6.62}$$

Note that both $v_\perp$ and v_z are constants of the motion. This implies that $f_0(\mathbf{v}') = f_0(\mathbf{v})$, because f_0 is only a function of $v_\perp$ and v_z. Since $v_\perp = ({v_x'}^2 + {v_y'}^2)^{1/2}$, we can write

$$\frac{\partial f_0}{\partial v_x'} = \frac{\partial v_\perp}{\partial v_x'}\frac{\partial f_0}{\partial v_\perp} = \frac{v_x'}{v_\perp}\frac{\partial f_0}{\partial v_\perp} = \cos\,[\Omega\,(t'-t)+\theta\,]\,\frac{\partial f_0}{\partial v_\perp}, \tag{6.63}$$

$$\frac{\partial f_0}{\partial v_y'} = \frac{\partial v_\perp}{\partial v_y'}\frac{\partial f_0}{\partial v_\perp} = \frac{v_y'}{v_\perp}\frac{\partial f_0}{\partial v_\perp} = \sin\,[\Omega\,(t'-t)+\theta\,]\,\frac{\partial f_0}{\partial v_\perp}, \tag{6.64}$$

$$\frac{\partial f_0}{\partial v_z'} = \frac{\partial f_0}{\partial v_z}. \tag{6.65}$$

Let us assume an $\exp[\,i\,(\mathbf{k}\cdot\mathbf{r}-\omega\,t)]$ dependence of all perturbed quantities, with $\mathbf{k}$ lying in the x-z plane. Equation (6.57) yields

$$f_1 = -\frac{e}{m}\int_{-\infty}^{t}\left[(E_x + v_y'\,B_z - v_z'\,B_y)\,\frac{\partial f_0}{\partial v_x'} + (E_y + v_z'\,B_x - v_x'\,B_z)\,\frac{\partial f_0}{\partial v_y'}\right.$$

$$+(E_z + v'_x\,B_y - v'_y\,B_x)\,\frac{\partial f_0}{\partial v'_z}\Bigg]\,\exp\left[\,i\{\mathbf{k}\cdot(\mathbf{r}'-\mathbf{r}) - \omega\,(t'-t)\}\right]\,dt'. \tag{6.66}$$

Making use of Eqs. (6.59)–(6.65), and the identity

$$e^{\,i\,a\,\sin x} \equiv \sum_{n=-\infty}^{\infty} J_n(a)\,e^{\,i\,n\,x}, \tag{6.67}$$

Eq. (6.66) gives

$$\begin{aligned} f_1 \;=\; & -\frac{e}{m}\int_{-\infty}^{t}\Bigg[(E_x - v_z\,B_y)\,\cos\chi\,\frac{\partial f_0}{\partial v_\perp} + (E_y + v_z\,B_x)\,\sin\chi\,\frac{\partial f_0}{\partial v_\perp} \\ & +(E_z + v_\perp\,B_y\,\cos\chi - v_\perp\,B_x\,\sin\chi)\,\frac{\partial f_0}{\partial v_z}\Bigg]\,\sum_{n,m=-\infty}^{\infty} J_n\left(\frac{k_\perp\,v_\perp}{\Omega}\right)\,J_m\left(\frac{k_\perp\,v_\perp}{\Omega}\right) \\ & \times\exp\{\,i\,[(n\,\Omega + k_z\,v_z - \omega)\,(t'-t) + (m-n)\,\theta\,]\,\}\,dt', \end{aligned} \tag{6.68}$$

where

$$\chi = \Omega\,(t-t') + \theta. \tag{6.69}$$

Maxwell's equations yield

$$\mathbf{k}\times\mathbf{E} \;=\; \omega\,\mathbf{B}, \tag{6.70}$$

$$\mathbf{k}\times\mathbf{B} \;=\; -i\,\mu_0\,\mathbf{j} - \frac{\omega}{c^2}\,\mathbf{E} = -\frac{\omega}{c^2}\,\mathbf{K}\cdot\mathbf{E}, \tag{6.71}$$

where $\mathbf{j}$ is the perturbed current, and $\mathbf{K}$ is the dielectric permittivity tensor introduced in Sect. 4.2. It follows that

$$\mathbf{K}\cdot\mathbf{E} = \mathbf{E} + \frac{i}{\omega\,\epsilon_0}\,\mathbf{j} = \mathbf{E} + \frac{i}{\omega\,\epsilon_0}\,\sum_s e_s\int\mathbf{v}\,f_{1\,s}\,d^3\mathbf{v}, \tag{6.72}$$

where $f_{1\,s}$ is the species-s perturbed distribution function.

After a great deal of rather tedious analysis, Eqs. (6.68) and (6.72) reduce to the following expression for the dielectric permittivity tensor:

$$K_{ij} = \delta_{ij} + \sum_s \frac{{e_s}^2}{\omega^2\,\epsilon_0\,m_s}\,\sum_{n=-\infty}^{\infty}\int\frac{S_{ij}}{\omega - k_z\,v_z - n\,\Omega_s}\,d^3\mathbf{v}, \tag{6.73}$$

where

$$S_{ij} = \begin{pmatrix} v_\perp\,(n\,J_n/a_s)^2\,U & i\,v_\perp\,(n/a_s)\,J_n\,J'_n\,U & v_\perp\,(n/a_s)\,J_n^2\,U \\ -i\,v_\perp\,(n/a_s)\,J_n\,J'_n\,U & v_\perp\,{J'_n}^2\,U & -i\,v_\perp\,J_n\,J'_n\,W \\ v_z\,(n/a_s)\,J_n^2\,U & i\,v_z\,J_n\,J'_n\,U & v_z\,J_n^2\,W \end{pmatrix}, \tag{6.74}$$

and

$$U = (\omega - k_z v_z)\frac{\partial f_{0s}}{\partial v_\perp} + k_z v_\perp \frac{\partial f_{0s}}{\partial v_z}, \tag{6.75}$$

$$W = \frac{n\Omega_s v_z}{v_\perp}\frac{\partial f_{0s}}{\partial v_\perp} + (\omega - n\Omega_s)\frac{\partial f_{0s}}{\partial v_z}, \tag{6.76}$$

$$a_s = \frac{k_\perp v_\perp}{\Omega_s}. \tag{6.77}$$

The argument of the Bessel functions is a_s. In the above, $'$ denotes differentiation with respect to argument.

The dielectric tensor (6.73) can be used to investigate the properties of waves in just the same manner as the cold plasma dielectric tensor (4.36) was used in Sect. 4. Note that our expression for the dielectric tensor involves singular integrals of a type similar to those encountered in Sect. 6.2. In principle, this means that we ought to treat the problem as an initial value problem. Fortunately, we can use the insights gained in our investigation of the simpler unmagnetized electrostatic wave problem to recognize that the appropriate way to treat the singular integrals is to evaluate them as written for $\mathrm{Im}(\omega) > 0$, and by analytic continuation for $\mathrm{Im}(\omega) \leq 0$.

For Maxwellian distribution functions, we can explicitly perform the velocity space integral in Eq. (6.73), making use of the identity

$$\int_0^\infty x\, J_n^2(s\,x)\, e^{-x^2}\, dx = \frac{e^{-s^2/2}}{2} I_n(s^2/2), \tag{6.78}$$

where I_n is a modified Bessel function. We obtain

$$K_{ij} = \delta_{ij} + \sum_s \frac{\omega_{ps}^2}{\omega}\left(\frac{m_s}{2T_s}\right)^{1/2}\frac{e^{-\lambda_s}}{k_z}\sum_{n=-\infty}^{\infty} T_{ij}, \tag{6.79}$$

where

$$T_{ij} = \begin{pmatrix} n^2 I_n Z/\lambda_s & i n (I_n' - I_n) Z & -n I_n Z'/(2\lambda_s)^{1/2} \\ -i n (I_n' - I_n) Z & (n^2 I_n/\lambda_s + 2\lambda_s I_n - 2\lambda_s I_n') Z & i\lambda_s^{1/2}(I_n' - I_n) Z'/2^{1/2} \\ -n I_n Z'/(2\lambda_s)^{1/2} & -i\lambda_s^{1/2}(I_n' - I_n) Z'/2^{1/2} & -I_n Z' \xi_n \end{pmatrix}. \tag{6.80}$$

Here, λ_s, which is the argument of the Bessel functions, is written

$$\lambda_s = \frac{T_s k_\perp^2}{m_s \Omega_s^2}, \tag{6.81}$$

whilst Z and Z' represent the plasma dispersion function and its derivative, both with argument

$$\xi_n = \frac{\omega - n\Omega_s}{k_z}\left(\frac{m_s}{2T_s}\right)^{1/2}. \tag{6.82}$$

Let us consider the cold plasma limit, $T_s \rightarrow 0$. It follows from Eqs. (6.81) and (6.82) that this limit corresponds to $\lambda_s \rightarrow 0$ and $\xi_n \rightarrow \infty$. From Eq. (6.46),

$$Z(\xi_n) \rightarrow -\frac{1}{\xi_n}, \tag{6.83}$$

$$Z'(\xi_n) \rightarrow \frac{1}{\xi_n^2} \tag{6.84}$$

as $\xi_n \rightarrow \infty$. Moreover,

$$I_n(\lambda_s) \rightarrow \left(\frac{\lambda_s}{2}\right)^{|n|} \tag{6.85}$$

as $\lambda_s \rightarrow 0$. It can be demonstrated that the only non-zero contributions to K_{ij}, in this limit, come from $n = 0$ and $n = \pm 1$. In fact,

$$K_{11} = K_{22} = 1 - \frac{1}{2}\sum_s \frac{\omega_{ps}^2}{\omega^2}\left(\frac{\omega}{\omega-\Omega_s}+\frac{\omega}{\omega+\Omega_s}\right), \tag{6.86}$$

$$K_{12} = -K_{21} = -\frac{i}{2}\sum_s \frac{\omega_{ps}^2}{\omega^2}\left(\frac{\omega}{\omega-\Omega_s}-\frac{\omega}{\omega+\Omega_s}\right), \tag{6.87}$$

$$K_{33} = 1 - \sum_s \frac{\omega_{ps}^2}{\omega^2}, \tag{6.88}$$

and $K_{13} = K_{31} = K_{23} = K_{32} = 0$. It is easily seen, from Sect. 4.3, that the above expressions are identical to those we obtained using the cold-plasma fluid equations. Thus, in the zero temperature limit, the kinetic dispersion relation obtained in this section reverts to the fluid dispersion relation obtained in Sect. 4.

6.7 Parallel Wave Propagation

Let us consider wave propagation, though a warm plasma, *parallel* to the equilibrium magnetic field. For parallel propagation, $k_\perp \rightarrow 0$, and, hence, from Eq. (6.81), $\lambda_s \rightarrow 0$. Making use of the asymptotic expansion (6.85), the matrix T_{ij} simplifies to

$$T_{ij} = \begin{pmatrix} [Z(\xi_1)+Z(\xi_{-1})]/2 & i\,[Z(\xi_1)-Z(\xi_{-1})]/2 & 0 \\ -i\,[Z(\xi_1)-Z(\xi_{-1})]/2 & [Z(\xi_1)+Z(\xi_{-1})]/2 & 0 \\ 0 & 0 & -Z'(\xi_0)\,\xi_0 \end{pmatrix}, \tag{6.89}$$

where, again, the only non-zero contributions are from $n = 0$ and $n = \pm 1$. The dispersion relation can be written [see Eq. (4.10)]

$$\mathbf{M}\cdot\mathbf{E} = \mathbf{0}, \tag{6.90}$$

where

$$M_{11} = M_{22} = 1 - \frac{k_z^2\, c^2}{\omega^2} + \frac{1}{2}\sum_s \frac{\omega_{ps}^2}{\omega\; k_z v_s}\left[Z\left(\frac{\omega-\Omega_s}{k_z\, v_s}\right) + Z\left(\frac{\omega+\Omega_s}{k_z\, v_s}\right)\right], \tag{6.91}$$

$$M_{12} = -M_{21} = \frac{i}{2}\sum_s \frac{\omega_{ps}^2}{\omega\, k_z v_s}\left[Z\left(\frac{\omega-\Omega_s}{k_z\, v_s}\right) - Z\left(\frac{\omega+\Omega_s}{k_z\, v_s}\right)\right], \tag{6.92}$$

$$M_{33} = 1 - \sum_s \frac{\omega_{ps}^2}{(k_z\, v_s)^2}\, Z'\left(\frac{\omega}{k_z\, v_s}\right), \tag{6.93}$$

and $M_{13} = M_{31} = M_{23} = M_{32} = 0$. Here, $v_s = \sqrt{2\, T_s/m_s}$ is the species-s thermal velocity.

The first root of Eq. (6.90) is

$$1 + \sum_s \frac{2\,\omega_{ps}^2}{(k_z\, v_s)^2}\left[1 + \frac{\omega}{k_z\, v_s}\, Z\left(\frac{\omega}{k_z\, v_s}\right)\right] = 0, \tag{6.94}$$

with the eigenvector $(0, 0, E_z)$. Here, use has been made of Eq. (6.40). This root evidentially corresponds to a longitudinal, electrostatic plasma wave. In fact, it is easily demonstrated that Eq. (6.94) is equivalent to the dispersion relation (6.50) that we found earlier for electrostatic plasma waves, for the special case in which the distribution functions are Maxwellians. Recall, from Sects. 6.3–6.5, that the electrostatic wave described by the above expression is subject to significant damping whenever the argument of the plasma dispersion function becomes less than or comparable with unity: *i.e.*, whenever $\omega \lesssim k_z\, v_s$.

The second and third roots of Eq. (6.90) are

$$\frac{k_z^2\, c^2}{\omega^2} = 1 + \sum_s \frac{\omega_{ps}^2}{\omega\; k_z v_s}\, Z\left(\frac{\omega+\Omega_s}{k_z\, v_s}\right), \tag{6.95}$$

with the eigenvector $(E_x, i\, E_x, 0)$, and

$$\frac{k_z^2\, c^2}{\omega^2} = 1 + \sum_s \frac{\omega_{ps}^2}{\omega\; k_z v_s}\, Z\left(\frac{\omega-\Omega_s}{k_z\, v_s}\right), \tag{6.96}$$

with the eigenvector $(E_x, -i\, E_x, 0)$. The former root evidently corresponds to a right-handed circularly polarized wave, whereas the latter root corresponds to a left-handed circularly polarized wave. The above two dispersion relations are essentially the same as the corresponding fluid dispersion relations, (4.89) and (4.90), except that they explicitly contain collisionless *damping* at the cyclotron resonances. As before, the damping is significant whenever the arguments of the plasma dispersion functions are less than or of order unity. This corresponds to

$$\omega - |\Omega_e| \lesssim k_z\, v_e \tag{6.97}$$

for the right-handed wave, and

$$\omega - \Omega_i \lesssim k_z v_i \tag{6.98}$$

for the left-handed wave.

The collisionless cyclotron damping mechanism is very similar to the Landau damping mechanism for longitudinal waves discussed in Sect. 6.3. In this case, the resonant particles are those which gyrate about the magnetic field with approximately the same angular frequency as the wave electric field. On average, particles which gyrate slightly faster than the wave lose energy, whereas those which gyrate slightly slower than the wave gain energy. In a Maxwellian distribution there are less particles in the former class than the latter, so there is a net transfer of energy from the wave to the resonant particles. Note that in kinetic theory the cyclotron resonances possess a *finite* width in frequency space (*i.e.*, the incident wave does not have to oscillate at exactly the cyclotron frequency in order for there to be an absorption of wave energy by the plasma), unlike in the cold plasma model, where the resonances possess *zero* width.

6.8 Perpendicular Wave Propagation

Let us now consider wave propagation, through a warm plasma, *perpendicular* to the equilibrium magnetic field. For perpendicular propagation, $k_z \rightarrow 0$, and, hence, from Eq. (6.82), $\xi_n \rightarrow \infty$. Making use of the asymptotic expansions (6.83)–(6.84), the matrix T_{ij} simplifies considerably. The dispersion relation can again be written in the form (6.90), where

$$M_{11} = 1 - \sum_s \frac{\omega_{ps}^2}{\omega} \frac{e^{-\lambda_s}}{\lambda_s} \sum_{n=-\infty}^{\infty} \frac{n^2 I_n(\lambda_s)}{\omega - n\Omega_s}, \tag{6.99}$$

$$M_{12} = -M_{21} = i \sum_s \frac{\omega_{ps}^2}{\omega} e^{-\lambda_s} \sum_{n=-\infty}^{\infty} \frac{n\,[I_n'(\lambda_s) - I_n(\lambda_s)]}{\omega - n\Omega_s}, \tag{6.100}$$

$$M_{22} = 1 - \frac{k_\perp^2 c^2}{\omega^2} - \sum_s \frac{\omega_{ps}^2}{\omega} \frac{e^{-\lambda_s}}{\lambda_s} \sum_{n=-\infty}^{\infty} \frac{\left[n^2 I_n(\lambda_s) + 2\lambda_s^2 I_n(\lambda_s) - 2\lambda_s^2 I_n'(\lambda_s)\right]}{\omega - n\Omega_s}, \tag{6.101}$$

$$M_{33} = 1 - \frac{k_\perp^2 c^2}{\omega^2} - \sum_s \frac{\omega_{ps}^2}{\omega} e^{-\lambda_s} \sum_{n=-\infty}^{\infty} \frac{I_n(\lambda_s)}{\omega - n\Omega_s}, \tag{6.102}$$

and $M_{13} = M_{31} = M_{23} = M_{32} = 0$. Here,

$$\lambda_s = \frac{(k_\perp \rho_s)^2}{2}, \tag{6.103}$$

where $\rho_s = v_s/|\Omega_s|$ is the species-s Larmor radius.

The first root of the dispersion relation (6.90) is

$$n_\perp^{\,2} = \frac{k_\perp^{\,2} c^2}{\omega^2} = 1 - \sum_s \frac{\omega_{ps}^{\,2}}{\omega}\, e^{-\lambda_s} \sum_{n=-\infty}^{\infty} \frac{I_n(\lambda_s)}{\omega - n\,\Omega_s}, \tag{6.104}$$

with the eigenvector $(0, 0, E_z)$. This dispersion relation obviously corresponds to the electromagnetic plasma wave, or *ordinary* mode, discussed in Sect. 4.10. Note, however, that in a warm plasma the dispersion relation for the ordinary mode is strongly modified by the introduction of *resonances* (where the refractive index, $n_\perp$, becomes infinite) at all the *harmonics* of the cyclotron frequencies:

$$\omega_{n\,s} = n\,\Omega_s, \tag{6.105}$$

where n is a non-zero integer. These resonances are a *finite Larmor radius* effect. In fact, they originate from the variation of the wave phase across a Larmor orbit. Thus, in the cold plasma limit, $\lambda_s \rightarrow 0$, in which the Larmor radii shrink to zero, all of the resonances disappear from the dispersion relation. In the limit in which the wave-length, λ, of the wave is much larger than a typical Larmor radius, ρ_s, the relative amplitude of the nth harmonic cyclotron resonance, as it appears in the dispersion relation (6.104), is approximately $(\rho_s/\lambda)^{|n|}$ [see Eqs. (6.85) and (6.103)]. It is clear, therefore, that in this limit only low-order resonances [*i.e.*, $n \sim O(1)$] couple strongly into the dispersion relation, and high-order resonances (*i.e.*, $|n| \gg 1$) can effectively be neglected. As $\lambda \rightarrow \rho_s$, the high-order resonances become increasigly important, until, when $\lambda \lesssim \rho_s$, all of the resonances are of approximately equal strength. Since the ion Larmor radius is generally much larger than the electron Larmor radius, it follows that the ion cyclotron harmonic resonances are generally more important than the electron cyclotron harmonic resonances.

Note that the cyclotron harmonic resonances appearing in the dispersion relation (6.104) are of *zero width* in frequency space: *i.e.*, they are just like the resonances which appear in the cold-plasma limit. Actually, this is just an artifact of the fact that the waves we are studying propagate *exactly perpendicular* to the equilibrium magnetic field. It is clear from an examination of Eqs. (6.80) and (6.82) that the cyclotron harmonic resonances originate from the zeros of the plasma dispersion functions. Adopting the usual rule that substantial damping takes place whenever the arguments of the dispersion functions are less than or of order unity, it is clear that the cyclotron harmonic resonances lead to significant damping whenever

$$\omega - \omega_{n\,s} \lesssim k_z\, v_s. \tag{6.106}$$

Thus, the cyclotron harmonic resonances possess a *finite* width in frequency space provided that the parallel wave-number, k_z, is non-zero: *i.e.*, provided that the wave does not propagate exactly perpendicular to the magnetic field.

The appearance of the cyclotron harmonic resonances in a warm plasma is of great practical importance in plasma physics, since it greatly increases the number of resonant frequencies at which waves can transfer energy to the plasma. In magnetic fusion these resonances are routinely exploited to heat plasmas via externally launched electromagnetic

waves. Hence, in the fusion literature you will often come across references to "third harmonic ion cyclotron heating" or "second harmonic electron cyclotron heating."

The other roots of the dispersion relation (6.90) satisfy

$$\left(1-\sum_s \frac{\omega_{ps}^2}{\omega}\frac{e^{-\lambda_s}}{\lambda_s}\sum_{n=-\infty}^{\infty}\frac{n^2\,I_n(\lambda_s)}{\omega-n\,\Omega_s}\right)\left(1-\frac{k_\perp^2\,c^2}{\omega^2}\right.$$
$$\left.-\sum_s \frac{\omega_{ps}^2}{\omega}\frac{e^{-\lambda_s}}{\lambda_s}\sum_{n=-\infty}^{\infty}\frac{\left[n^2\,I_n(\lambda_s)+2\,\lambda_s^2\,I_n(\lambda_s)-2\,\lambda_s^2\,I_n'(\lambda_s)\right]}{\omega-n\,\Omega_s}\right)$$
$$=\left(\sum_s \frac{\omega_{ps}^2}{\omega}\,e^{-\lambda_s}\sum_{n=-\infty}^{\infty}\frac{n\,[I_n'(\lambda_s)-I_n(\lambda_s)]}{\omega-n\,\Omega_s}\right)^2, \tag{6.107}$$

with the eigenvector $(E_x, E_y, 0)$. In the cold plasma limit, $\lambda_s \rightarrow 0$, this dispersion relation reduces to that of the *extraordinary* mode discussed in Sect. 4.10. This mode, for which $\lambda_s \ll 1$, unless the plasma possesses a thermal velocity approaching the velocity of light, is little affected by thermal effects, except close to the cyclotron harmonic resonances, $\omega = \omega_{n\,s}$, where small thermal corrections are important because of the smallness of the denominators in the above dispersion relation.

However, another mode also exists. In fact, if we look for a mode with a phase velocity much less than the velocity of light (*i.e.*, $c^2\,k_\perp^2/\omega^2 \gg 1$) then it is clear from (6.99)–(6.102) that the dispersion relation is approximately

$$1-\sum_s \frac{\omega_{ps}^2}{\omega}\frac{e^{-\lambda_s}}{\lambda_s}\sum_{n=-\infty}^{\infty}\frac{n^2\,I_n(\lambda_s)}{\omega-n\,\Omega_s}=0, \tag{6.108}$$

and the associated eigenvector is $(E_x, 0, 0)$. The new waves, which are called *Bernstein waves* (after I.B. Bernstein, who first discovered them), are clearly slowly propagating, longitudinal, electrostatic waves.

Let us consider electron Bernstein waves, for the sake of definiteness. Neglecting the contribution of the ions, which is reasonable provided that the wave frequencies are sufficiently high, the dispersion relation (6.108) reduces to

$$1-\frac{\omega_p^2}{\omega}\frac{e^{-\lambda}}{\lambda}\sum_{n=-\infty}^{\infty}\frac{n^2\,I_n(\lambda)}{\omega-n\,\Omega}=0, \tag{6.109}$$

where the subscript s is dropped, since it is understood that all quantities relate to electrons. In the limit $\lambda \rightarrow 0$ (with $\omega \neq n\,\Omega$), only the $n = \pm 1$ terms survive in the above expression. In fact, since $I_{\pm 1}(\lambda)/\lambda \rightarrow 1/2$ as $\lambda \rightarrow 0$, the dispersion relation yields

$$\omega^2 \rightarrow \omega_p^2 + \Omega^2. \tag{6.110}$$

It follows that there is a Bernstein wave whose frequency asymptotes to the upper hybrid frequency [see Sect. 4.10] in the limit $k_\perp \rightarrow 0$. For other non-zero values of n, we have

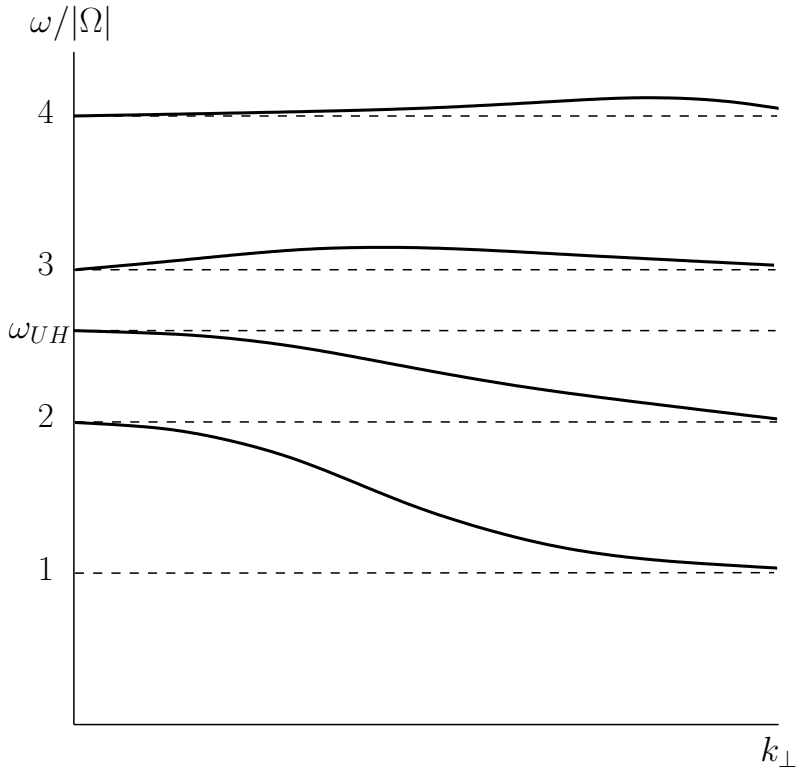

Figure 6.8: *Dispersion relation for electron Bernstein waves in a warm plasma.*

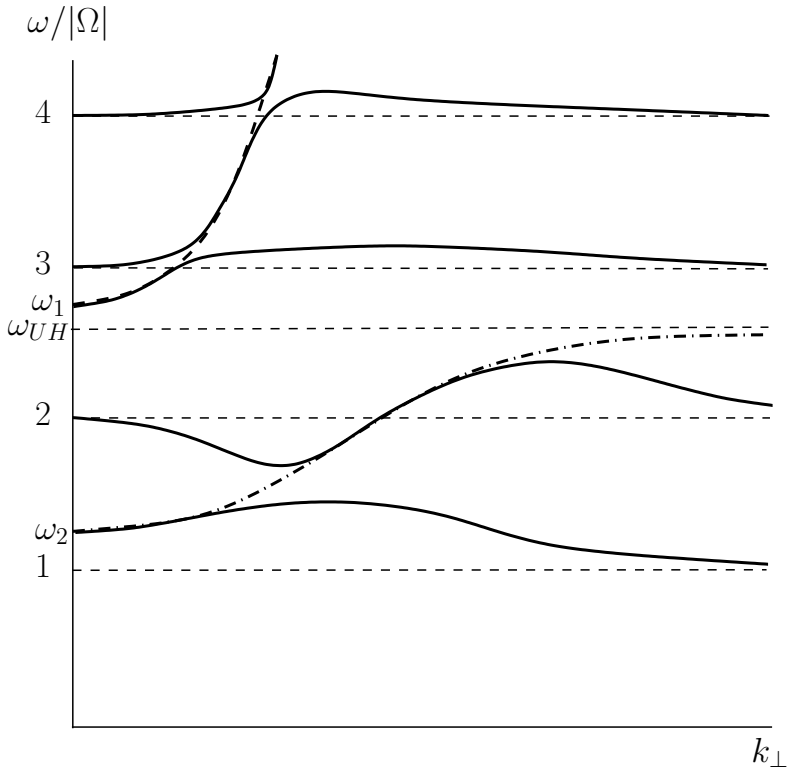

Figure 6.9: *Dispersion relation for electron Bernstein waves in a warm plasma. The dashed line indicates the cold plasma extraordinary mode.*

$I_n(\lambda)/\lambda \to 0$ as $\lambda \to 0$. However, a solution to Eq. (6.108) can be obtained if $\omega \to n\,\Omega$ at the same time. Similarly, as $\lambda \to \infty$ we have $e^{-\lambda} I_n(\lambda) \to 0$. In this case, a solution can only be obtained if $\omega \to n\,\Omega$, for some n, at the same time. The complete solution to Eq. (6.108) is sketched in Fig. 6.8, for the case where the upper hybrid frequency lies between $2\,|\Omega|$ and $3\,|\Omega|$. In fact, wherever the upper hybrid frequency lies, the Bernstein modes above and below it behave like those in the diagram.

At small values of $k_\perp$, the phase velocity becomes large, and it is no longer legitimate to neglect the extraordinary mode. A more detailed examination of the complete dispersion relation shows that the extraordinary mode and the Bernstein mode cross over near the harmonics of the cyclotron frequency to give the pattern shown in Fig. 6.9. Here, the dashed line shows the cold plasma extraordinary mode.

In a lower frequency range, a similar phenomena occurs at the harmonics of the ion cyclotron frequency, producing ion Bernstein waves, with somewhat similar properties to electron Bernstein waves. Note, however, that whilst the ion contribution to the dispersion relation can be neglected for high-frequency waves, the electron contribution cannot be neglected for low frequencies, so there is not a complete symmetry between the two types of Bernstein waves.